Multigenerational Regulation of the *C. elegans* Chromatin Landscape by Germline Small RNAs

by

Natasha Elise Weiser

**A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Cellular and Molecular Biology)
in the University of Michigan
2017**

Doctoral Committee:

**Associate Professor John K. Kim, Johns Hopkins University, Co-Chair
Associate Professor JoAnn Sekiguchi, Co-Chair
Associate Professor Ken Inoki
Assistant Professor Shigeki Iwase
Associate Professor Sundeep Kalantry**

"All human wisdom can be contained in these two words—wait and hope."

~Alexandre Dumas, *The Count of Monte Cristo*

ACKNOWLEDGEMENTS

I am grateful to all my colleagues in the Kim lab for their support, technically and personally, over the last four years. From the beginning, I looked up to Allison Billi, Ting Han, and Vishal Khivansara who were wonderful role models for scientific rigor and grace under fire. Amelia Alessi has also been a great resource and always generous with her time and expertise. Danny Yang initiated the MORC-1 project in the lab and was always a source of good ideas. Mallory Freeberg contributed her bioinformatics expertise to my projects throughout my time in the lab. My thesis committee, JoAnn Sekiguchi, Ken Inoki, Shigeki Iwase, Sundeep Kalantry, and formerly Aaron Goldstrohm and Patrick Hu, have provided valuable scientific insights and technical help. I have been very lucky to have found wonderful mentors at every stage of my education, especially Brian Gauvin, whose tenth grade chemistry class first sparked my interest in science, and Amy Cheng Vollmer, my undergraduate research advisor. I am so grateful to my advisor, John Kim, for providing such a rigorous training environment and giving me the independence to take full advantage of it. My time in the Kim lab has been the most intellectually engaging experience I have ever had; sometimes I worry that everything else will be boring by comparison.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

Table

LIST OF ABBREVIATIONS AND ACRONYMS

C. elegans: Caenorhabditis elegans

ChIP: chromatin immunoprecipitation

CNV: copy number variation

CRISPR: clustered, regularly interspaced short palindromic repeats

crRNA: CRISPR targeting RNA

DAPI: 4',6-diamidino-2-phenylindole

DCC: dosage compensation complex

DSB: double-stranded break

dsRNA: double-stranded RNA

endo-: endogenous

ENU: N-ethyl-N-nitrosourea

Eri: enhanced RNA interference (phenotype)

exo-: exogenous

FDR: false discovery rate

FISH: fluorescence in situ hybridization

GHKL: gyase, Hsp90, histidine kinase, MutL

H3K27me3: histone 3 lysine 27 trimethyl

H3K36me3: histone 3 lysine 36 trimethyl

H3K4me2/3: histone 3 lysine 4 di/trimethyl

H3K9me2/3: histone 3 lysine 9 di/trimethyl

Him: high incidence of males (phenotype)

Hrde: heritable RNAi defective (phenotype)

IP-MS: immunoprecipitation-mass spectrometry

kb: kilobase

miRNA: microRNA

Morc: microrchidia

Mrt: germline mortal (phenotype)

Mut: mutator (phenotype)

Nrde: nuclear RNAi defective (phenotype)

nt: nucleotide

piRNA: Piwi-interacting RNA

qRT-PCR: quantitative real-time polymerase chain reaction

qPCR: quantitative polymerase chain reaction

Rde: RNAi defective (phenotype)

RdRP: RNA-dependent RNA polymerase

RISC: RNA induced silencing complex

RNA Pol II: RNA polymerase II

RNAa: RNA-induced transactivation

RNAe: RNA-induced epigenetic silencing

RNAi: RNA interference

RPKM: Reads Per Kilobase of transcript per Million mapped reads

siRNA: small interfering RNA

SNP: single nucleotide polymorphism

snRNP: small nuclear ribonucleo protein

TOFU: twenty-one U fouled up

UTR: untranslated region

WAGO: worm-specific Argonaute

WT: wildtype

ABSTRACT

In evolutionary terms, propagation of a species is the only biological imperative. In sexually reproducing species, species survival requires the formation of specialized gametes that exist in a perpetual cycle of fertilization and establishment of totipotency, differentiation into the next generation of gametes, and then fertilization and the re-establishment of totipotency. This makes the germline an immortal cell lineage. To successfully repeat this cycle at every generation, the germline must maintain its replicative immortality. Equally important is the accurate transmission of genetic and epigenetic information to the next generation to facilitate the correct execution of the developmental program. Thus, pathways that protect genome fidelity and ensure appropriate gene expression at every generation are essential for germline immortality.

In animals, small non-coding RNAs that are expressed in the germline and transmitted to progeny control gene expression to promote fertility. Germline-expressed small RNAs, including endogenous siRNAs (endo-siRNAs) and Piwi-interacting RNAs (piRNAs), drive the repression of deleterious transcripts such as transposons, repetitive elements, and pseudogenes. Small RNA pathways are highly conserved in metazoans and have been best described in the model organism *Caenorhabditis elegans*. Many features of *C. elegans* make it ideal for the study of small RNA biology and germline maintenance, including genetic tractability, rapid development, an invariant cell lineage, and propagation via self-fertilization.

In *C. elegans*, endo-siRNAs are deposited from the maternal germline into the embryo, where they trigger the amplification of another generation of endo-siRNAs as well as heritable heterochromatin formation at target genes. Transgenerational inheritance of endo-siRNAs is critical for germline maintenance, as mutants that are *h*eritable *R*NAi *de*fective (Hrde) are germline mortal (progressively sterile). How endo-siRNAs control chromatin state and germline maintenance is unknown. I have identified *morc-1*, the sole *C. elegans* homolog of the highly conserved Microrchidia family of

chromatin-binding proteins, as a crucial link between endo-siRNAs and multigenerational chromatin organization. I also identify and describe the first suppressor of the germline mortality phenotype of *hrde* pathway mutants. This work establishes a critical role for endo-siRNAs and MORC-1 in the control of transgenerational chromatin architecture to promote fertility.

CHAPTER 1

Introduction to *C. elegans* small RNAs

INTRODUCTION

In the years since the discovery of RNA interference (RNAi), the robust silencing of homologous genes caused by double stranded RNAs (dsRNAs) (Fire et al., 1998), small non-coding RNA molecules have emerged as major determinants of gene expression. Exogenous RNAi (exo-RNAi) has become a widely used tool for research, allowing for targeted gene-knockdown by delivery of exogenous dsRNA and subsequent processing into small interfering RNAs (siRNAs). The proteins that execute the biogenesis and silencing function of siRNAs also regulate gene expression during organismal development through the generation and effector function of endogenous small non-coding RNAs. Deep sequencing of small RNAs from *C. elegans* and subsequent studies have uncovered numerous classes of small RNAs regulating diverse processes during development, including microRNAs (miRNAs), 26G and 22G endogenous siRNAs (endo-siRNAs), and 21U Piwi-interacting RNAs (piRNAs). While miRNAs primarily regulate somatic development, endo-siRNAs and piRNAs are expressed in the *C. elegans* germline and are required for fertility.

Here, I will discuss the salient features of each germline small RNA pathway. The endo-siRNAs can be broadly broken up into two major 26G RNA pathways and two major 22G RNA pathways. Each pathway is defined by a signature Argonaute protein, the core component of all RNA-induced silencing complexes (RISC), which guide small RNAs to their targets. The Argonaute proteins all share two characteristic RNA-binding domains, PAZ and PIWI. The PIWI domain is RNAse H-related and some Argonautes have retained the catalytic activity of this domain, indicated by the presence of a catalytic triad of two aspartates and one histidine, the DDH motif, that enables the

cleavage or "slicing" of target RNAs. I will follow the convention of classifying small RNAs by their associated Argonaute to discuss, in turn, ERGO-1- and ALG-3/4-class 26G RNAs and WAGO- and CSR-1-class 22G RNAs. piRNAs are bound by the PRG-1 Argonaute and are classified as a single group. Finally, I will discuss emerging models of how different small RNA pathways interact to coordinate germline gene expression.

ENDOGENOUS siRNAs

The first large-scale studies identifying endo-siRNAs utilized high-throughput sequencing of mixed-stage *C. elegans* as well as purified germ cells. These studies identified two pools of antisense endo-siRNAs with a 5' bias for guanosine, enriched at lengths of 26nt and 22nt, now respectively called 26G RNAs and 22G RNAs (Ambros et al., 2003; Ruby et al., 2006; Yigit et al., 2006). The focus of this chapter is entirely on the endogenous small RNAs, however, the endo-siRNA pathway is tightly linked to the exo-RNAi pathway, and therefore a brief overview of key features of the exogenous pathway is instructive.

Many of the protein factors that regulate endo-siRNAs also contribute to robust transgene silencing and RNAi-mediated gene silencing. Thus, many of the components of the endo-siRNA pathways have been identified through genetic screens for genes affecting RNAi sensitivity and expression of transgenes (Grishok et al., 2000; Kennedy et al., 2004; Kim et al., 2005; Robert et al., 2005; Tabara et al., 1999). As the endogenous and exogenous siRNA pathways compete for shared protein components, loss of factors that are shared between the endogenous and exogenous pathways, such are DCR-1 and RDE-4, leads to an RNAi deficient (Rde) phenotype. Loss of factors that are unique to the endogenous pathway, such as ERI-1, frees the shared components to process exogenous siRNAs, leading to the generation of more exo-siRNAs and thus to enhanced RNAi sensitivity (Eri phenotype) (Duchaine et al., 2006; Kennedy et al., 2004; Lee et al., 2006; Yigit et al., 2006).

Both endogenous and exogenous RNAi involve two sequential rounds of siRNA biogenesis (Gent et al., 2010; Yigit et al., 2006). In the exogenous RNAi pathway, long

dsRNAs are processed by DCR-1, RDE-4, and DRH-1 into 22nt primary siRNAs and loaded onto the Argonaute RDE-1 (Tabara et al., 1999; 2002; Thivierge et al., 2012). DRH-1 is a DEAD-box helicase whose paralog, DRH-3, functions in the endo-siRNA pathways (Tabara et al., 2002). DCR-1 and RDE-4 are discussed in greater detail below. The exo-siRNAs trigger the production of secondary 22G RNAs by the WAGO 22G RNA pathway (Yigit et al., 2006). In the endogenous pathway, which I explain here in detail, the primary siRNAs are the 26G RNAs (Gent et al., 2010). In association with either ERGO-1 or ALG-3/4 RISC, the 26G RNAs trigger the biogenesis of WAGO 22G RNAs to regulate their target genes (Ambros et al., 2003; Gent et al., 2010; Vasale et al., 2010; Yigit et al., 2006).

26G RNAs

26G RNAs are expressed in the germline and maternally deposited in embryos (Han et al., 2009). 26G RNA expression is temporally regulated, with distinct 26G RNAs and distinct 26G-associated Argonautes expressed in spermatogenic versus oogenic germlines (Han et al., 2009). Spermatogenesis-enriched and oogenesis-enriched 26G RNAs regulate distinct target genes (Han et al., 2009). The ALG-3/4-class 26G RNAs are made during spermatogenesis and their target genes are enriched for transcripts expressed during spermatogenesis (Conine et al., 2010; 2013; Han et al., 2009). As a result, loss of ALG-3/4 class 26G RNAs causes temperature sensitive (ts) sterility of spermatogenic origin (Conine et al., 2010; Han et al., 2009). The ALG-3/4 pathway is not thought to compete with exo-RNAi for shared cofactors, as exclusive loss of ALG-3/4-class 26GRNAs (such as in *alg-3(-);alg-4(-)* mutants) does not result in an RNAi sensitivity phenotype (Han et al., 2009). The oogenesis-enriched 26G RNAs associate with ERGO-1 RISC (Han et al., 2009; Vasale et al., 2010). The ERGO-1-class 26G RNAs do not target germline-enriched genes; thus, exclusive loss of ERGO-1-class 26G RNAs (such as in *ergo-1(-)* mutants) does not significantly impair fertility (Han et al., 2009). Exclusive loss of the ERGO-1-class 26G RNAs causes an Eri phenotype due to increased availability of cofactors that are also involved in exo-RNAi (Duchaine et al.,

2006; Han et al., 2009; Kennedy et al., 2004; Lee et al., 2006; Yigit et al., 2006). Accordingly, loss of factors that are shared between the ERGO-1-class and the ALG-3/4 class causes both Eri and ts sterility phenotypes. The ERGO-1 and ALG-3/4 pathways share an identical set of biogenesis factors and then diverge at Argonaute loading. As a result, loss of any of the factors that act upstream of Argonaute causes depletion of both ALG-3/4- and ERGO-1-class 26G RNAs and therefore both Eri and ts sterility. Here, I will first discuss the shared 26G RNA biogenesis cofactors. I will then discuss Argonaute loading and effector function of the ALG-3/4-class and the ERGO-1-class separately.

Biogenesis of 26G RNAs

The 26G RNAs are transcribed by RRF-3, one of the four *C. elegans* RNA-dependent RNA polymerases (RdRP) (Conine et al., 2010; Duchaine et al., 2006; Gent et al., 2010; 2009; Han et al., 2009; Pavelec et al., 2009; Smardon et al., 2000; Vasale et al., 2010). As the 26G RNAs bind their targets with perfect complementarity, the transcripts that serve as templates for 26G RNA biogenesis are also targeted by 26G RNAs for silencing (Han et al., 2009). Some 26G RNAs span exon-exon junctions, indicating that they originate from spliced templates (Han et al., 2009; Ruby et al., 2006). *rrf-3(-)* mutants have reduced fertility and a high incidence of males (Him) compared to wildtype worms. These phenotypes are rescued by mating to wildtype male worms, indicating that they arise from defects in spermatogenesis (Gent et al., 2009). As *rrf-3* is dispensable for exogenous RNAi, *rrf-3(-)* mutants are also Eri (Simmer et al., 2002).

The 26G RNAs are 5' monophosphorylated, indicating that they are processed by the sole *C. elegans* Dicer homolog, DCR-1, an RNase III endonuclease that is specific to dsRNA (Gent et al., 2010; Grishok et al., 2001; Han et al., 2009). DCR-1-mediated cleavage generates short RNA duplexes; deep-sequencing studies have shown that these products contain the antisense 26G RNA in a duplex with a sense-oriented passenger strand of 23nt and a 3nt 3' overhang (Ruby et al., 2006). The DCR-

1 helicase domain allows for the processive trimming of blunt ends and 5' overhangs to generate 3' overhangs of two or three nucleotides (Pavelec et al., 2009; Welker et al., 2011; 2010). Thus, helicase-dead DCR-1 variants cannot process endo-siRNAs (Welker et al., 2010; 2011). Importantly, DCR-1 also processes exogenous dsRNA delivered by RNAi and miRNAs. As a result, *dcr-1(-)* mutants exhibit pleiotropic phenotypes: Rde due to loss of exo-RNAi, impaired fertility due to loss of ALG-3/4 26G RNAs, and developmental timing (heterochronic) defects due to loss of miRNAs (Grishok et al., 2001).

Immunopurification experiments have shown that RRF-3 and DCR-1 function as part of a complex, termed the ERI complex, whose other components are also essential for 26G RNA biogenesis. The ERI complex includes an RdRP subcomplex, or module, containing the RdRP RRF-3, the Tudor domain protein ERI-5, and the Dicer-like helicase DRH-3 (Duchaine et al., 2006; Thivierge et al., 2012). As each of these factors (i.e. RRF-3, ERI-5, and DRH-3) is dependent on the other two for its association with DCR-1, these three components are considered part of a distinct module within the ERI complex (Thivierge et al., 2012). A paralog of ERI-5, EKL-1, contributes to 22G RNA accumulation and can function in the ERI complex in *eri-5(-)* mutants; accordingly, *eri-5(-)* mutants only exhibit about 50% depletion of 26G RNAs (Thivierge et al., 2012). DRH-3 is an essential component of all RdRP modules, thus *drh-3(-)* mutants show depletion of many classes of endo-siRNAs and are Rde because they cannot generate exo-RNAi-triggered secondary 22G RNAs (Aoki et al., 2007; Claycomb et al., 2009; Duchaine et al., 2006; Gu et al., 2009; Nakamura et al., 2007). Other phenotypes of *drh-3(-)* mutants will be discussed in the section on CSR-1-class 22G RNAs. Recent biochemical studies have shown that DRH-3 functions as a homodimer that preferentially binds RNA duplexes of 22nt or longer (Fitzgerald et al., 2014). As DRH-3 is essential for the association of the RdRP module with DCR-1, the preference of DRH-3 for substrates of at least 22nt likely prevents DCR-1 from binding and processing shorter duplexes (i.e. less than 22nt). This may contribute to the long-recognized lower limit of 22nt for DCR-1 products.

Other members of the ERI complex include ERI-1b, ERI-3, and RDE-4. The DEDDh 3'-5' exonuclease ERI-1 was first identified in a screen for mutants that

enhance neuronal sensitivity to RNAi (Kennedy et al., 2004). *eri-1* encodes two isoforms, both of which function in the processing of 5.8S rRNA; only the ERI-1b isoform interacts with DCR-1 and contributes to endo-siRNA biogenesis (Duchaine et al., 2006; Gabel and Ruvkun, 2008). Although the precise function of ERI-1b has not been described, it is thought to process double-stranded RRF-3 products to make blunt ends that can then be processed by the DCR-1 helicase domain (Welker et al., 2011). *eri-1(-)* mutants are depleted of all 26G RNAs and exhibit the phenotypes characteristic of these defects: ts sterility due to defective spermatogenesis and Eri (Duchaine et al., 2006; Gabel and Ruvkun, 2008; Han et al., 2009; Kennedy et al., 2004; Pavelec et al., 2009). Although the function of ERI-3, which has no annotated domains, is unknown, it is essential for formation of the ERI complex, suggesting that it may recruit other components or stabilize their interactions (Duchaine et al., 2006). The contribution of RDE-4 is also unknown. RDE-4 functions in both endo-siRNA and exo-siRNA pathways, and thus the *rde-4(-)* mutant is ts sterile and Rde (Blanchard et al., 2011; Tabara et al., 2002).

26G RNA biogenesis is temporally regulated, occurring in two stages during worm development and resulting in the accumulation of two distinct classes of 26G RNAs (Figure 1.1) (Han et al., 2009). During spermatogenesis, in the fourth larval stage, 26G RNAs are loaded onto a RISC complex containing one of two redundant Argonaute proteins ALG-3 and ALG-4 (Han et al., 2009). The ALG-3/4-class 26G RNAs template mRNAs are enriched for transcripts expressed during spermatogenesis, this likely contributes to the temporal regulation of ALG-3/4-class 26G RNAs (Conine et al., 2010; 2013; Han et al., 2009). The restriction of ALG-3 expression to the spermatogenic germline probably also contributes to the temporal regulation of its associated 26G RNAs (Conine et al., 2010). When worms reach adulthood, the differentiation of germline nuclei shifts to an oogenic fate. When the germline switches to oogenesis, accumulation of an oogenic-and embryo-specific class of 26G RNAs occurs (Han et al., 2009). The female 26G RNAs are bound by the Argonaute ERGO-1 and its associated RISC complex (Han et al., 2009; Vasale et al., 2010). The targets of the ERGO-1-class 26G RNAs are not enriched for germline-expressed transcripts.

ERGO-1-class 26G RNAs

The biological significance of the oogenic class of 26G RNAs is somewhat mysterious, as exclusive loss of this class, as in *ergo-1(-)* mutants, is not associated with a significant fertility defect or any phenotype other than Eri (Han et al., 2009; Pavelec et al., 2009; Yigit et al., 2006). Characterization of the mRNA targets of this pathway indicates that ERGO-1 targets are depleted for spermatogenesis, oogenesis, and germline-intrinsic genes, suggesting that the ERGO-1-class 26G RNAs do not contribute to the tuning of germline gene expression (Han et al., 2009). Many targets are gene duplications and other non-coding transcripts; additionally, many ERGO-1 targets are misannotated or do not map uniquely to the genome (Fischer et al., 2011). The fact that over half of ERGO-1 targets are not conserved in the closely related nematode *C. briggsae* suggests that this pathway may target recently evolved genes (Fischer et al., 2011).

ERGO-1 is robustly expressed in the oogenic germline of adult worms beginning at pachytene exit (Billi et al., 2012; Vasale et al., 2010). Its expression persists in embryos and early larval stages (Billi et al., 2012; Vasale et al., 2010). There is no detectable ERGO-1 expression in the third and fourth larval stages (Vasale et al., 2010). Because 26G RNA stability requires Argonaute, the pattern of temporal regulation of the ERGO-1-class 26G RNAs mirrors the temporal regulation of the ERGO-1 protein (Vasale et al., 2010).

ERGO-1 is one of the few *C. elegans* Argonaute proteins in which the DDH motif, the catalytic triad that is required for slicing activity, is conserved (Fischer et al., 2011; Yigit et al., 2006). Like *ergo-1* deletion mutants, worms expressing a putative slicer-dead *ergo-1* variant are Eri, suggesting that slicing is required for ERGO-1 effector function (Fischer et al., 2011). While ERGO-1 slicing function does not contribute significantly to target silencing, it is thought to allow the dissociation of the siRNA-mRNA duplex (Fischer et al., 2011). As described below, the secondary 22G RNAs are the major effectors of target silencing in this pathway.

Although biogenesis of the ERGO-1-class 26G RNAs is occurs primarily in the adult germline, the enrichment of 26G RNA transcripts in embryos indicates that they are stably deposited and maintained in the developing embryo. Perdurance of 26G RNAs in embryos requires protection against exonucleolytic degradation, achieved by 2' O-methylation at the 3' end of the 26G RNA by the methyltransferase HENN-1 (Billi et al., 2012; Kamminga et al., 2012; Montgomery et al., 2012). The increased stability afforded by HENN-1-mediated 2' O-methylation allows 26G RNA expression to persist in embryos and early larvae (Billi et al., 2012; Kamminga et al., 2012; Montgomery et al., 2012). In *ergo-1(-)* mutants, the few residual 26G RNAs are not 2' O-methylated, suggesting that HENN-1 recruitment and subsequent siRNA methylation are dependent on ERGO-1 binding (Billi et al., 2012). Methylation by HENN-1 is not unique to the ERGO-1-class 26G RNAs, but is shared by 21U piRNAs and possibly some 22G RNAs (Billi et al., 2012; Kamminga et al., 2012; Montgomery et al., 2012). As in other animals, *C. elegans* miRNAs are not 2' O-methylated (Billi et al., 2012; Montgomery et al., 2012).

A number of other factors contribute to accumulation of ERGO-1-class 26G RNAs, although their precise roles have not been defined. These include ERI-6/7, ERI-9, and several Mutator proteins. ERI-6/7 is a single helicase protein encoded by two antiparallel genes, *eri-6* and *eri-7*, that are trans-spliced into a single mRNA (Fischer et al. 2008; Fischer et al. 2011). ERI-9 is conserved only in closely related *Caenorhabditis* species and is required for the accumulation of ERGO-1-class 26G RNAs and ERGO-1-dependent 22G RNAs (Pavelec et al., 2009). ERI-6/7 and ERI-9 are not required for fertility at elevated temperatures, as would be expected of ALG-3/4-class regulators, and thus are thought to act exclusively in the ERGO-1 pathway (Fischer et al., 2011; Pavelec et al., 2009). MUT-16 is critical for ERGO-1-class 26G RNA accumulation and target silencing (Zhang et al., 2011). Other Mutator proteins, including MUT-2, 7, 8, 14, and 15 also contribute to accumulation of ERGO-1-class 26G RNAs but not ALG-3/4-class 26G RNAs (Zhang et al., 2011). These factors will be discussed further in the WAGO 22G RNA section, as this is the context in which their function is best described. Some *mut-16* mutants are ts sterile, which is likely a reflection of its broader role in endo-siRNA accumulation, as these mutants are also depleted of WAGO and CSR-1

class of 22G RNAs and ALG-3/4 class of 22G (but not 26G) endo-siRNAs (Zhang et al., 2011).

ALG-3/4-class 26G RNAs

The ALG-3 and ALG-4 Argonautes are expressed in spermatogenic germline beginning at pachytene exit and persisting up to but not including mature spermatids (Conine et al., 2010). Due to their functional redundancy, *alg-3(-);alg-4(-)* double mutants but not *alg-3(-)* or *alg-4(-)* single mutants are depleted of the associated 26G RNAs and exhibit ts sterility. Unlike the ERGO-1-class 26G RNAs, the ALG-3/4-class 26G RNAs are not 2' O-methylated (Billi et al., 2012; Kamminga et al., 2012; Montgomery et al., 2012). This is consistent with the model that stabilization via methylation is needed for 26G RNA expression during embryogenesis, as ALG-3/4-class 26G RNAs are not detected in mature sperm or in embryos (Billi et al., 2012; Conine et al., 2010).

The ALG-3/4 class of 26G RNAs regulates a large suite of spermatogenesis genes (Conine et al., 2013; Han et al., 2009; Ruby et al., 2006). This regulation is critical for spermatogenesis, as indicated by the ts sterility of *alg-3(-);alg-4(-)* mutants that can be rescued by mating to wildtype males (Conine et al., 2010; 2013; Han et al., 2009). Interestingly, high-throughput sequencing studies have elucidated two sets of ALG-3/4 target genes; one set that is highly upregulated by ALG-3/4 (positively regulated) and a second that is repressed by ALG-3/4 (negatively regulated) (Conine et al., 2013).

Occupancy of RNA polymerase II (RNA Pol II) at the positively-regulated targets is *alg-3/4*-dependent, indicating that ALG-3/4 contributes to co-transcriptional positive regulation, or licensing, of its target genes (Conine et al., 2013). The contribution of ALG-3/4 to RNA Pol II occupancy at these targets is more robust at 25°C than at 20°C, suggesting that the ALG-3/4 26G RNAs promote the expression of spermatogenesis genes at elevated temperatures (Conine et al., 2013). The enhanced downregulation of ALG-3/4 target genes in *alg-3(-);alg-4(-)* mutants grown at elevated temperatures is

likely a major contributor to the ts sterility observed in *alg-3(-);alg-4(-)* mutants. In agreement with a role for ALG-3/4-class 26G RNAs promoting the expression of their target genes, H3K4me2, a marker of euchromatin, is depleted by immunostaining in the germlines of *alg-3(-);alg-4(-)* double mutants compared to wildtype (Conine et al., 2013). ALG-3/4-directed gene licensing is thought to occur via the CSR-1 22G RNA pathway, the details of which are explained below.

Like the ERGO-1-class 26G RNAs, the ALG-3/4-class 26G RNAs trigger the biogenesis WAGO 22G RNAs (Conine et al., 2010; 2013). The secondary WAGO-class 22G RNAs are required for ALG-3/4 effector function at negatively-regulated targets (i.e. post-transcriptional gene silencing) (Conine et al., 2013). Unlike the *ergo-1(-)* mutant, the *alg-3(-);alg-4(-)* double mutant does not have an RNAi sensitivity phenotype (Han et al., 2009). This suggests that the ALG-3/4 pathway does not provide a large enough workload to the WAGO pathway to compete with the exo-RNAi pathway for shared protein cofactors.
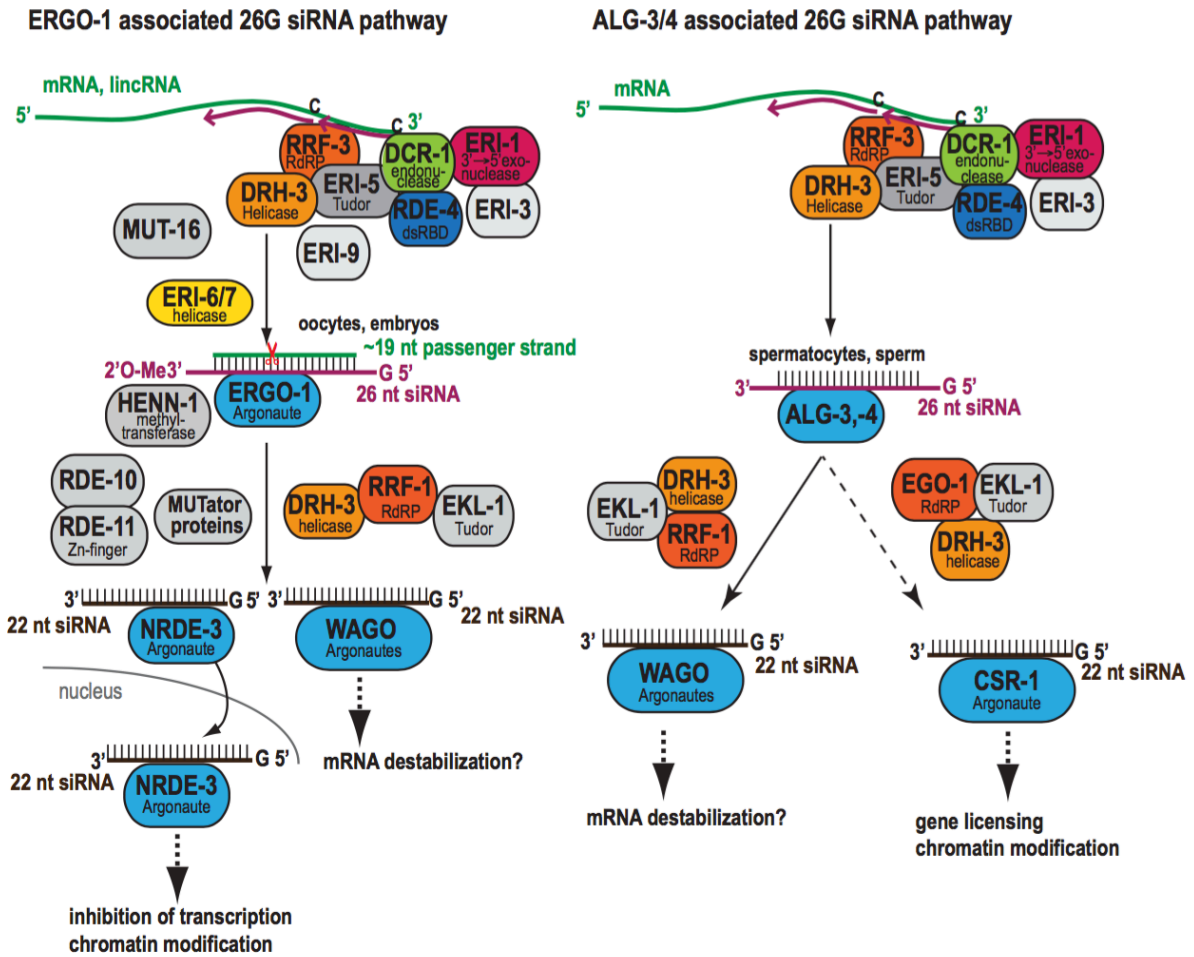
Figure 1.1. Overview of the 26G RNA pathways. ERGO-1 and ALG-3/4 class 26G RNAs are generated by the ERI complex, which includes the RdRP module (RRF-3, DRH-3, ERI-5) in association with DCR-1, RDE-4, ERI-1, and ERI-3 (Duchaine et al., 2006; Gent et al., 2009; Han et al., 2009; Tabara et al., 2002; Thivierge et al., 2012; Vasale et al., 2010). ERI-6/7, ERI-9, and MUT-16 contribute to biogenesis of ERGO-1-class 26G RNAs only (Fischer et al., 2011; Pavelec et al., 2009; Zhang et al., 2011). ERGO-1-class 26G RNAs are 2'O-methylated by HENN-1 (Billi et al., 2012; Kamminga et al., 2012; Montgomery et al., 2012). Both ERGO-1-class and ALG-3/4-class 26G RNAs trigger the biogenesis of WAGO 22G RNAs by the RRF-1 RdRP module (RRF-1, DRH-3, EKL-1) (Aoki et al., 2007; Gu et al., 2009; Han et al., 2009; Vasale et al., 2010). ERGO-1-dependent 22G RNAs also require RDE-10, RDE-11, and Mutator proteins (Fischer et al., 2013; Phillips et al., 2012; Yang et al., 2012; Zhang et al., 2011; 2012). ALG-3/4-class 26G RNAs also act upstream of CSR-1-class 22G RNAs to promote target gene expression (Conine et al., 2013). Figure adapted from (Billi et al., 2014).

11

22G RNAs

The 22G RNAs are comprised of two distinct classes. The WAGO 22G RNAs bind a partially redundant set of 12 worm-specific Argonaute proteins (WAGOs) to target protein-coding genes, pseudogenes, transposons, and other repetitive elements (Ashe et al., 2012; Buckley et al., 2012; Gu et al., 2009; Guang et al., 2008; Ni et al., 2014; Shirayama et al., 2012; Sijen and Plasterk, 2003). A separate worm-specific Argonaute, CSR-1, binds a distinct complement of 22G RNAs targeting germline-expressed genes (Claycomb et al., 2009; Gu et al., 2009). The WAGO-class and CSR-1-class 22G RNAs share several common features including a length of 22nt and bias for a 5'G (Ambros et al., 2003). Like the 26G RNAs, the 22G RNAs are antisense transcripts. Unlike 26G RNAs, which have a 5' monophosphate group, 22G RNAs harbor a triphosphate group on the 5' end, a hallmark of unprocessed RdRP products (Ambros et al., 2003; Gent et al., 2010; Pak and Fire, 2007; Ruby et al., 2006). The biogenesis, processing, and effector function of these pathways are discussed in turn below.

WAGO 22G RNAs

The WAGO 22G RNAs are enriched in the germline, including mature gametes, and embryos (Conine et al., 2010; Han et al., 2009). Biogenesis of the WAGO 22G RNAs requires triggering by primary siRNAs (26G RNAs or exo-RNAi) (Pak and Fire, 2007; Sijen et al., 2001; 2007; Vasale et al., 2010) or by piRNAs (Lee et al., 2012), thus the WAGO 22G RNAs are considered secondary siRNAs. Accordingly, depletion of 26G RNAs or piRNAs, such as by deletions of the genes encoding their biogenesis machinery or Argonaute, also causes depletion of the secondary 22G RNAs (Bagijn et al., 2012; Das et al., 2008; Lee et al., 2012; Vasale et al., 2010). The WAGO 22G RNAs comprise a much larger fraction of germline endo-siRNAs than the 26G RNAs and are not always templated at the exact same position as the corresponding 26G RNAs, indicating that they are not the products of 26G RNA degradation (Ruby et al., 2006;

Vasale et al., 2010). The binding of the primary siRNAs or piRNAs, in association with RISC, to their mRNA targets triggers 22G RNA biogenesis utilizing the targeted mRNA as the biogenesis template (Bagijn et al., 2012; Das et al., 2008; Lee et al., 2012; Pak and Fire, 2007; Sijen et al., 2001; 2007; Vasale et al., 2010). Thus, 22G RNAs facilitate the amplification of 26G RNA- and piRNA-directed gene silencing. Similarly, WAGO 22G RNAs are triggered by exo-RNAi to execute the majority of target repression (Sijen et al., 2001; 2007; Yigit et al., 2006).

Like the 26G RNAs, the WAGO 22G RNAs are made by an RdRP module comprised of an RdRP, a Dicer-related helicase protein, and a Tudor domain protein (Figure 1.2). In this case, the module includes the RdRP RRF-1, DRH-3 (see above), and the Tudor domain protein ELK-1 (Gu et al., 2009). The 5' triphosphate of the 22G RNAs suggests that they are not processed by DCR-1 (Ambros et al., 2003; Gent et al., 2010; Ruby et al., 2006). Furthermore, *in vitro* experiments have shown that 22G RNA accumulation is DCR-1-independent (Aoki et al., 2007).

Amplification of the WAGO 22G RNAs

The recent characterization of the 22G RNA pathway has led to the identification of a novel subcellular structure, the Mutator focus, where many 22G biogenesis factors are concentrated (Phillips et al., 2012; 2014; Zhang et al., 2011). The Mutator foci serve as amplification centers for WAGO 22G RNAs downstream of 26G RNAs, piRNAs, and exo-RNAi. Many of the components of these structures were originally identified in genetic screens for mutants with high rates of germline mutations (Mutator or Mut) (Collins et al., 1987; Ketting et al., 1999; Vastenhouw et al., 2003). The Mutator phenotype arises from increased transposition of DNA transposons such as Tc1 and Tc3 elements (Collins et al., 1987; Ketting et al., 1999; Vastenhouw et al., 2003). Transposon silencing is also piRNA-dependent (Batista et al., 2008; Tabara et al., 1999), thus the increased transposition observed in *mutator* mutants highlights the essential role of the WAGO 22G RNAs in piRNA-mediated transposon silencing (Das et al., 2008; Gu et al., 2009; Lee et al., 2012). Loss of the Mutator focus components is

also associated with resistance to RNAi, indicating their involvement in the amplification of secondary 22G RNAs triggered by exo-RNAi (Ketting et al., 1999; Tabara et al., 1999).

Mutator foci are localized in the nuclear periphery near the nuclear pore and adjacent to P granules and are thought to serve as centers for surveillance of mRNAs as they are exported from the nucleus (Phillips et al., 2012). These foci may also interact with or facilitate the shuttling of transcripts to P granules, which are major centers for post-transcriptional RNA modifications and processing in the *C. elegans* germline (Phillips et al., 2012; Sheth et al., 2010). However, Mutator foci and P granules are not interdependent for their formation or stability (Phillips et al., 2012). Many Mutator components are required for accumulation of not only 22G RNAs, but also ERGO-1-class 26G RNAs (Zhang et al., 2011). ERGO-1 also physically interacts with MUT-16 (Phillips et al., 2014). These findings suggest that Mutator foci may also contribute to 26G RNA biogenesis, processing and/or Argonaute loading.

The components of the Mutator foci include RRF-1, MUT-2, -7, -8, -14, -15, -16, SMUT-1, and RDE-8 (Phillips et al., 2012; 2014; Tsai et al., 2015). Of these, only MUT-16, which contains a prion-like Q/N-rich region, is essential for focus formation (Phillips et al., 2012; 2014). *mut-16(-)* mutants are Rde in the germline and soma, ts sterile, and depleted of WAGO 22G RNAs and some CSR-1 22G RNAs (Phillips et al., 2012; 2014; Zhang et al., 2011). Interestingly, the siRNAs that are present in *mut-16(-)* mutants are enriched for the presence of 3' nontemplated uridines, which are characteristic of CSR-1-class 22G RNAs, suggesting that in the absence of MUT-16 and Mutator foci, WAGO-class 22G RNAs may be misrouted to the CSR-1 pathway (Phillips et al., 2014).

Recent evidence suggests that components of Mutator foci may contribute to cleavage and uridylation of mRNA targets to facilitate RdRP recruitment (Talsky and Collins, 2010; Tsai et al., 2015). Target uridylation precedes 22G RNA synthesis and is dependent on the endoribonuclease RDE-8 (Tsai et al., 2015). Biochemical studies of RDE-8 have primarily focused on its role in exo-RNAi, but based on the WAGO 22G RNA depletion observed in *rde-8(-)* mutants, it may function similarly in endo-siRNA biogenesis (Tsai et al., 2015). Taken together, these studies suggest that RDE-8 is recruited to an mRNA target by primary siRNAs (both primary exo-siRNAs and ERGO-

1-class 26G RNAs) in association with RISC (Tsai et al., 2015). The interaction between RDE-8 and the mRNA target may be stabilized by other components of Mutator foci including MUT-15 (Tsai et al., 2015). RDE-8 then cleaves the mRNA target and the liberated 3' end is uridylated, possibly by the ß-nucleotidyltransferase MUT-2/RDE-3, which then recruits RRF-1 (Tsai et al., 2015). It is interesting to note that while RDE-8 also appears to be required for ERGO-1-class 26G RNA accumulation, its catalytic function is not required, suggesting that its function in 26G accumulation does not involve cleavage of the target mRNA (Tsai et al., 2015).

The functions of the other components of Mutator foci are not well described. MUT-8/RDE-2 is not required for the integrity of Mutator foci, but is needed to recruit MUT-7 to the foci (Phillips et al., 2012). MUT-7 is a 3' to 5' exonuclease. *mut-7(-)* mutants are characterized by Tc1 upregulation, Rde, and ts sterility caused by spermatogenic defects (Ketting et al., 1999; Tabara et al., 1999). MUT-14 and SMUT-1 are DEAD-box helicases that function redundantly to regulate ERGO-1-dependent 22G RNAs, possibly by shuttling mRNAs from P granules to Mutator foci (Phillips et al., 2014; Tijsterman et al., 2002).

The accumulation of WAGO 22G RNAs is also dependent on several factors that are not components of Mutator foci. RDE-10, which has no annotated domains, and RDE-11, a RING-type zinc finger, form a complex (Yang et al., 2012; Zhang et al., 2012). During exo-RNAi, RDE-1 recruits RDE-10 to the target mRNA. RDE-11 then contributes to deadenylation and degradation of the RDE-10-bound targets (Yang et al., 2012; Zhang et al., 2012). How this contributes to secondary siRNA accumulation is not yet clear, but one proposed model is that the RDE-10/11 complex may stabilize the degrading mRNA target to facilitate RRF-1 binding and transcription of the secondary 22G RNAs (Yang et al., 2012). RDE-10 and RDE-11 physically interact with ERGO-1 and are required for accumulation of some ERGO-1-class 26G endo-siRNAs through an as-yet undescribed mechanism (Yang et al., 2012; Zhang et al., 2012).

The binding of RDE-10 to target mRNA is also dependent on the DEAD-box helicase RDE-12 (Yang et al., 2014). Like RDE-10, RDE-12 binds mRNAs that are targeted by RNAi in an RDE-1-dependent manner (Shirayama et al., 2014; Yang et al., 2014). RDE-12 physically interacts with ERGO-1 and WAGO-1 but is only required for

accumulation of 22G RNAs, not 26G RNAs (Fischer et al., 2013; Shirayama et al., 2014; Yang et al., 2014).

RDE-12 colocalizes primarily with P granules (Shirayama et al., 2014; Yang et al., 2014) as well as with RSD-6, which functions in a complex with RSD-2 (Tijsterman et al., 2004; Yang et al., 2014). RSD-2/RSD-6 are dispensable for 26G RNA expression but required for biogenesis and/or maintenance of 22G RNAs (Sakaguchi et al., 2014; Zhang et al., 2012). *rsd-2(-)* and *rsd-6(-)* mutants share a ts progressive sterility (germline mortality or Mrt) phenotype (Sakaguchi et al., 2014). This corresponds with depletion of 22G RNAs targeting transposons and expression of repetitive elements at non-permissive temperature (Sakaguchi et al., 2014). This fertility defect cannot be rescued by mating with wildtype males and thus is not exclusively spermatogenic in origin (Sakaguchi et al., 2014). Therefore, the germline mortality defect of *rsd-2(-)* and *rsd-6(-)* mutants cannot be explained by loss of 26G-dependent 22Gs (because the ALG-3/4 pathway affects spermatogenesis only and the ERGO-1 pathway is dispensable for fertility). Finally, the ATP-binding cassette (ABC) transporter protein HAF-6 is also required for accumulation of ERGO-1-dependent 22G RNAs (Zhang et al., 2012), although its function has yet to be determined.

Worm-specific Argonautes

The 12 WAGOs that function in this pathway are largely redundant. Worms with mutations in all 12 *wago*s (so-called *mago12* mutants) are ts sterile, Him, and Rde; in single mutants, these phenotypes are absent or not fully penetrant (Conine et al., 2010; Gu et al., 2009; Tijsterman et al., 2002; Yigit et al., 2006). Genetic screens for factors involved specifically in siRNA-mediated silencing in the nucleus have identified two specialized WAGOs, WAGO-12/NRDE-3 and WAGO-9/HRDE-1, hereafter called NRDE-3 and HRDE-1 respectively. The specialized functions of HRDE-1 and NRDE-3 are discussed below.

Nuclear RNAi

RNAi targeting of transcripts expressed in the maternal germline has been observed to cause multi-generational gene silencing. For example, *gfp* RNAi of worms expressing a GFP-tagged histone 2B reporter (GFP::H2B) from a transgene under the control of a germline-specific promoter (*pie-1)* results in highly penetrant silencing of *gfp* in F1 and F2 worms that have never been exposed to exogenous *gfp* RNAi (Ashe et al., 2012; Buckley et al., 2012). High-throughput sequencing studies have shown that in the P0 generation grown on *gfp* RNAi, the siRNAs targeting *gfp* are 22nt long with no 5' nucleotide bias, indicating that they are DCR-1-dependent primary siRNAs (Ashe et al., 2012). After four generations with no exposure to exogenous *gfp* RNAi, there are only DCR-1-independent 22G RNAs at the *gfp::h2b* target gene. In addition, the secondary siRNAs have spread upstream and downstream of the original siRNA trigger (Ashe et al., 2012). This indicates that exogenous *gfp* RNAi in the P0 generation triggers the amplification of secondary 22G RNAs in subsequent generations (Ashe et al., 2012). Furthermore, leveraging the ability of 22G RNAs to both spread from the initial trigger and also to silence in *trans*, there is now direct evidence that secondary 22G RNAs can themselves serve as triggers for additional rounds of amplification and spreading, i.e. the generation of *tertiary* siRNAs (Sapetschnig et al., 2015). Taken together, these findings suggest that at each generation, maternally inherited siRNAs are amplified in the progeny and loaded into the zygote, thus transmitting a multigenerational memory of siRNA-mediated gene silencing.

The core machinery that transmits this memory has been identified as the RISC complex defined by the Argonaute WAGO-9/HRDE-1 (named for its heritable RNAi defective phenotype) (Ashe et al., 2012; Buckley et al., 2012; Gu et al., 2012a; Shirayama et al., 2012). The non-Argonaute RISC components NRDE-1, -2, and -4 (named for their nuclear RNAi defective phenotype) are also required for multigenerational siRNA-mediated silencing (Buckley et al., 2012); their roles will be discussed in greater depth in the context of somatic nuclear RNAi below (Figure 1.3). The assembled nuclear RISC mediates target silencing in two ways. First, RISC is guided by bound siRNAs to nascent pre-mRNA, where it induces the stalling of RNA Pol

II (Guang et al., 2010). Second, RISC recruits histone modifying enzymes to mediate the trimethylation of histone 3 at lysine 9 (H3K9) and lysine 27 (H3K27) residues at the corresponding targeted genomic loci; H3K9 and H3K27 methylation are hallmarks of heterochromatin (Burkhart et al., 2011; Burton et al., 2011; Guang et al., 2010; Mao et al., 2015). The siRNA-induced H3K9me3 marks can extend far beyond the footprint of the siRNA itself, potentially up to 9kb from the target sequence (Gu et al., 2012a).

Recent studies have elucidated some of the chromatin factors that act downstream of nuclear RNAi. Methylation of H3K9 is catalyzed by the histone methyltransferases SET-25 and MET-2 and also requires the H3K9me3 reader HPL-2, a homolog of the mammalian H3K9 methyl reader HP1 (Ashe et al., 2012; Mao et al., 2015). Methylation of H3K27 is dependent on MES-2 and MES-6 (Mao et al., 2015). Together with MES-3, MES-2 and MES-6 form a complex orthologous to the *Drosphila* Polycomb repressive complex 2 (PRC2). The H3K36 histone methyltransferase MES-4, an established regulator of MES-2, -3, -6 activity, is also required for H3K27 methylation downstream of RNAi (Gaydos et al., 2012; Mao et al., 2015). Interestingly, in *hrde-1(-)* mutants, target genes exhibit progressive loss of H3K9me3 over generations, suggesting that nuclear RNAi is required for maintenance as well as establishment of H3K9me3 at target genes (Buckley et al., 2012).

The memory of gene silencing transmitted by germline nuclear RNAi is critical for proper germline development. *hrde-1(-)* and *nrde-1(-), -2(-), and -4(-)* mutants all exhibit a Mrt phenotype (Buckley et al., 2012). In *hrde-1(-)* mutants, the Mrt phenotype is ts, occurring only at 25°C (Buckley et al., 2012). The precise cause of the *hrde-1(-)* Mrt phenotype is unknown, as late generation *hrde-1(-)* mutants display an array of germline defects affecting both sperm and oocytes (Buckley et al., 2012). Interestingly, the *hrde-1(-)* Mrt phenotype is reversible: after several generations of growth at the non-permissive temperature, fertility can be restored by downshifting to a permissive temperature (Ni et al., 2016). This indicates that the underlying cause of the Mrt phenotype must also be reversible. Nuclear RNAi likely regulates germline maintenance in the same pathway as *rsd-2* and *rsd-6*, as the Mrt phenotype of *rsd-2(-)* and *rsd-6(-)* mutants is not enhanced by *nrde-2(-)* (Sakaguchi et al., 2014)*.

In somatic tissues, nuclear RNAi utilizes the same machinery with the exception of HRDE-1, which is expressed exclusively in the germline (Buckley et al., 2012; Burkhart et al., 2011; Burton et al., 2011; Guang et al., 2010). The WAGO for somatic nuclear RNAi is NRDE-3 (Guang et al., 2008). NRDE-3 binds 22G RNAs in the cytoplasm and then translocates to the nucleus and targets nascent pre-mRNAs (Guang et al., 2008). This association recruits NRDE-2 (Guang et al., 2010). Both NRDE-3 and NRDE-2 are required to recruit NRDE-1 and NRDE-4 (Burkhart et al., 2011). NRDE-1 binds the target pre-mRNA and is the only nuclear RISC component that has been shown to directly bind chromatin (Burkhart et al., 2011). Interestingly, NRDE-1 requires NRDE-4 to bind chromatin but not pre-mRNA (Burkhart et al., 2011). This suggests that NRDE-1 is recruited to NRDE-2 and NRDE-3-bound pre-mRNA, and, in turn, recruits NRDE-4 to the complex, which facilitates NRDE-1 binding to the genomic target, thus recruiting chromatin modifying enzymes described above to direct methylation of H3K9 and H3K27 (Burkhart et al., 2011; Burton et al., 2011; Mao et al., 2015). It is likely NRDE-1, -2, and -4 play analogous roles in germline nuclear RNAi, but this has not yet been explicitly demonstrated.

Because nuclear RISC functions exclusively in the nucleus, it is dispensable for silencing of cytoplasmic mRNAs (Guang et al., 2008). Thus, *nrde* mutants do not exhibit classic cytoplasmic Rde phenotypes, as they are exclusively nuclear RNAi defective. Nrde phenotypes can be elucidated using exogenous RNAi targeting a polycistronic nuclear pre-mRNA and looking for a phenotype indicative of nuclear, but not cytoplasmic, silencing. For example, *lir-1* is expressed in an operon with the essential gene *lin-26*. Because *lir-1* is dispensable for development, loss of cytoplasmic *lir-1* mRNA does not cause any discernable phenotype. Loss of the polycistronic pre-mRNA containing both *lir-1* and *lin-26* by exogenous RNAi of *lir-1* in the nucleus results in larval arrest and lethality in wildtype worms but not in *nrde* mutants (Guang et al., 2010).
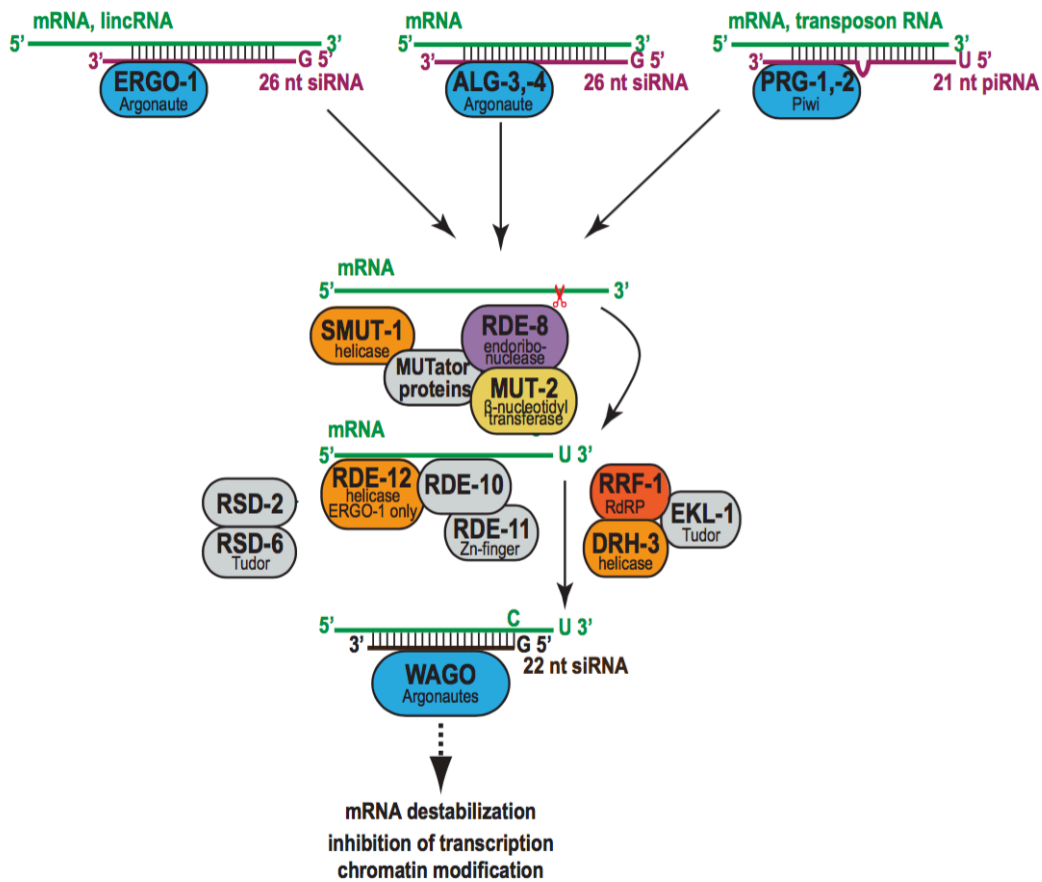
Figure 1.2. Overview of the WAGO 22G RNA pathway. 22G RNA biogenesis is triggered by 26G RNAs and piRNAs (Ashe et al., 2012; Bagijn et al., 2012; Luteijn et al., 2012; Shirayama et al., 2012; Vasale et al., 2010) . In association with the Mutator focus, the target mRNA is cleaved and 3' uridylated, recruiting the RRF-1 RdRP module to transcribe the 22G RNAs (Gent et al., 2010; Phillips et al., 2012; Tsai et al., 2015). 22G RNA biogenesis also requires RDE-10,-11,-12 and RSD-2,-6 (Fischer et al., 2013; Sakaguchi et al., 2014; Yang et al., 2012; 2014; Zhang et al., 2012). Figure adapted from (Billi et al., 2014).
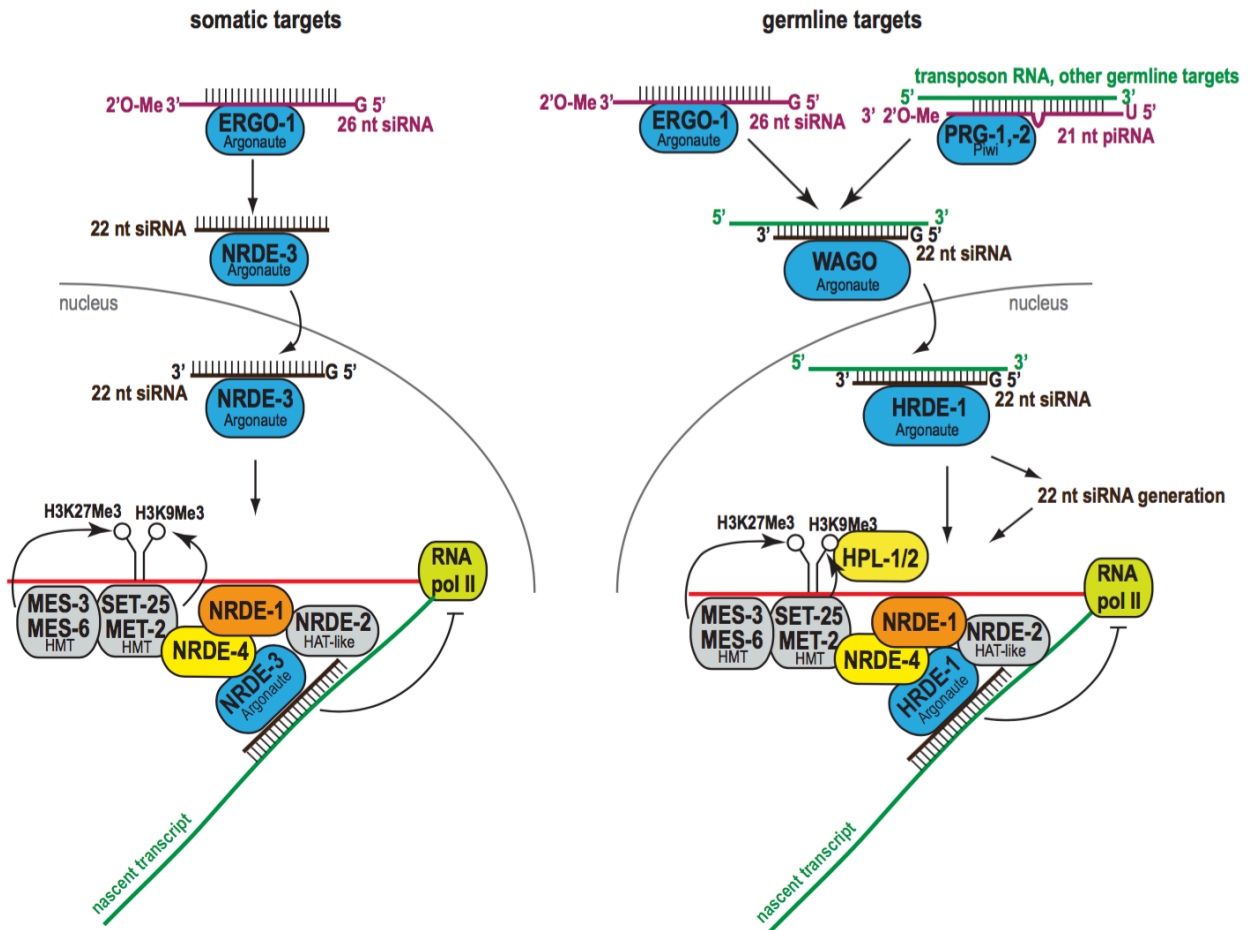
Figure 1.3. Overview of the nuclear RNAi pathways. Nuclear RNAi is mediated by NRDE-3 in the cytoplasm (Guang et al., 2008) and HRDE-1 in the germline (Buckley et al., 2012). Other members of the nuclear RISC are NRDE-1,-2,-4 (Burkhart et al., 2011; Burton et al., 2011; Guang et al., 2010). Both nuclear RISC complexes associate with nascent pre-mRNAs (Buckley et al., 2012; Guang et al., 2010) and direct H3K9 methylation by SET-25 and MET-2 (Ashe et al., 2012; Buckley et al., 2012; Gu et al., 2012a) and H3K27 methylation by MES-3,-6 (Mao et al., 2015). Figure adapted from (Billi et al., 2014).

CSR-1 22G RNAs

Although CSR-1 is also a worm-specific Argonaute protein, it functions differently from the other WAGOs and therefore defines its own class of 22G RNAs (Figure 1.4). One key difference between CSR-1 and WAGO pathways is that the CSR-1-class 22G RNAs are not secondary siRNAs and therefore their biogenesis is not dependent on the 26G RNAs or piRNAs. While some CSR-1-class 22G RNAs share common targets with the ALG-3/4 26G RNAs (Conine et al., 2013), there is no direct evidence to suggest that they are secondary 22G RNAs. Many of the biogenesis factors for CSR-1-class 22G RNAs have been identified, but how they are recruited to a specific set of mRNA templates is unknown. The existing evidence suggests that CSR-1 plays a number of important roles, including modulating germline gene expression, by both gene licensing and gene silencing, and regulating chromatin organization to promote correct chromosome segregation in the early embryo (Avgousti et al., 2012; Cecere et al., 2014; Claycomb et al., 2009; Gerson-Gurwitz et al., 2016; Seth et al., 2013; van Wolfswinkel et al., 2009; Wedeles et al., 2013).

Biogenesis of CSR-1 22G RNAs

CSR-1-class 22G RNAs are produced in the germline by an RdRP module comprised of the RdRP EGO-1, EKL-1, and DRH-3 (Claycomb et al., 2009). These 22G RNAs target germline-expressed genes, a feature that is conserved in the CSR-1 pathway of the sister species *C. briggsae* (Tu et al., 2015).

While CSR-1 and the WAGOs bind distinct 22G RNAs, how the 22G RNAs are sorted to the appropriate Argonaute is an outstanding question in the field. One possibility is that post-transcriptional modifications of the 22G RNAs may direct loading to a specific RISC. One feature that may help distinguish CSR-1 22G RNAs from WAGO 22G RNAs is the presence of at least one non-templated uridine at the 3' end of many CSR-1 22G RNAs (Claycomb et al., 2009). Importantly, 60% of CSR-1-bound 22G RNAs are not uridylated, therefore 3' uridylation is not necessary to direct loading

onto CSR-1 (Claycomb et al., 2009; van Wolfswinkel et al., 2009). 3' uridylation of CSR-1-class 22G RNAs is catalyzed by the ß-nucleotidyltransferase CDE-1 (van Wolfswinkel et al., 2009). Interestingly, *cde-1(-)* mutants express increased levels of endo-siRNAs compared to wildtype, suggesting that 3' uridylation may function to destabilize 22G RNAs, perhaps to maintain the relative low abundance of CSR-1-class 22G RNAs compared to the WAGO-class 22G RNAs (Claycomb et al., 2009; Gu et al., 2009; van Wolfswinkel et al., 2009). In fact, endo-siRNAs that are upregulated in the *cde-1(-)* mutant correspond with downregulation of their target genes (van Wolfswinkel et al., 2009). Most evidence suggests that CSR-1-class 22Gs do not play a major role in target silencing (Claycomb et al., 2009), therefore the observed target downregulation observed in *cde-1(-)* mutants may indicate that the stabilized CSR-1-class 22G RNAs are misrouted to the WAGO pathway to mediate target silencing (van Wolfswinkel et al., 2009). Alternatively, recent evidence suggests that CSR-1 mediates the silencing of genes targeted by very abundant 22G RNAs (Gerson-Gurwitz et al., 2016), thus the increased abundance of the non-uridylated 22G RNAs in *cde-1(-)* mutants may promote more CSR-1-mediated target silencing.

