ABSTRACT

Title of Dissertation:	INVESTIGATING GENE REGULATORY ARCHITECTURES THAT DICTATE TRANSGENERATIONAL EPIGENETIC EFFECTS IN C. ELEGANS
	Mary Somontha Chey, Doctor of Philosophy, 2023
Dissertation directed by:	Associate Professor, Dr. Antony Merlin Jose, Department of Cell Biology and Molecular Genetics

The form and function of an organism rely on the recreation of similar gene expression patterns in every generation. The information for these expression patterns is stored in a single cell (e.g., zygote) in two forms – the genome sequence and the spatial arrangements of gene regulators. The interactions of regulators and the genome form intricate networks with different regulatory architectures. However, a change in the environment can impact gene expression by disrupting the physical and/or chemical properties of regulators or interactions without mutations in DNA sequence. Such epigenetic information can be transmitted across generations, but how long these effects can last is unclear. Here, we investigate the regulatory elements that promote transient or permanent epigenetic effects by analyzing the properties of a recombinant two-gene operon that expresses the fluorescent proteins mCherry and GFP and is susceptible to long-term RNA silencing in the nematode *C. elegans*. We reveal that 1)

multiple mechanisms regulate transgenerational gene silencing and 2) the presence of the *mCherry* sequence can perturb RNA regulation within the germline to facilitate heritable epigenetic changes. Previous studies showed that the Argonaute protein HRDE-1 is required for the maintenance of silencing in the germline initiated by double-stranded (ds)RNAs and that poly-UG (pUG)-RNAs are key intermediates generated from the target mRNA. We found that loss of HRDE-1 can selectively rescue the expression of one cistron in a two-gene operon, suggesting that the two cistrons are not regulated by the same silencing pathway, but rather by a chromatin-independent mechanism that requires an unknown regulator. Surprisingly, we detected distinct populations of pUG-RNAs associated with expressed and silenced genes, suggesting that pUG-RNAs could potentially prime expressed genes for long-term silencing. Consistently, total RNA sequencing revealed trace amounts of anti-sense RNAs against *mCherry* and *gfp* that could trigger the production of pUG-RNAs. Examining the endogenous genes perturbed by the presence of *mCherry* suggests that long-term RNA silencing relies on the synergy between the sensing and processing of dsRNAs. Together, our results provide insights into the regulatory architectures and mechanisms of heritable gene silencing that occur without genetic mutations.

INVESTIGATING GENE REGULATORY ARCHITECTURES THAT DICTATE TRANSGENERATIONAL EPIGENETIC EFFECTS IN *C. ELEGANS*

by

Mary Somontha Chey

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2023

Advisory Committee:
Associate Professor Antony M. Jose, Chair
Professor Kan Cao
Professor Karen Carleton
Professor Eric Haag
Professor Zhongchi Liu

Dedication

I'd like to dedicate this thesis to all the loved ones I lost throughout this journey - paternal grandfather (February 1936 - July 2023), paternal grandmother (October 1939 - May 2016), maternal grandfather (1922 - March 2003), and maternal grandmother (March 1925 - April 2021).

To my beloved father (January 1963 - January 2022) and beloved mother (October 1953 -).

My parents are my biggest strengths in life and no words can describe how truly grateful I am for them and their sacrifices.

Acknowledgments

It has been a pleasure to have met so many brilliant and amazing individuals throughout my time in graduate school. I have grown exponentially as a scientist, scholar, and individual. The journey was not easy, but I am grateful to all those who have helped me arrive here today.

I would like to start by thanking the committee members for all their invaluable feedback throughout the years – Dr. Kan Cao, Dr. Karen Carleton, Dr. Eric Haag, and Dr. Zhongchi Liu.

I am very grateful to all the past and present members of the Jose Lab, especially Victoria Lowers. Victoria Lowers has gone above and beyond the role of a good lab citizen and demonstrates great kindness, strength, and intelligence that is unique and astounding. I am so honored to have met such a beautiful individual as herself. I am also honored and privileged to have mentored three bright individuals - Julianna Gross, Victoria Murphy, and Riley Nolan. I'd also like to thank my colleagues and friends – Ebony Argaez, Monica Cornejo, Mark de Jesus, Pablo Diaz, Spandan Pathak, Vanessa Rubio, Saul Sarria, Arvind Jaya Shankar, Anshuman Swain, Asan and Husan Turdiev, Liz Valverde, and Zaak Walton.

Thank you to the Baltimore Worm Club, the BISI office, especially Elizabeth Pepper and Dr. Zarna Rajeshkumar Pala, and the 2nd floor BRB custodians especially. Rosalba. I am also thankful for Rafael Antonio Segovia-Rodriguez and my beloved family and relatives.

I appreciate everyone's unwavering support, guidance, encouragement, love, and kindness throughout these years. I am grateful for all the beautiful friendships and relationships that have blossomed from this journey.

I'd like to thank Victoria Lowers, Vinashya Venkatasubramani, Zaak Walton, Mark de Jesus, Dr. Monica Cornejo, Dr. Sindhuja Devanapally, Dr. Narayanaswamy Bharathan, and Dr. Antony Jose for their comments, feedback, and suggestions on the writing.

I would like to give special thanks to Dr. Narayanaswamy Bharathan and Dr. Antony Jose.

Dr Narayanswamy Bharathan has been with me every step of the way since I was an undergraduate at Indiana University of Pennsylvania (IUP). He never stopped believing in me, even when I didn't believe in myself. He inspired me to begin this journey, to work hard, and to never give up. He encouraged me to be the best version of myself and to achieve my goals even in my darkest times. Thank you Dr. Bharathan for your mentorship, support, and compassion throughout the years. I am forever grateful.

Dr. Antony Jose, it has been a great honor to work with him and to be a part of his lab. I appreciate your patience, knowledge, professionalism, mentorship, and training throughout these years. Antony's persona exemplifies the key aspects of science – think critically, solve efficiently, write effectively, and communicate clearly. Antony is an amazing scientist and can only hope to have acquired some of his skills. Thank you for being a wonderful role model in the field.

Table of Contents

List of Tables

List of Figures

Chapter 1: *Caenorhabditis elegans* – a simple model organism for studying complex heritable epigenetic effects

1.1. Transgenerational effects can be triggered by the environment

The environment can trigger transgenerational effects that can be transmitted across generations. Exposure to stress or toxins and changes in diet have been reported to impact normal growth and development without any alterations to the DNA sequence. These epigenetic changes can be stored as changes in DNA methylation, histone methylation, or small RNA abundance and can be inherited (Cavalli & Heard, 2019).

Stress can trigger an epigenetic change by DNA methylation. In mice, when parents were exposed to an odor paired with an electric shock before conception, they observed enlarged glomeruli and hypomethylation of the odorant receptor gene *Olfr151* (Dias & Ressler, 2014). Through bisulfite sequencing of sperm in offspring never exposed to electrical stimuli, the hypomethylated gene was inherited and had enlarged glomeruli (up to F2). These offspring also showed enhanced sensitivity and not necessarily fear of the odor. Their data suggest that DNA methylation could be a potential regulator of epigenetic changes transmitted across generations.

Exposure to a toxic substance can also lead to changes in DNA methylation that affect sperm quality in descendants. One study observed that gestating female rats exposed to endocrine disrupters had >90% male descendants with decreased sperm quality and a higher incidence of infertility (observed up to F4) (Anway et. al., 2005). Such effects were also correlated with DNA methylation patterns in the male germ line.

One study analyzed the effects of ancestral food supply on descendants in an isolated community in northern Sweden. They observed that grandparents with either better or poor access to food supply during their "slow growth" period, the years leading up to the prepubertal growth spurt (males ages 9-12 and females ages 8-10), could either increase or decrease the mortality risk among grand progeny respectively; these heritable effects were also sex-specific (Pembrey et. al., 2005; Pembrey et. al., 2014). Paternal grandfather with poor access to food supply resulted in grandsons with a decreased mortality risk ratio and not granddaughters. Similarly, the grandmother's food supply affected only granddaughters and not grandsons. Although ancestral diet can affect the health risks in grand progeny, the molecular evidence for this effect is unavailable.

DNA methylation has been implicated as a potential regulator or remnant of an epigenetic change that can be transmitted across generations. However, another study revealed that a high-fat diet can impact the health of descendants without changes to DNA methylation, suggesting that DNA methylation alone cannot explain such heritable effects. This study subjected female mice to a high-fat diet and observed that the paternal lineage of male and female descendants (up to F3) was obese, resistant to insulin, and had addictive behaviors (Sarker et. al., 2018). To determine whether such heritable defects resulted in changes to DNA methylation and if those changes were inherited, they performed methylome profiling of sperm in F1s and F2s. They did not find any obvious differences in methylated DNA of offspring who were affected by their mother's diet and those that were not, suggesting that the inherited epigenetic

information affected by diet could be stored as chromatin modifications or changes in small RNAs that are independent of DNA methylation.

In these case studies, it is difficult to completely rule out DNA mutations within these complex systems. To expand on past research, simple organisms such as *C. elegans* will be used as a proxy to examine similar regulators and mechanisms used by mammalian systems.

1.2. The short lifespan of C. elegans makes it an attractive model for transgenerational studies

C. elegans is a simple transparent worm with a defined lineage. The exact pattern of divisions from zygote to adult can be traced and the primordial germ cells separate early from somatic cells during embryogenesis (Strome et. al., 2005). These animals grow to an average of 1 mm in length and exist predominantly as hermaphrodites. However, a frequency of one in 500 self-progenies can result in males (Hodgkin et. al., 1979). Hermaphrodites can produce their sperm and oocyte and thus self-fertilize making it easy to maintain strains of specific genotypes. The low frequency of males also makes it easy to introduce new genes and assess the gain or loss of function of a gene. Throughout development, the worm undergoes different larval stages which can be used to isolate worms of a specific stage (Cox et. al., 1981). Therefore, it makes it easy to distinguish one generation of progeny from the next. For example, isolating hermaphrodites at the L4 stage can give rise to the next generation of L4-staged hermaphrodites ~4 days later. Thus *C. elegans* is ideal for studying heritable effects of a single gene expression pattern.

1.3. Epigenetic inheritance needs to overcome transgenerational homeostasis

In most organisms, every generation begins at the single-cell bottleneck stage called the zygote. The information needed for the development of a similar organism in every generation consists of two kinds: (1) genetic information stored in the DNA sequence and (2) epigenetic information stored as regulators and their interactions such as proteins, RNAs, methylation marks, etc. (Fig. 1-1; Jose AM, 2020. Jr of the Royal Society Interface; Jose AM, 2020. *BioEssays*; Jose AM, 2018). These two stores of information presumably create a three-dimensional arrangement like the previous generation (Fig. 1A). This network of information can be defined as the regulatory architecture and must exist as positive feedback loops to ensure the recreation of an organism with similar form and function in every generation (Fig. 1-1B, C; Jose AM., 2018). Developmental reprogramming mechanisms are examples of positive feedback loops that are critical for the recreation of a zygote from one generation to the next (reviewed in Cantone et. al. 2013 Feng et. al., 2010).

Throughout the life cycle of an organism, the regulatory architectures can be altered by the environment without direct effects on the DNA sequence. These can be defined as epigenetic changes. However, some epigenetic changes can evade reprogramming mechanisms and can be transmitted to the next generation. For example, some loci reported in humans that are associated with neurological disorders can resist developmental reprogramming and can be inherited by the next generation (Tang et. al., 2015). It is unclear how some genes can escape these reprogramming mechanisms, but for any epigenetic change to cause long-lasting effects in descendants, these changes need to be integrated as a change in a positive feedback loop (Jose AM,

2023). To identify the extra-genomic molecules that affect these positive feedback loops, a single gene expression pattern that is susceptible to stable silencing needs to be extensively investigated.

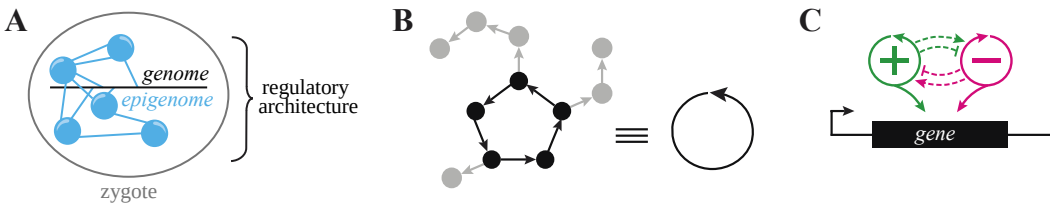

Figure 1-1. The faithful recreation of the single-cell zygote in every generation relies on the positive feedback loops in any regulatory architecture.

(**A**) The single-cell zygote contains the minimal information necessary for the reproducibility of a similar organism in every generation. This information is stored in two forms - the DNA sequence (genome) and the 3D arrangement of regulators (blue nodes) and their interactions (blue lines) that make up the epigenome. Together their interactions define the regulatory architecture(s) of an organism. (**B**) These regulatory architectures must exist as positive feedback loops (black nodes and black lines, equivalent to a black circle with an arrow on the right). Any changes outside of the feedback loops (gray nodes and gray lines) cannot be perpetuated across generations (as shown in Chey & Jose, 2021). (**C**) The expression of a gene can be affected by positive (green) and negative (pink) regulators that exist as independent positive feedback loops. Competition between two such feedback loops (dotted line with an arrow or dotted inhibition line) can determine the outcome of a gene – expressed or silenced (as shown in Chey & Jose, 2021).

1.4. Inherited RNA silencing in C. elegans can be used to investigate regulatory architectures that control gene expression across generations

Several advances in *C. elegans* have been made toward explaining heritable epigenetic changes of a single gene expression across generations. Discovery of RNA interference (RNAi) by injecting double-stranded (ds)RNA into the germline of *C. elegans* revealed a post-transcriptional mRNA-specific silencing mechanism (Fire et. al., 1998). Furthermore, transcriptomic analyses revealed another parallel silencing pathway regulated by small, 21 nucleotides (nts) RNAs with a 5' uracil bias end (21U-RNAs) called piwi-interacting (pi)RNAs (Batista et. al., 2008). Both dsRNA and piRNA triggered gene silencing is dependent on the continuous cycle of a few known regulators (Fig. 1-2; reviewed in Frolows & Ashe (2021) and Ketting & Cochella (2021)):

1) Poly-UG RNAs, which are a product of cleaved mRNA transcripts that are appended with dinucleotides - uracil and guanosine.
2) Secondary (2°)small RNAs, referred to as 22G-RNAs because of their 22 nt length and 5' guanosine bias. These are a product of unprimed amplification by an RNA-dependent RNA-polymerase (RdRp).
3) HRDE-1 bound 22G-RNAs, which can enter the nucleus and target pre-mRNA transcripts. HRDE-1 is a secondary Argonaute protein required for the maintenance of germline gene silencing across generations.

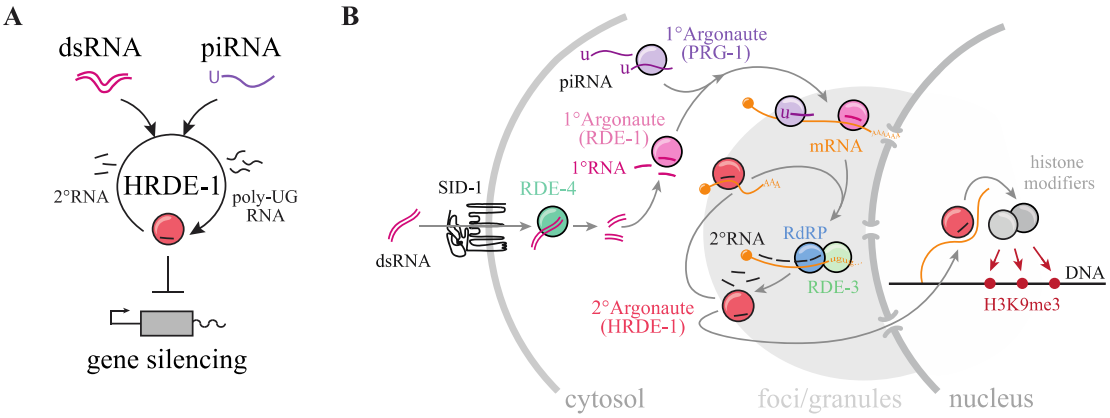

Figure 1- 2. Gene silencing by dsRNA and piRNA in the *C. elegans* germline.

(**A**) Simplified model depicting stable gene silencing that can be triggered by both exogenous double-stranded (ds)RNAs and endogenous piwi-interacting (pi)RNAs. Both pathways require the perpetuation of a positive feedback loop that is dependent on the secondary Argonaute HRDE-1 to maintain gene silencing across generations. Key intermediates of this cycle are secondary RNAs (2°RNA; e.g 22G-RNAs) and poly-UG (pUG)-RNAs. (**B**) Detailed schematic of the dsRNA- and piRNA-induced silencing pathways in the germline. Regulators of the dsRNA pathway include the following but are not limited to SID-1, RDE-4, RDE-1, RdRp, RDE-3, and HRDE-1. Regulators of the piRNA pathway include the following but are not limited to PRG-1, RdRp, RDE-3, and HRDE-1. Reviewed in Frolows & Ashe (2021) and Ketting & Cochella (2021).

Two more stable RNA silencing phenomena have been revealed in *C. elegans* which require similar dsRNA and piRNA regulators (Fig. 1-3). Although these phenomena have not been well-studied, both types of heritable silencing depend on the HRDE-1 silencing loop; loss of HRDE-1 could rescue expression. Although the triggers of such stable RNA silencing are still unknown, mating animals expressing a specific recombinant gene (Fig. 1-3A) can induce heritable RNA silencing and exposure to a silenced gene (Fig. 1-3B) can also. These two phenomena are referred to

as mating-induced silencing (Devanapally et. al., 2021) and *trans* silencing (Shirayama et. al, 2012).

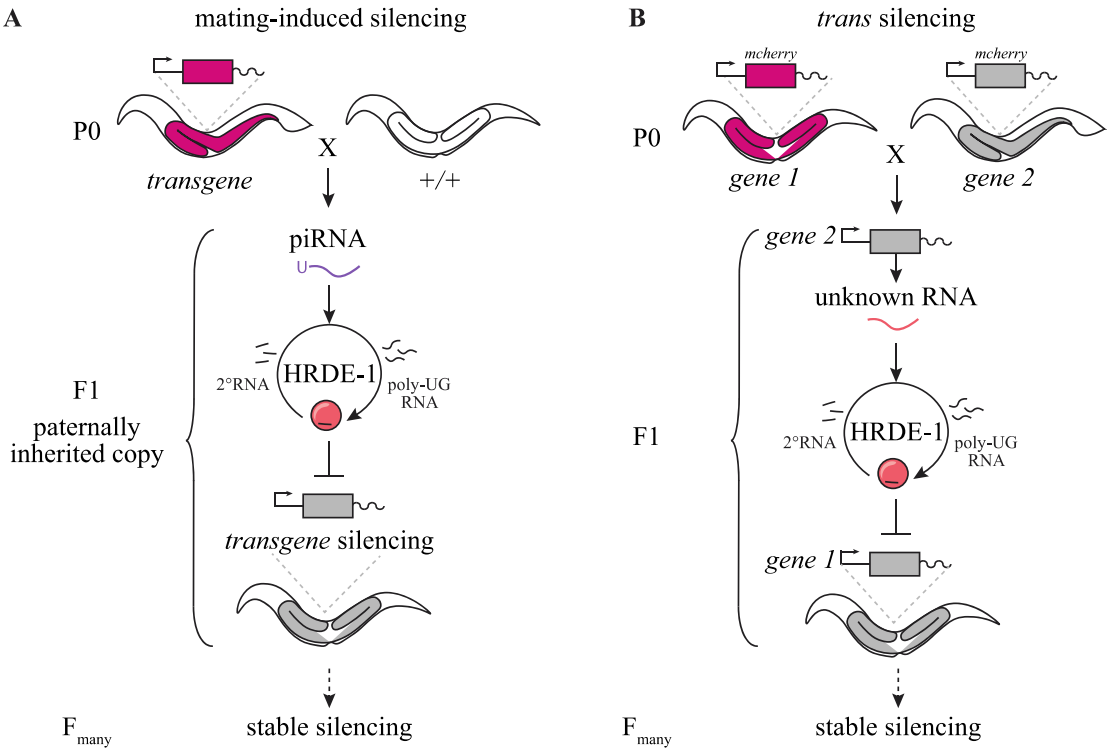

Figure 1- 3. Understudied stable RNA silencing phenomena in *C. elegans*.

Two stable RNA silencing phenomena reported in *C. elegans* that share similar regulators as dsRNA- and piRNA-induced silencing. (**A**) Simplified model of mating-induced silencing reported by Devanapally et. al., (2021). Males expressing a recombinant gene (*transgene*) crossed with wild type (+/+) yield F1 progeny that are silenced for the *transgene*. Although this silencing requires piRNA binding in addition to mating for the initiation of silencing, it is still unclear what other regulators can initiate this silencing and why silencing cannot occur in the reciprocal cross. Silencing, however, can be maintained across generations by the HRDE-1-dependent silencing loop. (**B**) Model of *trans* silencing. Two genes (*gene 1* and *gene 2*) share the same reporter gene *mCherry*. However, *gene 1* is expressed and *gene 2* is not. When animals containing the expressed *gene 1* are mated with animals containing the unexpressed *gene 2*, *gene 1* can now turn off by *trans* silencing signals emitted by the unexpressed *gene 2*. This stable RNA silencing can also be maintained by the HRDE-1-dependent silencing loop.

Changes to the environment of the worm can affect these types of silencing pathways. For example, exposure to high temperatures (Ni et. al., 2016) or starvation (Rechavi et. al., 2014) can increase small RNA abundance normally suppressed by RNAi regulators (e.g. HRDE-1 bound siRNAs) and increase the life span of worms across generations. Different types of chromatin-modifying enzymes can also impact the health and lifespan of worms in a context-dependent manner (Guillermo et. al. 2021). For example, active chromatin has been associated with H3K4me3 and inactive chromatin has been associated with H3K27me3 and H3K9me3. Loss of function of the H3K27 demethylase UTX-1 resulted in elevated levels of H3K27me3 marks and increased lifespan in descendants (Jin et. al., 2011; Maures et. al. 2011). Other molecular silencing candidates include modified mRNA fragments and interactions between nascent transcripts and small RNAs.

Despite such diversity of regulators, an unexplained feature of epigenetic mechanisms is the persistence or duration of the effect across generations. Three key questions described in Chey & Jose (2021) must be considered to mechanistically explain any heritable epigenetic effect:

i) How many generations can the effect last?

ii) In each generation, what regulatory molecules are required for the observable effect?

iii) Which regulatory interactions are required to explain the effect?

Addressing these questions will identify positive and negative regulators that impact at least one closed loop of mutual production (Fig. 1B and C). Any indefinite process relies on this type of positive feedback loop. Therefore, for any epigenetic change to persist, the overall activity of the loops formed by both positive and negative regulators must be affected. If the change affects regulators emanating the closed loops, then the effect cannot be permanent (Fig. 1B, Chey & Jose, 2021). Therefore, investigating a single gene expression pattern susceptible to heritable RNA silencing perturbed in different ways (by dsRNA, piRNA, mating, or exposure to a stably silenced epiallele) in *C. elegans* can provide insight into the positive feedback loops constrained by the dynamics of regulatory architectures.

Chapter 2: Evidence for multiple forms of heritable RNA silencing

2.1. Preface

Some of the gene sequences were obtained from Wormbase and the *sun-1::gfp* sequence for strain OCF62 from the Jantsch Laboratory (University of Vienna). Some strains were obtained from the *Caenorhabditis elegans* Genetic Stock Center (CGC), the Cohen-Fix Laboratory (National Institute of Health), and the Seydoux Laboratory (Johns Hopkins University). I contributed the remaining work in this chapter.

Mating can initiate stable RNA silencing in *C. elegans*

Part of this work builds on the findings published by: Devanapally S.*, Raman* P., **Chey M.**, Allgood S., Ettefa F., Diop M., Lin Y., Cho YE., Jose AM. Mating can initiate stable RNA silencing that overcomes epigenetic recovery. *Nature Communications* 12:4239. My contribution to this publication involved measuring pre-mRNA and mRNA levels of *mcherry* and *gfp* in different backgrounds using RT-qPCR and sequence verification of the *gfp* sequence in OCF62 (*sun-1::gfp)* (supplemental Fig. 2c and supplemental Fig. 6c in Devanapally et al., 2021). I also tested animals expressing mCherry fused to the N-terminus of the endogenous *mex-5* coding sequence for mating-induced silencing published in Fig. 1e in Devanapally et. al., (2021) and Figure 2-1 of this chapter. I used the strains generated in this publication to generate new data shown in this chapter.

Feeding RNAi can have different transgenerational effects

Pravrutha Raman and Farida Ettefa performed feeding RNAi experiments against the transgene *T* to generate data for Fig. 2-2. Rui Yin cloned three different two-gene constructs (i.e. *Tswap, Tcherry∆h2b, and Tind_adj*) which I injected them into *C. elegans* to generate integrations via MosSCI. These strains are in Figures 2-5 and 2-6 of this chapter. I also helped Farida clone the *mcherry* sequence to create a plasmid that expresses dsRNA against the half-length and full-length of the *mcherry* coding sequence. I imaged F1s upon *mcherry* feeding RNAi and these data were used to generate Fig. 2-2 of this chapter.