With the exception of EKL-1, the major components of the CSR-1 pathway, CSR-1, DRH-3, EGO-1, and CDE-1, all localize to perinuclear P granules (Claycomb et al., 2009; Gerson-Gurwitz et al., 2016; van Wolfswinkel et al., 2009). While EKL-1 has not been detected in P granules, it is essential for P granule integrity, as are EGO-1, DRH-3, and CSR-1 (Claycomb et al., 2009; Vought et al., 2005). In mature gametes and embryos, CSR-1 localizes to the nucleus (Claycomb et al., 2009). Biochemical studies have shown that CSR-1 binds to nascent pre-mRNAs, chromatin, and RNA Pol II (Cecere et al., 2014; Claycomb et al., 2009; Wedeles et al., 2013) and can be immunopurified from purified sperm and oocyte chromatin (Chu et al., 2006). However, a recent study utilizing a single-copy CSR-1 transgene detected its localization at P granules but no association with mitotic chromosomes in embryos (Gerson-Gurwitz et al., 2016).

The *csr-1(-)* mutant is Rde and sterile; the few embryos that are produced by *csr-1(-)* mutants are inviable due to chromosome segregation defects (Yigit et al., 2006). Mutation or knock-down of the other CSR-1 pathway components (e.g. *drh-3, elk-1,* and

cde-1) leads to similar phenotypes, although in *cde-1(-)* mutants the defects are less severe (Claycomb et al., 2009; van Wolfswinkel et al., 2009). The details of these chromosome segregation defects are discussed below.

CSR-1 tunes germline gene expression and regulates chromatin organization

The initial characterization of the CSR-1 pathway found that its targets were largely unaffected or slightly downregulated in *csr-1(-)* mutants compared to wildtype (Claycomb et al., 2009). For example, a majority of the genes encoding histones are downregulated in the *csr-1(-)* mutant (Avgousti et al., 2012). Additionally, global nuclear run-on followed by high-throughput sequencing (GRO-seq) studies have demonstrated decreased occupancy of RNA Pol II at CSR-1 target genes in both *csr-1* partial loss-of-function, or hypomorphic, mutants and *drh-3(-)* mutants (Cecere et al., 2014). As RNA Pol II occupancy at WAGO targets is unaffected in the *drh-3(-)* mutant, this effect is thought to be specific to the CSR-1-class 22G RNAs (Cecere et al., 2014). Taken together, these findings suggest that CSR-1 acts co-transcriptionally to positively regulate its targets.

Further study of the CSR-1 pathway in the germlines of male worms supports the model of CSR-1 positively regulating its targets. By immunostaining, *csr-1(-)* mutant males show decreased levels of RNA Pol II and H3K4me2 in the germline, similar to *alg-3(-);alg-4(-)* double mutants (Conine et al., 2013). While CSR-1-class 22G RNAs are not thought to be triggered by 26G RNAs, CSR-1 appears to function downstream of ALG-3/4 in the male germline, targeting over 80% of ALG-3/4 target genes (Conine et al., 2013). The precise mechanism by which the ALG-3/4 pathway regulates CSR-1 activity is unclear. One possibility is that ALG-3/4 recruits CSR-1 to its nuclear targets. In support of this model, CSR-1 expression is retained at P granules but lost at chromatin in *alg-3(-);alg-4(-)* double mutants, suggesting that ALG-3/4 are dispensable for CSR-1-class 22G RNA biogenesis and loading onto CSR-1 RISC, as both of these steps occur at the P granule (Conine et al., 2013). CSR-1 also binds the genomic targets of ALG-3/4 26G RNAs and this binding is *alg-3/4*-dependent (Conine et al.,

2013). It is important to note that studies in hermaphrodites have shown that *csr-1* RNAi causes upregulation of spermatogenesis genes, suggesting that in the hermaphrodite, CSR-1 contributes to gene silencing (Campbell and Updike, 2015). This could indicate that CSR-1 positively regulates targets in males but negatively regulates them in hermaphrodites. Alternatively, the observed upregulation of spermatogenesis genes upon RNAi knockdown of *csr-1* could be an artifact of disrupted P granules, as P granule disruption is a consequence of loss of the CSR-1 pathway (Campbell and Updike, 2015). P granule disruption is also associated with upregulation of spermatogenesis genes independent of the CSR-1 pathway (Campbell and Updike, 2015).

Further evidence for gene licensing by CSR-1 has been described in the context of piRNA-mediated multigenerational silencing and is discussed in the piRNA section below (Seth et al., 2013; Wedeles et al., 2013). Taken together, these data support a model in which CSR-1 promotes the expression of germline genes.

The existing evidence suggests that CSR-1 acts to repress gene expression by two distinct mechanisms. First, CSR-1 has an intact catalytic triad and is capable of slicing its targets *in vitro* (Aoki et al., 2007). The recent finding that many 22G RNA target genes are upregulated in worms expressing CSR-1(SIN), a slicer-dead CSR-1 variant, compared to wildtype CSR-1, indicates that (1) CSR-1 slices target mRNAs *in vivo* and (2) CSR-1-mediated slicing contributes to target gene silencing (Gerson-Gurwitz et al., 2016). The severity of target upregulation in the CSR-1(SIN) mutant correlates strongly with the density of corresponding 22G RNA reads from CSR-1 immunopurification (Gerson-Gurwitz et al., 2016). This suggests that CSR-1 may preferentially slice highly expressed genes (Gerson-Gurwitz et al., 2016).

A second CSR-1-mediated silencing mechanism may explain how CSR-1 silences some genes that are expressed at very low levels in wildtype worms. In a *csr-1* hypomorphic mutant, RNA Pol II occupancy is increased at weakly expressed genes (Cecere et al., 2014). This is consistent with a subset of targets that are upregulated in the CSR-1(SIN) mutant but are not normally highly expressed or germline-enriched; interestingly, these targets do not correspond to identified CSR-1-bound 22G RNAs (Gerson-Gurwitz et al., 2016). How could this potentially 22G RNA-independent

silencing of low-expressing genes be achieved? One explanation is that CSR-1 may be important for the integrity of heterochromatin in the vicinity of its targets and that loss of CSR-1 impairs heterochromatin formation or stability leading to ectopic expression of normally repressed transcripts. In support of this model, both imaging and biochemical studies have shown that the centromeric histone variant CENP-A is disorganized in *csr-1(-)* or hypomorphic mutants, indicative of a heterochromatin defect (Cecere et al., 2014; Claycomb et al., 2009). In *csr-1* hypomorphic mutants, normally silent loci that flank CSR-1 target genes lose their enrichment for two repressive histone marks, H3K27me3 and CENP-A, and are upregulated (Cecere et al., 2014). Taken together, these findings suggest that CSR-1 regulates the chromatin landscape in the vicinity of its targets to maintain the silenced, heterochromatic state. Loss of these repressive marks in *csr-1* hypomorphic mutants and *drh-3(-)* mutants may explain the global increase in antisense transcription observed in these mutants (Cecere et al., 2014).

In addition to its contribution to CENP-A and H3K27me3 organization, CSR-1 is required for correct loading of condensins and cohesins onto mitotic chromosomes in early embryos (Claycomb et al., 2009). The widespread disorganization of these factors is thought to cause the chromosome segregation defects, such as anaphase bridging and accumulation of abnormal nuclei, observed in *csr-1(-)* mutants (Claycomb et al., 2009). While the consequences of this misregulation is most apparent in early embryos, where defects arise at the first mitotic division and continue until embryonic arrest at the 50-cell stage (Claycomb et al., 2009), they are not limited to the early embryo. In male *csr-1(-), drh-3(-),* and *ekl-1(-)* mutants, H3K9me2 is more broadly distributed along meiotic chromosomes compared to wildtype (She et al., 2009). Meiotic chromosomes also show decreased H3K4me2 staining and aberrant overlap of H3K9me2 and H3K4me2 marks, suggesting that the boundaries between heterochromatin and euchromatin are blurred (She et al., 2009). Loss of heterochromatin-euchromatin boundaries may also contribute to the finding that male *csr-1(-)* and *alg-3/4(-)* mutant spermatocytes have more condensed nuclei than wildtype spermatocytes (Conine et al., 2013), although there are no reports on the distribution of heterochromatic marks in male *csr-1(-)* mutants.
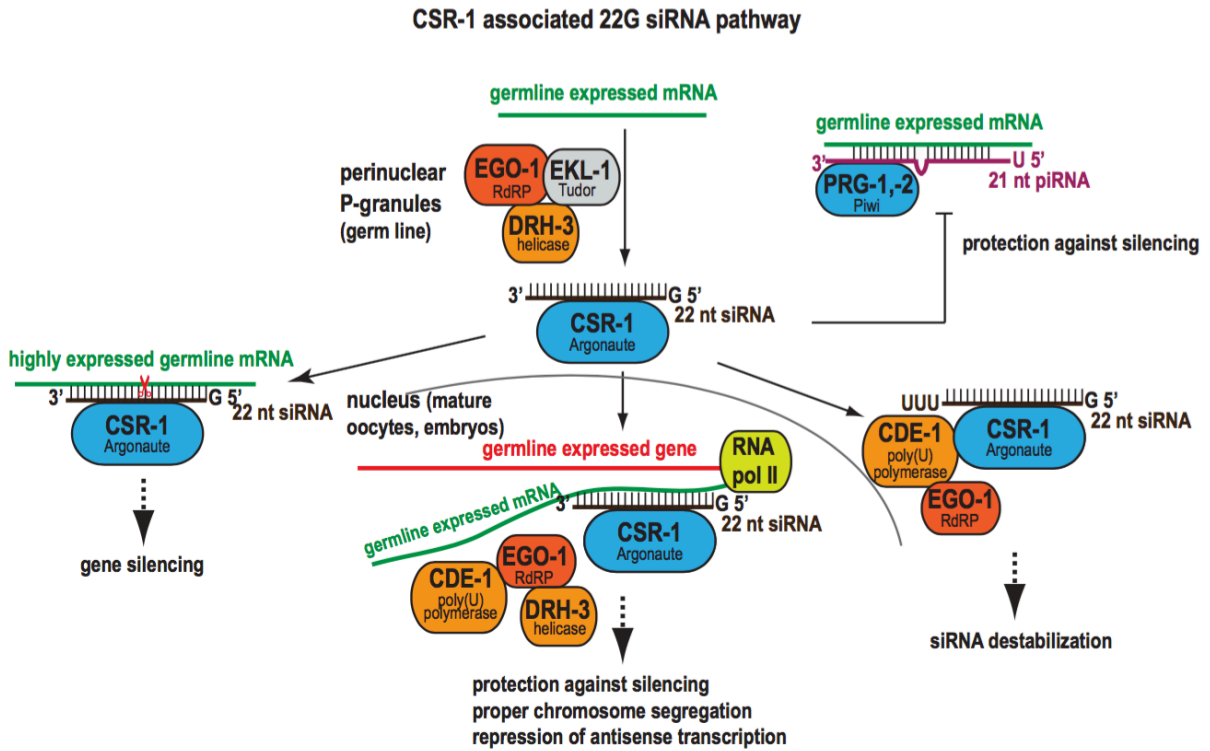
Figure 1.4. Overview of the CSR-1 22G RNA. CSR-1 22G RNAs are transcribed by the EGO-1 RdRP module (Claycomb et al., 2009). CSR-1 silences highly expressed germline mRNAs by slicing (Aoki et al., 2007; Gerson-Gurwitz et al., 2016). In the nucleus, CSR-1 associates with nascent transcripts to license gene expression and repress antisense transcription (Cecere et al., 2014; Claycomb et al., 2009; Seth et al., 2013; Wedeles et al., 2013). Some CSR-1 22G RNAs are destabilized via uridylation by CDE-1 (van Wolfswinkel et al., 2009). Figure adapted from (Billi et al., 2014).

21U piRNAs

The *C. elegans* piRNAs comprise a large class of germline-enriched sense and antisense 5' monophosphorylated RNAs with a characteristic length of 21nt and a bias for a 5' uridine (Batista et al., 2008; Ruby et al., 2006). Over 15,000 unique piRNAs are transcribed as autonomous units from broad regions on chromosome IV (Billi et al., 2013; Cecere et al., 2012; Ruby et al., 2006). Each piRNA transcriptional unit is generally AT rich and nucleosome depleted (Cecere et al., 2012; Ruby et al., 2006). In metazoans, the expression of piRNAs in the germline is a common strategy for transposon repression (Aravin et al., 2006; Batista et al., 2008; Brennecke et al., 2007; Das et al., 2008; Girard et al., 2006; Grivna et al., 2006a; 2006b; Yigit et al., 2006). Transposon repression is critical for germline function, thus *C. elegans* that are depleted of piRNAs have reduced brood sizes compared to wildtype worms and are ts sterile at 25°C (Batista et al., 2008; Das et al., 2008; Wang and Reinke, 2008). While transposon repression is a common function of metazoan piRNAs, the mechanism for repression in *C. elegans* is unique; piRNAs are required to initiate target silencing, but the primary effectors of piRNA-directed silencing are the downstream secondary 22G RNAs (Ashe et al., 2012; Bagijn et al., 2012; Lee et al., 2012; Luteijn et al., 2012; Shirayama et al., 2012).

Biogenesis of piRNAs

While the piRNA sequences are not themselves conserved in other nematodes, the clusters are syntenic with the 21U RNA-encoding clusters in other *Caenorhabditis* species (de Wit et al., 2009; Ruby et al., 2006). Also conserved is the enrichment of an eight nucleotide core motif: CTGTTTCA, known as the Ruby motif (Batista et al., 2008; Ruby et al., 2006). This motif, separated from the piRNA by an AT-rich spacer of 20-60nt, is required to drive piRNA expression (Billi et al., 2013; Cecere et al., 2012; Ruby et al., 2006). The conservation of the Ruby motif in other *Caenorhabditis* species suggests that the regulatory mechanisms that control the piRNA pathway are likely

conserved, even though the actual piRNA sequences are not. The precise role of the Ruby motif has not yet been described, although it has been proposed to be a transcription factor binding site (Cecere et al., 2012). The Ruby motif also contributes to sex-specific expression of piRNAs, as the 5' cytosine of the motif is enriched upstream of piRNAs that are expressed during spermatogenesis, whereas piRNAs expressed during oogenesis have no nucleotide bias at that position (Billi et al., 2013). While many piRNAs are expressed sex-specifically, enriched either in spermatogenic or oogenic germline, no differences in target selection or effector function between the two groups have been identified.

piRNAs are transcribed by RNA Pol II as a 23-30nt precursor species (Figure 1.5) during the fourth larval stage and into adulthood, when germline expansion and maturation are most active (Billi et al., 2013; Cecere et al., 2012; Gu et al., 2012b; Weick et al., 2014). piRNA transcription requires the transcription factor GEI-11/SNPC-4, which binds throughout the piRNA clusters (Goh et al., 2014; Kasper et al., 2014). *gei-11(-)* mutants are depleted for piRNAs and sterile (Kasper et al., 2014). As GEI-11 binding strength does not correlate with piRNA expression level, it is not thought to act as a canonical transcription factor at these loci (Kasper et al., 2014). GEI-11 colocalizes with PRDE-1, a casein kinase family member (Kasper et al., 2014). PRDE-1 is required for accumulation of piRNA precursors, indicating that it acts upstream of the piRNA Argonaute, PRG-1 (Weick et al., 2014). PRDE-1 is expressed exclusively in the germline where it localizes specifically to chromosome IV (Weick et al., 2014). Also required for piRNA precursor accumulation are three uncharacterized factors that were identified in a screen for piRNA biogenesis factors or Twenty-One-U's Fouled Up (TOFU): TOFU-3, -4, and -5 (Goh et al., 2014).

A number of recently identified factors are involved in piRNA processing. Like the ERGO-1 class of 26G RNAs, piRNAs are 2' O-methylated by HENN-1 which contributes to their stability and facilitates piRNA maintenance in the developing embryo (Billi et al., 2012; Kamminga et al., 2012; Montgomery et al., 2012). The exonuclease PARN-1 is required for 3' trimming of piRNAs (Tang et al., 2016). In *parn-1(-)* mutants, the untrimmed piRNAs are stably expressed, 2' O-methylated, and bound by PRG-1, indicating that 3' trimming is not required for piRNA stability, methylation, or Argonaute

loading (Tang et al., 2016). As *parn-1(-)* mutants have decreased fertility compared to wildtype and are defective for piRNA-dependent gene silencing, 3' trimming is likely required for piRNA effector function, perhaps by recruiting additional components of the silencing machinery (Tang et al., 2016).

Two uncharacterized TOFU factors, TOFU-1 and TOFU-2, are required for accumulation of mature piRNAs (Goh et al., 2014). In *tofu-1(-)* and *tofu-2(-)* mutants, piRNA precursors accumulate over three-fold compared to wildtype, suggesting that they are involved in piRNA maturation (Goh et al., 2014). PID-1, which has no annotated domains other than putative nuclear localization and export signals, is also required for accumulation of mature piRNAs but not for piRNA precursors, suggesting a role for PID-1 in piRNA maturation or stability (de Albuquerque et al., 2014).

The Piwi Argonaute PRG-1 is essential for fertility at 25°C and silencing of Tc3 transposons (Batista et al., 2008; Das et al., 2008; Yigit et al., 2006). Worms expressing a putative slicer-dead PRG-1 variant have reduced fertility, suggesting that PRG-1 slicing contributes to its function; however, *prg-1*-dependent secondary 22G RNA biogenesis and target silencing are largely slicing-independent (Lee et al., 2012).

piRNA effector function and crosstalk with endo-siRNAs

Here I will discuss two emerging models of piRNA function in the *C. elegans* germline. These models highlight the importance of cross-regulation between the piRNA pathway and both WAGO-class and CSR-1-class 22G RNAs. First, I will discuss the role of piRNAs in establishing extremely stable transgene silencing that is maintained by the WAGO 22G RNAs and opposed by the CSR 22G RNAs. Second, I will discuss two recent studies that have elucidated a role for piRNAs in the correct allocation of 22G RNAs into WAGO and CSR-1 pathways.

piRNAs establish multigenerational transgene silencing

Recent studies have described an extremely stable, completely penetrant form of transgene silencing termed RNA induced epigenetic silencing (RNAe) (Luteijn et al., 2012; Shirayama et al., 2012). Once RNAe is initiated, it is thought to be permanently maintained over generations, as spontaneous reversion has never been observed (Luteijn et al., 2012; Shirayama et al., 2012). RNAe requires PRG-1, and is thus considered piRNA-mediated (Luteijn et al., 2012; Shirayama et al., 2012). Interestingly, once a transgene has been silenced by RNAe, *prg-1* is dispensable for both gene silencing and the expression of piRNA-triggered secondary 22G RNAs (Luteijn et al., 2012; Shirayama et al., 2012). Maintenance of silencing requires *mut-7, rde-3, hrde-1,* and *wago-10*, but not *rde-1*, indicating that it is mediated by the WAGO class of 22G RNAs (Luteijn et al., 2012; Shirayama et al., 2012). The loss of silencing observed in both *hrde-1(-)* and *wago-10(-)* mutants is not completely penetrant, suggesting that they may act redundantly (Luteijn et al., 2012; Shirayama et al., 2012). Both pre-mRNA and mRNA of the silenced transcript are downregulated during RNAe, indicating that this repression is co-transcriptional (Luteijn et al., 2012; Shirayama et al., 2012). There is also a chromatin component to RNAe, as RNAe induces *hrde-1/nrde-1*-dependent H3K9me3 deposition at the target and requires *mes-3, mes-4,* and *hpl-2* (Luteijn et al., 2012; Shirayama et al., 2012).

When RNAe is inherited through the maternal germline, it can act *in trans* to suppress active transgenes, indicating that the secondary 22G RNAs are inherited (Luteijn et al., 2012). Interestingly, while deep sequencing of these secondary 22G RNAs shows that they spread upstream of the original piRNA trigger, they are limited to parts of the transgene that encode foreign sequences such as *gfp* and do not target the transgene-borne endogenous sequences (Luteijn et al., 2012; Shirayama et al., 2012). Furthermore, RNAe-mediated co-suppression of other sequences *in trans* is limited to other foreign sequences and does not occur at endogenous genes, indicating that endogenous sequences are likely protected from RNAe-targeting (Shirayama et al., 2012).

Some expressed transgenes are capable of mediating RNA-induced transactivation (RNAa) to desilence sequences *in trans* (Shirayama et al., 2012). While the sequence features that allow this transactivation are unknown, it is dependent on the CSR-1 22G RNA pathway, as indicated by the identification of 22G RNAs corresponding to the activating allele and the finding that RNAa does not occur on *csr-1* RNAi (Seth et al., 2013). The activation of an RNAe allele by RNAa can also be transmitted over generations, but this requires multiple generations of exposure to the RNAa allele (Seth et al., 2013). RNAa is heritable but not permanent, as trans-activated RNAe alleles always revert back to the silenced state even after multiple generations of RNAa-mediated activation (Seth et al., 2013). As with *de novo* establishment of RNAe, this re-establishment requires *prg-1* (Seth et al., 2013)*.* These data suggest that CSR-1 acts in opposition to piRNAs to promote expression of its targets. The recent use of the phage λ–BoxB system to tether CSR-1 to target transcripts has shown that CSR-1 tethering can protect a transcript from silencing by RNAe and that with multiple generations of CSR-1-tethering, a silenced allele can be activated (Wedeles et al., 2013). Taken together, these studies suggest that licensing by CSR-1 protects transcripts from silencing by piRNAs to promote the expression of germline genes in the developing embryo (Wedeles et al., 2013). As both CSR-1 and PRG-1 function primarily at germline P granules, these may be the sites of mRNA surveillance and sorting into endogenous or "self" transcripts that are CSR-1-bound and foreign or "non-self" transcripts that are targeted by piRNAs and shuttled to adjacent Mutator foci to template 22G RNA amplification.

piRNAs differentiate WAGO-class and CSR-1-class 22G RNAs

To what extent are studies of RNAe and RNAa applicable to endogenous sequences? While this has yet to be fully addressed, two recent studies suggest that piRNAs contribute to the establishment of a genetic memory between "self" and "non-self" that is transmitted to the next generation by WAGO-class and CSR-1-class 22G RNAs. By crossing mutants of two different WAGO 22G RNA pathway components, de

Albuquerque *et al.* assessed *de novo* 22G RNA establishment in the F1 generation. For example, *mut-7(-/-)* and *mut-16(-/-)* homozygous mutants are strongly depleted of WAGO 22G RNAs, but their cross progeny, which are heterozygous *mut-7(+/-);mut-16(+/-),* are able to accumulate WAGO 22G RNAs (de Albuquerque et al., 2015). Phillips *et al.* took a similar approach, crossing *mut-14(-/-);smut-1(-/-)* homozygous mutants with *mut-16(-/-)* mutants (Phillips et al., 2015). For clarity, I will refer to both approaches interchangeably as "22G RNA resetting". While 22G RNA resetting in a *prg-1(+/+)* background results in about 30% sterility in the F1 generation, 22G RNA resetting in a piRNA-depleted mutant background (i.e. *prg-1(-/-)* or *pid-1(-/-)*) results in 100% sterility in the F1 generation (de Albuquerque et al., 2015; Phillips et al., 2015). This indicates that resetting the WAGO 22G RNAs in the absence of piRNAs is detrimental for fertility. Interestingly, if the same experiments are performed while the WAGO pathway is compromised in the F1 generation, such as by growth on *mut-16* RNAi or in a *hrde-1(-/-)* background, the sterility from resetting in the *prg-1(-/-)* background can be partially rescued (de Albuquerque et al., 2015; Phillips et al., 2015). This indicates that it is the reintroduction of WAGO 22G RNAs in the absence of piRNAs that causes F1 sterility. Immunopurification of HRDE-1 and WAGO-1 after 22G RNA resetting without piRNAs shows that these WAGOs are enriched for both WAGO-class and CSR-1-class 22G RNAs (de Albuquerque et al., 2015; Phillips et al., 2015). After resetting in a *prg-1(+/+)* background, however, immunopurified HRDE-1 is enriched for WAGO-class 22G RNAs only (Phillips et al., 2015). These studies provide compelling evidence that piRNAs establish a "self" versus "non-self" distinction for germline transcripts that is enforced by the 22G RNAs.
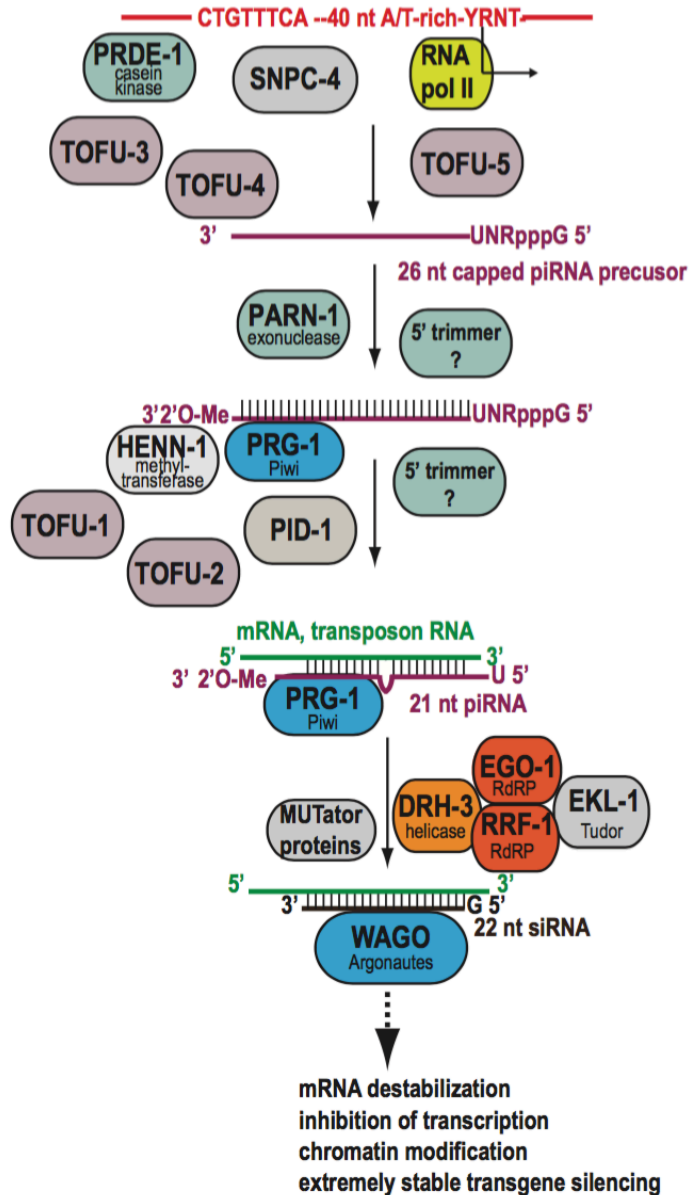
Figure 1.5. Overview of the 21U piRNA pathway. piRNAs are transcribed as 23-30nt precursors by RNA Pol II (Gu et al., 2012b). piRNA transcription also requires SNPC-4, PRDE-1, and TOFU-3,-4,-5 (Goh et al., 2014; Kasper et al., 2014; Weick et al., 2014). piRNA maturation requires trimming at the 5' end by an unknown factor. At the 3' end, piRNA precursors are trimmed by PARN-1 (Tang et al., 2016) and 2' O-methylated by HENN-1 (Billi et al., 2012; Kamminga et al., 2012; Montgomery et al., 2012). In association with PRG-1 RISC, piRNAs bind their targets with imperfect complementarity to trigger the production of secondary WAGO-class 22G RNAs (Das et al., 2008; Lee et al., 2012). Figure adapted from (Billi et al., 2014).

REMAINING QUESTIONS

While many of the protein cofactors for germline small RNA pathways have been elucidated as described above, there are many unanswered mechanistic questions. How target selection is achieved is still a major gap in the field, as we do not know why some mRNAs serve as 26G RNA or CSR-1-class 22G RNA templates and some do not. How 26G RNAs or piRNAs actually trigger 22G RNA amplification is also unknown. As piRNAs appear to bind targets with imperfect complementarity (Lee et al., 2012), a robust target set for this pathway has not yet been defined. Finally, untangling how all the classes of germline small RNAs act together to regulate gene expression and how they relate to other epigenetic mechanisms such as chromatin organization is currently a major focus of the field.

Additionally, recent studies have highlighted the contributions of small RNA pathways to transgenerational inheritance. Transgenerational effects of small RNAs include the triggering of *de novo* siRNA biogenesis by maternal inherited siRNAs and piRNAs and the establishment and multigenerational maintenance of chromatin modifications downstream of nuclear RNAi. How siRNAs enforce an epigenetic memory of gene silencing that is remembered at the chromatin level is unknown, but the finding that the progressive loss of H3K9 methylation in nuclear RNAi mutants is concurrent with progressive sterility highlights the biological importance of this epigenetic memory.

The major goal of my thesis work is to understand how siRNAs regulate chromatin structure and how that regulation is transmitted to the next generation. The majority of my work has focused on the role of *morc-1,* the *C. elegans* Microrchidia homolog, in the regulation of nuclear RNAi effector function at the chromatin level. In Chapter Two, I discuss (1) the characterization of *morc-1* and its role in integrating nuclear RNAi and transgenerational chromatin architecture and (2) the identification of mutations in *met-1*, which encodes a histone methyltransferase, that suppress the transgenerational fertility defect of *morc-1(-)* and nuclear RNAi mutants. In Chapter Three, I discuss ongoing studies and future directions to further elucidate the role of *morc-1* in transgenerational epigenetic inheritance. In Chapter Four, I discuss ongoing studies regarding the epigenetic regulation of small RNA biogenesis.

THE MICRORCHIDIA GHKL ATPases

The GHKL (gyrase, Hsp90, histidine kinase, MutL) ATPase domain is shared by a diverse set of prokaryotic and eukaryotic enzymes that function as molecular clamps to manipulate their substrates. These enzymes share a common structure, the Bergerat fold, but are divergent at the animo acid level (Dutta and Inouye, 2000). The ATPase domain is typically fused to a dimerization domain and the ATPase domain itself dimerizes upon ATP binding, entrapping the substrate; ATP hydrolysis drives manipulation of the substrate and release of the dimerized ATPase domains and the substrate (Figure 1.6A) (Dutta and Inouye, 2000). Many GHKL ATPases regulate chromatin structure, including topoisomerase II, MutL-family members, and Structural maintenance of chromosomes flexible hinge domain containing 1 (Smchd1) (Iyer et al., 2008a; 2008b).

The family of Microrchidia GHKL ATPases are important regulators of gene expression in both plants and animals. The first MORC family member, mammalian MORC1 is highly expressed in the germline of male mice where it is required for chromosome pairing during meiosis I, transposon silencing, DNA methylation, and fertility (Inoue et al., 1999; Pastor et al., 2014; Watson et al., 1998). Male *Morc1(-)* mutant mice develop normally with the exception of a small testes, or microrchidia, phenotype (Watson et al., 1998). Subsequent studies have implicated three additional mammalian MORC proteins as well as MORC homologs in *A. thaliana* and *C. elegans* in various aspects of epigenetic regulation. The MORC proteins share a common domain structure, with an N-terminal GHKL ATPase domain and a C-terminal coiled coil domain. The animal MORC proteins also contain a CW zinc finger that is predicted to bind histone 3 (H3) (Figure 1.6B). Recent biochemical studies have shown that mammalian MORC3 and MORC4 preferentially bind H3 that is methylated at lysine 4 (H3K4me2/3) but have not detected direct interactions between H3 and MORC1 or MORC2 (Andrews et al., 2016; Li et al., 2016; Liu et al., 2016). These biochemical findings suggest that (1) the CW domain is not sufficient to direct H3 binding and (2) some MORC proteins may regulate chromatin structure without directly binding chromatin. For instance, the *A. thaliana* MORC proteins are required for

heterochromatin compaction and some DNA methylation but lack a CW domain (Harris et al., 2016; Moissiard et al., 2014; 2012).

The mammalian MORC proteins play diverse roles in chromatin regulation. As described above, MORC1 is required for chromosome pairing during meiosis I, DNA methylation, and transposon silencing (Inoue et al., 1999; Pastor et al., 2014; Watson et al., 1998). Both MORC2 and MORC3 function in DNA damage response. MORC2 promotes chromatin decompaction in response to DNA damage (Li et al., 2012). MORC3 is required for p53 activation following DNA damage (Takahashi et al., 2007). The role of MORC4 has not yet been described.

In *A. thaliana*, there are seven annotated MORCs. AtMORC6 forms heterodimers with AtMORC1 and AtMORC2 to regulate heterochromatin compaction and transposon silencing (Moissiard et al., 2012; 2014). The AtMORC1/2/6 heterodimers form nuclear bodies adjacent to highly compacted clusters of centromeric and pericentromeric heterochromatin called chromocenters (Moissiard et al., 2012; 2014). AtMORC4 and AtMORC7 form homodimers and redundantly regulate the expression of protein-coding genes involved in pathogen response (Harris et al., 2016). Expression of AtMORC4 and AtMORC7 dimers is diffuse throughout the nucleus and enriched adjacent to chromocenters (Harris et al., 2016). The two remaining MORCs make little to no contribution to gene expression: AtMORC3 is thought to be a pseudogene and AtMORC5 is expressed exclusively in pollen (Harris et al., 2016).

In *C. elegans*, the single MORC homolog, MORC-1, is largely undescribed. RNAi-based genetic screens have identified *morc-1* as a candidate RNAi factor (Guang et al., 2008; Kim et al., 2005). Subsequently, knockdown of *morc-1* by RNAi was found to cause transgene desilencing, suggesting that *morc-1* might function in siRNA-directed gene silencing (Moissiard et al., 2012). As the contribution of MORC proteins to gene silencing appears to be highly conserved, they likely play an ancient role in epigenetic regulation and genome defense, however, their specific function in the regulation of chromatin architecture and transposon silencing is not yet understood.
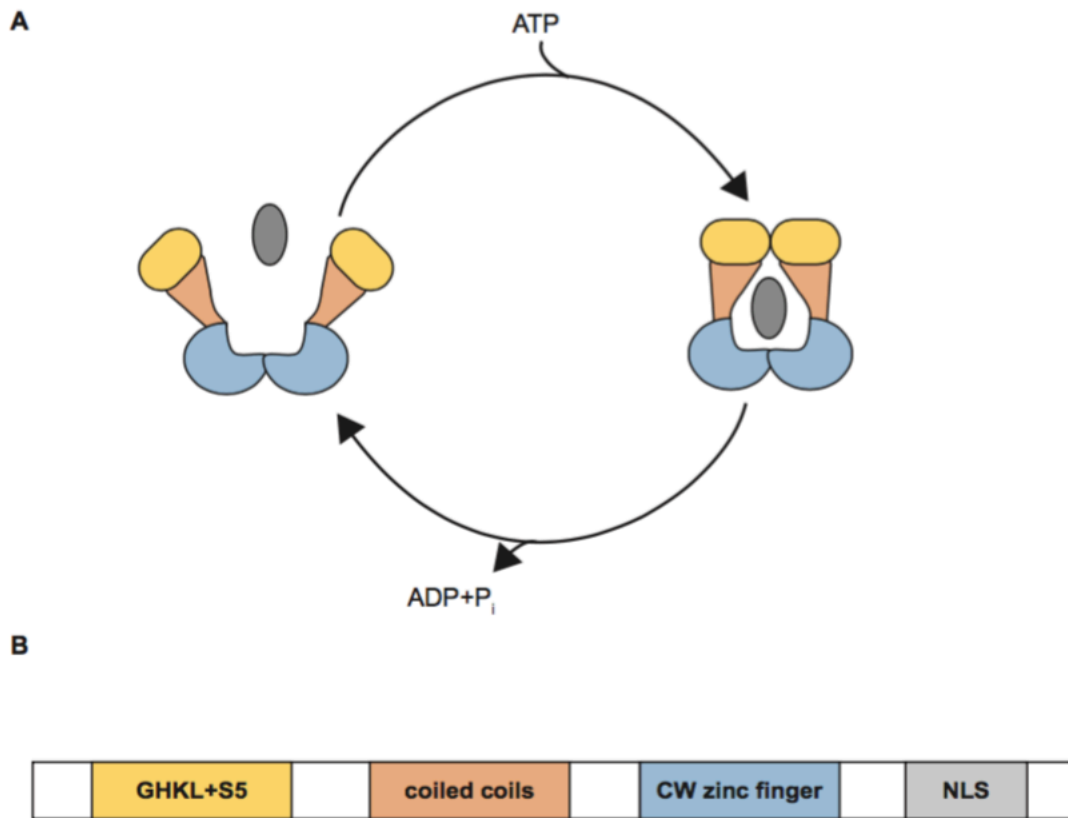
Figure 1.6. The Microrchidia GHKL ATPases. (A) The GHKL ATPases function as molecular clamps. ATP binding drives dimerization of the ATPase domains and ATP hydrolysis drives substrate manipulation and release of the ATPase domains and the substrate (Dutta and Inouye, 2000). **(B)** Domain structure of *C. elegans* MORC-1.

REFERENCES

Ambros, V., Lee, R.C., Lavanway, A., Williams, P.T., and Jewell, D. (2003). MicroRNAs and other tiny endogenous RNAs in C. elegans. Curr. Biol. *13*, 807–818.

Andrews, F.H., Tong, Q., Sullivan, K.D., Cornett, E.M., Zhang, Y., Ali, M., Ahn, J., Pandey, A., Guo, A.H., Strahl, B.D., et al. (2016). Multivalent Chromatin Engagement and Inter-domain Crosstalk Regulate MORC3 ATPase. Cell Rep *16*, 3195–3207.

Aoki, K., Moriguchi, H., Yoshioka, T., Okawa, K., and Tabara, H. (2007). In vitro analyses of the production and activity of secondary small interfering RNAs in C. elegans. Embo J. *26*, 5007–5019.

Aravin, A., Gaidatzis, D., Pfeffer, S., Lagos-Quintana, M., Landgraf, P., Iovino, N., Morris, P., Brownstein, M.J., Kuramochi-Miyagawa, S., Nakano, T., et al. (2006). A novel class of small RNAs bind to MILI protein in mouse testes. Nature *442*, 203–207.

Ashe, A., Sapetschnig, A., Weick, E.-M., Mitchell, J., Bagijn, M.P., Cording, A.C., Doebley, A.-L., Goldstein, L.D., Lehrbach, N.J., Le Pen, J., et al. (2012). piRNAs can trigger a multigenerational epigenetic memory in the germline of C. elegans. Cell *150*, 88–99.

Avgousti, D.C., Palani, S., Sherman, Y., and Grishok, A. (2012). CSR-1 RNAi pathway positively regulates histone expression in C. elegans. Embo J. *31*, 3821–3832.

Bagijn, M.P., Goldstein, L.D., Sapetschnig, A., Weick, E.-M., Bouasker, S., Lehrbach, N.J., Simard, M.J., and Miska, E.A. (2012). Function, targets, and evolution of Caenorhabditis elegans piRNAs. Science *337*, 574–578.

Batista, P.J., Ruby, J.G., Claycomb, J.M., Chiang, R., Fahlgren, N., Kasschau, K.D., Chaves, D.A., Gu, W., Vasale, J.J., Duan, S., et al. (2008). PRG-1 and 21U-RNAs interact to form the piRNA complex required for fertility in C. elegans. Mol. Cell *31*, 67–78.

Billi, A.C., Alessi, A.F., Khivansara, V., Han, T., Freeberg, M., Mitani, S., and Kim, J.K. (2012). The Caenorhabditis elegans HEN1 ortholog, HENN-1, methylates and stabilizes select subclasses of germline small RNAs. PLoS Genet. *8*, e1002617.

Billi, A.C., Fischer, S.E.J., and Kim, J.K. (2014). Endogenous RNAi pathways in C. elegans. WormBook 1–49.

Billi, A.C., Freeberg, M.A., Day, A.M., Chun, S.Y., Khivansara, V., and Kim, J.K. (2013). A conserved upstream motif orchestrates autonomous, germline-enriched expression of Caenorhabditis elegans piRNAs. PLoS Genet. *9*, e1003392.

Blanchard, D., Parameswaran, P., Lopez-Molina, J., Gent, J., Saynuk, J.F., and Fire, A. (2011). On the nature of in vivo requirements for rde-4 in RNAi and developmental

pathways in C. elegans. RNA Biol *8*, 458–467.

Brennecke, J., Aravin, A.A., Stark, A., Dus, M., Kellis, M., Sachidanandam, R., and Hannon, G.J. (2007). Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. Cell *128*, 1089–1103.

Buckley, B.A., Burkhart, K.B., Gu, S.G., Spracklin, G., Kershner, A., Fritz, H., Kimble, J., Fire, A., and Kennedy, S. (2012). A nuclear Argonaute promotes multigenerational epigenetic inheritance and germline immortality. Nature *489*, 447–451.

Burkhart, K.B., Guang, S., Buckley, B.A., Wong, L., Bochner, A.F., and Kennedy, S. (2011). A pre-mRNA-associating factor links endogenous siRNAs to chromatin regulation. PLoS Genet. *7*, e1002249.

Burton, N.O., Burkhart, K.B., and Kennedy, S. (2011). Nuclear RNAi maintains heritable gene silencing in Caenorhabditis elegans. Proc. Natl. Acad. Sci. U.S.a. *108*, 19683–19688.

Campbell, A.C., and Updike, D.L. (2015). CSR-1 and P granules suppress sperm-specific transcription in the C. elegans germline. Development *142*, 1745–1755.

Cecere, G., Hoersch, S., O'Keeffe, S., Sachidanandam, R., and Grishok, A. (2014). Global effects of the CSR-1 RNA interference pathway on the transcriptional landscape. Nat. Struct. Mol. Biol. *21*, 358–365.

Cecere, G., Zheng, G.X.Y., Mansisidor, A.R., Klymko, K.E., and Grishok, A. (2012). Promoters recognized by forkhead proteins exist for individual 21U-RNAs. Mol. Cell *47*, 734–745.

Chu, D.S., Liu, H., Nix, P., Wu, T.F., Ralston, E.J., Yates, J.R., and Meyer, B.J. (2006). Sperm chromatin proteomics identifies evolutionarily conserved fertility factors. Nature *443*, 101–105.

Claycomb, J.M., Batista, P.J., Pang, K.M., Gu, W., Vasale, J.J., van Wolfswinkel, J.C., Chaves, D.A., Shirayama, M., Mitani, S., Ketting, R.F., et al. (2009). The Argonaute CSR-1 and its 22G-RNA cofactors are required for holocentric chromosome segregation. Cell *139*, 123–134.

Collins, J., Saari, B., and Anderson, P. (1987). Activation of a transposable element in the germ line but not the soma of Caenorhabditis elegans. Nature *328*, 726–728.

Conine, C.C., Batista, P.J., Gu, W., Claycomb, J.M., Chaves, D.A., Shirayama, M., and Mello, C.C. (2010). Argonautes ALG-3 and ALG-4 are required for spermatogenesis-specific 26G-RNAs and thermotolerant sperm in Caenorhabditis elegans. Proc. Natl. Acad. Sci. U.S.a. *107*, 3588–3593.

Conine, C.C., Moresco, J.J., Gu, W., Shirayama, M., Conte, D., Yates, J.R., and Mello, C.C. (2013). Argonautes promote male fertility and provide a paternal memory of

germline gene expression in C. elegans. Cell *155*, 1532–1544.

Das, P.P., Bagijn, M.P., Goldstein, L.D., Woolford, J.R., Lehrbach, N.J., Sapetschnig, A., Buhecha, H.R., Gilchrist, M.J., Howe, K.L., Stark, R., et al. (2008). Piwi and piRNAs act upstream of an endogenous siRNA pathway to suppress Tc3 transposon mobility in the Caenorhabditis elegans germline. Mol. Cell *31*, 79–90.

de Albuquerque, B.F.M., Luteijn, M.J., Cordeiro Rodrigues, R.J., van Bergeijk, P., Waaijers, S., Kaaij, L.J.T., Klein, H., Boxem, M., and Ketting, R.F. (2014). PID-1 is a novel factor that operates during 21U-RNA biogenesis in Caenorhabditis elegans. Genes Dev. *28*, 683–688.

de Albuquerque, B.F.M., Placentino, M., and Ketting, R.F. (2015). Maternal piRNAs Are Essential for Germline Development following De Novo Establishment of Endo-siRNAs in Caenorhabditis elegans. Dev. Cell *34*, 448–456.

de Wit, E., Linsen, S.E.V., Cuppen, E., and Berezikov, E. (2009). Repertoire and evolution of miRNA genes in four divergent nematode species. Genome Res. *19*, 2064–2074.

Duchaine, T.F., Wohlschlegel, J.A., Kennedy, S., Bei, Y., Conte, D., Pang, K., Brownell, D.R., Harding, S., Mitani, S., Ruvkun, G., et al. (2006). Functional proteomics reveals the biochemical niche of C. elegans DCR-1 in multiple small-RNA-mediated pathways. Cell *124*, 343–354.

Dutta, R., and Inouye, M. (2000). GHKL, an emergent ATPase/kinase superfamily. Trends Biochem. Sci. *25*, 24–28.

Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E., and Mello, C.C. (1998). Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature *391*, 806–811.

Fischer, S.E.J., Montgomery, T.A., Zhang, C., Fahlgren, N., Breen, P.C., Hwang, A., Sullivan, C.M., Carrington, J.C., and Ruvkun, G. (2011). The ERI-6/7 helicase acts at the first stage of an siRNA amplification pathway that targets recent gene duplications. PLoS Genet. *7*, e1002369.

Fischer, S.E.J., Pan, Q., Breen, P.C., Qi, Y., Shi, Z., Zhang, C., and Ruvkun, G. (2013). Multiple small RNA pathways regulate the silencing of repeated and foreign genes in C. elegans. Genes Dev. *27*, 2678–2695.

Fitzgerald, M.E., Vela, A., and Pyle, A.M. (2014). Dicer-related helicase 3 forms an obligate dimer for recognizing 22G-RNA. Nucleic Acids Res. *42*, 3919–3930.

Gabel, H.W., and Ruvkun, G. (2008). The exonuclease ERI-1 has a conserved dual role in 5.8S rRNA processing and RNAi. Nat. Struct. Mol. Biol. *15*, 531–533.

Gaydos, L.J., Rechtsteiner, A., Egelhofer, T.A., Carroll, C.R., and Strome, S. (2012).

Antagonism between MES-4 and Polycomb repressive complex 2 promotes appropriate gene expression in C. elegans germ cells. Cell Rep *2*, 1169–1177.

Gent, J.I., Lamm, A.T., Pavelec, D.M., Maniar, J.M., Parameswaran, P., Tao, L., Kennedy, S., and Fire, A.Z. (2010). Distinct phases of siRNA synthesis in an endogenous RNAi pathway in C. elegans soma. Mol. Cell *37*, 679–689.

Gent, J.I., Schvarzstein, M., Villeneuve, A.M., Gu, S.G., Jantsch, V., Fire, A.Z., and Baudrimont, A. (2009). A Caenorhabditis elegans RNA-directed RNA polymerase in sperm development and endogenous RNA interference. Genetics *183*, 1297–1314.

Gerson-Gurwitz, A., Wang, S., Sathe, S., Green, R., Yeo, G.W., Oegema, K., and Desai, A. (2016). A Small RNA-Catalytic Argonaute Pathway Tunes Germline Transcript Levels to Ensure Embryonic Divisions. Cell *165*, 396–409.

Girard, A., Sachidanandam, R., Hannon, G.J., and Carmell, M.A. (2006). A germline-specific class of small RNAs binds mammalian Piwi proteins. Nature *442*, 199–202.

Goh, W.-S.S., Seah, J.W.E., Harrison, E.J., Chen, C., Hammell, C.M., and Hannon, G.J. (2014). A genome-wide RNAi screen identifies factors required for distinct stages of C. elegans piRNA biogenesis. Genes Dev. *28*, 797–807.

Grishok, A., Pasquinelli, A.E., Conte, D., Li, N., Parrish, S., Ha, I., Baillie, D.L., Fire, A., Ruvkun, G., and Mello, C.C. (2001). Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control C. elegans developmental timing. Cell *106*, 23–34.

Grishok, A., Tabara, H., and Mello, C.C. (2000). Genetic requirements for inheritance of RNAi in C. elegans. Science *287*, 2494–2497.

Grivna, S.T., Beyret, E., Wang, Z., and Lin, H. (2006a). A novel class of small RNAs in mouse spermatogenic cells. Genes Dev. *20*, 1709–1714.

Grivna, S.T., Pyhtila, B., and Lin, H. (2006b). MIWI associates with translational machinery and PIWI-interacting RNAs (piRNAs) in regulating spermatogenesis. Proc. Natl. Acad. Sci. U.S.a. *103*, 13415–13420.

Gu, S.G., Pak, J., Guang, S., Maniar, J.M., Kennedy, S., and Fire, A. (2012a). Amplification of siRNA in Caenorhabditis elegans generates a transgenerational sequence-targeted histone H3 lysine 9 methylation footprint. Nat. Genet. *44*, 157–164.

Gu, W., Lee, H.-C., Chaves, D., Youngman, E.M., Pazour, G.J., Conte, D., and Mello, C.C. (2012b). CapSeq and CIP-TAP Identify Pol II Start Sites and Reveal Capped Small RNAs as C. elegans piRNA Precursors. Cell *151*, 1488–1500.

Gu, W., Shirayama, M., Conte, D., Vasale, J., Batista, P.J., Claycomb, J.M., Moresco, J.J., Youngman, E.M., Keys, J., Stoltz, M.J., et al. (2009). Distinct argonaute-mediated 22G-RNA pathways direct genome surveillance in the C. elegans germline. Mol. Cell

36, 231–244.

Guang, S., Bochner, A.F., Burkhart, K.B., Burton, N., Pavelec, D.M., and Kennedy, S. (2010). Small regulatory RNAs inhibit RNA polymerase II during the elongation phase of transcription. Nature *465*, 1097–1101.

Guang, S., Bochner, A.F., Pavelec, D.M., Burkhart, K.B., Harding, S., Lachowiec, J., and Kennedy, S. (2008). An Argonaute transports siRNAs from the cytoplasm to the nucleus. Science *321*, 537–541.

Han, T., Manoharan, A.P., Harkins, T.T., Bouffard, P., Fitzpatrick, C., Chu, D.S., Thierry-Mieg, D., Thierry-Mieg, J., and Kim, J.K. (2009). 26G endo-siRNAs regulate spermatogenic and zygotic gene expression in Caenorhabditis elegans. Proc. Natl. Acad. Sci. U.S.a. *106*, 18674–18679.

Harris, C.J., Husmann, D., Liu, W., Kasmi, F.E., Wang, H., Papikian, A., Pastor, W.A., Moissiard, G., Vashisht, A.A., Dangl, J.L., et al. (2016). Arabidopsis AtMORC4 and AtMORC7 Form Nuclear Bodies and Repress a Large Number of Protein-Coding Genes. PLoS Genet. *12*, e1005998.

Inoue, N., Hess, K.D., Moreadith, R.W., Richardson, L.L., Handel, M.A., Watson, M.L., and Zinn, A.R. (1999). New gene family defined by MORC, a nuclear protein required for mouse spermatogenesis. Hum. Mol. Genet. *8*, 1201–1207.

Iyer, L.M., Abhiman, S., and Aravind, L. (2008a). MutL homologs in restriction-modification systems and the origin of eukaryotic MORC ATPases. Biol. Direct *3*, 8.

Iyer, L.M., Anantharaman, V., Wolf, M.Y., and Aravind, L. (2008b). Comparative genomics of transcription factors and chromatin proteins in parasitic protists and other eukaryotes. Int. J. Parasitol. *38*, 1–31.

Kamminga, L.M., van Wolfswinkel, J.C., Luteijn, M.J., Kaaij, L.J.T., Bagijn, M.P., Sapetschnig, A., Miska, E.A., Berezikov, E., and Ketting, R.F. (2012). Differential impact of the HEN1 homolog HENN-1 on 21U and 26G RNAs in the germline of Caenorhabditis elegans. PLoS Genet. *8*, e1002702.

Kasper, D.M., Wang, G., Gardner, K.E., Johnstone, T.G., and Reinke, V. (2014). The C. elegans SNAPc component SNPC-4 coats piRNA domains and is globally required for piRNA abundance. Dev. Cell *31*, 145–158.

Kennedy, S., Wang, D., and Ruvkun, G. (2004). A conserved siRNA-degrading RNase negatively regulates RNA interference in C. elegans. Nature *427*, 645–649.

Ketting, R.F., Haverkamp, T.H., van Luenen, H.G., and Plasterk, R.H. (1999). Mut-7 of C. elegans, required for transposon silencing and RNA interference, is a homolog of Werner syndrome helicase and RNaseD. Cell *99*, 133–141.

Kim, J.K., Gabel, H.W., Kamath, R.S., Tewari, M., Pasquinelli, A., Rual, J.-F., Kennedy,

S., Dybbs, M., Bertin, N., Kaplan, J.M., et al. (2005). Functional genomic analysis of RNA interference in C. elegans. Science *308*, 1164–1167.

Lee, H.-C., Gu, W., Shirayama, M., Youngman, E., Conte, D., and Mello, C.C. (2012). C. elegans piRNAs mediate the genome-wide surveillance of germline transcripts. Cell *150*, 78–87.

Lee, R.C., Hammell, C.M., and Ambros, V. (2006). Interacting endogenous and exogenous RNAi pathways in Caenorhabditis elegans. Rna *12*, 589–597.

Li, D.-Q., Nair, S.S., Ohshiro, K., Kumar, A., Nair, V.S., Pakala, S.B., Reddy, S.D.N., Gajula, R.P., Eswaran, J., Aravind, L., et al. (2012). MORC2 signaling integrates phosphorylation-dependent, ATPase-coupled chromatin remodeling during the DNA damage response. Cell Rep *2*, 1657–1669.

Li, S., Yen, L., Pastor, W.A., Johnston, J.B., Du, J., Shew, C.J., Liu, W., Ho, J., Stender, B., Clark, A.T., et al. (2016). Mouse MORC3 is a GHKL ATPase that localizes to H3K4me3 marked chromatin. Proc. Natl. Acad. Sci. U.S.a. 201609709.

Liu, Y., Tempel, W., Zhang, Q., Liang, X., Loppnau, P., Qin, S., and Min, J. (2016). Family-wide Characterization of Histone Binding Abilities of Human CW Domain-containing Proteins. J. Biol. Chem. *291*, 9000–9013.

Luteijn, M.J., van Bergeijk, P., Kaaij, L.J.T., Almeida, M.V., Roovers, E.F., Berezikov, E., and Ketting, R.F. (2012). Extremely stable Piwi-induced gene silencing in Caenorhabditis elegans. Embo J. *31*, 3422–3430.

Mao, H., Zhu, C., Zong, D., Weng, C., Yang, X., Huang, H., Liu, D., Feng, X., and Guang, S. (2015). The Nrde Pathway Mediates Small-RNA-Directed Histone H3 Lysine 27 Trimethylation in Caenorhabditis elegans. Curr. Biol. *25*, 2398–2403.

Moissiard, G., Bischof, S., Husmann, D., Pastor, W.A., Hale, C.J., Yen, L., Stroud, H., Papikian, A., Vashisht, A.A., Wohlschlegel, J.A., et al. (2014). Transcriptional gene silencing by Arabidopsis microrchidia homologues involves the formation of heteromers. Proc. Natl. Acad. Sci. U.S.a. *111*, 7474–7479.

Moissiard, G., Cokus, S.J., Cary, J., Feng, S., Billi, A.C., Stroud, H., Husmann, D., Zhan, Y., Lajoie, B.R., McCord, R.P., et al. (2012). MORC family ATPases required for heterochromatin condensation and gene silencing. Science *336*, 1448–1451.

Montgomery, T.A., Rim, Y.-S., Zhang, C., Dowen, R.H., Phillips, C.M., Fischer, S.E.J., and Ruvkun, G. (2012). PIWI associated siRNAs and piRNAs specifically require the Caenorhabditis elegans HEN1 ortholog henn-1. PLoS Genet. *8*, e1002616.

Nakamura, M., Ando, R., Nakazawa, T., Yudazono, T., Tsutsumi, N., Hatanaka, N., Ohgake, T., Hanaoka, F., and Eki, T. (2007). Dicer-related drh-3 gene functions in germ-line development by maintenance of chromosomal integrity in Caenorhabditis elegans. Genes Cells *12*, 997–1010.

Ni, J.Z., Chen, E., and Gu, S.G. (2014). Complex coding of endogenous siRNA, transcriptional silencing and H3K9 methylation on native targets of germline nuclear RNAi in C. elegans. BMC Genomics *15*, 1157.

Ni, J.Z., Kalinava, N., Chen, E., Huang, A., Trinh, T., and Gu, S.G. (2016). A transgenerational role of the germline nuclear RNAi pathway in repressing heat stress-induced transcriptional activation in C. elegans. Epigenetics Chromatin *9*, 3.

Pak, J., and Fire, A. (2007). Distinct populations of primary and secondary effectors during RNAi in C. elegans. Science *315*, 241–244.

Pastor, W.A., Stroud, H., Nee, K., Liu, W., Pezic, D., Manakov, S., Lee, S.A., Moissiard, G., Zamudio, N., Bourc'his, D., et al. (2014). MORC1 represses transposable elements in the mouse male germline. Nat Commun *5*, 5795.

Pavelec, D.M., Lachowiec, J., Duchaine, T.F., Smith, H.E., and Kennedy, S. (2009). Requirement for the ERI/DICER complex in endogenous RNA interference and sperm development in Caenorhabditis elegans. Genetics *183*, 1283–1295.

Phillips, C.M., Brown, K.C., Montgomery, B.E., Ruvkun, G., and Montgomery, T.A. (2015). piRNAs and piRNA-Dependent siRNAs Protect Conserved and Essential C. elegans Genes from Misrouting into the RNAi Pathway. Dev. Cell *34*, 457–465.

Phillips, C.M., Montgomery, B.E., Breen, P.C., Roovers, E.F., Rim, Y.-S., Ohsumi, T.K., Newman, M.A., van Wolfswinkel, J.C., Ketting, R.F., Ruvkun, G., et al. (2014). MUT-14 and SMUT-1 DEAD box RNA helicases have overlapping roles in germline RNAi and endogenous siRNA formation. Curr. Biol. *24*, 839–844.

Phillips, C.M., Montgomery, T.A., Breen, P.C., and Ruvkun, G. (2012). MUT-16 promotes formation of perinuclear mutator foci required for RNA silencing in the C. elegans germline. Genes Dev. *26*, 1433–1444.

Robert, V.J.P., Sijen, T., van Wolfswinkel, J., and Plasterk, R.H.A. (2005). Chromatin and RNAi factors protect the C. elegans germline against repetitive sequences. Genes Dev. *19*, 782–787.

Ruby, J.G., Jan, C., Player, C., Axtell, M.J., Lee, W., Nusbaum, C., Ge, H., and Bartel, D.P. (2006). Large-scale sequencing reveals 21U-RNAs and additional microRNAs and endogenous siRNAs in C. elegans. Cell *127*, 1193–1207.

Sakaguchi, A., Sarkies, P., Simon, M., Doebley, A.-L., Goldstein, L.D., Hedges, A., Ikegami, K., Alvares, S.M., Yang, L., LaRocque, J.R., et al. (2014). Caenorhabditis elegans RSD-2 and RSD-6 promote germ cell immortality by maintaining small interfering RNA populations. Proc. Natl. Acad. Sci. U.S.a. *111*, E4323–E4331.

Sapetschnig, A., Sarkies, P., Lehrbach, N.J., and Miska, E.A. (2015). Tertiary siRNAs mediate paramutation in C. elegans. PLoS Genet. *11*, e1005078.

Seth, M., Shirayama, M., Gu, W., Ishidate, T., Conte, D., and Mello, C.C. (2013). The C. elegans CSR-1 argonaute pathway counteracts epigenetic silencing to promote germline gene expression. Dev. Cell *27*, 656–663.

She, X., Xu, X., Fedotov, A., Kelly, W.G., and Maine, E.M. (2009). Regulation of heterochromatin assembly on unpaired chromosomes during Caenorhabditis elegans meiosis by components of a small RNA-mediated pathway. PLoS Genet. *5*, e1000624.

Sheth, U., Pitt, J., Dennis, S., and Priess, J.R. (2010). Perinuclear P granules are the principal sites of mRNA export in adult C. elegans germ cells. Development *137*, 1305–1314.

Shirayama, M., Seth, M., Lee, H.-C., Gu, W., Ishidate, T., Conte, D., and Mello, C.C. (2012). piRNAs initiate an epigenetic memory of nonself RNA in the C. elegans germline. Cell *150*, 65–77.

Shirayama, M., Stanney, W., Gu, W., Seth, M., and Mello, C.C. (2014). The Vasa Homolog RDE-12 engages target mRNA and multiple argonaute proteins to promote RNAi in C. elegans. Curr. Biol. *24*, 845–851.

Sijen, T., Fleenor, J., Simmer, F., Thijssen, K.L., Parrish, S., Timmons, L., Plasterk, R.H., and Fire, A. (2001). On the role of RNA amplification in dsRNA-triggered gene silencing. Cell *107*, 465–476.

Sijen, T., and Plasterk, R.H.A. (2003). Transposon silencing in the Caenorhabditis elegans germ line by natural RNAi. Nature *426*, 310–314.

Sijen, T., Steiner, F.A., Thijssen, K.L., and Plasterk, R.H.A. (2007). Secondary siRNAs result from unprimed RNA synthesis and form a distinct class. Science *315*, 244–247.

Simmer, F., Tijsterman, M., Parrish, S., Koushika, S.P., Nonet, M.L., Fire, A., Ahringer, J., and Plasterk, R.H.A. (2002). Loss of the putative RNA-directed RNA polymerase RRF-3 makes C. elegans hypersensitive to RNAi. Curr. Biol. *12*, 1317–1319.

Smardon, A., Spoerke, J.M., Stacey, S.C., Klein, M.E., Mackin, N., and Maine, E.M. (2000). EGO-1 is related to RNA-directed RNA polymerase and functions in germ-line development and RNA interference in C. elegans. Curr. Biol. *10*, 169–178.

Tabara, H., Sarkissian, M., Kelly, W.G., Fleenor, J., Grishok, A., Timmons, L., Fire, A., and Mello, C.C. (1999). The rde-1 gene, RNA interference, and transposon silencing in C. elegans. Cell *99*, 123–132.

Tabara, H., Yigit, E., Siomi, H., and Mello, C.C. (2002). The dsRNA binding protein RDE-4 interacts with RDE-1, DCR-1, and a DExH-box helicase to direct RNAi in C. elegans. Cell *109*, 861–871.

Takahashi, K., Yoshida, N., Murakami, N., Kawata, K., Ishizaki, H., Tanaka-Okamoto, M., Miyoshi, J., Zinn, A.R., Shime, H., and Inoue, N. (2007). Dynamic regulation of p53

subnuclear localization and senescence by MORC3. Mol. Biol. Cell *18*, 1701–1709.

Talsky, K.B., and Collins, K. (2010). Initiation by a eukaryotic RNA-dependent RNA polymerase requires looping of the template end and is influenced by the template-tailing activity of an associated uridyltransferase. J. Biol. Chem. *285*, 27614–27623.

Tang, W., Tu, S., Lee, H.-C., Weng, Z., and Mello, C.C. (2016). The RNase PARN-1 Trims piRNA 3' Ends to Promote Transcriptome Surveillance in C. elegans. Cell *164*, 974–984.

Thivierge, C., Makil, N., Flamand, M., Vasale, J.J., Mello, C.C., Wohlschlegel, J., Conte, D., and Duchaine, T.F. (2012). Tudor domain ERI-5 tethers an RNA-dependent RNA polymerase to DCR-1 to potentiate endo-RNAi. Nat. Struct. Mol. Biol. *19*, 90–97.

Tijsterman, M., May, R.C., Simmer, F., Okihara, K.L., and Plasterk, R.H.A. (2004). Genes required for systemic RNA interference in Caenorhabditis elegans. Curr. Biol. *14*, 111–116.

Tijsterman, M., Pothof, J., and Plasterk, R.H.A. (2002). Frequent germline mutations and somatic repeat instability in DNA mismatch-repair-deficient Caenorhabditis elegans. Genetics *161*, 651–660.

Tsai, H.-Y., Chen, C.-C.G., Conte, D., Moresco, J.J., Chaves, D.A., Mitani, S., Yates, J.R., Tsai, M.-D., and Mello, C.C. (2015). A ribonuclease coordinates siRNA amplification and mRNA cleavage during RNAi. Cell *160*, 407–419.

Tu, S., Wu, M.Z., Wang, J., Cutter, A.D., Weng, Z., and Claycomb, J.M. (2015). Comparative functional characterization of the CSR-1 22G-RNA pathway in Caenorhabditis nematodes. Nucleic Acids Res. *43*, 208–224.

van Wolfswinkel, J.C., Claycomb, J.M., Batista, P.J., Mello, C.C., Berezikov, E., and Ketting, R.F. (2009). CDE-1 affects chromosome segregation through uridylation of CSR-1-bound siRNAs. Cell *139*, 135–148.

Vasale, J.J., Gu, W., Thivierge, C., Batista, P.J., Claycomb, J.M., Youngman, E.M., Duchaine, T.F., Mello, C.C., and Conte, D. (2010). Sequential rounds of RNA-dependent RNA transcription drive endogenous small-RNA biogenesis in the ERGO-1/Argonaute pathway. Proc. Natl. Acad. Sci. U.S.a. *107*, 3582–3587.

Vastenhouw, N.L., Fischer, S.E.J., Robert, V.J.P., Thijssen, K.L., Fraser, A.G., Kamath, R.S., Ahringer, J., and Plasterk, R.H.A. (2003). A genome-wide screen identifies 27 genes involved in transposon silencing in C. elegans. Curr. Biol. *13*, 1311–1316.

Vought, V.E., Ohmachi, M., Lee, M.-H., and Maine, E.M. (2005). EGO-1, a putative RNA-directed RNA polymerase, promotes germline proliferation in parallel with GLP-1/notch signaling and regulates the spatial organization of nuclear pore complexes and germline P granules in Caenorhabditis elegans. Genetics *170*, 1121–1132.

Wang, G., and Reinke, V. (2008). A C. elegans Piwi, PRG-1, regulates 21U-RNAs during spermatogenesis. Curr. Biol. *18*, 861–867.

Watson, M.L., Zinn, A.R., Inoue, N., Hess, K.D., Cobb, J., Handel, M.A., Halaban, R., Duchene, C.C., Albright, G.M., and Moreadith, R.W. (1998). Identification of morc (microrchidia), a mutation that results in arrest of spermatogenesis at an early meiotic stage in the mouse. Proc. Natl. Acad. Sci. U.S.a. *95*, 14361–14366.

Wedeles, C.J., Wu, M.Z., and Claycomb, J.M. (2013). Protection of germline gene expression by the C. elegans Argonaute CSR-1. Dev. Cell *27*, 664–671.

Weick, E.-M., Sarkies, P., Silva, N., Chen, R.A., Moss, S.M.M., Cording, A.C., Ahringer, J., Martinez-Perez, E., and Miska, E.A. (2014). PRDE-1 is a nuclear factor essential for the biogenesis of Ruby motif-dependent piRNAs in C. elegans. Genes Dev. *28*, 783–796.

Welker, N.C., Maity, T.S., Ye, X., Aruscavage, P.J., Krauchuk, A.A., Liu, Q., and Bass, B.L. (2011). Dicer's helicase domain discriminates dsRNA termini to promote an altered reaction mode. Mol. Cell *41*, 589–599.

Welker, N.C., Pavelec, D.M., Nix, D.A., Duchaine, T.F., Kennedy, S., and Bass, B.L. (2010). Dicer's helicase domain is required for accumulation of some, but not all, C. elegans endogenous siRNAs. Rna *16*, 893–903.

Yang, H., Vallandingham, J., Shiu, P., Li, H., Hunter, C.P., and Mak, H.Y. (2014). The DEAD box helicase RDE-12 promotes amplification of RNAi in cytoplasmic foci in C. elegans. Curr. Biol. *24*, 832–838.

Yang, H., Zhang, Y., Vallandingham, J., Li, H., Li, H., Florens, L., and Mak, H.Y. (2012). The RDE-10/RDE-11 complex triggers RNAi-induced mRNA degradation by association with target mRNA in C. elegans. Genes Dev. *26*, 846–856.

Yigit, E., Batista, P.J., Bei, Y., Pang, K.M., Chen, C.-C.G., Tolia, N.H., Joshua-Tor, L., Mitani, S., Simard, M.J., and Mello, C.C. (2006). Analysis of the C. elegans Argonaute family reveals that distinct Argonautes act sequentially during RNAi. Cell *127*, 747–757.

Zhang, C., Montgomery, T.A., Fischer, S.E.J., Garcia, S.M.D.A., Riedel, C.G., Fahlgren, N., Sullivan, C.M., Carrington, J.C., and Ruvkun, G. (2012). The Caenorhabditis elegans RDE-10/RDE-11 complex regulates RNAi by promoting secondary siRNA amplification. Curr. Biol. *22*, 881–890.

Zhang, C., Montgomery, T.A., Gabel, H.W., Fischer, S.E.J., Phillips, C.M., Fahlgren, N., Sullivan, C.M., Carrington, J.C., and Ruvkun, G. (2011). mut-16 and other mutator class genes modulate 22G and 26G siRNA pathways in Caenorhabditis elegans. Proc. Natl. Acad. Sci. U.S.a. *108*, 1201–1208.

CHAPTER 2

MORC-1 integrates nuclear RNAi and transgenerational chromatin architecture to promote germline immortality

AUTHORS: Natasha E. Weiser, Danny X. Yang, Suhua Feng, Natallia Kalinava, Kristen C. Brown, Jayshree Khanikar, Mallory A. Freeberg, Martha J. Snyder, Gyorgyi Csankovszki, Raymond C. Chan, Sam G. Gu, Taiowa A. Montgomery, Steven E. Jacobsen, and John K. Kim

AUTHOR CONTRIBUTIONS: Unless otherwise stated, all experiments and analysis were performed by me. Other contributions are as follows:

- RNAi sensitivity and inheritance (Figures 2.4, 2.6, 2.7): Danny X. Yang (University of Michigan)
- Library preparation and sequencing: Suhua Feng and Steven E. Jacobsen (University of California – Los Angeles)
- Computational analysis
  - ChIP-seq: Natallia Kalinava and Sam G. Gu (Rutgers University)
  - mRNA-seq and sRNA-seq: Kristen C. Brown and Taiowa A. Montgomery (Colorado State University)
  - Genome screen: Mallory A. Freeberg (Johns Hopkins University)
- Heterochromatin localization (Figure 2.21): Jayshree Khanikar and Raymond C. Chan (University of Michigan)
- X chromosome compaction and male rescue (Figure 2.22): with Martha J. Snyder and Gyorgyi Csankovszki (University of Michigan)

ABSTRACT

Germline-expressed endogenous siRNAs (endo-siRNAs) transmit multigenerational epigenetic information to ensure fertility in subsequent generations. In *C. elegans*, nuclear RNAi ensures robust inheritance of endo-siRNAs and deposition of H3K9me3 marks at target loci. How target silencing is maintained in subsequent generations is unknown. We discovered that *morc-1* is essential for transgenerational fertility and acts as an effector of endo-siRNAs. Unexpectedly, *morc-1* is dispensable for siRNA inheritance, but required for target silencing and maintenance of siRNA-dependent chromatin organization. A forward genetic screen identified mutations in *met-1*, which encodes a H3K36 methyltransferase, as potent suppressors of *morc-1(-)* and nuclear RNAi mutant phenotypes. Further analysis of nuclear RNAi and *morc-1(-)* mutants revealed a progressive, *met-1*-dependent enrichment of H3K36me3, suggesting that robust fertility requires repression of MET-1 activity at nuclear RNAi targets. In the absence of MORC-1 and nuclear RNAi, MET-1-mediated encroachment of euchromatin leads to detrimental decondensation of germline chromatin and germline mortality.

INTRODUCTION

Germline proliferation and maintenance are key processes in organismal development. From an evolutionary perspective, the production of functional gametes is the only imperative for species propagation. The faithful transmission of genetic and epigenetic information to the next generation is essential to this process, enabling the production of functional gametes and of the subsequent generation. The ability of germ cells to undergo infinite cellular divisions makes the germline an immortal cell lineage (Smelick and Ahmed, 2005). How this immortality is achieved remains a central question in developmental biology.

The rapid development and genetic tractability of the model organism *C. elegans* makes it an ideal system in which to study the molecular mechanisms of germline

immortality. Recent studies in *C. elegans* have identified epigenetic regulators that are critical for transgenerational maintenance of fertility. These include regulators of DNA damage repair and telomere maintenance as well as histone modifying enzymes (Ahmed and Hodgkin, 2000; Andersen and Horvitz, 2007; Greer et al., 2014; Katz et al., 2009; Meier et al., 2009; Zeller et al., 2016). Epigenetic regulation, such as post-translational modification of histones or DNA methylation, appears to be a conserved mechanism for repression of repetitive elements and germline maintenance in metazoans (Hajkova et al., 2002; Santos and Dean, 2004; Smelick and Ahmed, 2005).

Although repression of deleterious genomic elements, such as transposons, is critical for germline immortality, how repression is targeted and maintained at these loci is unclear. Epigenetic regulators must target these elements for germline repression and transmit this targeting to progeny. In recent years, germline-enriched small-non coding RNAs, such as endogenous siRNAs (endo-siRNAs) and Piwi-interacting RNAs (piRNAs) have been implicated in transposon repression (Batista et al., 2008; Ketting et al., 1999; Tabara et al., 2002). In *C. elegans*, endo-siRNAs and piRNAs collectively regulate genome defense by targeting deleterious or foreign transcripts such as transposons, pseudogenes, and transgenes for repression (Ashe et al., 2012; Batista et al., 2008; Das et al., 2008; de Albuquerque et al., 2015; Han et al., 2009; Kim et al., 2005; Lee et al., 2012; Luteijn et al., 2012; Phillips et al., 2015; Shirayama et al., 2012; Tabara et al., 1999). Furthermore, these pathways are able to regulate gene expression transgenerationally via germline nuclear RNAi (Ashe et al., 2012; Buckley et al., 2012; Grishok et al., 2000; Gu et al., 2012; Luteijn et al., 2012; Shirayama et al., 2012), and thus are thought to transmit a memory of genomic "self" and "non-self" elements.

Recent genetic screens for nuclear RNAi defective (Nrde) and heritable RNAi defective (Hrde) mutants have identified the core factors that transmit this memory (Ashe et al., 2012; Buckley et al., 2012; Burkhart et al., 2011; Gu et al., 2012; Guang et al., 2010; 2008; Luteijn et al., 2012; Shirayama et al., 2012). These factors (HRDE-1/WAGO-9, NRDE-1,2,4) comprise the germline nuclear RNA-induced silencing complex (RISC) which maintains transgenerational gene silencing through heritable siRNA-mediated mechanisms (Ashe et al., 2012; Buckley et al., 2012; Gu et al., 2012; Ni et al., 2016; Shirayama et al., 2012). The germline nuclear RISC is guided by endo-

siRNAs to nascent pre-mRNA transcripts to mediate RNA polymerase II stalling and deposition of H3K9 methylation at the corresponding genomic loci (Buckley et al., 2012; Burkhart et al., 2011; Gu et al., 2012; Guang et al., 2010). Loss of the germline nuclear RNAi pathway, by mutation of *hrde-1* or *nrde-1,2,4*, leads to defective inheritance of silencing by nuclear RNAi, germline mortality, and a progressive depletion of H3K9me3 marks at endogenous target sites (Buckley et al., 2012). Although previous studies have implicated the histone methyltransferases SET-25 and MET-2 in siRNA-directed H3K9 methylation (Ashe et al., 2012; Mao et al., 2015), the mechanisms by which siRNAs regulate heterochromatin deposition and maintenance remain unclear.

The single *C. elegans* homolog of the highly conserved family of microrchidia genes, *morc-1*, is required for transgene silencing (Moissiard et al., 2012) and has been identified in screens for regulators of exogenous RNAi (Guang et al., 2008; Kim et al., 2005). The upregulation of transposons observed in both mouse and plant *Morc* mutants suggests that Morc proteins play an ancient role in the repression of toxic genetic elements (Harris et al., 2016; Moissiard et al., 2012; 2014; Pastor et al., 2014). Morc proteins in plants, worms, and mammals are homologous to ATPases that alter chromatin superstructure (Dutta and Inouye, 2000; Iyer et al., 2008), and animal Morc proteins share a CW zinc finger that binds histone H3 (Andrews et al., 2016; Li et al., 2016; Liu et al., 2016). In *A. thaliana,* Morc proteins are required for compaction and silencing of pericentromeric heterochromatin, revealing a role for Morc proteins in regulation of chromatin architecture (Moissiard et al., 2012).

Here, we uncover *morc-1* as an effector of the nuclear RNAi pathway that acts downstream of the canonical *nrde* factors to regulate chromatin organization. We establish a critical role for *morc-1* in siRNA-mediated enforcement of transgenerational chromatin organization and germline immortality. Like *hrde-1(-)/nrde-1,2,4(-)* mutants, *morc-1(-)* mutants display a germline mortal phenotype. Further, we performed a forward genetic screen and identified four alleles of *met-1*, which encodes an H3K36 histone methyltransferase (Andersen and Horvitz, 2007), that suppress germline mortality in *morc-1(-)* and nuclear RNAi mutants . Our data support a model in which loss of nuclear RNAi or MORC-1 triggers ectopic, MET-1-dependent H3K36 trimethylation and chromatin decondensation that compromise germline maintenance.

We show that integration of endo-siRNAs with germline chromatin organization via MORC-1, and the subsequent antagonism of MET-1 activity, are critical for transgenerational germline maintenance.

RESULTS

morc-1 is required for nuclear RNAi

We previously demonstrated that MORC-1 is required for transgene silencing in worms (Moissiard et al., 2012). Given that transgene silencing depends on many RNAi pathway genes (Dernburg et al., 2000; Ketting and Plasterk, 2000), we hypothesized that MORC-1 may also play a role in the germline nuclear RNAi pathway and multigenerational epigenetic inheritance. First, we investigated MORC-1 expression and subcellular localization by introducing a 3xFlag tag at the C- terminus of MORC-1 using the CRISPR-Cas9 system (Paix et al., 2015). Immunofluorescence confirmed that MORC-1::3xFlag is expressed in the nuclei of germline and somatic cells (Figures 2.1 and 2.2), consistent with a possible germline function for MORC-1. Indeed, we found that *morc-1(-)* mutants exhibit a temperature-sensitive germline mortality phenotype, with a progressive decline in fertility with each generation at 25°C until they become completely sterile after 6-7 generations (Figure 2.3). This defect is shared by mutants of the germline nuclear RISC complex such as *hrde-1(-)* (Figure 2.3)*.*

To determine if MORC-1 participates in nuclear RNAi, we tested whether *morc-1(-)* mutant worms recapitulate key mutant phenotypes of canonical nuclear RNAi factors in addition to germline mortality, including resistance to nuclear RNAi and defective RNAi inheritance (Buckley et al., 2012; Burton et al., 2011; Guang et al., 2008). We evaluated nuclear RNAi function by using siRNAs that target nuclear, polycistronic pre-mRNAs. For instance, the non-essential gene *lir-1* is transcribed within a polycistronic pre-mRNA that includes the essential gene *lin-26*. siRNAs targeting *lir-1* lead to silencing of the *lir-1/lin-26* polycistronic pre-mRNA via nuclear RNAi, and trigger larval arrest and lethality due to loss of *lin-26* (Guang et al., 2008)*.* We found that *morc-*

1(-) mutants were highly resistant to *lir-1* nuclear RNAi, phenocopying *nrde-2(-)* mutants. This resistance was rescued with a *morc-1::gfp* transgene (Figure 2.4A). Furthermore, the genes *lin-15a* and *lin-15b* are expressed together within a polycistronic pre-mRNA, and loss of both genes is required to produce a multi-vulva (Muv) phenotype. RNAi against *lin-15b* is sufficient to cause Muv with almost complete penetrance in the enhanced RNAi mutant *eri-1(-); this* phenotype is fully suppressed in the nuclear RNAi mutant *nrde-2(-)* (Guang et al., 2010)*. Like *nrde-2(-)*, *morc-1(-)* fully suppressed the enhanced sensitivity of *eri-1(-)* mutants to *lin-15b* RNAi (Figure 2.4B).

We next tested whether *morc-1* is required for RNAi inheritance using worms expressing a germline *gfp::h2b* reporter. Exposure to *gfp* RNAi for only one generation is sufficient to silence this reporter for many subsequent generations (Ashe et al., 2012; Buckley et al., 2012) (Figure 2.5). Although *gfp* silencing was 100% penetrant in the P0 generation of *morc-1(-)* mutants, the maintenance of *gfp* silencing was defective, with most worms (67-95%) activating *gfp* expression in the F1 generation and nearly all worms (96-99%) expressing *gfp* by the F2 generation (Figures 2.6A and 2.6B). Taken together, these data indicate that MORC-1 functions in the germline nuclear RNAi pathway.

morc-1 mediates endo-siRNA effector function

Based on our finding that MORC-1 is required for transgenerational silencing, we next wanted to determine where in the nuclear RNAi pathway MORC-1 functions. To determine if MORC-1 affects siRNA inheritance, we fed *gfp* dsRNA to adult wildtype, *hrde-1(-),* and *morc-1(-)* mutants expressing the germline *gfp::h2b* reporter (Figure 2.5). We detected 22G siRNAs against *gfp* in parental (P0) and in the first and second inheriting generations (F1 and F2) by qRT-PCR. As expected, *hrde-1(-)* mutants displayed a small but consistent loss of anti-*gfp* siRNAs at the F2 generation (Figure 2A, right). In contrast, *morc-1(-)* mutants expressed wild-type levels of siRNAs in all three generations (Figures 2.7A and 2.7B). These data suggest that unlike *hrde-1,* which is required for inheritance of siRNA expression in progeny, *morc-1* is dispensable for

siRNA biogenesis and inheritance. Thus, *morc-1* functions downstream of exogenous siRNAs and the germline nuclear RISC complex.

To investigate whether *morc-1* functions as an effector of endo-siRNAs, we performed mRNA-seq of wildtype, *hrde-1(-)*, and *morc-1(-)* worms. To capture changes in gene expression as germline function declines, we collected worms after two generations at 25°C (F1 or "early generation") and after five generations at 25°C (F4/F5 or "late generation") (Figure 2.8). We compared mRNA-seq libraries from these samples using a significance threshold of p<0.05. Early generation *morc-1(-)* and *hrde-1(-)* worms showed upregulation of 284 and 379 mRNAs respectively (Figure 2.9), with significant enrichment for targets of the *mutator* pathway, which is required for amplification of 22G endo-siRNAs (Phillips et al., 2014; 2012; Zhang et al., 2011) (*hrde-1* targets: $p<6.6 \times 10^{-16}$, *morc-1* targets: p=0.00253, Fisher's test) (Figures 2.10A and 2.10B). Gene upregulation was more severe in late generation mutant worms, with 1,316 targets upregulated in *morc-1(-)* and 1,506 targets upregulated in *hrde-1(-)* (Figure 2.9). As with the early generation targets, these *morc-1(-)* and *hrde-1(-)*-upregulated genes were significantly enriched for targets of the *mutator*-dependent endo-siRNAs (*hrde-1* targets: $p=1.0758 \times 10^{-8}$, *morc-1* targets: $p<6.6 \times 10^{-16}$) (Figures 2.10C and 2.10D). In contrast, genes that were downregulated in *morc-1(-)* and *hrde-1(-)* mutants were not enriched for *mutator* targets (Figures 2.11A and 2.11B). These data indicate that MORC-1 is required for silencing some endo-siRNA target genes.

We wondered whether the upregulation of the *mutator* targets observed in *morc-1(-)* mutants might be explained by a loss of the corresponding 22G endo-siRNAs. We performed 5'-independent small RNA-seq and compared 22G endo-siRNA levels in wildtype worms to *morc-1(-)* and *hrde-1(-)* mutants. We quantified 22G endo-siRNA reads across the *mutator* target mRNAs that were significantly upregulated in *morc-1(-)* mutants at early or late generation. Although some of these 22G endo-siRNAs were downregulated in *morc-1(-)* mutants compared to wildtype (Figure 2.12), *hrde-1(-)* mutants exhibited a substantial reduction in more of the 22G endo-siRNAs corresponding to these mRNAs (Figure 2.12). Next, we quantified the 22G endo-siRNAs corresponding to *mutator* targets that were upregulated in the *hrde-1(-)* mutants. Again, loss of *morc-1* had much less of an effect among these 22G endo-siRNAs compared to

the loss of *hrde-1* (Figure 2.13). Thus, we conclude that the desilencing of *mutator* targets in *morc-1(-)* mutants cannot be fully explained by a loss of endo-siRNAs, consistent with our findings regarding exogenous RNAi. Analysis of global levels of *mutator*-dependent 22G endo-siRNAs, irrespective of mRNA regulation, showed some endo-siRNA depletion in late generation *morc-1(-)* mutants compared to wildtype (Figure 2.14). We posit that this reduction reflects the relative loss of germline tissue in *morc-1(-)* mutants compared to wildtype, rather than a specific role for *morc-1* in regulating endo-siRNA levels. Together, our data suggest that *morc-1* is required for target silencing but dispensable for siRNA biogenesis and inheritance, and thus must function downstream of RISC.

morc-1 is required for maintenance of H3K9me3 marks at a subset of *hrde-1* targets

Given that endo-siRNA levels were largely unaffected in *morc-1(-)* mutants, we asked if *morc-1* might regulate siRNA effector function at the level of chromatin. Based on our model that MORC-1 functions downstream of HRDE-1, we looked specifically at whether MORC-1 is (1) required for siRNA-directed methylation of histone 3 lysine 9 (H3K9) and (2) regulates a common set of targets with HRDE-1. We performed H3K9me3 chromatin immunoprecipitation followed by deep sequencing (ChIP-seq) on early and late generation wildtype, *morc-1(-)*, and *hrde-1(-)* worms. We analyzed the H3K9me3 signal over 1-kb windows across the entire genome. Using a 1.5-fold threshold and false-discovery rate (FDR) of <0.05, we identified 57 1-kb regions that were depleted of H3K9me3 in early generation *morc-1(-)* mutants compared to wildtype (Figure 2.15A). Late generation (F4) *morc-1(-)* mutants exhibited a much larger defect; 206 1-kb windows were depleted of H3K9me3 in late generation *morc-1(-)* compared to wildtype (Figures 2.15A, 2.16A, and 2.16B). We refer to these 206 regions as *morc-1-* dependent H3K9me3 targets. Importantly, these targets are strongly enriched for endo-siRNAs, confirming that MORC-1 regulates H3K9me3 at endo-siRNA targets (Figure 2.17A).

We observed a progressive increase in the number of H3K9me3-depleted sites from early to late generation *morc-1(-)* mutants. Furthermore, most of the H3K9me3-depleted regions in the early generation were also depleted in the late generation (40 of 57, Figure 2.15B). Relative to F1 wildtype, H3K9me3 levels at *morc-1*-dependent targets decreased significantly from F1 to F4 generation in *morc-1(-)* mutants (p=8.0x10^{-8}, Welch's t-test), supporting our conclusion that H3K9me3 loss in *morc-1(-)* mutants is progressive (Figure 2.18). These findings indicate that *morc-1* is required for maintenance of H3K9me3 at these targets.

Strikingly, of the 206 *morc-1*-dependent H3K9me3 targets, 194 were also *hrde-1(-)*-dependent, demonstrating that *morc-1* regulates a subset of *hrde-1* targets (Figure 2.17B). Relative to F1 wildtype, these targets were more severely H3K9me3-depleted in *hrde-1(-)* mutants than in *morc-1(-)* mutants but the effect was not progressive (Figure 2.18). Thus, in *morc-1(-)* mutants, a subset of *hrde-1* targets are progressively depleted of H3K9me3 (Figures 2.18, 2.19, and 2.20).

Similar to our mRNA-seq results, *hrde-1* affects a larger number of targets than *morc-1* in both generations (Figures 2.15A and 2.15B). We identified 744 *hrde-1*-dependent H3K9me3 targets from a comparison of late generation *hrde-1(-)* and wildtype (Figures 2.15A, 2.16A, and 2.16B). The majority of these (550) were *morc-1*-independent (Figure 2.17B). These targets exhibited progressive H3K9me3 depletion from F1 to F4 generation in *hrde-1(-)* mutants compared to wildtype (p<2.2x10^{-16}, Welch's t-test) but were unaffected in *morc-1(-)* mutants (Figures 2.18 and 2.20B). Taken together, these results indicate that *morc-1* is required for maintenance of H3K9me3 at a subset of endogenous nuclear RNAi targets.

morc-1 is required for heterochromatin localization and compaction

To further elucidate the role of *morc-1* and nuclear RNAi in chromatin organization, we investigated the localization of a heterochromatin reporter in embryos (Towbin et al., 2012). The high-copy *gfp-lacI::lacO* array is normally heterochromatinized and localized at the nuclear periphery in embryos; its proper

localization is dependent on methylation at H3K9 by SET-25 and MET-2 (Towbin et al., 2012). The ubiquitously expressed GFP-LacI fusion protein binds to the *lacO* sites, revealing the position of the array relative to the nuclear periphery (Meister et al., 2010). We examined the localization of GFP-LacI by fluorescence microscopy in *morc-1(-)* (F4 generation at 25°C)*, nrde-2(-)* (F1 generation at 25°C)*,* and *hrde-1(-)* (F1 generation at 25°C) mutant embryos compared to *set-25(-)* and *met-2(-);set-25(-)* controls. Approximately 90% of GFP-LacI foci were localized to the outermost third of the nucleus in wild-type embryos, whereas only ~70% of GFP-LacI foci were correctly localized in *morc-1(-), nrde-2(-),* and *hrde-1(-)* mutant embryos (Figures 2.21A-D). The degree of displacement of the array from the nuclear periphery observed in these mutants is similar to the defect that we and others observed in *set-25(-)* mutants (Figures 2.21A and 2.21B) (Towbin et al., 2012)*.* These data reveal that MORC-1 and nuclear RNAi are required for proper localization of heterochromatin to the nuclear periphery.

To examine the effects of *morc-1* and nuclear RNAi on the chromatin structure of endogenous sequences, we performed X chromosome paint fluorescence in situ hybridization (FISH) to evaluate compaction of the X chromosome. In hermaphrodite worms, both X chromosomes are targeted by the dosage compensation complex (DCC) to decrease their transcriptional activity by half (Meyer and Casson, 1986). As a result, the X chromosomes in hermaphrodites are highly compacted, occupying about 10% of the nuclear volume despite containing 18% of the genome (Lau et al., 2014). This compaction is also dependent on methylation at H3K9 (Snyder et al., 2016). We utilized this unique feature of the hermaphrodite X chromosomes to determine whether chromatin compaction is defective in *morc-1(-)* and *hrde-1(-)* mutants. The X chromosomes in *morc-1(-)* and *hrde-1(-)* hermaphrodites each occupied 17% and 20%, respectively, of the nuclear volume compared to 10% in wildtype hermaphrodites, indicating a significant compaction defect (*morc-1(-)* vs. wildtype: $p=1.7 \times 10^{-5}$, *hrde-1(-)* vs. wildtype $p=9.6 \times 10^{-6}$, student's t-test) (Figures 2.22A and 2.22B). Furthermore, *morc-1* and *hrde-1* RNAi caused a small but significant rescue of the male lethality triggered by aberrant targeting of the DCC to the single X chromosome in males (Carmi et al., 1998; Miller et al., 1988) (*morc-1* vs vector: p=0.0053, *hrde-1* vs vector: p=0.0018, student's t-test) (Figure 2.22C). The magnitude of rescue achieved by loss of MORC-1

or HRDE-1 is similar to that observed with loss of H3K9 methylation (Snyder et al., 2016). RNAi against *dpy-27,* which encodes a DCC component, induced a more substantial rescue of male lethality, as expected (Meyer and Casson, 1986). Thus, MORC-1 and HRDE-1 regulate condensation of endogenous chromatin.

Our attempts to directly identify MORC-1 genomic targets by ChIP-seq of MORC-1::3XFlag did not yield sufficient enrichment over background, suggesting that MORC-1 may interact with chromatin indirectly or in a highly transient manner. Intriguingly, co-immunofluorescence revealed that MORC-1::3xFlag localizes adjacent to, but not overlapping with, H3K9me3 in the distal germline (Figure 2.23). Furthermore, in pachytene nuclei which have highly condensed chromosomes that can be visualized with 4',6-diamidino-2-phenylindole (DAPI) staining, we observe MORC-1::3xFlag at the nuclear periphery and adjacent to, but not overlapping with, the condensed chromosomes, seemingly as a barrier around condensed chromatin (Figure 2.23). Together, these data indicate that MORC-1 regulates heterochromatin modifications, localization and compaction, although it may be physically excluded from highly condensed chromatin.

Mutations in the gene encoding MET-1 suppress *morc-1(-)* and *hrde-1(-)* germline mortality

To elucidate the mechanism behind the progressive fertility defect of *morc-1(-)* mutants (Figure 2.3), we performed a forward genetic screen to identify suppressors of *morc-1(-)* germline mortality. ENU-mutagenized *morc-1(-)* mutants were grown for 14 generations at 25°C (Figure 2.24A), thus selecting for suppressor mutations that enable robust fertility far beyond when unmutagenized *morc-1(-)* mutants become sterile (i.e. 6-7 generations at 25°C). We generated clonal lines from the most fertile worms and performed whole genome sequencing. We identified four independent alleles *(xk1-4)* causing missense or nonsense mutations in the gene encoding the H3K36me3 histone methyltransferase MET-1 (Andersen and Horvitz, 2007). Two of these alleles introduce

substitutions in the SET domain (*xk1* and *xk4).* The *xk4* allele contains two additional substitutions outside the SET domain (Figure 2.24B and Table 2.1).

We outcrossed the four *met-1* alleles ten times and then crossed them into *morc-1(-)*. All four *met-1* alleles potently suppressed *morc-1(-)* germline mortality (Figures 2.25A-C). To determine whether *met-1* also suppresses the canonical nuclear RNAi mutants, we crossed *met-1* alleles *xk2* and *xk4* into *hrde-1(-)* mutants. We found that both *met-1* alleles suppressed *hrde-1(-)* germline mortality (Figures 2.26A and 2.26B). However, the two publicly available *met-1* deletions alleles, *n4337* and *ok2172,* did not rescue *morc-1(-)* germline mortality (Figure 2.27A). Both *met-1* deletion alleles, but not *met-1* alleles *xk2* and *xk4*, severely compromised fertility at 25°C (Figures 2.25A, 2.25B, and 2.27A). We propose that a hypomorphic allele of *met-1*, rather than a complete loss-of-function allele, might be required to rescue *morc-1(-)* germline mortality. This is consistent with our finding that depletion of *met-1* by RNAi also suppressed *morc-1(-)* germline mortality (Figure 2.27B). These data suggest that at 25°C the germline can tolerate partial, but not complete, loss of *met-1* activity. Thus, depletion of MET-1 rescues germline mortality caused by loss of MORC-1 or of nuclear RNAi.

MET-1 mediates H3K36 trimethylation of some endo-siRNA target genes in the absence of MORC-1

We hypothesized that MET-1 might oppose nuclear RNAi by marking nuclear RNAi target genes with trimethylated H3K36, a hallmark of euchromatin. To test this hypothesis, we profiled H3K36me3 genome-wide in early and late generation wildtype, *morc-1(-),* and *hrde-1(-)* worms by ChIP-seq. We found that late generation *morc-1(-)* and *hrde-1(-)* mutants exhibited increased H3K36me3 compared to wildtype (Figure 2.28A). Using a 1.5-fold threshold and FDR<0.05, we identified 49 1-kb regions that are enriched for H3K36me3 in late generation *morc-1(-)* compared to wildtype (Figures 2.28A and 2.28B). We refer to these as *morc-1(-)*-dependent H3K36me3 regions. As with *morc-1* regulation of H3K9me3, *morc-1* regulates H3K36me3 at a subset of the loci that are also regulated by *hrde-1* (Figure 2.28C)*.* We observed a significant increase in

H3K36me3 levels at *morc-1(-)*-dependent sites from early to late generations in *morc-1(-)* mutants (p=6.44x10^{-11}, Welch's t-test) (Figures 2.28A and 2.29), similar to the progressive loss of H3K9 methylation in *morc-1(-)* mutants. These sites also exhibited elevated H3K36me3 levels at both generations in *hrde-1(-)* mutants, but there was no significant increase from early to late generation (Figure 2.29).

Comparing H3K36me3 levels in *hrde-1(-)* to wildtype we identified 384 *hrde-1(-)*-dependent H3K36me3-enriched regions (Figures 2.28A and 2.28B). Of these, the majority (350) are *morc-1(-)*-independent (Figure 2.28C). At these *morc-1(-)*-independent targets, there was a significant gain of H3K36me3 from early to late generation in both *hrde-1(-)* and *morc-1(-)* mutants (*hrde-1(-)*: p<2.2x10^{-16}, *morc-1(-)*: p=0.0012, Welch's t-test) (Figures 2.28A and 2.29). This suggests that *morc-1* regulates a larger proportion of *hrde-1(-)*-dependent sites than we have identified, but not severely enough to meet our 1.5-fold threshold for defining *morc-1(-)*-dependent sites. As a control, we used regions that have high levels of H3K36me3 in F1 wildtype; these regions did not show regulation by *hrde-1* or *morc-1* at either generation (Figure 2.29). Taken together, these data demonstrate that *morc-1(-)* and *hrde-1(-)* mutants progressively gain H3K36me3 at select loci.

Most regions that gained H3K36me3 in mutant backgrounds also lost H3K9me3. Of the 49 *morc-1(-)*-dependent H3K36me3 regions, 27 were depleted of H3K9me3 by >1.5-fold in *morc1(-)* mutants (Figures 2.30A and 2.30B). In *hrde-1(-)* mutants, 291 of the 384 H3K36me3-enriched regions were also H3K9me3-depleted (Figures 2.30A and 2.30B). For example, the endo-siRNA target *bath-45* was depleted of H3K9me3 and enriched for H3K36me3 in early and late generation *hrde-1(-)* mutants and in late generation *morc-1(-)* mutants relative to wildtype (Figure 2.31, top). Control regions showed no change in H3K36me3 enrichment (Figure 2.31, bottom). Thus, most nuclear RNAi targets that gain H3K36me3 also lose H3K9me3 in *morc-1(-)* and *hrde-1(-)* mutants (Figure 2.32).

We next asked how *morc-1* and *hrde-1* affect H3K9me3 and H3K36me3 marks over larger genomic intervals. To do this, we plotted relative H3K9me3 and H3K36me3 coverage 20kb up and downstream from the 5' end of the 744 *hrde-1*-dependent H3K9me3 targets (Figure 2.33). For visual clarity, we assign the 5' end of the 1kb

61

targets to position zero and plot average coverage at 10bp. We performed the same analysis with the 206 *morc-1*-dependent H3K9me3 targets (Figure 2.34). We observed that in the wildtype background, both *hrde-1* and *morc-1* targets occupy local maxima of H3K9me3 enrichment and local minima of H3K36me3 enrichment. In *morc-1(-)* and *hrde-1(-)* mutants, we observed strong depletion of H3K9me3 over a span of several kb and concurrent enrichment of H3K36me3, to a level similar to the flanking regions or to a local maximum (Figures 2.33 and 2.34). Taken together, these data show that loci that lose H3K9me3 in *hrde-1(-)* and *morc-1(-)* mutants also gain H3K36me3, suggesting that HRDE-1 and MORC-1 may repress encroachment of H3K36me3 from neighboring loci.

To determine whether the enrichment of H3K36me3 in *morc-1(-)* and *hrde-1(-)* backgrounds is *met-1*-dependent, we performed H3K36me3 ChIP-seq in early (F1) and late generation (F4) *met-1(xk4), met-1(xk4);morc-1(-)*, and *met-1(xk4);hrde-1(-)* mutants. First, we compared H3K36me3 levels at the 49 sites with *morc-1(-)*-dependent H3K36me3 enrichment. Indeed, the H3K36me3 enrichment observed at these loci in late generation *morc-1(-)* mutants was partially suppressed in *met-1(xk4);morc-1(-)* mutants (p=2.50x10^{-5}, Welch's t-test) (Figure 2.35A). Second, we compared H3K36me3 levels at the 384 loci with *hrde-1(-)*-dependent H3K36me3 enrichment. At these loci, the H3K36me3 enrichment observed in late generation *hrde-1(-)* mutants is also partially suppressed in *met-1(xk4);hrde-1(-)* mutants (F1: p=3.28x10^{-14}, F4: p=1.15x10^{-10}, Welch's t-test) (Figure 2.35B). These data indicate that MET-1 contributes to the gain of H3K36me3 upon loss of MORC-1 or HRDE-1.

MORC-1 restricts MET-1-dependent H3K36 trimethylation genome-wide

Having shown that the ectopic H3K36me3 observed in *morc-1(-)* and *hrde-1(-)* mutants is partially *met-1*-dependent, we next wanted to identify *met-1*-dependent H3K36 methylated regions in wildtype worms with an intact nuclear RNAi pathway. Therefore, we compared H3K36me3 in *met-1(xk4)* and wildtype worms by ChIP-seq. We observed a progressive loss of H3K36me3 from early to late generation in *met-1(xk4)* mutants compared to wildtype (Figure 2.36A), confirming that the *xk4* allele is in

fact hypomorphic and suggesting that *met-1* participates in transgenerational maintenance H3K36me3 marks. We identified 424 loci that were depleted of H3K36me3 in *met-1(xk4)* compared to wildtype (Figure 2.36B). Further, we observed a progressive increase in H3K36me3 in *morc-1(-)* and *hrde-1(-)* mutants compared to F1 wildtype at these regions (*morc-1(-)* F1 vs. F4: $p<2.2\times10^{-16}$, *hrde-1(-)* F1 vs F4: $p<2.2\times10^{-16}$, Welch's t-test) (Figure 2.37). The increased H3K36me3 at these loci is *met-1-*dependent (Figure 2.37). Thus, MET-1 activity at its normal targets in wildtype worms is increased upon loss of MORC-1 or HRDE1.

Next, we wanted to determine if the distribution of *met-1*-dependent H3K36 methylated regions are altered upon loss of MORC-1 or HRDE-1. Therefore, we performed ChIP-seq for H3K36me3 in *met-1(xk4);morc-1(-), morc-1(-), met-1(xk4);hrde-1(-)* and *hrde-1(-)* mutant backgrounds. We identified 1,942 regions that were depleted of H3K36me3 in *met-1(xk4);morc-1(-)* compared to *morc-1(-)* and 2,449 loci that were depleted of H3K36me3 in *met-1(xk4);hrde-1(-)* compared to *hrde-1(-)* (Figures 2.38A and 2.38B). These data indicate that loss of MORC-1 or HRDE-1 dramatically increases the number of MET-1-dependent loci (Figure 2.39). Taken together, these data suggest that the chromatin decompaction induced by loss of MORC-1 and HRDE-1 is associated with an increase in MET-1-dependent H3K36 trimethylation throughout the genome.

MET-1 depletion restores chromatin organization in *morc-1(-)* mutants

We next asked how the increased H3K36 methylation in *morc-1(-)* affected germline chromatin. We performed whole-mount DAPI stating of adult wildtype, *morc-1(-),* and *met-1(xk4);morc-1(-)* worms at the F5 generation at 25°C. All the *morc-1(-)* mutants were sterile and exhibited pleiotropic germline defects. Consistent with the decondensation of the X chromosome in intestinal nuclei of *morc-1(-)* hermaphrodites (Figure 2.22A), distal germline nuclei were enlarged and lacked the uniformity or regular spacing of wildtype germline nuclei (Figure 2.40). In the loop region and the proximal germline, we did not observe the highly condensed chromosomes or the single-file line

of nuclei that are characteristic of diplotene and diakinetic nuclei, indicative of failure to properly complete meiotic prophase (Figure 2.40). In contrast, these defects were largely restored in *met-1(xk4);morc-1(-)* mutants (Figure 2.40). Thus, in the absence of MORC-1 a partial loss of the histone H3K36 methyltransferase MET-1 and global reduction of H3K36me3 is sufficient to compensate for reduced transgenerational maintenance of H9K9me3, and to restore germline immortality and chromatin organization.

MORC-1::3xFlag

| vector RNAi | *morc-1* RNAi |

DAPI — DAPI

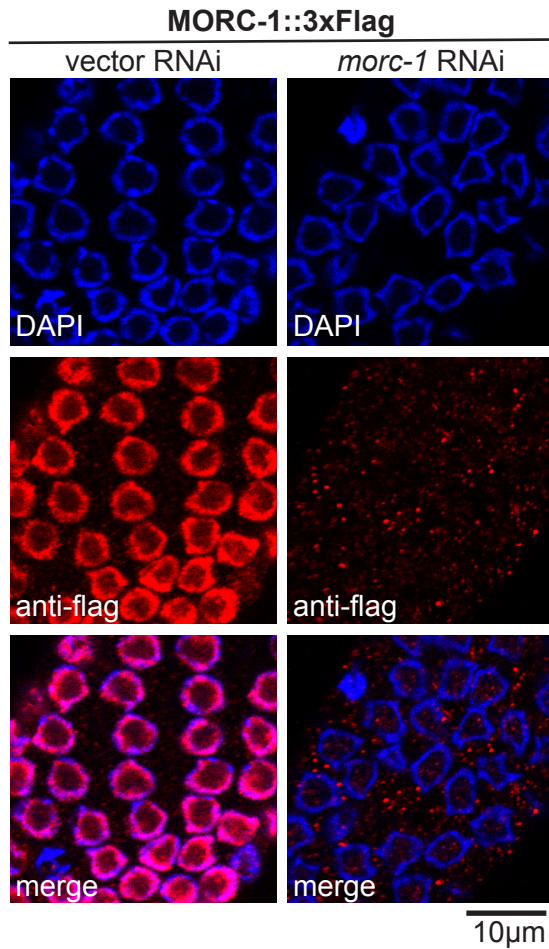anti-flag — anti-flag

merge — merge

10µm

Figure 2.1. MORC-1 is expressed in germline nuclei. Anti-Flag immunofluorescence (middle panels) and DAPI staining (top panels) of worms expressing *morc-1::3xFlag* at the endogenous *morc-1* locus and grown for two generations on empty vector RNAi (left panels) or on *morc-1* RNAi (right panels).

MORC-1::3xFlag

vector RNAi | *morc-1* RNAi

DAPI

DAPI

anti-flag

anti-flag

merge

merge

10μm
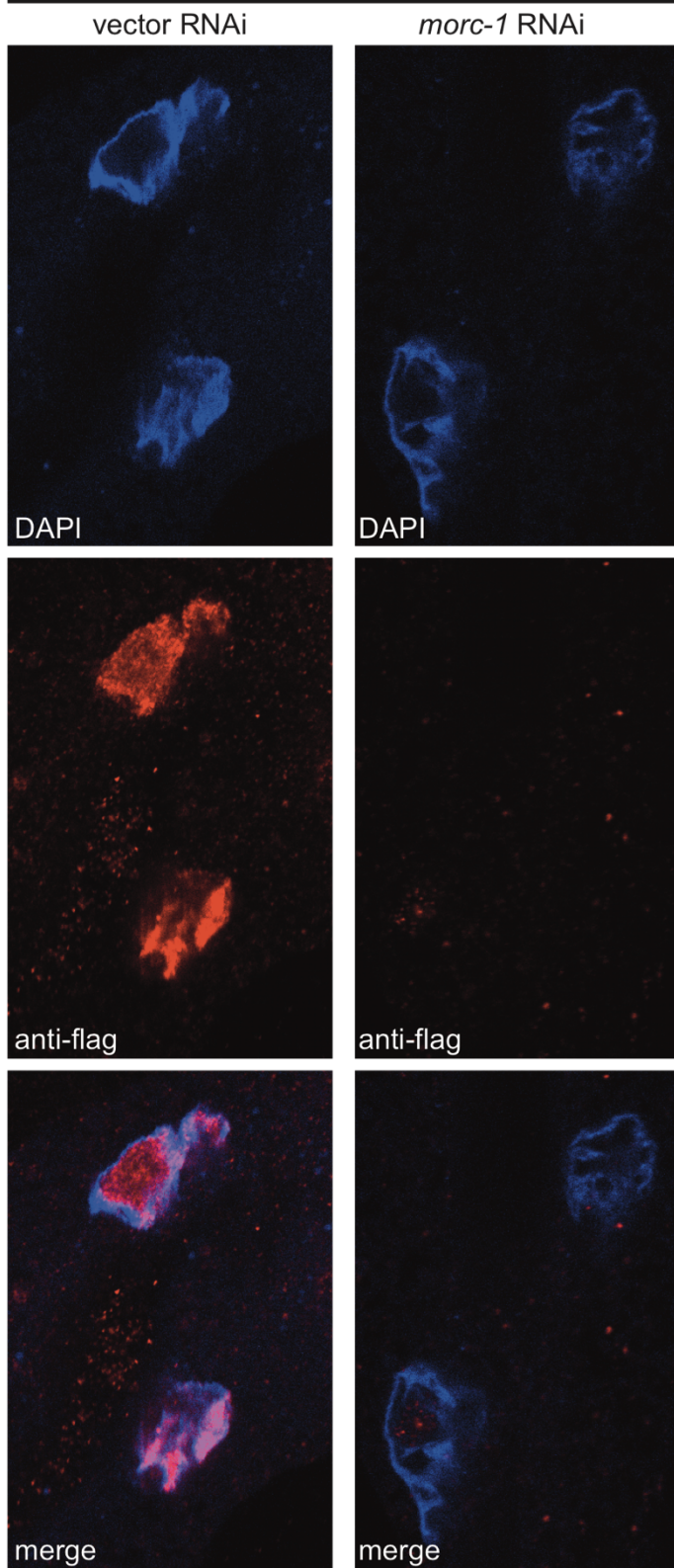
Figure 2.2. MORC-1 is expressed in intestinal nuclei. Anti-flag immunofluorescence (middle panels) and DAPI staining (top panels) of worms expressing *morc-1::3xflag* at the endogenous *morc-1* locus and grown for two generations on empty vector RNAi (left panels) or on *morc-1* RNAi (right panels).

Figure 2.3. *morc-1* **is required for germline immortality.** *morc-1(-)* (left) and *hrde-1(-)* (right) are germline mortal at 25°C. Worms were grown at 25°C beginning with starved L1s in the P0 generation. Symbols indicate total number of live progeny for individual, self-fertilized worms at the indicated generation. Error bars represent mean +/- standard deviation

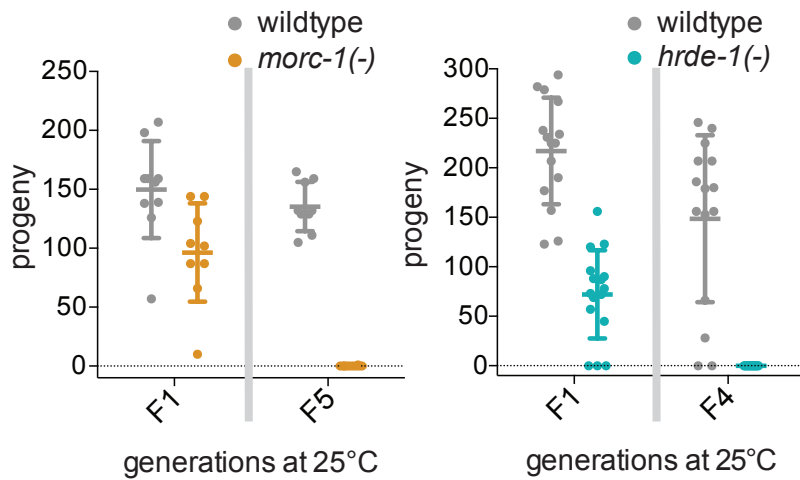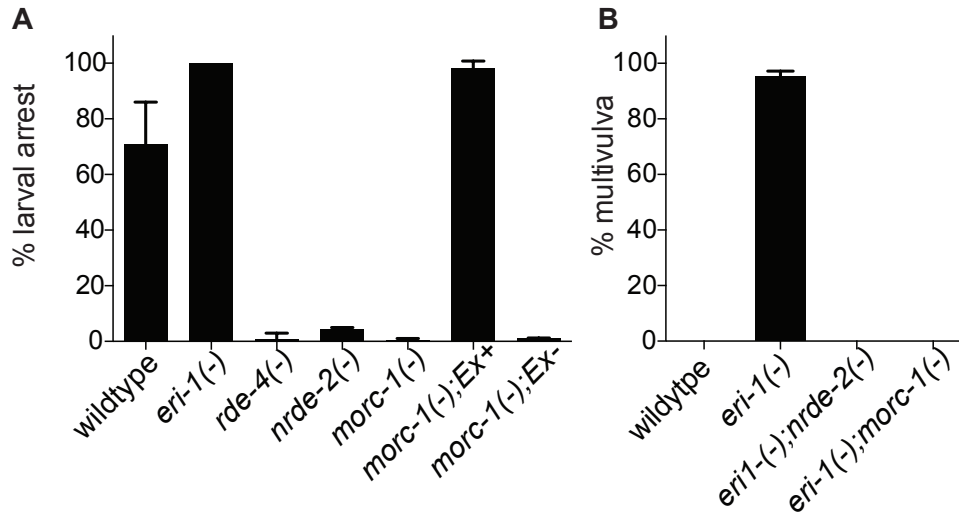**Figure 2.4. *morc-1* is required for nuclear RNAi. (A)** *morc-1(-)* mutants are resistant to *lir-1* RNAi. Expression of a *morc-1::gfp* extrachromosomal array driven by the ubiquitous *dpy-30* promoter rescues *morc-1(-)* resistance to *lir-1* RNAi. Non-transgenic siblings are resistant to *lir*-1 RNAi. Percent larval arrest indicates the proportion of worms that are arrested or dead. **(B)** *morc-1(-)* suppresses the sensitivity of *eri-1(-)* to *lin-15b* RNAi after two generations on *lin-15b* RNAi. Percent multivulva indicates the proportion of F1 adults with multiple vulvas.
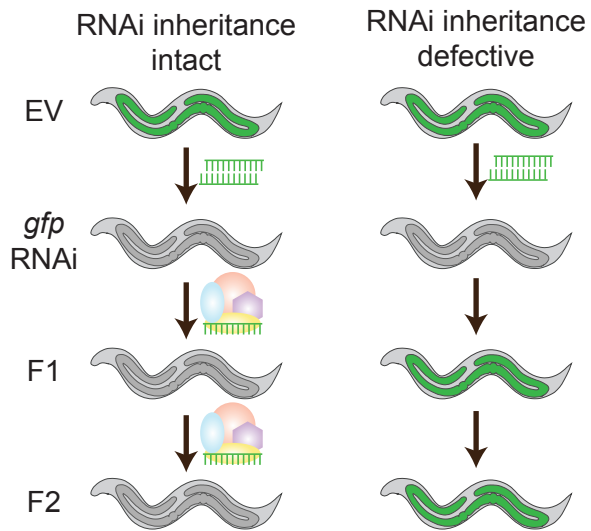
Figure 2.5 Model for RNAi inheritance experiments. Worms expressing *gfp::h2b* in the germline are grown on *gfp* RNA for one generation. Their progeny are grown for two generations on OP50 food. When RNAi inheritance is intact, *gfp* silencing is maintained in the F1 and F2 progeny. When RNAi inheritance is defective, *gfp* is silenced in the P0 generation but that silencing is not inherited in the progeny.
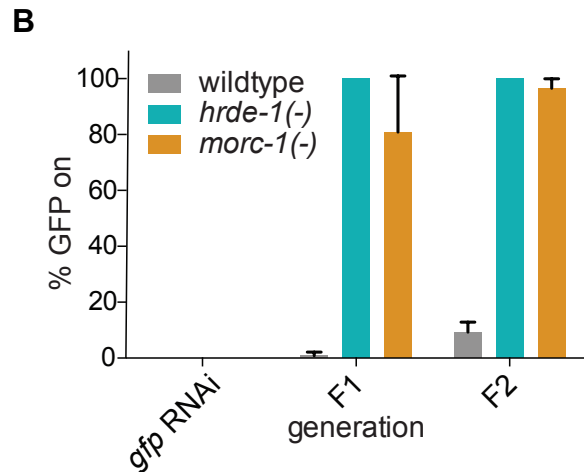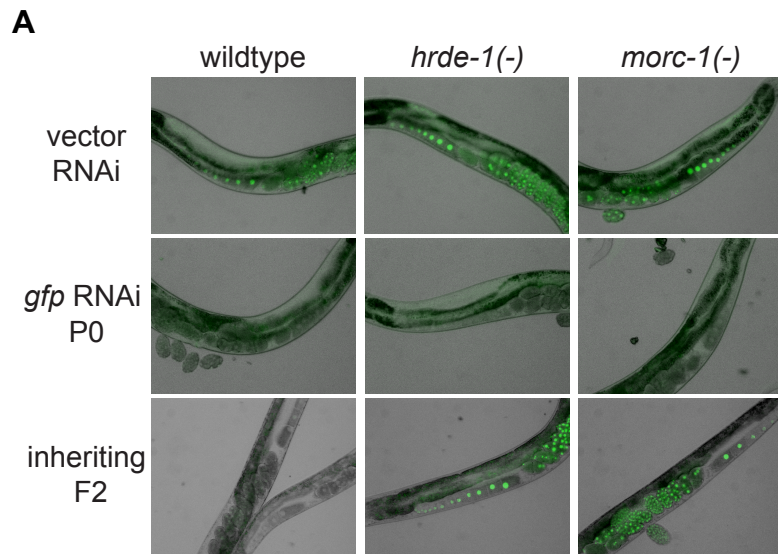
Figure 2.6. MORC-1 is required for RNAi inheritance. (A) *morc-1(-)* and *hrde-1(-)* are resistant to RNAi inheritance. On *gfp* RNAi, silencing of *gfp* is completely penetrant. Both *morc-1(-)* and *hrde-1(-)* fail to maintain *gfp* silencing in F1 and F2 generations. Quantification is shown in **(B)**.
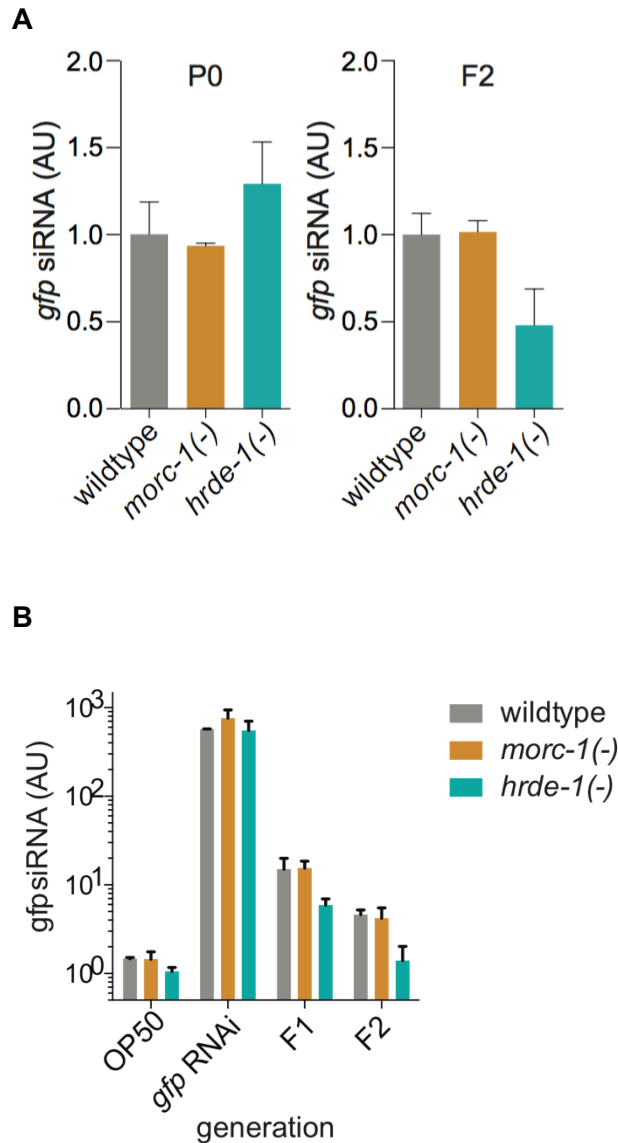
Figure 2.7 MORC-1 is dispensable for siRNA biogenesis and inheritance. (A) Quantification of 22G siRNAs targeting *gfp* based on Taqman qRT-PCR normalized to U18 at P0 (left) and F2 (right) generations. 22G levels are show in arbitrary units with wildtype type levels set to one. **(B)** Levels of anti-gfp siRNAs in adult worms grown on *gfp* RNAi and in F1 and F2 generations grown on OP50. siRNA levels measured by Taqman qRT-PCR and normalized to U18.
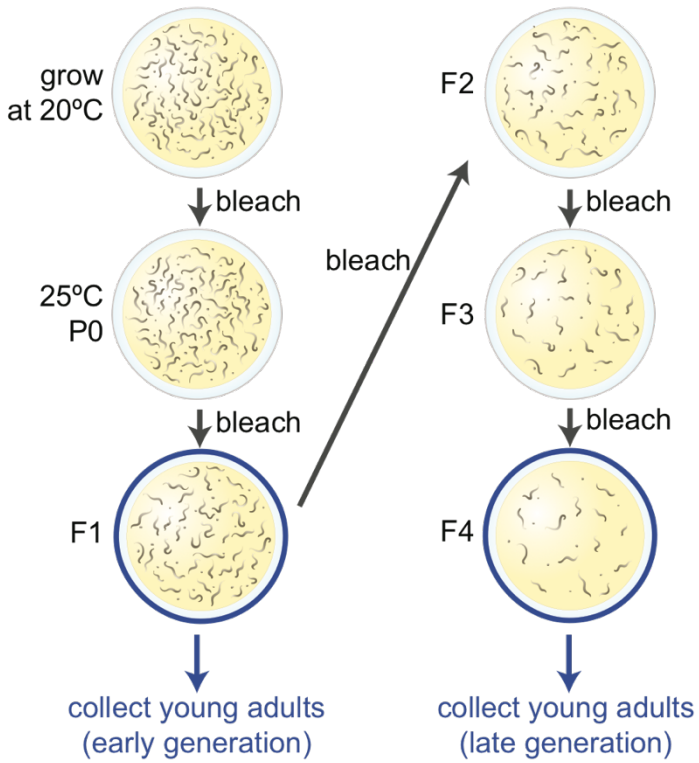
Figure 2.8 Sample collection schematic for RNA-seq and ChIP-seq experiments. Worms were shifted to 25°C as synchronized, starved L1s (P0 generation). Every generation, adult worms were bleached to isolate embryos. Young adults at F1 and F4 generations were collected for total RNA extraction and chromatin IP.