2.2. Introduction

Gene expression is spatially and temporally controlled by both genetic and epigenetic mechanisms. The experiences acquired in one generation can impact gene expression and yet a similar organism develops in the next generation because developmental processes can erase epigenetic information from the previous generation (Cavalli & Heard, 2019). However, it is unclear which epigenetic interactions and networks can evade reprogramming mechanisms to initiate and maintain gene silencing. Often the epigenetic machinery is broken up into three categories: 1) methylation of DNA, 2) change in chromatin structure, and 3) interaction with RNAs and RNA binding proteins (Mazzone et. al., 2019). However, the presence of one of these associated epigenetic marks such as methylation of CpG sites (C^5 position of cytosine residues) in DNA, methylation of histone tails such as lysine 9 on histone 3 (H3K9), or sequence-specific interaction with short interfering (e.g. microRNAs) or long non-coding RNAs (lncRNAs), cannot explain neurological,

behavioral, or genetic disorders observed in animals (Moosavi et. al. 2016; Mazzone et. al., 2019) or other phenomena such as paramutation in plants (Hollick et. al., 2017).

In humans, missense mutations of the highly conserved methyl-binding domain of the X-linked gene that encodes for methyl CpG-binding protein 2 (MeCP2) cause a cognitive disability called Rett Syndrome (Amir et. al.,1999). MeCP2 selectively binds CpG dinucleotides to transcriptionally repress the interaction between histone deacetylase and corepressor SIN3A. Loss of MeCP2 can vary in penetrance among females due to X inactivation, but in males, it can lead to severe congenital encephalopathies and early death (Mattick & Amaral, 2023). Such interplay between both DNA and histone methylation can critically impact brain development and function. To understand how the epigenetic machinery can impact gene expression, it is important to consider the interaction of regulators at all levels – DNA, RNA, and protein - throughout development.

In this chapter, we reveal three cases in which two genes of an operon can be regulated independently of one another. Such regulation, therefore, is likely to be chromatin-independent. The first case in which we observed such differences was upon feeding RNAi against *mCherry* and *gfp* in animals expressing the two-gene operon *T*. We unexpectedly observed non-sequence specific mCherry silencing in F3 upon anti-*gfp* dsRNA, suggesting that only the epigenetic mechanisms that initiate silencing seen in parents and early inheriting generations (up to F2) are a result of mRNA-specific silencing. In subsequent generations following F2, mCherry silencing suggests pre-mRNA silencing presumably because the two genes are expressed under the same germline promoter *mex-5*, and thus transcribed as a single pre-mRNA transcript. Our

data suggests that the transient effects in early inheriting generations and the transgenerational effects in later generations rely on different mechanisms. Therefore, further investigation is warranted.

We observed the second case of independent silencing of two cistrons through *trans* silencing. Prior work revealed that an epiallele with two silenced cistrons can turn off a homologous sequence that is expressed (Shirayama et. al., 2012). However, such silencing can be temporary. By F4 generation, the homologous sequence can recover its expression, and reactivate the expression of only the homologous cistron in the silenced epiallele. Our data suggests that the signals maintaining the silencing of each cistron in the epiallele are regulated differently and that such signals are mRNA-specific.

Loss of the regulator HRDE-1 represents the last case of independent silencing we observed. Previous studies have shown that this secondary Argonaute is required for gene silencing in the germline and loss of function can prevent silencing (Reviewed by Perez et. al., 2019). However, in this case, we observed that loss of HRDE-1 cannot completely rescue the expression of a germline gene. When examining the features of a silenced two-gene operon, the expression of only one cistron was rescued upon loss of HRDE-1 while the other cistron remained silenced. Our results suggest that the epigenetic mechanisms regulating one cistron are dependent on HRDE-1 but those regulating the other are not and require some unknown regulator.

2.3. Materials and Methods

2.3.1. Strains and oligonucleotides

Table 2-1. Strains

Strain	Shorthand name	Genotype/Allele	5'→ 3' sequence
N2		wild type	
EG6787	*T*	*oxSi487*	*mex-5p::mCherry::h2b::tbb-2 3'utr::gpd-2 operon::sl2::gfp::h2b::cye-1 3'utr + unc-119(+) II; unc-119(ed3) III?*
GE1708	*dpy-2(-)*	*dpy-2(e8) unc-4(e120)II*	
JH3270	*pgl-1::gfp*		*pgl-1p::pgl-1::gfp::pgl-1 3'utr*
AMJ799	*pie-1::gfp::PH*	*ltIs38*	*(pAA1) pie-1::GFP::PH(PLC1delta1) + unc-119(+)*
AMJ552	*iT*	*oxSi487 dpy-2(jam33) II; unc-119(ed3)? III*	*mex-5p::mCherry::h2b::tbb-2 3'utr::gpd-2 operon::sl2::gfp::h2b::cye-1 3'utr + unc-119(+) II; unc-119(ed3) III?*
AMJ581	*T dpy-2(-)*	*oxSi487 dpy-2(e8) II; unc-110(ed3) III*	*mex-5p::mCherry::h2b::tbb-2 3'utr::gpd-2 operon::sl2::gfp::h2b::cye-1 3'utr + unc-119(+) II; unc-119(ed3) III?*
AMJ844	*iT dpy-2(-)*	*oxSi487 dpy-2(e8) II; unc-119(ed3)? III*	*mex-5p::mCherry::h2b::tbb-2 3'utr::gpd-2 operon::sl2::gfp::h2b::cye-1 3'utr + unc-119(+) II; unc-119(ed3) III?*
AMJ1208	*mcherry::mex-5*	*jam197*	*mex-5p::mCherry::mex-5::mex-5 3'utr IV*
AMJ1209	*TcherryΔpi*	*jamSi39*	*mex-5p::mCherry (without piRNA sites)::cye-1 3' utr] II; unc-119(ed3) III*
AMJ1550	*gfp::mex-5*	*jam211*	*mex-5p::gfp::mex-5::mex-5 3'utr ch IV*
AMJ1557-59	*Tind_far; hrde-1(-)*	*Tcherry(jamSi56);*	*mex-5p::mcherry::cye-1 3'utr unc-119(+) ch I; mex-5p::gfp::cye-1 3'utr unc-*

		Tgfp(jamSi59); hrde-1(jam220)	*119(+) chII ; hrde-1[Gln234*C>T] unc-119(ed3) III*
AMJ1652/1653	*iT_re-inserted*	*jamSi83*	*mex-5p::mcherry::h2b::tbb-2 3'utr::gpd-2 operon::sl2::gfp::h2b::cye-1 3'utr + unc-119(+) II; unc-119(ed3) III?*
AMJ1628	*iTcherryΔh2b*	*jamSi78*	*mex-5p::mCherry::tbb-2 3'utr::gpd-2 operon::sl2::mCherry::h2b::cye-1 3'utr + unc-119(+) II; unc-119(ed3) III?*
AMJ1629	*iTswap*	*jamSi79*	*mex-5p::gfp::h2b::tbb-2 3'utr::gpd-2 operon::sl2::mCherry::h2b::cye-1 3'utr + unc-119(+) II; unc-119(ed3) III?*
AMJ1630	*iTind_adj*	*jamSi80*	*mex-5p::mCherry::h2b::tbb-2 3'utr gfp::h2b::cye-1 3'utr + unc-119(+) II; unc-119(ed3) III?*
AMJ1665	*T_re-inserted; hrde-1(-)*	*jamSi83; hrde-1(jam303)*	*mex-5p::mcherry::h2b::tbb-2 3'utr::gpd-2 operon::sl2::gfp::h2b::cye-1 3'utr + unc-119(+) II; unc-119(ed3) III?*
AMJ1666	*TcherryΔh2b; hrde-1(-)*	*jamSi67; hrde-1(jam304)*	*mex-5p::mCherry::tbb-2 3'utr::gpd-2 operon::sl2::mCherry::h2b::cye-1 3'utr + unc-119(+) II; unc-119(ed3) III?*
AMJ1429-31	*Tswap; hrde-1(-)*	*jamSi67; hrde-1(tm1200)*	*mex-5p::gfp::h2b::tbb-2 3'utr::gpd-2 operon::sl2::mCherry::h2b::cye-1 3'utr + unc-119(+) II; unc-119(ed3) III?*
AMJ1435-37	*Tind_adj; hrde-1(-)*	*jamSi68 ; hrde-1(tm1200)*	*mex-5p::mCherry::h2b::tbb-2 3'utr gfp::h2b::cye-1 3'utr + unc-119(+) II; unc-119(ed3) III?*

Table 2-2. Oligonucleotides

Name	Sequence 5'→3'	Genotyping

P1	cttaccggaaaccaaaggacg	Left MosSCI junction FOR
P2	taaggagttccacgcccag	Right MosSCI junction REV
P3	aattccacagttgctccgac	Left MosSCI junction REV
P4	tcacctagtctgtgccatttc	Right MosSCI junction FOR
P5	ataaggagttccacgcccag	G52
P6	ctagtgagtcgtattataagtg	G53
P7	tgaagacgacgagccacttg	G54
P8	gtcaaattcctgagctgcatgg	ttTi4348 Left Arm FOR
P9	ggcgcaagaactgcgattag	ttTi4348 Left Arm REV
P10	gagcagtacgagagagctgag	mcherry FOR
P12	gggaucagguaguggcaaag	*mex-5* CDS crRNA
P13	aaugaggauuacuuacaauu	*hrde-1* Q234* crRNA
P16	gaaaguuucaacgcguuuua	*rde-3/mut-2* ARG34* crRNA
P17	cauuauuauaggcgccgucu	*znfx-1* W250* crRNA
P18	cagacguuuggcuauacgcc	*deps-1* W157* crRNA
P19	cgaaaaguuuguugagcuca	*pgl-1* E154* crRNA
P20	attaattttatcgataatcaattgaatgtttcagacagagaatggtctc caagggagagg	*mcherry::mex-5* HRT FOR
P21	caagggatcaggtagtggcaaagcggcatcaaatagtgtctcgtc ggccggaggatcagt	*mcherry::mex-5* HRT REV
P22	tgatcctccggccgacgagacactatttgatgccgctttttgtatag ttcatccatg	*gfp::mex-5* HRT FOR
P23	aattttatcgataatcaattgaatgtttcagacagagaatgagtaaag gagaagaagttt	*gfp::mex-5* HRT REV
P24	aaagctcaacaatgaggattacttataatttggtaatgccacattccc acgtctttctcc	*hrde-1* Q234* HRT
P25	cgtttcacaaatttaacggaaagtttcaatgagttttaagggatcacg aggacgatttca	*rde-3/mut-2* ARG34* HRT
P26	agtaaacattattataggcgccgtcttagcatgatttctcgatggcttt cgatattgctg	*znfx-1* W250* HRT
P27	gcgacccgtgcggagccagacgtttgactatacgcctggattcga ttcgaaactaccatg	*deps-1* W157* HRT

P28	atgatgccgccgaaaagtttgtttagctcacggaacttgtcagcca actggatctggact	*pgl-1* E154* HRT
P29	gctatggctgttctcatggcggcgtcgccatattctacttcacacac acacacacaca	Universal RT primer with adapters for pUG PCR
P30	atggtctccaagggagag	*mCherry* FOR1 pUG PCR
P31	atgagtaaaggagaagaacttttca	*gfp* FOR1 pUG PCR
P32	gagttctacgatcacattct	*gsa-1* FOR1 pUG PCR
P33	gctatggctgttctcatggc	Adapter 1 pUG PCR 1 REV 1
P34	gagaggaggataacatggct	*mCherry* FOR2 pUG PCR
P35	Ttcactggagttgtccca	*gfp* FOR2 pUG PCR
P36	cacttgctggaaagacaagg	*gsa-1* FOR2 pUG PCR
P37	ggcgtcgccatattctactt	Adapter 2 pUG PCR 2 REV2

2.3.2. Feeding RNAi against *mcherry* and *gfp*

RNAi experiments were performed at 20°C on nematode growth media (NGM) plates supplemented with 1 mM IPTG (Omega Bio-Tek) and 25g/ml Carbenicillin (MP Biochemicals) (RNAi plates). Age-matched P0 animals were fed control RNAi (L4440) or *gfp*-dsRNA for 24 hrs (post-L4) and controls were scored alongside as a control. P0 animals and their untreated descendants were analyzed for mCherry and GFP expression as indicated in the schematic. L4 animals were scored in each generation by imaging and L4 siblings were passaged to obtain progeny for the next generation; except for F2 animals upon *gfp* RNAi, which was scored by eye.

2.3.3. Genetic crosses

Mating-induced silencing cross

About 9-12 L4-staged males containing either *gfp::mex-5* (AMJ1550) or *mcherry::mex-5* (AMJ1208) were placed on NG plates seeded with OP50 bacteria. Three L4-staged wild-type hermaphrodites were placed on the same plates containing the males. Plates were kept at 20°C and animals were allowed to mate. About 4 to 5 days later, L4 cross progeny males were picked and imaged. Similarly, 9-12 L4 N2 males were placed on seeded NG plates and 3 L4 hermaphrodites with *gfp::mex-5* or *mcherry::mex-5* were placed on the same plates and L4 cross progeny were picked and imaged. In the case in which cross-progeny hermaphrodites were imaged, *dpy-2*(-) (GE1708) hermaphrodites were used to distinguish self-progeny from cross-progeny. L4- staged cross progeny hermaphrodites were isolated some were imaged and others were imaged 24-48 hours post-L4.

Protection against mating-induced silencing cross

About 9-12 L4-staged males expressing *T* (EG6787) were allowed to mate with 3 hermaphrodites expressing mCherry::MEX-5 (AMJ1208). About 4-5 days later, L4-staged cross-progeny males were isolated and imaged.

Trans silencing cross of pgl-1::gfp and pie-1::gfp::PH

About 12 L4 males containing *pgl-1::gfp* were crossed with hermaphrodites containing the silenced state of *T* referred to as *iT*. *iT* was linked with a recessive *dpy-2(-)* mutation (AMJ844) for ease of genotyping. F1 L4-staged cross-progeny males

were isolated and imaged. F1 L4 nonDpy hermaphrodites were isolated and allowed to lay progeny. F2-F4 generations were genotyped for homozygosity of *pgl-1::gfp;iTdpy-2(-)*. F5 L4 homozygous hermaphrodites were isolated for imaging and others were isolated for the second cross where they were allowed to mate with 9-12 L4 males containing *pie-1::gfp::PH*. These F1 L4-staged nonDpy hermaphrodites were isolated and imaged 24-48 hours post L4.

Trans silencing cross of gfp::mex-5 and mcherry::mex-5

About 9 L4 males containing either *gfp::mex-5* or *mcherry::mex-5* were crossed with 3 hermaphrodites containing either *dpy-2(-)* (GE1708) or *iT dpy-2(-)* (AMJ844). F1 L4-staged nonDpy hermaphrodites and males were isolated and either imaged at L4 or 24-48 hours post L4. F1 nonDpy hermaphrodites were also singled out and allowed to lay self-progeny. F2 hermaphrodites were singled out and allowed to lay progeny. About 3.5-4 days later F2 moms were lysed and genotyped for homozygosity of *gfp::mex-5; iT dpy-2(-), mcherry::mex-5; iT dpy-2(-), iT dpy-2(-), gfp::mex-5* and *mcherry::mex-5*. F3s were singled out, allowed to lay progeny, and re-genotyped to confirm homozygosity. F4 L4-staged hermaphrodites were isolated and imaged at either L4 or 24-72 hrs post-L4. After confirming homozygosity, 3 L4 hermaphrodites were passaged in subsequent generations and L4-s were picked and imaged 24-72 hrs post L4.

Cross to generate animals expressing Tcherry I; Tgfp II

Both *Tcherry I* (AMJ1338) and *Tgfp II* (AMJ1195) are susceptible to mating-induced silencing. To prevent silencing, males expressing either *Tcherry* or *Tgfp* were crossed with hermaphrodites expressing *T*. Three F1 cross progeny males which should be heterozygous for *Tcherry/+; T/+* or *Tgfp/T* were used to set up a second cross with one heterozygous hermaphrodite containing either *Tgfp/T* or *Tcherry/+; T/+* respectively. A few days later after time was given for them to mate and lay progeny, the F2 moms were lysed and genotyped using triplex PCR for either *Tgfp* (P5, P6, and P7) or *Tcherry* (P8, P9, and P10). Subsequent generations were singled out and genotyped for homozygosity. After ~5 generations, animals now homozygous for *Tcherry; Tgfp* were imaged. Initially, I observed the expression of both reporters. However, some progeny in these generations exhibited no mCherry expression. After several generations of passaging, progeny were isolated and imaged. All homozygotes lacked mCherry expression but retained GFP expression.

2.3.3. Genome editing by MosSCI and CRISPR

Settings for needle pulling:

Program 10, P=500, Heat=510, Pull=60, Vel.=60, Time=250. Needle pull time ranged between 7-9 seconds.

Mos-1 Single-Copy Insertion

About 10-25 L4-staged animals containing the Mos1 gene on *ttTi5605* and *unc-119(ed4)* (EG4322) were isolated. Animals were injected 24 hours post-L4. The injection mix of 10 μl included the following: pCFJ601 (50 ng/ul), pMA122 (10 ng/ul),

and plasmid-containing gene of interest (55-60 ng/ul). The injection mix was centrifuged at 14,000 rpm for 30 min. 1.5 ul of the mix was loaded onto the needle. A needle pressure of 30 psi was used. After injections, animals were placed at 20°C to recover. After ~10 days when animals starved, injected plates were screened for nonUncs. nonUncs were subject to heat shock in a water bath at 34°C for 2.5 hrs. They were then placed at 20°C to recover. After 2-3 days, nonUnc heat-shock survivors were isolated onto freshly seeded NG plates and allowed to lay progeny. Moms were lysed and lysates were split three ways to genotype for the left junction (P1 and P3) and right junction (P4 and P2) of integration, as well as homozygosity (P1 and P2). After successful genotyping, gDNA was prepped from a whole plate of worms. Worms were washed with M9 buffer 3X and spun down at 14,000 rpm for 30 sec between each wash. Phenol-chloroform was used to extract gDNA. Different primer sets were used to amplify the entire sequence by PCR. PCR products were confirmed by gel-electrophoresis and PCR products were purified by nitrocellulose membrane. Samples were sent for Sanger Sequencing by Genewiz from Azenta Life Sciences.

Cas-9-mediated genome editing

The injection mix of 10 µl included the following: 0.4 ug/ul of tracrRNA, ~4.2 pmol/ul crRNA, 0.35 pmol/ul *dpy-10* crRNA or 40 ng/ul pRF4 *rol-6,* and 1.6 pmol/µl Cas9 protein (PNA Bio catalog no. CP01). Homology templates to generate *dpy-10(-)* were single-stranded DNA oligos. However, homology templates to generate mCherry and GFP N-terminus fusion proteins were amplified from plasmids containing the reporter sequence using Phusion High Fidelity polymerase (New England Biolabs

catalog no. M0530S) and gene-specific primers with ~35 bp of overhang homologous to the site of integration. PCR products were purified using NucleoSpin Gel and PCR Clean-up (Macherey-Nagel, catalog no. 740609.250).

2.3.4. Imaging

Animals of the L4 stage or between 24-72 hours after the L4 stage were mounted on a slide after paralyzing the worm using 3 mM levamisole (Sigma-Aldrich, Cat# 196142), imaged under non-saturating conditions (Nikon AZ100 microscope and Photometrics Cool SNAP HQ^2 camera), and binned into two (on or off) or three (bright, dim and off) groups. In the case of scoring mCherry::MEX-5 expression upon mating-induced silencing or *trans* silencing, the LUTs were used as a cut-off for scoring. Expression visible between the range of 0-5,000 was categorized as "on"; expression outside of this range was categorized as "off'. C-HGFI Intensilight Hg Illuminator was used to excite GFP (filter cube: 450 to 490 nm excitation, 495 dichroic, and 500 to 550 nm emission) or mCherry (filter cube: 530 to 560 nm excitation, 570 dichroic, and 590 to 650 nm emission). Sections of the gonad that are not obscured by autofluorescence from the intestine were examined to classify GFP and mCherry fluorescence. Autofluorescence was more appreciable when imaging GFP than it was when imaging mCherry. In the case of scoring for expression of GFP::MEX-5 upon *trans* silencing, even after adjusting the LUTs, it was still difficult to determine whether the expression should be categorized as dim or off.

In some cases, fluorescence intensity within the germline was scored by eye without imaging at the L4 stage in Fig. (F2 animals only) at fixed magnification and zoom using the Olympus MVX10 fluorescent microscope.

2.3.4. Statistics

For each figure, $\chi 2$ test was used to compare data as indicated in figure legends except in cases where only one category (on or off) was present in both datasets being compared, in which case Wilson's estimates for proportions were used. For RT-qPCR, significance was compared using a student's t-test of unequal variance. P-values are indicated in the legend for each figure.

2.3.5. RT-PCR of poly-UG RNAs

Worms containing different genotypes were grown on 35 mm NG plates at 20°C with 100 μl of seeded OP50. Animals were allowed to grow and crowd. Just before starvation, worms were washed with M9 buffer 3X and centrifuged at 14,000 rpm for 30 sec each time. About 50 μl of worm pellets was left. 1 mL of TRIzol (Fisher Scientific) was added to the worm pellet and worms were subject to freeze and thaw 3X using liquid nitrogen. RNA was extracted according to manufacturing protocol using chloroform. RNA was precipitated in 10 ug/ul of glycogen. Reverse transcription was performed using Superscript III (Invitrogen) and RT primer P29 according to manufacturing protocol. A total of ~1000-5000 ng of total RNA was used. About 2 μl of cDNA was used for nested PCR. The Phusion Taq Polymerase was used (NEB) for the first round of PCR. A final reaction volume of 20 μl was used and the PCR cycling condition was set for 20-25 cycles. The first set of gene-specific forward primers were used for each target gene (*mcherry* P30, *gfp* P31, and *gsa-1* P32) with the adapter 1 reverse primer (P33). The PCR product from round 1 PCR was diluted by 1:100. About 1 μl of the diluted PCR product was used for the second round of PCR in a total reaction volume of 50 μl. The second set of gene-specific forward primers (*mcherry* P34, *gfp*

P35, and *gsa-1* P36) were used with adapter 2 reverse primer (P37). About 10-20 μl of PCR was loaded with 6X loading dye on a 1% agarose gel and imaged using gel documentation (BioRad).

2.4. Results

Prior work in the Jose Lab serendipitously discovered a new phenomenon termed mating-induced silencing (Devanapally et. al. 2021). Here, permanent RNA silencing was initiated when progeny inherited only the paternal copy of the single-copy transgene referred to as *T*. This transgene was integrated into the genome on chromosome II using a method called Mos-1-mediated Single-Copy Insertion (MosSCI) (Frojkaer-Jensen et al. 2012; Frojkaer-Jensen et al. 2008). It is an operon that is expressed in the germline by an additional copy of the **m**uscle **ex**cess (*mex)-5* promoter sequence. It contains two cistrons with different reporter genes: a) *mCherry* fused to *h2b* followed by the *tbb-2 3'utr* sequence and *gpd-2* operon sequence containing the splice leader head 2 (SL2), and b) *gfp* fused to *h2b* followed by the *cye-1 3'utr sequence* (Fig. 2-1A). However, even if a part of the coding sequence is present in the hermaphrodites used to cross with males expressing the transgene, mating cannot initiate silencing in progeny despite inheriting only the paternal copy. Dissecting the sequences to create a minimal version of the transgene revealed that the presence of either *mCherry* or *gfp* alone under the control of the *mex-5* promoter could still leave the gene susceptible to such permanent RNA silencing, even with changes to the 3'utr sequence. However, the *mCherry* or *gfp* sequence alone cannot explain such susceptibility, because fusing the same *mCherry* and *gfp coding sequence* to the N-

terminus of the endogenous *mex-5* locus resulted in transgenic animals that were resistant to mating-induced silencing (Fig. 2-1A and B). Furthermore, we observed that despite resistance to mating-induced silencing, the endogenous mCherry tag (*mCherry::mex-5*) could still protect against silencing (Fig. 2-1C, bottom cross). Finally, the removal of the *mCherry* open-reading frame present in hermaphrodites could not protect progeny that inherits only the paternal copy from silencing (Devanapally et. al., 2021).

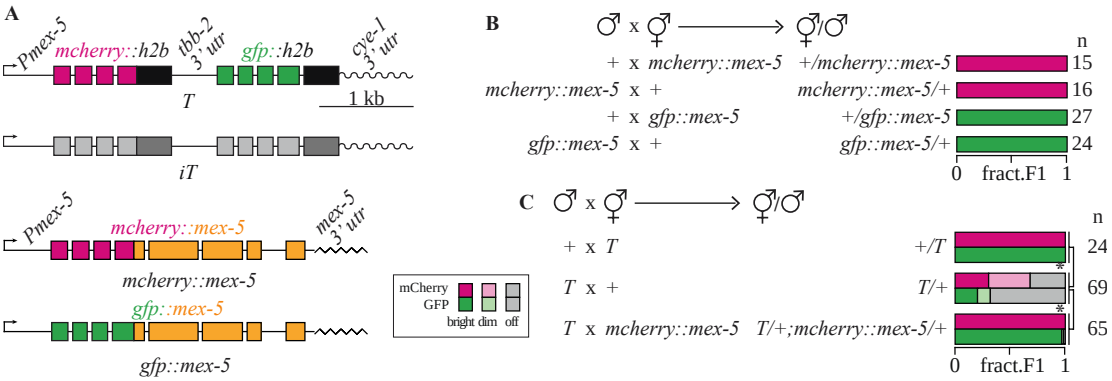

Figure 2-1. *Cis*-regulatory sequences influence susceptibility to transgenerational RNA silencing upon mating.