A mRNAs, F1

hrde-1(-) vs wildtype   *morc-1(-)* vs wildtype

log$_2$(fold change)

normalized average FPKM

B mRNAs, F4

hrde-1(-) vs wildtype   *morc-1(-)* vs wildtype

log$_2$(fold change)

normalized average FPKM
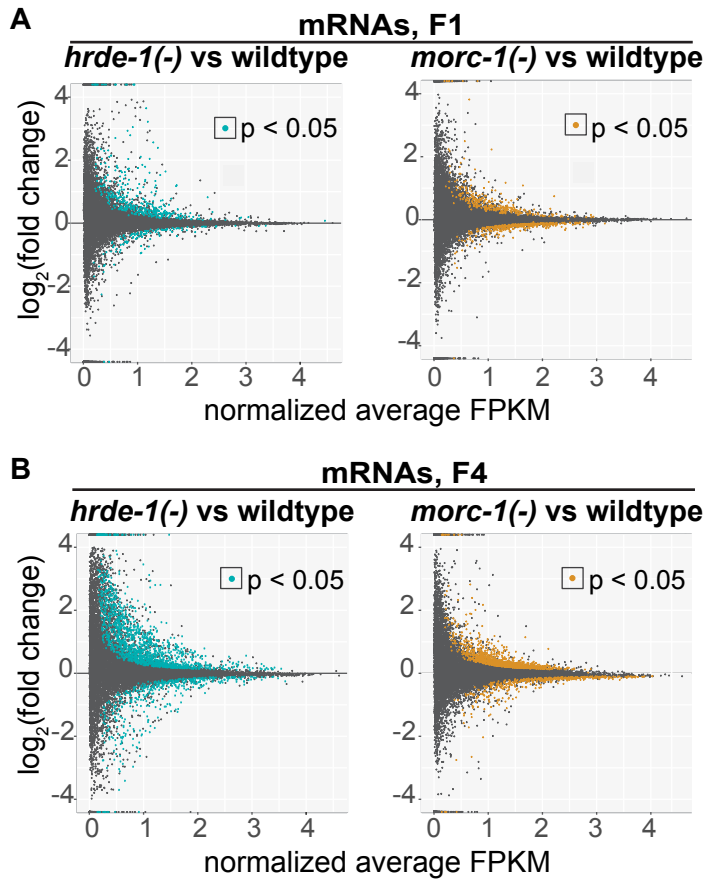
Figure 2.9 Identification of *morc-1*-regulated mRNAs. (A) Log2(fold change) as a function of normalized average RPKM *hrde-1(-)* (left) and *morc-1(-)* (right) compared to wildtype in F1 generation. mRNAs that are up- or down-regulated with p<0.05 in *hrde-1(-)* or *morc-1(-)* are indicated in blue and yellow, respectively. **(B)** Log2(fold change) as a function of normalized average RPKM *hrde-1(-)* (left) and *morc-1(-)* (right) compared to wildtype in F4 generation. mRNAs that are up- or down-regulated with p<0.05 in *hrde-1(-)* or *morc-1(-)* are indicated in blue and yellow, respectively.
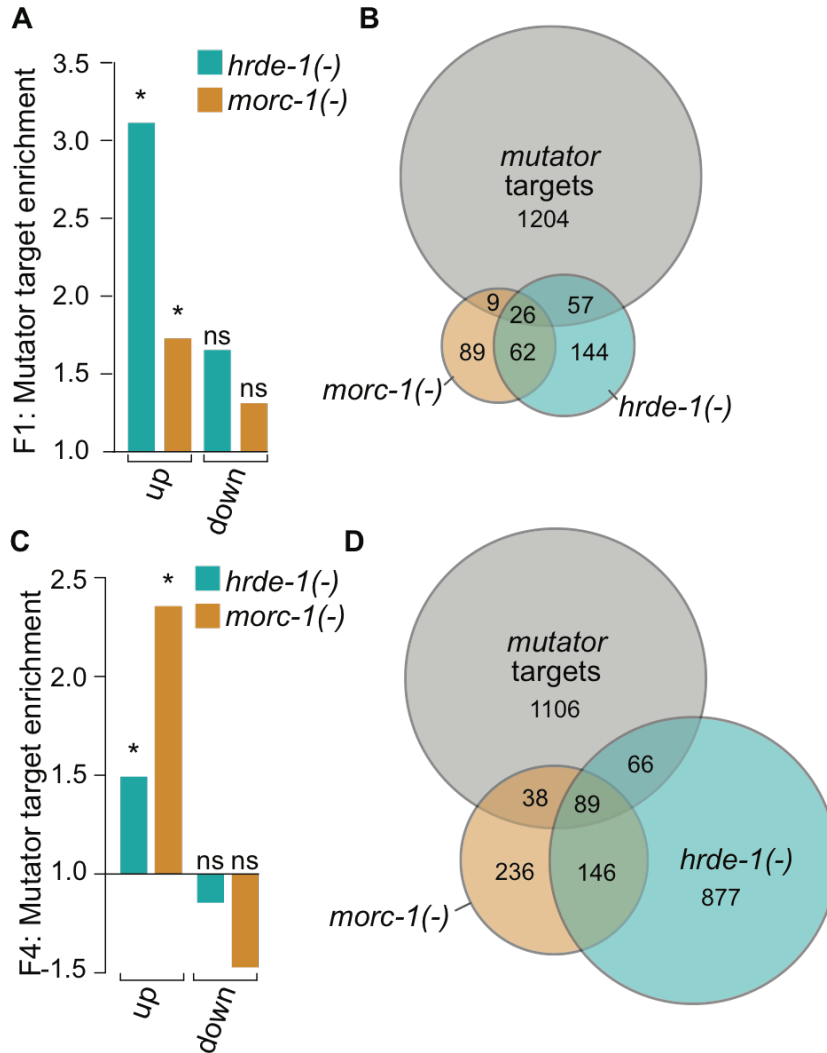
Figure 2.10. *morc-1(-)* upregulated mRNAs are enriched for *mutator* targets. (A) Upregulated mRNAs in F1 *hrde-1(-)* and *morc-1(-)* mutants are significantly enriched for targets of the *mutator* pathway (*hrde-1(-)* targets: $p<6.6\times10^{-16}$, *morc-1(-)* targets: $p=0.00253$, Fisher's test). Downregulated genes are not enriched (*hrde-1(-)* targets: $p=0.263$, *morc-1(-)* targets: $p=0.785$, Fisher's test). **(B)** Venn diagram of overlap of *hrde-1(-)* and *morc-1(-)* upregulated targets with all testable *mutator* targets in F1 libraries. **(C)** Upregulated mRNAs in F4 *hrde-1(-)* and *morc-1(-)* mutants are significantly enriched for targets of the *mutator* pathway (*hrde-1(-)* targets: $p=1.0758\times10^{-8}$, *morc-1(-)* targets: $p<6.6\times10^{-16}$, Fisher's test). Downregulated genes are not enriched (*hrde-1(-)* targets: $p=1$, *morc-1(-)* targets: $p=1$, Fisher's test) **(D)** Venn diagram of overlap of *hrde-1(-)* and *morc-1(-)* upregulated targets with all testable *mutator* targets in F4 libraries.
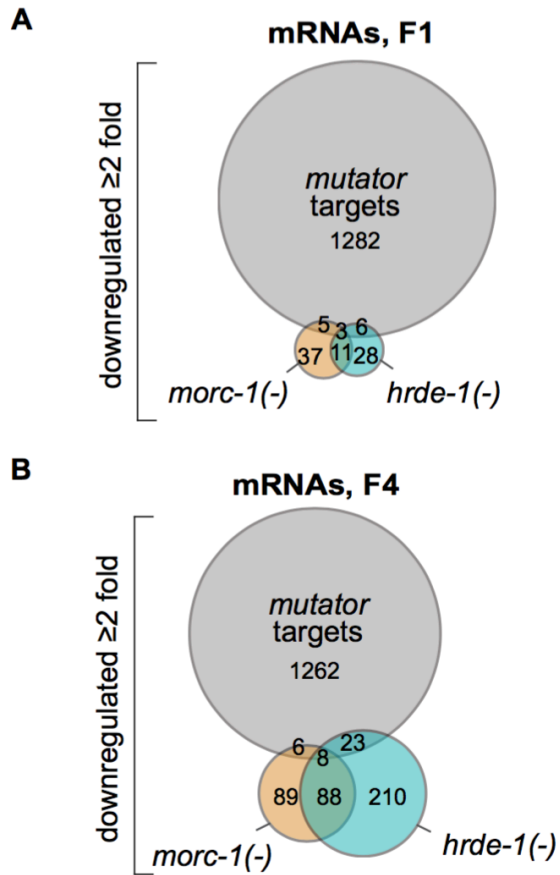
Figure 2.11. *morc-1(-)* downregulated targets are not enriched for *mutator* targets.
(A) Overlap of testable *mutator* targets with downregulated mRNAs in *morc-1(-)* and *hrde-1(-)* in F1and **(B)** F4.
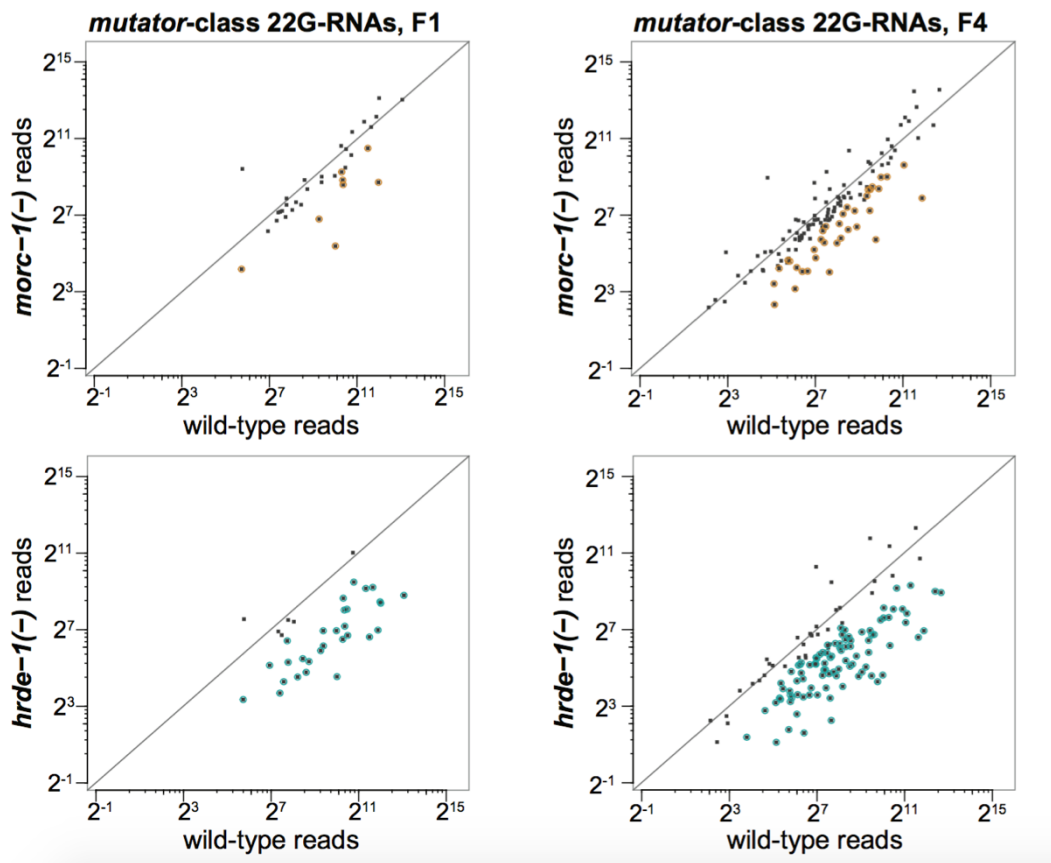
Figure 2.12. *morc-1* does not regulate 22G endo-siRNAs targeting *morc-1(-)* upregulated mRNAs. Quantification of 22G endo-siRNA reads across *mutator* targets that are upregulated in *morc-1(-)* mutants at F1 and F4. Upper panels, 22G endo-siRNA levels in *morc-1(-)* vs. wildtype. Lower panels, 22G endo-siRNA reads along the same targets in *hrde-1(-)* vs. wildtype. 22G endo-siRNAs that are depleted in *morc-1(-)* relative to wildtype are highlighted in yellow. 22G endo-siRNAs that are depleted in *hrde-1(-)* relative to wildtype are highlighted in blue.
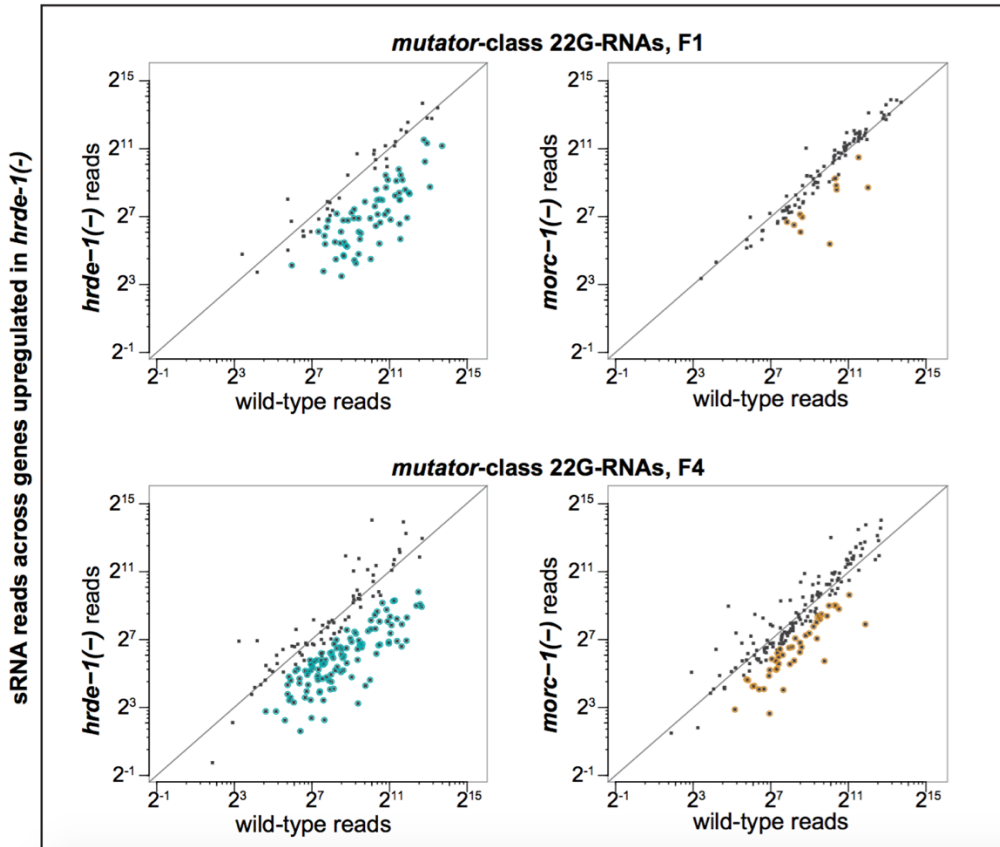
Figure 2.13. *morc-1* does not regulate 22G endo-siRNAs targeting *hrde-1(-)* upregulated mRNAs. 22G endo-siRNA levels corresponding to *mutator* targets that are upregulated in *hrde-1(-)* at F1 (top) and F4 (bottom) in *hrde-1(-)* and *morc-1(-)* relative to wildtype. Significantly depleted 22G endo-siRNAs in *hrde-1(-)* are highlighted in blue, significantly depleted 22G endo-siRNAs in *morc-1(-)* are highlighted in yellow.
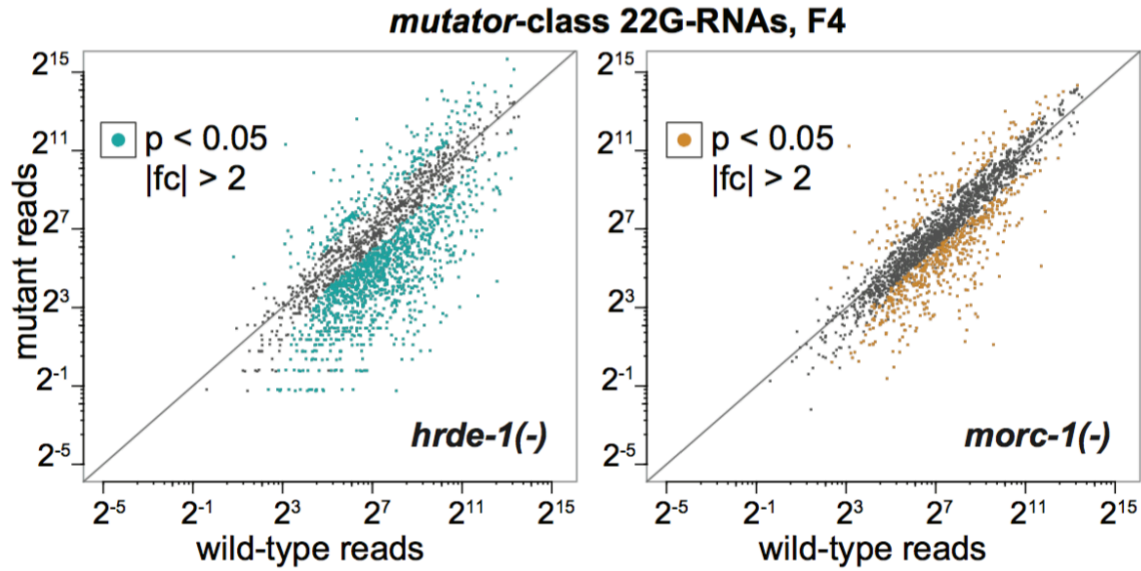
Figure 2.14. ***mutator*-class 22G endo-siRNA levels in late generation *morc-1(-)* mutants.** Levels of all *mutator*-class 22G endo-siRNAs in *hrde-1(-)* and *morc-1(-)* relative to wildtype. Significantly depleted 22G endo-siRNAs in *hrde-1(-)* are highlighted in blue, significantly depleted 22G endo-siRNAs in *morc-1(-)* are highlighted in yellow.

A

H3K9me3 ChIP-seq

● *morc-1*-dependent H3K9me3, F4
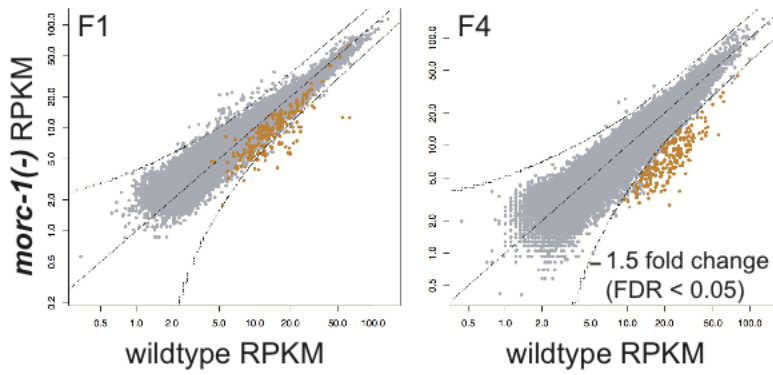
F1 ... F4 — 1.5 fold change (FDR < 0.05)

morc-1(-) RPKM vs wildtype RPKM

● *hrde-1*-dependent H3K9me3, F4

F1 ... F4 — 1.5 fold change (FDR < 0.05)

hrde-1(-) RPKM vs wildtype RPKM

B

morc-1 H3K9me3 targets

F1 targets   F4 targets

17 — 40   166

hrde-1 H3K9me3 targets

F1 targets   F4 targets

154 — 478 — 266
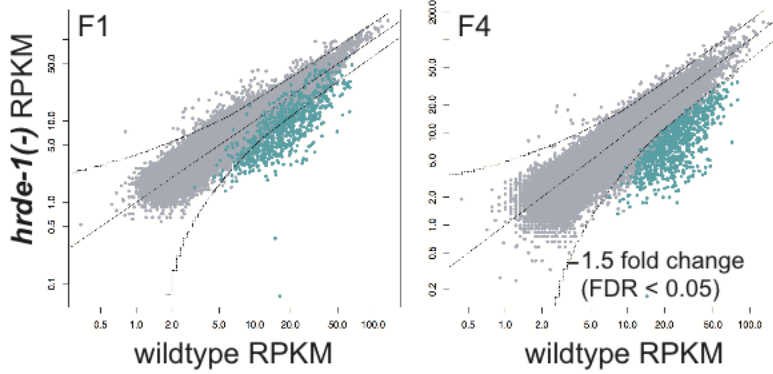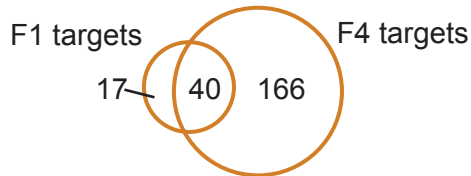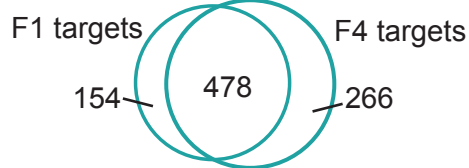
Figure 2.15. *morc*-1 regulates transgenerational H3K9me3 maintenance. (A) RPKM of 1kb windows in wildtype (x-axis) vs mutant (y-axis) H3K9me3 ChIP-seq libraries in F1 (left) and F4 (right) generations. Indicated in yellow and blue are 1kb regions that are more than 1.5-fold-depleted in F4 *morc-1(-)* and *hrde-1(-)* respectively with false discovery rate of <0.05 in two biological replicates. **(B)** Overlap of regions that are depleted of H3K9me3 in F1 and F4 generations in *morc-1(-)* (top) and *hrde-1(-)* (bottom) compared to wildtype.
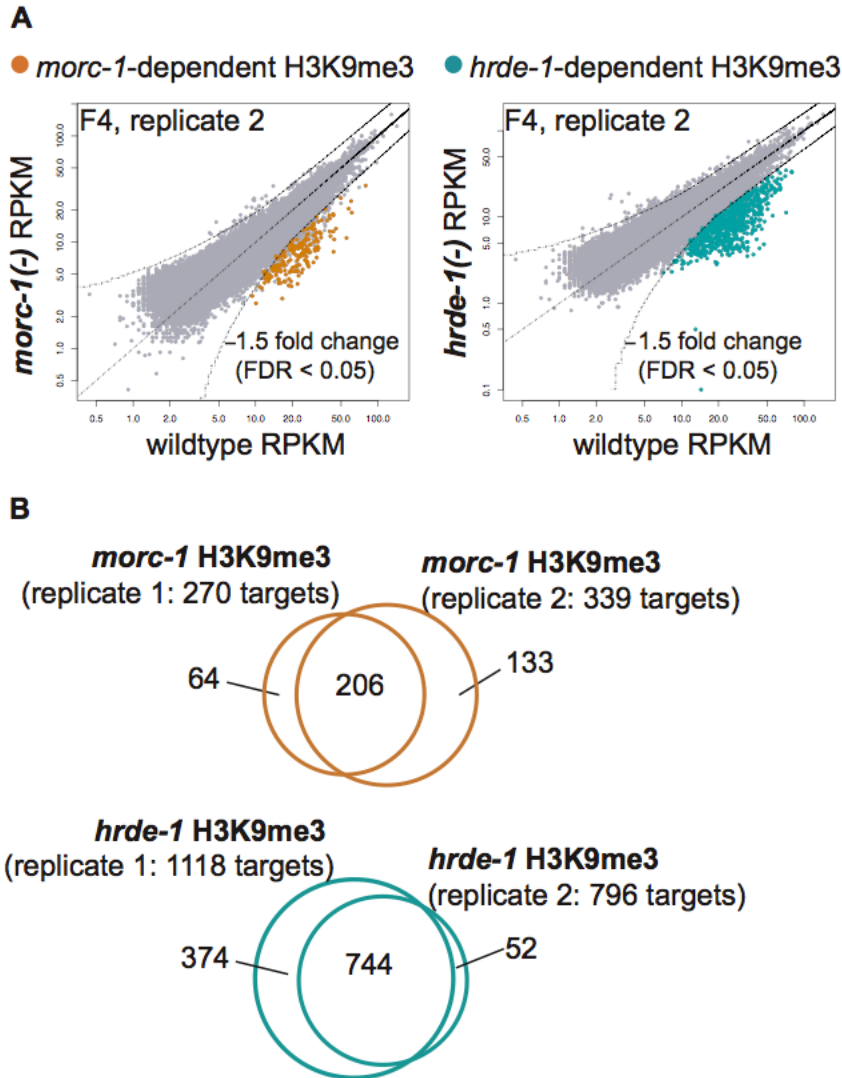
Figure 2.16. Replicate overlap for *morc-1-* and *hrde-1-*dependent H3K9me3 targets. (A) Scatter plots showing H3K9me3 levels of 1kb regions in F4 *morc-1(-)* and *hrde-1(-)* mutants compared to wildtype in the second biological replicate. Highlighted in yellow and blue are targets that are more than 1.5-fold depleted in *morc-1(-)* and *hrde-1(-)* mutants, respectively, in both biological replicates. **(B)** Venn diagrams showing overlap between replicates of H3K9me3-depleted regions in F4 mutant compared to wildtype (top: *morc-1(-)* vs wildtype, bottom: *hrde-1(-)* vs wildtype). Subsequent analyses were performing using targets that were identified in both biological replicates.
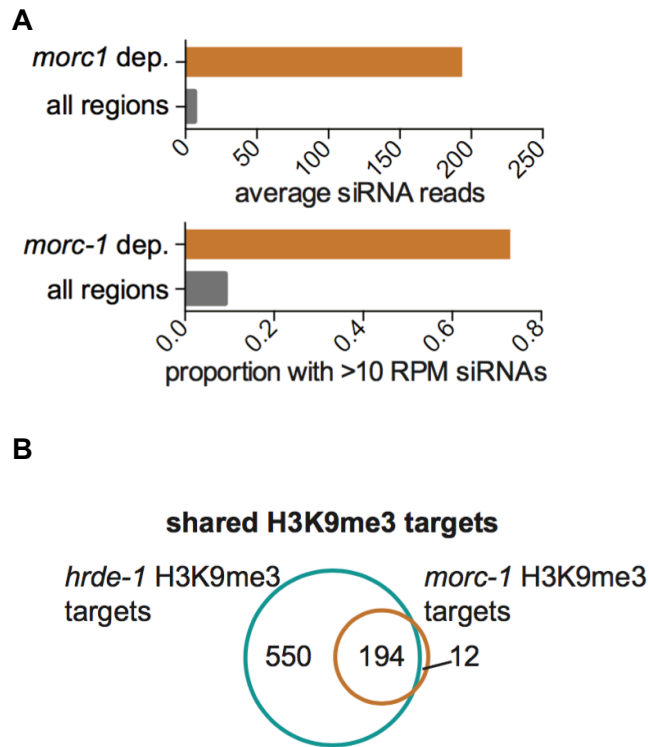
Figure 2.17. *morc-1*-dependent **H3K9me3 loci are targets of endo-siRNAs and** *hrde-1.* **(A)** *morc-1*-dependent H3K9me3 regions are enriched for siRNAs. Top: average wildtype siRNA reads across *morc-1*-dependent H3K9me3 1kb regions compared to all 1kb regions in the genome. Bottom: proportion of *morc-1*-dependent H3K9me3 1kb regions with wildtype siRNA reads over 10 RPM compared to proportion of 1kb regions over the entire genome with wildtype siRNA reads over 10 RPM. **(B)** Overlap of *morc-1(-)* and *hrde-1(-)* H3K9me3-depleted regions shows that MORC-1 regulates a subset of HRDE-1 targets.
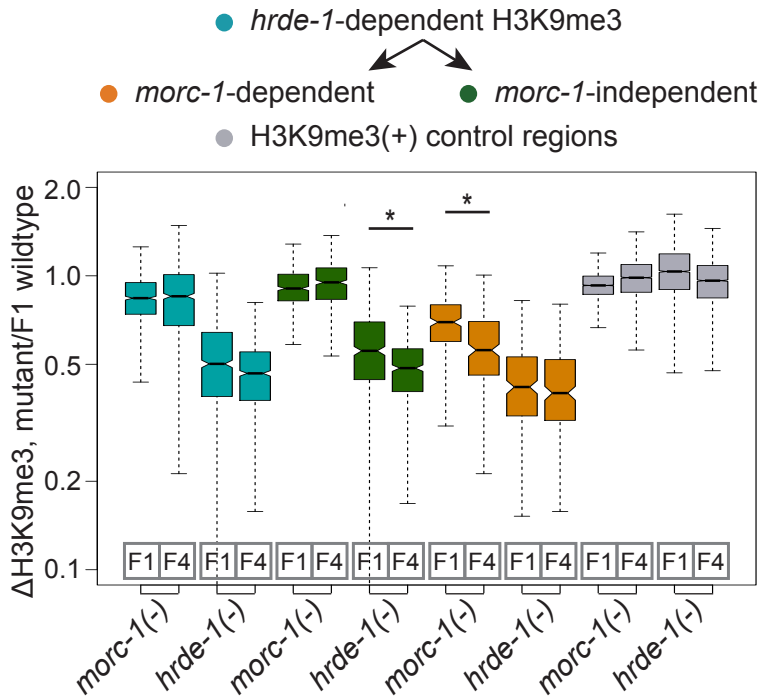
Figure 2.18. H3K9m3 loss in *morc-1(-)* and *hrde-1(-)* mutants is progressive. Box plot representing the ratio of recovered reads of the specified target group in the indicated mutant at the indicated generation compared to reads wildtype (F1). Blue boxes indicate mutant/F1 wildtype ratios in all *hrde-1*-dependent regions. Green boxes indicate mutant/F1 wildtype ratios in *hrde-1*-dependent, *morc-1*-independent regions, which show progressive loss of H3K9me3 from F1 to F4 in *hrde-1(-)* mutants (p<2.2x10^{-16}, Welch's t-test). Yellow boxes indicate mutant/F1 wildtype ratios in *hrde-1*-dependent, *morc-1*-dependent regions. These targets are significantly more depleted of H3K9me3 in *morc-1(-)* F4 compared to *morc-1(-)* F1 (p=8.0x10^{-8}, Welch's t-test). Gray boxes show H3K9me3 positive control regions that were previously found to be the 5% most H3K9me3-enriched regions in the entire genome.
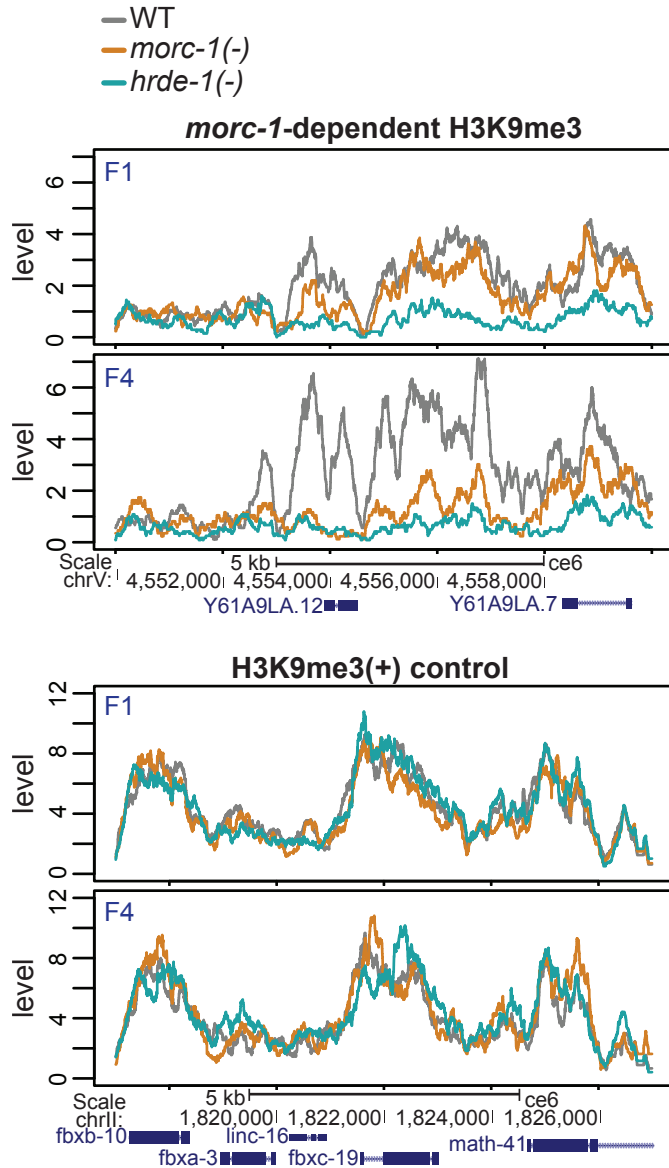
Figure 2.19. Exemplary *morc-1*-dependent H3K9me3 and control regions. H3K9me3 levels in the indicated genetic background at F1 and F4 in a *morc-1*-dependent region (top) and an unregulated control region (bottom).
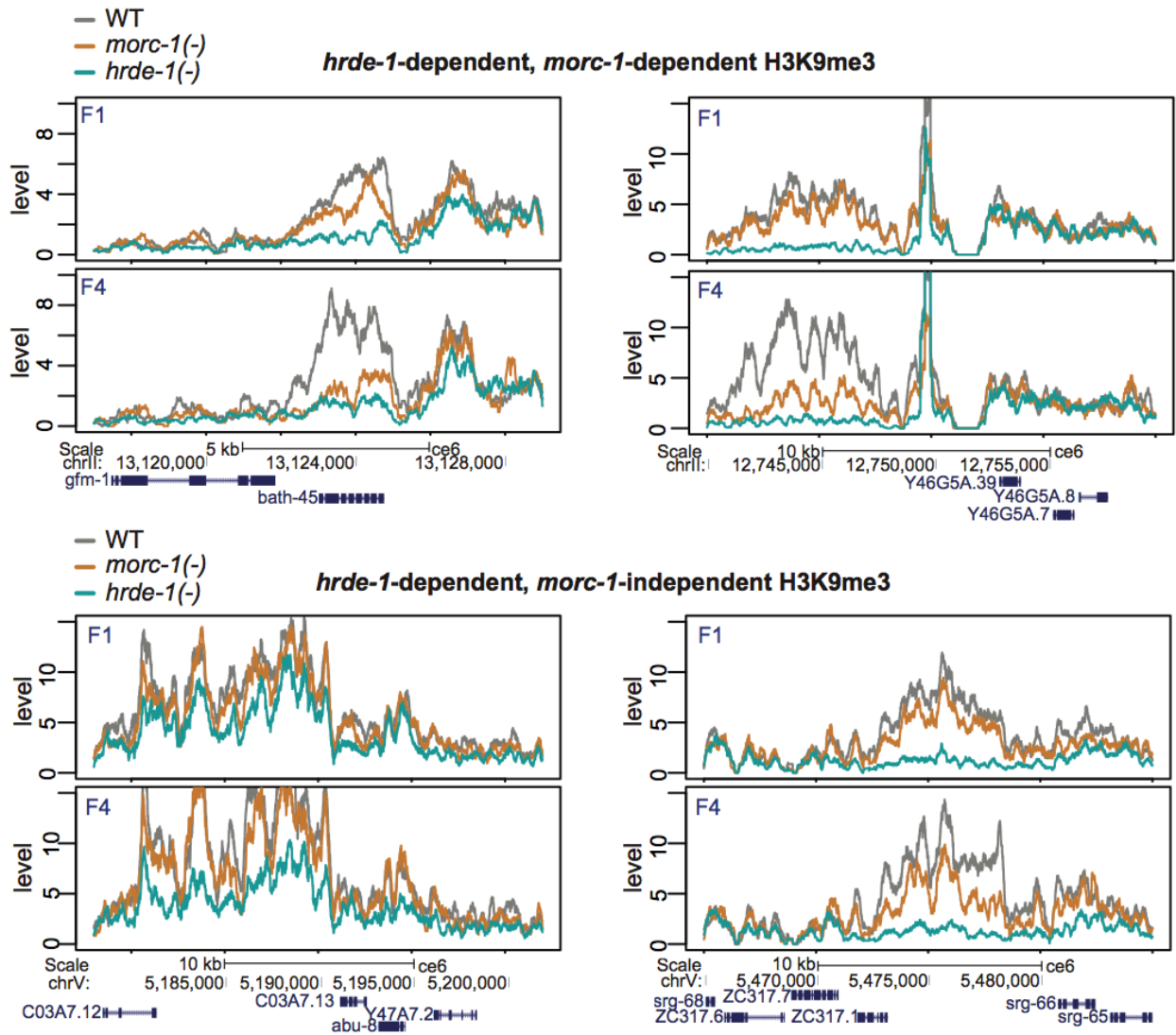
Figure 2.20. Exemplary *morc-1*-dependent and *morc-1*-independent H3K9me3 targets. H3K9me3 levels at exemplary *morc-1*-dependent H3K9me3 targets at early and late generation (top). H3K9me3 levels at exemplary *hrde-1*-dependent, *morc-1*-independent H3K9me3 targets at early and late generation (bottom).

85

A

wildtype   *morc-1(-)*   *set-25(-)*

GFP
DAPI

5µm

B

region 1
region 2
region 3

% GFP foci

100
80
60
40
20
0

wildtype   *morc-1(-)*   *set-25(-)*

C

wildtype   *hrde-1(-)*   *nrde-2(-)*   *met-2(-);set-25(-)*

GFP
DAPI

5µm

D

% GFP foci

100
80
60
40
20
0

wildtype   *nrde-2(-)*   *hrde-1(-)*   *met-2(-); set-25(-)*

86
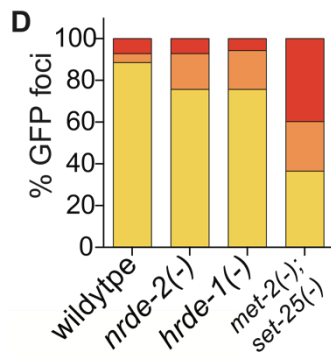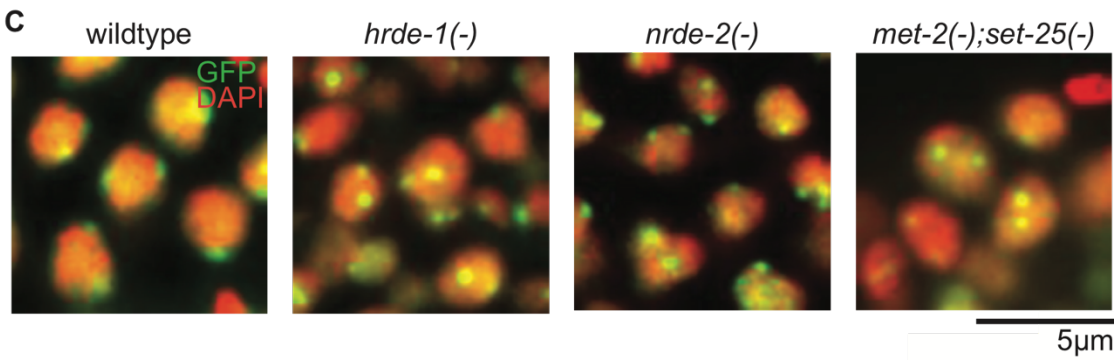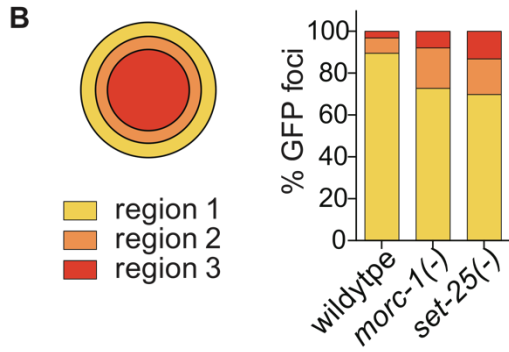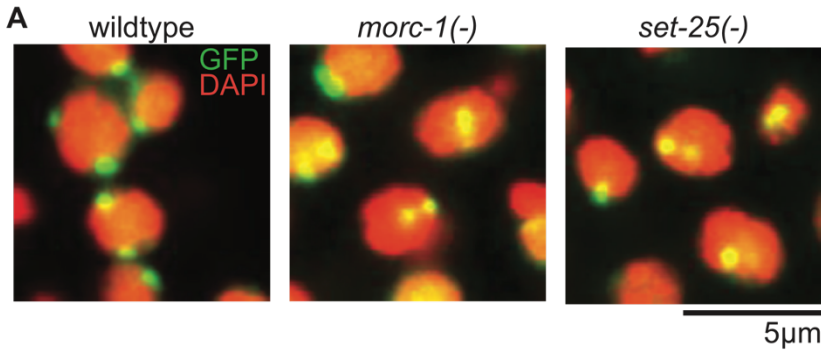
Figure 2.21. *morc-1* **is required for heterochromatin localization. (A)** Heterochromatin localization is defective in *morc-1(-)* mutants. DAPI stained embryo nuclei of the indicated genotype expressing a high-copy *gwIs4[gfp::lacI::lacO]* array at F4 generation at 25°C. *set-25(-)* is a control for mislocalization **(B)** Quantification is performed by determining the proportion of GFP foci contained in each of three regions of equal volume. **(C)** Localization of *gwIs4* in *hrde-1(-)* and *nrde-2(-)* mutant embryos at F1 generation at 25°C using *met-2(-);set-25(-)* as a control for mislocalization. Quantification in **(D)**.
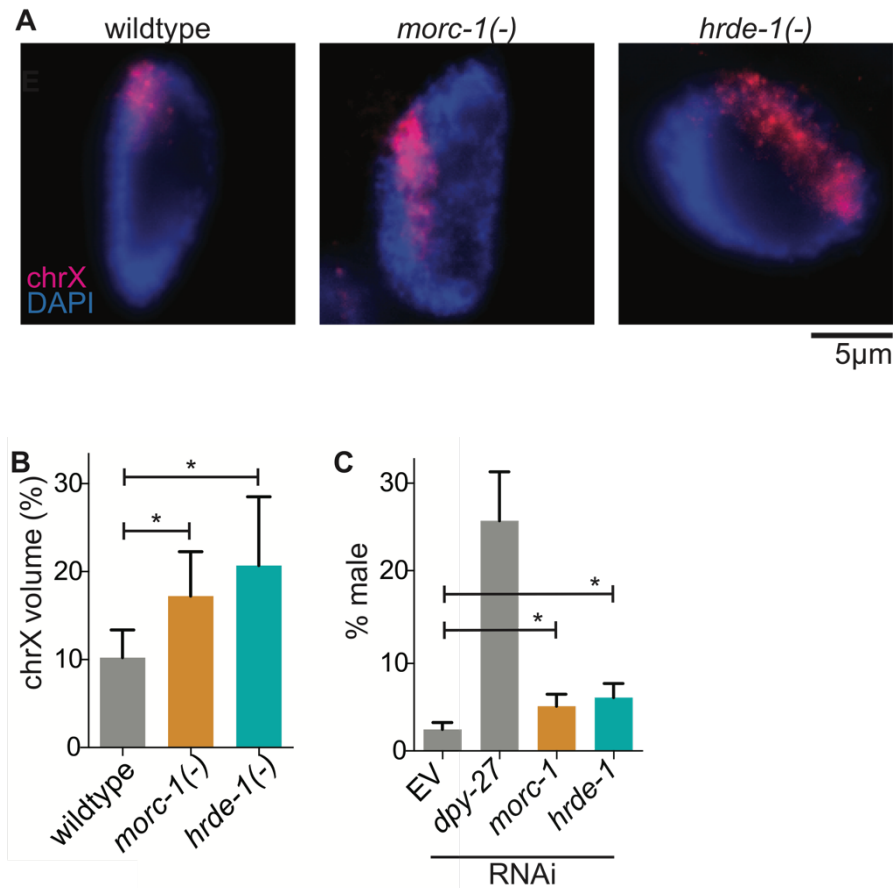
Figure 2.22. ***morc-1*** **is required for X chromosome compaction and dosage compensation. (A)** Defective X chromosome compaction in *morc-1(-)* and *hrde-1(-)* mutants. DNA FISH using a probe targeting the X chromosome and DAPI staining shows increased X chromosome volume in *morc-1(-)* and *hrde-1(-)* intestinal nuclei. **(B)** X chromosome volume is significantly increased in *morc-1(-)* and *hrde-1(-)* intestinal nuclei (*morc-1(-)* vs wildtype p=1.7×10^{-5}, *hrde-1(-)* vs wildtype p=9.6×10^{-6}, student's t-test). **(C)** RNAi against *morc-1* and *hrde-1* partially rescues *xol-1(-);sex-1(-);him-8(-)* male lethality. *xol-1(-);sex-1(-);him-8(-)* hermaphrodites were grown on the indicated RNAi clone. Bars indicate the average percentage of live progeny that are male from five biological replicates. Error bars represent standard deviation. Both *morc-1* and *hrde-1* RNAi cause significant increases in percentage of live male progeny (*morc-1* vs empty vector p=0.0053, *hrde-1* vs empty vector p=0.0018, student's t-test).
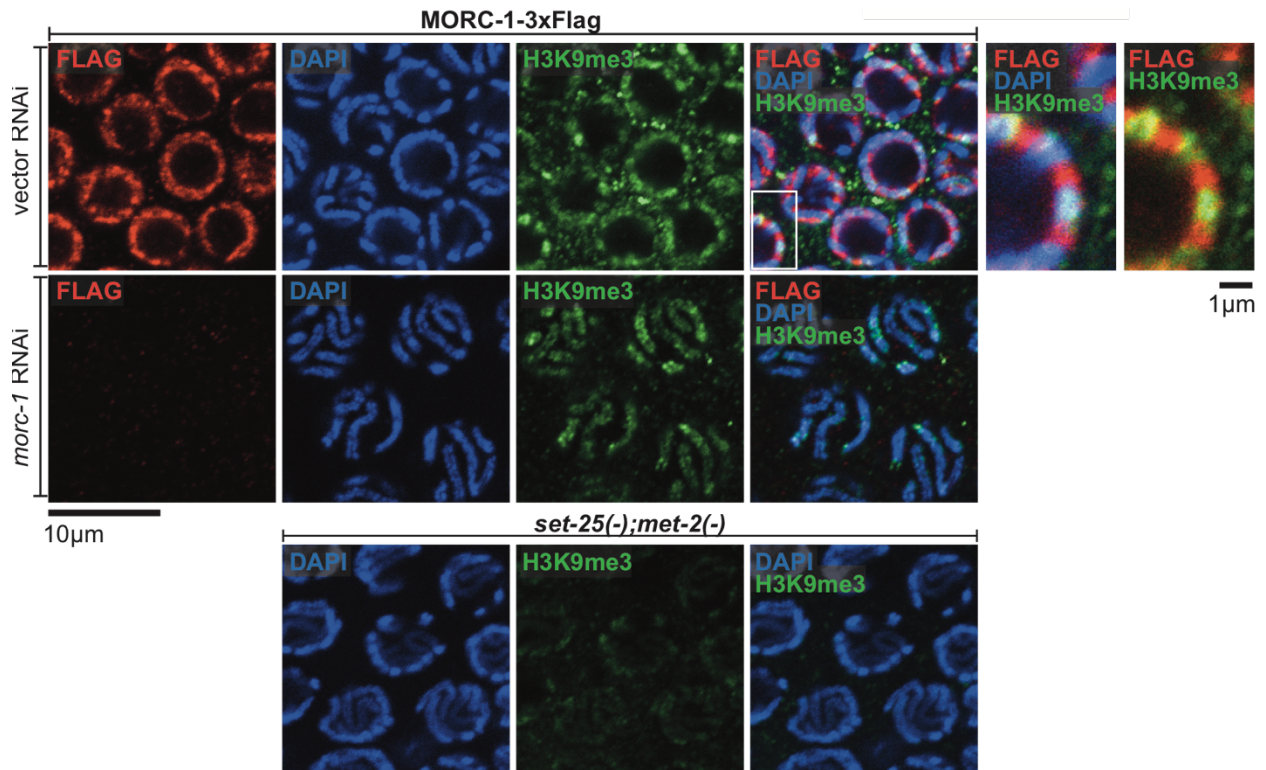
Figure 2.23. MORC-1 is excluded from H3K9me3 in pachytene nuclei.
Immunostaining against H3K9me3 and Flag in a strain expressing MORC-1::3xFlag on empty vector RNAi (top row) shows that MORC-1::3xFlag localizes to the nuclear periphery adjacent to H3K9me3. Right most images show a high-magnification image of the region indicated by the white box with and without the DAPI channel. Anti-Flag staining is specific as indicated by decreased signal when worms are fed RNAi against *morc-1* (middle row). *met-2;set-25(-)* mutants serve as a negative control for H3K9me3 staining (bottom row).
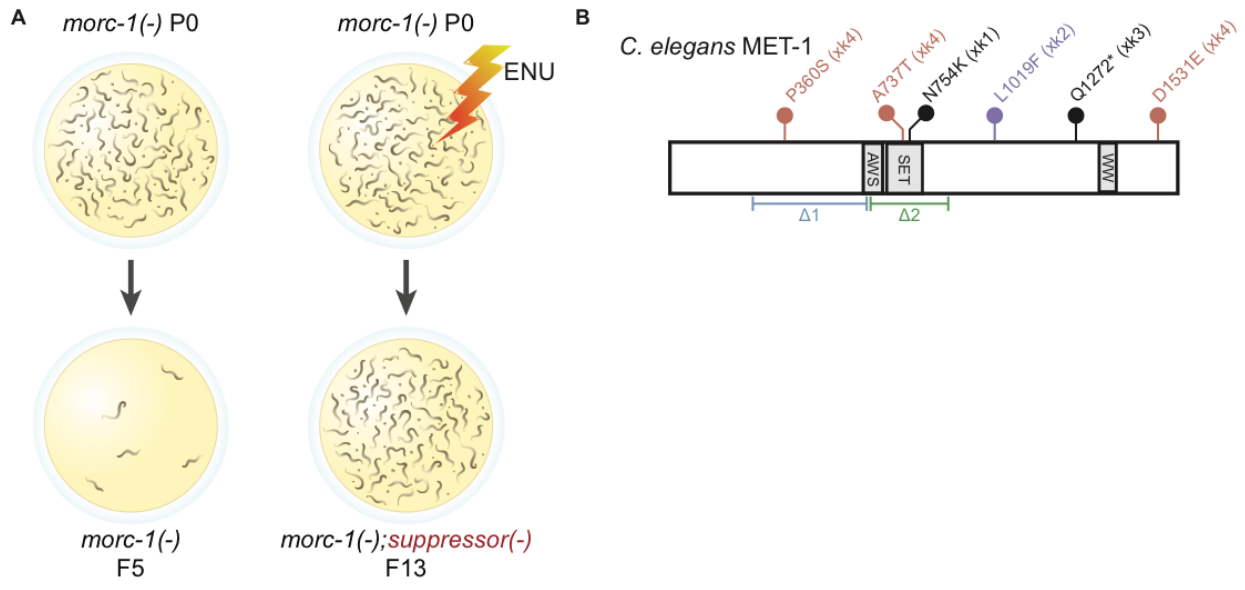
Figure 2.24. A forward genetic screen identifies mutations that suppress *morc-1(-)* germline mortality. **(A)** Schematic of a forward genetic screen for suppressors of *morc-1(-)* germline mortality. **(B)** Schematic of *C. elegans met-1* alleles (isoform b). Flags indicate the alleles generated by our screen *(xk1-4)*. The *xk4* allele contains three missense mutations. Indicated below are the pre-existing deletion alleles, *n4337(Δ1)*, and *ok2172(Δ2)*.

strain	genotype	amino acid change (MET-1 isoform b)
QK89	*met-1(xk1) I*	N754K
QK90	*met-1(xk2) I*	L1019F
QK91	*met-1(xk3) I*	Q1272*
QK92	*met-1(xk4) I*	P360S/A737T/D1531E

Table 2.1. *met-1* alleles generated from this study.
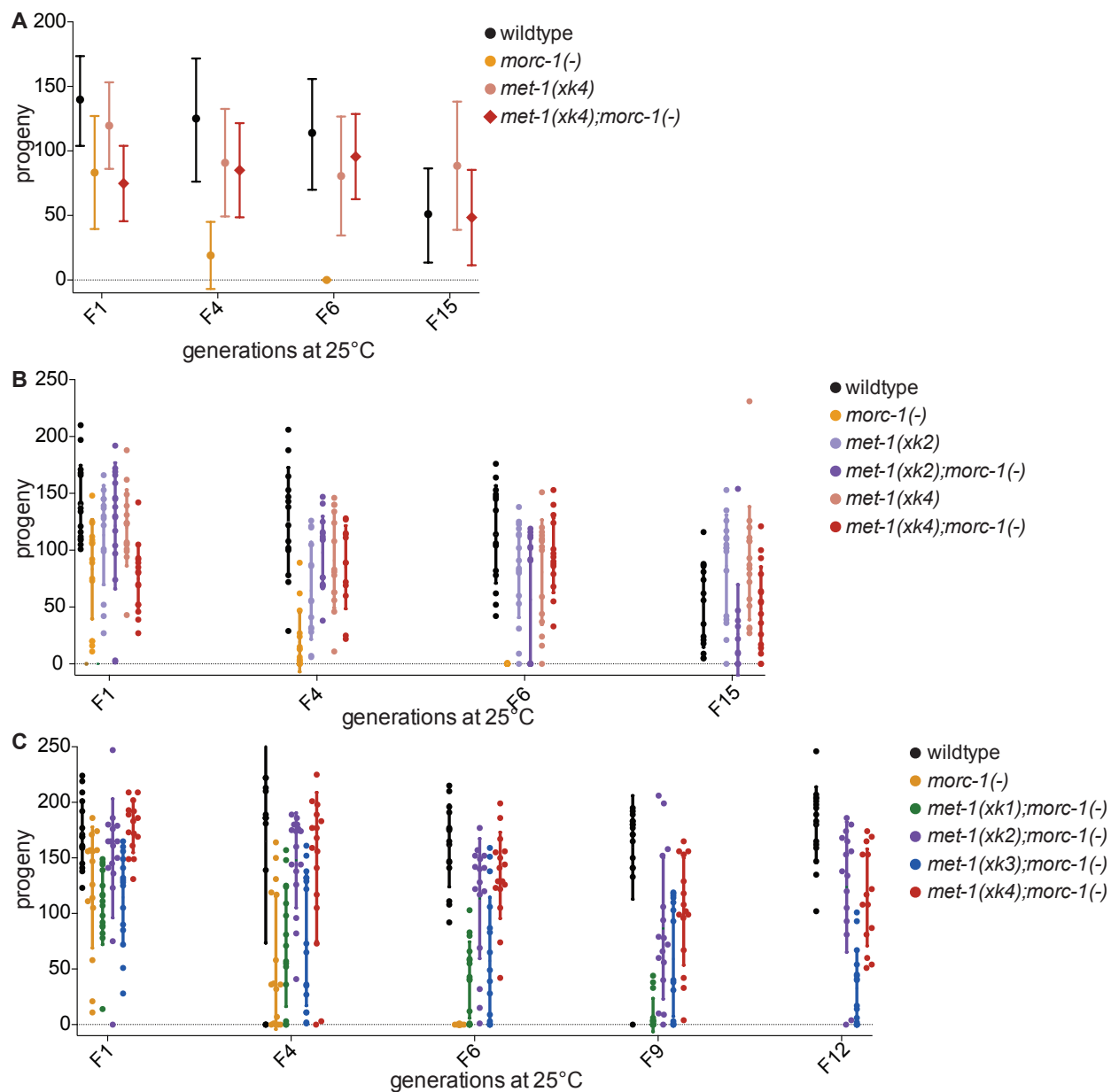
Figure 2.25. Mutations in *met-1* rescue *morc-1(-)* germline mortality. (A) *met-1(xk4)* rescues *morc-1(-)* germline mortality. Mean fertility +/- standard is plotted for each genetic background at each generation. **(B)** *met-1(xk4)* and *met-1(xk2)* suppress *morc-1(-)* mortality. *Met-1(xk4)* and *met-1(xk2)* do not exhibit a fertility defect at 25°C. Circles indicate fertility counts of individual worms. Error bars indicate mean +/- standard deviation. **(C)** All four *met-1* alleles generated from the *morc-1* suppressor screen (*xk1-4*) rescue *morc-1(-)* germline mortality. Circles indicate fertility counts of individual worms. Error bars indicate mean +/- standard deviation.
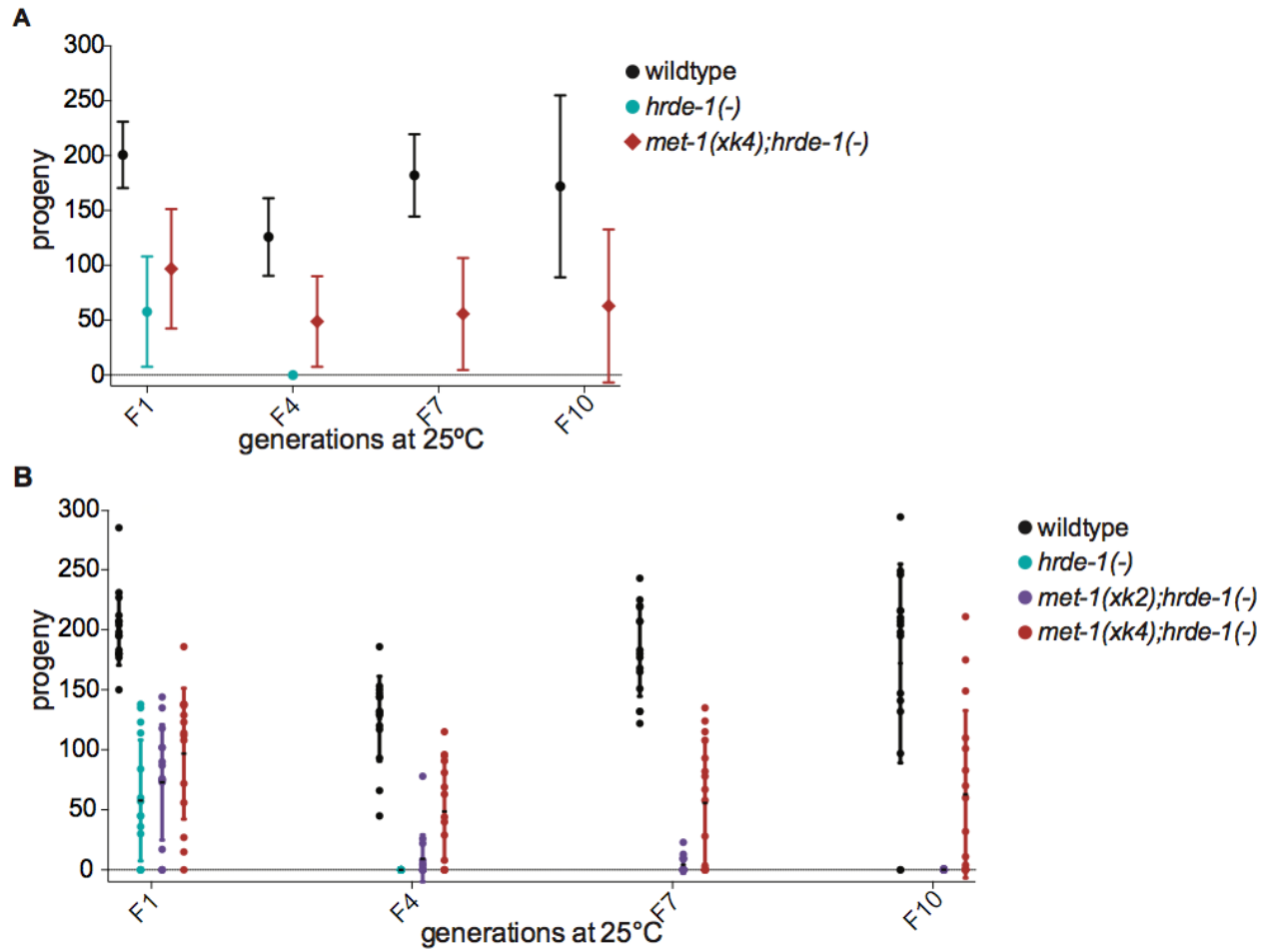
Figure 2.26. Mutations in ***met-1* rescue *hrde-1(-)* germline mortality. (A)** *met-1(xk4)* rescues *hrde-1(-)* germline mortality. Mean fertility +/- standard deviation is plotted for each genetic background at each generation. **(B)** *met-1* alleles *xk2* and *xk4* rescue *hrde-1(-)* germline mortality. Circles indicate fertility counts of individual worms. Error bars indicate mean +/- standard deviation.
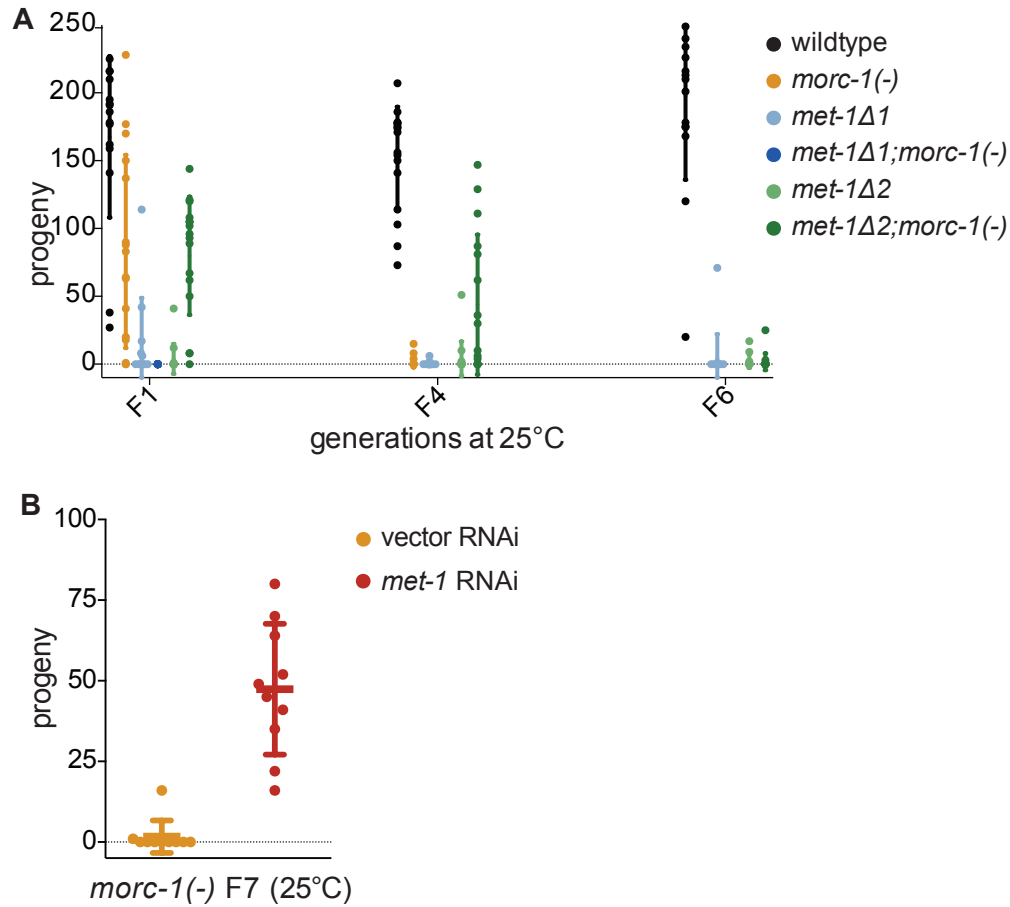
Figure 2.27. Partial loss of *met-1* function is required for *morc-1(-)* suppression.
(A) Deletion alleles of *met-1* cause severely reduced fertility at 25°C and do not rescue *morc-1(-)* germline mortality. Circles indicate fertility counts of individual worms. Error bars indicate mean +/- standard deviation. **(B)** RNAi against *met-1* rescues the *morc-1(-)* fertility defect at F7.