(**A**) Schematics of the two-gene operon referred to as *T*, its silenced epiallele *iT*, and fusion tag of *mCherry* and *gfp* to the N-terminus of the endogenous *mex-5* locus. (**B**) Genetic cross to test the susceptibility of *mCherry::mex-5* and *gfp::mex-5* to mating-induced silencing. Progeny that inherits only the paternal copy show silencing. However, in both gene tags, none of the progeny that inherit only the paternal copy (second and fourth cross) are silenced. A similar expression was seen for both genes in the control cross (first and third cross). (**C**) Genetic cross to test protection against mating-induced silencing. Progeny that inherits the paternal copy of *T* are susceptible to silencing (second cross). However, the presence of *mCherry* in hermaphrodites can prevent the silencing of *T* in progeny (third cross). In (B) and (C) cross progenies were imaged and scored for mCherry and GFP expression. Expression was categorized as either bright, dim, or off (box). Asterisks indicates p-value <0.05 (χ^2 test). Number of progenies assayed = n.

2.4.1. dsRNA can trigger both mRNA sequence-specific and pre-mRNA sequence-specific transgenerational silencing of a two-gene operon

Double-stranded (ds)RNA can silence a gene of matching sequence when delivered by injection (Fire et al. 1998) or ingestion (Timmons et. al., 2001; Timmons et al., 1998). However, such silencing can either last for a few or many generations despite using the same coding sequence and trigger (Devanapally et. al., 2021). It is unclear which factors regulate the duration of heritable silencing. By observing a single gene expression pattern across generations, we can begin to understand these factors.

The transgene *T* is susceptible to heritable RNA silencing by feeding RNA interference (RNAi) where silencing of GFP was observed to persist for more than five generations (Devanapally et. al., 2021). Feeding RNAi relies on animals ingesting an engineered bacterium that expresses dsRNA against a target gene. The bacterial strain is a non-pathogenic derivative of *E.coli* which lacks the dsRNA-specific endonuclease (RNaseIII) activity encoded by the *rnc* gene (Timmons et. al., 2001). This prevents the degradation of dsRNAs within the bacterial cells to increase optimum dsRNA expression for ingestion by the nematode. This bacterium contains an engineered plasmid with two convergent T7 polymerase promoters in opposite orientations that require isopropyl β-D-1-thiogalactopyranoside (IPTG) for induction (Timmons et. al., 2001). Between the T7 promoters is a multiple cloning site and the gene sequence to be targeted which will create two strands in both directions and form a dsRNA sequence. To improve bacterial growth and dsRNA induction, the presence of both tetracycline and ampicillin is also included in the plasmid. Tetracycline selects against *rnc*$^{+}$ in host cells and ampicillin selects against cells that lost the plasmid. Upon seeding

the bacterium on nutrient grow medium (NGM) plates infused with IPTG and the respective antibiotics, the bacteria will grow and express dsRNA against the target sequence.

We observed two different heritable RNA silencing effects upon feeding RNAi targeting the exon sequences of *mCherry* and *gfp* in animals expressing *T* (Fig. 2-2A). The first effect was seen when animals were fed *mCherry* RNAi for 24 hours. Both mCherry and GFP expression was off in parents and such silencing could persist across generations up to F8 (Fig. 2-2B). The second effect was seen when animals were fed *gfp* RNAi. GFP expression was off in parents and progeny up to F2, but mCherry expression began to diminish by F3 (Fig. 2-2C). The fraction of animals that were either bright, dim, or off for both mCherry and GFP were similar across generations. However, in two out of the three parent lines tested, progeny showed recovery of gene expression and in the other case, the majority of progeny remained stably off (Fig. 2-2C).

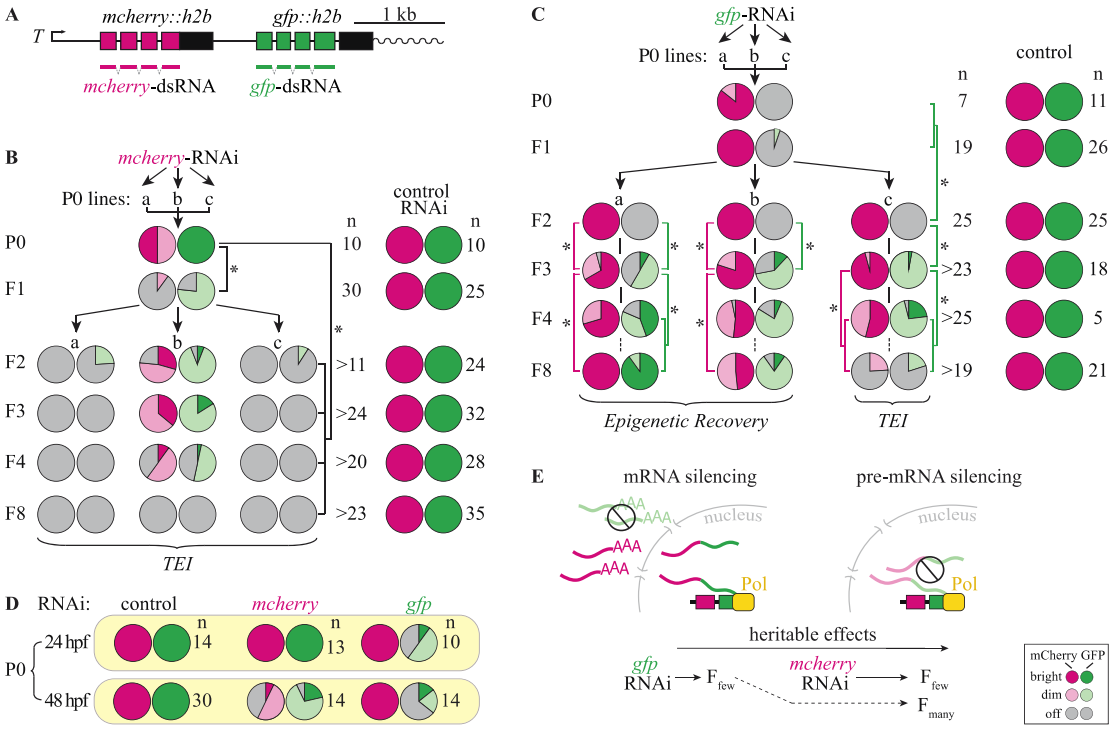

Figure 2-2. Feeding RNAi against one cistron reveals a switch from co-transcriptional to post-transcriptional mechanisms across generations.

(**A**) Schematic of *T* and dsRNA matching the exon sequences of *mcherry* (maroon) and *gfp* (green). (**B-D**) P0s expressing *T* were exposed to L4440 control RNAi and *mcherry* RNAi or *gfp* RNAi for 24 hrs. P0s and descendants were imaged and analyzed for both mCherry and GFP expression. Expression was binned as either bright, dim, or off (box). In (B) three animals were blindly propagated in every generation. In (**B**) and (**C**) F1s were pooled from three separate P0 parent plates (a, b, c). In (**D**) P0s were imaged an additional 24hrs post feed (48 hpf). (**E**) Model depicting post-transcriptional mechanism that inhibits mRNA transcripts and co-transcriptional mechanism that inhibits pre-mRNA transcripts. Feeding RNAi against *gfp* results in mRNA-level silencing observed between P0-F2. Loss of mCherry expression in F3 suggests a switch to pre-mRNA silencing in F3 from mRNA silencing in F2. In contrast, upon *mcherry* RNAi, expression of both mCherry and GFP expression was reduced immediately from F1, suggesting that transgenerational silencing engaged pre-mRNA regulation earlier. All progenies were propagated in an unbiased manner. Number of animals assayed = n. Asterisks indicate p-value <0.05 from χ^2 tests. The population of descendants that recovered mCherry and GFP expression are labeled with "Epigenetic Recovery" and the population of descendants that remain silenced are labeled with "Transgenerational Epigenetic Inheritance (TEI)".

Our data suggests that before splicing, the RNA transcribed contains both *mCherry* and *gfp* as a single pre-mRNA transcript. Upon ingestion of anti-*mcherry* dsRNAs, there is a higher degree of pre-mRNA silencing, therefore we see no expression for both mCherry and GFP (seen 48 hrs post feeding RNAi because of protein perdurance, Fig. 2-2D). Exposure to anti-*gfp* dsRNAs resulted in mRNA silencing because there was no GFP expression in earlier generations (P0-F2). We speculate that there could be some degree of pre-mRNA silencing within P0-F2 but not to the extent where both genes are affected. By F3, there is evidence for pre-mRNA silencing because both genes were unexpressed. Therefore, both post-transcriptional and co-transcriptional level silencing are present upon feeding RNAi, but this switch in regulation is seen more clearly upon feeding RNAi against the second cistron. However, it is unclear which regulators function between generations to control mRNA and pre-mRNA level silencing. This raises the question – why do we see such co-transcriptional silencing when the first cistron in the operon is targeted and not when the second cistron is targeted? What regulators are recruited for *gfp* mRNA silencing upon *gfp*-dsRNA that are then accompanied by *mCherry*-dependent regulators? How are the *mCherry* regulators upon *gfp*-RNAi different from the regulators upon *mCherry*-RNAi?

2.4.2. A gene that can recover expression can re-activate the expression of a homologous epiallele that has been silenced for many generations

Silencing by dsRNA can induce heritable silencing of the target gene. Another type of heritable silencing known in the field is referred to as *trans* silencing (Fig 1s-3B; Shirayama et. al., 2012). A gene is susceptible to silencing when exposed to an

inactive gene that has homologous sequences. The extent of such silencing is dependent on the percentage of homology. For example, when the gene that is stably silenced referred to as *iT* was crossed to the red fluorescent protein (RFP) that had the longest match of 13 nt and percent identity of ~68% to the *mCherry* sequence in *iT*, a few to no progeny were off for RFP. However, in the case where *iT* was crossed to animals expressing *gfp* tagged to the N-terminus of the *gtbp-1* coding sequence with the longest nt match of 698 and 99.9 percent identity to *gfp* in *iT*, more than 50% of progeny were off for *gtbp-1::gfp.* Additionally, we observed that animals expressing *pgl-1::gfp* were also susceptible to *trans* silencing in more than 50% of the progeny (Fig. 2-3A, also observed in Devanapally et. al., 2021 Fig. 5B). In both cases, silencing was transient. By F5, when animals no longer contain the presence of the silenced epiallele, homozygous animals expressing either gene could recover expression (Fig 2-3A, recovery seen as early as F3 in Devanapally et. al., 2021 Fig. 5B).

These data suggest that when *trans* silencing is initiated in F1, the gene can recover expression and resist further silencing despite the presence of the trigger *iT* in the background. Potentially the epigenetic regulators maintaining the silencing of *iT* are different from the silencing signals transmitted to the locus of the other homologous gene. Whether those *trans* silencing signals were lost by the time the gene recovered expression or if other positive regulators protected the homologous signal from silencing the gene is unclear.

To determine whether *iT* can re-initiate *trans* silencing despite the recovery of PGL-1::GFP expression and resistance to re-silencing, we crossed animals homozygous for *pgl-1::gfp; iT* to animals expressing another homologous *gfp*

sequence, *pie-1::gfp::PH*. We observed that the silenced *iT* could no longer initiate silencing of the other homologous expressed gene *pie-1::gfp::PH* when in the recovered and resistant background of *pgl-1::gfp* (Fig. 2-3A). But *iT* can initiate *trans* silencing of *pie-1::gfp::PH* if it did not go through a first round of *trans* silencing event (Fig. 2-3B, bottom cross). Our data suggests that the *trans* silencing signals were lost after F1 and that *trans* silencing signals are different from the signals that maintain the silencing of *iT* when initiated by mating. We hypothesize that these *trans* silencing signals are regulating gene expression at the RNA level.

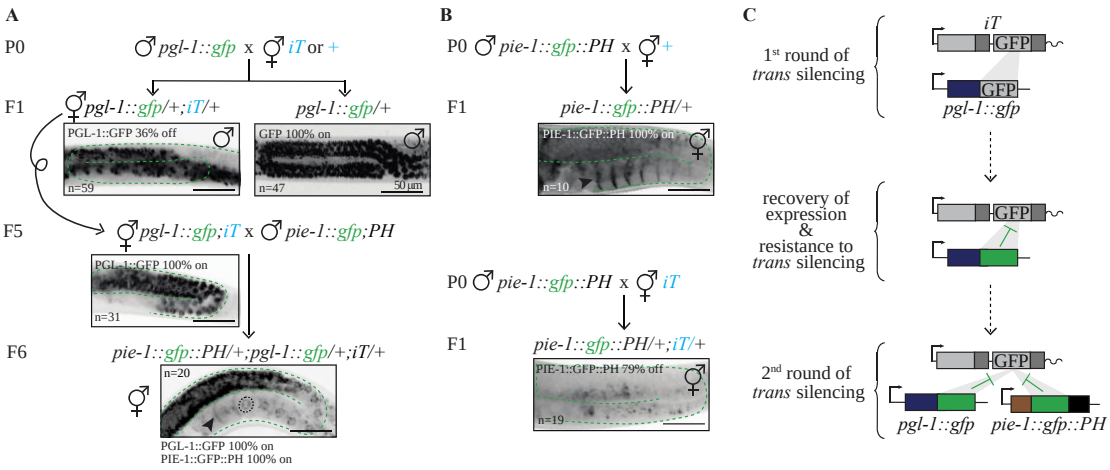

Figure 2-3. Evidence for the reduction in the potency of *trans* silencing by a silenced gene upon exposure to an expressed gene with homologous sequences.

(**A**) Males expressing *pgl-1::gfp* were crossed with hermaphrodites containing *iT* or with wild-type hermaphrodites. 36% of F1s that inherited a copy of *pgl-1::gfp* and *iT* were off for PGL-1::GFP. Whereas 100% of F1s that inherit only a copy of *pgl-1::gfp/+* remained on. nonDpy heterozygous F1 hermaphrodites were allowed to have self-progeny to obtain *pgl-1::gfp; iT* homozygotes. F5 homozygotes showed recovery of PGL-1::GFP expression despite the unexpressed state of *iT* in the background. *pgl-1::gfp* resisted further *trans* silencing by *iT*. F5 homozygotes were crossed with another homologous gene *pie-1::gfp::PH* to observe whether *iT* in a resistant background can initiate the silencing of another gene. F6 heterozygous cross progeny did not show *trans* silencing of PIE-1::GFP::PH despite complete recovery of PGL-1::GFP expression and continued lack of expression from the epiallele *iT*. (**B**) *Top cross*, males expressing PIE-1::GFP::PH were crossed with wild-type hermaphrodites. F1 heterozygous hermaphrodites imaged 24 hrs post-L4 did not show silencing. *Bottom cross*, males expressing PIE-1::GFP::PH crossed with *iT* hermaphrodites, however, did show silencing *in trans* in about 79% of progeny. (C) Schematic depicting *trans* silencing of the homologous GFP sequence of *pgl-1::gfp* in F1, recovery of PGL-1::GFP expression in Fn despite the unexpressed state of *iT*, and the inability for *iT* to re-initiate *trans* silencing of a newly introduced homologous GFP sequence (*pie-1::gfp::PH*). Scale bars, 50 µm.

Although *trans* silencing can be transient, this is not the case all the time. We observed transgenerational *trans* silencing effects when *iT* was exposed to different variants of the *T* sequence. For example, the two variants referred to as *Tcherry* and *TcherryΔpi* contain the same promoter *mex-5*, *mCherry* sequence (except *TcherryΔpi* contains 14 silent mutations), and *cye-1 3'utr*. In these cases, we observed silencing that could persist up to F3 or F4 generation respectively. About 60% of homozygous *Tcherry* and ~100% of *TcherryΔpi* animals recovered expression by F7. The only case in which we observed long-term silencing up to F8 was when progeny was passaged as heterozygotes for *iT* and *Tcherry.*

We continued to investigate the extent of *trans* silencing for two endogenous tags *gfp::mex-5* and *mCherry::mex-5* that were not susceptible to mating-induced silencing (Fig. 2-4A). Both endogenous tags have homologous sequences to *iT* (e.g. promoter and reporter gene). When males expressing *gfp::mex-5* were crossed to *iT* hermaphrodites, we observed that 100% of cross-progeny males and ~92% of hermaphrodites resisted *trans* silencing (Fig. 2-4B, F1s). About 8% of F1 hermaphrodites showed a reduction in GFP expression. These data suggest *trans* silencing is not dependent on sequence homology alone because, despite ~100% *gfp* sequence homology, *gfp::mex-5* showed little to no *trans* silencing. Perhaps such resistance to silencing could be influenced by the presence of the *mex-5* coding sequence. To test this hypothesis, we performed the same cross with animals containing *mCherry::mex-5.*

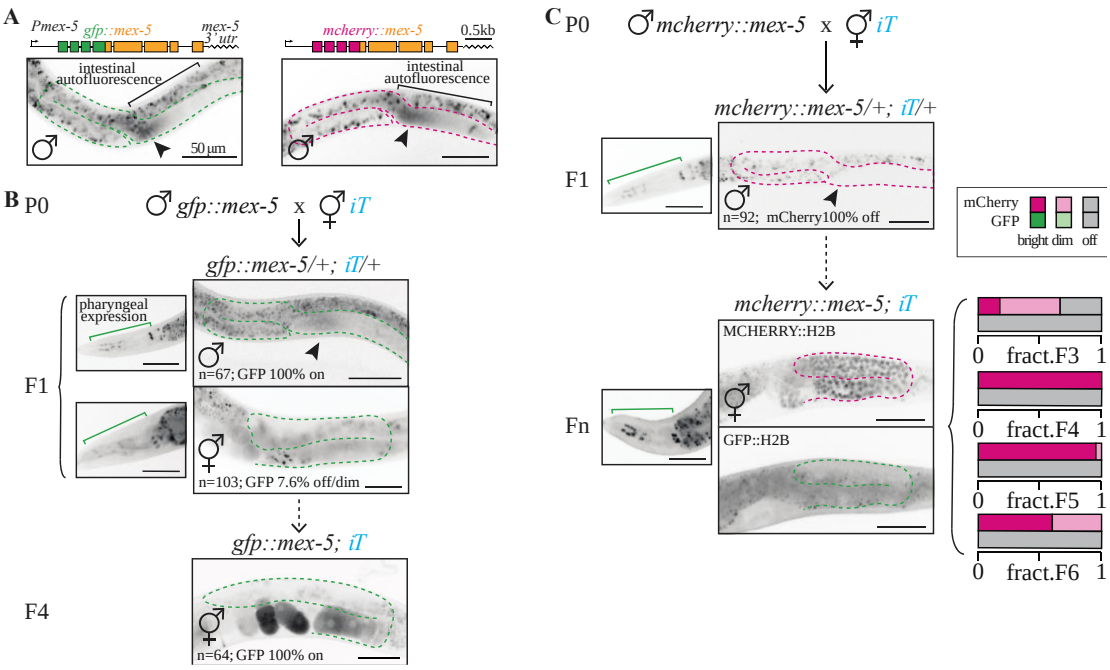

Figure 2-4. A gene can recover expression after *trans* silencing and can re-activate an unexpressed homologous gene.

(**A**) Representative images of L4 males expressing either *gfp::mex-5* (left) or *mCherry::mex-5* (right). (**B**) Males expressing *gfp::mex-5* were crossed with *iT* hermaphrodites which yielded F1 males and hermaphrodites that resist *trans* silencing. GFP::MEX-5 expression remained on in 100% of F1 males. About 8% of hermaphrodites were off for GFP::MEX-5. By F4, homozygous *gfp::mex-5; iT* recovered complete expression. (C) Males expressing *mCherry::mex-5* were crossed with *iT* hermaphrodites. All F1 males were off for mCherry::MEX-5. By F3, homozygous *mCherry::mex-5; iT* showed re-activation of mCherry::H2B (top bars) from *iT*. Expression was stably re-activated in the subsequent generation, despite the lack of re-activation of GFP::H2B (bottom bar).

In contrast to *gfp::mex-5*, when we crossed males expressing *mCherry::mex-5* to *iT* males, we observed that 100% of F1 males were off for mCherry::MEX-5 (Fig. 2-4C, F1s). This suggests that the *mex-5* coding sequence alone cannot explain the resistance to *trans* silencing we see with *gfp::mex-5*. We continued to look across generations of animals containing both *mCherry::mex-5* and *iT* to observe the transgenerational silencing effect seen with *iT/Tcherry* observed in Devanapally et. al. (2021). Curiously, we observed re-activation of the homologous *mCherry::h2b* sequence in *iT* (Fig. 2-4C, Fn). The re-activation remained restricted to only mCherry::H2B when measured from F3 through F6 and did not occur with GFP::H2B within *iT* (Fig. 2-4C, F4). This re-activation of mCherry::H2B expression in *iT* made it difficult to observe the recovery of mCherry::MEX-5 expression at L4- or adult-staged hermaphrodites (Fig. 2-4C, Fn). However, we did observe recovery of expression from F3-F6 when looking at embryos (data not shown), suggesting that the endogenous *mCherry::mex-5* tag can recover expression by F3 and such recovery is stable.

Our results suggest after *trans* silencing, the unexpressed homologous sequence can recover expression and rescue the expression of the stably silenced epiallele *iT*. It is possible that the *trans* silencing signals that act on *mCherry:mex-5* were lost either through competition between positive regulators that protect the gene from silencing and negative regulators that silence genes, or simply via dilution of such silencing signals. Both protective and silencing signals must be actively maintained at the RNA level and not chromatin-dependent, because GFP actively remains silenced. These *trans* silencing signals are also different from the silencing signals generated upon mating and are sequence-specific. These signals can only be transmitted to other

homologous loci. Unlike the recovery of both genes after feeding RNAi, only one cistron can recover expression upon interaction with *mCherry::mex-5*, and the other cistron can remain actively silenced. This independent regulation of the two cistrons within the two-gene operon reveals the second case of a different form of heritable RNA silencing.

Prior work in the Laboratory of Craig Mello showed a similar re-activation phenomenon seen in *trans*, but this type of regulation is different from the chromatin-independent re-activation we see with the two-gene operon *T* (Shirayama et. al., 2012). They tested two different transgenes expressing GFP in the germline, *gfp::wrm-1*, and *oma-1::gfp*. Both transgenes could recover expression by F3 upon *gfp* RNAi. They tested if these transgenes were also susceptible to *trans* silencing and if the silencing could persist across generations. Not only did they find that both genes could resist *trans* silencing in F1 when exposed to a homologous transgene that is not expressed, but these genes could also dominantly activate the expression of the homologous transgene that was silenced (Fig. 2-4B, F4). However, the number of F1s that did show re-activation was relatively low compared to the number of re-activated F1s (100% F1s) we saw in *iT* upon exposure to *mCherry::mex-5*. Although *gfp::mex-5* could resist *trans* silencing, it cannot re-activate the expression of the homologous *gfp* gene in *iT*, different from what was observed in the Mello where the resistant *gfp*-containing genes could re-activate the silenced homologous gene.

They also saw complete recovery by F3 when propagating animals containing the activated transgene in the presence of the resistant transgenes *gfp::wrm-1* or *oma-1::gfp* similar to what we observed with *mCherry::mex-5* in the presence of *iT*.

However, from the Mello lab, re-isolating the transgene that was initially off resulted in 100% re-silencing or stable expression depending on the gene context of the inactive gene. Together, our results suggest the presence of protective and silencing mechanisms that compete with one another. Although the types of regulation reported by the Mello Lab and our lab are very different from one another, it is unclear what factors can determine the outcome of a gene – expressed by *trans* activation or unexpressed by *trans* silencing.

2.4.3. Loss of the Argonaute HRDE-1 can selectively re-activate the expression of one cistron in a two-gene operon

In *C. elegans* a common gene-editing method for integrating a single-copy transgene is called MosSCI developed by the Jorgensen Lab (Frokjaer-Jensen et. al., 2008). The principles behind this method were first demonstrated by the Bessereau Lab for a method they later called MosTIC (**Mos**1 excision-induced **t**ransgene-**i**nstructed gene **c**onversion) which showed successful tag inserts and deletions for specific genes that were dependent on the presence of a Mos1 insert at a gene locus (Bessereau et. al., 2001). A high-throughput method was used to create a library of MosI inserts in the *C. elegans* genome at defined sites so that it could be used by all labs to easily integrate their gene of choice (NemaGENETAG consortium, Duverger et. al., 2007; Bazopoulou et. al. 2009). The Jorgensen Lab adapted this method and chose a site that would not disrupt the function of neighboring genes, nor nearby promoters and enhancers. They settled on the locus called *ttTi5605* which is present on chromosome II. To identify insertions, they used a positive selection marker by crossing *ttTi5605* Mos1 to animals lacking the *unc-119* gene (*unc-119(ed3)).* These mutants are typically small, almost

paralyzed, incapable of forming Dauer upon starvation, and have a relatively lower brood size than wild type. To rescue this phenotype upon integration, they incorporated the wild-type, *C. briggsae, unc-119* sequence into a plasmid that contains the gene of insert flanked by sequences homologous to *ttTi5605*. This plasmid acts as both the homology-repair template and positive selection marker. To excise the Mos1 element from *ttTi5605*, the Mos1 transposase is expressed as a helper plasmid under the promoter *eft-3* that is injected together with the homology repair plasmid. Upon injection into the germline, the transposase is activated and cuts the Mos1 element out of the genome leaving a double-stranded DNA break. Additionally, they developed a negative selection marker to facilitate better identification of the insert after injection. This plasmid contains the toxin PEEL-1 driven by the heat-shock promoter *hsp-1*. Upon heat shock, any false-positive transgenic animals would be killed by the toxin. They demonstrated successful integration of different reporter genes, one of which is called *oxSi487* which is the two-gene operon that expresses mCherry and GFP in the germline that we referred to as *T.*

Previous studies have shown that upon integration of a gene of interest using this method, the gene could turn off spontaneously (Shirayama et. al., 2012; Ashe et. al., 2012; Devanapally et. al., 2021). Some labs have seen that the gene can recover expression after passaging for a few generations and other labs have seen more permanent effects. The Mello lab showed that spontaneous silencing by MosSCI injection is a result of piRNA-induced silencing which depends on the primary Argonaute called PRG-1 (Shirayama et. al., 2012). piRNAs, which are endogenously transcribed in the *C. elegans* genome can bind to the primary Argonaute, PRG-1, to

target a gene of complementary sequence and result in a cascade of events that requires the secondary Argonaute HRDE-1 for silencing. They observed that injecting a transgene containing *gfp* into the Mos1 background results in spontaneous long-term silencing. However, injection into a *prg-1(-)* background can prevent spontaneous silencing. Furthermore, they showed that loss of HRDE-1 can rescue the expression of the gene susceptible to such silencing.

We wanted to determine if the transgene *T* was susceptible to spontaneous pi-RNA silencing upon MosSCI integration. We re-cloned the sequences of *T* into a plasmid containing the *cbr unc-119* rescue sequence flanked by homologous sequences to the *ttTi5605* locus. After injection, we screened for transgenic animals. Successfully integrated animals were referred to as *T_re-inserted*. There was spontaneous silencing of both cistrons and upon removal of HRDE-1, there was re-expression of both mRNAs (Fig.2-5A and B). This suggests that *T_re-inserted* is susceptible to spontaneous silencing regulated by HRDE-1.

We next considered how would re-constructing the sequences of *T* affect susceptibility to mating-induced silencing, *trans* silencing, and dsRNA-induced silencing. Therefore, we created various two-gene operon constructs of the gene sequences present in *T*. The first two-gene variant was designed to make it easier to see mCherry and GFP expression in the germline. Because both mCherry and GFP are fused to H2B, there could be homologous sequences from *h2b* of one cistron influencing expression or silencing of *h2b* in the other cistron. To create a different pattern, we deleted the *h2b* sequence from *mcherry* and kept *gfp* fused to *h2b*. We refer to this variant as *Tcherry∆h2b* (Fig. 2-5A second gene). We expect to see mCherry

expression throughout the entire germline and GFP to remain expressed in the nucleus. Indeed, that was what we observed when we successfully integrated the gene into the genome (Fig. 2-5B, second row). However, like *T_re-inserted*, it was off upon integration and HRDE-1 loss rescued the expression of both cistrons (Fig. 2-5B, second row).

The second two-gene operon we created is referred to as *Tswap* (Fig. 2-5A, third gene). This new construct instead contains *gfp* as the first cistron and *mcherry* as the second. Upon injection, it too was silenced spontaneously, suggesting that this gene variant is susceptible to piRNA-induced silencing and therefore loss of HRDE-1 would re-activate the expression we saw with *T_re-inserted* and *Tcherry∆h2b*. However, that was not what we observed. The expression of the second cistron, mCherry::H2B recovered, but GFP::H2B remained stably off (Fig. 2-5B, third row), even after one year of passaging. Our data reveals a chromatin-independent epigenetic mechanism that acts independently of HRDE-1 or in addition to HRDE-1.

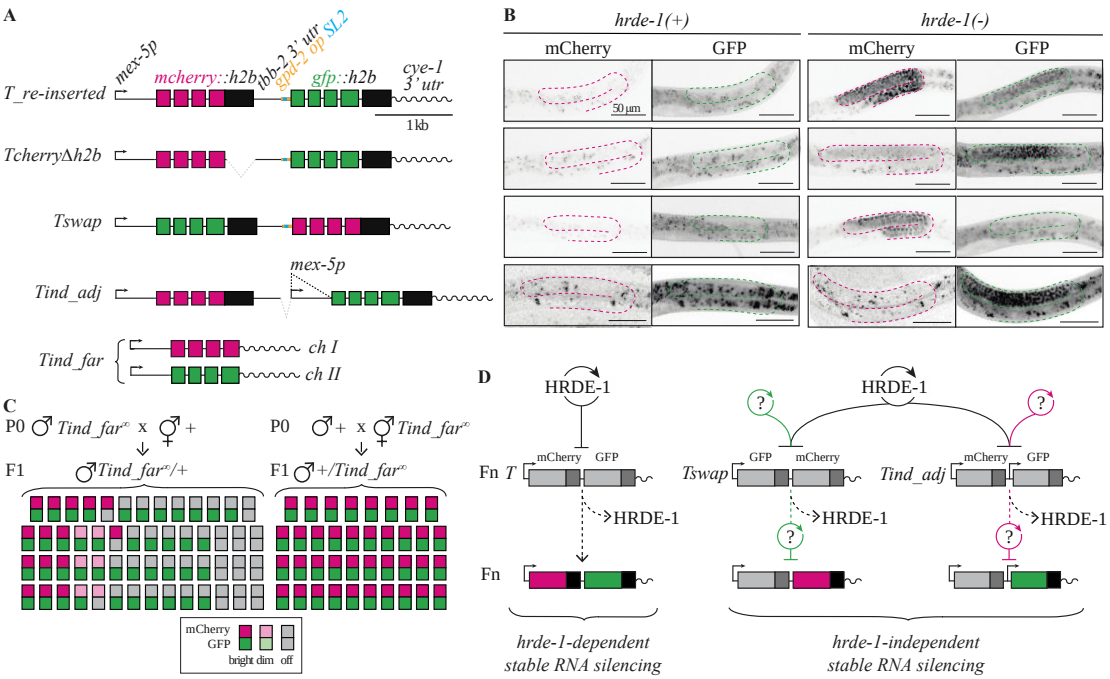

Figure 2-5. A two-gene operon is regulated by HRDE-1-dependent and HRDE-1-independent mechanisms.

(**A**) Schematic of different genes that contain both *mcherry* and *gfp* expressed under the same promoter *mex-5* integrated into the genome using MosSCI. (**B**) Representative images of gene constructs in (A). *Left*, upon MosSCI integration, the genes were not expressed. *Right*, loss of *hrde-1(-)* rescued expression of one or both unexpressed genes in different cases. All scale bars = 50µm. (**C**) Mating-induced silencing test of *Tind_far∞*. Males expressing *Tind_far∞* crossed with hermaphrodites that lack the gene can show silencing of mCherry and GFP in F1. mCherry and GFP expression was scored as either bright, dim or off (box). Each rectangle indicates one worm. ∞ indicates the presence of *Tind_far* in a *hrde-1(-)* background. (**D**) model depicting two types of stable RNA silencing. Far left gene is susceptible to heritable RNA silencing regulated by HRDE-1-dependent positive feedback loop. Middle and right depicts both HRDE-1-dependent feedback loop silencing of one gene whereas silencing of the other gene is dependent on an unknown factor.

Lastly, we expressed *mCherry* and *gfp* as two different genes either on the same or different chromosomes and denoted these constructs with "ind" (short for independent). One of the two independent genes we designed express *mcherry* on chromosome I and *gfp* on chromosome II and therefore we refer to this gene variant as *Tind_far* (where "*T"* references the sequences of the two-gene operon and "*far"* refers to the two reporter genes being on different chromosomes) (Fig. 2-5A, fifth gene). Both *mcherry* and *gfp* contain the same *mex-5* promoter and *cye-1 3'utr* sequence (*mex-5p::mcherry::cye-1 3'utr; mex-5p::gfp::cye-1 3'utr*). The second independent design contains *mcherry* and *gfp* next to one another on the *ttTi5605* locus of chromosome II and we refer to this gene variant as *Tind_adj* (where *"adj"* stands for adjacent) (Fig. 2-5A, fourth gene). Both reporter genes contain their own respective *mex-5* promoter sequence but different 3'utr sequences (*mex-5p::mcherry::cye-1 3'utr mex-5p::gfp::tbb-2 3'utr*). Presumably, these two independent gene variants should be transcribed at the same rate, but potentially translated at different rates.

We created *Tind_far* by careful genetic crossing to prevent mating-induced silencing and noticed that after a few generations of propagating animals expressing both transgenes, mCherry spontaneously turned off and GFP remained stably on. This suggests that *mcherry* is more susceptible to perturbations than *gfp*. Loss of HRDE-1 rescued the expression of mCherry. We tested if *Tind_far; hrde-1(-)* is susceptible to mating-induced silencing by crossing males expressing both transgenes to hermaphrodites lacking the gene. Indeed, both reporters were susceptible to mating-induced silencing (Fig. 2-5C). However, as seen with spontaneous silencing over time, mCherry was more susceptible than GFP.

Tind_adj, however, was created by MosSCI. Both genes were off after successful injection and integration into the genome (Fig. 2-5B, fourth row). Unlike *T_re-inserted*, *TcherryΔh2b,* and *Tind_far*, the loss of HRDE-1 did not rescue the expression of both genes despite having independent promoter sequences. Instead, similar to what we observed with *Tswap*, one of the genes came back on (GFP::H2B recovered expression and mCherry::H2B did not) (Fig. 2-5B, fourth gene). Our data reinforces the idea that the two genes within an operon or not are regulated independently of one another and require both HRDE-1-independent and HRDE-1-dependent mechanisms.

2.4.4. Molecular and genetic requirements for different epigenetic states

Integrating *Tswap* and *Tind_far* by MossCI reveals different HRDE-1 requirements for piRNA-induced spontaneous silencing. To begin understanding the molecular differences between each two-gene variant, we measured the presence of pUGylated *mcherry* and *gfp* RNAs. The Kennedy Lab showed that upon dsRNA-and piRNA-induced silencing, the cleaved target transcripts are stabilized by the addition of UG-dinucleotides through a nucleotidyltransferase called RDE-3/MUT-2 (Shukla et. al., 2020; Shukla et. al., 2021). Through RT-PCR they showed that pUGylated RNAs can only be detectable upon silencing and that such intermediates can be detectable across generations that remain silenced. Using this method, we wanted to determine the silencing signature of pUGylated *mCherry* and *gfp* RNAs in *iT*, the epiallele of *T* that has been off for hundreds of generations.

We used a universal reverse transcription primer with AC repeats (AC_9) for cDNA synthesis. Presumably, transcripts with UG-dinucleotides will only be selected.

As a positive control, we used the gene, *gsa-1*, which has an 18nt long gnomically encoded pUG repeat in its 3'utr. As a negative control, we used both wild type and *T*. As expected, we detected pUGylated *mcherry* and *gfp* RNAs in *iT* and no pUGylated transcripts in wild type (Fig. 2-6, compare *iT*, *T* and wild type). To our surprise, we also detected pUGylated transcripts of different patterns in *T*, the gene that is expressed. Our data suggests that pUGylated transcripts can be detectable in both the on and off states of a gene and that perhaps the presence of such RNAs in the expressed state could affect its susceptibility to heritable RNA silencing through mechanisms unknown.

We next measured *mcherry* and *gfp* pUG-RNAs from the two-gene transgene with different epigenetic states (*Tswap; hrde-1(+)* and *Tswap; hrde-1(-)*, Fig. 2-6). The epigenetic state refers to the unexpressed or expressed state of a gene. For example, *T*, *T_re-inserted*, *Tcherry∆h2b*, and *Tswap* were spontaneously silenced by piRNAs upon genome-editing. Therefore, the unexpressed state of these genes counts as one type of epigenetic state. Impact of HRDE-1 loss on gene expression counts as a different epigenetic state. In *T* and *T_re-inserted*, both cistrons were re-expressed, but in the case of *Tswap*, one of the cistron was re-expressed and the other remained unexpressed. The unexpressed state of *T_re-inserted* will be referred to as *iT** because it has the exact same sequence as *iT*. However, silencing of *iT** occurred upon genome-editing via MosSCI and silencing of *iT* occurred upon mating. Therefore, the asterisk refers to the differences in methods that led to the stable unexpressed state of *T*. Slight differences in pUGylated patterns for *mcherry* and *gfp* was observed between both off states of *T* (Fig. 2-6B. *iT* and *iT**). These differences could be a result of how silencing was initiated and how it is maintained. Although both types of silencing were initiated by

piRNAs, the two silencing states are different and have different silencing regulators that distinguish the two. Likewise, the two phenomena require HRDE-1 to maintain silencing across generations, but HRDE-1 requirement alone cannot explain for these differences in heritable silencing states. Overall, the reproducibility of these pUGylated patterns between samples of biological replicates could suggest that the same population of pUG-RNAs exist for the same epigenetic state.

Furthermore, we looked at the presence of *mcherry* and *gfp* pUG-RNAs in animals with *Tswap* subject to spontaneous silencing and HRDE-1 loss. As expected, we detected pUGylated RNAs for *Tswap* in which the two cistrons are off. However, the pattern of pUGylated transcripts was different from the patterns we saw with *T_re-inserted* where both cistrons were off upon injection, suggesting that despite similar triggers, the order in which the sequences are arranged within a gene can dictate different silencing effects. Upon loss of HRDE-1 where one of the cistron mCherry::H2B recovered expression, and GFP::H2B remained off, we noticed that some of the larger *mcherry* pUGylated RNAs were similar to the pUGylated *mcherry* RNAs in the expressed state of *T.* However, the off state of GFP::H2B showed a different *gfp* pUG RNA pattern from the silenced state of GFP::H2B in both *iT* and *iT** suggesting even further that the epigenetic silencing mechanism for *gfp* in *Tswap* requires HRDE-1 independent regulators that still need to be identified.

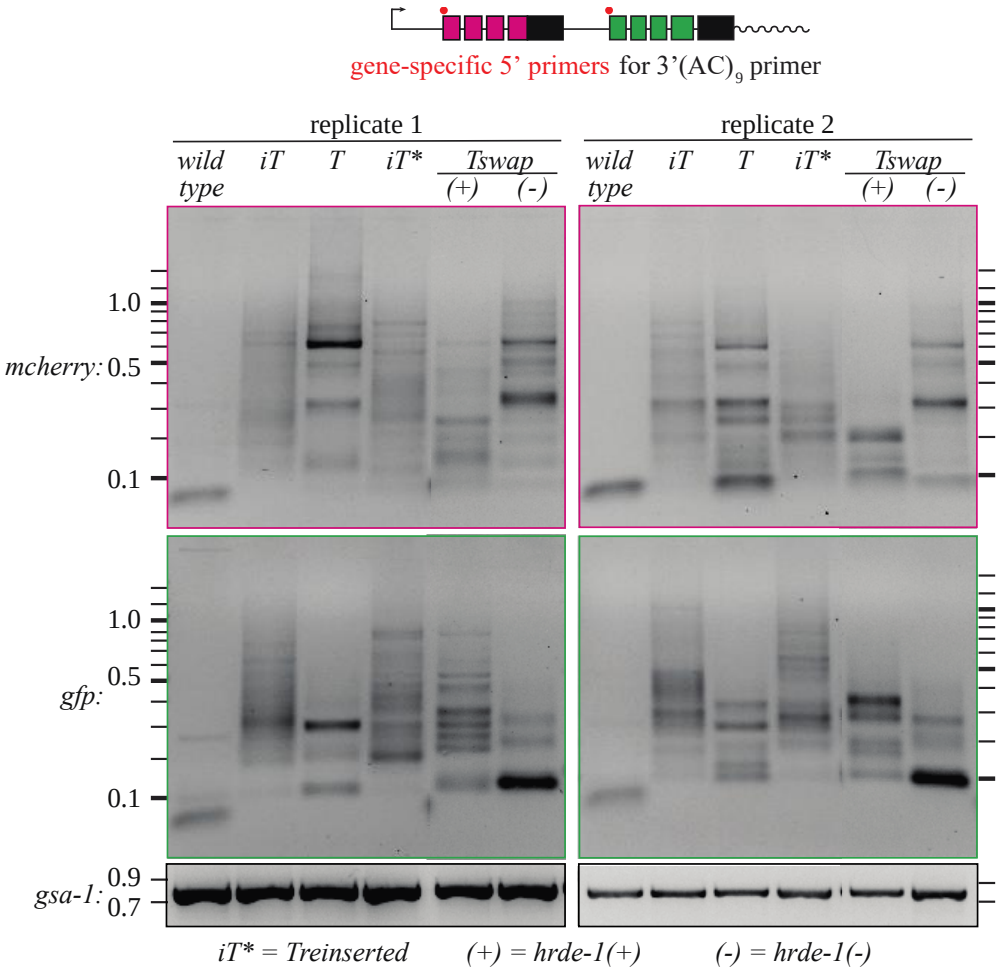

Figure 2-6. Apparently similar heritable epigenetic states can be associated with different molecular intermediates.

Top, schematic depicting location of gene-specific forward primers at the start of *mcherry* and *gfp*. *Bottom*, *mcherry* (maroon box), *gfp* (green box) pUG-RNAs measured in wild type and in animals containing *T* or *iT*. Samples were prepared from two biological replicates (1 & 2). *gsa-1* (black box) with a genomically encoded stretch of AC dinucleotides was used as a control (black box). *iT** represents *T_re-inserted*, which remained unexpressed upon MosSCI integration. pUG RNAs were measured in *Tswap;hrde-1(+)* (indicated as +), where both mCherry and GFP expression are off and in *Tswap;hrde-1(-)* (indicated as -), where mCherry is re-expressed and GFP remains off.

To determine which regulators could be acting independently of HRDE-1 or in combination with HRDE-1 to maintain the silencing of one cistron despite the presence of the expressed cistron in *Tswap*, we created nonsense mutations of three regulators known to function at different steps in the RNAi pathway. These regulators are RDE-3/MUT-2, ZNFX-1, or DEPS-1. RDE-3/MUT-2 is required for the formation of the mutator foci. This is a site of pUGylation activity (where mRNA transcripts are targeted and cleaved) and the production of secondary small RNAs (22G-RNAs) by an RNA-dependent RNA polymerase (RdRp) (Sijen et. al., 2001). ZNFX-1 is required for the formation of Z granules. These are membraneless germ granules that act as a hub of 22G-RNA interaction with the secondary Argonaute called WAGO-4 (Xu et. al., 2018). This pathway is thought to act in parallel to HRDE-1-dependent silencing where HRDE-1 loading of 22G-RNA presumably occurs in a different germ granule called P granules (Reviewed Marnik & Updike, 2019). Lastly, DEPS-1 has been reported to function in both P and Z granules and is required for piRNA-dependent silencing (Suen et. al., 2020). Loss of DEPS-1 results in failure to produce secondary siRNAs for piRNA targets (Suen et. al., 2020).

We found that the loss of either MUT-2, ZNFX-1, or DEPS-1 could not recover the expression of both cistrons in *Tswap*. Instead, only mCherry::H2B recovered and GFP::H2B remained off, similar to what we observed with HRDE-1 loss (Fig. 2-7A). We next created double mutants. Upon loss of MUT-2, ZNFX-1, and DEPS-1 in the absence of HRDE-1, GFP::H2B remained off and mCherry::H2B remained on (Fig. 2-7B). These results suggest a few things: one, cleavage of *gfp* transcripts and production of secondary short-interfering (siRNAs) could be occurring somewhere else in the

cytoplasm; two, secondary siRNAs which could be of different species are loaded to an unknown regulator; or three, the hub of interaction between secondary siRNAs and Argonautes is different from P and Z granules; or that yet unknown mechanisms are required for selective silencing. Further investigation into these properties is necessary to determine HRDE-1-independent regulators of heritable RNA silencing. Thus, the third case of a different heritable RNA silencing form is revealed. Two genes in an operon are regulated independent of one another. One gene is regulated by HRDE-1 and the other gene is regulated independent of HRDE-1. This type of regulation occurs at the RNA level and is independent of chromatin modifications.

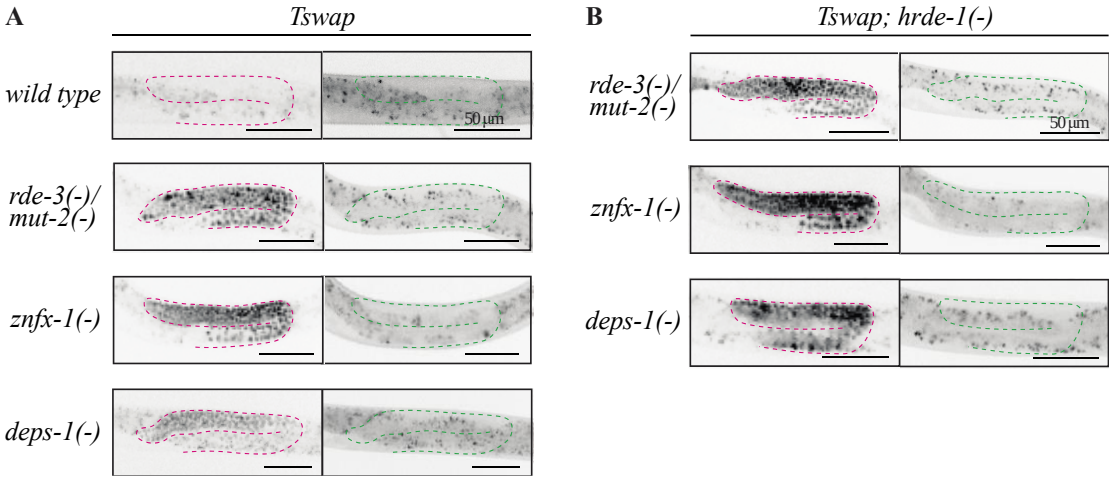

Figure 2-7. Stable RNA silencing of one cistron in a two-gene operon is independent of the regulators required for the silencing of the other cistron.

(**A**) *Tswap* in either *rde-3(-)/mut-2(-), znfx-1(-), or deps-1(-)* single mutants showing expression of mCherry (left) but not GFP (right). (**B**) *rde-3(-)/mut-2(-), znfx-1(-),* and *deps-1(-)* mutants generated in *Tswap;hrde-1(-),* showing the same pattern as in (A). All scale bars 50 um.

2.5. Discussion

Perturbing the environment of a two-gene operon expressed in the germline reveals different modes of heritable RNA silencing. The first mode of silencing was clearly observed upon exposure to *mCherry*- and *gfp*-dsRNA. Although *mCherry* RNAi suggests pre-mRNA level silencing of both cistrons, it was not clear how two cistrons were regulated differently across generations. *gfp* RNAi reveals tight regulation between co-transcriptional and post-transcriptional mechanisms within generations.

The second mode of silencing was seen *in trans* between two genes containing a homologous *mCherry* sequence. The two-gene operon which contains both *mCherry* and *gfp* can maintain an unexpressed state across many generations and turn off a gene with a homologous *mCherry* sequence. This phenomenon has been reported in previous studies. However, such *trans* silencing is temporary. In turn, the homologous *mCherry* gene can recover expression and re-activate the expression of the homologous cistron in the operon. Thus, revealing that two cistrons in an operon are not susceptible to the same type of regulation but instead contain separate *trans* activating and *trans* silencing mechanisms.

The third mode of silencing revealed HRDE-1- dependent and -independent silencing feedback loops. Genome editing by the excision of a transposable element can result in stochastic silencing of the integrated gene through the piRNA/PRG-1 pathway. This form of silencing relies on the Argonaute HRDE-1. Without HRDE-1 function, the integrated gene can recover stable expression. Using a two-gene operon that contains two reporter genes, we reveal that both cistrons are susceptible to

stochastic silencing upon injection. However, stable RNA silencing of one cistron is dependent on HRDE-1 and the other cistron is not. Loss of other known regulators of silencing such as *rde-3, znfx-1, deps-1* reveals a different RNA silencing mechanism that could share similarities with the dsRNA- and piRNA- pathways in *C. elegans*. Furthermore, different patterns of molecular intermediates such as pUGylated RNAs reveal that the unexpressed state of a gene perturbed by mating or stochastic silencing is not the same.

Chapter 3: Stable RNA silencing relies on synergy between sensing and processing of double-stranded (ds)RNA

3.1. Preface

Some strains were obtained from the *Caenorhabditis elegans* Genetic Stock Center (CGC), the Cohen-Fix Laboratory (National Institute of Health), and the Seydoux Laboratory (Johns Hopkins University). Some of the gene sequences were obtained from Wormbase.

Victoria Murphy, an undergraduate student I mentored, performed experiments to generate Fig. 3-1. The *mCherry*-containing gene she used was generated by Sindhuja Devanapally and published in Devanapally S.*, Raman* P., Chey M., Allgood S., Ettefa F., Diop M., Lin Y., Cho YE., Jose AM. Mating can initiate stable RNA silencing that overcomes epigenetic recovery. Nature Communications 12:4239.