A

H3K36me3 ChIP-seq

● *morc-1(-)*-dependent H3K36me3, F4
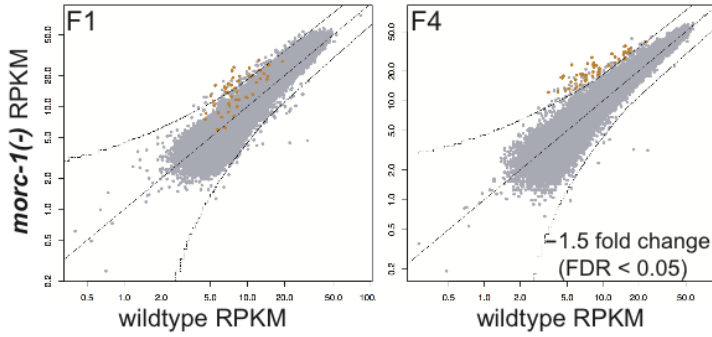
● *hrde-1(-)*-dependent H3K36me3, F4
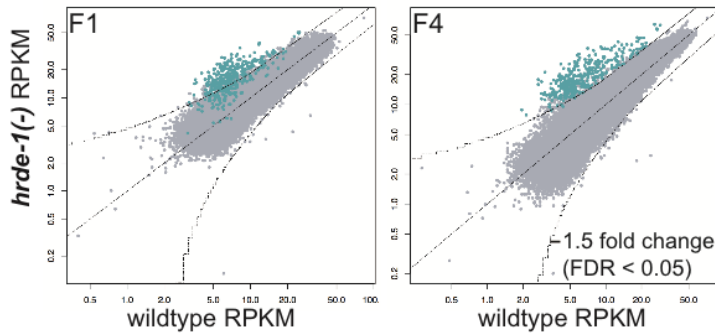
B

morc-1(-) H3K36me3
(replicate 1: 104 targets)

morc-1(-) H3K36me3
(replicate 2: 90 targets)

55   49   41
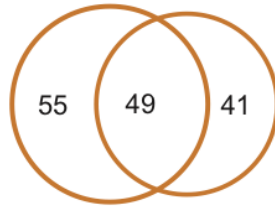
hrde-1(-) H3K36me3
(replicate 1: 525 targets)

hrde-1(-) H3K36me3
(replicate 2: 480 targets)

141   384   96
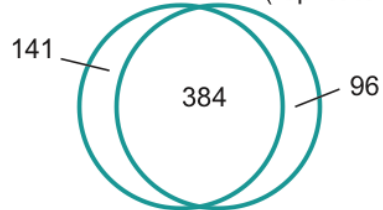
C

shared H3K36me3 targets

hrde-1(-) F4 targets

morc-1(-) F4 targets

350   34   15

95
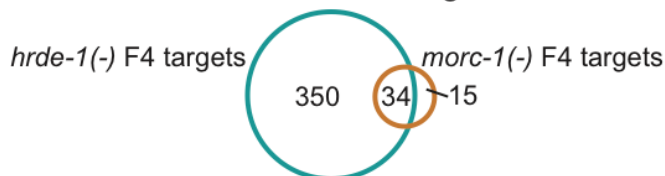
Figure 2.28. Loss of *morc-1* and *hrde-1* leads to enrichment of H3K36me3. (A) RPKM of 1kb windows in wildtype (x-axis) vs mutant (y-axis) H3K36me3 ChIP libraries in F1 (left) and F4 (right) generations. Indicated in yellow and blue are 1kb regions that are more than 1.5-fold-enriched in F4 *morc-1(-)* and *hrde-1(-)* respectively with FDR of <0.05 in two biological replicates. **(B)** Venn diagrams showing overlap between replicates of H3K36me3-enriched regions in F4 mutant compared to wildtype (top: *morc-1(-)* vs wildtype, bottom: *hrde-1(-)* vs wildtype). Subsequent analyses were performing using targets that were identified in both biological replicates. **(C)** Overlap of *morc-1(-)* and *hrde-1(-)* H3K36me3-enriched regions shows that MORC-1 regulates a subset of HRDE-1 targets.
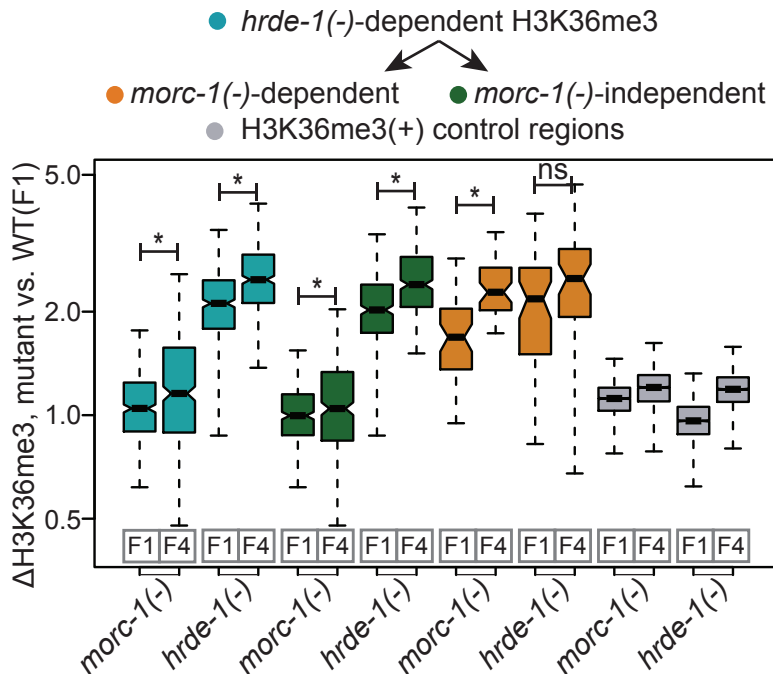
Figure 2.29. H3K36me3 enrichment in *morc-1(-)* and *hrde-1(-)* mutants is progressive. Box plot representing the ratio of recovered reads of the specified target group in H3K36me3 IP in the indicated mutant at the indicated generation compared to reads in F1 wildtype. *hrde-1(-)*-dependent regions (blue) show significant increase in H3K36me3 from F1 to F4 in both *morc-1(-)* (p=3.07x10^{-7}, Welch's t-test) and *hrde-1(-)* (p<2.2x10^{-16}, Welch's t-test) mutants. *hrde-1(-)*-dependent, *morc-1(-)*-independent regions (green) show progressive gain of H3K36me3 from F1 to F4 in *hrde-1(-)* mutants (p<2.2x10^{-16}, Welch's t-test) and *morc-1(-)* mutants (p=0.0012, Welsh's t-test). *hrde-1-*dependent, *morc-1*-dependent regions (yellow) are significantly more enriched in *morc-1(-)* F4 compared to *morc-1(-)* F1 (p=6.44x10^{-11}) but not significantly different in *hrde-1(-)* F1 and F4 (p=0.011). Gray boxes show H3K36me3 levels at the regions that are most H3K36me3-enriched in wildtype.
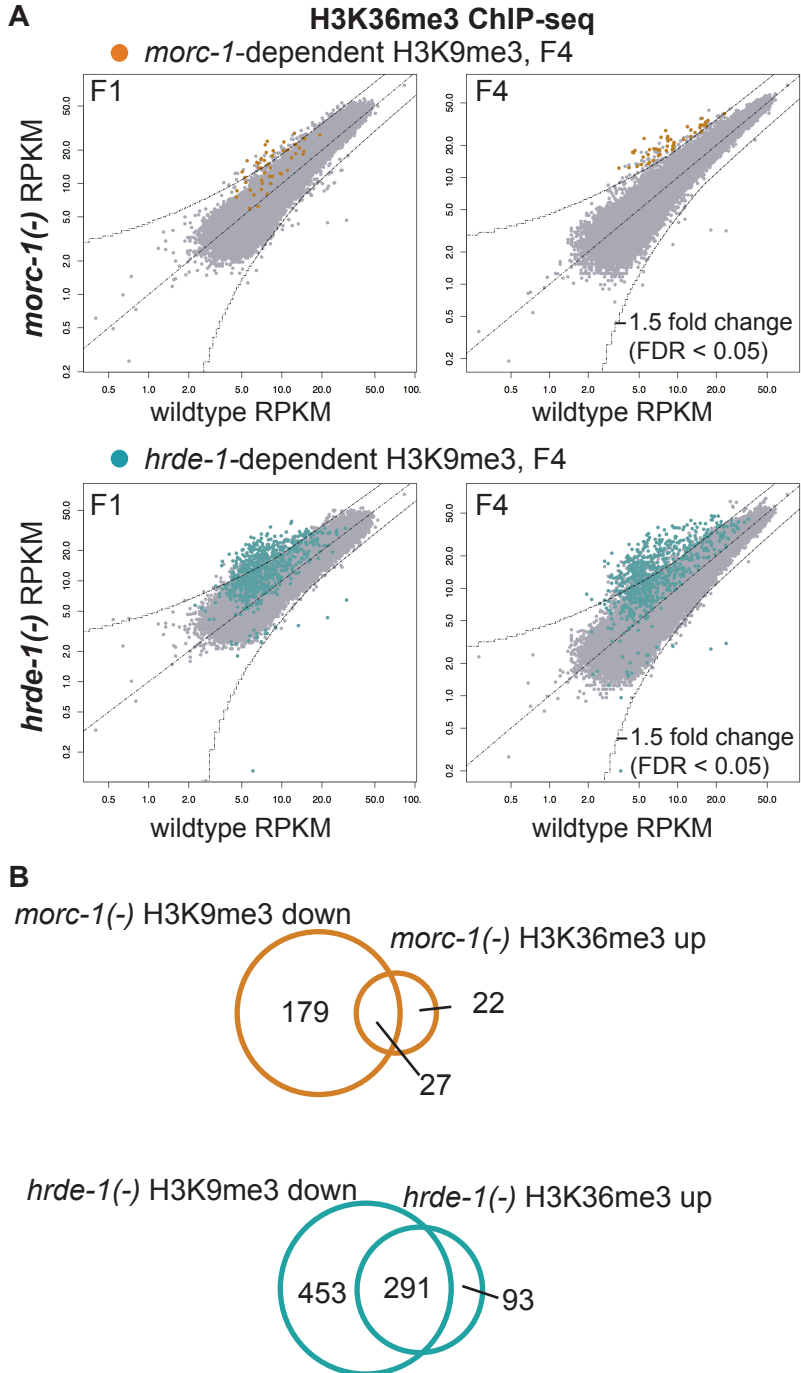
Figure 2.30. H3K9me3-depleted regions are H3K36me3-enriched in *hrde-1(-)* and *morc-1(-)* mutants. (A) Scatter plots showing H3K36me3 levels in F1 and F4 *morc-1(-)* (top) and *hrde-1(-)* (bottom) versus wildtype. Highlighted in yellow are regions that are depleted of H3K9me3 in F4 *morc-1(-)* compared to wildtype in two biological replicates. Highlighted in blue are regions that are depleted of H3K9me3 in F4 *hrde-1(-)* compared to wildtype in two biological replicates. **(B)** Venn diagrams showing overlap between targets that are depleted of H3K9me3 and targets that are enriched for H3K36me3 in late generation *morc-1(-)* (top) and *hrde-1(-)* (bottom) compared to wildtype.
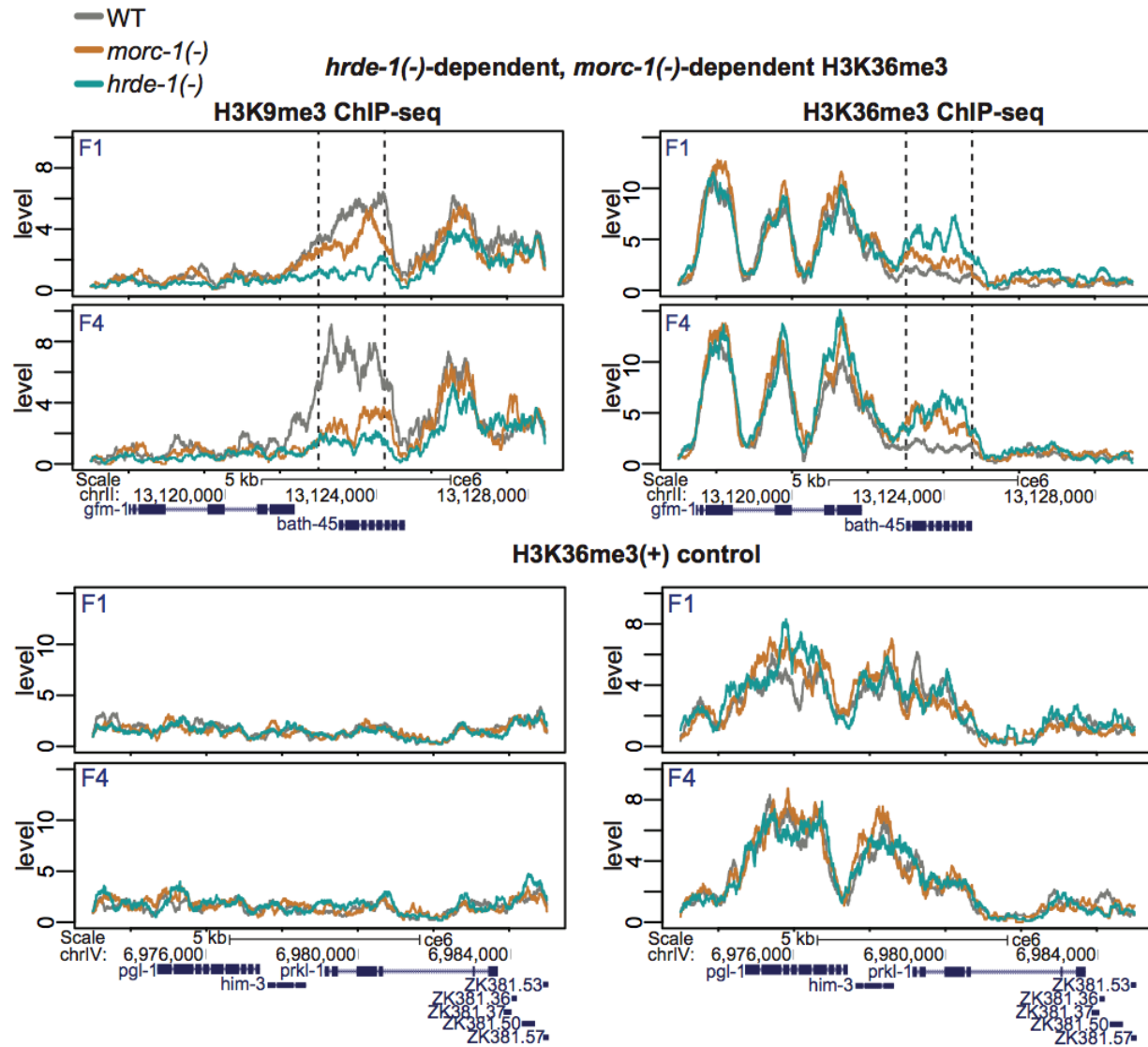
Figure 2.31. At endo-siRNA target *bath-45*, H3K9me3 depletion is concomitant with H3K36me3 enrichment. H3K9me3 (left) and H3K36me3 (right) levels in the indicated genetic background at F1 and F4 at *morc-1*-dependent region containing *bath-45* (top, *bath-45* indicated by dotted lines) and an unregulated control region (bottom).

99
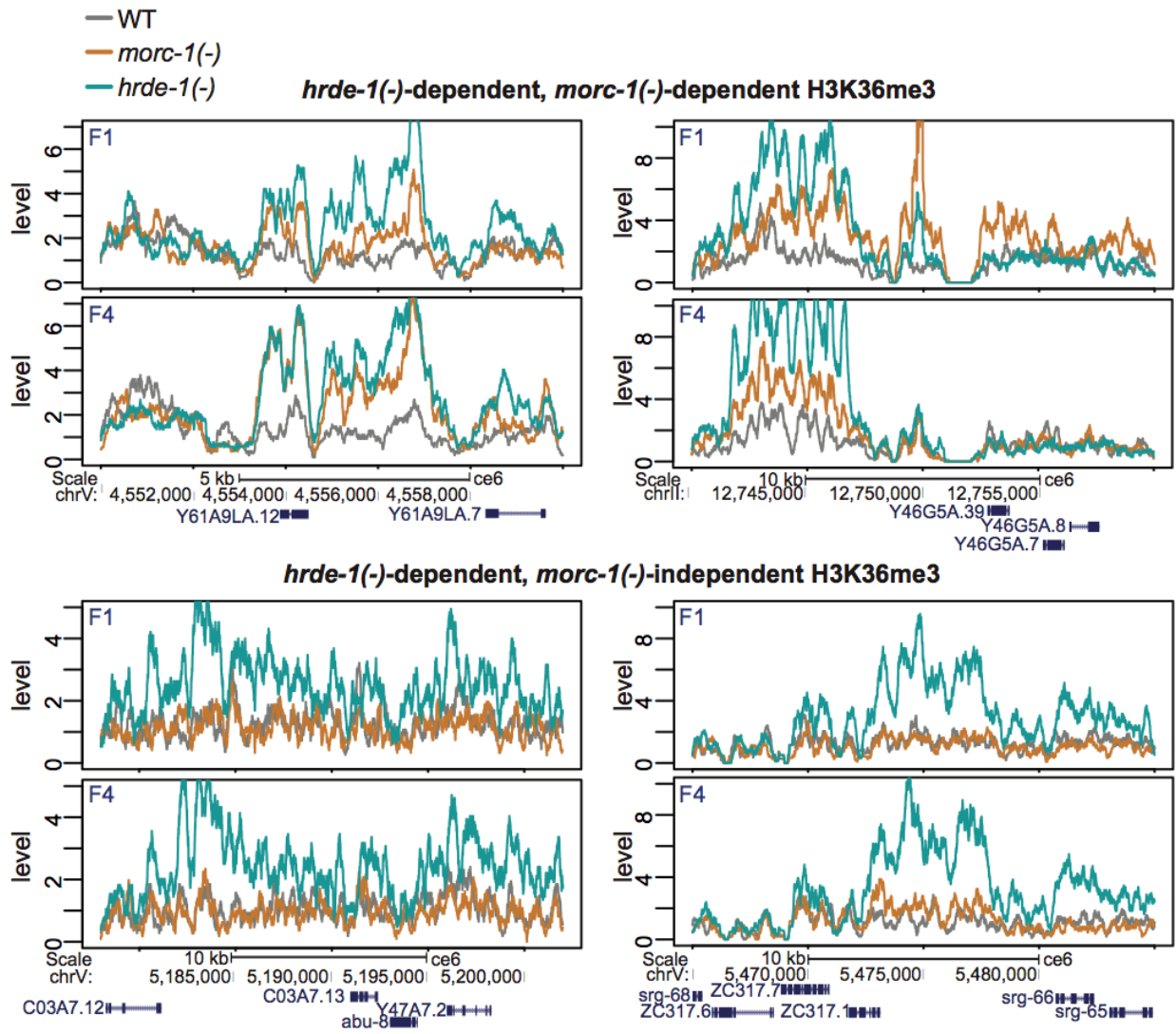
Figure 2.32. Exemplary *morc-1(-)*-dependent and *morc-1(-)*-independent H3K36me3 regions. H3K36me3 levels at exemplary *morc-1(-)*-dependent H3K36me3 targets at early and late generation (top). H3K36me3 levels at exemplary *hrde-1(-)*-dependent, *morc-1(-)*-independent H3K36me3 targets at early and late generation (bottom).

100
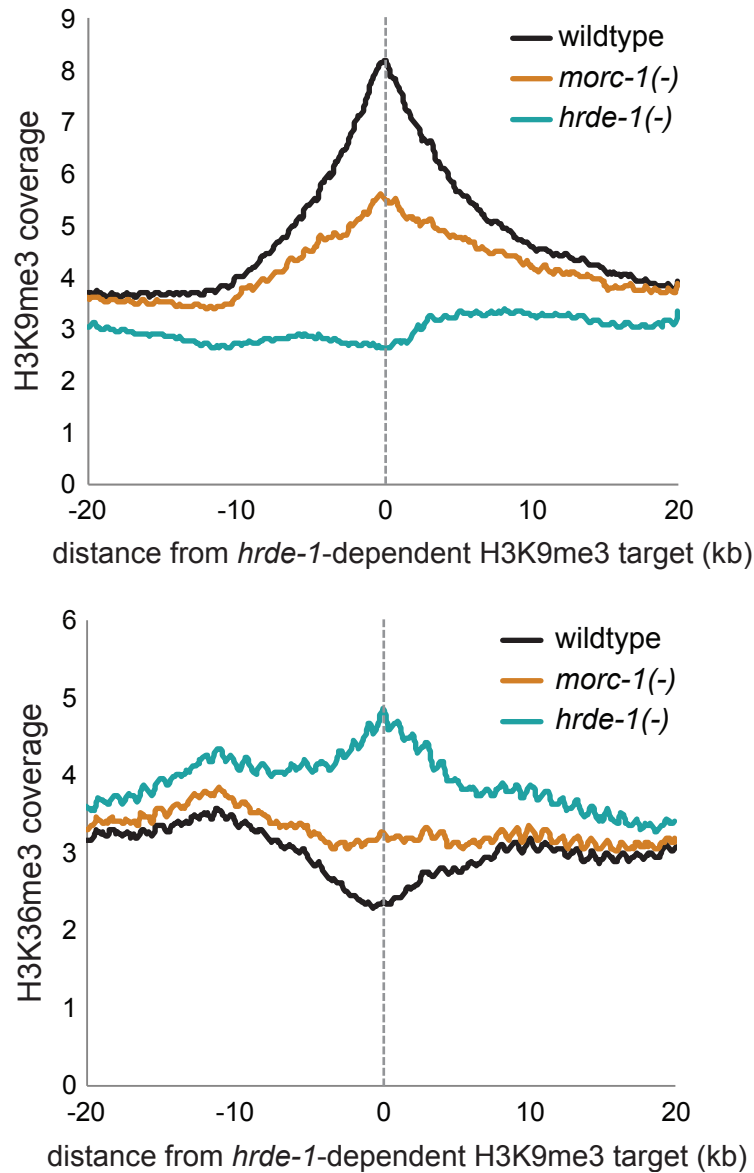
Figure 2.33. *hrde-1*-dependent H3K9me3 regions occupy local maxima of H3K9me3. H3K9me3 (top) and H3K36me3 (bottom) levels 20kb upstream and downstream of *hrde-1*-dependent H3K9me3 targets in wildtype, *morc-1(-)*, and *hrde-1(-)* worms at F4 generation.
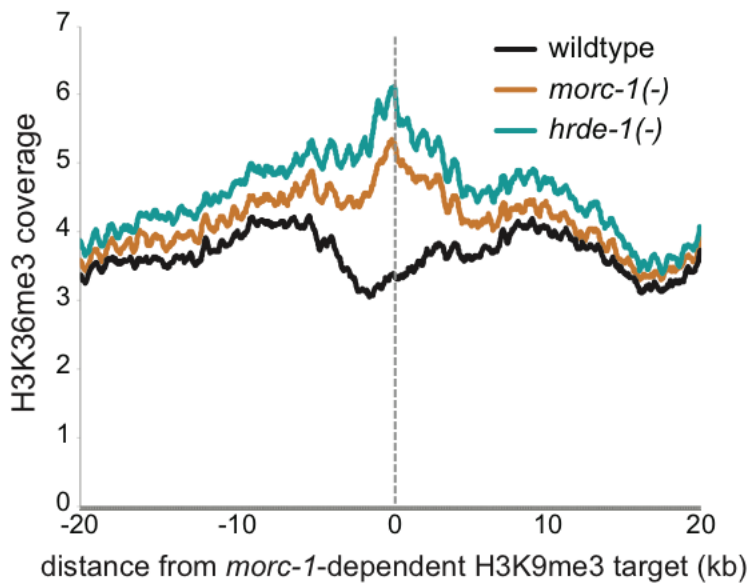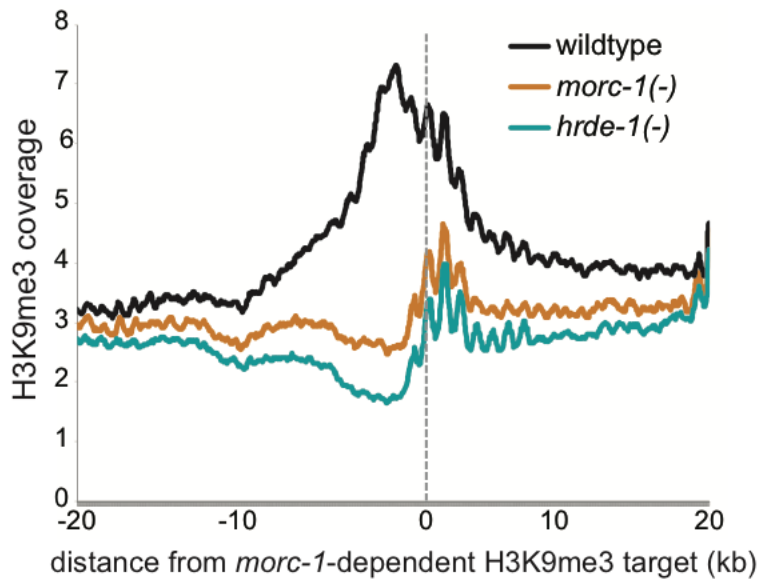
Figure 2.34. *morc-1*-dependent H3K9me3 regions occupy local maxima of H3K9me3. H3K9me3 (top) and H3K36me3 (bottom) levels 20kb upstream and downstream of *morc-1*-dependent H3K9me3 targets in wildtype, *morc-1(-)*, and *hrde-1(-)* worms at F4 generation.
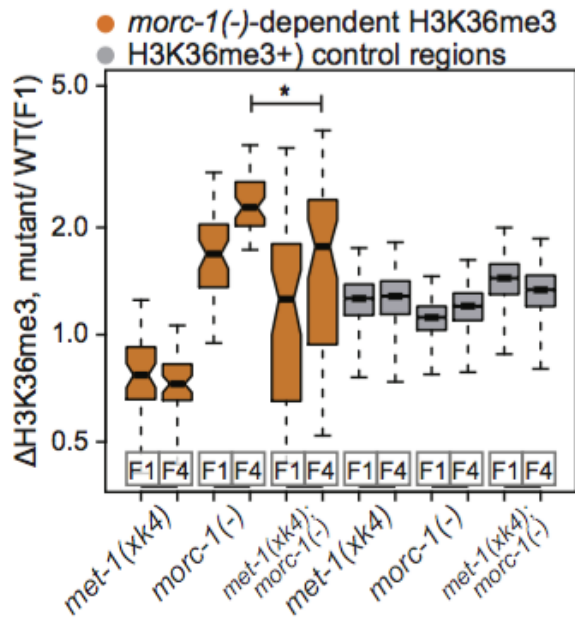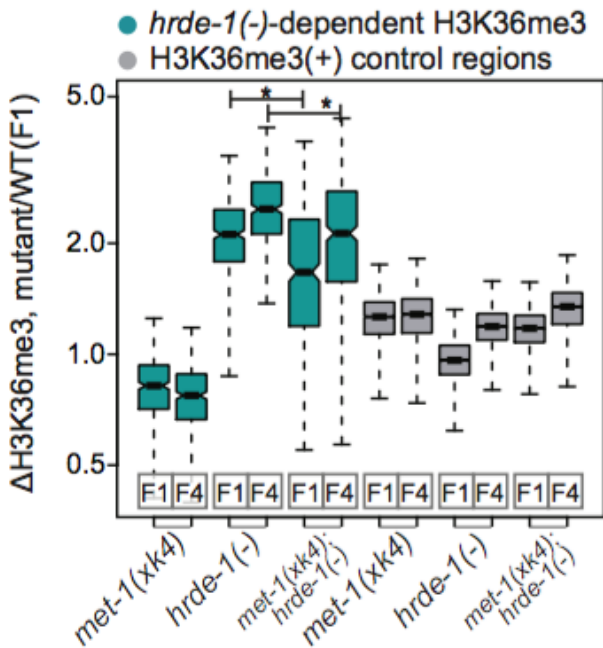
Figure 2.35. H3K36me3 gain in *morc-1(-)* and *hrde-1(-)* mutants is *met-1-dependent*. (A) *morc-1(-)* dependent H3K36me3 regions are significantly depleted of H3K36me3 in *met-1(xk4);morc-1(-)* compared to *morc-1(-)* at F4 generation (p=2.50x10^{-5}, Welch's t-test). Gray boxes show H3K36me3 levels at the most H3K36me3 regions in wildtype. **(B)** H3K36me3 levels at *hrde-1(-)*-dependent H3K36me3 targets in the indicated mutant background and generation relative to wildtype F1. These regions are significantly depleted in *met-1(xk4);hrde-1(-)* compared to *hrde-1(-)* at both generations (F1 p=3.28x10^{-14}, F4 p=1.15x10^{-10}, Welch's t-test). Gray boxes show H3K36me3 levels at the most H3K36me3 regions in wildtype.
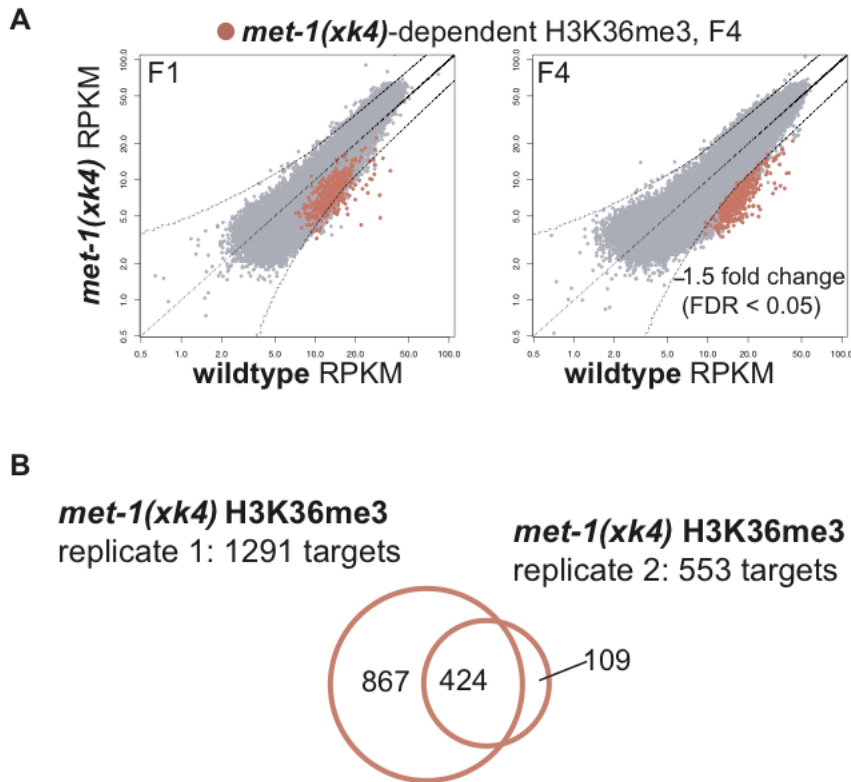
Figure 2.36 *met-1* is required for transgenerational H3K36me3 maintenance. (A) Scatter plots showing H3K36me3 levels in *met-1(xk4)* compared to wildtype at early and late generation. Highlighted in pink are targets that are at least 1.5-fold depleted in F4 *met-1(xk4)* compared to wildtype with FDR<0.05 in two biological replicates. **(B)** Overlap of F4 targets that are at least 1.5-fold depleted of H3K36me3 in two biological replicates of *met-1(xk4)* vs. wildtype.
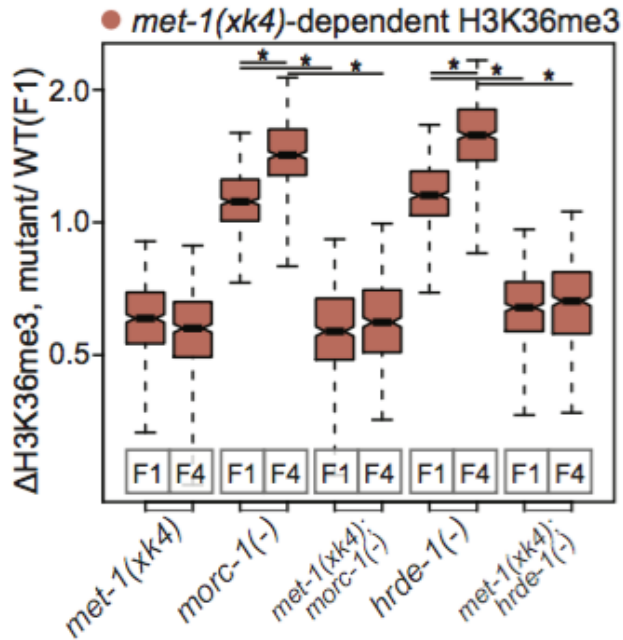
Figure 2.37 MET-1 activity is increased in *morc-1(-)* and *hrde-1(-)* mutants. At *met-1(xk4)*-dependent regions, H3K36me3 levels are elevated in *morc-1(-)* and *hrde-1(-)* mutants (*morc-1(-)* F1 vs. F4: $p<2.2\times10^{-16}$, *hrde-1(-)* F1 vs F4: $p<2.2\times10^{-16}$, Welch's t-test). H3K36me3 gain at these sites is *met-1(xk4)*-dependent.

A

met-1(xk4);morc-1(-)-dependent H3K36me3, F4

met-1(xk4);hrde-1(-)-dependent H3K36me3, F4
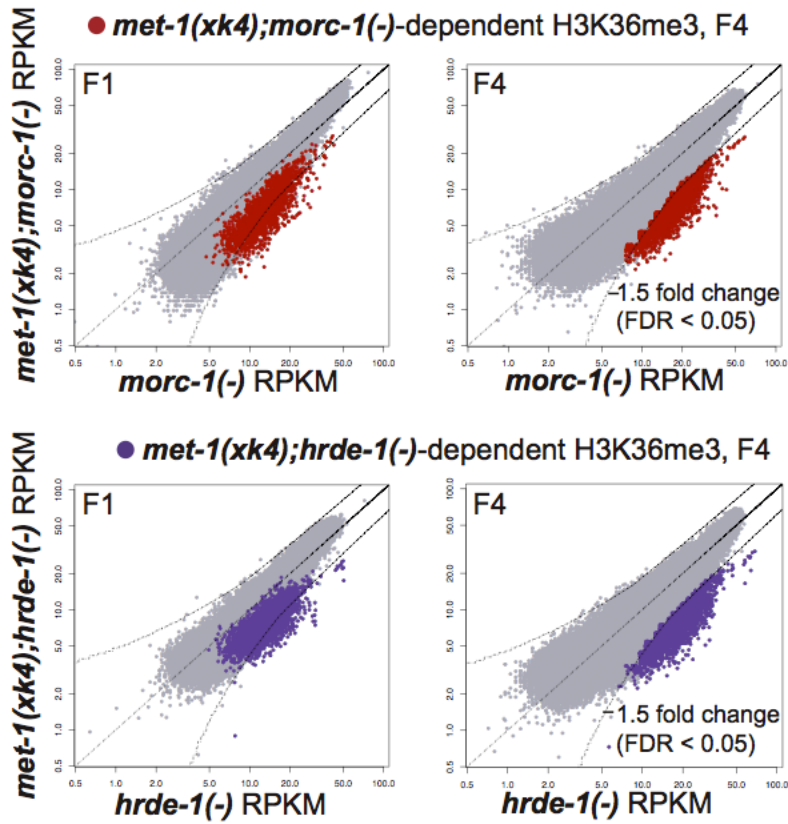
B

met-1(xk4);morc-1(-) H3K36me3
replicate 1: 2249 targets

307   1942   1344
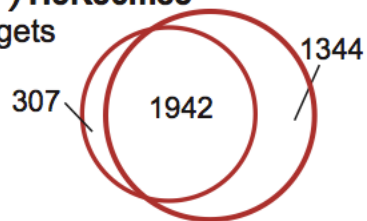
met-1(xk4);morc-1(-) H3K36me3
replicate 2: 3286 targets

met-1(xk4);hrde-1(-) H3K36me3
replicate 1: 2912 targets

463   2449   1366
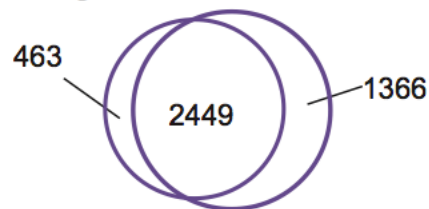
met-1(xk4);hrde-1(-) H3K36me3
replicate 2: 3815 targets

Figure 2.38. Identification of *met-1*-dependent loci in *morc-1(-)* and *hrde-1(-)* mutants. (A) Scatter plots showing H3K36me3 levels in *met-1(xk4);morc-1(-)* compared to *morc-1(-)* at F1 and F4 generation (top) and in *met-1(xk4);hrde-1(-)* compared to *hrde-1(-)* (bottom). Regions that are more than 1.5-fold depleted of H3K36me3 with FDR<0.05 in the *met-1(xk4)* background at F4 are highlighted in red (*met-1(xk4);morc-1(-)*-dependent) and purple (*met-1(xk4);hrde-1(-)*-dependent). **(B)** Overlap of F4 targets that are at least 1.5-fold depleted of H3K36me3 in two biological replicates of *met-1(xk4);morc-1(-)* vs. *morc-1(-)* (top) and *met-1(xk4);hrde-1(-)* vs *hrde-1(-)* (bottom).
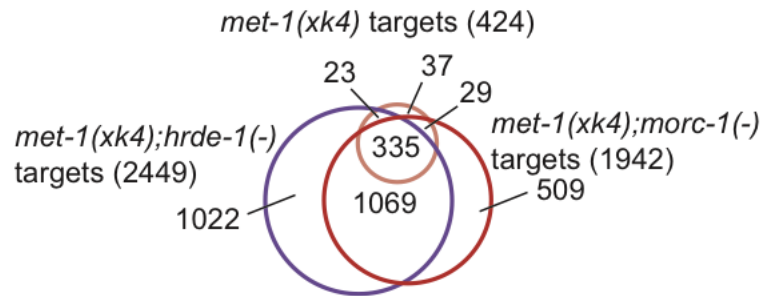
Figure 2.39. MORC-1 and HRDE-1 restrict MET-1 target selection. Venn diagram of 1kb regions that are H3K36me3-depleted in *met-1(xk4)* compared to wildtype, *met-1(xk4);morc-1(-)* compared to *morc-1(-)*, and *met-1(xk4);hrde-1(-)* compared to *hrde-1(-).*
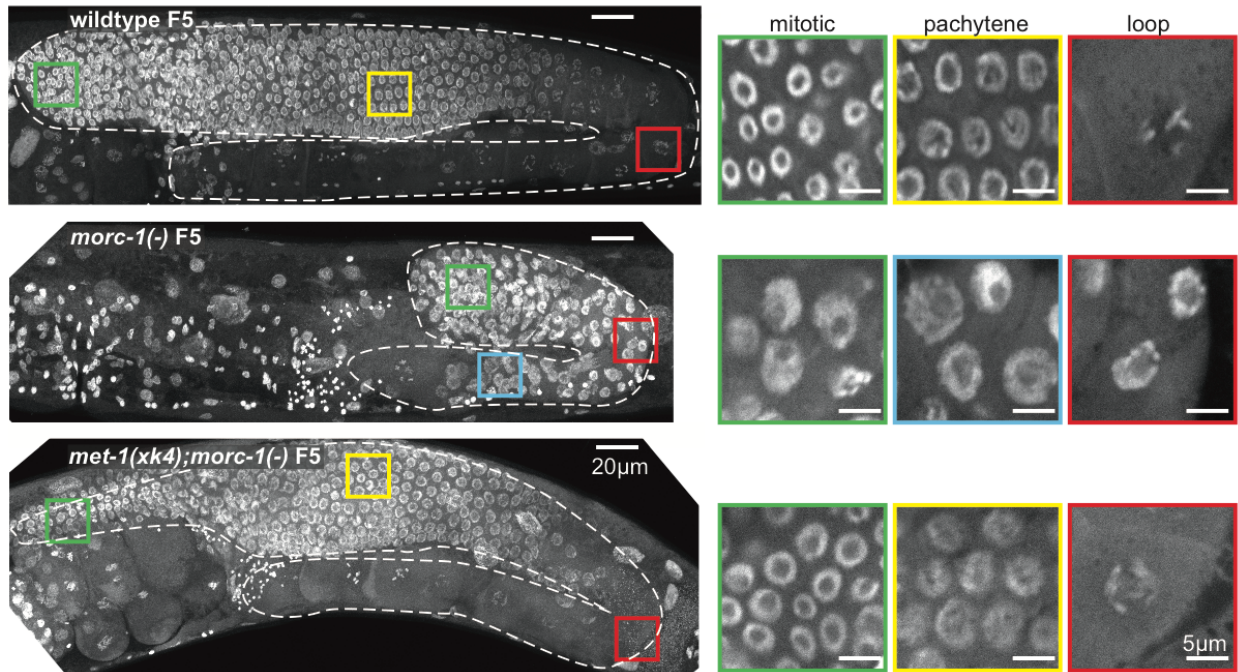
Figure 2.40. *met-1(xk4)* rescues *morc-1(-)* germline chromatin disorganization. Whole mount DAPI staining of wildtype, *morc-1(-)*, and *met-1(xk4);morc-1(-)* at F5 (25°C) shows enlarged nuclei in distal germline and failure to execute meiotic prophrase in sterile *morc-1(-)* mutants. These defects are rescued by *met-1(xk4).*

DISCUSSION

Here we demonstrated that *morc-1* functions as a critical link between endo-siRNAs and heritable chromatin modifications to ensure germline integrity. We found that *morc-1* is essential for H3K9me3 maintenance at a subset of endogenous nuclear RNAi targets. In addition, we reveal that germline nuclear RNAi and *morc-1* are required for proper heterochromatin localization, chromatin compaction, and suppression of H3K36 methylation. These defects likely contribute to the germline mortality phenotype of *morc-1(-)* and *hrde-1(-)* mutants. In support of this, we identified mutations in the gene encoding the H3K36me3 histone methyltransferase MET-1 as potent suppressors of *morc-1(-)* and *hrde-1(-)* germline mortality. We propose a model in which MORC-1 facilitates the maintenance of H3K9me3 at endo-siRNA targets to control chromatin architecture. Without the protection afforded by MORC-1, heterochromatin is susceptible to loss of H3K9 methylation, decompaction, mislocalization, and encroachment of neighboring H3K36me3 (Figure 2.41). Our study highlights an essential role for nuclear RNAi and MORC-1 in transgenerational control of chromatin structure. Furthermore, our finding that loss of H3K36me3 via depletion of MET-1 is sufficient to restore germline immortality in *morc-1(-)* and *hrde-1(-)* mutants emphasizes the critical role of chromatin architecture in multigenerational germline maintenance (Figure 2.42).

Our ChIP-seq studies indicate that MORC-1 regulates H3K9me3 maintenance almost exclusively at a subset of *hrde-1*-targets. Why do some targets require MORC-1 for H3K9me3 maintenance while others do not? Our data indicate that *hrde-1(-)* mutants are more severely depleted of H3K9me3 at *morc-1*-dependent targets than at *morc-1*-independent targets. Perhaps *morc-1*-dependent targets are more prone to encroachment of euchromatin and thus require additional machinery to maintain the heterochromatic state. Based on our anchored analysis showing that *morc-1*-dependent H3K9me3 targets reside in local maxima of H3K9me3 and local minima of H3K36me3, we speculate that proximity to euchromatin may contribute to susceptibility to loss of heterochromatin marks, thus necessitating a protective factor to preserve these boundaries. A role in maintenance of heterochromatin-euchromatin boundaries would

be consistent with the striking MORC-1 localization pattern we observe in pachytene nuclei where MORC-1 appears to be excluded from H3K9me3-marked regions. MORC-1 may function as a scaffold around compact chromatin to maintain its compaction and localization.

In support of a role for MORC-1 in regulation of heterochromatin-euchromatin boundaries, we observe a progressive increase in H3K36me3 in *morc-1(-)* mutants compared to wildtype. While the number of sites that meet our 1.5-fold threshold to be designated as enriched for H3K36me3 in *morc-1(-)* mutants is small (49), we posit that this is reflective of a more pervasive trend for several reasons. First, we observe a dramatic increase in the number of *met-1(xk4)*-dependent loci in the *morc-1(-)* background compared to in the wildtype background. This would be consistent with these loci exhibiting a small amount of *met-1*-dependent H3K36me3 in wildtype, gain of H3K36me3 in *morc-1(-)* mutants, and depletion to below wildtype levels in *met-1(xk4);morc-1(-)* mutants (Figure S7C). In such cases, it is possible that only the depletion in *met-1(xk4);morc-1(-)* would be large enough to meet our 1.5-fold threshold. Second, our sample collection method for ChIP-seq studies selected against mutants with the strongest enrichment of H3K36me3. Due to the large populations of worms required to perform robust ChIP, it was necessary to isolate embryos at each generation through hypochlorite treatment of gravid worms. As the penetrance of germline mortality is quite variable in intermediate generations, this strategy enriched our samples for the most fertile worms. The potent suppression of *morc-1(-)* germline mortality by four *met-1* alleles highlights the tremendous impact that H3K36me3 imparts on the fertility of *morc-1(-)* mutants. For experiments that did not require such large late generation samples, such as germline mortality assays and DAPI staining for germline morphology, we propagated worms by picking progeny to new plates before they reached adulthood. At these larval stages we could not distinguish fertile and sterile worms and therefore did not introduce any selective pressure to enrich our mutant populations for the most fertile worms.

We have identified not only a downstream effector of the nuclear RNAi pathway but also the first suppressor of a fertility defect of nuclear RNAi mutants, *met-1.* Our findings indicate that germline mortality may be actively regulated, not merely a

stochastic decline in germline function. Our study suggests that *met-1* is an important regulator of the chromatin landscape in *C. elegans* even when the nuclear RNAi pathway is intact, as we observe a dramatic increase in the number of loci that are H3K36me3-depleted in *met-1(xk4)* compared to wildtype from F1 to F4 generation. This suggests that MET-1 contributes to transgenerational maintenance of H3K36me3, a role that previously has been ascribed to the sole other H3K36 histone methyltransferase, MES-4 (Furuhashi et al., 2010; Rechtsteiner et al., 2010). Although *mes*-4 did not emerge from our screen as a candidate *morc-1(-)* suppressor, it may also contribute to the increase in H3K36me3 in *morc-1(-)* and *hrde-1(-)* mutants. Because loss of *mes-4* causes maternal-effect sterility, it is unlikely that our screen, which was based on rescue of germline mortality, would have identified *mes-4*. Due to the progressive nature of the *morc-1(-)* mutant phenotypes and the complete penetrance of the maternal-effect-sterility resulting from mutation or RNAi knockdown of *mes-4*, we are currently unable to address whether *mes-4* antagonizes *morc-1*. MES-4 and two H3K27 methyltransferases, MES-3 and MES-6, were recently implicated in nuclear RNAi-dependent H3K27 trimethylation (Mao et al., 2015).

Perhaps the most striking defect we observe in *morc-1(-)* mutants is the extensive disorganization of germline chromatin in the late generation. Our data suggest that the severely decondensed nuclei in the distal germline are not able to complete meiotic prophase. In fact, the entire germline is so abnormal that we cannot accurately distinguish the mitotic and meiotic zones in the distal germline. These defects are likely to be major contributors to germline mortality in *morc-1(-)* mutants. Although we cannot distinguish whether chromatin decompaction is the cause or the consequence of euchromatic encroachment in *morc-1(-)* mutants, the restoration of germline architecture in late generation *met-1(xk4);morc-1(-)* mutants suggests that at least some degree of decompaction is driven by *met-1*-mediated H3K36me3 deposition. The dramatic increase in the number of loci that are regulated by *met-1* in *morc-1(-)* and *hrde-1(-)* mutants as well as the increased H3K36me3 at targets that are normally regulated by *met-1* suggests that nuclear RNAi and *morc-1* repress *met-1* activity at a large number of sites. In fact, our data suggest that *morc-1* represses *met-1* at a much

larger number of targets than it regulates via H3K9me3 maintenance, highlighting a new role for nuclear RNAi in regulation of euchromatic marks.

Finally, our work suggests that endo-siRNAs are required not only for silencing discrete genetic loci, but also to maintain chromatin organization. We have identified MORC-1 as a major effector of this function and MET-1 as a major antagonist. We expect that future studies will uncover additional factors connecting endo-siRNAs to global regulation of chromatin architecture.
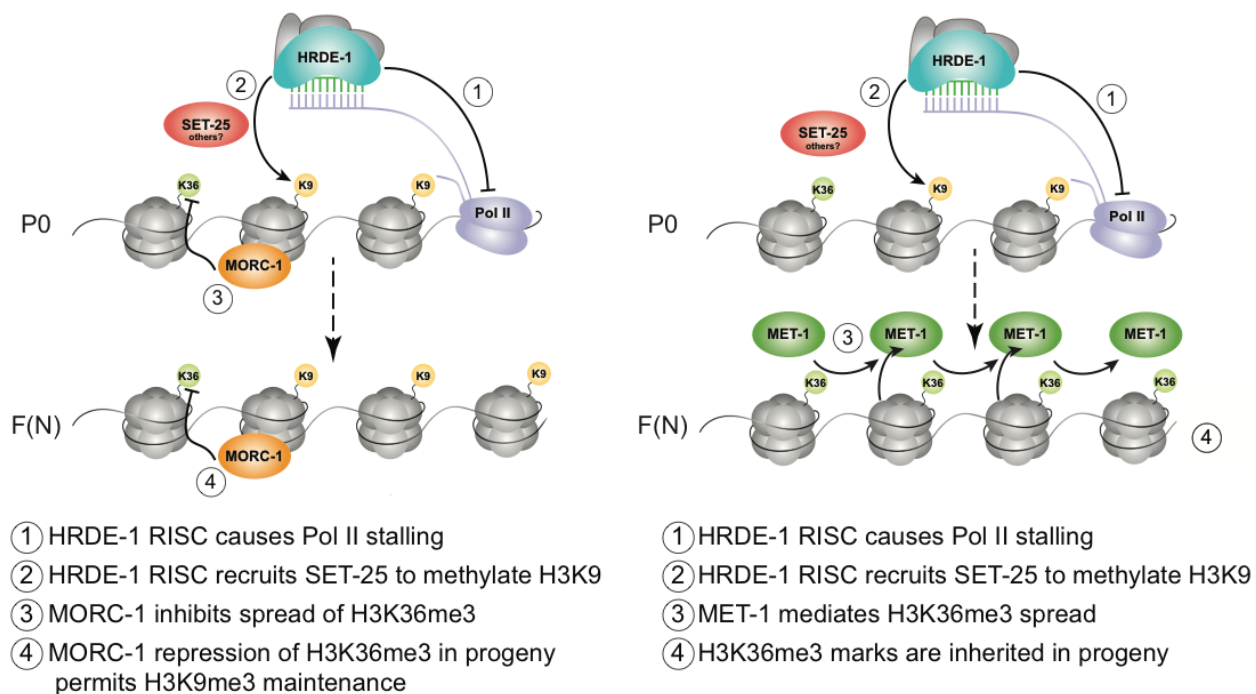
Figure 2.41. Model for regulation of transgenerational chromatin marks at *hrde-1-*target loci by MORC-1. MORC-1 protects H3K9-methylated HRDE-1 targets from encroachment of euchromatin. In the absence of MORC-1, H3K9me3 marks are deposited at these targets but cannot be maintained in the absence of repression of MET-1 by MORC-1. MET-1 mediates H3K36 methylation at these targets.
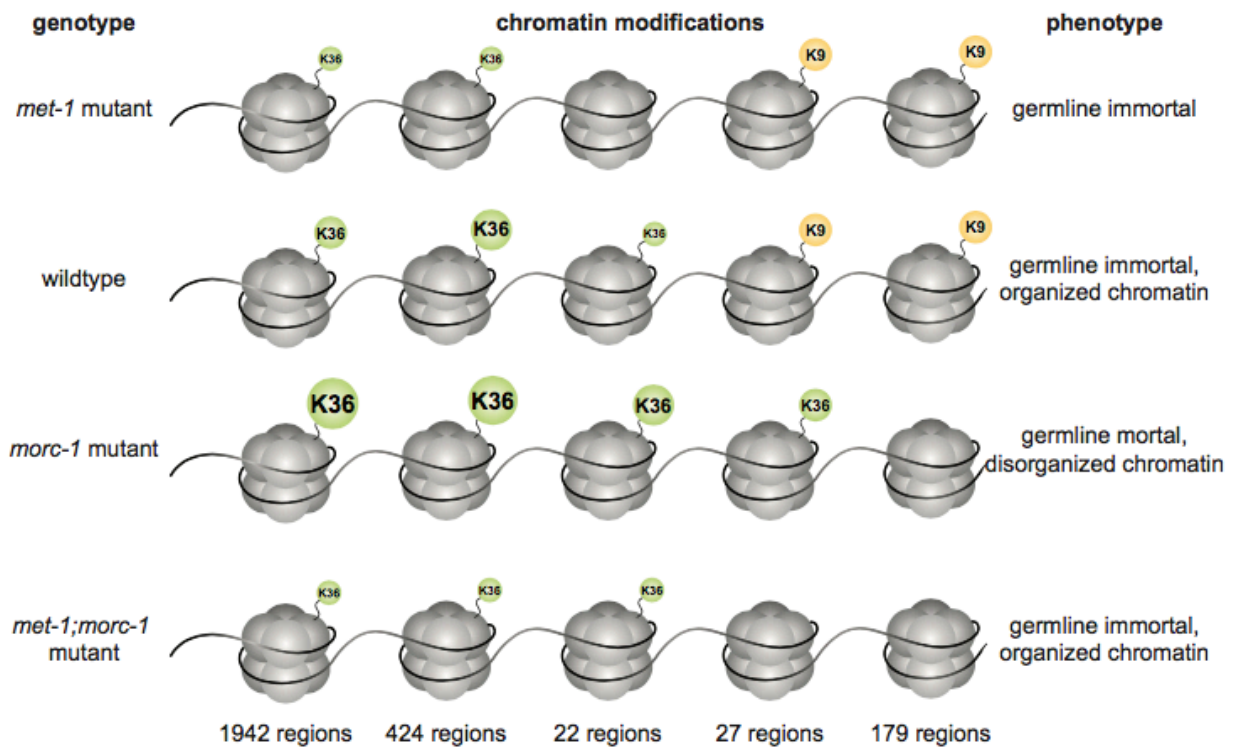
Figure 2.42 Model of MORC-1 and MET-1 effects on chromatin modifications and fertility.

MATERIALS AND METHODS

Strains and worm maintenance

C. elegans were maintained using standard procedures at 20°C unless otherwise noted. Bristol N2 strain was used for wildtype control. All other strains used in this study are listed in Table 2.2. The ubiquitous *morc-1::gfp* extrachromosomal array was made using a 4.1kb *dpy-30* promoter, 2.9kb of *morc-1* genomic coding sequence with no termination codon, 0.9kb *gfp* coding sequence, and 0.6kb of *tbb-2* 3'UTR. This was injected at a concentration of 1.0ng/μl with 2.0ng/μl *myo2::rfp* co-injection marker. The *morc-1::3xflag* strain was made by standard Crispr/Cas9 methods using crRNA: AATATTGAGCACCAAATCAA and repair template: GTCGAGGAGAAACTGAAAAATATCGAACATCAGATTAAAGGAAAGAAAAAAGACT ACAAAGACCATGACGGTGATTATAAAGATCATGATATCGATTACAAGGATGACGA TGACAAGTAATATTCCAGCTCGCCCTGTTTTCCCAGAAGGACGTACTTGTAC (Paix et al., 2015). Worms were fed OP50 *E. coli* for all experiments that did not involve RNAi. RNAi experiments were performed with HT115 *E. coli* and the empty vector *L4440* as a control. The *met-1* RNAi clone was generated by inserting 713bp of *met-1* coding sequence into the *L4440* vector using SacII and KpnI restriction enzymes. The *met-1* fragment was made with primers TGAATGGTACCTCAAATCCACTCTGC and ATCATCCGCGGCGCTTCAACTGCTAA.

RNAi sensitivity assays

Bacterial RNAi clones were grown from the Ahringer library(Kamath et al., 2003). Synchronized L1 worms were plated on the indicated RNAi clone or on the empty vector *L4440*. Sensitivity to *lir-1* RNAi was determined by counting arrested or dead worms 96 hours after plating (to account for any developmental delay). Sensitivity to *lin-15b* RNAi was determined by counting adult worms showing multiple vulvas after two generations on *lin-15b* RNAi.

Germline mortality assays

Worms were maintained at 20ºC, embryos were isolated through hypochlorite preparation and hatched overnight in M9. Synchronized L1 worms were plated on OP50 at 25ºC, designated as the P0 generation. Twenty of the progeny (F1 generation) of these worms were transferred to a new plate at the L2 or L3 stage. This was repeated for each generation. To test fertility in a given generation, L2 or L3 hermaphrodites were singled to new OP50 plates and their live progeny were counted.

RNAi inheritance assays

Synchronized L1 worms (P0 generation) were plated on the *L4440* empty vector or *gfp* RNAi. P0 gravid adults were subjected to hypochlorite egg prep to generate F1 embryos. Synchronized L1 worms from F1 and subsequent generations were plated on OP50. At each generation, gravid adults were imaged with an Olympus BX61 microscope or collected in TriReagent for RNA extraction. Germline nuclear GFP brightness was categorically scored as on or off.

RNA extraction and sequencing

RNA extraction from gravid hermaphrodites was performed as previous described (Billi et al., 2012). In brief, worms were collected in TriReagent (Ambion) and subjected to three freeze-thaw cycles. After phase-separating using 1-bromo-3-chloropropane (BCP), RNA was precipitated in isopropanol at -80°C for one hour. RNA was pelleted by centrifugation at 15,000g for 30 minutes at 4°C. After two washes in 70% ethanol and one wash in 75% ethanol, the pellet was resuspended in water. mRNA libraries were made using Illumina TruSeq Stranded mRNA Library Prep Kit for NeoPrep and an Illumina NeoPrep Library Prep machine according to manufacturer's instructions. Phosphate independent small RNA libraries were made as described in (Phillips et al., 2014) using RNA 5' polyphosphatase (Illumina) and the Illumina TruSeq Small RNAi Library Prep Kit. Libraries were sequenced on the Illumina HiSeq 2000 platform.

Small RNA sequencing data analysis

Raw qseq files were converted to fastq format using custom Awk scripts. Small RNAs were parsed from adapter sequences and reads 18-32 nt long passing a Q30 threshold were aligned to the *C. elegans* genome (WS230) using CASHX v. 2.3 (Fahlgren et al., 2009). Total read counts for *Mutator*-class siRNAs (Phillips et al., 2014; Zhang et al., 2011) were analyzed for differential expression using DESeq2 (Love et al., 2014). Scatter plots were generated using R.

mRNA sequencing data analysis

Raw qseq files were converted to fastq format using custom Awk scripts. Reads were adapter-trimmed and quality filtered (Q30) using Trimmomatic v. 0.35 (Bolger et al., 2014)and aligned to the *C. elegans* genome and transcriptome (WS230) using TopHat2 (Kim et al., 2013). Differential gene expression analysis was done using CuffDiff2 (Trapnell et al., 2013). Plots were generated using R and CummeRBund v. 2.12.1. Gene enrichment analysis was done in R using Fisher's exact test and p-values were corrected for multiple comparisons using the Bonferroni adjustment. Venn diagrams were generated using BioVenn (Hulsen et al., 2008).

Quantitative RT-PCR

Taqman small RNA probes were made by Applied Biosystems. cDNA was made using 50ng of total RNAi using Multiscribe Reverse Transcriptase (Applied Biosystems). Analysis was performed using a Realplex thermocycler (Eppendorf) with TaqMan Universal PCR Master Mix, No AmpErase UNG (Applied Biosystems). A Taqman small RNA probe detecting U18 was used for normalization of *gfp* 22G levels.