Daphne Knudsen performed poly-A+ RNA-sequencing on animals expressing *Tcherry::cye-1 3'utr* and I plotted these data as a volcano plot in Fig. 3-2A.

Dr. Antony Jose performed bioinformatic analysis to curate the total-RNA sequencing results for the *mcherry* strains tested in this chapter and these data are plotted as a heat-map for Fig.3-3-A. Dr. Jose also compared our dataset with other RNA-seq datasets from other labs to make Fig. 3-4. He also mapped the sense and anti-sense reads of a few genes tested to generate Fig. 3-5B.

The remaining work in this chapter was generated by me.

3.2. Introduction

In all systems, there are thousands of diverse RNA species that play an important role in development, behavior, metabolism, and response to physiological and environmental changes (Jones & Baylin, 2002; Boyko et. al., 2011.) However, recognition of RNA as a regulatory molecule did not occur until the first decade of the 21st century when transcriptomic analyses used to define the proteome discovered that the number of protein-coding genes is similar across animals. *C. elegans*, a simple worm that grows to an average size of 1 mm, contains ~1000 somatic cells and humans, being far more complex have ~$30x10^{12}$ somatic cells. Amazingly, only ~20,000 protein-coding genes are present in both systems (Shaye et. al., 2011). Therefore, protein alone could not explain all biological phenomena or disorders.

One of the earliest examples that hinted at the diverse roles of non-coding RNAs in gene regulation was the observation made with *"meiRNAs"* (meiotic RNAs) for its regulation of meiosis I in the yeast *Schizosaccharomyces pombe*; loss of *meiRNAs* results in meiotic cell cycle arrest *(*Watanabe et. al., 1994; Shichino et. al., 2014; Yamashita et. al., 2019). However, the role of RNAs as essential regulators of gene expression did not come to fruition until the three most abundant classes of non-coding RNAs were discovered: microRNAs (miRNAs), PIWI-interacting RNAs (piRNAs), and small-interfering RNAs (siRNAs). Only then did the field of epigenetics begin to expand and RNA became appreciated for its role in underlying diseases or biological phenomena that cannot simply be explained by DNA alone.

The first endogenous small RNAs of 22 nts *lin-4* and *let-7* were discovered in eukaryotes by the groups of Victor Ambros and Gary Ruvkun respectively (Lim et. al.,

2003). The Ambros group showed that *lin-4* could temporally control post-embryonic development (normal transition between larval stage L1-to-L2) in *C. elegans*. *Lin-4* maps to the intron of a long spliced non-coding RNA and its transcript is processed into two overlapping RNA fragments of 61 nts (rare) and 22 nts (most abundant) in length (Lee et. al., 1993). The 22 nts RNAs can target the 3' untranslated region (utr) of *lin-14* transcripts, a heterochronic gene that encodes a nuclear protein. Loss of *lin-4* results in *lin-14* gain-of-function post-L2 development and leads to retarded development. Ruvkun group showed that *let-7* 22 nt RNAs expressed in the hypodermal seam cells of L3 and L4 staged in *C. elegans* could temporally control the later stages of development (L4-to-adult) by targeting the 3'utr of multiple target genes, one of which is a heterochronic gene *lin-41* (Slack et. al., 2000). Loss of *let-7* results in continuous seam cell division, failure to transition to the adult stage, and thus death Reinhart et. al., 2000). Both *lin-4* and *let-7* small RNAs could target the 3'utr of their target mRNAs through partial complementary binding, preventing translation, thereby reducing the level of protein product without affecting transcript abundance (Ambros, 2001). It was unclear what functions non-coding RNAs had in gene regulation at the time. Little did they know, they discovered a new class of antisense RNAs called miRNAs conserved in both plants and animals. This post-transcriptional mechanism can regulate diverse processes such as cell fate specification, apoptosis, and metabolism (Bartel & Chen, 2004). To date, there are ~112 confirmed miRNA genes present in *C. elegans*, but gene estimates predict it could be as high as 300 (Ruby et. al., 2006; Grad et. al., 2003; Lim et al., 2003).

One of the largest and most recent classes of small non-coding RNAs to be discovered in animals is piRNAs. piRNAs are notably 20-35 nts in length, contain a 5' monophosphate end with a strong bias for Uracil, and a 2*'-O*-methyl modified the 3'end, which alters periodate degradation (Bagijn et. al., 2008). They help to promote genome stability and maintain fertility by repressing transposable elements in the germline. piRNAs can bind to PIWI-clade proteins that belong to an animal-specific subgroup of the Argonaute family and are guided to a target transcript of complementary sequence which leads to a cascade of silencing events (Izumi & Tomari, 2014; Perez-Borrajero et. al. 2021). Piwi-like proteins are also present in bacteria and archaea, suggesting that this small RNA-based immune response is of ancient origin.

The PIWI/piRNA pathway was first discovered in the P-element induced wimpy testis of *D. melanogaster* by the Gvozdev group (Aravin et. al., 2001; Aravin et. al., 2004). They observed homology-dependent RNAi-related silencing mechanism of the X-linked tandem testis-specific gene *Stellate (Ste)*. This silencing was regulated by the Y-linked gene *Suppresor of Stellate (Su(Ste))*, present as abundant homologous tandem repeats that form sense and anti-sense short RNAs (Adashev et. al., 2021). Mutations in *Su(Ste)* lead to the de-repression of *Ste*, long terminal repeats (LTRs) and non-LTR retrotransposons, and an accumulation of crystalline aggregates in spermatocytes (Adashev et. al., 2021). Numerous meiotic defects were observed including male sterility.

In *C. elegans*, piRNAs are 21 nts in length (21U-RNA) and are endogenously transcribed between protein-coding genes and within introns from two clusters on

chromosome IV of two million and four million base pairs respectively (Ruby 2006). These autonomously expressed small RNAs share a large upstream motif referred to as the Ruby motif, which is conserved among other nematodes (Ruby 2006). Although these worm-specific piRNAs resemble mammalian pachytene piRNAs expressed from large genomic clusters (Aravin et al., 2006; Batista et al., 2008;Lee et. al., 2012) these 21U-RNA loci are uniquely interspersed on both strands with an average density of 21U-RNA locus every 200-300 bp. There is little to no overlap with each other, repetitive elements, or protein-coding regions (Batista et. al. 2008). More than 15,000 piRNAs can be transcribed and could potentially target all germline genes because of their diverse sequence composition and small 21nt length (Batista et. al. 2008). Yet, there are no obvious strong piRNA targets except for the Tc3 family of Class II DNA transposons (Batista et. al. 2008). In total, 12% of the *C. elegans* genome is comprised of six active DNA transposon families (Tc1-5 and Tc7). The Tc3 family which has about 22 copies, are the only transposons targeted by piRNAs (Fischer et. al., 2003). Thus, the nature, function, and target of piRNAs remain a mystery in the worm.

Foreign elements introduced into the genome such as reporter genes are susceptible to piRNA targeting. piRNAs can tolerate a few mismatches within the first 8-nt (position 2-9) at the 5' end referred to as the seed region, but prefer perfect pairing (Zhang et. al., 2018). There are two PIWI proteins present in *C. elegans*, PRG-1 and PRG-2. Although PRG-2 shares sequence homology with PRG-1, loss of PRG-2 does not affect piRNAs abundance and the loss of PRG-1 does (Bagjin et al 2012, Lee et al 2012; Perez-Borrajero et. al. 2021). Loss of PRG-1 also results in fertility defects. Therefore, piRNAs require the primary Argonaute PRG-1 for recognition of a target

transcript. Once PRG-1/piRNAs are bound to the target, an RNA-dependent RNA-polymerase (RdRp) is recruited and results in an unprimed amplification of 5'-triphosphorylated Guanosine of 22 nts in length. These secondary small RNAs referred to as 22G-RNAs can bind to the worm-specific Argonaute (WAGO) HRDE-1 (Bagjin et al 2012; Perez-Borrajero et. al., 2021). HRDE-1 contains a nuclear localization signal that guides the 22G-RNAs into the nucleus. This can trigger chromatin modifications by the histone methyltransferases SET-25 (H3K9) and SET-32 (K3K23me3), and heterochromatin protein 1 ortholog, HPL-2 (Ashe et. al., 2012). piRNAs can initiate silencing of foreign genes but are not required in subsequent generations to maintain silencing. HRDE-1 is required for the maintenance of silencing across generations. Such epigenetic regulation is necessary to prevent instability and maintain germline integrity across generations.

siRNAs are another class of non-coding RNAs identified for their role in co-transcriptional and post-transcriptional gene silencing in plants and fungi. In 1998, Mello and Fire demonstrated that exogenous dsRNA in *C. elegans* can act as a trigger for siRNA-mediated silencing now referred to as RNA interference (RNAi), and this silencing could be inherited (Ghildiyal & Zamore, 2009). However, the mechanism was unclear until it was demonstrated in *D. melanogaster* that siRNAs can guide endonucleolytic cleavage of target sequences (Schwarz et. al., 2002) and DICER, a dsRNA-specific RNase-III ribonuclease, can cut dsRNA into a ~21 nt siRNA duplex with 2 nt overhangs at the 3' ends (Berstein et al 2001). Each strand contains a 5'phosphate and 3' hydroxyl group. One strand of the duplex, the "passenger" is degraded while the other, the "guide", can target a transcript of complementary

sequence and lead to a cascade of Argonaute-small RNA interactions that result in gene silencing. This process is generically referred to as RNA interference and the Argonaute-siRNA complex is referred as **R**NAi-**i**nduced **s**ilencing **c**omplex (RISC).

In *C. elegans* dsRNA can enter the cytosol through the conserved transmembrane protein SID-1. The dsRNA-binding protein RDE-4 can bind to the exogenous dsRNA and recruit DICER. dsRNA is cleaved into shorter duplexes, the primary Argonaute RDE-1 binds to the guide strand and targets an mRNA of complementary sequence. At the site of interaction between the guide and target, the endoribonuclease RDE-8 cleaves the mRNA transcript into two fragments, one end containing the 5'cap and the other with the 3' poly-A tail. The ribonucleotidyltransferase RDE-3/MUT-2 can stabilize the 5' fragment by adding UG-dinucleotides to the open 3' end of the transcript while the 3' poly-A fragment is degraded. An RdRp is recruited to these stabilized poly-UG (pUG) RNAs and can catalyze the production of complementary 22 nt RNAs from the 3' to 5' direction. This hub of interaction takes place in the mutator foci regulated by both RDE-3 and MUT-16. These secondary 22 nt long RNAs which have a 5' guanosine bias (22G-RNA), can bind to worm-specific Argonaute proteins (WAGOs). As mentioned earlier, HRDE-1 is a secondary WAGO protein that functions in the germline to maintain silencing. HRDE-1-bound 22G-RNAs can enter the nucleus, bind to nascent transcripts, and recruit chromatin modifiers such as SET-25 (e.g., deposition of repressive H3K9me3 marks) and other regulators. The discovery of the RNAi process identifies regulators of gene-specific co-transcriptional and post-transcriptional silencing. These regulators shared by both the dsRNA- and piRNA-silencing pathways can be used to investigate

the regulatory architectures that promote or inhibit change in gene expression patterns of a single gene across generations.

Here we investigate the properties of a single reporter gene *mcherry* that is susceptible to long-term silencing by both piRNA and dsRNA triggers. Previously we observed that MosSCI injections could result in spontaneous silencing of mCherry but loss of HRDE-1 can rescue expression. Restoring HRDE-1 function (*hrde-1(+))* can result in stochastic re-silencing through mechanisms unknown. Furthermore, poly-A selected RNA-seq analysis reveal upregulation of the Argonautes ALG-3/4 and WAGO-10 in animals expressing a transgene containing mCherry. ALG-3/4 and WAGO-10 are male-enriched and are known regulators of the endogenous *C. elegans* 26G-RNA pathway (Charlesworth et. al., 2021). We hypothesize that such upregulation promotes the perturbation upon mating and could aid in the establishment of long-term RNA silencing of the paternally inherited copy. Ribo-depleted total RNA-seq analysis of six genes that contain *mcherry* reveals 66 differentially expressed genes. 12 of the 66 misregulated genes were highly affected by the loss of regulators of RNA silencing within the germline. Prior to this work, RNA-seq analysis of mutants that lack the dsRNA importer SID-1, revealed 4 misregulated genes that are shared with the above list (Shugarts et. al., 2023). Therefore, we refer to these subsets of genes as *sid-1-dependent genes* (*sdgs)*. We also found that there are trace amounts of sense and anti-sense reads that mapped to the *mcherry* sequence suggesting that there are low levels of dsRNAs made against *mcherry*. We hypothesize that these *sdg* genes play an important role in the sensing and processing of dsRNAs. It is possible that such trace amounts of *mcherry* dsRNA could suppress the activity of HRDE-1-dependent

silencing loops, thereby reducing transgenerational gene silencing. Then upon mating, such activity is perturbed and therefore we see long-term RNA silencing.

3.3. Materials and Methods

3.3.1. Strains and oligonucleotides

Table 3-1. Strains

Strain	Shorthand name	Genotype/Allele	5'→ 3' sequence
N2		wild type	
EG6787	*T*	*oxSi487*	*mex-5p::mCherry::h2b::tbb-2 3'utr::gpd-2 operon::sl2::gfp::h2b::cye-1 3'utr + unc-119(+) II; unc-119(ed3) III?*
AMJ552	*iT*	*oxSi487 dpy-2(jam33) II; unc-119(ed3)? III*	*mex-5p::mCherry::h2b::tbb-2 3'utr::gpd-2 operon::sl2::gfp::h2b::cye-1 3'utr + unc-119(+) II; unc-119(ed3) III?*
AMJ577	*hrde-1(-)*	*hrde-1(tm1200)*	*hrde-1(tm1200)[4x]*
AMJ1208	*mcherry::mex-5*	*jam197*	*mex-5p::mCherry::mex-5::mex-5 3'utr IV*
AMJ1209	*TcherryΔpi*	*jamSi39*	*mex-5p::mCherry (without piRNA sites)::cye-1 3' utr] II; unc-119(ed3) III*
AMJ1170	*Tcherry::cye-1 3'utr*	*jamSi37 II; unc-119(ed3) III*	*mex-5p::mcherry::cye-1 3'utr ch II; unc-119(ed3) III*
AMJ1272-75	*Tcherry::mex-5 3'utr; hrde-1(-)*	*jamSi45-48 II; unc-119(ed3) hrde-1(tm1200)*	*mex-5p::mcherry::mex-5 3'utr ch II; unc-119(ed3) III hrde-1(tm1200) III*
AMJ1414-20	*Tcherry::mex-5 3'utr*	*jamSi45 II*	*mex-5p::mcherry::mex-5 3'utr ch II; unc-119(ed3) III*

Table 3-2. Oligonucleotides

Name	Sequence 5'→3'	Genotyping
P1	gct atg gct gtt ctc atg gc ggc gtc gcc ata ttc tac tt cacacacacacacacaca	RT-primer pUG PCR
P2	atggtctccaagggagag	*mcherry* FOR 1 pUG PCR

P3	gagaggaggataacatggct	*mcherry* FOR 2 pUG PCR
P4	atgagtaaaggagaagaacttttca	*gfp* FOR 1 pUG PCR
P5	ttcactggagttgtccca	*gfp* FOR 2 pUG PCR
P6	gag ttc tac gat cac att ct	*gsa-1* FOR 1 pUG PCR
P7	cac ttg ctg gaa aga caa gg	*gsa-1* FOR 2 pUG PCR
P8	gacaggccaattccgataaaag	Alg-3 RT primer
P9	ccaatggatctttatcacctcc	*alg-3* FOR
P10	ccaattccgataaaagaagaggg	*alg-3* REV
P12	gacaagccaactccgataag	*alg-4* RT primer
P13	catctcgtccagcgtcag	*alg-4 FOR*
P16	ccaactccgataagagaagagg	*alg-4 REV*
P17	gtggagcgggtcttttc	*wago-10 RT primer*
P18	atgccgactaccgtcca	*wago-10 FOR*
P19	ggagcgggtcttttcgtt	*wago-10 REV*
P20	cgtggcacatactttccgttgttg	*tbb-2 RT primer*
P21	gtcatctccgacgagcac	*tbb-2 FOR*
P22	ttccgttgttggcttcgttg	*tbb-2 REV*

3.3.2. Pol-A+ RNA sequencing and analysis

Daphne Knudsen grew N2 and AMJ1170 animals on NG plates. Just before animals were starved, she washed the worms in M9 and collected them as worm pellets for total RNA extraction using TRIzol (Fisher Scientific). PolyA+ RNAs were purified and converted to DNA libraries by Omega Bioservices using the Illumina TruSeq Library Preparation Kit. FASTQ files were processed using the command "cutadapt -j 0 -a AGATCGGAAGAGCACACGTCTGAACTCCAGTCA-AAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT- m 20 -q 20 -o cutread.gz fasta1.gz" (Martin, 2011). Reads were assigned Transcript IDs and counted using the

command “salmon quant -i celegans.index -l A -r cutread.gz -p 8 –validateMappings -o quant_file” (Patro et al., 2017). Transcript IDs were converted to gene IDs using tximport (Soneson et al., 2015) in R and genes with > 0.1 counts per million for at least 2 samples were used for subsequent analysis. Samples were normalized using the trimmed mean of M-values method (Robinson & Oshlack, 2010). Differential expression analysis was performed using limma(voom) (Law et al., 2014) in R (R script in Supplementary File 2). Volcano plots of differential expression for all genes compared between N2 and AMJ1170 were plotted using custom R scripts with genes having an adjusted *p*-value threshold (*q*-value) less than 0.05 in maroon or blue and those greater than 0.05 in grey.

3.3.3 Ribo-depleted total RNA sequencing without poly-A+ selection

Mixed-stage animals were washed from 6 plates in biological triplicates about 5 days after passaging L4-stage animals. Animals were washed with M9 buffer 3x and pelleted for RNA extraction. Total RNA was extracted from the worm pellets using TRIzol (Fisher Scientific) and sent for sequencing by the University of Maryland Brain & Behavior Institute Advance Genomic Technologies Core. Illumina Stranded Total RNA w/ Ribo-Zero Plus kit was used for library prep. IDT oligos for C. elegans-specific rRNA depletion was used for ribo depletion. TapeStation and Quibit plus was used to assess the quality of the RNA before sequencing.

3.3.4. RNA-seq analysis from total RNA

The fastq files obtained after total RNA-seq were analyzed as in Shugarts et al., (BioRxiv, 2023). Adapters were removed using cutadapt, reads were assigned transcript IDs and counted using Salmon. Genes with more than 0.1 counts per million for at least 3 samples were used for subsequent analysis. Differential expression

analysis was performed using limma(voom) in R (example available at AntonyJose-Lab/Shugarts_et_al_2023 on github). Mapping to the ce11 genome or the plasmid pCFJ359 was performed using Bowtie2 with default options. Mapped reads were then analyzed using custom R scripts. For identifying the genes that are changed in strains that have mCherry, lists of genes with >2-fold change in expression when compared with expression in wild-type (N2) animals were analyzed to select genes that are changed in four or more comparisons (total 66 genes). For comparing genes changed in the presence of mCherry with those identified in previous studies where RNA silencing within the germline is perturbed, the past lists were overlapped with the mCherry-dependent genes and sorted to indicate the genes most affected. In addition to genes from previous studies, genes co-regulated with F07B7.2/W09B7.2 as indicated by the SPELL database were also included with an arbitrary value of 2-fold up regulation (similarly arbitrary values of 2-fold were used when past studies did not declare the exact changes in expression in their >2-fold changed data tables).

3.3.5. Analysis for mapped sense and antisense reads

After mapping sense and antisense reads from wild type (N2) and strains containing mCherry, the coverage was calculated after mapping using bowtie2. Antisense transcripts were counted by mapping to the corresponding antisense transcriptome. The extent of reads mapping to each gene (*mCherry, gfp, and h2b*) were normalized for the different strains based on the number of reads mapping to the *ce11* sense genome before representing as heatmaps.

3.3.6. RT-qPCR and statistical analysis

L4-staged strains were passaged into 6-10 NG plates and allowed to crowd. When the animals reached near starvation by depleting most of the bacteria, plates were washed 3X with M9 buffer and spun at 14,000 rpm for 1 min. 6 plates of worms were pooled into a final worm pellet of 50-100 µl. RNA was extracted from these worm pellets using TRIzol according to the manufacturers protocol. About 1000-5000ng of total RNA was used for cDNA synthesis using SuperScript III (Invitrogen). Gene specific RT primers were used for reverse transcription. qPCR was performed on BioRad qPCR thermocycler. Total reaction consists of 1-2 ul of cDNA, SYBR green (Roche) and gene-specific forward and reverse primers. Raw qPCR Ct values were exported into excel. The standard 2^{-Ct} method was used for plotting. Median of Ct values were used and Student's t-test for unequal variance was used to determine statistical significance.

3.4. Results

3.4.1. Restoring HRDE-1 function does not guarantee re-silencing of a transgene susceptible to spontaneous silencing by piRNAs

In Chapter 2, we discussed stochastic silencing by MosSCI. This phenomenon requires piRNA/PRG-1 for initiating silencing after injection but not for maintaining silencing in subsequent generations. Instead, HRDE-1 is required to maintain heritable RNA silencing. In Devanapally et. al., (2021), they discovered that a few MosSCI transgenes which contain the same *mcherry* and germline promoter sequence *mex-5* are susceptible to spontaneous silencing. However, removal of HRDE-1 could rescue

expression. They tested the expressed transgenes in a *hrde-1(-)* background for mating-induced silencing and revealed that it too is susceptible to such silencing despite changes to the 3'utr sequence. Victoria Murphy, an undergraduate I mentored, showed that upon restoring the loss of function of HRDE-1 in one of the transgenes *Tcherry::mex-5 3'utr*, stochastic re-silencing can occur. Some progeny remained expressed, and others remained silenced. It is unclear why stochastic silencing occurs in the first place, but these data suggest that that presence of HRDE-1 does not guarantee silencing. Other regulators that interact with HRDE-1 are perturbed by mating and these unknown determinants could be distributed at random among siblings. Perhaps there is a competition between protective and silencing regulators. Further investigation is required to understand this phenomenon.

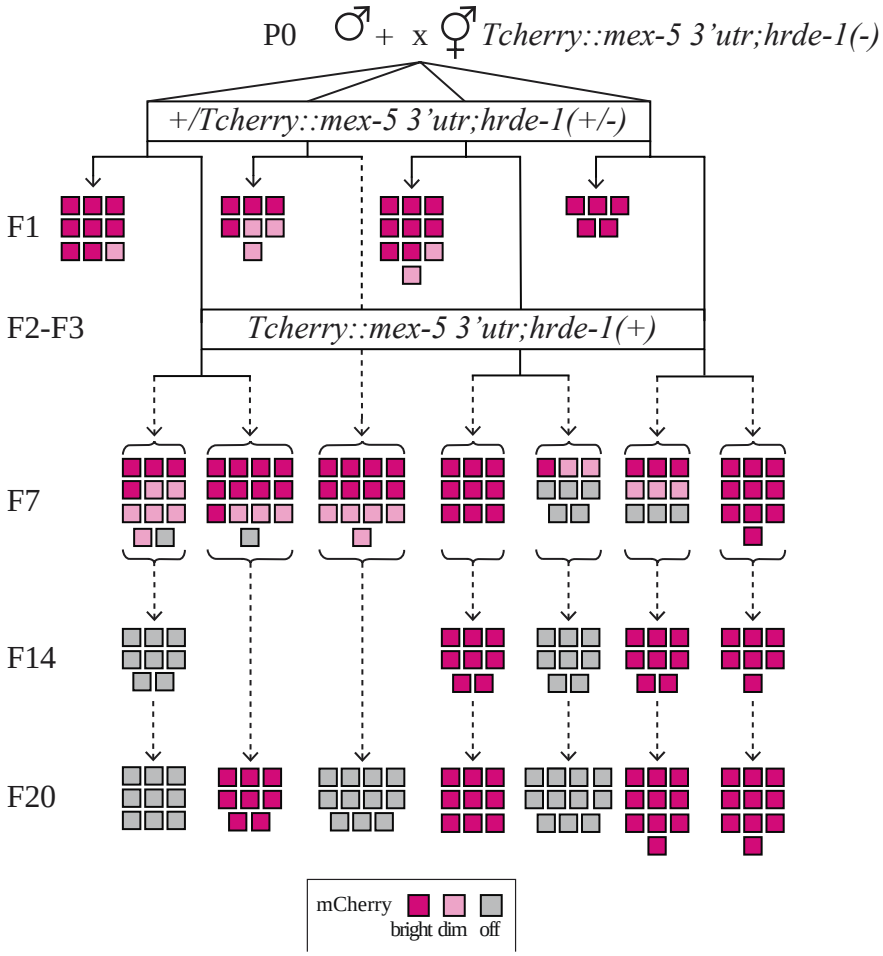

Figure 3-1. Restoring the function of HRDE-1 can result in stochastic silencing of genes susceptible to piRNA-dependent silencing.

Wild-type males were crossed with hermaphrodites expressing *Tcherry::mex-5 3'utr; hrde-1(-)*. Each line segment branching from a cross represents one isolate. F1 cross progeny males were imaged and scored. F2 and F3s were genotyped for homozygosity of both *Tcherry::mex-5 3'utr* and a wild-type copy of *hrde-1(+)*. Subsequent progeny from each homozygous line were passaged in an unbiased manner, imaged, and scored as either bright, dim, or off for mCherry expression. Each square represents a single animal.

3.4.2. Expression of a transgene susceptible to mating-induced silencing is associated with downregulation of sperm-enriched Argonautes

One of the *mCherry*-containing genes referred to as *Tcherry::cye-1 3'utr* is susceptible to mating-indued silencing. Progeny that only inherit a paternal copy of the transgene are susceptible to silencing. To determine which genes are misregulated in this background, poly-A RNA-sequencing was performed on total RNA extracted from a pool of animals expressing *Tcherry::cye-1 3'utr*. In comparison to wild type, animals expressing *Tcherry* showed downregulation of three sperm-enriched Argonautes, ALG-3, ALG-4, and WAGO-10. To verify the downregulation of these genes, I performed RT-qPCR from total RNA without poly-A selection. In three biological replicates, I found that there was a downregulation of these genes in *Tcherry* compared to wild type. However, the extent of the fold-decrease of ALG-3, ALG-4, and WAGO-10 was different from the fold-decrease observed using the poly-A+ RNA-seq analysis.

Our data suggests that perhaps the difference in poly-A selection could account for the differences in fold-change. How down-regulation of ALG-3, ALG-4, and WAGO-10 might increase susceptibility to mating-induced silencing is unclear. Both ALG-3 and ALG-4 are required for 26G small RNA pathway native to *C. elegans*. Loss of AGL-3/-4 results in a decrease of 26G RNAs and upregulation of spermatogenesis-expressed transcripts (Conine et. al., 2010). An uncharacterized Ago, WAGO-10, was also found to target spermatogenic genes with CSR-1a isoform (Charlesworth et al 2021). CSR-1 is an Argonaute that is often associated with the 22G-RNA pathway (Chaves et. al., 2021). It has the potential to play as both a positive (licensing) and negative (silencing) regulator. There are two isoforms – a and b. Although CSR-1b has

been extensively studied in the germline of hermaphrodites, new information on CSR-1a reveals its role in spermatogenesis. Both CSR-1a and WAGO-10 integrate with ALG-3/ 4 in the 26G RNA pathway to promote male fertility. We hypothesize that the upregulation of spermatogenic genes in animals containing *Tcherry::cye-1 3'utr* could contribute to the susceptibility of stable RNA silencing by mating. Mating somehow perturbs the levels of spermatogenic genes by unknown interactions between the piRNA/PRG-1 pathway and 26G RNA pathway to enhance susceptibility to long-term RNA silencing.

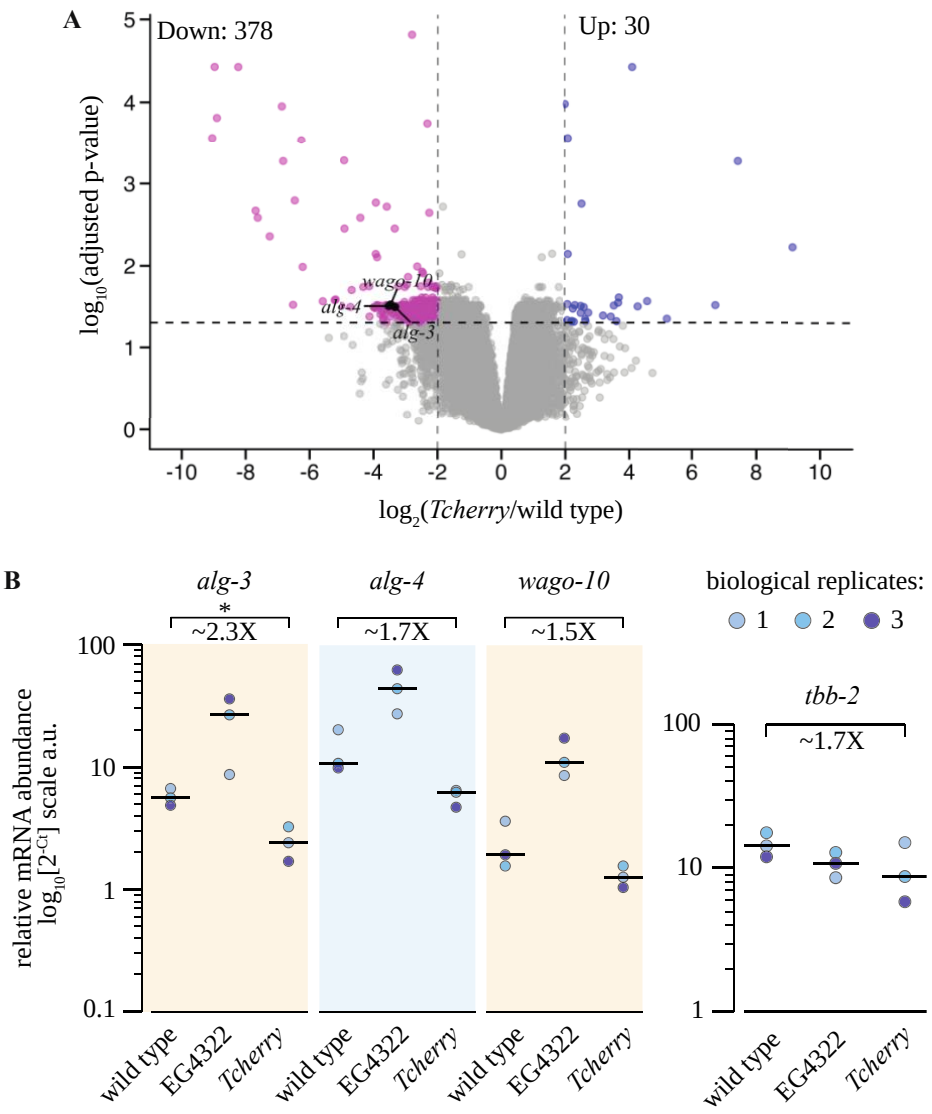

Figure 3-2. Poly-A+ RNA sequencing of animals with an *mCherry* gene reveals downregulation of sperm-enriched Argonautes.

(**A**) Volcano plot of $\log_2$ fold-changes(FCs) in abundance of polyA+ RNA in animals expressing *Tcherry* vs wild type animals. Magenta *q*-values <0.05 and blue log2FC(<2). *alg-3* ($\log_2$FC = -3.3), *alg-4* ($\log_2$FC = -3.5), and *wago-10* ($\log_2$FC = -3.5) are among the genes reduced in animals expressing *Tcherry*. (**B**) *left*, RT-qPCR measurement of *alg-3*, *alg-4*, and *wago-10* mRNAs in wild-type animals, the background into which *Tcherry* was inserted (EG4322), and animals expressing *Tcherry*. Relative mRNA abundance plotted. Fold-change plotted above. * p-values <0.05. *Right*, RT-qPCR of the control gene *tbb-2* in wild type animals, strain EG4322, and animals containing *Tcherry*. Each dot represents one biological replicate.

3.4.3. Total RNA-sequencing analysis reveals a potential role for dsRNA regulators in mating-induced silencing

Devanapally et al. (2021) revealed *mCherry*-containing genes with different sequence contexts that are susceptible and resistant to stable RNA silencing by mating. Table 3-1 summarizes a list of these candidate genes. To reveal unknown regulators that promote or inhibit stable RNA silencing we performed total RNA sequencing after ribosomal RNA depletion on animals with this list of candidate genes (Table 3-1). Between the expressed and unexpressed state of *T,* we expect to find misregulated genes that are different in the off state compared to the on state. Similarly, we'd expect to find such difference between *mCherry* variants that are resistant to silencing such as *Tcherry∆pi* which lack the piRNA-binding sites required for initiating mating-induced silencing, and *mcherry::mex*-5, the endogenous sequence context that can resist silencing and variants that are susceptible to silencing such as *Tcherry::cye-1 3'utr* and *Tcherry::mex-5 3'utr*. However, that was not the case.

Despite no obvious differences of misregulated genes between those that resist silencing and those that cannot, sixty-six genes were differentially expressed in at least four or more of the six *mCherry*-containing genes compared to wild type. We wanted to verify the misregulation of a few candidate genes that showed a greater $\log_2$-fold change by RT-qPCR, so we chose *cTel55X.1*, *far-3*, *F07B7.2,* and *scrm-4* respectively in two different *mCherry backgrounds* (Fig. 3-3B). From the RNA-seq data, *cTel55X.1*, and *far-3* were highly upregulated in both *Tcherry* and *mCherry::mex-5* compared to wild type. However, by RT-qPCR, *cTel55X.1*, and *far-3* were not significantly upregulated in these *mCherry* backgrounds compared to wild type. Similarly, we did

not observe any significant differences for two downregulated genes, *F07B7.2* and *scrm-4* in the *mCherry* backgrounds compared to wild type. Our data suggest that RT-qPCR may not be a reliable method.

Table 3-3. List of mcherry containing genes that can resist or are susceptible to stable RNA silencing by mating

Gene	Epigenetic state	Susceptible to mating-induced silencing?
T	On	Yes
iT	Off	-
Tcherry::cye-1 3'utr ; hrde-1(-)	On	Yes
Tcherry::mex-5 3'utr; hrde-1(+)	On	Yes
TcherryΔpi	On	No
mcherry::mex-5	On	No

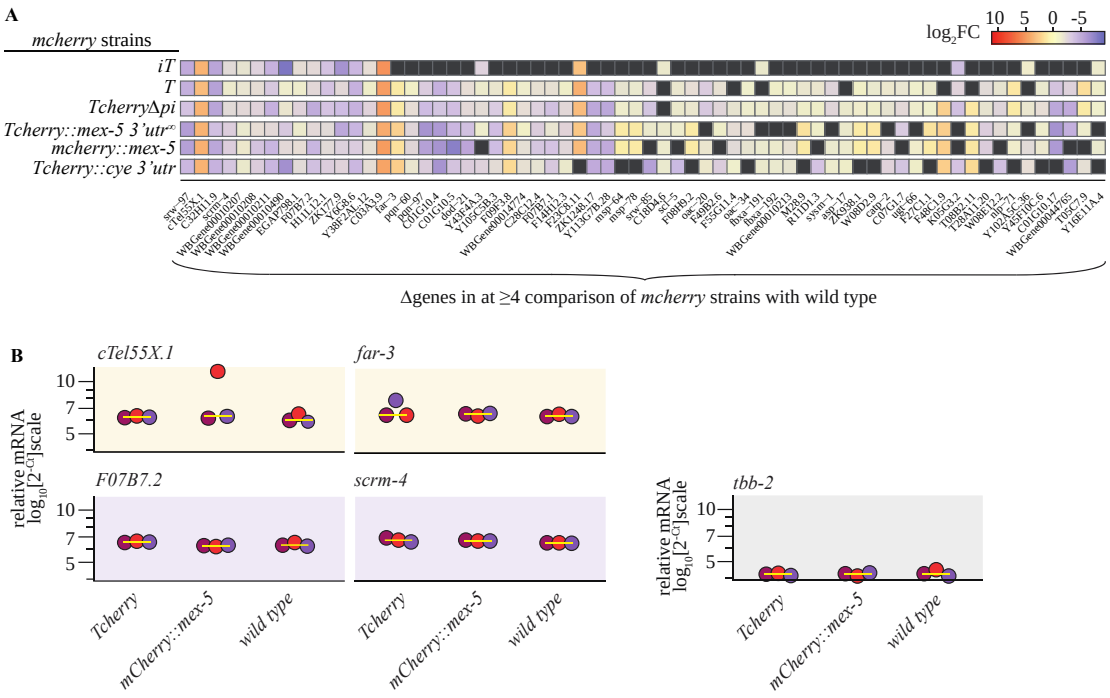

Figure 3-3. Differentially expressed genes in six strains with *mCherry* identified using RNA-seq and verification by RT-qPCR.

(**A**) Heat-map of 66 genes with >2-fold change in at least four or more *mCherry* backgrounds compared to wild type. Black box indicates adjusted p-values >0.05 or not present in corresponding dataset. (**B**) RT-qPCR plotting relative mRNA abundance of four candidate genes (*cTel55X.1*, *far-3*, *F07B7.2*, and *scrm-4*) in two different *mcherry* strains – *Tcherry* and *mCherry::mex-5* and in wild type animals. *tbb-2* mRNAs were used as a control. Student's t-test for unequal variance was used to test for significant differences.

The majority of the 66 misregulated genes were also misregulated upon loss of different known RNAi regulators (Fig 3-5., comparison with 21 different RNA-seq datasets). Loss of PRG-1 results in a decrease of these differentially expressed transcripts, but in an increase in 22G RNAs suggests that piRNAs and PRG-1 could prevent the silencing of *mCherry*-containing genes. However, loss of HRDE-1 results in an upregulation of these transcripts suggesting that these downregulated genes in *mCherry* variants are targets of HRDE-1. HRDE-1 acts to repress these transcripts to some extent such that mCherry is still expressed when the mRNA levels of these genes are reduced. Furthermore, there are different species of 22G-RNAs that are dependent on different RNAi regulators. The production of these HRDE-1-bound populations of 22G RNAs requires the mutator foci, regulated by MUT-16, and P granules, regulated by MEG-3, MEG-4, and DEPS-1. Consistently, the loss of these regulators results in a downregulation of 22G RNAs that target these differentially expressed transcripts.

Curiously, the loss of two dsRNA interacting regulators results in an upregulation of these *mCherry*-dependent RNA transcripts suggesting that a subset of these differentially expressed genes are dependent on the dsRNA pathway. Shugarts et. al., (2023) revealed that one of these genes, *F07B7.2* (that also came up from this dataset) is dependent on SID-1, a transmembrane protein required for the import of dsRNA into the cytosol. Therefore, they renamed this gene *sdg-1* (*sid-1 dependent gene-1*) and we will refer to this subset of *mCherry*-dependent genes as *sdg*s. Our data suggests a unifying hypothesis whereby these *sdgs* could be acting as processors of dsRNAs and these processed dsRNAs require HRDE-1. Shugarts et. al., (2023) also revealed that SDG-1 localizes with Z granules and such localization could be used for

sensing dsRNAs. Therefore, we propose that susceptibility to stable RNA silencing could depend on a synergy between the sensing and processing of dsRNAs.

Interestingly, mapping sense and anti-sense RNAs from the total-RNA seq results reveal trace amounts of dsRNA against both cistrons in *T*. Additionally dsRNA against *mCherry* was also observable in all *mCherry*-containing genes suggesting that these trace amounts of dsRNAs against *mCherry* could be what *sdgs* are sensing and processing. Furthermore, because of the presence of pUGylated RNAs in both the on and off state of *T*, it also supports the idea that dsRNAs are made in the presence of *mCherry*.

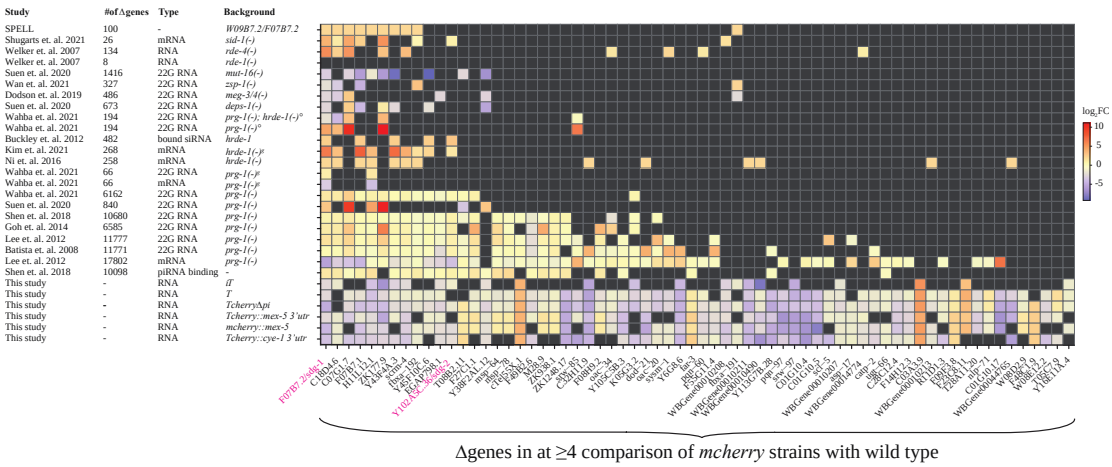

Figure 3-4. Comparison of 66 differentially expressed genes in six *mCherry* strains to different published RNA-seq datasets.

Left, the number of genes showing differences (# of Δgenes) in 23 different RNA-seq datasets generated using small RNA-seq, mRNA-seq, Crosslinking, Ligation, and Sequencing of Hybrids (CLASH, piRNA binding), microarray analysis (*rde-4(-)* and *rde-1(-)*), or Serial Pattern of Expression Levels Locator analysis (SPELL, W09B7.2/F07B7.2). *Right*, heat-map of 66 genes with >2-fold change compared to RNA-seq datasets from previous studies. The pink text highlights two genes recently identified as dependent on SID-1. *sdg-1* was most frequently affected among the studies considered. SPELL (Serial Pattern of Expression Levels Locator) uses a gene list to identify other genes with similar expression profiles across many experiments. In this case, F07/B7.2/W09B7.2 was used as a query and the top 100 genes with similar expression profiles were included. The 66 genes were ordered based on the frequency of occurrence in the different datasets considered. The black box indicates adjusted p-values >0.05 or not present in the corresponding dataset.

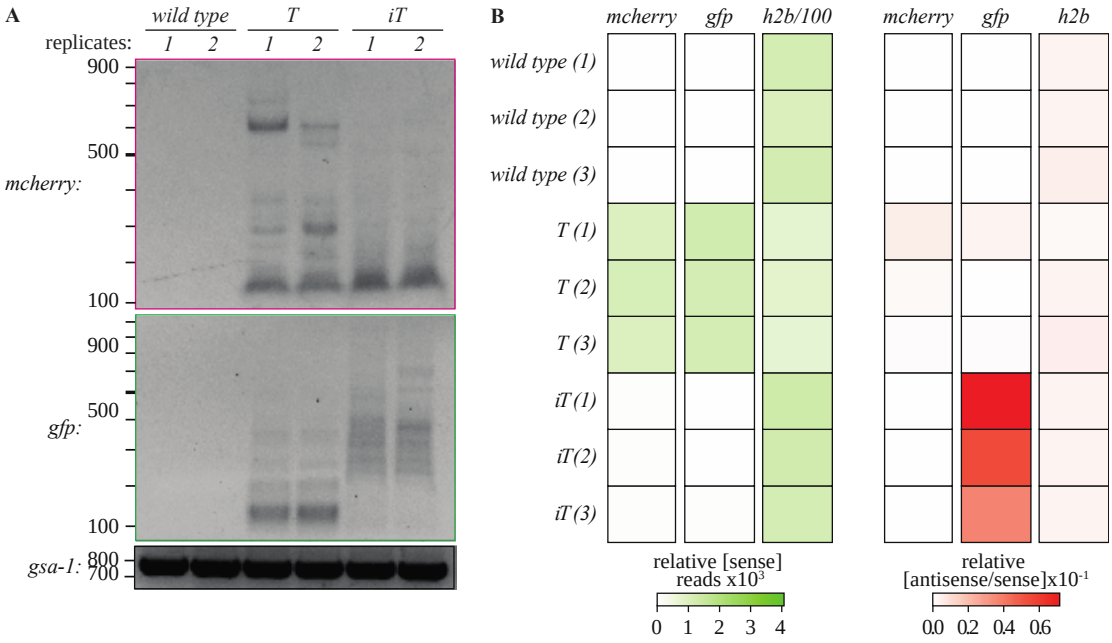

Figure 3-5. Antisense RNAs made from an expressed gene could prime the gene for long-term RNA silencing.

(**A**) pUG PCR against *mcherry*, *gfp* and *gsa-1* in wild type, animals with *T* or with the epiallele *iT*. *mCherry* and *gfp* pUG-RNAs are detectable in both the unexpressed and expressed state of *T* but not in wild type, which lack sequences for *gfp* and *mCherry*. *gsa-1* was used as a control and results in a single product of ~750 nt. RT-PCR was performed on total RNA extracted from two biological replicates (*1* and *2*). (**B**) Relative sense (green) and antisense (red) reads mapping to the open-reading frames of *mcherry* and *gfp* in wild type (none expected), and in animals containing *T* or *iT*. Downsampling of mapped *h2b* reads by 100x was necessary to generate the heatmap where the sense reads can be displayed together for all sequences (left heatmap). Relative antisense reads were quantified separately (right heatmap). Rate of antisense RNAs was uniform for *mcherry*, *gfp* and *h2b* in all backgrounds except in animals containing *iT* which had 10x more relative antisense RNAs. This increase was driven by a reduction in sense RNA rather than a change in the amounts of antisense RNA.

3.5. Discussion

Mating can perturb the environment and result in epigenetic changes that can last almost indefinitely. Devanapally et. al., (2021) revealed a few gene contexts that enhances the susceptibility of a single reporter gene, *mcherry,* to heritable RNA silencing. *mcherry* is susceptible to spontaneous silencing upon MosSCI injections because of the presence of piRNA binding sites predicted throughout the coding sequence, but the presence of piRNA binding sites alone doesn't guarantee silencing. Not all *mcherry* genes are silenced upon injection. We and other labs show that loss of HRDE-1 can rescue the expression of a gene susceptible to such spontaneous silencing upon injection. However, restoring the function of HRDE-1 can also result in stochastic re-silencing through mechanisms unknown. It is unclear what regulators and interactions are perturbed upon restoring a mutation, but our data suggests that the gene is in a different epigenetic state although it appears wild type. We hypothesize that there is a threshold of interactions between regulators such as piRNAs, PRG-1 and HRDE-1 that is distributed at random among progeny, which dictates whether mCherry is expressed or not.

We also reveal that upon the presence of an *mCherry*-containing transgene (*Tcherry::cye-1 3'utr*) there is a downregulation of sperm-enriched Argonautes. Perhaps such downregulation is important for why we observe mating-induced silencing in one direction – when males expressing the transgene are crossed with hermaphrodites lacking the gene. These data suggest that sperm-enriched mRNA transcripts could be elevated in addition to the number of increased endogenous 26G RNAs in males expressing the transgene. Hermaphrodites, which do not contain the

transgene may not have these elevated levels. Upon mating, perhaps the activity of these Agos is increased and in conjunction with the piRNA pathway, silence the transgene.

Total-RNA sequencing of six strains containing the *mcherry* sequence revealed a few differentially expressed genes that significantly overlap with regulators of the dsRNA pathway. Furthermore, there is evidence of trace amounts of dsRNA production against the reporter gene *mcherry*. Additionally, these antisense *mcherry* fragments could be why we see remnants of pUGylated *mcherry* transcripts (pUG PCR data), suggesting that at some rate, *mcherry* transcripts are being cleaved, appended with UG nts, and are leading to the production of secondary small RNAs. However, not to an extent we see silencing in these *mcherry* strains because they remain expressed. These subset of differentially expressed genes now provisionally renamed *sdg*s could be acting at both the RNA and protein levels to enhance the susceptibility to long-term RNA silencing through mechanisms unknown. We hypothesize that these *sdgs* could be acting as sensors and processers of dsRNAs and mating could enhance this synergy between sensing and processing such that heritable RNA silencing is achieved.

Chapter 4: General Discussion

4.1 Introduction

The effects of the environment have been correlated with molecular changes in animals and in their descendants. However, due to the complex interplay between regulators and their interactions with DNA, it is difficult to clearly explain the epigenetic machinery required for both intergenerational (parent to progeny) and transgenerational (parent to grand progeny) effects. In this dissertation, we investigate a single gene expression pattern across generations. We demonstrate that 1) two similar gene contexts driven by the same promoter can have different epigenetic outcomes despite similar perturbations and 2) the presence of non-endogenous elements is detected and processed by parallel RNA interfering pathways that can enhance the duration of RNA silencing across generations.

4.2 Implications of independent transgenerational regulation of each cistron in a two-gene operon

The most notable operon, the *lac* operon was first coined by Jacob and Monod in the bacteria *E. coli*. This revealed a level of gene organization that is effective for function. Since their discovery, gene clusters that are operon-like have been reported in eukaryotic systems such as filamentous fungi, plants, and animals. *C. elegan*s, however, is one member of the animal kingdom that remarkably resembles the structure of prokaryotic operons (Blumenthal et. al 2005-2018 Wormbook). About 15% of the nematode's genes (2,600 genes) are operons (Blumenthal & Gleason, 2003) which are comprised of linked or clusters of 2-8 functionally similar but non-homologous genes

under the control of a single promoter. Unlike prokaryotic operons, *C. elegans* operons are transcribed as polycistronic pre-mRNA transcripts and processed by trans-splicing events lead by two types of splice leader (SL) sequences – SL1 and SL2 (Blumenthal 2005-2018 Wormbook). This results in mono-cistronic mRNA transcripts, that are translated separately. Although most operons are regulated by SL2 splicing, SL1 splicing occurs in more than half of the mRNAs.

About 70% of mRNAs in *C. elegans* contain a 22nt sequence transcribed from a different sequence and spliced onto the mRNA, called the splice leader (SL) sequence (Blumenthal 2005-2018 Wormbook). This sequence is donated by a 100 nt RNA called the SL RNA by trans-splicing and exists as small nuclear ribonucleoproteins (snRNPs). SL1 trans-splices at the 3' splice site following outrons. Outrons are intron-like sequences present at the 5' end of a primary transcript. It contains an "Ou" element that is capable of base pairing to the 5' splice site of the SL snRNP but lacks the functional 5' splice site upstream. The presence of an outron determines whether the gene is trans-spliced or not.