Heterochromatin localization assays

Worms expressing the *gwIs4* reporter in indicated backgrounds were maintained at 20ºC. L4 hermaphrodites were picked to plates at 25°C to lay the P0 generation embyros. P0 L4s were picked to new plates to produce the F1 generation. At the specified generation (F4 for *morc-1(-),* F1 for *nrde-2(-)*and *hrde-1(-))*, embryos were

dissected from day 1 adults, freeze-cracked, and DAPI stained as described in (Seydoux and Dunn, 1997). Embryos were imaged in GFP and DAPI channels with an Olympus BX61 microscope at 60X magnification, across 10um of Z-stacks at 250nm intervals. Deconvolution was performed in Huygens Essential software. DAPI-labeled nuclear diameters were measured in ImageJ by averaging 3 diameter measurements, and nuclei were divided into three zones of equal volume. The distance from nuclear boundary to the center of GFP foci was measured in ImageJ, and array localization was assigned to Zone 1 (peripheral third), Zone 2 (intermediate third), or Zone 3 (interior third).

X chromosome staining

DNA FISH probes were prepared and utilized as described in (Csankovszki et al., 2004). In brief, adult worms were dissected in sperm salts (50mM PIPES pH=7, 25mM KCl, 1mM $MgSO_4$, 45mM NaCl, 2mM $CaCl_2$) and fixed with 2% paraformaldehyde (PFA) for five minutes in a humid chamber. After freeze-cracking, slides were washed three times in PBST (1x PBS, 1mM EDTA, 0.5% Triton X-100) for 10 minutes. Slides were then incubated in increasing concentrations of ethanol, two minutes each in 70%, 80%, 95%, 100% ethanol. Ten microliters of FISH probe was added to each slide and then it was denatured on a 95°C heat block for three minutes. Slides were incubated in a humid chamber at 37°C overnight. The following washes were performed at 39°C: five minutes in 2XSSC/50%formamide (three times), five minutes in 2XSSC (three times), 10 minutes in 1XSSC (once). The final wash was performed in 4XSSC with $0.01\mu g/ml$ DAPI for 15 minutes at room temperature. Slides were mounted with Vectashield (Vector Labs) and imaged on a Olympus BX61 microscope at 60X magnification The volume of intestinal nuclei and the X chromosome were quantified as described in (Lau et al., 2014) using Slidebook.

Male rescue

On day one, *him-8(-);xol-1(-);sex-1(-)* embryos were isolated with hypochlorite and grown on the indicated RNAi clone or empty vector RNAi (*L4440*) and grown at 20°C. On day four, young adults were transferred to a new plate of the same RNAi

clone (three worms per plate, four plates per RNAi clone) and allowed to lay embryos overnight. The morning of day five, the adult worms were removed and embryos were counted. On day six, unhatched (dead) embryos were counted. On days eight and nine, male worms were counted. Rescue of male lethality was calculated as the proportion of live embryos that developed into male worms.

Chromatin immunoprecipitation and sequencing

500μl frozen worm pellets were ground into powder and crosslinked with 2% formaldehyde in RIPA buffer (1XPBS, 1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS). Crosslinked chromatin was sonicated using a Diagenode Bioruptor for three eight-minute cycles on high amplitude, 30 seconds on/off. Chromatin was nutated overnight at 4°C with 2μg of the designated antibody and then for two hours with 50μl dynabeads. After three washes with 800μl LiCl buffer (100mM Tris-Cl pH7.5, 500mM LiCl, 1% NP-40, 1% sodium deoxycholate), chromatin was decrosslinked in worm lysis buffer (0.1M Tris-Cl pH7.5, 0.1M NaCl, 50mM EDTA, 1% SDS) for at least four hours at 65°C. DNA was extracted by phenolchloroform and dissolved in TE buffer. RNAse treatment was performed for at least one hour at 37°C. Antibodies used were Abcam ab8898 (H3K9me3) and WAKO 300-95289 (H3K36me3). For each ChIP sample, an input library was generated from 10% of the amount of chromatin in the IP. These samples were decrosslinked and DNA extraction of input samples was performed as described above. ChIP-seq libraries were made using NuGen Ovation Ultralow Library Systems v2 according to manufacturer's instructions and sequenced on the Illumina HiSeq 2000 platform.

ChIP-seq analysis

Thirty-nine DNA ChIP-seq libraries were sequenced and used in this study. DNA ChIP-seq libraries for different samples were labeled with 6-nt index, pooled together and sequenced. All libraries were sequenced using Illumina HiSeq 2000 platform: 50-nt single-end run with index sequencing. De-multiplexed raw data was provided by the sequencing provider in fastq format, which was deposited in NCBI (GEO accession number: GSE89887).

H3K9me3 and H3K36me3 data analysis: H3K9me3 or H3K36me3 ChIP-seq sequencing results in fastq format were aligned to the WS190(ce6) genome using bowtie 0.12.7. All analysis was performed using Python and R. RPKM value for each 1kb was calculated by the number of reads that perfectly aligned to both strands of 1kb-non-overlaping regions, normalized by the sequencing depth of the library in millions. All analysis excluded misannotated regions and *morc-1(-)*-duplication region (chromosome V position 2350000-2570000bp). All types of targets were identified as 1.5 Fold change with FDR<0.05 (Maniar and Fire, 2011) and as an overlap of two replicates. A table with RPKM values and lists of targets was deposited in NCBI (GEO accession number: GSE89887). Two-region Venn diagrams were generated using VennDiagram package in R; three-region Venn diagram was generated using http://www.benfrederickson.com/venn-diagrams-with-d3.js/. Exemplary regions were plotted at 1bp resolution, normalized by the sequencing depth of the corresponding library. Anchor plots were calculated +/- 20kb around start of each 1kb region of corresponding target list and plotted with smoothing window of 10bp.

Immunofluorescence microscopy

Dissected adult worms were fixed in 2% formaldehyde and 1x sperm salts for five minutes in a humid chamber. Samples were then freeze-cracked and fixed in methanol for 10 minutes at -20°C and washed three times in PBST (PBS+0.2% tween-20). Slides were incubated with primary antibody in a humid chamber overnight at 4°C. Slides were washed three times in PBST and then incubated with secondary antibody in a humid chamber for one hour at room temperature. Slides were washed three times in PBST and mounted using Vectashield with DAPI (Vectorlabs H-1200). Slides were imaged at 63X magnification on a Zeiss LSM700 confocal microscope. Antibodies used were Abcam ab8898 (rabbit anti-H3K9me3), Sigma F1804 (mouse anti-flag), Alexa anti-rabbit 674, Alexa anti-mouse 555. All washes were performed at room temperature.

Genetic screen

Starting with a *morc-1(-);pie1P-gfp-h2b* strain, worms in the fourth larval stage were mutagenized in 0.6mM N-ethyl-N-nitrosourea (ENU) for four hours. After three washes in M9, 100 worms (P0 generation) were transferred to each of 32 plates. Worms were grown for five days at 20°C. Adult worms (F1 generation) were bleached to isolate embryos (F2 generation) which were plated, grown at 25°C, and then bleached to isolate embryos (F3 generation). Worms were maintained at 25°C and bleached at adulthood for every generation until F14. Ten larval worms were singled from each of the remaining fertile pools (29 of 32) and their progeny were counted. The most fertile worms of each pool were then propagated as clonalized lines. Genomic DNA was extracted from 20 clonalized lines representing 17 pools (three pools were represented by two lines each) using the Gentra Puregene Tissue kit (Qiagen). Genomic paired-end DNA libraries were made using the Illumina Truseq Nano DNA LT Library Prep Kit according to manufacturer's instructions and sequenced on the Illumina Hi Seq 2000 platform.

Genome sequence data processing

Raw paired-end fastq files were demultiplexed based on 6-nt barcodes and aligned to the *C. elegans* genome (version WS220) using Bowtie2 (Langmead and Salzberg, 2012) and the following parameters: -q --phred64 -N 1 --end-to-end. Alignment SAM files were filtered using samtools (Li et al., 2009) to only include alignments with a bit FLAG containing 2, which indicates that each segment aligned properly, and then converted to BAM format. Alignment BAM files were sorted (SortSam), named (AddOrReplaceReadGroups), and indexed (BuildBamIndex) using functions from the PICARD suite of tools for manipulating high-throughput sequencing data (https://github.com/broadinstitute/picard). Alignment BAM files were realigned to improve indel calling using RealignerTargetCreator and IndeRealigner from the Genome Analysis ToolKit (GATK) suite of tools (McKenna et al., 2010). Duplicate reads were marked using PICARD MarkDuplicates, and the resulting BAM files were used with GATK's UnifiedGenotyper command to call variants.

Identification of suppressor mutations

Variants (SNPs) were combined from replicate unmutagenized *morc-1(-)* strains using GATK's CombineVariants. To identify variants in each mutagenized strain that do not appear in the unmutagenized *morc-1(-)* strain, GATK's SelectVariants function was used to subtract combined unmutagenized *morc-1(-)* variants from each mutagenized strain. Unique variants from all mutagenized strains were filtered using SnpSift (Cingolani et al., 2012) to retain only variants with a high quality score (QUAL≥30) and covered by at least five reads (DP≥5). To determine the effect of each variant in each strain on encoded proteins, SnpEff (Cingolani et al., 2012) was used to annotate each variant and predict its impact on any encoded protein. Genes of interest in the genetic screen were defined as genes that contained different variants with "high" impact on the protein-coding sequence (*e.g.* non-synonymous replacement, start codon loss/gain) across two or more of the mutagenized strains. *met-1*, for example, contained four different variants in six different mutagenized strains that originated from four of the original 32 pools.

Whole-mount DAPI staining

Worms were nutated in M9 for 15 minutes at room temperature, then centrifuged at 5,000rpm for 30 seconds. The supernatant was discarded and the pellet was resuspended in 1ml ice cold methanol and nutated at room temperature for 10 minutes. Worms were washed twice with 1ml PBST (1xPBS with 0.1% Tween-20). Worms were centrifuged at 5,000rpm for 30 seconds and then resuspended in 1ml PBST with 100 μg/ml DAPI and nutated at room temperature for 30 minutes. Worms were then centrifuged at 5,000rpm for 30 seconds and then pipetted onto a microscope slide and mounted with Vectashield. Slides were imaged with a Zeiss LSM700 confocal microscope at 63X magnification.

Strain	Genotype	Reference
N2	wildtype, Bristol isolate	
QK80	*morc-1(tm6048) III*	this study
GR1373	*eri-1(mg366) IV*	CGC
YY186	*nrde-2(gg91) II*	Burkhart et al, 2011
WM49	*rde-4(ne301) III*	CGC
EKM89	*hrde-1(tm1200) III*	Buckley et al 2012
QK81	*xkEx50[dpy-30p::morc-1::gfp::tbb2 3'UTR]*	this study
QK82	*morc-1(tm6048) III;xkEx50[dpy-30p::morc-1::gfp::tbb2 3'UTR, myo-2p::rfp]*	this study
YY151	*nrde-2(gg91) II;eri-1(mg366) IV*	Guang et al 2010
QK83	*morc-1(tm6048) III;eri-1(mg366) IV*	this study
QK84	*morc-1::3xflag*	this study
YY513	pkIS32[pie-1p::gfp::h2b]	Buckley et al 2012
YY528	*hrde-1(tm1200) III; pkIS32[pie-1p::gfp::h2b]*	Buckley et al 2012
QK85	*morc-1(tm6048) III;pkIS32[pie-1p::gfp::h2b]*	this study
GW76	*GwIs4[baf-1p::GFP-lacI::let-858 3'UTR; myo-3p::RFP] X*	Meister et al 2010
GW640	*set-25(n5021) III;GwIs4[baf-1p::GFP-lacI::let-858 3'UTR; myo-3p::RFP] X*	Towbin et al 2012
GW637	*met-2(n4256) III;set-25(n5021) III;GwIs4[baf-1p::GFP-lacI::let-858 3'UTR; myo-3p::RFP] X*	Towbin et al 2012
QK86	*morc-1 (tm6048) III;GwIs4[baf-1p::GFP-lacI::let-858 3'UTR; myo-3p::RFP] X*	this study
QK87	*hrde-1 (tm1200) III;GwIs4[baf-1p::GFP-lacI::let-858 3'UTR; myo-3p::RFP] X*	this study
QK88	*nrde-2 (gg91) II;GwIs4[baf-1p::GFP-lacI::let-858 3'UTR; myo-3p::RFP] X*	this study
TY4403	*him-8 (e1489) IV;xol-1 (y9) X;sex-1 (y263) X*	Petty et al 2009
MT16973	*met-1(n3227) I*	CGC
VC1666	*met-1(ok2172) I*	CGC
QK93	*met-1(n3227) I; morc-1 (tm6048) III*	this study
QK94	*met-1(ok2172) I; morc-1 (tm6048) III*	this study
QK95	*met-1(xk1) I; morc-1 (tm6048) III*	this study
QK96	*met-1(xk2) I; morc-1 (tm6048) III*	this study
QK97	*met-1(xk3) I; morc-1 (tm6048) III*	this study
QK98	*met-1(xk4) I; morc-1 (tm6048) III*	this study
QK99	*met-1(xk2) I; hrde-1 (tm1200) III*	this study
QK100	*met-1(xk4) I; hrde-1 (tm1200) III*	this study
QK101	*met-1(n3227) I; pkIS32[pie-1p::gfp::h2b]*	this study
QK102	*met-1(n3227) I; morc-1 (tm6048) III; pkIS32[pie-1p::gfp::h2b]*	this study
QK103	*met-1(xk4) I; pkIS32[pie-1p::gfp::h2b]*	this study
QK104	*met-1(xk4) I; morc-1 (tm6048) III; pkIS32[pie-1p::gfp::h2b]*	this study

Table 2.2. Additional strains used in this study.

ACKNOWLEDGEMENTS

REFERENCES

Ahmed, S., and Hodgkin, J. (2000). MRT-2 checkpoint protein is required for germline immortality and telomere replication in C. elegans. Nature *403*, 159–164.

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

Andrews, F.H., Tong, Q., Sullivan, K.D., Cornett, E.M., Zhang, Y., Ali, M., Ahn, J., Pandey, A., Guo, A.H., Strahl, B.D., et al. (2016). Multivalent Chromatin Engagement and Inter-domain Crosstalk Regulate MORC3 ATPase. Cell Rep *16*, 3195–3207.

Ashe, A., Sapetschnig, A., Weick, E.-M., Mitchell, J., Bagijn, M.P., Cording, A.C., Doebley, A.-L., Goldstein, L.D., Lehrbach, N.J., Le Pen, J., et al. (2012). piRNAs can trigger a multigenerational epigenetic memory in the germline of C. elegans. Cell *150*, 88–99.

Batista, P.J., Ruby, J.G., Claycomb, J.M., Chiang, R., Fahlgren, N., Kasschau, K.D., Chaves, D.A., Gu, W., Vasale, J.J., Duan, S., et al. (2008). PRG-1 and 21U-RNAs interact to form the piRNA complex required for fertility in C. elegans. Mol. Cell *31*, 67–78.

Billi, A.C., Alessi, A.F., Khivansara, V., Han, T., Freeberg, M., Mitani, S., and Kim, J.K. (2012). The Caenorhabditis elegans HEN1 ortholog, HENN-1, methylates and stabilizes select subclasses of germline small RNAs. PLoS Genet. *8*, e1002617.

Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics *30*, 2114–2120.

Buckley, B.A., Burkhart, K.B., Gu, S.G., Spracklin, G., Kershner, A., Fritz, H., Kimble, J., Fire, A., and Kennedy, S. (2012). A nuclear Argonaute promotes multigenerational epigenetic inheritance and germline immortality. Nature *489*, 447–451.

Burkhart, K.B., Guang, S., Buckley, B.A., Wong, L., Bochner, A.F., and Kennedy, S. (2011). A pre-mRNA-associating factor links endogenous siRNAs to chromatin regulation. PLoS Genet. *7*, e1002249.

Burton, N.O., Burkhart, K.B., and Kennedy, S. (2011). Nuclear RNAi maintains heritable gene silencing in Caenorhabditis elegans. Proc. Natl. Acad. Sci. U.S.a. *108*, 19683–19688.

Carmi, I., Kopczynski, J.B., and Meyer, B.J. (1998). The nuclear hormone receptor SEX-1 is an X-chromosome signal that determines nematode sex. Nature *396*, 168–173.

Cingolani, P., Platts, A., Wang, L.L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Lu, X.,

and Ruden, D.M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly (Austin) *6*, 80–92.

Csankovszki, G., McDonel, P., and Meyer, B.J. (2004). Recruitment and spreading of the C. elegans dosage compensation complex along X chromosomes. Science *303*, 1182–1185.

Das, P.P., Bagijn, M.P., Goldstein, L.D., Woolford, J.R., Lehrbach, N.J., Sapetschnig, A., Buhecha, H.R., Gilchrist, M.J., Howe, K.L., Stark, R., et al. (2008). Piwi and piRNAs act upstream of an endogenous siRNA pathway to suppress Tc3 transposon mobility in the Caenorhabditis elegans germline. Mol. Cell *31*, 79–90.

de Albuquerque, B.F.M., Placentino, M., and Ketting, R.F. (2015). Maternal piRNAs Are Essential for Germline Development following De Novo Establishment of Endo-siRNAs in Caenorhabditis elegans. Dev. Cell *34*, 448–456.

Dernburg, A.F., Zalevsky, J., Colaiácovo, M.P., and Villeneuve, A.M. (2000). Transgene-mediated cosuppression in the C. elegans germ line. Genes Dev. *14*, 1578–1583.

Dutta, R., and Inouye, M. (2000). GHKL, an emergent ATPase/kinase superfamily. Trends Biochem. Sci. *25*, 24–28.

Fahlgren, N., Sullivan, C.M., Kasschau, K.D., Chapman, E.J., Cumbie, J.S., Montgomery, T.A., Gilbert, S.D., Dasenko, M., Backman, T.W.H., Givan, S.A., et al. (2009). Computational and analytical framework for small RNA profiling by high-throughput sequencing. Rna *15*, 992–1002.

Furuhashi, H., Takasaki, T., Rechtsteiner, A., Li, T., Kimura, H., Checchi, P.M., Strome, S., and Kelly, W.G. (2010). Trans-generational epigenetic regulation of C. elegans primordial germ cells. Epigenetics Chromatin *3*, 15.

Greer, E.L., Beese-Sims, S.E., Brookes, E., Spadafora, R., Zhu, Y., Rothbart, S.B., Aristizábal-Corrales, D., Chen, S., Badeaux, A.I., Jin, Q., et al. (2014). A histone methylation network regulates transgenerational epigenetic memory in C. elegans. Cell Rep *7*, 113–126.

Grishok, A., Tabara, H., and Mello, C.C. (2000). Genetic requirements for inheritance of RNAi in C. elegans. Science *287*, 2494–2497.

Gu, S.G., Pak, J., Guang, S., Maniar, J.M., Kennedy, S., and Fire, A. (2012). Amplification of siRNA in Caenorhabditis elegans generates a transgenerational sequence-targeted histone H3 lysine 9 methylation footprint. Nat. Genet. *44*, 157–164.

Guang, S., Bochner, A.F., Burkhart, K.B., Burton, N., Pavelec, D.M., and Kennedy, S. (2010). Small regulatory RNAs inhibit RNA polymerase II during the elongation phase of transcription. Nature *465*, 1097–1101.

Guang, S., Bochner, A.F., Pavelec, D.M., Burkhart, K.B., Harding, S., Lachowiec, J., and Kennedy, S. (2008). An Argonaute transports siRNAs from the cytoplasm to the nucleus. Science *321*, 537–541.

Hajkova, P., Erhardt, S., Lane, N., Haaf, T., El-Maarri, O., Reik, W., Walter, J., and Surani, M.A. (2002). Epigenetic reprogramming in mouse primordial germ cells. Mech. Dev. *117*, 15–23.

Han, T., Manoharan, A.P., Harkins, T.T., Bouffard, P., Fitzpatrick, C., Chu, D.S., Thierry-Mieg, D., Thierry-Mieg, J., and Kim, J.K. (2009). 26G endo-siRNAs regulate spermatogenic and zygotic gene expression in Caenorhabditis elegans. Proc. Natl. Acad. Sci. U.S.a. *106*, 18674–18679.

Harris, C.J., Husmann, D., Liu, W., Kasmi, F.E., Wang, H., Papikian, A., Pastor, W.A., Moissiard, G., Vashisht, A.A., Dangl, J.L., et al. (2016). Arabidopsis AtMORC4 and AtMORC7 Form Nuclear Bodies and Repress a Large Number of Protein-Coding Genes. PLoS Genet. *12*, e1005998.

Hulsen, T., de Vlieg, J., and Alkema, W. (2008). BioVenn - a web application for the comparison and visualization of biological lists using area-proportional Venn diagrams. BMC Genomics *9*, 488.

Iyer, L.M., Abhiman, S., and Aravind, L. (2008). MutL homologs in restriction-modification systems and the origin of eukaryotic MORC ATPases. Biol. Direct *3*, 8.

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.

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

Ketting, R.F., and Plasterk, R.H. (2000). A genetic link between co-suppression and RNA interference in C. elegans. Nature *404*, 296–298.

Ketting, R.F., Haverkamp, T.H., van Luenen, H.G., and Plasterk, R.H. (1999). Mut-7 of C. elegans, required for transposon silencing and RNA interference, is a homolog of Werner syndrome helicase and RNaseD. Cell *99*, 133–141.

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

Kim, J.K., Gabel, H.W., Kamath, R.S., Tewari, M., Pasquinelli, A., Rual, J.-F., Kennedy, S., Dybbs, M., Bertin, N., Kaplan, J.M., et al. (2005). Functional genomic analysis of RNA interference in C. elegans. Science *308*, 1164–1167.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359.

Lau, A.C., Nabeshima, K., and Csankovszki, G. (2014). The C. elegans dosage compensation complex mediates interphase X chromosome compaction. Epigenetics Chromatin *7*, 31.

Lee, H.-C., Gu, W., Shirayama, M., Youngman, E., Conte, D., and Mello, C.C. (2012). C. elegans piRNAs mediate the genome-wide surveillance of germline transcripts. Cell *150*, 78–87.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078–2079.

Li, S., Yen, L., Pastor, W.A., Johnston, J.B., Du, J., Shew, C.J., Liu, W., Ho, J., Stender, B., Clark, A.T., et al. (2016). Mouse MORC3 is a GHKL ATPase that localizes to H3K4me3 marked chromatin. Proc. Natl. Acad. Sci. U.S.a. 201609709.

Liu, Y., Tempel, W., Zhang, Q., Liang, X., Loppnau, P., Qin, S., and Min, J. (2016). Family-wide Characterization of Histone Binding Abilities of Human CW Domain-containing Proteins. J. Biol. Chem. *291*, 9000–9013.

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

Luteijn, M.J., van Bergeijk, P., Kaaij, L.J.T., Almeida, M.V., Roovers, E.F., Berezikov, E., and Ketting, R.F. (2012). Extremely stable Piwi-induced gene silencing in Caenorhabditis elegans. Embo J. *31*, 3422–3430.

Maniar, J.M., and Fire, A.Z. (2011). EGO-1, a C. elegans RdRP, modulates gene expression via production of mRNA-templated short antisense RNAs. Curr. Biol. *21*, 449–459.

Mao, H., Zhu, C., Zong, D., Weng, C., Yang, X., Huang, H., Liu, D., Feng, X., and Guang, S. (2015). The Nrde Pathway Mediates Small-RNA-Directed Histone H3 Lysine 27 Trimethylation in Caenorhabditis elegans. Curr. Biol. *25*, 2398–2403.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., et al. (2010). The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. *20*, 1297–1303.

Meier, B., Barber, L.J., Liu, Y., Shtessel, L., Boulton, S.J., Gartner, A., and Ahmed, S. (2009). The MRT-1 nuclease is required for DNA crosslink repair and telomerase activity in vivo in Caenorhabditis elegans. Embo J. *28*, 3549–3563.

Meister, P., Towbin, B.D., Pike, B.L., Ponti, A., and Gasser, S.M. (2010). The spatial

dynamics of tissue-specific promoters during C. elegans development. Genes Dev. *24*, 766–782.

Meyer, B.J., and Casson, L.P. (1986). Caenorhabditis elegans compensates for the difference in X chromosome dosage between the sexes by regulating transcript levels. Cell *47*, 871–881.

Miller, L.M., Plenefisch, J.D., Casson, L.P., and Meyer, B.J. (1988). xol-1: a gene that controls the male modes of both sex determination and X chromosome dosage compensation in C. elegans. Cell *55*, 167–183.

Moissiard, G., Bischof, S., Husmann, D., Pastor, W.A., Hale, C.J., Yen, L., Stroud, H., Papikian, A., Vashisht, A.A., Wohlschlegel, J.A., et al. (2014). Transcriptional gene silencing by Arabidopsis microchidia homologues involves the formation of heteromers. Proc. Natl. Acad. Sci. U.S.a. *111*, 7474–7479.

Moissiard, G., Cokus, S.J., Cary, J., Feng, S., Billi, A.C., Stroud, H., Husmann, D., Zhan, Y., Lajoie, B.R., McCord, R.P., et al. (2012). MORC family ATPases required for heterochromatin condensation and gene silencing. Science *336*, 1448–1451.

Ni, J.Z., Kalinava, N., Chen, E., Huang, A., Trinh, T., and Gu, S.G. (2016). A transgenerational role of the germline nuclear RNAi pathway in repressing heat stress-induced transcriptional activation in C. elegans. Epigenetics Chromatin *9*, 3.

Paix, A., Folkmann, A., Rasoloson, D., and Seydoux, G. (2015). High Efficiency, Homology-Directed Genome Editing in Caenorhabditis elegans Using CRISPR-Cas9 Ribonucleoprotein Complexes. Genetics *201*, 47–54.

Pastor, W.A., Stroud, H., Nee, K., Liu, W., Pezic, D., Manakov, S., Lee, S.A., Moissiard, G., Zamudio, N., Bourc'his, D., et al. (2014). MORC1 represses transposable elements in the mouse male germline. Nat Commun *5*, 5795.

Phillips, C.M., Brown, K.C., Montgomery, B.E., Ruvkun, G., and Montgomery, T.A. (2015). piRNAs and piRNA-Dependent siRNAs Protect Conserved and Essential C. elegans Genes from Misrouting into the RNAi Pathway. Dev. Cell *34*, 457–465.

Phillips, C.M., Montgomery, B.E., Breen, P.C., Roovers, E.F., Rim, Y.-S., Ohsumi, T.K., Newman, M.A., van Wolfswinkel, J.C., Ketting, R.F., Ruvkun, G., et al. (2014). MUT-14 and SMUT-1 DEAD box RNA helicases have overlapping roles in germline RNAi and endogenous siRNA formation. Curr. Biol. *24*, 839–844.

Phillips, C.M., Montgomery, T.A., Breen, P.C., and Ruvkun, G. (2012). MUT-16 promotes formation of perinuclear mutator foci required for RNA silencing in the C. elegans germline. Genes Dev. *26*, 1433–1444.

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*, e1001091.

Santos, F., and Dean, W. (2004). Epigenetic reprogramming during early development in mammals. Reproduction *127*, 643–651.

Seydoux, G., and Dunn, M.A. (1997). Transcriptionally repressed germ cells lack a subpopulation of phosphorylated RNA polymerase II in early embryos of Caenorhabditis elegans and Drosophila melanogaster. Development *124*, 2191–2201.

Shirayama, M., Seth, M., Lee, H.-C., Gu, W., Ishidate, T., Conte, D., and Mello, C.C. (2012). piRNAs initiate an epigenetic memory of nonself RNA in the C. elegans germline. Cell *150*, 65–77.

Smelick, C., and Ahmed, S. (2005). Achieving immortality in the C. elegans germline. Ageing Res. Rev. *4*, 67–82.

Snyder, M.J., Lau, A.C., Brouhard, E.A., Davis, M.B., Jiang, J., Sifuentes, M.H., and Csankovszki, G. (2016). Anchoring of Heterochromatin to the Nuclear Lamina Reinforces Dosage Compensation-Mediated Gene Repression. PLoS Genet. *12*, e1006341–33.

Tabara, H., Sarkissian, M., Kelly, W.G., Fleenor, J., Grishok, A., Timmons, L., Fire, A., and Mello, C.C. (1999). The rde-1 gene, RNA interference, and transposon silencing in C. elegans. Cell *99*, 123–132.

Tabara, H., Yigit, E., Siomi, H., and Mello, C.C. (2002). The dsRNA binding protein RDE-4 interacts with RDE-1, DCR-1, and a DExH-box helicase to direct RNAi in C. elegans. Cell *109*, 861–871.

Towbin, B.D., González-Aguilera, C., Sack, R., Gaidatzis, D., Kalck, V., Meister, P., Askjaer, P., and Gasser, S.M. (2012). Step-wise methylation of histone H3K9 positions heterochromatin at the nuclear periphery. Cell *150*, 934–947.

Trapnell, C., Hendrickson, D.G., Sauvageau, M., Goff, L., Rinn, J.L., and Pachter, L. (2013). Differential analysis of gene regulation at transcript resolution with RNA-seq. Nat. Biotechnol. *31*, 46–53.

Wang, D., Kennedy, S., Conte, D., Kim, J.K., Gabel, H.W., Kamath, R.S., Mello, C.C., and Ruvkun, G. (2005). Somatic misexpression of germline P granules and enhanced RNA interference in retinoblastoma pathway mutants. Nature *436*, 593–597.

Wu, X., Shi, Z., Cui, M., Han, M., and Ruvkun, G. (2012). Repression of germline RNAi pathways in somatic cells by retinoblastoma pathway chromatin complexes. PLoS Genet. *8*, e1002542.

Zeller, P., Padeken, J., van Schendel, R., Kalck, V., Tijsterman, M., and Gasser, S.M. (2016). Histone H3K9 methylation is dispensable for Caenorhabditis elegans development but suppresses RNA:DNA hybrid-associated repeat instability. Nat. Genet.

1–13.

Zhang, C., Montgomery, T.A., Gabel, H.W., Fischer, S.E.J., Phillips, C.M., Fahlgren, N., Sullivan, C.M., Carrington, J.C., and Ruvkun, G. (2011). mut-16 and other mutator class genes modulate 22G and 26G siRNA pathways in Caenorhabditis elegans. Proc. Natl. Acad. Sci. U.S.a. *108*, 1201–1208.

CHAPTER 3

MORC-1 future directions

AUTHORS: Natasha E. Weiser, Suhua Feng, Kristen C. Brown, Mallory A. Freeberg, Taiowa A. Montgomery, Steven E. Jacobsen, and John K. Kim

INTRODUCTION

The studies described in Chapter Two demonstrate a role for MORC-1 in the regulation of nuclear RNAi effector function downstream of RISC. These data led us to a model in which MORC-1 contributes to H3K9me3 maintenance at a subset of targets by repressing euchromatic encroachment, mediated in part by MET-1. In this chapter, I will discuss ongoing work to address several remaining questions.

First, I will discuss ongoing structure-function analysis of MORC-1. The family of GHKL ATPases are thought to function as molecular clamps, forming constitutive

dimers at the C-terminus and dimerizing upon ATP binding at the N-terminus (Dutta and Inouye 2000; Corbett and Berger 2003; Corbett and Berger 2005; Li et al. 2016). ATP hydrolysis induces physical manipulation of the substrate and release of the dimer (Dutta and Inouye 2000; Corbett and Berger 2003; Corbett and Berger 2005; Li et al. 2016). We predict that the ATPase domain of *C. elegans* MORC-1 is required for its function in the nuclear RNAi pathway. We expect that MORC-1 function also requires chromatin-binding via the CW zinc finger. In this chapter, I discuss preliminary data and provide an outline of future studies to elucidate the roles of the ATPase domain and CW zinc-finger.

Second, I will discuss some of the pitfalls of our original *morc-1(-)* suppressor screen and propose a modified screen that I expect will elucidate additional chromatin factors that suppress *morc-1(-).*

Third, I will discuss preliminary findings suggesting that MORC-1 is essential for genome stability. Genome instability is a common feature of germline mortal strains and has recently been found to occur in the absence of global H3K9 methylation (Ahmed and Hodgkin 2000; Meier et al. 2009; Zeller et al. 2016). Accumulation of DNA damage could contribute to the progressive nature of the *morc-1(-)* fertility defect.

Finally, I will discuss evidence suggesting that MORC-1 interacts with the CSR-1 22G endo-siRNAs. I will outline further studies to elucidate this intriguing connection between MORC-1, a gene-silencing factor, and CSR-1, which is thought to promote gene expression.

STRUCTURE-FUNCTION ANALYSIS OF MORC-1

We have shown that MORC-1 regulates transgenerational chromatin organization at a subset of nuclear RNAi targets. We hope to further characterize the role of MORC-1 in regulation of germline chromatin by elucidating the role of the GHKL ATPase domain, characterizing the chromatin-binding properties of MORC-1, and identifying physical interactors.

Characterization of MORC-1 ATPase domain

Morc proteins are members of the GHKL ATPase superfamily, a diverse group of proteins that share a unique ATPase domain structure. Although the members of this family are divergent at the amino acid level, they share a characteristic structure called the Bergerat fold (Dutta and Inouye 2000). This superfamily is named for four canonical members: DNA gyrase, Hsp90, histidine kinase, and MutL. These enzymes show constitutive dimerization at the C-terminus and ATP-dependent dimerization at the N-terminus. The release of the two N-termini requires ATP hydrolysis. Thus, these enzymes are thought to function as molecular clamps to drive conformational changes in their substrates (Dutta and Inouye 2000; Iyer et al. 2008; Li et al. 2016). In the Microrchidia family, constitutive dimerization is driven by the C-terminal coiled-coil domain with the ATPase activity regulating N-terminal dimerization as in other GHKL ATPases (Li et al. 2016). Interestingly, condensins employ a similar mechanism of constitutive C-terminal dimerization and ATPase-mediated N-terminal dimerization, albeit with a structurally distinct ATPase domain, to entrap and manipulate chromatin (Hirano 2016).

Using CRISPR-Cas-9-mediated genome editing, we have generated MORC-1::3xFlag variants predicted to inactivate ATP-binding or ATP-hydrolysis at the endogenous *morc-1* locus (DD69-70AA or DDAA and E39A respectively). Based on what is known about other Morc proteins and GHKL ATPases generally, we hypothesize that ATP binding is required for N-terminal dimerization and ATP hydrolysis is required for dimer dissolution and release of MORC-1 from chromatin. We will investigate the functional contribution of the ATPase domain by testing whether the MORC-1(DDAA)::3xFlag and/or MORC-1(E39A)::3xFlag variants recapitulate any of the *morc-1(-)* mutant phenotypes. Our preliminary data suggests that both the DDAA and E39A variants may be hypomorphic. Like the *morc-1(-)* deletion mutant, the *morc-1* variants cause Mrt at 25°C, but with a slower decline in fertility: while the *morc-1(-)* mutants were sterile at F6, the DDAA mutants were sterile at F7, and the E39A mutants were sterile at F11 (Figure 3.1). The *morc-1(WT)::3xFlag* strain maintained wildtype-levels of progeny throughout this experiment, confirming that the Flag tag does not

interfere with MORC-1 function (Figure 3.1). The Mrt phenotype of worms expressing the DDAA or E39A MORC-1 variants suggests that the ATPase activity is required for MORC-1-mediated chromatin regulation. To evaluate the requirement for MORC-1 ATPase activity in exogenous RNAi, we will perform the RNAi inheritance and *lir-1* sensitivity assays with DDAA and E39A variants. We hypothesize that one or both of these variants will cause a defect in both RNAi inheritance and *lir-1* sensitivity.

The activity of the ATPase domain could modulate how MORC-1 interacts with and regulates chromatin. Based on previous studies of condensins, cohesins, and other GHKL ATPases, we predict that dimerization of the N-terminus of MORC-1 forms a ring structure that entraps chromatin (Hirano 2016). If ATP hydrolysis opens this structure to facilitate manipulation of the target chromatin and dissociation of MORC-1 from chromatin, then impairing this activity could increase the strength of MORC-1-chromatin binding. We propose to leverage this in order to identify the genomic targets of MORC-1. Our previous attempts to perform MORC-1(WT)::3xFlag ChIP-seq did not recover enough reads to detect binding over background, which may indicate that interactions between MORC-1 and chromatin are too transient to efficiently capture using standard ChIP-seq methods. We will use the E39A variant to stabilize the interactions of MORC-1 with chromatin and use ChIP-seq of MORC-1(E39A)::3xFlag to identify MORC-1 targets. We will also profile H3K4me3, H3K9me3, and H3K36me3 marks genome wide in the *morc-1(E39A)::3xFlag* strain. We will evaluate these data sets to see whether MORC-1 preferentially binds to chromatin marked with one of these modifications. If ChIP-seq of the E39A variant is unsuccessful, this might indicate that MORC-1 interacts with chromatin indirectly and thus cannot be efficiently crosslinked to chromatin by formaldehyde. To address problems with crosslinking efficiency, we will use a crosslinking agent that crosslinks over greater distances than formaldehyde, such as ethylene glycol bis(sulfosuccinimidyl succinate) (EGS).

Our previous ChIP-seq studies identified about 200 1kb loci that require MORC-1 for H3K9me3 maintenance. We presume that MORC-1 physically interacts with chromatin at or near these loci. However, we do not yet know how MORC-1 is targeted. Our proposed MORC-1(E39A)::3xFlag ChIP-seq, in parallel with ChIP-seq of H3K9me3, H3K4me3, and H3K36me3 marks, may provide clues as to what histone marks MORC-

1 may preferentially bind. This will be addressed further as part of the characterization of the MORC-1 CW zinc-finger.

Characterization of MORC-1 CW zinc-finger domain

Recent studies have shown that the mammalian MORC3 and MORC4 CW zinc-finger domains preferentially bind methylated H3K4 (Andrews et al. 2016; Li et al. 2016; Liu et al. 2016). MORC1 and MORC2, which are more closely related to *C. elegans* MORC-1, were not found to directly bind H3 despite harboring canonical CW zinc-finger motifs (Liu et al. 2016). *C. elegans* MORC-1 contains the two tryptophan residues and three of the four cysteines that are characteristic of CW zinc fingers, however, biochemical studies of mammalian MORC1 and MORC2 suggest that these residues are not sufficient to direct H3 binding (Perry and Zhao 2003; Liu et al. 2016). We have generated a putative zinc-finger dead MORC-1::3xFlag variant by replacing the two conserved tryptophans, W509 and W518, of the CW zinc-finger with alanines to generate MORC-1(WWAA)::3xFlag. If the function of MORC-1 is dependent on its ability to bind chromatin as we expect, then worms expressing the MORC-1(WWAA)::3xFlag variant will phenocopy the *morc-1(-)* deletion mutant and show defects in nuclear RNAi sensitivity and RNAi inheritance and will be Mrt. We also expect the Mrt phenotype to be suppressed by the *met-1* alleles generated from our genetic screen as described in Chapter Two. If MORC-1 does not directly bind chromatin and associates indirectly to regulate chromatin structure, the CW zinc-finger may be dispensable for MORC-1 function. If this is the case, we hope to characterize MORC-1 binding partners (described below), as these may interact with chromatin directly as part of a MORC-1-complex.

Based on the relatively small number of sites that are regulated by MORC-1, we hypothesize that additional factors may control MORC-1 targeting to chromatin. Because we know that in terms of H3K9me3 maintenance and H3K36me3 repression, MORC-1 regulates almost exclusively a subset of HRDE-1 targets, one obvious candidate for driving MORC-1 targeting to chromatin is HRDE-1 RISC. We hypothesize

that MORC-1 localization to the nuclear periphery in the distal germline is dependent on both targeting by HRDE-1 RISC and either targeting or reinforcement from the resulting H3K9me3 marks at the target loci. Thus, we expect that loss of either HRDE-1 RISC or H3K9me3 will result in diffuse nuclear staining of MORC-1 throughout the germline. We will test the genetic requirements for MORC-1 localization to the nuclear periphery by immunofluorescence of MORC-1::3xFlag in several genetic backgrounds. Because *hrde-1* and *morc-1* are tightly linked, we cannot cross the *hrde-1(-)* mutation into any of our CRISPR-tagged *morc-1::3xFlag* strains. Instead, we will evaluate the localization of MORC-1-3xFlag in a *nrde-2(-)* background. As NRDE-2 is shared between germline and somatic nuclear RISCs (Guang et al. 2010; Buckley et al. 2012), we can simultaneously evaluate the contribution of nuclear RISC to MORC-1 localization in the germline and in the soma using the *nrde-2(-)* mutants. It has been reported that *hrde-1* sometimes functions redundantly with other WAGOs including *wago-10* (Shirayama et al. 2012)*.* To account for possible redundancy among WAGOs regulating MORC-1 localization, we will also evaluate localization in *eri-1(-)* and *rrf-1(-)* backgrounds in which all WAGO-class 22G RNAs are depleted. We will also test the contribution of H3K9me3 to MORC-1 localization independent of nuclear RISC function by performing immunofluorescence of MORC-1::3xFlag in genetic mutants for bona fide H3K9 histone methyltransferases, *set-25* and *met-2.*

Identification of the MORC-1 complex

We hypothesize that MORC-1 functions as part of a complex. We will use two parallel approaches to identify MORC-1 interactors. First, we will perform immunoprecipitation and mass spectrometry (IP-MS) using the strain expressing MORC-1::3xFlag. We will also stabilize interactions between MORC-1 and its binding partners by formaldehyde crosslinking. As a complementary approach, we will use BioID2, a novel method employing a promiscuous biotin ligase fused to MORC-1. Upon supplementation of worm media with biotin, the MORC-1-fused biotin ligase will biotinylate nearby proteins, including direct interactors, which can then be affinity

purified with streptavidin beads (Roux et al. 2012; Kim et al. 2016). By performing *in vivo* labeling of MORC-1 binding partners, we will be able to immunoprecipitate members of the MORC-1 complex even if the complex is not stable enough to immunoprecipitate directly.

Both IP-MS and BioID2 approaches will likely identify many proteins that are not true MORC-1 interactors. We will filter the results to generate a high-priority list of putative interactors, prioritizing candidates that were identified using both methods, those with known chromatin-binding domains, and those that have been previously implicated in small RNA pathways through genetic screens. To experimentally validate putative interactors, we will utilize CRISPR-Cas9 to epitope-tag these candidates and then perform co-immunoprecipitation (co-IP) experiments with MORC-1::3xFlag. We will perform co-IP experiments with and without DNAse I treatment to determine whether any putative interactions are DNA-dependent. We will either generate by CRISPR-Cas9 or obtain deletion mutants for candidate genes to test for nuclear RNAi phenotypes, specifically resistance to *lir-1* RNAi, defective RNAi inheritance, and germline mortality.

Summary and Significance

The studies proposed here will address several remaining questions regarding MORC-1 function in the nuclear RNAi pathway. By characterizing the roles of the MORC-1 ATPase and CW domains, we will further elucidate the mechanism by which MORC-1 regulates nuclear RNAi effector function. Importantly, we will reveal whether MORC-1 interacts directly with chromatin by several complementary strategies, including characterization of a putative zinc-finger dead MORC-1 variant, MORC-1 ChIP-seq to identify genomic targets, and two parallel strategies to identify MORC-1 interactors. Our identification of the MORC-1 complex will likely reveal new regulators of nuclear RNAi effector function at the chromatin level.
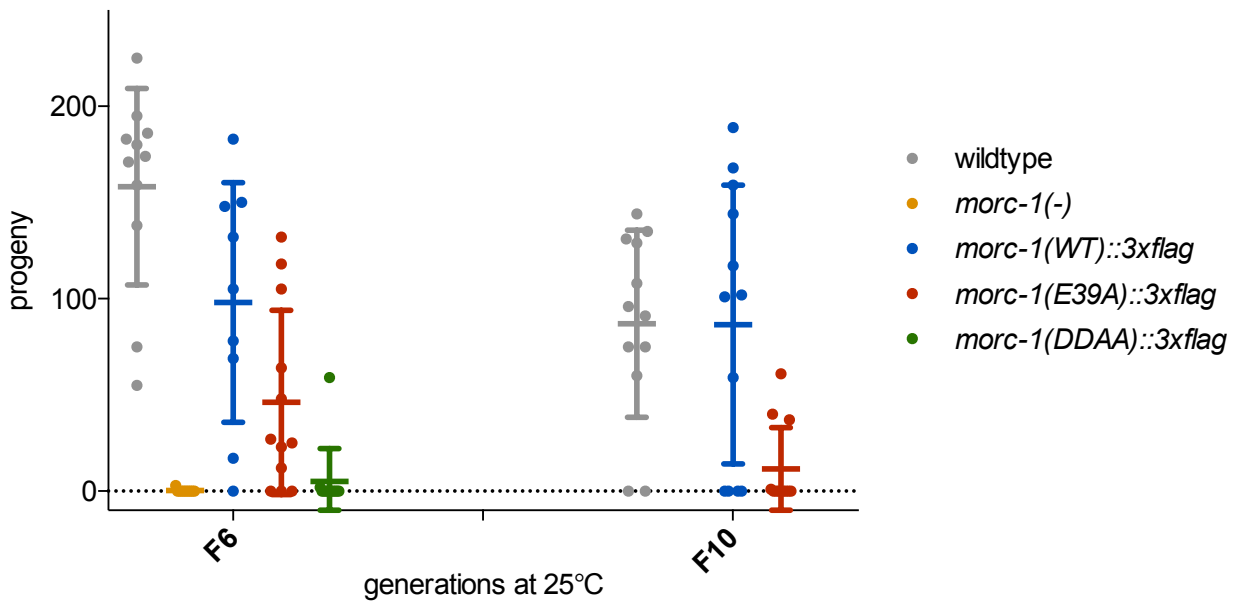
Figure 3.1. ATP-binding and hydrolysis dead MORC-1 variants are hypomorphic. Worms expressing an ATP-binding dead variant, *morc-1(DDAA)::3xFlag*, are sterile after 8 generations (F7) of growth at 25°C. Worms expressing an ATP-hydrolysis dead variant, *morc-1(E39A)::3xFlag,* are nearly sterile after 11 generations (F10) of growth at 25°C. Wildtype and *morc-1(WT)::3xFlag* worms are indefinitely fertile at 25°C. *morc-1(-)* mutants are sterile after 7 generations (F6) at 25°C.

IDENTIFICATION OF NOVEL *MORC-1(-)* SUPRESSORS

We have described the identification of *met-1* as a potent suppressor of *morc-1(-)* germline mortality through a forward genetic screen. Though we speculate that a number of other chromatin-modifying or chromatin-binding proteins could suppress *morc-1* activity, *met-1* was the only high-confidence candidate that emerged from our screen with any established or predicted connection to chromatin regulation. Two additional candidates from the screen, which have no annotated role in chromatin regulation, are discussed below. It is possible that a larger screen would identify more candidates, but this would not solve the larger issue with the design of our screen. Many proteins that interact with and regulate germline chromatin are critical for fertility, especially at elevated temperatures (Kelly 2014). As a result, our screen for mutations that restore *morc-1(-)* germline maintenance over the course of 13 generations at 25°C may have selected against *morc-1(-)* suppressors that independently compromise fertility. For example, the second *C. elegans* H3K36 methyltransferase, MES-4, may antagonize MORC-1 but because it is essential for fertility, any loss-of-function mutations in *mes-4* generated by our screen were likely incompatible with robust fertility over 13 generations at 25°C. More details on a possible role for MES-4 in antagonism of MORC-1 were presented in Chapter Two. Furthermore, our finding that *met-1* suppresses *morc-1(-)* germline mortality but not resistance to nuclear RNAi or defective RNAi inheritance suggests that endogenous and exogenous nuclear RNAi may be antagonized by distinct mechanisms. Thus, a genetic screen for suppressors of the *morc-1(-)* defects relating to exogenous RNAi may identify a distinct set of suppressors.

To address the limitations of our prior screen, we will perform a new *morc-1(-)* suppressor (*smorc*) screen to identify suppressors of the *morc-1(-)* RNAi inheritance defect. This will not require extensive propagation at elevated temperatures and thus may identify suppressing mutations that cause fertility defects in addition to suppression of the *morc-1(-)* RNAi inheritance defect. We will mutagenize *morc-1(-);pkIS32[pie-1p::gfp::h2b]* worms and evaluate the resulting mutant lines for suppression of *morc-1(-)* defective RNAi inheritance (Figure 3.2). In brief, we will mutagenize L4 worms with ENU (P0 generation), separate them into 50 pools, and then allow them to lay embryos for 24

hours (F1 generation). We will allow the F1 worms to self-fertilize in order to homozygose the mutations from the ENU treatment. We will then isolate F2 embryos by hypochlorite treatment of F1 gravid worms. We will grow the F2 generation on *gfp* RNAi. We will then score the F2 adult worms for *gfp* expression. Because we are interested in mutations that affect RNAi inheritance specifically, rather than general RNAi sensitivity, we will single F2 adults that do not express GFP, and therefore have intact RNAi sensitivity, to new plates with standard OP50 food. We will single 20 F2 adults from each of the 50 pools. The F2 worms will self-fertilize to generate the F3 generation which will be scored as adults for GFP expression. We will select plates on which fewer than 50% of the F3 worms express GFP, which would be indicative of a putative *morc-1(-)* suppressing mutation. We will perform whole genome sequencing of these putative suppressor lines and analyze the data as described in Chapter Two to identify genes with non-synonymous SNPs or nonsense mutations in multiple suppressor lines.

If whole genome sequencing generates a large number of putative suppressors, we will prioritize candidate genes with annotated chromatin-binding or chromatin-modifying domains. We will validate candidates by obtaining deletion mutants when available and testing whether these alleles suppress the *morc-1(-)* RNAi inheritance defect. We will also test these alleles for suppression of *hrde-1(-)* defective RNAi inheritance. For candidates with no available deletion alleles, we will backcross the alleles generated from the screen, cross them back into the *morc-1(-);pkIs32[pie-1p::gfp::h2b]* background, and test for rescue of defective RNAi inheritance. These validation methods will yield a suite of bona fide suppressors of the *morc-1(-)* RNAi inheritance defect.

We will test the suppressors for rescue of other phenotypes that are characteristic of nuclear RNAi mutants including suppression of *morc-1(-)* and *nrde-2(-)* resistance to *lir-1* RNAi and *morc-1(-)* and *hrde-1(-)* Mrt. Finally, we will perform H3K9me3 ChIP-qPCR of previously identified *morc-1*-dependent H3K9me3 regions in wildtype, *morc-1(-)*, *smorc(-)*, and *smorc(-);morc-1(-)* mutants to test whether target H3K9 methylation is restored in *smorc(-);morc-1(-)* mutants.

In our previous screen for mutations that suppress *morc-1(-)* germline mortality, we identified multiple alleles of *fxba-164* and *srj-24* that were heterozygous. Although no

deletion alleles are available for either gene, we tested RNAi knock-down of both genes for suppression of *morc-1(-)* germline mortality. We found that knockdown of either *fbxa-164* or *srj-24* partially restores *morc-1(-)* fertility at the F7 generation at 25°C (Figure 3.3). Both genes are uncharacterized; based on amino acid sequence, *fbxa-164* encodes one of over 300 F box proteins in *C. elegans*. *srj-24* encodes a G-protein coupled receptor (GPCR). Further study, including genetic validation, will be required to determine if either *fbxa-164* or *srj-24* are true *morc-1(-)* suppressors.

These studies will identify novel regulators of nuclear RNAi. We expect to characterize a number of chromatin-modifying or chromatin-binding factors that will further elucidate the mechanisms by which nuclear RNAi establishes and maintains heterochromatin at target genes.

morc-1(-)
P0

morc-1(-)
F1

morc-1(-)
F2
+
gfp RNAi

morc-1(-)
F3

morc-1(-);suppressor(-)

RNAi inheritance
defective

whole genome sequencing
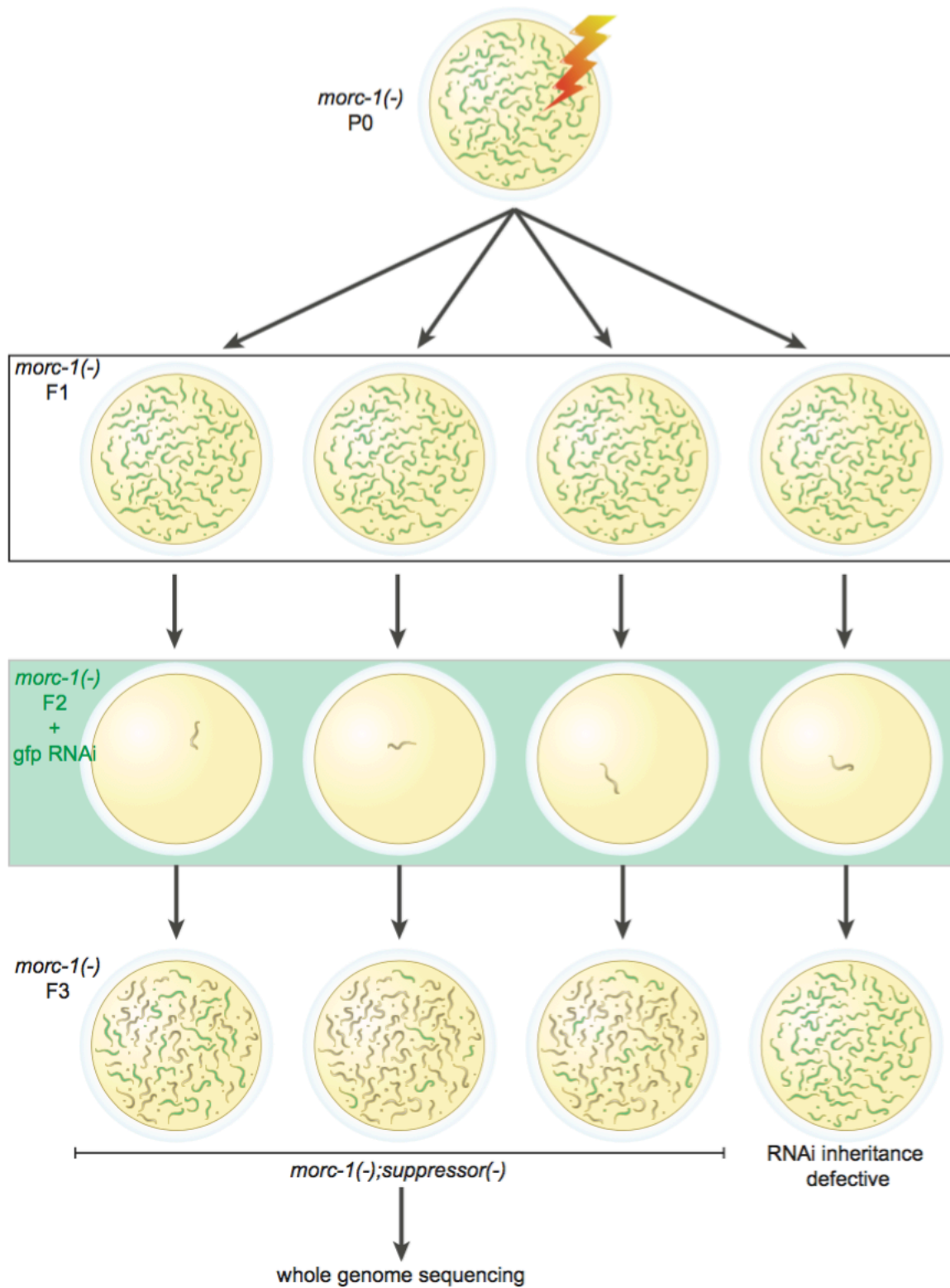
Figure 3.2. Overview of modified *morc-1(-)* suppressor (*smorc*) screen. *morc-1(-)* mutant worms expressing a multicopy *pie1p::h2b::gfp* transgene are ENU mutagenized. Their progeny (F1) self-fertilize to homozygose mutations. F2 worms are plated as L1 larva onto *gfp* RNAi. F2 adult worms that do not express *gfp* are singled to new plates. F2 worms lay eggs on OP50 food and their progeny (F3) are scored for *gfp* expression. F3 plates on which fewer than half of the worms express *gfp* are collected for genomic DNA extraction and whole-genome sequencing to identify putative suppressors of *morc-1(-)* RNAi inheritance defect.

Figure 3.3. RNAi against *fbxa-164* and *srj-24* suppresses the *morc-1(-)* fertility defect. *morc-1(-)* mutants were grown on vector RNAi (EV), or RNAi against *met-1, fbxa-164,* or *srj-24* for eight generations at 25°C. In the eighth generation (F7), worms were singled and their total progeny were counted. Circles represent total progeny from a single hermaphrodite. Lines indicate mean +/- standard deviation.

MORC-1 IS REQUIRED FOR GENOME STABILITY

Previous studies have established a critical role for H3K9 methylation in maintaining genome stability in worms, flies, and mammals (Peters et al. 2001; Peng and Karpen 2006; Zeller et al. 2016). In worms, complete loss of H3K9 methylation in the *met-2(-);set-25(-)* double mutant leads to only modest gene upregulation (fewer than 10% of H3K9me-marked genes) but more substantial upregulation of H3K9me-marked repetitive elements throughout the genome, even in the absence of a nearby gene (Zeller et al. 2016). In the absence of H3K9 methylation, the ectopic transcription of repetitive elements is associated with the formation of RNA:DNA hybrids (R loops) (Zeller et al. 2016). The R loops cause the replication fork to stall, leaving the DNA susceptible to damage via double-stranded breaks (DSBs) (Zeller et al. 2016). Consequently, *met-2(-);set-25*(-) mutants acquire DNA damage, in the form of small insertions and deletions (indels), at the loci that encode the upregulated repetitive elements (Zeller et al. 2016). The accumulation of genomic damage in *met-2(-);set-25(-)* mutants shows that global H3K9 methylation is essential for maintaining genome fidelity.

Intriguingly, in the studies described in Chapter Two, we found evidence of genome instability in *morc-1(-)* mutants. In our genomic DNA sequencing, we observed that *morc-1(-)* mutants grown for many generations at 25°C accumulated hundreds of SNPs compared to *morc-1(-)* mutants that were propagated at 20°C. In our ChIP-seq studies, we identified a 200kb span of chromosome V that had twice as much read coverage as neighboring regions in the input samples in *morc-1(-)* samples but not in *hrde-1(-)* or wildtype libraries (Figure 3.4). Although we were not able to map the duplication due to the short read-length from these libraries, we concluded that at some point during its propagation, the *morc-1(-)* mutant strain acquired a duplication of this part of chromosome V. Interestingly, this duplication most likely occurred at 20°C, as it was present in both F1 and F4 libraries, suggesting that even at lower temperatures, *morc-1(-)* mutants may still accumulate DNA damage. Taken together, these findings suggest that *morc-1* is required for genome stability. Here, we propose a series of studies to investigate (1) the precise role of MORC-1 in maintaining genomic fidelity and

(2) what characteristics make genomic loci vulnerable to DNA damage in *morc-1(-)* mutants. We will investigate several potential, non-mutually exclusive explanations for increased DNA damage in *morc-1(-)* mutants to identify the specific defects that lead to genomic instability in *morc-1(-)* mutants and the sites that are most susceptible to damage.

In mammals, MORC proteins make two important contributions to genome stability: transposon repression and activation of p53 and p21 in response to DNA damage (Takahashi et al. 2007; Pastor et al. 2014). Based on the function of mammalian MORC proteins in p53 activation and recruitment, we hypothesize that DNA damage repair mechanisms are impaired in *morc-1(-)* mutants. First, we will ask whether *morc-1(-)* mutants are hypersensitive to genotoxic stress. In order to narrow down the defect leading to genome instability in *morc-1(-)* mutants, we will test sensitivity to different genotoxic stressors that trigger distinct repair pathways. We will expose wildtype and *morc-1(-)* larva to ionizing radiation (IR) to evaluate DSB repair, UV radiation to evaluate nucleotide excision repair, and to hydroxyurea to evaluate sensitivity to replication stress. We will also include *hrde-1(-)* and *met-2(-);set-25(-)* mutants as controls. To date, no evidence of genome instability in *hrde-1(-)* mutants has been reported. Given the shared phenotypic defects of *morc-1(-)* and *hrde-1(-)* and our finding that, at the level of chromatin modification, *morc-1* regulates almost exclusively a subset of *hrde-1* targets, we expect that *hrde-1(-)* mutants will show similar defects to *morc-1(-)* mutants. Recently, *met-2(-);set-25(-)* mutants were found to be hypersensitive to hydroxyurea but not UV stress, consistent with increased replication stress due to R loop formation (Zeller et al. 2016). If *morc-1(-)* mutants show a similar phenotype, it would suggest that the *morc-1(-)* genome instability is the result of H3K9me3 loss and may be due to derepression of repetitive elements and resultant R loops. If *morc-1(-)* mutants are hypersensitive to UV irradiation, this would suggest a defect in nucleotide excision repair and thus a distinct mechanism of genome instability from that caused by the global loss of H3K9me3 in *met-2(-);set-25(-)* mutants. DSBs occur in response to a number of genotoxic insults, including IR, exposure to chemical mutagens, oxidative damage, and replication stress (O'Neil and Rose 2006). Notably, the *C. elegans* germline chromosomes normally undergo some degree of DSBs to promote meiotic

recombination, therefore not all DSBs are pathological (O'Neil and Rose 2006). We will test the efficacy of DSB repair pathways by evaluating sensitivity of *morc-1(-)* and wildtype worms to IR.

The consequences of a defect in DNA damage repair pathways would be enhanced by increased occurrence of DNA lesions such as DSBs. The severely decompacted chromatin we observe in *morc-1(-)* germline nuclei may be more vulnerable to damage and DSBs. To evaluate the occurrence of DSBs in *morc-1(-)* mutants compared to wildtype worms, we will perform immunostaining against RAD-51 with and without IR. If *morc-1(-)* mutants exhibit a higher level of DSBs at baseline, we will observe more RAD-51 foci in *morc-1(-)* germlines compared to wildtype. If *morc-1(-)* mutants fail to recruit RAD-51 to DSBs, we will observe fewer RAD-51 in *morc-1(-)* mutants after exposure to IR compared to wildtype.