SL2 is donated by the SL2 snRNPs and becomes the 5'end of the mature mono-cistronic transcript. Trans-splicing requires the Ur element found within the 100 bp sequence located between each cistron in the operon. The Ur element base pairs with the 5' splice site of the SL2 snRNP and protects the downstream gene from degradation following the 3' end cleavage of the upstream gene. The "AAUAAA" sequence at the 3'end of a cistron acts as a signal for 3' end cleavage and polyadenylation (Blumenthal 2005-2018 Wormbook). The polycistronic pre-mRNA is converted then to mono-cistronic mRNAs. It is likely these trans-splicing events are what led to the abundance

of operons present in nematodes which is different from operons in prokaryotes. The origins of trans-splicing is unclear, but operons have been identified in *Drosophila* which lack this tans-splicing regulation. Instead, operons are transcribed into di-cistronic mRNAs that are translated directly.

Few operons are conserved across bacteria and lack homology. It has been reported in bacteria that operons can split apart or cistrons can be deleted which can impact gene expression. The destruction of operons has also been seen in nematodes. In *C. elegans* about 96% of the 900 known operons are also conserved in the closely related species *C. briggsae* which is thought to have diverged 50-100 million years ago (Blumenthal et. al 2005-2018 Wormbook). Rearrangements between *C. elegans* and *C. briggsae* chromosomes do not occur more frequently for genes within operons than it does for other protein-coding genes and there is evidence for operons that have formed in *C. elegans* which have been destroyed in *C. briggsae*. There are other types of operons present in *c. elegans* such as "hybrid" and "alternative" operons. There are two types of hybrid operons. One is subject to a higher rate of SL1 splicing due to an additional internal promoter within the operon structure. The second contains a long (typically >500bp and rarely 2kb) inter-cistronic region and downstream genes are SL2 splicing. Alternative operons can result in three types of products: 1) full-length mRNAs that match the entire gene, 2) mRNAs that match the 5' end of the gene due to 3'end formation within the gene, and 3) mRNAs that match the 3'end of the gene due to SL2 trans-splicing downstream of the 3'end formation site.

These different types of operons reveal the effects of trans-splicing on gene expression. It is possible that in the case for which we see one cistron expressed while

the other remains off could be due to aberrant trans-splicing events in addition to the differences in RNAi regulators.

4.2 Different gene silencing phenomena have similar but different epigenetic regulators and interactions that promote heritable effects

The expressed or unexpressed state of a gene is spatially and temporally controlled by different epigenetic mechanisms. Not all gene silencing is an effect of the same type of regulation, and this could be dictated by differences in the regulatory sequences of a gene. Importantly, the type of perturbation and record of the perturbations throughout development and across generations is important for understanding why some genes become more susceptible to permanent silencing than others. Identifying which regulators and interactions occur within different silencing and protective mechanisms provides insight into the past events or experiences of a gene.

As mentioned in earlier chapters, the two well-studied RNA germline silencing pathways in *C. elegans* can defined by the source of the trigger – dsRNAs or piRNAs. The dsRNA and piRNA pathway share similar regulators such as RDE-3/MUT-2 and HRDE-1 (Chapter 2 and 3). Additionally, there is the 26G RNA pathway which can be categorized into two different classes – class I are sperm-specific 26G RNAs and class II are highly enriched in oocyte/embryos (Han et. al., 2009). 26G RNAs are endogenously transcribed from spliced mRNAs by an RNA-dependent RNA-polymerase (RdRp) *rrf-3* and partially *ego-1* (Billi et. al., 2005-2018 Wormbook). These 26G RNAs are also processed by Dicer and result in a 5’ monophosphate guanosine as opposed to the 5’ triphosphorylated 22G RNAs (another byproduct of

RdRps). Agos ALG-3/-4 are required for 26G RNAs in sperm and ERGO-1 is required for 26G RNAs in oocytes and embryos.

One known licensing pathway is often referred to as the CSR-1 pathway. CSR-1 is an Argonaute that binds 22G RNAs and protects transcripts from piRNA-induced silencing (Charlesworth et. al., 2021). It is possible that upon a different type of perturbation, these protective and silencing mechanisms compete with one another. Whether the gene sequence becomes susceptible to silencing could be dictated by past events. For example, in the case of the *mcherry* containing genes, some of these genes are susceptible to silencing. The way these genes were created and the experiences or events these genes had undergone could influence whether the gene becomes susceptible to permanent silencing. Investigating which silencing and protective regulators interact with these *mCherry*-containing genes could provide an explanation for why some are susceptible to RNA silencing by mating and others can resist.

Chapter 5: Future directions

5.1 Preface

Thoughts presented in this chapter are largely influenced by ideas published in: Chey, M. and Jose A.M. Heritable epigenetic changes at single genes: challenges and opportunities in *Caenorhabditis elegans*. Trends in Genetics 38(2):16-119.

Victoria Murphy designed and cloned transgenes containing *mcherry*. Riley Nolan, another undergraduate student I mentored, also helped integrate these *mcherry* genes into the *C. elegans* genome. Vinashya Venkatasubramani, a graduate rotation student helped add stop codons to the integrated constructs using Cas9-mediated genome editing.

Daphne Knudsen made *mut-16* single, and *mut-16; rde-10* double mutants that I used to cross with animals containing *Tswap*. The rest of the work presented in this chapter was performed by me.

5.2 Introduction

This dissertation reveals two things about heritable RNA silencing: 1) there are at least two different chromatin-independent RNA silencing mechanisms that can selectively regulate two cistrons of an operon and 2) long-term RNA silencing is modulated by dsRNA regulators. This chapter will highlight unanswered questions that could inspire future work that builds on these discoveries.

5.3 Different forms of heritable RNA silencing can be used to investigate regulatory architectures transmitted across generations

The behavior of *Tswap* is a puzzling case. This two-gene operon is susceptible to spontaneous silencing upon injection like *T.* Therefore, loss of any of the piRNA-dependent regulators such as of HRDE-1 could rescue expression as we and other labs have seen with any gene susceptible to spontaneous silencing (with the other exception of *Tind_adj*). Loss of HRDE-1 in *Tswap* resulted in re-expression of the second cistron mCherry::H2B and not the first cistron GFP::H2B. This atypical observation suggests there are two parallel silencing pathways that regulate gene expression upon injection of DNA. This second pathway is chromatin-independent and does not require HRDE-1 and other known RNAi regulators like RDE-3, DEPS-1, and ZNFX-1 (Chapter 2). The *C. elegans* genome encodes 27 Argonautes, 12 of which are worm-specific Agos (WAGOs) (Yigit et al 2006). We speculate that WAGOs can function in multiple RNAi pathways. Other candidate regulators to test would be the loss of PRG-1, WAGO-1, WAGO-4, WAGO-10, MUT-16, and RDE-10 and in different combinations with other known regulators.

Preliminarily I tested whether loss of MUT-16 or both MUT-16 and RDE-10 could result in re-expression of both cistrons, and it did not (data not shown). In both cases, similar to HRDE-1 loss, mCherry::H2B recovered expression and GFP::H2B did not. PRG-1 is a primary Argonaute required for piRNA-binding and initiation of silencing upon spontaneous silencing. To determine whether GFP silencing is dependent on piRNA/PRG-1 silencing as is mCherry, we could inject the DNA construct for *Tswap* into *prg-1(-)* mutants. WAGO-1 is a cytoplasmic germline Ago

protein (Gu et al 2009) that can partially rescue the expression of a gene subject to spontaneous silencing (Shirayama et al 2012). WAGO-1 and HRDE-1 double mutants can also partially rescue the expression of a gene. However, WAGO-10 and HRDE-1 double mutants can fully rescue the expression. WAGO-10 is a sperm-enriched Ago protein that functions with ALG-3/-4 in the 26G RNA pathway that target spermatogenic genes. It is possible that silencing upon injection relies on both the piRNA/PRG-1 and 26G RNA pathway. WAGO-4 is another cytoplasmic Ago previously shown to be required for dsRNA-induced silencing of germline genes. Similar to HRDE-1, WAGO-4 can bind 22G RNAs. These 22G RNAs are uridylated with 3'Us by the nucleotidyltransferase CDE-1 which can be bound by the Ago CSR-1 (van Wolfswinkel et. al., 2009). The CSR-1 pathway has been implicated for its role in licensing germline genes.

One quick way to induce ectopic expression of a gene is through heat-shock. In our lab, we've observed expression of mCherry::H2B in the intestine of animals expressing *T* upon heat-shock. Therefore, as a test of whether GFP::H2B can show any expression under any condition, I subjected animals containing the unexpressed *Tswap* in different mutant backgrounds to heat-shock. Preliminarily, I observed the expression of GFP from *Tswap* in a few large nuclei in the unexpressed state, *Tswap* in *rde-3, hrde-1,* and *znfx-1* single mutants and *Tswap* in *hrde-1;rde*-3 double mutant (data not shown). However, it is still unclear whether this expression is related to GFP::H2B or not. Because I only subjected the animals to a 4 hr 32°C heat shock and placed them into a water bath at 25°C for the remaining 24 hrs, it is possible that wider expression could be detectable with more extensive heat-shock.

Additionally, sequencing of the *gfp* coding sequence did not reveal any point mutations in the coding sequence that could disrupt the protein function. However, we did find a single point mutation in the *h2b* sequence following *gfp* which is not present in the *h2b* sequence following *mcherr*y. This single nt change from C to T results in an Alanine to Valine amino acid change (p.A104V). It is possible that such change could result in the misfolding or degradation of the GFP protein. To determine whether this SNP is the sole cause for why GFP is unexpressed, I will correct this point mutation using Cas9-mediated genome editing and look at the expression. If there are no causal point mutations that could explain why GFP is unexpressed, a forward genetic screen can be used to look for mutants that result in the expression of GFP. This could result in the identification of potentially new regulators of RNA silencing.

Trans silencing and *trans* activation of *T* revealed another form of RNA regulation. As mentioned previously, loss of HRDE-1 in animals containing *Tswap* could re-activate mCherry expression but not GFP. It is possible that the unexpressed GFP sequence can initiate *trans* silencing of the *gfp::mex-5* tag that was resistant to *trans* silencing when exposed to *iT* (Chapter 2). It would be interesting to look in subsequent generations to determine if the expressed *gfp::mex-5* tag could re-activate the unexpressed GFP::H2B in *Tswap; hrde-1(-).*

Furthermore, if we can get both cistrons in *Tswap* to be expressed, we could subject these animals to feeding RNAi against *mCherry* and *gfp*. We've shown that feeding *gfp* RNAi in animals expressing *T* results in the silencing of GFP followed by silencing of mCherry. If we feed *mCherry* RNAi in animals expressing *Tswap,* could

we see mCherry silencing followed by GFP silencing? How would *Tswap* behave in comparison to *T* upon feeding RNAi and how long would silencing persist?

5.4 Effect o*f cis*-regulatory sequences on susceptibility to long-term RNA silencing

We have identified a few recombinant gene contexts that are susceptible to heritable RNA silencing by mating in *C elegans*. At a minimum, these transgenes share a common open reading frame *mCherry*. However, not all *mCherry*-containing genes are susceptible to heritable RNA silencing. We hypothesize that in addition to the locus, the *cis*-regulatory sequences (e.g. promoter and 3'utr) could affect whether the gene is susceptible to mating-induced silencing or not. Future work entails creating transgenes that have different promoter and 3'utr sequences adjacent to the *mCherry* coding sequence. Four MosSCI gene constructs have been designed and cloned by Victoria Murphy (*gtbp-1p::mCherry::gtbp-1 3'utr*, *pgl-1p::mcherry::pgl-1 3'utr*, *oma-1p::mCherry::oma-1 3'utr*, *sun-1p::mCherry::sun-1 3'utr*). Riley Nolan, helped screen for integrants after injecting two of these constructs (*pgl-1p::mCherry* and *oma-1p::mcherry*) and verified the sequences of these integrants. However, these variants lacked the stop codon because of an initial design flaw and future work requires the addition of a stop codon sequence at the end of the *mcherry* sequences to assess how changes in *cis*-regulatory sequences could affect susceptibility to long-term sRNA silencing. If these transgenes are on, they can be tested for their susceptibility to mating-induced silencing. However, if they are off, removal of *hrde-1(-)* will be used to rescue the expression before testing.

Another interesting find revealed by Devanapally et. al. (2021) was when they were testing if the P granule component, PGL-1, is required for the initiation of mating-induced silencing of *T*. By crossing *pgl-1(-)* mutants with hermaphrodites expressing *T* (not the mating-induced silencing cross, but the reciprocal direction), they saw that *T* turned off in subsequent generations. To determine if *pgl-1* is required for the expression of *T*, I introduced a *pgl-1* nonsense mutation in animals expressing *T* using genome editing. After successful editing, I initially observed that *T* was expressed in the germline. However, after unbiased passaging for a few generations, I noticed that both mCherry and GFP expression in *T* turned off (after >several generations, data not shown). These data suggest that *pgl-1* is required for the expression of *T*. Future work could revert the *pgl-1* mutation to wild type and observe if the expression is rescued.

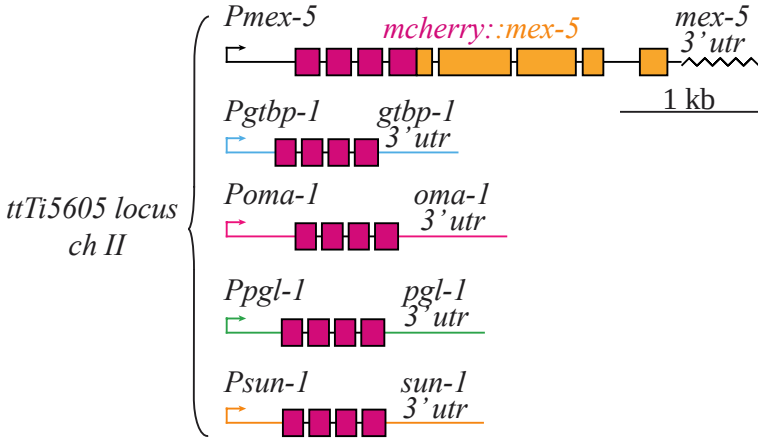

Figure 5-1. Schematics of transgenes driving expression of mCherry in the germline.

These *mCherry*-containing genes have different promoter and 3'utr sequences. Testing the susceptibility of animals with these genes to mating-induced silencing could reveal cis-regulatory sequences required for susceptibility to silencing.

5.5 Impact of dsRNA-interacting regulators on susceptibility to long-term RNA silencing

Total RNA-seq results reveal candidate regulators of RNA silencing (Chapter 2). The role of *sdg-1* has been investigated in part by Shugarts et. al. (*bioRxiv,* 2023). Candidate genes I began to investigate include the following: *sdg-1(-)/F07B7.2*, *C07G1.7, ZK177.9,* and *C18D4.6.* I've designed nonsense mutations for these genes and successfully created a few mutations in the wild type and *T* background. These can be used to test whether loss of one or a combination of these candidate genes could affect mating-induced silencing (either preventing or enhancing silencing). Additionally, I designed and integrated a couple of *mCherry* tags to express mCherry fusion proteins. These will allow us to see where these candidate genes are expressed and how that could affect heritable RNA silencing induced by mating. Preliminarily, one of the candidate genes (C18D4.6) is only detectable in neurons and not the germline. As future directions, total-RNA sequencing of other two-gene variants discussed earlier could provide insight into how dsRNA regulators could potentially enhance the direction of heritable RNA silencing.

pUGylated_RNAs can be used to define the epigenetic status of a gene. We observed in Chapter 2 and 3 that these RNAs are detectably associated with both the expressed and unexpressed state of a gene. By qualitative measuring and sequencing the pUGylated RNAs in different epigenetic states we could determine which portion of the RNAs are cleaved and pUGylated. It would be interesting to test how loss of SID-1, RDE-4, RDE-1 and the other differentially expressed genes F07B7.2, C07G1.7, ZK177.9, and C18D4.6 would affect the different pUGylation patterns we see in *T* and

iT. Would loss of any of the dsRNA regulators like SID-1 or RDE-4 affect the detectable amounts of antisense RNAs we mapped to *T*? How would this enhance or prevent stable RNA silencing?

In Devanapally et. al., (2021) they mapped the location of the protective signal where animals inheriting the paternal copy can be protected against mating-induced silencing. Similar to mapping the location of the protective signal, we could map the susceptibility signal of *T* by crossing it with recessive linkers *unc-4* and *dpy-2*. I have attempted work in this direction, but the sample number was relatively low and would need to be repeated more carefully before any conclusions can be made.

Appendices

6.1 Preface

Julianna Gross, a high school student I mentored, performed part of the work presented in this section. This was used to generate data for Fig.6-2 and Fig.6-3. One part of the appendices builds on the forward genetic screen and feeding RNAi experiments performed by Pravrutha Raman and Farida Ettefa to generate Figs.6-1.

The rest of the work presented in this section was performed by me.

6.2 Introduction

Prior work in the Jose Lab identified a single recombinant gene susceptible to heritable RNA silencing called *T.* As mentioned in previous chapters, it is an operon that expresses two reporter genes, mCherry and GFP, under the germline promoter *mex-5*. Upon mating males that express the transgene to hermaphrodites that lack the gene, the gene can turn off in subsequent generations. This heritable RNA silencing can persist for more than 300 generations. The epiallele of *T* that shows stable silencing is referred to as *iT* (where "i" denotes "inactive"). One of the known regulators of RNAi is the intrinsically disordered protein MUT-16. This regulator is required for the formation of the mutator foci, a hub of interaction between RNAs and RNA binding proteins. MUT-16 requirement for the initiation and maintenance of mating-induced silencing has been tested (Devanapally, et. al., 2021). However, it was unclear if such maintenance of silencing is maternally, paternally, or zygotically required. To test this requirement, I crossed males containing *iT* with *mut-16(-)* hermaphrodites. I segregated homozygous progeny of three different genotypes: *iT* only, *mut-16(+/-);iT*, and *mut-16(-/-);iT*. As expected, progeny with *iT* only, which has functional MUT-16, did not

re-express mCherry and GFP. Similarly having one functional copy of MUT-16 could not rescue the expression of mCherry and GFP in *iT*. However, complete loss of MUT-16 did. These data suggest that MUT-16 is zygotically and maternally required for the maintenance of stable RNA silencing initiated by mating (Fig. 6-1).

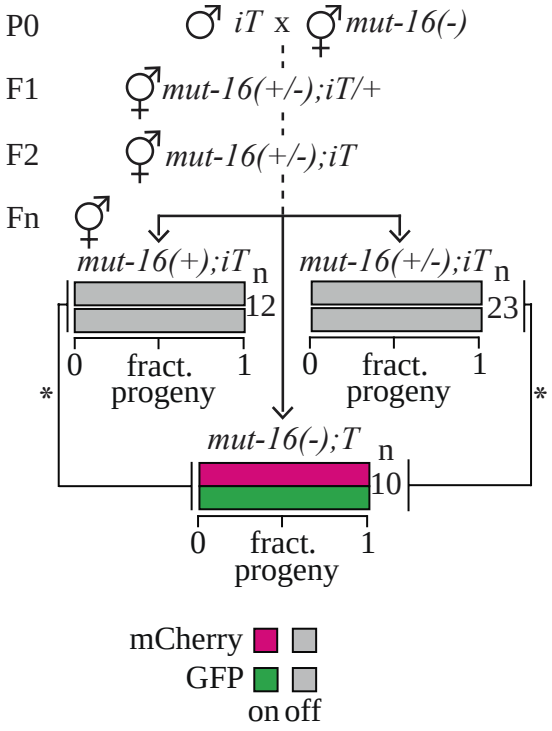

Figure 6-1. MUT-16 is maternally and zygotically required for the maintenance of silencing initiated by mating.

Males containing the unexpressed transgene *iT* was crossed with *mut-16(-)* hermaphrodites. F1s heterozygous for both *mut-16(+/-)* and *iT*/+ were isolated. F2s heterozygous for *mut-16(+/-)* and homozygous for *iT* were selected. Maternally there is one functional copy of *mut-16* that could be inherited to the next generation. Descendants with the following genotypes were segregated and scored for on or off expression for both mCherry and GFP: *mut-16(+); iT*, *mut-16(+/-);iT*, and *mut-16(-/-);iT*. Asterisks indicates p-value <0.05 (χ^2 test). Number of progenies assayed = n.

A forward genetic screen was performed on *iT* after ~150 generations of silencing by Pravrutha Raman and Farida Ettefa. Together, they mutagenized ~72,000 haploid genomes and looked for **r**e-**a**ctivation of **g**ene **e**xpression (RAGE). A total of 42 mutants, now referred to as *rage* mutants, were isolated by F3. Several defects were noted. 15 *rage* mutants were fertile, and the remaining were sterile. Of the 15 fertile mutants, some fraction of isolates had a shorter germline, enlarged nuclei, and ectopic expression. These defects could reveal native epigenetic mechanisms that rely on persistent silencing across generations in *C. elegans*.

6.3 Results

To begin understanding which regulators are responsible for maintaining heritable RNA silencing initiated by mating, the mutants were subjected to feeding RNAi against *unc-22*, *bli-1*, and *gfp*. These target genes were chosen because they could be used to categorize the *rage* mutants based on their site of action in different tissues (soma and/or germline) and subcellular locations (cytoplasm and/or nucleus). For example, both *unc-22* and *bli-1* are expressed in somatic tissues, but *unc-22* expressed in the muscles requires cytoplasmic RNAi and *bli-1* expressed in the hypodermis requires the nuclear RNAi. This difference is based on their requirement for the somatic nuclear Argonaute NRDE-3 for *bli-1* silencing but not *unc-22* silencing. On the other hand, *gfp* is expressed in the germline because of the promoter *mex-5* and can be silenced with a dependence on the nuclear Argonaute HRDE-1 or independent of it depending on the amount of dsRNA delivered.

After two generations of feeding RNAi against these three different target genes, the 15 *rage* mutants showed different susceptibility to silencing for these targets

which could be categorized into five different classes of silencing (Fig.6-1). Three classes in particular, Class III, IV, and V showed resistance to *bli-1* silencing. There is likely more than one *rage* gene affected by this screen, therefore, to begin identifying which *rage* genes are required for transgenerational silencing, we looked at the 10 *rage* mutants resistant to *bli-1* silencing because it is the easiest to score.

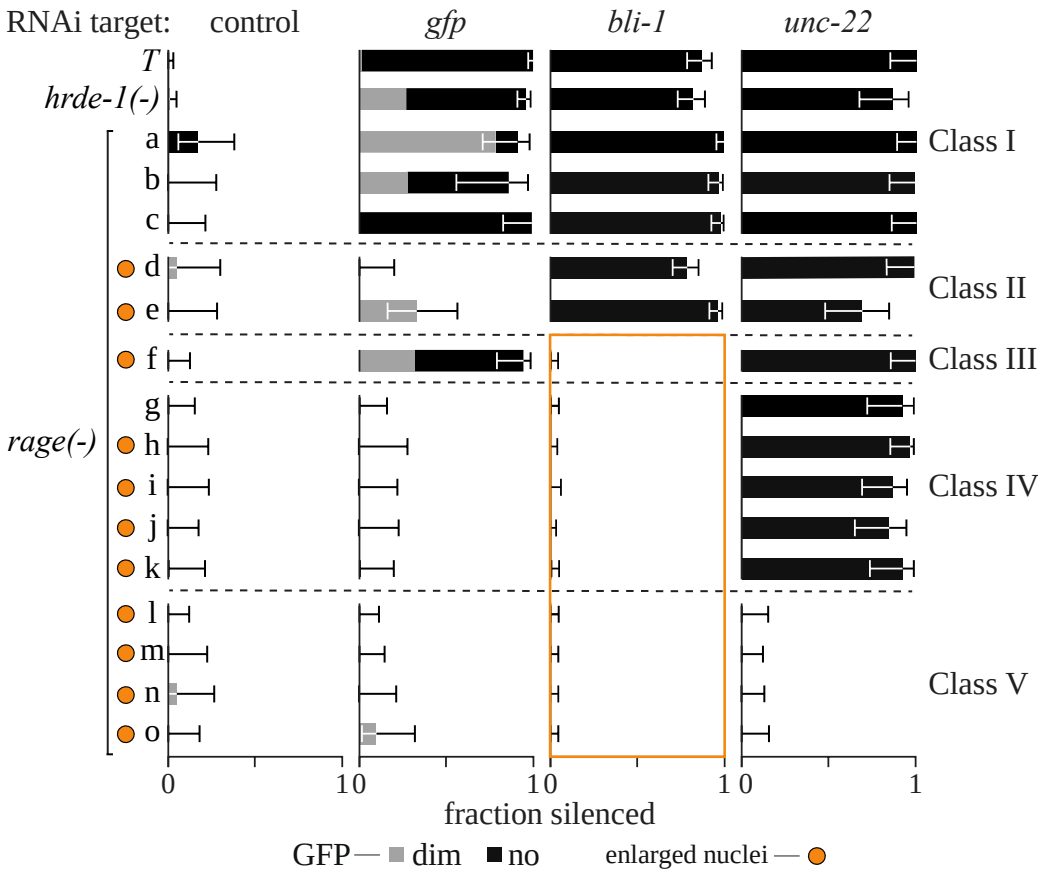

Figure 6-2. Feeding RNAi against three different target genes reveals different classes of *rage* mutants.