Another mechanism by which *morc-1(-)* mutants might accumulate DNA damage is by defects in germline apoptosis, thus failing to eliminated damaged genomes from the developing germline. In *C. elegans*, some degree of physiological apoptosis occurs during oogenesis (Gumienny et al. 1999). Physiological apoptosis is not thought to eliminate damaged germ cells, but rather to recycle cellular contents to facilitate the dramatic increase in cell size following pachytene exit (Gartner et al. 2008). In response to irreparable DNA damage, non-physiological germ cell apoptosis can cull damaged nuclei from the developing germline (Gartner et al. 2000). Inducible germ cell apoptosis requires the same cellular machinery as physiological apoptosis, including the transmembrane receptor CED-1 (Zhou et al. 2001). Using a well-established CED-1::GFP fusion protein as a reporter for germline apoptosis, we will evaluate apoptosis levels in wildtype, *morc-1(-), hrde-1(-)*, and *met-2(-);set-25(-)* mutants with and without UV irradiation. *met-2(-);set-25(-)* have previously been shown to exhibit elevated levels of germ cell apoptosis in the absence of genotoxic stress (Zeller et al. 2016). Increased CED-1::GFP expression in *morc-1(-)* mutants would be additional evidence in support of elevated DNA damage in these mutants. If CED-1::GFP levels in *morc-1(-)* are elevated or equivalent to wildtype following UV irradiation, the DNA damage checkpoint is likely intact and inducing apoptosis to cull damaged nuclei from the germline. CED-1::GFP levels that are lower in *morc-1(-)* than in wildtype following UV irradiation would suggest

that the checkpoint is impaired and that damaged nuclei are proceeding through oogenesis and may pass on accumulated genomic damage to their progeny.

Finally, while our previously described studies have provided preliminary evidence for genome instability in *morc-1(-)* mutants, we have not yet characterized the precise types of DNA damage or the susceptible locations in *morc-1(-)* mutants. Characterizing the most prevalent types of DNA damage that occur in *morc-1(-)* mutants will provide insight into how *morc-1* protects genome fidelity. By comparing the sites of DNA damage in *morc-1(-)* mutants to our previous identified sites of gene upregulation, H3K9me3 depletion, and H3K36me3 enrichment, we will determine whether *morc-1* protects endo-siRNA target sites from accumulating DNA damage.

Based on the other progressive defects that we observe in *morc-1(-)* mutants (i.e. germline mortality, endo-siRNA target gene upregulation, loss of H3K9me3, H3K36me3 enrichment), we expect that the accumulation of DNA damage will also be progressive. To characterize progressive damage, we will compare whole genome sequencing data from worms grown at 25°C for a single generation to worms grown at 25°C for 10 generations. To ensure that there is no pre-existing genetic variation in the starting population, we will propagate both wildtype and *morc-1(-)* mutants starting from a single worm at 20°C (P-1 generation). When the P-1 worms are adults, we will perform an overnight egg lay at 25°C and then remove the adult worm, leaving behind only embryos (P0 generation). When the P0 worms reach the second larval stage, we will single 16 P0 worms to new plates. These will be propagated as distinct lines for 10 generations (to F9 generation). The remaining P0 worms will be hypochlorite-treated as adults to yield synchronized F1 embryos. When the F1 embryos reach adulthood they will be collected for genomic DNA extraction as early generation *morc-1(-)* mutants and wildtype controls. When the F9 progeny of the 16 originally singled lines reach adulthood, they will be collected for genomic DNA extraction. For each genotype, this will effectively generate 16 replicates of late-generation genomic DNA to compare to the initial F1 DNA. We will perform high throughput sequencing to identify single nucleotide polymorphisms (SNPs) and small insertions and deletions (indels). We will compare the sites of DNA damage to our previously identified *morc-1*-dependent H3K9me3 and H3K36me3 targets to determine whether sites that are epigenetically misregulated in

150

morc-1(-) mutants are prone to DNA damage. Additionally, we will look for evidence of copy-number variations (CNVs) in our sequencing data. Based on the 200kb duplication that we have already identified, we expect to see evidence of large insertions and deletions or chromosomal rearrangements. Any samples that show evidence of this type of sequence variation will be resequenced using a MinION sequencing platform (Oxford Nanopore), a novel sequencing technology that can generate reads that are tens of kilobases long, enabling easy detection of large chromosomal rearrangements (Bolisetty et al. 2015; Goodwin et al. 2015). One potential problem with MinION sequencing is that it generates lower read coverage than other sequencing approaches. We expect that the progressive acquisition of DNA lesions in *morc-1(-)* mutants will generate genetically heterogeneous populations in the late generation samples. Identifying rare variants within these heterogeneous populations may not be possible with the low read coverage of the MinION platform. Sequencing via the Pacific Biosciences (PacBio) platform is a higher capacity method for sequencing very long reads (an average read-length of 10kb) that we can use as an alternative approach (Rhoads and Au 2015).

The transmission of accurate genomic information to the next generation is essential for species survival. There are many mechanisms that collaborate to maintain genome fidelity from one generation to the next. Our preliminary data indicate that MORC-1 is essential for genome stability. The studies proposed here will explore the mechanism by which MORC-1 contributes to genome stability and evaluate what kinds of DNA damage result from loss of MORC-1.
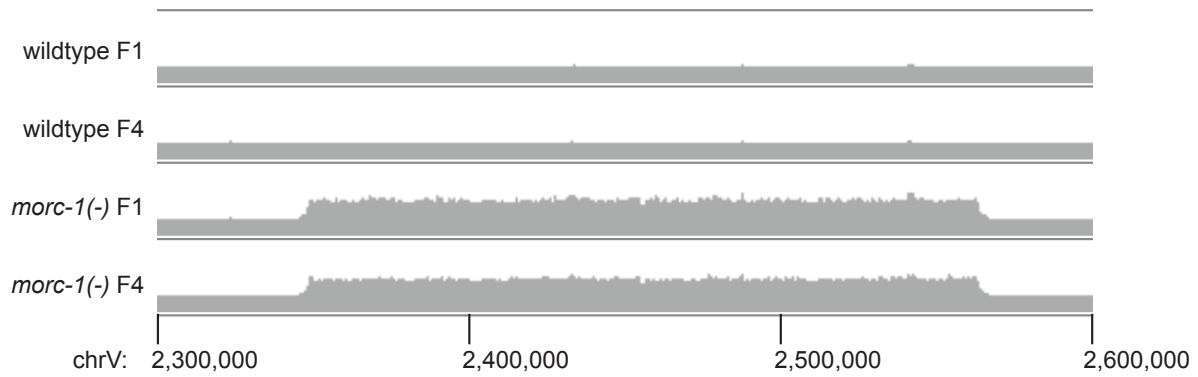
Figure 3.4. A 200kb duplication of chromosome V indicates that *morc-1(-)* mutants have unstable genomes. RPKM across the indicated coordinates on chromosome V in early and late generation wildtype and *morc-1(-)* worms. In *morc-1(-)* mutants, this region of chromosome V has ~2-fold greater read coverage than surrounding regions on chromosome V or the same position in wildtype worms.

CHARACTERIZATION OF MORC-1 FUNCTION IN THE CSR-1 22G ENDO-siRNA PATHWAY

We have described a role for MORC-1 in the regulation of a subset of targets of the WAGO 22G endo-siRNA pathway. Our findings are consistent with the previously described roles of other members of the Morc family in repression of transgenes and transposable elements. Unexpectedly, we found that MORC-1 also regulates targets of the CSR-1 22G endo-siRNA pathway, which is thought to license germline-expressed genes and promote sense transcription (Claycomb et al. 2009; Seth et al. 2013; Wedeles et al. 2013; Cecere et al. 2014). Our mRNA-seq results revealed that targets of the CSR-1 22G endo-siRNA pathway were significantly enriched among genes that were upregulated in late generation *morc-1(-)* mutants compared to wildtype (Figure 3.5, p=9.6x10^{-14}, Fisher's test). There was no enrichment for CSR-1 target genes among early generation upregulated genes or downregulated genes at either generation (Figure 3.5). Intriguingly, *hrde-1*-regulated mRNAs were not enriched for CSR-1 target genes at either generation, suggesting that the role of MORC-1 on the CSR-1 pathway is *independent* of HRDE-1 (Figures 3.5 and 3.6).

To determine the step at which MORC-1 functions in the CSR-1 pathway, we asked whether levels of CSR-1-class 22G RNAs are misregulated in *morc-1(-)* mutants. Using the 5'-independent small RNA-seq libraries described in Chapter Two, we quantified the levels of CSR-1-class 22G RNAs in wildtype, *morc-1(-),* and *hrde-1(-)* mutants. We found that CSR-1-class 22G RNA levels were not affected in *morc-1(-)* mutants at either early or late generation (Figure 3.7), suggesting that MORC-1-mediated regulation of CSR-1 targets occurs downstream of 22G RNA biogenesis and loading onto CSR-1 RISC. Based on the role we established for MORC-1 in regulation of chromatin organization in Chapter Two, MORC-1-mediated repression of CSR-1 target genes is likely occurring at the chromatin level.

As both the HRDE-1 and the CSR 22G endo-siRNA pathways converge on chromatin regulation, these data suggest an intriguing model in which, through its effects on chromatin structure, MORC-1 impinges upon both pathways. Although there are no reports to date of coordinated regulation of chromatin structure between HRDE-1

and CSR-1, a number of recent studies suggest that there is more crosstalk between the HRDE-1/WAGO and CSR-1 pathways than previously appreciated. At the level of Argonaute loading, piRNAs direct the sorting of 22G RNAs into CSR-1 or WAGO RISCs (de Albuquerque et al. 2015; Phillips et al. 2015). Without the epigenetic memory of "self" and "non-self" provided by the piRNAs, CSR-1-22G endo-siRNAs are misloaded onto HRDE-1 RISC and vice versa, indicating that the sorting of 22G RNAs into HRDE-1 versus CSR-1 RISCs is actively regulated (de Albuquerque et al. 2015; Phillips et al. 2015). At the level of post-transcriptional regulation of gene expression, CSR-1-bound mRNAs are protected from silencing by RNAe, which is initiated by piRNAs and maintained transgenerationally by HRDE-1 22G RNAs (Shirayama et al. 2012; Seth et al. 2013; Wedeles et al. 2013). Taken together, these studies show that piRNAs, HRDE-1 22G RNAs, and CSR-1 22G RNAs comprise an integrated, regulatory network network that orchestrates proper gene expression in the developing germline and embryo.

Hypomorphic *csr-1* mutants exhibit increased transcription of normally silent transcripts that are adjacent to active genes, corresponding to a depletion of the repressive H3K27me3 mark and the centromeric H3 variant CENP-A (Cecere et al. 2014). Furthermore, in *csr-1(-)* meiotic chromosomes, immunostaining shows broadly distributed H3K9me2 staining over the chromosomes and overlapping signal from H3K4me2 and H3K9me2, which are normally mutually exclusive (She et al. 2009). These data support a role for CSR-1 in maintenance of heterochromatin-euchromatin boundaries. We propose that MORC-1 and CSR-1 collaborate to protect the integrity of a shared set of heterochromatin-euchromatin boundaries.

In Chapter Two, we presented data suggesting that MORC-1 prevents the spread of euchromatic marks into silent loci. We hypothesize that CSR-1 may contribute to the chromatin alterations in *morc-1(-)* mutants by aberrantly licensing *morc-1* target genes for transcriptional activation and euchromatinization. One possible mechanism for this is that CSR-1 is acting upstream of MET-1 to promote euchromatic encroachment into silent loci. We propose the following studies to test this model.

Does *morc-1(-)* phenocopy the other *csr-1(-)* phenotypes?

To test the model that MORC-1 functions downstream of CSR-1 RISC, we will evaluate *morc-1(-)* mutants for defects in P granule formation, chromosome segregation, and trans-activation. The initial steps of the CSR-1 pathway occur at P granules, where many pathway components are localized (e.g. CSR-1, EGO-1, DRH-3, and CDE-1) (Claycomb et al. 2009; van Wolfswinkel et al. 2009). Additionally, CSR-1 is required for P granule stability (Updike and Strome 2009). The 22G RNAs are loaded onto CSR-1 RISC at the P granules in the distal germline (Claycomb et al. 2009). As the germ cells differentiate into oocytes, CSR-1 translocates into the nucleus to license nascent transcripts for expression and to simultaneously repress inactive genes by enforcing CENP-A and CENP-C distribution at silent loci (Claycomb et al. 2009; Updike and Strome 2009; Cecere et al. 2014). Based on our model that MORC-1 functions downstream of CSR-1 at the chromatin level, we hypothesize that *morc-1(-)* mutants will not exhibit the loss of P granules that is characteristic of depletion of CSR-1-class 22G RNAs. We will test this using well-characterized P granule reporters such as PGL-1::GFP and PGL-3::GFP and evaluate germline expression levels in wildtype and *morc-1(-)* backgrounds compared to *csr-1* RNAi control.

Partial loss of *csr-1* or *drh-3* activity leads to a high incidence of males (Him) and embryonic lethality (Emb), which are phenotypes associated with impaired segregation of the X chromosome (causing Him) or of autosomes (causing Emb) (Claycomb et al. 2009). The most striking defects appear in early embryos, where centromeric H3 variants CENP-A/HPC-3 and CENP-C/HPC-4 are loaded onto the chromosomes but not arranged correctly (Claycomb et al. 2009). Similarly, condensins (CAPG-1 and KLE-2) and cohesins (SCC-1 and SCC-3) show disorganized loading on mitotic chromosomes of worms fed *csr-1* or *drh-3* RNAi (Claycomb et al. 2009).

We expect that MORC-1 may be among the chromatin factors that are disrupted by *csr*-1 RNAi. We will use immunofluorescence against MORC-1::3xFlag to evaluate the effects of *csr-1* and *drh-3* RNAi on MORC-1 localization in oocytes and early embryos. Disrupted MORC-1 localization would indicate either that (1) MORC-1 functions downstream of CSR-1 and is recruited to chromatin by CSR-1 or a CSR-1-

regulated factors such as CENP-A or CENP-C or that (2) MORC-1 requires the correct demarcation of heterochromatin in the genome and thus the disorganized chromatin caused by *csr-1* RNAi is be sufficient to disrupt MORC-1 binding.

We hypothesize that *morc-1(-)* mutation will phenocopy *csr*-1 RNAi, leading to disorganization of CENP-A, CENP-C, and KLE-2. Using transgenic constructs expressing tubulin::GFP and mCherry-tagged CENP-A, CENP-C, and KLE-2, we will evaluate the effects of *morc-1(-)* on the loading and organization of these components on mitotic chromosomes. As controls we will use wildtype worms grown on empty vector *csr-1,* and *drh-3* RNAi. Effects of MORC-1 on KLE-2 distribution could be due to direct regulation of condensin activity or simply a by-product of insufficiently condensed chromatin.

If MORC-1 functions in parallel to CSR-1 to regulate chromatin structure in similar ways, we expect that *morc-1(-)* mutants will phenocopy *csr-1* RNAi but *csr-1* RNAi will not affect MORC-1 localization. To confirm that MORC-1 and CSR-1 function in parallel, we will test whether *csr-1* knockdown by RNAi enhances the *morc-1(-)* mitotic defects (e.g. disorganized CENP-A, CENP-B, and KLE-2 organization). If the chromatin disorganization caused by deletion of *morc-1(-)* and depletion of *csr-1* by RNAi is in fact additive, *morc-1(-)* worms grown on *csr-1* RNAi may not be viable. We will address this possibility by (1) feeding *morc-1(-)* mutants dilute *csr-1* RNAi to titrate CSR-1 activity and (2) performing these experiments in worms expressing an ATP-hydrolysis dead *morc-1* variant (*morc-1(E39A)::3xFlag*). Worms expressing this hypomorphic *morc-1* allele may be more tolerant of *csr-1* depletion than *morc-1(-)* deletion mutants.

Alternatively, CSR-1 and MORC-1 may function in parallel pathways to regulate the same target genes through distinct effects on chromatin structure and organization, in which case *morc-1(-)* mutants would not be expected to phenocopy the effects of *csr-1* RNAi on CENP-A, CENP-C, or KLE-2 organization.

Does MORC-1 antagonize CSR-mediated transactivation?

As described in Chapter One, CSR-1-mediated RNAa antagonizes piRNA-initiated, HRDE-1-maintained RNAe (Shirayama et al. 2012; Seth et al. 2013; Wedeles et al. 2013). While RNAa and RNAe clearly act in opposition to eachother, the mechanism by which RNAa can reverse RNAe and vice versa are unknown. As both pathways direct chromatin modifications at target loci, the shifting of heterochromatin-euchromatin boundaries may contribute to the interplay between RNAa and RNAe. Based on our finding that MORC-1 represses CSR-1 target genes, we predict that MORC-1 also antagonizes CSR-1-mediated gene licensing during RNAa.

First, we will establish whether MORC-1, like HRDE-1, is required for RNAe (Figure 3.8). The basic framework to test RNAe involves crossing a silent, single copy *gfp::cdk-1(RNAe)* transgene into a strain that is actively expressing an identical single copy transgene, *gfp::cdk-1(+).* For clarity I will refer to the silent transgene as *gfp(RNAe)* and the active transgene as *gfp(+)*. First, we will test whether *morc-1* is required for establishment of RNAe. In the P0 generation, we will cross *morc-1(+/-);gfp(RNAe)* into *morc-1(+/-);gfp(+)*. The resulting *morc-1(+/+)*, *morc-1(+/-),* and *morc-1(-/-)* F1 worms will all be hemizygous for both *gfp(RNAe)* and *gfp(+).* Because RNAe is fully penetrant, any GFP expression in the F1 generation is indicative of a defect in RNAe establishment. Thus, if *morc-1* is required to establish RNAe, the *morc-1(-/-)* F1 worms will express GFP. We predict that, like *hrde*-1, *morc-1* is dispensable for establishment of RNAe but required for transgenerational maintenance of RNAe; therefore 100% of the F1 progeny should not express *gfp*, indicating that the *gfp(+)* transgene has been silenced by RNAe.

To test the contribution of *morc*-1 to maintenance of RNAe, we will allow the F1 *morc-1(+/-)* heterozygotes to self-fertilize and score GFP expression in their *morc-1(-/-)* homozygous progeny (F2 generation) compared to their *morc-1(+/+)* siblings. If *morc-1* is required for maintenance of RNAe, we will observe GFP expression in the *morc-1(-/-)* homozygotes but not in their *morc-1(+/+)* siblings (Figure 3.8). As controls for this experiment, we will use *hrde-1(+/-)* and *prg-1(+/-)* mutants in the P0 generation (Ashe et al. 2012; Luteijn et al. 2012; Shirayama et al. 2012).

Second, we will ask if *csr-1* potentiates the contribution of *morc-1* and *hrde-1* to RNAe by performing the experiment described above with or without *csr-1* RNAi (Figure 3.8). We predict that the activation of the previously silent transgene in *morc-1(-/-)* or *hrde-1(-/-)* F2 worms is at least partially *csr-1*-dependent. Accordingly, the degree of GFP activation in *morc-1(-/-)* and *hrde-1(-/-)* F2 worms grown on *csr-1* RNAi will be decreased compared to worms grown on empty vector, indicating that RNAe has been restored.

Third, we will test whether MORC-1 antagonizes CSR-1-directed gene licensing. Previous studies have found that tethering CSR-1 to a silent transcript is sufficient to overcome piRNA-mediated silencing and reactivate the target gene via RNAa (Seth et al. 2013; Wedeles et al. 2013). RNAa is discussed in greater detail in Chapter One. RNAa can be evaluated by fusing *gfp* to phage λ BoxB hairpins (*gfp::boxb*) and CSR-1 to the phage λN anti-termination protein fragment (CSR-1::λN). For simplicity, I will refer to this strategy as CSR-1 tethering. Importantly, the *gfp::boxb* transcript is susceptible to silencing by *gfp::cdk(RNAe)* (Wedeles et al. 2013). To confirm that silencing of *gfp::boxb(RNAe)* is *morc-1*-dependent RNAe, we will initiate RNAe of *gfp::boxb* strain crossing *gfp::boxb* with *gfp(RNAe)*. I will refer to the silencing *gfp::boxb* allele as *gfp::boxb(RNAe)*. We will cross *gfp::boxb(RNAe)* into *morc-1(-/-)* mutants to generate F1 worms that are *morc-1(+/-)* heterozygotes and hemizygous for *gfp::box (RNAe)*. The F1s will self-fertilize and we will score GFP expression in their *gfp::boxb (RNAe)(+/+); morc-1(-/-)* progeny (F2 generation). Higher GFP expression in *morc-1(-/-)* F2 worms compared to their *morc-1(+/-)* or *morc-1(+/+)* siblings will confirm that maintenance of *gfp::boxb(RNAe)* is *morc-1*-dependent (Figure 3.9A).

To test the contribution of *morc-1* to CSR-1-mediated RNAa, we will compare how many generations of CSR-1 tethering are required to induce RNAa in a *morc-1*(-/-) background compared *to morc-1*(+/+). We will tether CSR-1 *gfp::boxb(RNAe)* for many generations in a *morc-1(+/-)* heterozygous background. At each generation, the *morc-1(+/-)* worms will self-fertilize. We will monitor GFP expression in the *morc-1(-/-)* and *morc-1(+/+)* progeny to determine when RNAa is established (Figure 3.9B). Based on our model that *morc-1(-)* antagonizes CSR-1-mediated gene licensing, we hypothesize that activation of the *gfp::boxb(RNAe)* allele will occur at an earlier generation in the

morc-1(-/-) background than in the *morc-1(+/+)* background. Based on published reports, we expect that activation in the *morc-1(+/+)* background will occur at approximately the F4 generation (Wedeles et al. 2013). These studies will allow us to take a targeted approach to dissect the relative contributions of *morc-1* and *csr-1* to the expression of a specific transcript.

Do MORC-1 and CSR-1 regulate an overlapping set of chromatin marks at the same targets?

Next, we will investigate the overlap of MORC-1 and CSR-1 functions at the chromatin level. To evaluate whether MORC-1 and CSR-1 regulate chromatin at a shared set of endogenous targets, we will perform ChIP-seq to elucidate genome-wide patterns of H3K9me3, H3K36me3, H3K4me3, H3K27me3, and CENP-A in early and late generation wildtype and *morc-1(-)* worms with or without *csr-1* RNAi. We expect that RNAi against *csr-1* will partially rescue the effects of *morc-1(-)* at heterochromatin-euchromatin boundaries, and restore heterochromatic marks (H3K9me3, H3K27me3, and/or CENP-A) and repress ectopic euchromatic (H3K36me3 and H3K4me3) marks. One caveat to this experiment is that loss of *morc-1* or *csr-1* independently compromises fertility, thus it may not be possible to grow enough *morc-1(-)* worms on *csr-1* RNAi for more than one generation to perform ChIP experiments. To address the issue of additive fertility defects, we will profile the same chromatin marks by ChIP-seq separately in wildtype and *morc-1(-)* worms grown on *csr-1* RNAi to identify *csr-1*-regulated and *morc-1*-regulated loci. Significant overlap of *csr-1*-regulated and *morc-1*-regulated loci will indicate that CSR-1 and MORC-1 do, in fact, regulate chromatin structure at a shared set of targets.

Does CSR-1 function upstream of MET-1?

Intriguingly, CSR-1 tethering causes gain of H3K4 methylation and depletion of H3K9 methylation at the target locus (Wedeles et al. 2013). This is reminiscent of our finding that *morc-1(-)* mutants exhibit H3K36me3 gain and loss of H3K9me3 at target loci. We hypothesize that CSR-1 licensing will drive the activity of MET-1 at MORC-1-dependent sites. A previous study found high correlation between CSR-1 22G targets and H3K36me3 enrichment but no depletion of H3K36me3 in a *csr-1* hypomorphic mutant (Cecere et al. 2014). However, H3K36 trimethylation in the parental germline is re-established, or "remembered" in the progeny by the H3K36 histone methyltransferase MES-4 (Rechtsteiner et al. 2010). CSR-1 might contribute to the initial H3K36me3 deposition at a given target and be completely dispensable for transgenerational maintenance of the H3K36me3 mark. Investigation of the role of CSR-1 in *de novo* establishment of H3K36me3 genome wide would require the erasure of the H3K36me3 from prior generations by mutation of *mes-4*. However, as *mes-4(-)* mutants are maternal-effect sterile, growing sufficient samples for ChIP-seq in a total loss-of-function background is not feasible. To address the role of CSR-1 in directing *de novo* H3K36 methylation, we must look at acquisition of new H3K36me3 marks without erasing all H3K36me3 in the genome. Fortunately, the ectopic gain of H3K36me3 in late generation *morc-1(-)* mutants allows us to examine a role for CSR-1 in *de novo* H3K36 trimethylation in a genetic background that is healthy enough to allow for large-scale sample collection.

We will test whether CSR-1 functions upstream of MET-1 using both targeted and genome-wide approaches. First, we will test whether RNAa is MET-1-dependent by tethering CSR-1 to the *gfp::boxb(RNAe)* reporter in wildtype, *met-1* hypomorph, and *met-1* null backgrounds (Figure 3.10) and comparing (1) the number of generations required to establish RNAa in wildtype versus *met-1* mutant backgrounds, (2) the level of GFP expression in each genetic background when RNAa is initiated, and (3) enrichment of H3K36me3 of the *gfp* locus by ChIP-qPCR. We expect that *met-1* hypomorph and null backgrounds, establishment of RNAa (i.e. GFP expression) will be repressed compared to wildtype and this will be reflected in H3K36me3 ChIP-qPCR by

a loss of H3K36me3 at this locus in the mutant background. This would demonstrate that *met-1* is required to establish CSR-1-mediated gene licensing.

We will next test whether *met-1* contributes to maintenance of RNAa once the CSR-1-tether has been removed. We will initiate RNAa by tethering CSR-1 to *gfp::boxb(RNAe)* for 10 generations (Seth et al. 2013; Wedeles et al. 2013). We will then cross the activated *gfp::boxb* allele into either wildtype, *met-1* hypomorph, or *met-1* null mutants and allow these worms to self-fertilize in the absence of the *csr-1::λN* transgene. We will continue to propagate these strains through self-fertilization and score GFP expression at every generation (Figure 3.11). We expect that GFP will be silenced at earlier generations in the *met-1* null or hypomorphic backgrounds compared to wildtype. This would suggest that *met-1* contributes to maintenance of gene licensing downstream of CSR-1.

While the phage λ-BobB system allows us to dissect the genetics of gene licensing, this transgenic system will not necessarily reflect the regulation of endogenous CSR-1-targets. To address the contribution of MET-1 to CSR-1-mediated chromatin regulation, we will perform ChIP-seq for H3K36me3, H3K4me3, H3K27me3, and CENP-A in wildtype, *morc-1(-), met-1(xk4),* and *met-1(xk4);morc-1(-)* mutants fed empty vector or RNAi against *csr-1*. First, we will evaluate whether CSR-1 and MET-1 regulate the same loci in a *morc-1(WT)* background. We will compare the locations of CSR-1-regulated targets to those of MET-1-regulated targets to see if CSR-1 and MET-1 normally regulate the same targets. Second, our analysis will focus on the regions that we identified in Chapter Two as *met-1*-dependent. We will use these 424 loci to interrogate whether (1) CSR-1 contributes to H3K36 methylation at *met-1*-dependent sites (2) whether the ectopic *met-1*-dependent increase in H3K36me3 at these loci is also *csr-1*-dependent.

Does loss of MORC-1-dependent epigenetic memory lead to reprogramming of CSR-1-associated endo-siRNAs?

Our previously described studies elucidated a role for MORC-1 and HRDE-1 in repression of H3K36 trimethylation. In the absence of *morc-1* or *hrde-1,* H3K36me3

gain is progressive from F1 to F4 and partially MET-1-dependent. We hypothesize that this kind of transgenerational epigenetic reprogramming may also alter how genome defense mechanisms distinguish "self" from "non-self" elements. For example, the upregulation of *hrde-1*-targets genes without an associated loss of the 22G endo-siRNAs in a *morc-1(-)* background could result in misloading of these 22G endo-siRNAs onto CSR-1 RISC and, as a result, reinforcement of derepression via CSR-mediated gene licensing. In fact, CSR-targeting of derepressed genes could also promote their H3K36 methylation by MET-1, thus leading to a self-reinforcing loop of derepression, licensing by CSR-1, H3K36 trimethylation by MET-1, and further derepression. Because CSR-1 targets germline-expressed genes to promote their expression (Claycomb et al. 2009; Conine et al. 2013; Seth et al. 2013; Wedeles et al. 2013), we speculate that the progressive gain of *met-1*-dependent H3K36me3 in *morc-1(-)* mutants may be sufficient to rewrite epigenetic memory by aberrant loading of HRDE-1-associated 22G endo-siRNAs onto CSR-1 RISC to promote a memory of gene expression.

The recent finding that resetting the 22G endo-siRNAs in the absence of piRNAs leads to crossloading of HRDE-1- and CSR-1 22G endo-siRNAs suggests that the distinction between 22G endo-siRNAs that are loaded on to HRDE-1 RISC or CSR-1 RISC is not entirely intrinsic to the endo-siRNAs (de Albuquerque et al. 2015; Phillips et al. 2015). We propose a similar strategy to investigate a role for MORC-1 in maintaining the distinction between HRDE-1- and CSR-1-loaded 22G endo-siRNAs. We will reset the germline 22G endo-siRNAs by crossing *mut-14(-/-);smut-1(-/-)* homozygous double mutants into *mut-16(-/-)* homozygous mutants. Both parental germlines contributing to this cross will be severely depleted of 22G endo-siRNAs, but the progeny will be heterozygous for all three mutations (*mut-14(+/-);smut-1(+/-);mut-16(+/-)*), thus the 22G endo-siRNA pathway will be intact. We expect about one-third of the F1 progeny to be sterile, based on previous reports (Phillips et al. 2015). We will also reset 22G RNAs in the *morc-1(-)* homozygous mutant background. We will evaluate fertility in the F1 progeny and also transgenerationally at 25°C (i.e. test for a fertility defect two or more generations after resetting). As a control, we will reset the 22G RNAs in the *prg-1(-)* homozygous mutant background which is expected to cause 100% sterility in the F1 generation (de Albuquerque et al. 2015; Phillips et al. 2015). We expect that if *morc-1*

162

contributes to the memory of HRDE-1-associated versus CSR-1-associated 22G endo-siRNAs, F1 sterility will be enhanced in the *morc-1(-)* background and/or these worms will exhibit an accelerated germline mortality defect compared to *morc-1(-)* mutants whose 22G endo-siRNAs have not been reset. Finally, we will perform the same 22G endo-siRNA resetting experiment in a strain expressing *3xHA::csr-1* at the endogenous *csr-1* locus (generated by CRISPR-Cas9). Immunopurification of 3xHA::CSR-1 in the F1 progeny and sequencing of the bound small RNAs will identify the 22G RNAs that bind to CSR-1 RISC following 22G RNA resetting. We will compare the small RNA species that bind to CSR-1 following 22G endo-siRNA resetting in the wildtype and *morc-1(-)* backgrounds. We expect that in the *morc-1(-)* background, CSR-1 will bind its normal suite of 22G endo-siRNAs and also a set of new endo-siRNAs that correspond to HRDE-1-target genes, indicating that HRDE-1-associated 22G endo-siRNAs are misloaded onto CSR-1 in the absence of MORC-1. If we observe evidence of 22G misloading in *morc-1(-)* mutants, we will confirm CSR-1-targeting of HRDE-1-target genes by UV crosslinking of 3xHA::CSR-1 followed by immunoprecipitation and high-throughput sequencing of the bound RNAs (HITS-CLIP) in the wildtype and *morc-1(-)* backgrounds.

Summary and Significance

In recent years, untangling the connections between distinct germline small RNA pathways has become a major goal of in the *C. elegans* small RNA field. Thus far, investigations have focused on the convergence of the CSR-1, HRDE-1, and piRNA pathways at the level of Argonaute loading or of post-transcriptional gene silencing. How the different germline small RNA pathways are integrated at the chromatin level remains unexplored. Our findings that MORC-1 is required for repression of CSR-1-target genes and likely functions downstream of both CSR-1 RISC and HRDE-1 RISC suggests that MORC-1 may integrate upstream signals from both pathways to protect heterochromatin-euchromatin boundaries. The studies we propose here will elucidate the role of MORC-1 in the CSR-1 pathway and will explore a potential role for CSR-1 in MET-1-mediated euchromatic encroachment at MORC-1 target loci.
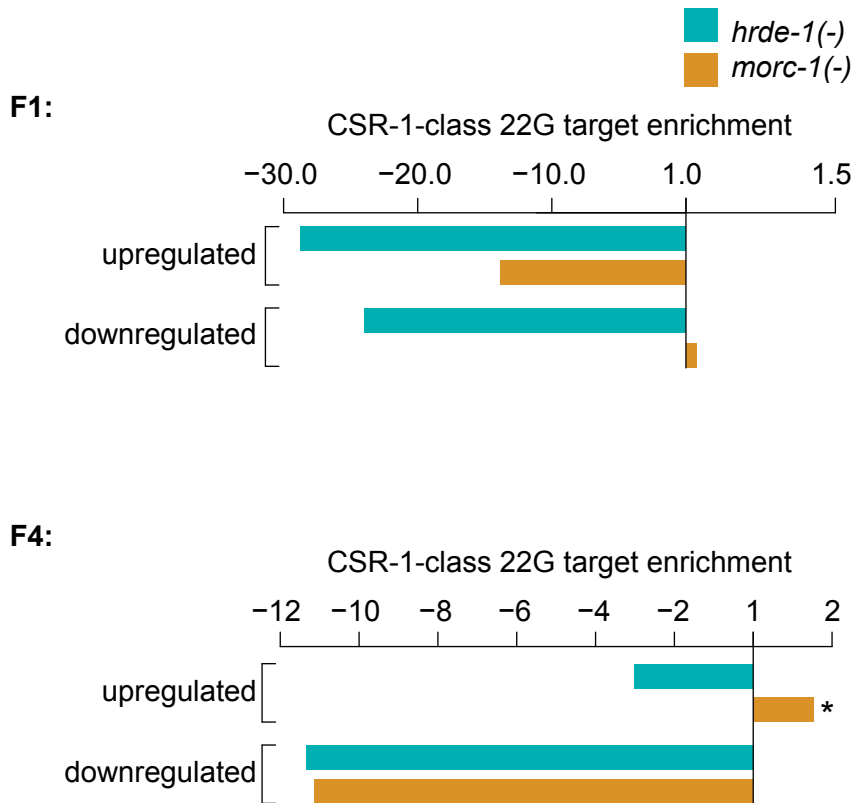
Figure 3.5. MORC-1 regulates targets of the CSR-1-class of 22G RNAs. mRNAs that are upregulated in *morc-1(-)* compared to wildtype at the F4 generation are significantly enriched for targets of the CSR-1 22G RNA pathway (p=9.591x10^{-14}, Fisher's test). There is no enrichment for CSR-1 targets among *morc-1(-)* upregulated targets at F1 generation or *morc-1(-)* downregulated targets at either generation. There is no enrichment for CSR-1 targets among *hrde-1(-)* up- or down-regulated targets at F1 or F4 generation.
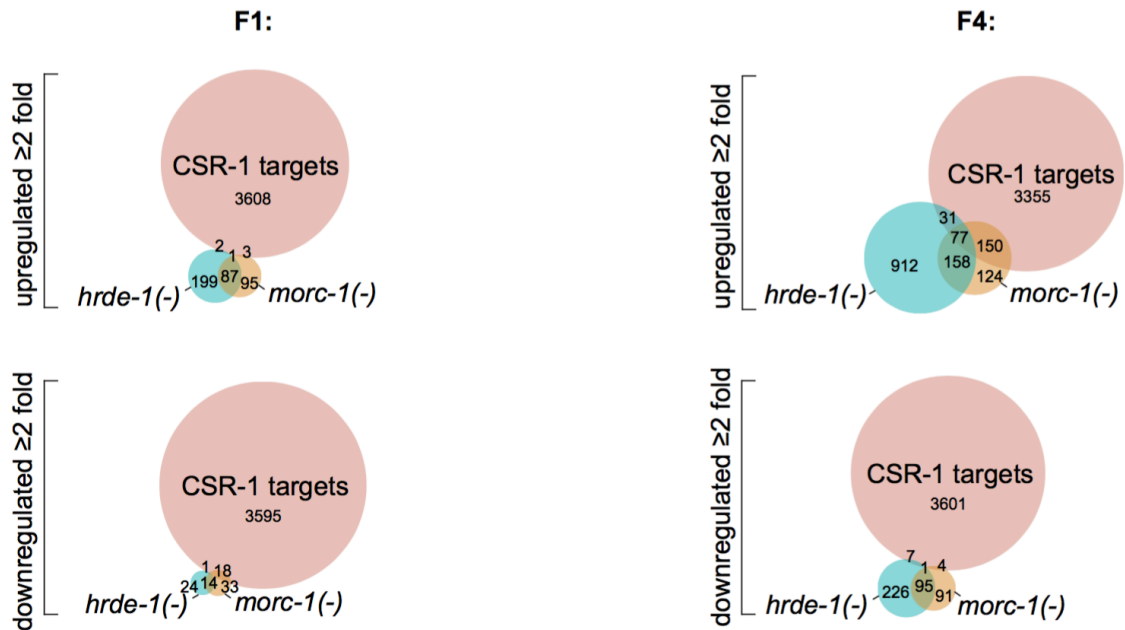
Figure 3.6. MORC-1 regulates CSR-1 targets independent of HRDE-1. Overlap of testable CSR targets with *hrde-1(-)* and *morc-1(-)* up- and down-regulated regulated targets at F1 (left) and F4 (right) generations. Only F4 *morc-1(-)* upregulated targets are enriched for CSR-1 targets.
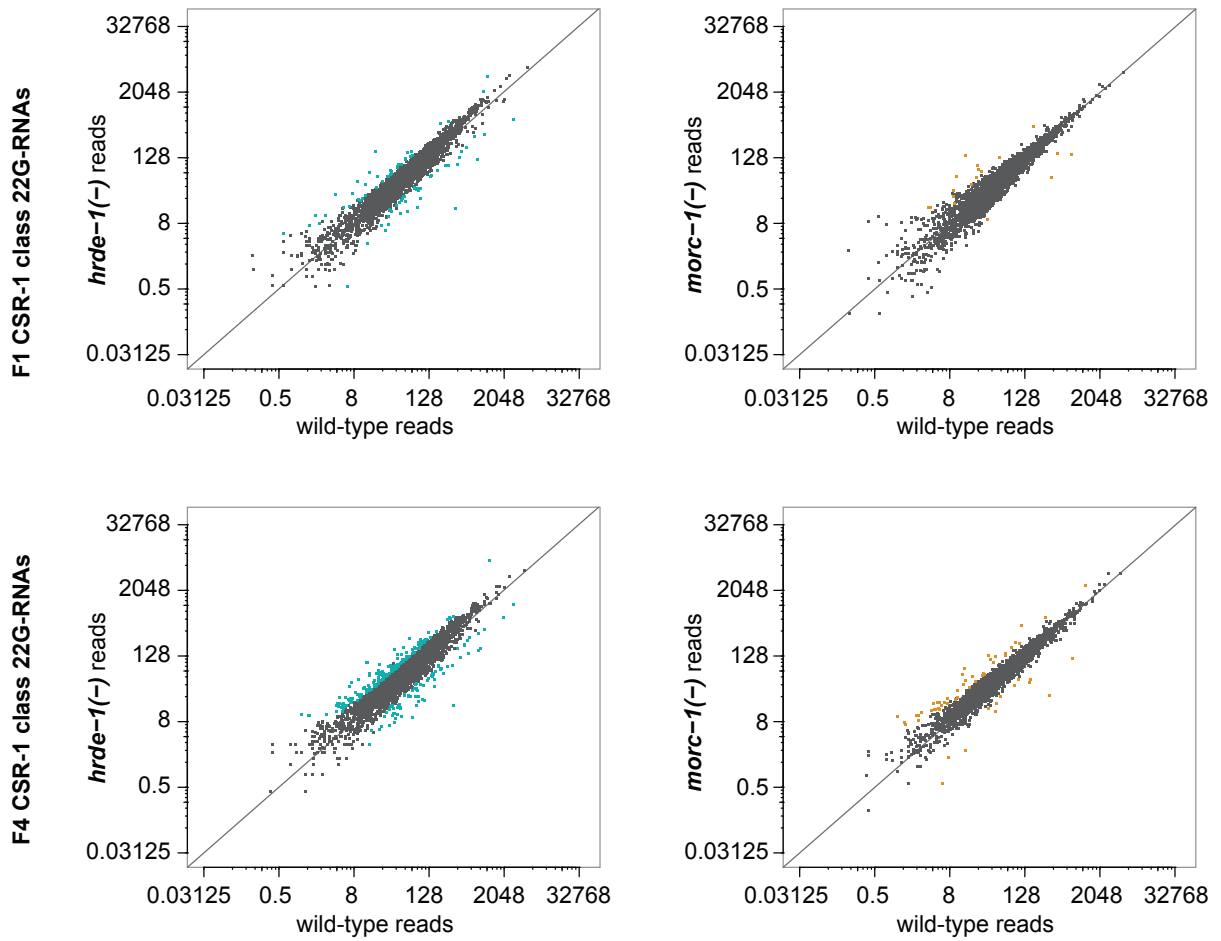
Figure 3.7. MORC-1 regulates CSR-1 target genes downstream of siRNAs. 22G RNA reads targeting annotated CSR-1 target mRNAs in *hrde-1(-)* (left) and *morc-1(-)* mutants (right) at early (top) and late (bottom) generations. 22G RNAs that are significantly enriched or depleted ($p<0.05$) in mutant compared to wildtype are highlighted in blue for *hrde-1(-)* and yellow for *morc-1(-).*
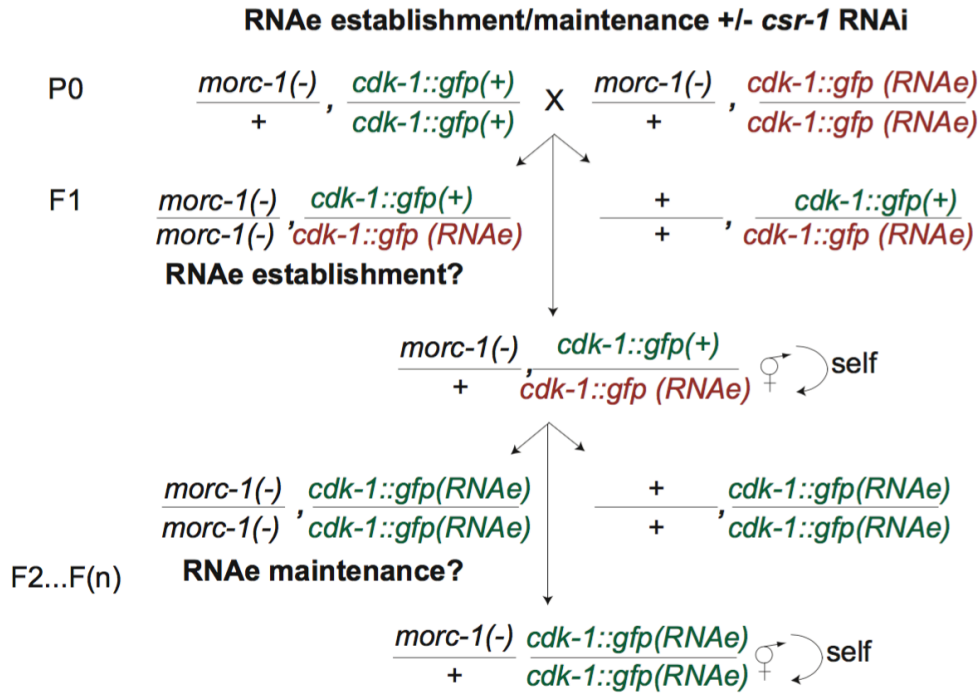
Figure 3.8. Diagram of genetic crosses to test RNAe in *morc-1(-)* mutants. Crosses will be performed with and without *csr-1* RNAi. Worms will be scored for GFP expression in the F1 generation to evaluate RNAe establishment and in the F2 and subsequent generations to evaluate RNAe maintenance.
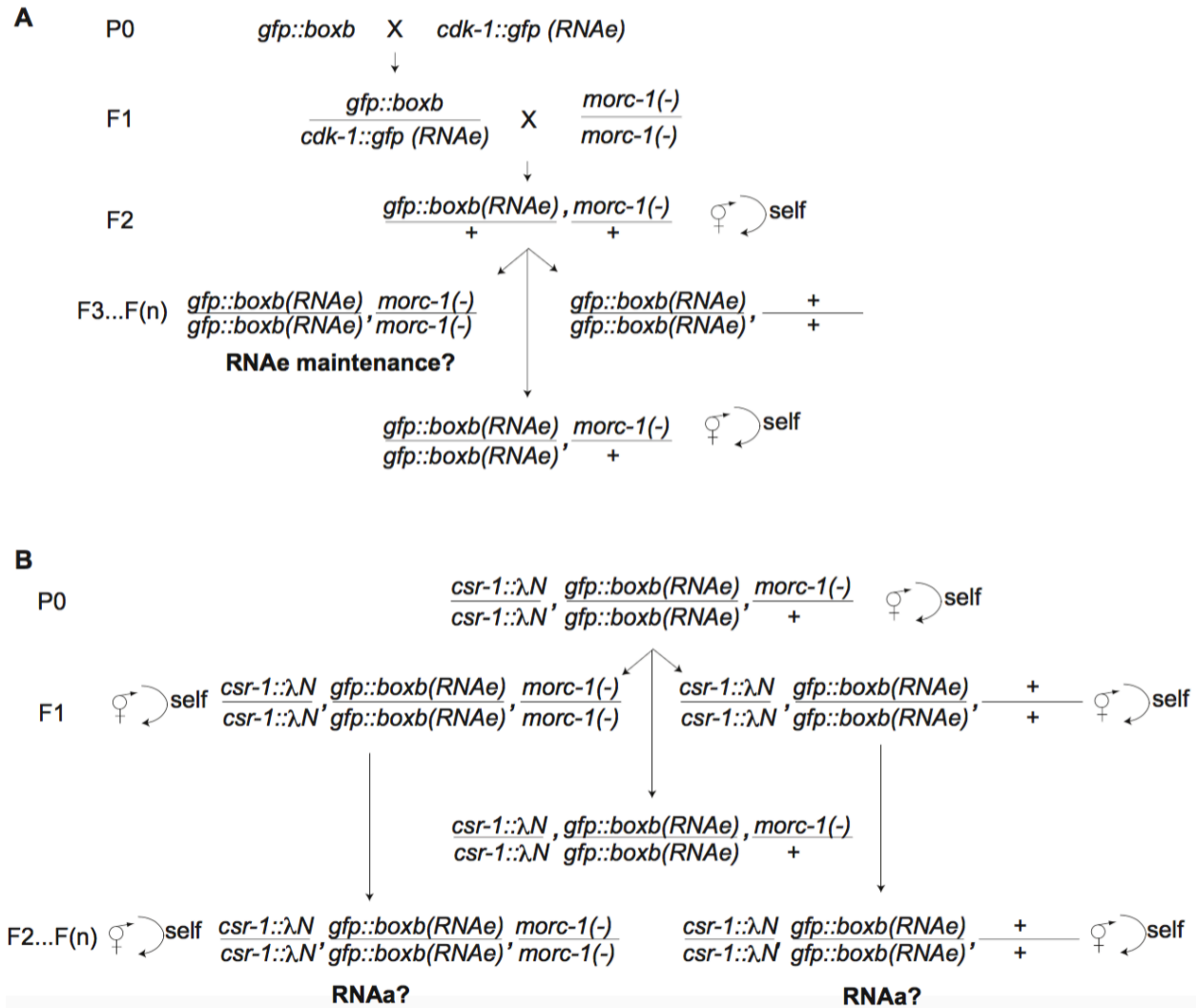
Figure 3.9. Diagram of genetic crosses to test *morc-1(-)* enhancement of RNAa.
(A) Evaluation of *morc-1* contribution to RNAe maintenance of *gfp::boxb.* Worms will be scored at F3 and subsequent generations for RNAe maintenance. **(B)** Evaluation of *morc-1* contribution to RNAa. Worms will be scored at F2 and subsequent generations for RNAa establishment.

P0 $\dfrac{csr\text{-}1::\lambda N \quad gfp::boxb(RNAe) \quad met\text{-}1(\text{-})}{csr\text{-}1::\lambda N' \; gfp::boxb(RNAe)' \quad +}$ ♂⤺self

F1   ♀⤺self $\dfrac{csr\text{-}1::\lambda N \quad gfp::boxb(RNAe) \quad met\text{-}1(\text{-})}{csr\text{-}1::\lambda N' gfp::boxb(RNAe)' \; met\text{-}1(\text{-})}$ $\dfrac{csr\text{-}1::\lambda N \quad gfp::boxb(RNAe) \quad +}{csr\text{-}1::\lambda N' gfp::boxb(RNAe)' \quad +}$ ♀⤺self

$\dfrac{csr\text{-}1::\lambda N \; , \; gfp::boxb(RNAe) \; , \; met\text{-}1(\text{-})}{csr\text{-}1::\lambda N \; \; gfp::boxb(RNAe) \quad +}$

F2...F(n) ♀⤺self $\dfrac{csr\text{-}1::\lambda N \quad gfp::boxb(RNAe) \quad met\text{-}1(\text{-})}{csr\text{-}1::\lambda N' gfp::boxb(RNAe)' \; met\text{-}1(\text{-})}$ $\dfrac{csr\text{-}1::\lambda N \quad gfp::boxb(RNAe) \quad +}{csr\text{-}1::\lambda N \; \; gfp::boxb(RNAe)' \quad +}$ ♀⤺self

RNAa?   **RNAa?**
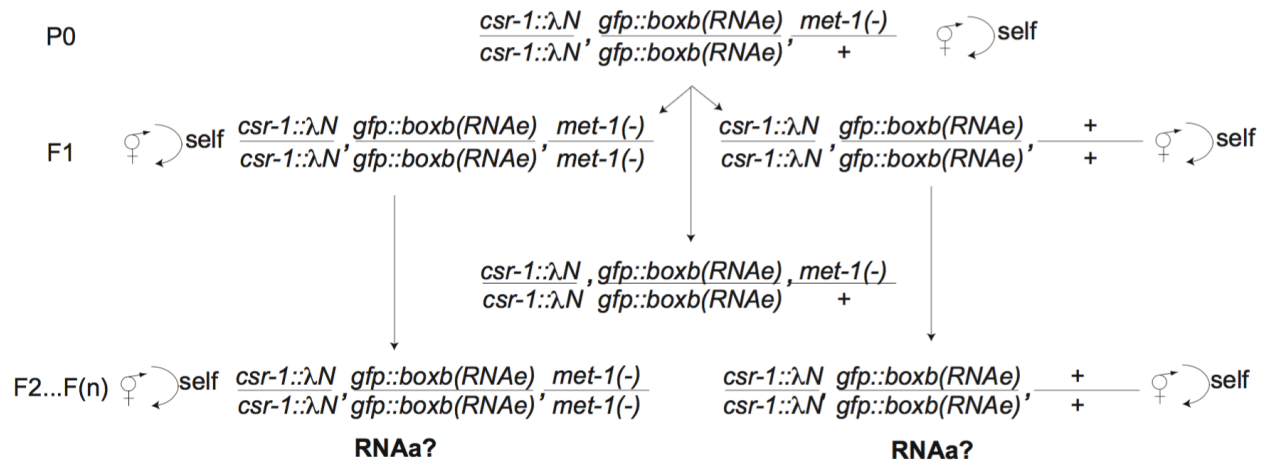
Figure 3.10. Diagram of genetic crosses to test *met-1* contribution to RNAa. Evaluation of the number of generations of CSR-1-tethering required to activate an RNAe-silenced *gfp::boxb* in wildtype versus *met-1* mutants. Worms will be scored for RNAa establishment at F2 and subsequent generations. Both *met-1(-)* deletion alleles and the *met-1(xk4)* hypomorphic allele will be tested.

Figure 3.11. Diagram of genetic crosses to test *met-1* contribution to maintenance of RNAa. The *gfp::boxb(RNAe)* allele will be activated by 10 generations of tethering to CSR-1. After removal of the CSR-1 tether, the number of generations required to re-silence in wildtype versus *met-1* alleles will be counted. F13 and subsequent generations will be scored for re-establishment of RNAe. Both *met-1(-)* deletion alleles and the *met-1(xk4)* hypomorphic allele will be tested.

REFERENCES

Ahmed S, Hodgkin J. 2000. MRT-2 checkpoint protein is required for germline immortality and telomere replication in C. elegans. Nature 403: 159–164.

Andrews FH, Tong Q, Sullivan KD, Cornett EM, Zhang Y, Ali M, Ahn J, Pandey A, Guo AH, Strahl BD, Costello JC, Espinosa JM, et al. 2016. Multivalent Chromatin Engagement and Inter-domain Crosstalk Regulate MORC3 ATPase. Cell Rep 16: 3195–3207.

Ashe A, Sapetschnig A, Weick E-M, Mitchell J, Bagijn MP, Cording AC, Doebley A-L, Goldstein LD, Lehrbach NJ, Le Pen J, Pintacuda G, Sakaguchi A, et al. 2012. piRNAs can trigger a multigenerational epigenetic memory in the germline of C. elegans. Cell 150: 88–99.

Bolisetty MT, Rajadinakaran G, Graveley BR. 2015. Determining exon connectivity in complex mRNAs by nanopore sequencing. Genome Biol. 16: 204.

Buckley BA, Burkhart KB, Gu SG, Spracklin G, Kershner A, Fritz H, Kimble J, Fire A, Kennedy S. 2012. A nuclear Argonaute promotes multigenerational epigenetic inheritance and germline immortality. Nature 489: 447–451.

Cecere G, Hoersch S, O'Keeffe S, Sachidanandam R, Grishok A. 2014. Global effects of the CSR-1 RNA interference pathway on the transcriptional landscape. Nat. Struct. Mol. Biol. 21: 358–365.

Claycomb JM, Batista PJ, Pang KM, Gu W, Vasale JJ, van Wolfswinkel JC, Chaves DA, Shirayama M, Mitani S, Ketting RF, Conte D, Mello CC. 2009. The Argonaute CSR-1 and its 22G-RNA cofactors are required for holocentric chromosome segregation. Cell 139: 123–134.

Conine CC, Moresco JJ, Gu W, Shirayama M, Conte D, Yates JR, Mello CC. 2013. Argonautes promote male fertility and provide a paternal memory of germline gene expression in C. elegans. Cell 155: 1532–1544.

Corbett KD, Berger JM. 2003. Structure of the topoisomerase VI-B subunit: implications for type II topoisomerase mechanism and evolution. EMBO J. 22: 151–163.

Corbett KD, Berger JM. 2005. Structural dissection of ATP turnover in the prototypical GHL ATPase TopoVI. Structure 13: 873–882.

de Albuquerque BFM, Placentino M, Ketting RF. 2015. Maternal piRNAs Are Essential for Germline Development following De Novo Establishment of Endo-siRNAs in Caenorhabditis elegans. Dev. Cell 34: 448–456.

Dutta R, Inouye M. 2000. GHKL, an emergent ATPase/kinase superfamily. Trends Biochem. Sci. 25: 24–28.

Gartner A, Boag PR, Blackwell TK. 2008. Germline survival and apoptosis. WormBook

1–20.

Gartner A, Milstein S, Ahmed S, Hodgkin J, Hengartner MO. 2000. A conserved checkpoint pathway mediates DNA damage--induced apoptosis and cell cycle arrest in C. elegans. Mol. Cell 5: 435–443.

Goodwin S, Gurtowski J, Ethe-Sayers S, Deshpande P, Schatz MC, McCombie WR. 2015. Oxford Nanopore sequencing, hybrid error correction, and de novo assembly of a eukaryotic genome. Genome Res. 25: 1750–1756.

Guang S, Bochner AF, Burkhart KB, Burton N, Pavelec DM, Kennedy S. 2010. Small regulatory RNAs inhibit RNA polymerase II during the elongation phase of transcription. Nature 465: 1097–1101.

Gumienny TL, Lambie E, Hartwieg E, Horvitz HR, Hengartner MO. 1999. Genetic control of programmed cell death in the Caenorhabditis elegans hermaphrodite germline. Development 126: 1011–1022.

Hirano T. 2016. Condensin-Based Chromosome Organization from Bacteria to Vertebrates. Cell 164: 847–857.

Iyer LM, Abhiman S, Aravind L. 2008. MutL homologs in restriction-modification systems and the origin of eukaryotic MORC ATPases. Biol. Direct 3: 8.

Kelly WG. 2014. Transgenerational epigenetics in the germline cycle of Caenorhabditis elegans. Epigenetics Chromatin 7: 6.

Kim DI, Jensen SC, Noble KA, Kc B, Roux KH, Motamedchaboki K, Roux KJ. 2016. An improved smaller biotin ligase for BioID proximity labeling. Mol. Biol. Cell 27: 1188–1196.

Li S, Yen L, Pastor WA, Johnston JB, Du J, Shew CJ, Liu W, Ho J, Stender B, Clark AT, Burlingame AL, Daxinger L, et al. 2016. Mouse MORC3 is a GHKL ATPase that localizes to H3K4me3 marked chromatin. Proc. Natl. Acad. Sci. U.S.A. 201609709.

Liu Y, Tempel W, Zhang Q, Liang X, Loppnau P, Qin S, Min J. 2016. Family-wide Characterization of Histone Binding Abilities of Human CW Domain-containing Proteins. J. Biol. Chem. 291: 9000–9013.

Luteijn MJ, van Bergeijk P, Kaaij LJT, Almeida MV, Roovers EF, Berezikov E, Ketting RF. 2012. Extremely stable Piwi-induced gene silencing in Caenorhabditis elegans. EMBO J. 31: 3422–3430.

Meier B, Barber LJ, Liu Y, Shtessel L, Boulton SJ, Gartner A, Ahmed S. 2009. The MRT-1 nuclease is required for DNA crosslink repair and telomerase activity in vivo in Caenorhabditis elegans. EMBO J. 28: 3549–3563.

O'Neil N, Rose A. 2006. DNA repair. WormBook 1–12.

Pastor WA, Stroud H, Nee K, Liu W, Pezic D, Manakov S, Lee SA, Moissiard G, Zamudio N, Bourc'his D, Aravin AA, Clark AT, et al. 2014. MORC1 represses transposable elements in the mouse male germline. Nat Commun 5: 5795.

Peng JC, Karpen GH. 2006. H3K9 methylation and RNA interference regulate nucleolar organization and repeated DNA stability. Nat Cell Biol 9: 25–35.

Perry J, Zhao Y. 2003. The CW domain, a structural module shared amongst vertebrates, vertebrate-infecting parasites and higher plants. Trends Biochem. Sci. 28: 576–580.

Peters AH, O'Carroll D, Scherthan H, Mechtler K, Sauer S, Schöfer C, Weipoltshammer K, Pagani M, Lachner M, Kohlmaier A, Opravil S, Doyle M, et al. 2001. Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. Cell 107: 323–337.

Phillips CM, Brown KC, Montgomery BE, Ruvkun G, Montgomery TA. 2015. piRNAs and piRNA-Dependent siRNAs Protect Conserved and Essential C. elegans Genes from Misrouting into the RNAi Pathway. Dev. Cell 34: 457–465.

Rechtsteiner A, Ercan S, Takasaki T, Phippen TM, Egelhofer TA, Wang W, Kimura H, Lieb JD, Strome S. 2010. The histone H3K36 methyltransferase MES-4 acts epigenetically to transmit the memory of germline gene expression to progeny. PLoS Genet. 6: e1001091.

Rhoads A, Au KF. 2015. PacBio Sequencing and Its Applications. Genomics Proteomics Bioinformatics 13: 278–289.

Roux KJ, Kim DI, Raida M, Burke B. 2012. A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. J. Cell Biol. 196: 801–810.

Seth M, Shirayama M, Gu W, Ishidate T, Conte D, Mello CC. 2013. The C. elegans CSR-1 argonaute pathway counteracts epigenetic silencing to promote germline gene expression. Dev. Cell 27: 656–663.

She X, Xu X, Fedotov A, Kelly WG, Maine EM. 2009. Regulation of heterochromatin assembly on unpaired chromosomes during Caenorhabditis elegans meiosis by components of a small RNA-mediated pathway. PLoS Genet. 5: e1000624.

Shirayama M, Seth M, Lee H-C, Gu W, Ishidate T, Conte D, Mello CC. 2012. piRNAs initiate an epigenetic memory of nonself RNA in the C. elegans germline. Cell 150: 65–77.

Takahashi K, Yoshida N, Murakami N, Kawata K, Ishizaki H, Tanaka-Okamoto M, Miyoshi J, Zinn AR, Shime H, Inoue N. 2007. Dynamic regulation of p53 subnuclear localization and senescence by MORC3. Mol. Biol. Cell 18: 1701–1709.

Updike DL, Strome S. 2009. A genomewide RNAi screen for genes that affect the stability, distribution and function of P granules in Caenorhabditis elegans. Genetics 183: 1397–1419.

van Wolfswinkel JC, Claycomb JM, Batista PJ, Mello CC, Berezikov E, Ketting RF. 2009. CDE-1 affects chromosome segregation through uridylation of CSR-1-bound siRNAs. Cell 139: 135–148.

Wedeles CJ, Wu MZ, Claycomb JM. 2013. Protection of germline gene expression by the C. elegans Argonaute CSR-1. Dev. Cell 27: 664–671.

Zeller P, Padeken J, van Schendel R, Kalck V, Tijsterman M, Gasser SM. 2016. Histone H3K9 methylation is dispensable for Caenorhabditis elegans development but suppresses RNA:DNA hybrid-associated repeat instability. Nat. Genet. 1–13.

Zhou Z, Hartwieg E, Horvitz HR. 2001. CED-1 is a transmembrane receptor that mediates cell corpse engulfment in C. elegans. Cell 104: 43–56.

CHAPTER 4

Crosstalk between endo-siRNAs, piRNAs, and splicing machinery regulates germline gene expression

AUTHORS: Natasha E. Weiser, Amelia F. Alessi, Suhua Feng, Kristen C. Brown, Taiowa A. Montgomery, Steven E. Jacobsen, and John K. Kim

AUTHOR CONTRIBUTIONS: Unless otherwise stated, all experiments and analysis were performed by me. Other contributions are as follows:

- F39E9.7 intronless experiment (Figure 4.3): Amelia F. Alessi (Johns Hopkins University)
- Library preparation and sequencing: Suhua Feng and Steven E. Jacobsen (University of California – Los Angeles)
- Computational analysis of sRNA-seq: Kristen C. Brown and Taiowa A. Montgomery (Colorado State University)

INTRODUCTION

The *C. elegans* germline is host to a diverse set of small non-coding RNAs that collectively regulate gene expression in the germline and developing embryo. As described in Chapter One, these include small RNAs that mediate gene silencing as well as separate classes that mediate gene licensing. The silencing small RNAs include piRNAs and oocyte-enriched primary 26G RNAs (ERGO-1-class 26G RNAs). Upon target binding, both piRNAs and 26G RNAs can trigger amplification of secondary

WAGO-class 22G RNAs (22G RNAs) (Batista et al. 2008; Das et al. 2008; Han et al. 2009; Vasale et al. 2010; Lee et al. 2012). As previously discussed, the ALG-3/ALG-4-class of spermatogenic primary 26G RNAs and the CSR-1-class of 22G RNAs contribute to gene licensing (Conine et al. 2013; Seth et al. 2013; Wedeles et al. 2013). Despite the widespread interest in these pathways over the last decade, there are many unanswered questions regarding how biogenesis of these small RNAs is specified and regulated. Here, I will discuss data and future directions of two studies investigating how biogenesis of ERGO-1-class 26G RNAs and piRNAs are regulated and integrated with other gene regulatory mechanisms.

First, I will discuss an ongoing study investigating how ERGO-1-class 26G RNAs are specified. To date, no primary sequence motif has been identified that might explain why some mRNAs serve as templates and targets of these 26G RNAs and some do not. We hypothesize that this template/target selection is mediated by splicing dynamics, including a novel 26G RNA biogenesis factor TCER-1.

Second, I will discuss an unexpected finding from the studies described in Chapter Two: We found through high-throughput 5'-dependent small RNA sequencing that late generation *hrde-1(-)* mutants are depleted of piRNAs. We speculate that HRDE-1-mediated regulation of chromatin state at the piRNA clusters may contribute to piRNA biogenesis and thus piRNAs and HRDE-1-class RNAs may function in a feed forward loop to control germline gene expression.

CONTROL OF ENDO-siRNA BIOGENESIS BY SPLICING MACHINERY

The *C. elegans* 26G RNAs are generated by the RdRP RRF-3 and its associated processing factors (Simmer et al. 2002; Duchaine et al. 2006; Gent et al. 2009; Han et al. 2009). ERGO-1-class and ALG-3/4-class 26G RNAs share a common set of biogenesis factors but are processed differently (for example, only the ERGO-1-class 26G RNAs are 2' O-methlyated by HENN-1) and are loaded onto distinct RISCs (Han et al. 2009; Conine et al. 2010; Billi et al. 2012; Kamminga et al. 2012; Montgomery et al.

2012). In this chapter, I will focus entirely on the biogenesis of ERGO-1-class 26G RNAs that are expressed during oogenesis and in embryos (Han et al. 2009).

The majority of 26G RNA mediated silencing is thought to arise not directly from ERGO-1-associated-26G RNA, but rather from the secondary 22G RNAs whose biogenesis is triggered by 26G RNA binding to its target (Gent et al. 2010; Vasale et al. 2010; Fischer et al. 2011). The details of this phenomenon are discussed in Chapter One. If not to execute target silencing, what is the role of the ERGO-1 pathway? One possibility is that the primary function of the ERGO-1 26G RNAs is to select the targets to be silenced. Because the 26G RNAs are RdRP products, their template mRNAs can be identified by their 100% sequence complementarity to the corresponding 26G RNAs (Han et al. 2009). As 26G RNAs also bind their targets with 100% sequence complementarity, the template mRNAs are also the targets (Han et al. 2009). There must be a mechanism by which the RdRP module selectively targets a distinct set of template mRNAs, but no consensus motif or other primary sequence feature has been found that can explain 26G RNA template selection, suggesting that template selection is determined by other factors.

There is emerging evidence that the splicing of endo-siRNA templates may contribute to siRNA biogenesis. Phylogenetic studies across 86 eukaryotic species have shown that many *C. elegans* RNAi factors share phylogenetic profiles with nine spliceosomal components, suggesting that these pathways may be functionally linked (Tabach et al. 2013). For example, since its divergence from *S. pombe, S. cerevisiae* has lost spliceosome components, introns, RdRP, and Argonaute (Tabach et al. 2013). The *S. pombe* genome, in contrast, has retained both introns and siRNA-mediated gene silencing (Tabach et al. 2013).

Recent studies in *Cryptococcus neoformans* and *S. pombe,* two yeast species with relatively intron-rich genomes and intact siRNA pathways, have shown convincing evidence of a mechanistic link between these pathways. In both species, components of the RNAi machinery physically interact with spliceosome components (Dumesic et al. 2013; Lee et al. 2013). Furthermore, a transcript must engage the spliceosome in order to produce siRNAs (Dumesic et al. 2013; Lee et al. 2013). In *C. neoformans,* the finding that spliceosome occupancy is higher at siRNA template mRNAs compared to the rest

of the genome has led to a model in which stalled spliceosomes recruit siRNA biogenesis machinery such as RdRP by increasing the availability of the mRNA template (Dumesic et al. 2013). This is supported by the finding that mRNAs that do not normally serve as siRNA templates can become siRNA templates with the introduction of mutations at canonical 3' splice acceptor sites (Dumesic et al. 2013). Only alterations at the 3' splice acceptor site, not at the 5' splice donor site, can turn a non-template mRNA into a template mRNA. Thus, ectopic siRNA accumulation is dependent on a functional 5' splice donor at the modified intron, highlighting that spliceosome engagement is required for siRNA biogenesis (Dumesic et al. 2013). In *S. pombe*, Nrl1, the homolog of *C. elegans* NRDE-2, physically interacts with components of the U2 and U5 small nuclear ribonucleo proteins (snRNPs) (Lee et al. 2013). In the *nrl1Δ* strain, decreased utilization of cryptic introns accompanies loss of non-centromeric siRNAs (Lee et al. 2013). Furthermore, when specific cryptic introns are deleted, the siRNAs that target the altered locus and neighboring genes are lost (Lee et al. 2013). These findings show that the utilization of cryptic introns drives siRNA formation.

TCER-1 is required for ERGO-1-class 26G RNA expression

In *C. elegans*, deep sequencing studies indicate that some 26G RNAs span exon-exon junctions, suggesting that they are templated by a spliced transcript (Ruby et al. 2006; Han et al. 2009). The identification of a number of splicing factors and nuclear pore components in screens for siRNA pathway components suggests that, as in yeast, siRNA accumulation and splicing may be functionally linked or regulated by a common set of factors (Kim et al. 2005; Robert et al. 2005; Montgomery et al. 2012). One such candidate is TCER-1, a putative transcription elongation factor (Kim et al. 2005; Montgomery et al. 2012). The mammalian homolog of TCER-1, TCERG/CA150, interacts with RNA Pol II and splicing machinery (Goldstrohm et al. 2001; Smith et al. 2004; Sánchez-Alvarez et al. 2006). In *C. elegans*, TCER-1 regulates expression of a diverse set of transcripts (Pushpa et al. 2013; Amrit et al. 2016). Some of the germline genes that are downregulated in *tcer-1(-)* compared to wildtype exhibit increased pre-

mRNA levels, suggesting that TCER-1 regulates gene expression post-transcriptionally (Pushpa et al. 2013). Yeast two-hybrid experiments indicate that TCER-1 also physically interacts with the snRNP SNR-3, which intriguingly has also been identified as a candidate DCR-1 interactor (Duchaine et al. 2006; Pushpa et al. 2013).

We tested whether TCER-1 is required for ERGO-1-class 26G RNA accumulation. We performed Taqman qRT-PCR to quantify expression levels of three ERGO-1 class 26G RNAs in wildtype, *eri-1(-)*, and *tcer-1(-)* backgrounds. We found that *tcer-1(-)* mutants are depleted of all three 26G RNAs (Figure 4.1). Because *tcer-1(-)* mutants have reduced fertility compared to wildtype at 20°C (Amrit et al. 2016) and are germline mortal at 22.5°C (Figure 4.2), we quantified the piRNA 21U-1848 as a control for global depletion of germline small RNAs due to impaired germline function. We found that *tcer-1(-)* mutants expressed wildtype levels of 21U-1848, suggesting that *tcer-1(-)* is specifically depleted for ERGO-1-class 26G RNAs (Figure 4.1). We are currently analyzing global RNA transcripts in *tcer-1(-)* mutant compared to wildtype to determine whether *tcer-1* regulates splicing. Our initial analysis suggests that *tcer-1(-)* mutants express about 300 novel exons that have not previously been annotated in wildtype worms (M. Freeberg, personal communication). Taken together, these data suggest that *tcer-1* is required for both 26G RNA accumulation and for canonical splicing. We are currently investigating how the intron/exon structure of 26G RNA target genes is affected in *tcer-1(-)* mutants. We hypothesize that *tcer-1* may modulate splicing kinetics and/or splice site utilization to promote 26G RNA biogenesis.

Spliceosome engagement contributes to 26G RNA accumulation

We next wanted to assess whether spliceosome engagement is required for 26G RNA biogenesis. We chose *F39E9.7,* the template and target of the highly-expressed 26G-O3, as a test case. *F39E9.7* is a pseudogene that does not encode protein (LeGendre et al. 2013). Importantly, the 26G-O3 sequence is exonic, so we can manipulate the intron independently of the 26G-O3 template sequence. We used CRISPR-Cas9 to delete the single intron from the *F39E9.7* locus, thus generating

F39E9.7(intronless) mutants. We quantified expression of *F39E9.7* mRNA and 26G-O3 in wildtype worms and *F39E9.7(intronless)* mutants grown on empty vector RNAi or *ergo-1* RNAi. On empty vector RNAi the *F39E9.7(intronless)* mutants expressed 26G-O3 at approximately 60% of wildtype levels (Figure 4.3A). In the *F39E9.7(intronless)* mutant, *F39E9.7* mRNA levels were about 1.5-fold overexpressed relative to wildtype (Figure 4.3A). We observed a similar degree of *F39E9.7* mRNA upregulation in wildtype worms grown on *ergo-1* RNAi (Figure 4.3A). 26G-O3 levels in both wildtype and *F39E9.7(intronless)* worms were severely depleted by *ergo-1* RNAi (Figure 4.3A). Taken together, these data suggest that 26G-O3 biogenesis is impaired in the *F39E9.7(intronless)* mutant and that the 26G-O3 mRNA target is upregulated accordingly. Because *F39E9.7* mRNA is the template as well as the target of 26G-O3, target upregulation in the absence of an siRNA defect would be expected to lead to increased 26G-O3 levels. When we calculated the ratio of 26G-O3 to *F39E9.7* mRNA, we found that *F39E9.79(intronless)* expressed about one-third of the wildtype amount of 26G-O3 relative to expression on the template (Figure 4.3B). These data support a model in which spliceosome engagement contributes to 26G RNA accumulation.

Ongoing studies

Our current studies are testing the requirement for spliceosomal engagement at additional endogenous 26G RNA targets using CRIPSR-Cas9 to selectively delete introns from 26G RNA target genes. We are also testing whether non-canonical splice donor and acceptor sequences can induce or repress 26G RNA biogenesis. Finally, we are inserting the *F37E9.7* intron sequence into a single copy, germline-expressed *gfp* reporter. If the introduction of the single intron is sufficient for 26G RNA biogenesis, then the reporter should be silenced in an *ergo-1-* and *eri-1*-dependent manner. These studies will elucidate the contribution of intronic sequences to 26G RNA biogenesis.

To elucidate the trans factors that link mRNA splicing and siRNA biogenesis, we will perform IP-MS of 3xFlag-tagged 26G RNA RdRP module components RRF-3, DRH-3, and ERI-5. We expect to identify physical interactions between the 26G RNA

biogenesis machinery and the spliceosome and/or nuclear pore. We will also perform HITS-CLIP of RRF-3::3xFlag to determine if 26G RNA template transcripts are fully spliced or retain introns. The retention of introns in RRF-3::3xFlag targets would suggest that 26G RNA biogenesis is concurrent with splicing.

In *S. pombe,* non-centromeric siRNAs are enriched at cryptic introns and their expression is thought to be dependent on utilization of the cryptic intron (Lee et al. 2013). If this is the case in *C. elegans*, we would expect that global changes in splice site utilization could lead to depletion of annotated 26G RNAs and expression of novel 26G RNA species. We propose to leverage *tcer-1(-)* mutants to investigate this model. We will perform mRNA-seq and small RNA-seq in wildtype, *tcer-1(-),* and *tcer-1(-);ergo-1(-)* worms. First, we will identify 26G RNAs that are uniquely expressed in *tcer-1(-)* and not in wildtype worms. To confirm that these are bona fide 26G RNAs and not degradation products, we will filter the *tcer-1(-)*-specific 26G RNAs for *ergo-1*-dependence, discarding any putative novel 26G RNA species that are not depleted in *tcer-1(-);ergo-1(-)* double mutants relative to *tcer-1(-).* Next, we will compare global mRNA levels in wildtype and *tcer-1(-)* mutants to identify loci where *tcer-1* regulates splicing. We will identify two types of *tcer-1*-regulated transcripts, based on (1) the expression of novel intron-exon structure in *tcer-1(-)* compared to wildtype or (2) altered relative ratios of different isoforms in *tcer-1(-)* mutants compared to wildtype. We predict that compared to non-regulated transcripts, *tcer-1*-regulated transcripts will be enriched for 26G RNA reads in wildtype or *tcer-1(-)* small RNA-seq libraries. In accordance with our model that splicing functions upstream of siRNA biogenesis, we expect that splice site utilization in *tcer-1(-);ergo-1(-)* mutants will resemble *tcer-1(-)* mutants and not wildtype. If this is not the case, siRNAs may act upstream of splicing. Taken together, these studies will establish novel roles for *tcer-1* in regulation of splicing and siRNA biogenesis.

Summary and significance

Endo-siRNAs promote genome defense by repressing deleterious genetic elements such as repetitive elements, pseudogenes, and other non-coding transcripts. How the endo-siRNA biogenesis machinery discerns which transcripts require silencing has been a long-standing question in the field. Here, I have presented preliminary evidence that the dynamics of transcript splicing may determine which transcripts serve as endo-siRNA templates, including the implication of a novel factor, TCER-1, that functions in both splicing and siRNA accumulation and the finding that deletion of an intron from a 26G RNA template reduces 26G RNA accumulation. The proposed experiments will further dissect what characteristics of introns promote or prevent endo-siRNA biogenesis, identify interactions between splicing components and endo-siRNA biogenesis machinery, and characterize the role of TCER-1 in the regulation of alternative splicing and in 26G RNA biogenesis. These studies will address the contribution of splicing to the selection of templates for siRNA biogenesis. One common feature of 26G RNA target genes is that they are poorly conserved and thus may be rapidly evolving (Fischer et al. 2011). Rapid sequence evolution may lead to the acquisition of subpar splice sites or the disruption of splice donor and acceptor sequences and thus promote recognition by the siRNA biogenesis machinery and siRNA-direct gene repression.
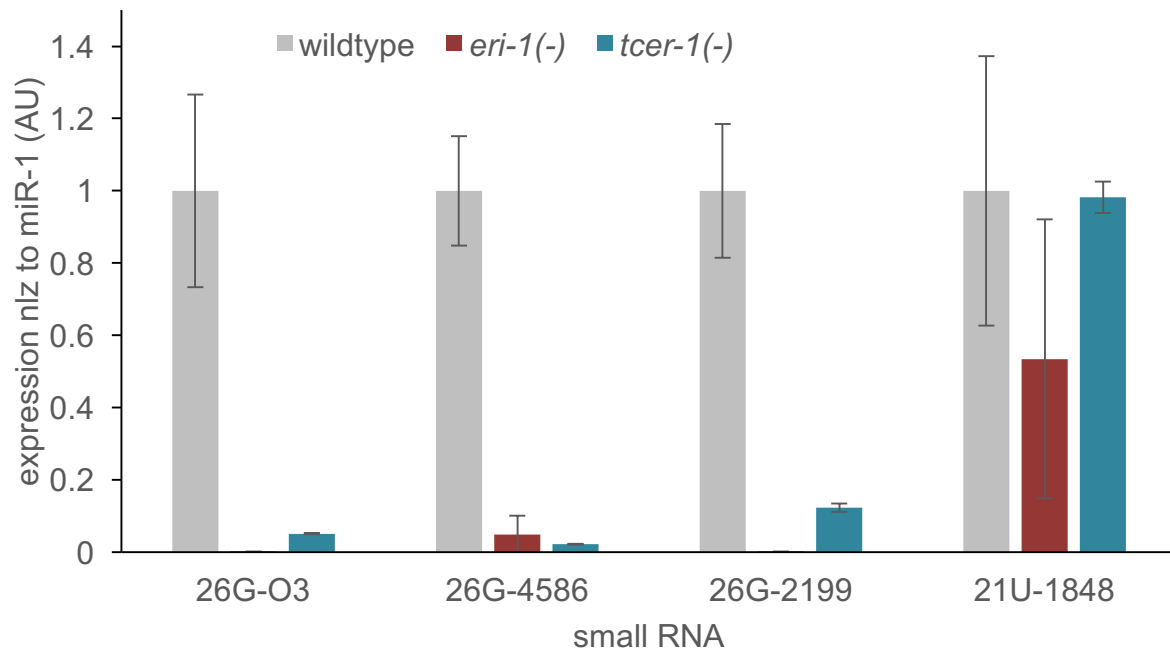
Figure 4.1. *tcer-1* is required for ERGO-1-class 26G RNA expression. Taqman qRT-PCR for ERGO-1 class 26G RNAs (O3, 4586, 2199) and the 21U-1848. *tcer-1(-)* mutants are depleted of 26G RNAs but not piRNAs. Bars indicate average expression between two replicates normalized to miR-1. Wildtype expression level for each small RNA is set to one. Error bars indicate standard deviation.
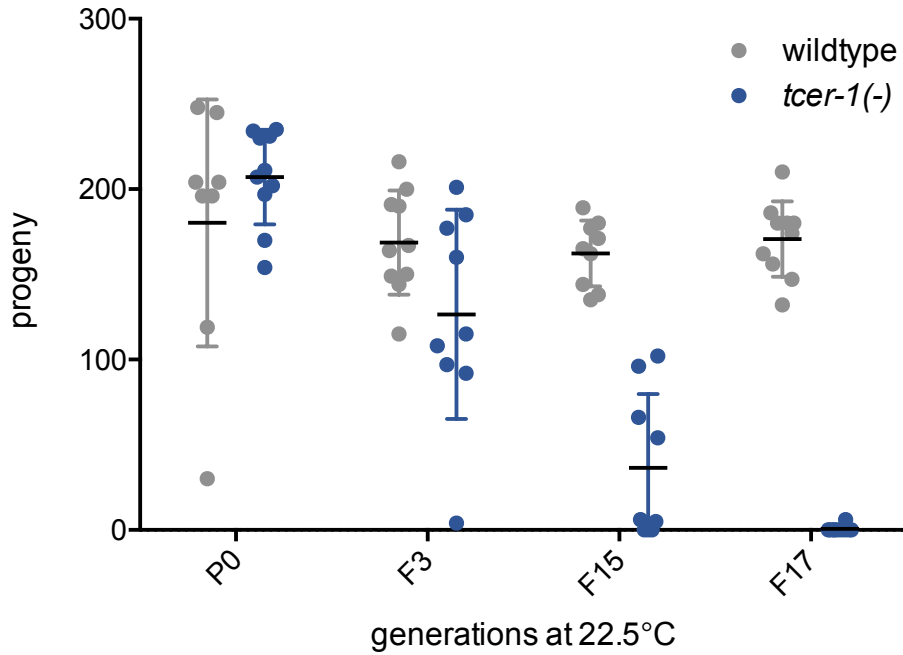
Figure 4.2. *tcer-1* **is required for germline immortality.** *tcer-1(-)* mutants are germline mortal at 22.5°C. Symbols indicate number of progeny from a single, self-fertilized hermaphrodite of the indicated genotype at the indicated generation. Error bars indicate average +/- standard deviation.
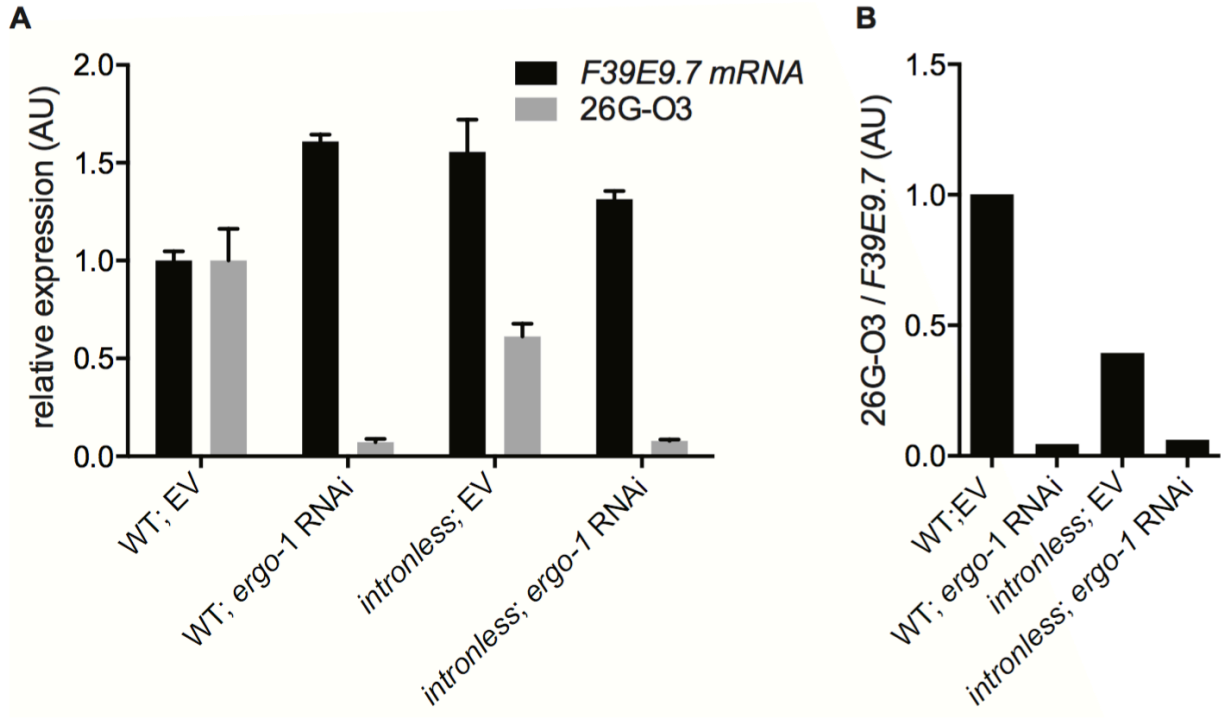
Figure 4.3. The _F37E9.7_ intron is required for 26G-O3 accumulation. (A) qRT-PCR for _F39E9.7_ and 26G-O3 in wildtype vs. _F39E9.7(intronless)_ worms. Worms were grown on empty vector RNAi (EV) or _ergo-1_ RNAi. _F39E9.7_ levels are normalized to _eft-2_, 26G-O3 levels are normalized to U18. Expression levels are shown as the mean between technical replicates in arbitrary units with expression in wildtype (EV) set to one. Error bars indicate standard deviation. **(B)** Ratio of 26G-O3 to _F39E9.7_ expression levels based on data in (A)_. F39E9.7(intronless)_ is depleted of 26G-O3 relative to wildtype.

HRDE-1 REGULATES piRNA EXPRESSION

In *C.elegans*, piRNAs are major contributors to genome defense against transposons and other "non-self" genetic elements (Batista et al. 2008; Das et al. 2008; Bagijn et al. 2012; Lee et al. 2012; Luteijn et al. 2012; Shirayama et al. 2012; de Albuquerque et al. 2015; Phillips et al. 2015). Thousands of autonomously transcribed piRNAs are expressed in the *C. elegans* germline from two clusters on chromosome IV (Ruby et al. 2006; Cecere et al. 2012; Gu et al. 2012; Billi et al. 2013). Although the actual sequences are poorly conserved, many features of piRNAs are shared among metazoans, including transcription from piRNA-rich genomic clusters, germline-specific expression, and important roles in transposon silencing (Aravin et al. 2006; Girard et al. 2006; Grivna et al. 2006; Watanabe et al. 2006; Aravin et al. 2007; Brennecke et al. 2007; Sakai et al. 2016).

In *C. elegans,* the effector function of piRNAs is mostly executed by secondary 22G RNAs bound to WAGOs and opposed by CSR-1-class 22G RNAs (Ashe et al. 2012; Bagijn et al. 2012; Lee et al. 2012; Luteijn et al. 2012; Shirayama et al. 2012; Seth et al. 2013; Wedeles et al. 2013). Interestingly, *csr-1*, *ego-1*, and *ekl-1,* part of the CSR-1 22G RNA pathway, have all been identified as suppressors of piRNA biogenesis (Goh et al. 2014). How these small RNA pathways collaborate to collectively orchestrate germline gene expression is an area of intense focus in the field.

We have identified a novel mechanism of cooperation between piRNAs and endo-siRNAs that implicates HRDE-1-bound 22G RNAs in promoting piRNA biogenesis. By high-throughput small RNA sequencing of wildtype and *hrde-1(-)* worms after two (F1) and five (F4) generations of growth at 25°C, we found that *hrde-1(-)* mutants exhibited a progressive depletion in piRNA levels from F1 to F4 (Figure 4.4A). Out of concern that this might be an artifact of the pervasive germline dysfunction in late generation *hrde-1(-)* mutants, we performed the same analysis normalizing piRNA read counts to levels of a germline-expressed miRNA, miR-35. Even with this normalization, we observed global depletion of piRNAs in late generation *hrde-1(-)* mutants, confirming that the late generation depletion of piRNAs in *hrde-1(-)* mutants is not an artifact of

overall lower germline gene expression (Figure 4.4B). Our findings suggest that *hrde-1* is required for transgenerational expression of piRNAs.

The possibility that *hrde-1* may function upstream of piRNAs is intriguing for a number of reasons. First, *hrde-1* is an important regulator of H3K9me3 deposition and maintenance and, like the piRNA depletion we observed, loss of H3K9me3 marks in *hrde-1(-)* mutants is progressive. Second, in *Drosophila melanogaster*, piRNA expression requires deposition of H3K9 methylation by *dSETDB1* (*egg*) as well as the H3K9 methyl reader Rhino (Klattenhoff et al. 2009; Rangan et al. 2011). Third, in *D. melanogaster* and *D. virilis*, piRNAs that are transcribed in the maternal germline are loaded into embryos where they (1) induce deposition of H3K9me3 marks at piRNA clusters and (2) trigger production of secondary piRNAs (Brennecke et al. 2008; Le Thomas et al. 2014a; Le Thomas et al. 2014b). The production of secondary piRNAs is dependent both on the primary piRNA trigger and also piRNA-directed H3K9me3 at the genomic locus (Le Thomas et al. 2014a; Le Thomas et al. 2014b). Thus, there is a positive feedback loop between piRNA expression and H3K9me3 deposition. A similar phenomenon has been described for siRNAs in *S. pombe*, where there is a positive feedback loop between siRNA expression and H3K9me3 deposition (Motamedi et al. 2004; Noma et al. 2004). In *C. elegans*, no such positive feedback loop between H3K9me3 and small RNAs has been identified. We propose that *C. elegans* may employ a hybrid system of sorts, with maternal piRNAs triggering the amplification of secondary 22G RNAs that act through the germline nuclear RNAi pathway (HRDE-1 RISC) to deposit H3K9me3 at target loci including at the piRNA clusters and that H3K9me3 promotes piRNA transcription. We propose several avenues by which to test this model:

Is hrde-1 required for piRNA transcription?

First, we will ask where in the piRNA pathway *hrde-1* functions. piRNAs are transcribed from the piRNA clusters as 23-30nt precursors and subsequently processed into mature 21U RNAs (Gu et al. 2012; Weick et al. 2014). Thus, loss of protein factors

that are required for piRNA transcription (e.g. GEI-11 and PRDE-1) causes depletion of both piRNA precursors and mature piRNAs, whereas downstream processing factors such as PID-1 and PARN-1 are only required for accumulation of mature piRNAs (de Albuquerque et al. 2014; Kasper et al. 2014; Weick et al. 2014; Tang et al. 2016). Based on our model that *hrde-1* promotes piRNA expression via H3K9 methylation of piRNA loci, we expect that *hrde-1* is required for transcription of piRNA precursors. Previous studies have shown that treatment of total RNA with tobacco acid pyrophosphatase (TAP) to remove 5' caps from RNA Pol II transcripts followed by 5'-monophosphate-dependent small RNA-seq allows the robust detection of piRNA precursors (Gu et al. 2012; de Albuquerque et al. 2014; Weick et al. 2014). We will use this method to quantify piRNA precursors, the majority of which will contain mature piRNA sequences with a 5' extension of 2nt, in wildtype, *hrde-1(-),* and *prde-1(-)* mutants (Gu et al. 2012; Weick et al. 2014). *prde-1* encodes an early piRNA biogenesis factor and is required for the accumulation of precursor and mature piRNAs (Weick et al. 2014). Wildtype and *hrde-1(-)* mutants will be collected at F1 and F4 generations at 25°C. Because of the severe fertility defect of *prde-1(-)* mutants (Weick et al. 2014), these samples will be collected from worms grown at 20°C and piRNA precursor levels will be compared to wildtype worms also grown at 20°C. We expect that, like *prde-1(-)* mutants, late generation *hrde-1(-)* mutants will be depleted of piRNA precursors compared to wildtype, confirming that HRDE-1 acts early in piRNA biogenesis, likely at the level of transcription.

Another explanation for global piRNA depletion in *hrde-1(-)* mutants could be that *hrde-1* negatively affects *prg-1* expression or function and that piRNA expression is lost due to depletion of PRG-1. Our early and late generation mRNA-seq data do not show any significant regulation of *prg-1* mRNA by *hrde-1*. We will quantify PRG-1 protein levels by western blot of a 3xFlag::PRG-1-expressing strain that we have generated by CRISPR-Cas9. It has previously been suggested that PRG-1 stability may require robust piRNA expression (de Albuquerque et al. 2014; Weick et al. 2014). Accordingly, a decrease in PRG-1 proteins levels could be due to piRNA depletion rather than a specific contribution of HRDE-1 to PRG-1 function. If we observe decreased PRG-1 expression in *hrde-1(-)* mutants by western blot, we will confirm that the remaining

PRG-1 is correctly expressed at P granules. We will perform immunostaining of 3xFlag::PRG-1 in early and late generation *hrde-1(-)* mutants expressing the P granule marker PGL-1::GFP. If PRG-1 depletion at the protein level is the result of piRNA depletion, we expect that any residual PRG-1 will be correctly localized in the P granule, as has been demonstrated in *prde-1(-)* mutants (Weick et al. 2014). If the piRNA depletion in *hrde-1(-)* mutants due to compromised PRG-1 function, we expect to see both decreased PRG-1 expression by western blot and mislocalization of PRG-1 by immunostaining.

If *hrde-1* functions primarily to regulate chromatin structure at the piRNA clusters, then we would expect it to regulate only those piRNAs that are expressed from the clusters. We will test this utilizing a strain in which we have inserted a single-copy, solitary piRNA by the Mos-1-mediated single copy insertion (MosSCI) method (Frøkjaer-Jensen et al. 2008). We previously demonstrated that the expression pattern of this solitary piRNA recapitulates that of endogenous piRNAs (Billi et al. 2013). We will cross the solitary piRNA into *hrde-1(-)* mutants and grow them for five generations at 25°C. We will then extract RNA from F4 adult worms and perform Taqman qRT-PCR for the transgenic piRNA and for a panel of endogenous piRNAs. We will us wildtype and *prg-1(-)* mutant worms as controls. If *hrde-1* is required exclusively for chromatin remodeling at the piRNA clusters, then only the endogenous, cluster-borne piRNAs will be depleted in *hrde-1(-)* mutants. The endogenous piRNAs that are not cluster-derived share the same regulatory elements, such as the Ruby motif, and biogenesis components. Thus, specificity of *hrde-1* regulation for cluster-borne piRNAs only would suggest that *hrde-1* regulates the genome architecture of the clusters to promote piRNA transcription.

Is piRNA transcription H3K9me3-dependent in C. elegans?

To test the next part of our model, we will investigate whether, as in *D. melanogaster,* the piRNA clusters must be H3K9 methylated to promote piRNA expression. First, we will quantify piRNA expression levels in the absence of global

H3K9 methylation by small RNA-seq in wildtype and *met-2(-);set-25(-)* mutants. *met-2* and *set-25* encode the two bonafide H3K9 histone methyltransferases in *C. elegans* (Towbin et al. 2012)*.* We anticipate that *met-2(-);set-25(-)* mutants will be depleted of piRNAs relative to wildtype. As piRNAs are critical for fertility, mutants that are depleted of piRNAs are also ts sterile, and thus any new factors implicated in the piRNA pathway should also be required for fertility. Indeed, *met-2(-);set-25(-)* mutants have a severe fertility defect at 25°C, consistent with a role for H3K9 methylation in piRNA expression (Zeller et al. 2016).

Second, we will ask whether the piRNA clusters are, in fact, enriched *for hrde-1-* dependent H3K9 methylation. Using our existing H3K9me3 ChIP-seq data sets from wildtype and *hrde-1(-)* worms at early and late generation at 25°C, we will answer the following questions: (1) are the piRNA clusters enriched for H3K9me3 relative to other parts of chromosome IV? (2) are the piRNAs clusters H3K9me3-depleted in late generation *hrde-1(-)* mutants versus wildtype or in late generation *hrde-1(-)* mutants versus early generation *hrde-1(-)* mutants? (3) Are *hrde-1*-dependent H3K9me3 loci enriched for piRNAs? Because *morc-1* is required for H3K9me3 maintenance exclusively at a subset of *hrde-1* targets but is not required for piRNA expression, we will use our *morc-1(-)* H3K9me3 ChIP-seq data as a control for this analysis. We will compare the average number of piRNAs at *hrde-1*-dependent, *morc-1*-independent loci compared to *morc-1*-dependent loci. We will also compare the proportion of *hrde-1-* dependent, *morc-1*-independent loci with over 10 piRNA reads per million (RPM) to the proportion of *morc-1*-dependent loci with over 10 piRNA RPM. If *hrde-1* regulates H3K9me3 levels at piRNA loci, we expect that both of these analyses will reveal much stronger enrichment of piRNAs targeting *hrde-1*-dependent, *morc-1*-independent H3K9me3 targets compared to *morc-1*-dependent H3K9me3 targets.

We expect to show that global H3K9 methylation is required for piRNA expression and that H3K9me3-enrichment at the piRNA clusters is *hrde-1*-dependent and *morc-1*-independent. Taken together, this will demonstrate that (1) *C. elegans* piRNAs, like *D. melanogaster* piRNAs, require H3K9 methylation for their expression and (2) a positive-feedback loop drives piRNA expression, siRNA biogenesis, and H3K9 trimethylation.

Possible H3K9me3-independent mechanisms for regulation of piRNAs by hrde-1

If *hrde-1* does not regulate H3K9me3 at piRNA loci, there are several additional mechanisms by which it could regulate piRNA expression. First, nuclear RNAi was recently shown to deposit H3K27me3 marks at target loci in addition to H3K9me3 (Mao et al. 2015). Thus, HRDE-1 could contribute to H3K27 methylation at the piRNA clusters. We will mine existing datasets from the modENCODE project to evaluate whether the piRNA clusters are H3K27me3-enriched compared to other parts of chromosome IV. We will also perform H3K27me3 ChIP-seq in early and late generation *hrde-1(-)* mutants and wildtype worms to evaluate the contribution of *hrde-1* to genome-wide H3K27 methylation. We will then look specifically at the piRNA clusters to see *hrde-1* is required for H3K27me3 across the piRNA clusters.

Second, our studies in Chapter Two showed that *hrde-1* is required for the localization of a heterochromatin reporter to the nuclear periphery. PRDE-1, which associates closely with the piRNA clusters, is also localized at the nuclear periphery in germline nuclei, suggesting that the piRNA clusters may be tethered to the nuclear periphery (Weick et al. 2014). *hrde-1* may contribute to piRNA expression by tethering or recruiting the piRNA clusters and/or piRNA biogenesis factors to the nuclear periphery to facilitate piRNA production. To determine whether *hrde-1* is required for localization of the piRNA clusters at the nuclear periphery, we will perform DNA FISH using probes to detect the piRNA clusters in wildtype, *hrde-1(-)*, and *morc-1(-)* mutants at early and late generation. To mark the nuclear periphery, we will perform these experiments in a LEM-2::GFP background. LEM-2 is a nuclear envelope component (Ikegami et al. 2010). As *morc-1(-)* mutants are not depleted of piRNAs, any defect in *hrde-1(-)* mutants that explains the observed piRNA depletion should not occur in *morc-1(-)* mutants. We showed in Chapter Two that *morc-1(-)* mutants are also required for localization of a heterochromatin reporter to the nuclear periphery. Thus, if we find that the piRNA clusters are similarly mislocalized in *hrde-1(-)* and *morc-1(-)* mutants, the mislocalization is likely due to a general defect in heterochromatin localization and not associated with piRNA depletion.

As a complementary approach, we will perform anti-LEM-2 ChIP-qPCR in wildtype, *hrde-1(-),* and *morc-1(-)* mutants and assay for LEM-2 enrichment at piRNA clusters compared to other parts of chromosome IV. If *hrde-1* contributes to piRNA cluster localization at the nuclear periphery, we would expect to see enrichment of piRNA cluster sequences in LEM-2 ChIPs in wildtype but not *hrde-1(-)* worms. We will also test whether *hrde-1* is required for the localization of established piRNA biogenesis factors to the nuclear periphery. Using CRISPR-Cas9, we will tag PRDE-1 and GEI-11 with mcherry and cross these reporters into the *lem-2::gfp* strain. We will examine localization of these factors in wildtype and *hrde-1(-)* mutants at early and late generations.

Third, HRDE-1 RISC could regulate RNA Pol II activity at the piRNA clusters. Nuclear RISC has been shown to cause RNA Pol II stalling at nascent transcripts (Guang et al. 2010). piRNAs are RNA Pol II-generated transcripts and, how RNA Pol II generates so many short (23-40nt) transcripts from these clusters is an outstanding question in the field (Cecere et al. 2012; Gu et al. 2012; Kasper et al. 2014). HRDE-1 RISC could regulate RNA Pol II activity at the piRNA clusters, potentially inducing stalling to promote transcriptional termination. We will test the contribution of *hrde-1* to RNA Pol II dynamics at the piRNA clusters by native elongating transcript sequencing (NET-seq), a high-resolution, crosslinking-independent method to deep-sequence nascent transcripts from their 3' end (Churchman and Weissman 2011). If *hrde-1* does indeed contribute to RNA Pol II stalling on nascent piRNA transcripts, we expect to see decreased RNA Pol II occupancy at the 3' end of annotated piRNAs (indicating decreased stalling) and potentially increased RNA Pol II occupancy at sequences that are directly downstream of piRNAs, indicating failed transcriptional termination in *hrde-1(-)* mutants.

Summary and significance

Our finding that *hrde-1(-)* mutants are depleted of piRNAs raises the intriguing possibility that HRDE-1 and its associated 22G RNAs promote piRNA expression,

creating a positive feedback loop between piRNAs and endo-siRNAs. The studies presented here will establish the mechanism by which HRDE-1 contributes to piRNA expression. Based on the established positive feedback loops between piRNAs and H3K9me3 in *D. melanogaster* and between siRNAs and H3K9me3 in *S. pombe,* we predict that HRDE-1 and its associated endo-siRNAs promote piRNA expression by directing H3K9 methylation, catalyzed by SET-25 and MET-2, at the piRNA clusters. The experiments proposed here will test this model, elucidating whether piRNA expression requires H3K9 methylation and addressing how *hrde-1* contributes to piRNA expression, including via chromatin modification over the piRNA clusters, tethering of the clusters to the nuclear periphery, and inducing RNA Pol II stalling during piRNA transcription. These studies will provide important insights into how piRNA and siRNA pathways are integrated to regulate germline gene expression, epigenetics, and genome defense.
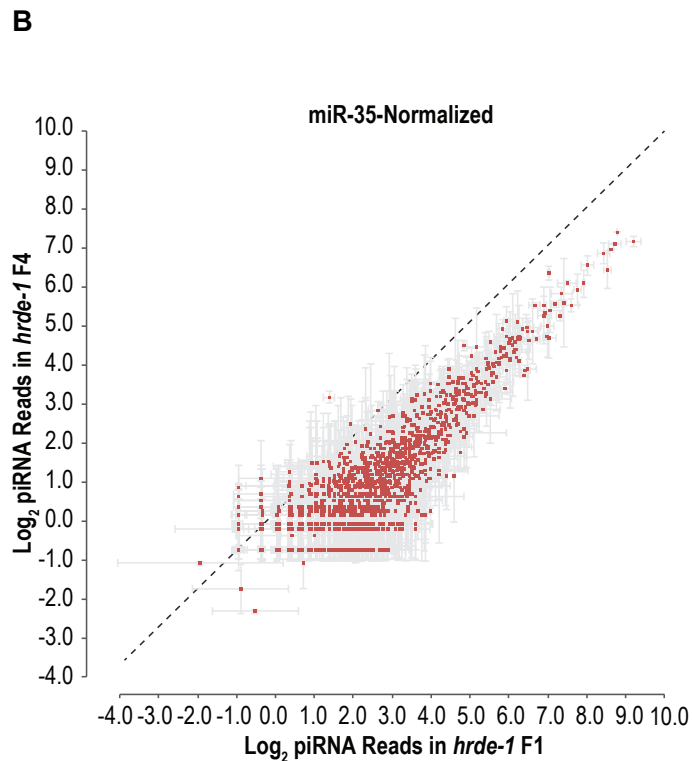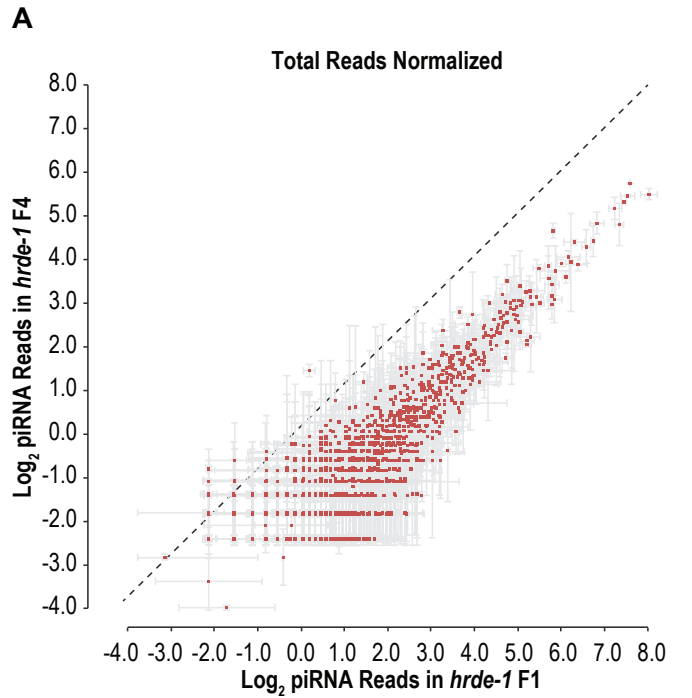
Figure 4.4. Progressive piRNA depletion in *hrde-1(-)* mutants. (A) piRNAs reads normalized to total library size (reads per kilobase of transcript per million mapped reads or RPKM) in late generation at 25°C vs early generation at 25°C. **(B)** piRNAs reads normalized to miR-35 levels in late generation at 25°C vs early generation at 25°C.

194

REFERENCES

Amrit FRG, Steenkiste EM, Ratnappan R, Chen S-W, McClendon TB, Kostka D, Yanowitz J, Olsen CP, Ghazi A. 2016. DAF-16 and TCER-1 Facilitate Adaptation to Germline Loss by Restoring Lipid Homeostasis and Repressing Reproductive Physiology in C. elegans. PLoS Genet. 12: e1005788.

Aravin A, Gaidatzis D, Pfeffer S, Lagos-Quintana M, Landgraf P, Iovino N, Morris P, Brownstein MJ, Kuramochi-Miyagawa S, Nakano T, Chien M, Russo JJ, et al. 2006. A novel class of small RNAs bind to MILI protein in mouse testes. Nature 442: 203–207.

Aravin AA, Sachidanandam R, Girard A, Fejes Toth K, Hannon GJ. 2007. Developmentally regulated piRNA clusters implicate MILI in transposon control. Science 316: 744–747.

Ashe A, Sapetschnig A, Weick E-M, Mitchell J, Bagijn MP, Cording AC, Doebley A-L, Goldstein LD, Lehrbach NJ, Le Pen J, Pintacuda G, Sakaguchi A, et al. 2012. piRNAs can trigger a multigenerational epigenetic memory in the germline of C. elegans. Cell 150: 88–99.

Bagijn MP, Goldstein LD, Sapetschnig A, Weick E-M, Bouasker S, Lehrbach NJ, Simard MJ, Miska EA. 2012. Function, targets, and evolution of Caenorhabditis elegans piRNAs. Science 337: 574–578.

Batista PJ, Ruby JG, Claycomb JM, Chiang R, Fahlgren N, Kasschau KD, Chaves DA, Gu W, Vasale JJ, Duan S, Conte D, Luo S, et al. 2008. PRG-1 and 21U-RNAs interact to form the piRNA complex required for fertility in C. elegans. Mol. Cell 31: 67–78.

Billi AC, Alessi AF, Khivansara V, Han T, Freeberg M, Mitani S, Kim JK. 2012. The Caenorhabditis elegans HEN1 ortholog, HENN-1, methylates and stabilizes select subclasses of germline small RNAs. PLoS Genet. 8: e1002617.

Billi AC, Freeberg MA, Day AM, Chun SY, Khivansara V, Kim JK. 2013. A conserved upstream motif orchestrates autonomous, germline-enriched expression of Caenorhabditis elegans piRNAs. PLoS Genet. 9: e1003392.

Brennecke J, Aravin AA, Stark A, Dus M, Kellis M, Sachidanandam R, Hannon GJ. 2007. Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. Cell 128: 1089–1103.

Brennecke J, Malone CD, Aravin AA, Sachidanandam R, Stark A, Hannon GJ. 2008. An epigenetic role for maternally inherited piRNAs in transposon silencing. Science 322: 1387–1392.

Cecere G, Zheng GXY, Mansisidor AR, Klymko KE, Grishok A. 2012. Promoters recognized by forkhead proteins exist for individual 21U-RNAs. Mol. Cell 47: 734–745.

Churchman LS, Weissman JS. 2011. Nascent transcript sequencing visualizes

transcription at nucleotide resolution. Nature 469: 368–373.

Conine CC, Batista PJ, Gu W, Claycomb JM, Chaves DA, Shirayama M, Mello CC. 2010. Argonautes ALG-3 and ALG-4 are required for spermatogenesis-specific 26G-RNAs and thermotolerant sperm in Caenorhabditis elegans. Proc. Natl. Acad. Sci. U.S.A. 107: 3588–3593.

Conine CC, Moresco JJ, Gu W, Shirayama M, Conte D, Yates JR, Mello CC. 2013. Argonautes promote male fertility and provide a paternal memory of germline gene expression in C. elegans. Cell 155: 1532–1544.

Das PP, Bagijn MP, Goldstein LD, Woolford JR, Lehrbach NJ, Sapetschnig A, Buhecha HR, Gilchrist MJ, Howe KL, Stark R, Matthews N, Berezikov E, et al. 2008. Piwi and piRNAs act upstream of an endogenous siRNA pathway to suppress Tc3 transposon mobility in the Caenorhabditis elegans germline. Mol. Cell 31: 79–90.

de Albuquerque BFM, Luteijn MJ, Cordeiro Rodrigues RJ, van Bergeijk P, Waaijers S, Kaaij LJT, Klein H, Boxem M, Ketting RF. 2014. PID-1 is a novel factor that operates during 21U-RNA biogenesis in Caenorhabditis elegans. Genes Dev. 28: 683–688.

de Albuquerque BFM, Placentino M, Ketting RF. 2015. Maternal piRNAs Are Essential for Germline Development following De Novo Establishment of Endo-siRNAs in Caenorhabditis elegans. Dev. Cell 34: 448–456.

Duchaine TF, Wohlschlegel JA, Kennedy S, Bei Y, Conte D, Pang K, Brownell DR, Harding S, Mitani S, Ruvkun G, Yates JR, Mello CC. 2006. Functional proteomics reveals the biochemical niche of C. elegans DCR-1 in multiple small-RNA-mediated pathways. Cell 124: 343–354.

Dumesic PA, Natarajan P, Chen C, Drinnenberg IA, Schiller BJ, Thompson J, Moresco JJ, Yates JR, Bartel DP, Madhani HD. 2013. Stalled spliceosomes are a signal for RNAi-mediated genome defense. Cell 152: 957–968.

Fischer SEJ, Montgomery TA, Zhang C, Fahlgren N, Breen PC, Hwang A, Sullivan CM, Carrington JC, Ruvkun G. 2011. The ERI-6/7 helicase acts at the first stage of an siRNA amplification pathway that targets recent gene duplications. PLoS Genet. 7: e1002369.

Frøkjaer-Jensen C, Davis MW, Hopkins CE, Newman BJ, Thummel JM, Olesen S-P, Grunnet M, Jorgensen EM. 2008. Single-copy insertion of transgenes in Caenorhabditis elegans. Nat. Genet. 40: 1375–1383.

Gent JI, Lamm AT, Pavelec DM, Maniar JM, Parameswaran P, Tao L, Kennedy S, Fire AZ. 2010. Distinct phases of siRNA synthesis in an endogenous RNAi pathway in C. elegans soma. Mol. Cell 37: 679–689.

Gent JI, Schvarzstein M, Villeneuve AM, Gu SG, Jantsch V, Fire AZ, Baudrimont A. 2009. A Caenorhabditis elegans RNA-directed RNA polymerase in sperm development and endogenous RNA interference. Genetics 183: 1297–1314.

Girard A, Sachidanandam R, Hannon GJ, Carmell MA. 2006. A germline-specific class of small RNAs binds mammalian Piwi proteins. Nature 442: 199–202.

Goh W-SS, Seah JWE, Harrison EJ, Chen C, Hammell CM, Hannon GJ. 2014. A genome-wide RNAi screen identifies factors required for distinct stages of C. elegans piRNA biogenesis. Genes Dev. 28: 797–807.

Goldstrohm AC, Albrecht TR, Suñé C, Bedford MT, Garcia-Blanco MA. 2001. The transcription elongation factor CA150 interacts with RNA polymerase II and the pre-mRNA splicing factor SF1. Mol. Cell. Biol. 21: 7617–7628.

Grivna ST, Beyret E, Wang Z, Lin H. 2006. A novel class of small RNAs in mouse spermatogenic cells. Genes Dev. 20: 1709–1714.

Gu W, Lee H-C, Chaves D, Youngman EM, Pazour GJ, Conte D, Mello CC. 2012. CapSeq and CIP-TAP Identify Pol II Start Sites and Reveal Capped Small RNAs as C. elegans piRNA Precursors. Cell 151: 1488–1500.

Guang S, Bochner AF, Burkhart KB, Burton N, Pavelec DM, Kennedy S. 2010. Small regulatory RNAs inhibit RNA polymerase II during the elongation phase of transcription. Nature 465: 1097–1101.

Han T, Manoharan AP, Harkins TT, Bouffard P, Fitzpatrick C, Chu DS, Thierry-Mieg D, Thierry-Mieg J, Kim JK. 2009. 26G endo-siRNAs regulate spermatogenic and zygotic gene expression in Caenorhabditis elegans. Proc. Natl. Acad. Sci. U.S.A. 106: 18674–18679.

Ikegami K, Egelhofer TA, Strome S, Lieb JD. 2010. Caenorhabditis elegans chromosome arms are anchored to the nuclear membrane via discontinuous association with LEM-2. Genome Biol. 11: R120.

Kamminga LM, van Wolfswinkel JC, Luteijn MJ, Kaaij LJT, Bagijn MP, Sapetschnig A, Miska EA, Berezikov E, Ketting RF. 2012. Differential impact of the HEN1 homolog HENN-1 on 21U and 26G RNAs in the germline of Caenorhabditis elegans. PLoS Genet. 8: e1002702.

Kasper DM, Wang G, Gardner KE, Johnstone TG, Reinke V. 2014. The C. elegans SNAPc component SNPC-4 coats piRNA domains and is globally required for piRNA abundance. Dev. Cell 31: 145–158.

Kim JK, Gabel HW, Kamath RS, Tewari M, Pasquinelli A, Rual J-F, Kennedy S, Dybbs M, Bertin N, Kaplan JM, Vidal M, Ruvkun G. 2005. Functional genomic analysis of RNA interference in C. elegans. Science 308: 1164–1167.

Klattenhoff C, Xi H, Li C, Lee S, Xu J, Khurana JS, Zhang F, Schultz N, Koppetsch BS, Nowosielska A, Seitz H, Zamore PD, et al. 2009. The Drosophila HP1 homolog Rhino is required for transposon silencing and piRNA production by dual-strand clusters. Cell 138: 1137–1149.

Le Thomas A, Marinov GK, Aravin AA. 2014a. A transgenerational process defines piRNA biogenesis in Drosophila virilis. Cell Rep 8: 1617–1623.

Le Thomas A, Stuwe E, Li S, Du J, Marinov G, Rozhkov N, Chen Y-CA, Luo Y, Sachidanandam R, Toth KF, Patel D, Aravin AA. 2014b. Transgenerationally inherited piRNAs trigger piRNA biogenesis by changing the chromatin of piRNA clusters and inducing precursor processing. Genes Dev. 28: 1667–1680.

Lee H-C, Gu W, Shirayama M, Youngman E, Conte D, Mello CC. 2012. C. elegans piRNAs mediate the genome-wide surveillance of germline transcripts. Cell 150: 78–87.

Lee NN, Chalamcharla VR, Reyes-Turcu F, Mehta S, Zofall M, Balachandran V, Dhakshnamoorthy J, Taneja N, Yamanaka S, Zhou M, Grewal SIS. 2013. Mtr4-like protein coordinates nuclear RNA processing for heterochromatin assembly and for telomere maintenance. Cell 155: 1061–1074.

LeGendre JB, Campbell ZT, Kroll-Conner P, Anderson P, Kimble J, Wickens M. 2013. RNA targets and specificity of Staufen, a double-stranded RNA-binding protein in Caenorhabditis elegans. J. Biol. Chem. 288: 2532–2545.

Luteijn MJ, van Bergeijk P, Kaaij LJT, Almeida MV, Roovers EF, Berezikov E, Ketting RF. 2012. Extremely stable Piwi-induced gene silencing in Caenorhabditis elegans. EMBO J. 31: 3422–3430.

Mao H, Zhu C, Zong D, Weng C, Yang X, Huang H, Liu D, Feng X, Guang S. 2015. The Nrde Pathway Mediates Small-RNA-Directed Histone H3 Lysine 27 Trimethylation in Caenorhabditis elegans. Curr. Biol. 25: 2398–2403.

Montgomery TA, Rim Y-S, Zhang C, Dowen RH, Phillips CM, Fischer SEJ, Ruvkun G. 2012. PIWI associated siRNAs and piRNAs specifically require the Caenorhabditis elegans HEN1 ortholog henn-1. PLoS Genet. 8: e1002616.

Motamedi MR, Verdel A, Colmenares SU, Gerber SA, Gygi SP, Moazed D. 2004. Two RNAi complexes, RITS and RDRC, physically interact and localize to noncoding centromeric RNAs. Cell 119: 789–802.

Noma K-I, Sugiyama T, Cam H, Verdel A, Zofall M, Jia S, Moazed D, Grewal SIS. 2004. RITS acts in cis to promote RNA interference-mediated transcriptional and post-transcriptional silencing. Nat. Genet. 36: 1174–1180.

Phillips CM, Brown KC, Montgomery BE, Ruvkun G, Montgomery TA. 2015. piRNAs and piRNA-Dependent siRNAs Protect Conserved and Essential C. elegans Genes from Misrouting into the RNAi Pathway. Dev. Cell 34: 457–465.

Pushpa K, Kumar GA, Subramaniam K. 2013. PUF-8 and TCER-1 are essential for normal levels of multiple mRNAs in the C. elegans germline. Development 140: 1312–1320.

Rangan P, Malone CD, Navarro C, Newbold SP, Hayes PS, Sachidanandam R, Hannon GJ, Lehmann R. 2011. piRNA production requires heterochromatin formation in Drosophila. Curr. Biol. 21: 1373–1379.

Robert VJP, Sijen T, van Wolfswinkel J, Plasterk RHA. 2005. Chromatin and RNAi factors protect the C. elegans germline against repetitive sequences. Genes Dev. 19: 782–787.

Ruby JG, Jan C, Player C, Axtell MJ, Lee W, Nusbaum C, Ge H, Bartel DP. 2006. Large-scale sequencing reveals 21U-RNAs and additional microRNAs and endogenous siRNAs in C. elegans. Cell 127: 1193–1207.

Sakai H, Sumitani M, Chikami Y, Yahata K, Uchino K, Kiuchi T, Katsuma S, Aoki F, Sezutsu H, Suzuki MG. 2016. Transgenic Expression of the piRNA-Resistant Masculinizer Gene Induces Female-Specific Lethality and Partial Female-to-Male Sex Reversal in the Silkworm, Bombyx mori. PLoS Genet. 12: e1006203.

Sánchez-Alvarez M, Goldstrohm AC, Garcia-Blanco MA, Suñé C. 2006. Human transcription elongation factor CA150 localizes to splicing factor-rich nuclear speckles and assembles transcription and splicing components into complexes through its amino and carboxyl regions. Mol. Cell. Biol. 26: 4998–5014.

Seth M, Shirayama M, Gu W, Ishidate T, Conte D, Mello CC. 2013. The C. elegans CSR-1 argonaute pathway counteracts epigenetic silencing to promote germline gene expression. Dev. Cell 27: 656–663.

Shirayama M, Seth M, Lee H-C, Gu W, Ishidate T, Conte D, Mello CC. 2012. piRNAs initiate an epigenetic memory of nonself RNA in the C. elegans germline. Cell 150: 65–77.

Simmer F, Tijsterman M, Parrish S, Koushika SP, Nonet ML, Fire A, Ahringer J, Plasterk RHA. 2002. Loss of the putative RNA-directed RNA polymerase RRF-3 makes C. elegans hypersensitive to RNAi. Curr. Biol. 12: 1317–1319.

Smith MJ, Kulkarni S, Pawson T. 2004. FF domains of CA150 bind transcription and splicing factors through multiple weak interactions. Mol. Cell. Biol. 24: 9274–9285.

Tabach Y, Billi AC, Hayes GD, Newman MA, Zuk O, Gabel H, Kamath R, Yacoby K, Chapman B, Garcia SM, Borowsky M, Kim JK, et al. 2013. Identification of small RNA pathway genes using patterns of phylogenetic conservation and divergence. Nature 493: 694–698.

Tang W, Tu S, Lee H-C, Weng Z, Mello CC. 2016. The RNase PARN-1 Trims piRNA 3' Ends to Promote Transcriptome Surveillance in C. elegans. Cell 164: 974–984.

Towbin BD, González-Aguilera C, Sack R, Gaidatzis D, Kalck V, Meister P, Askjaer P, Gasser SM. 2012. Step-wise methylation of histone H3K9 positions heterochromatin at the nuclear periphery. Cell 150: 934–947.

Vasale JJ, Gu W, Thivierge C, Batista PJ, Claycomb JM, Youngman EM, Duchaine TF, Mello CC, Conte D. 2010. Sequential rounds of RNA-dependent RNA transcription drive endogenous small-RNA biogenesis in the ERGO-1/Argonaute pathway. Proc. Natl. Acad. Sci. U.S.A. 107: 3582–3587.

Watanabe T, Takeda A, Tsukiyama T, Mise K, Okuno T, Sasaki H, Minami N, Imai H. 2006. Identification and characterization of two novel classes of small RNAs in the mouse germline: retrotransposon-derived siRNAs in oocytes and germline small RNAs in testes. Genes Dev. 20: 1732–1743.

Wedeles CJ, Wu MZ, Claycomb JM. 2013. Protection of germline gene expression by the C. elegans Argonaute CSR-1. Dev. Cell 27: 664–671.

Weick E-M, Sarkies P, Silva N, Chen RA, Moss SMM, Cording AC, Ahringer J, Martinez-Perez E, Miska EA. 2014. PRDE-1 is a nuclear factor essential for the biogenesis of Ruby motif-dependent piRNAs in C. elegans. Genes Dev. 28: 783–796.

Zeller P, Padeken J, van Schendel R, Kalck V, Tijsterman M, Gasser SM. 2016. Histone H3K9 methylation is dispensable for Caenorhabditis elegans development but suppresses RNA:DNA hybrid-associated repeat instability. Nat. Genet. 1–13.