Pravrutha and Farida fed 15 *rage* mutants (a-o) L4440 (control), *gfp*, *bli-1*, and *unc-22* RNAi for 24 hrs and scored them for the corresponding defects, if any. The differential susceptibility of these mutants to the different RNAi experiments can be used to place them in five different classes (I to V). Error bars represent 95% confidence intervals. Orange box highlights the 10 *rage* mutants with strong defects in silencing *bli-1*, which were further investigated by Julianna Gross and I.

Table 6-1. Summary of the response of 10 rage mutants when subject to RNAi targeting three different genes

rage mutant	Class	GFP	Bli-1	Unc-22
f	III	Yes	No	Yes
g	IV	No	No	Yes
h		No	No	Yes
i		No	No	Yes
j		No	No	Yes
k		No	No	Yes
l	V	No	No	No
m		No	No	No
n		No	No	No
o		No	No	No

6.3.1. Genetic cross reveal more than one *rage* gene

I worked with Julianna Gross who performed most of the assays to re-confirm that the mutants we were working with were indeed resistant to *bli-1* silencing. We confirmed that these mutants were resistant to *bli-1* silencing (Fig. 6-2A). We next performed a genetic cross to determine if the *rage* mutants were recessive or dominant. To do so, we crossed wild-type males with these *rage* mutants and fed F1 progeny *bli-1* RNAi. F1 progeny will have one functional copy of *rage.* If progeny develops blisters, then it behaves like wild type and therefore the *rage* mutation is recessive. However, if progeny shows little to no blisters, then the *rage* mutation could be

dominant or semi-dominant. Unfortunately, some of the mutants mated poorly due to sterility defects ('X' in Fig. 6-2 A). For those that mated successfully, we observed that 3/5 mutants (G, I, J, possibly four: K) belonging to Class IV are recessive (Fig.6-2A and B). 3/4 mutants (M, N, O) of Class V were also recessive. One mutant (F) in Class III was dominant or semi-dominant.

Next, we performed a complementation test by crossing each mutant with one another (Fig.6-3). Mutants that fail to complement suggest that they are alleles of the same *rage* gene because each parent provides a non-functional allele that will fail to support blister formation (i.e., fail to silence *bli-1*). If mutants complement each other, it suggests that two different *rage* genes are mutated and therefore the presence of one functional allele for each gene will result in blisters (silencing). Although it was difficult to score the number of F1s because of the low rate of mating between mutants, we observed that there are at least four different *rage* genes (Fig.6-3) - *rage-1*: mutant J failed to complement G and M; *rage-2*: mutant O complemented G, I, K, and M; *rage*-3: mutant K complemented J, L, M, O but failed to complement F; and *rage-4*: mutant F is dominant and failed to completement L, M, K, and O. These data were later found to be consistent with results from whole genome sequencing and in silico complementation by other members of the Jose Lab.

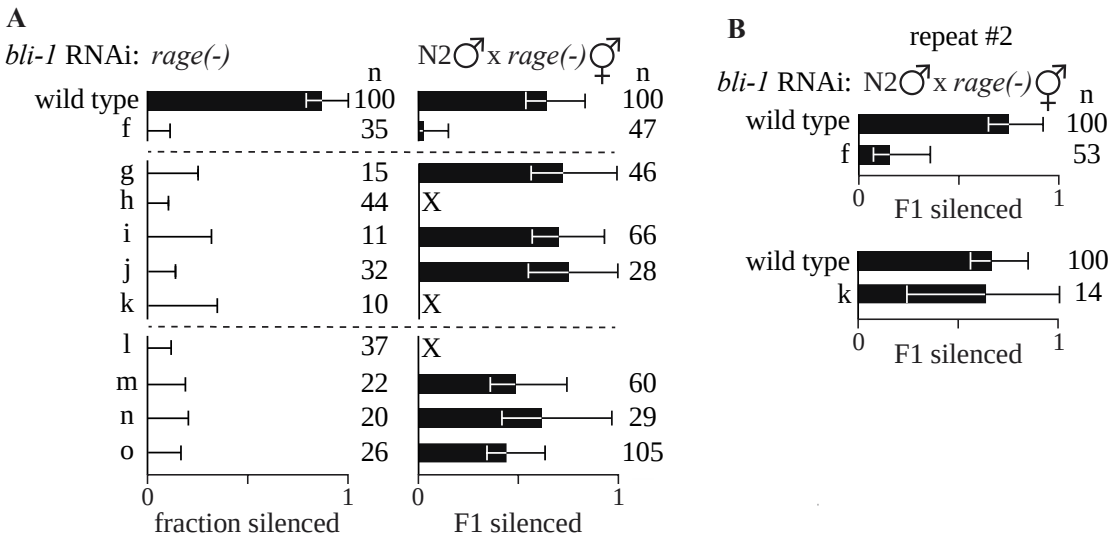

Figure 6-3. Isolates from the *rage* screen with defects in *bli-1* silencing include dominant and recessive mutants.

(**A**) *left*, 10 RAGE mutants (f-o) were fed *bli*-1 RNAi for 24 hrs and scored. Consistent with previous observations by Pravrutha and Farida, all fail to show *bli-1* silencing. *Right*, *bli-1* RNAi feed against F1 cross progeny to determine which mutants are dominant or recessive for *rage(-)*. "X" means N2 males failed to mate with *rage(-)* mutants. Isolate 'f' appears dominant. (**B**) repeat of *bli-1* RNAi feed against F1 cross progeny of some mutants in (A), *right*. Errors bars are 95% confidence intervals.

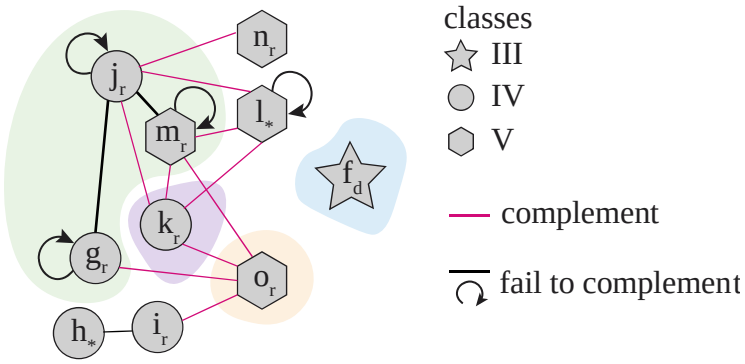

Figure 6 - 4. Complementation testing reveals the presence of multiple *rage* genes.

10 *rage* mutants of different RNAi classes (represented as shapes) were tested. Mutants that failed to complement other mutants by mating are connected by a black line. Mutants that did complement with other mutants are connected by pink lines. From these results, it can be inferred that there are different *rage* genes. Each colored cloud indicates one potential *rage* gene.

Glossary

Terms

Epiallele: an allele with an altered gene expression in the absence of DNA mutations.

Epigenetic change: a change that is heritable without any associated mutations in DNA.

Epigenetic recovery: the process by which a silenced gene becomes re-expressed in the absence of any new mutations.

Homozygous: the state of having two copies of the same allele of a gene.

Heterozygous: the state of having one copy each of two different alleles of a gene.

Transgene: an exogenous gene that has been introduced into the genome.

Transgenerational epigenetic inheritance (TEI): term used in the literature to describe an epigenetic change can persists for more than two generations.

Gene-editing tools

Cas9-mediated genome editing: method to introduce a single-nucleotide polymorphism (SNP), deletion or insertion using an RNA-guided endonuclease Cas9 which recognizes and cleaves the target DNA to create a double-stranded DNA break. This is followed by either homologous s or non-homologous repair.

MosSCI (Mos-1 Single-Copy Insertion): method to integrate a gene of choice using the excision of a transposable element that results in a double-stranded DNA break, followed by homology-directed repair.

Phenomena:

Double-stranded (ds)RNA-induced silencing: Silencing triggered by exposure to dsRNAs. It can result in sequence-specific silencing of the target gene post-

transcriptionally or co-transcriptionally. It requires the following regulators: SID-1, RDE-4, RDE-1, RNA-dependent RNA polymerases (RdRps) (e.g.), MUT-2/RDE-3, HRDE-1/WAGO-9 (for germline genes) or NRDE-3 (for somatic genes) which may lead to chromatin modifications regulated by SET-25 and SET-32 to name a few.

Mating-induced silencing: the process by which a gene is silenced in progeny when parent males expressing the gene is crossed to hermaphrodites lacking the gene, but is not silenced or less often silenced when the reciprocal cross is performed.

Piwi-interacting (pi) RNA-induced silencing: The silencing that is triggered by small RNAs that are 21 nts in length and contain a 5' uracil end (features of piRNAs in *C. elegans*). Requires binding to a PIWI-clade Argonaute called PRG-1 that guides it to a target transcript and result in silencing of the gene. However, such targeting can tolerate a few mismatches within the seed region of the piRNAs (position 2-9 nts from the 5'U end). The following regulators are required for silencing: 21U-RNAs, PRG-1, RdRp, MUT-2/RDE-3, and HRDE-1/WAGO-9.

Spontaneous silencing: the process by which a transgene, newly integrated into the genome via MosSCI, becomes silenced. requiring piRNA recognition and often leading to stable silencing.

***Trans* activation:** the process by which a gene that has been silenced in *trans* recovers expression and results in the re-expression of a silenced gene with homologous sequences.

***Trans* silencing:** the process by which an expressed gene is silenced upon exposure to a silenced gene with homologous sequences.

Bibliography

Adashev, V. E., Kotov, A. A., Bazylev, S. S., Shatskikh, A. S., Aravin, A. A., & Olenina, L. V. (2021). Stellate Genes and the piRNA Pathway in Speciation and Reproductive Isolation of Drosophila melanogaster. Frontiers in genetics, 11, 610665.

Ambros, V. (2001). microRNAs: tiny regulators with great potential. Cell, 107(7), 823-826.

Amir, R., Van den Veyver, I., Wan, M. et al. (1999). Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. Nat Genet 23, 185–188.

Anway, M. D., Cupp, A. S., Uzumcu, M., & Skinner, M. K. (2005). Epigenetic transgenerational actions of endocrine disruptors and male fertility. science, 308(5727), 1466-1469.

Aravin, AA., Naumova NM., Tulin AV., Vagin VV., Rozovsky YM., & Gvozdev, VA. (2001). Double-stranded RNA-mediated silencing of genomic tandem repeats and transposable elements in the *D. melanogaster* germline. Current Biology 11: 1017-1027.

Aravin, AA., Klenov MS., Vagin VV., Bantignies F., Cavalli G., & Gvozdev, VA. (2004). Dissection of a Natural RNA Silencing Process in the Drosophila melanogaster Germ Line. Molecular and Cellular Biology 24(15): 6742-6750

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., Pintacuda, G., Sakaguchi, A., Sarkies, P., Ahmed, S., & Miska, E. A. (2012). piRNAs can trigger a multigenerational epigenetic memory in the germline of C. elegans. Cell, 150(1), 88–99.

Bagijn, M. P., Goldstein, L. D., Sapetschnig, A., Weick, E. M., Bouasker, S., Lehrbach, N. J., Simard M., and Miska, EA. (2012). Function, targets, and evolution of Caenorhabditis elegans piRNAs. Science, 337(6094), 574-578.

Bartel, D. P., & Chen, C. Z. (2004). Micromanagers of gene expression: the potentially widespread influence of metazoan microRNAs. Nature reviews genetics, 5(5), 396-400.

Batista, P. J., Ruby, J. G., Claycomb, J. M., Chiang, R., Fahlgren, N., Kasschau, K. D., ... & Mello, C. C. (2008). PRG-1 and 21U-RNAs interact to form the piRNA complex required for fertility in C. elegans. Molecular cell, 31(1), 67-78.

Bazopoulou, D., & Tavernarakis, N. (2009). The NemaGENETAG initiative: large scale transposon insertion gene-tagging in Caenorhabditis elegans. Genetica, 137, 39-46.

Bernstein, E., Caudy, A. A., Hammond, S. M., & Hannon, G. J. (2001). Role for a bidentate ribonuclease in the initiation step of RNA interference. Nature, 409(6818), 363-366.

Bessereau, J. L., Wright, A., Williams, D. C., Schuske, K., Davis, M. W., & Jorgensen, E. M. (2001). Mobilization of a Drosophila transposon in the Caenorhabditis elegans germ line. Nature, 413(6851), 70-74.

Boyko, A., & Kovalchuk, I. (2011). Genetic and epigenetic effects of plant–pathogen interactions: an evolutionary perspective. Molecular Plant, 4(6), 1014-1023.

Billi AC, Fischer SEJ, Kim JK. Endogenous RNAi pathways in C. elegans. In: WormBook: The Online Review of C. elegans Biology [Internet]. Pasadena (CA): WormBook; 2005-2018. Available from: https://www.ncbi.nlm.nih.gov/books/NBK206718/

Blumenthal T. Trans-splicing and operons in C. elegans. In: WormBook: The Online Review of C. elegans Biology [Internet]. Pasadena (CA): WormBook; 2005-2018. Available from: https://www.ncbi.nlm.nih.gov/books/NBK19704/

Blumenthal T, Davis P, Garrido-Lecca A. Operon and non-operon gene clusters in the C. elegans genome. In: WormBook: The Online Review of C. elegans Biology [Internet]. Pasadena (CA): WormBook; 2005-2018. Available from: https://www.ncbi.nlm.nih.gov/books/NBK293639/

Blumenthal, T., Gleason, K. Caenorhabditis elegans operons: form and function. Nat Rev Genet 4, 110–118 (2003). https://doi.org/10.1038/nrg995

Cantone, I., & Fisher, A. G. (2013). Epigenetic programming and reprogramming during development. Nature structural & molecular biology, 20(3), 282-289.

Cavalli, G., Heard, E. (2019). Advances in epigenetics link genetics to the environment and disease. Nature 571, 489–499.

Charlesworth AG, Seroussi U, Lehrbach NJ, Renaud MS, Sundby AE, Molnar RI, Lao RX, Willis AR, Woock JR, Aber MJ, Diao AJ, Reinke AW, Ruvkun G, Claycomb JM. (2021). Two isoforms of the essential C. elegans Argonaute CSR-1 differentially regulate sperm and oocyte fertility. Nucleic Acids Res. 49(15):8836-8865.

Chaves, Daniel A., Dai, H., Li, L., Moresco J.J., Eun Oh M., Conte Jr. D.,Yates III J.R., Mello C., Gu W. (2021). The RNA phosphatase PIR-1 regulates endogenous small RNA pathways in C. elegans. Molecular Cell 81(3):546-557.

Chey M, Jose AM. (2021). Heritable epigenetic changes at single genes: challenges and opportunities in Caenorhabditis elegans. Trends in Genetics 38(2):116-119. 328(20):726-728.

Conine, C. C., Batista, P. J., Gu, W., Claycomb, J. M., Chaves, D. A., Shirayama, M., & Mello, C. C. (2010). Argonautes ALG-3 and ALG-4 are required for spermatogenesis-specific 26G-RNAs and thermotolerant sperm in Caenorhabditis elegans. Proceedings of the National Academy of Sciences, 107(8), 3588-3593.

Cox, G. N., Staprans, S., & Edgar, R. S. (1981). The cuticle of Caenorhabditis elegans: II. Stage-specific changes in ultrastructure and protein composition during postembryonic development. Developmental biology, 86(2), 456-470.

Devanapally, S.*, Raman, P.*, Chey, M., Allgood, S., Ettefa, F., Diop, M., Lin, Y., Cho, Y.E., Jose, AM. (2021). Mating can initiate stable RNA silencing that overcomes epigenetic recovery. Nature Communications 12:4239. (Online Jan 7, 2020, BioRxiv). *equal contribution.

Dias, B. G., & Ressler, K. J. (2014). Parental olfactory experience influences behavior and neural structure in subsequent generations. Nature neuroscience, 17(1), 89-96.

Duverger, Y., Belougne, J., Scaglione, S., Brandli, D., Beclin, C., & Ewbank, J. J. (2007). A semi-automated high-throughput approach to the generation of transposon insertion mutants in the nematode Caenorhabditis elegans. Nucleic acids research, 35(2), e11-e11.

Feng, S., Jacobsen, S. E., & Reik, W. (2010). Epigenetic reprogramming in plant and animal development. Science, 330(6004), 622-627.

Fire, A., Xu, S., Montgomery, M. et al. (1998). Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature 391, 806–811.

Fischer, S. E., Wienholds, E., & Plasterk, R. H. (2003). Continuous exchange of sequence information between dispersed Tc1 transposons in the Caenorhabditis elegans genome. Genetics, 164(1), 127–134.

Frøkjær-Jensen, C., Davis, M. W., Ailion, M., & Jorgensen, E. M. (2012). Improved Mos1-mediated transgenesis in C. elegans. Nature methods, 9(2), 117-118.

Frøkjær-Jensen, C., Wayne Davis, M., Hopkins, C. E., Newman, B. J., Thummel, J. M., Olesen, S. P., & Jorgensen, E. M. (2008). Single-copy insertion of transgenes in Caenorhabditis elegans. Nature genetics, 40(11), 1375-1383.

Frolows, N., & Ashe, A. (2021). Small RNAs and chromatin in the multigenerational epigenetic landscape of Caenorhabditis elegans. Philosophical Transactions of the Royal Society B, 376(1826), 20200112.

Ghildiyal, M., & Zamore, P. D. (2009). Small silencing RNAs: an expanding universe. Nature reviews genetics, 10(2), 94-108.

Grad, Y., Aach, J., Hayes, G. D., Reinhart, B. J., Church, G. M., Ruvkun, G., & Kim, J. (2003). Computational and experimental identification of C. elegans microRNAs. Molecular cell, 11(5), 1253-1263.

Gu, W., Shirayama, M., Conte, D., Vasale, J., Batista, P. J., Claycomb, J. M., ... & Mello, C. C. (2009). Distinct argonaute-mediated 22G-RNA pathways direct genome surveillance in the C. elegans germline. Molecular cell, 36(2), 231-244.

Guillermo, A. R., Chocian, K., Gavriilidis, G., Vandamme, J., Salcini, A. E., Mellor, J., & Woollard, A. (2021). H3K27 modifiers regulate lifespan in C. elegans in a context-dependent manner. BMC biology, 19(1), 1-17.

Han T., Manoharan AP., Harkins TT. Bouffard P., Fitzpatrick C., Chu DS., Thierry-Mieg D., Thierry-Mieg J., Kim J. (2009). 26G endo-siRNAs regulate spermatogenic and zygotic gene expression in Caenorhabditis elegans. PNAS 106(44):18674-18679.

Hodgkin, J., Horvitz, HR., Brenner,S. (1979). NONDISJUNCTION MUTANTS OF THE NEMATODE *CAENORHABDITIS ELEGANS*, *Genetics*, Volume 91, Issue 1: Pages 67–94.

Hollick, J.B. (2017). Paramutation and related phenomena in diverse species. Nat. Rev.Genet. 18, 5–23.

Izumi, N., & Tomari, Y. (2014). Diversity of the piRNA pathway for nonself silencing: worm-specific piRNA biogenesis factors. Genes & development, 28(7), 665–671.

Jin, C., Li, J., Green, C. D., Yu, X., Tang, X., Han, D., ... & Han, J. D. J. (2011). Histone demethylase UTX-1 regulates C. elegans life span by targeting the insulin/IGF-1 signaling pathway. Cell metabolism, 14(2), 161-172.

Jones, P., Baylin, S. (2002). The fundamental role of epigenetic events in cancer. Nat Rev Genet 3, 415–428.

Jose AM. (2023). Heritable epigenetic changes are constrained by the dynamics of regulatory architectures. bioRxiv.

Jose AM. (2020). A framework for parsing heritable information. Journal of the Royal Society Interface 17:20200154.

Jose AM. (2020). Heritable epigenetic changes alter transgenerational waveforms maintained by cycling stores of information. BioEssays 1900254.

Jose AM. (2018). Replicating and cycling stores of information perpetuate life. BioEssays. 40(4):1700161.

Ketting, R. F., & Cochella, L. (2021). Concepts and functions of small RNA pathways in C. elegans. Current topics in developmental biology, 144, 45-89.

Law, C. W., Chen, Y., Shi, W., Smyth, G. K. (2014). voom: precision weights unlock linear model analysis tools for RNA-seq read counts. Genome Biol. 15, R29.

Lee, H. C., Gu, W., Shirayama, M., Youngman, E., Conte, D., & Mello, C. C. (2012). C. elegans piRNAs mediate the genome-wide surveillance of germline transcripts. Cell, 150(1), 78-87.

Lee, R. C., Feinbaum, R. L., & Ambros, V. (1993). The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. cell, 75(5), 843-854.

Lim, L. P., Lau, N. C., Weinstein, E. G., Abdelhakim, A., Yekta, S., Rhoades, M. W., & Bartel, D. P. (2003). The microRNAs of Caenorhabditis elegans. Genes & development, 17(8), 991-1008.

Marnik, E. A., & Updike, D. L. (2019). Membraneless organelles: P granules in Caenorhabditis elegans. Traffic, 20(6), 373-379.
Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal 17(1), 10-12.

Mattick, J., & Amaral, P. (2023). RNA, the epicenter of genetic information (p. 400). Taylor & Francis

Maures, T. J., Greer, E. L., Hauswirth, A. G., & Brunet, A. (2011). The H3K27 demethylase UTX-1 regulates C. elegans lifespan in a germline-independent, insulin-dependent manner. Aging cell, 10(6), 980-990.

Mazzone, R., Zwergel, C., Artico, M. et al. (2019). The emerging role of epigenetics in human autoimmune disorders. Clin Epigenet 11(34).

Moosavi, A., & Motevalizadeh Ardekani, A. (2016). Role of Epigenetics in Biology and Human Diseases. Iranian biomedical journal, 20(5): 246–258.

Ni, J. Z., Kalinava, N., Chen, E., Huang, A., Trinh, T., & 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(1), 1-15.

Osbourn AE, Field B. Operons. (2009). Cell Mol Life Sci.(23):3755-75.

Patro, R., Duggal, G., Love, MI., Irizarry, RA., Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. Nat. Methods 14, 417-419.

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

Pembrey, M., Bygren, L., Kaati, G. et al. (2006). Sex-specific, male-line transgenerational responses in humans. Eur J Hum Genet 14, 159–166.

Perez, M. F., & Lehner, B. (2019). Intergenerational and transgenerational epigenetic inheritance in animals. Nature cell biology, 21(2), 143-151.

Perez-Borrajero, C., Podvalnaya, N., Holleis, K., Lichtenberger, R., Karaulanov, E., Simon, B., ... & Falk, S. (2021). Structural basis of PETISCO complex assembly during piRNA biogenesis in C. elegans. Genes & Development, 35(17-18), 1304-1323.

Rechavi, O., Houri-Ze'evi, L., Anava, S., Goh, W. S. S., Kerk, S. Y., Hannon, G. J., & Hobert, O. (2014). Starvation-induced transgenerational inheritance of small RNAs in C. elegans. Cell, 158(2), 277-287.

Reinhart, B. J., Slack, F. J., Basson, M., Pasquinelli, A. E., Bettinger, J. C., Rougvie, A. E., ... & Ruvkun, G. (2000). The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. nature, 403(6772), 901-906.

Robinson, M.D., Oshlack, A. (2010). A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* 11(3), R25.

Ruby, J. G., Jan, C., Player, C., Axtell, M. J., Lee, W., Nusbaum, C., ... & Bartel, D. P. (2006). Large-scale sequencing reveals 21U-RNAs and additional microRNAs and endogenous siRNAs in C. elegans. Cell, 127(6), 1193-1207.

Sarker, G., Berrens, R., von Arx, J., Pelczar, P., Reik, W., Wolfrum, C., & Peleg-Raibstein, D. (2018). Transgenerational transmission of hedonic behaviors and metabolic phenotypes induced by maternal overnutrition. Transl. Psychiatry 8, 195.

Schwarz, D. S., Hutvágner, G., Haley, B., & Zamore, P. D. (2002). Evidence that siRNAs function as guides, not primers, in the Drosophila and human RNAi pathways. Molecular cell, 10(3), 537-548.

Shaye DD, Greenwald I. (2011). OrthoList: a compendium of C. elegans genes with human orthologs. PLoS One. 6(5):e20085.

Shichino Y, Yamashita A and Yamamoto M. (2014). Meiotic long non-coding meiRNA accumulates as a dot at its genetic locus facilitated by Mmi1 and plays as a decoy to lure Mmi1. *Open Biol* 4:140022.

Shirayama M, Seth M, Lee HC, Gu W, Ishidate T, Conte D Jr, Mello CC. (2012). piRNAs initiate an epigenetic memory of nonself RNA in the C. elegans germline. Cell. Jul 6;150(1):65-77.

Shugarts N, Sathya A, Yi AL, Chan WM, Marré JA, Jose AM. (2023). SID-1 regulates a retrotransposon-encoded gene to tune heritable RNA silencing. bioRxiv.

Shukla, A., Perales, R., & Kennedy, S. (2021). piRNAs coordinate poly (UG) tailing to prevent aberrant and perpetual gene silencing. Current Biology, 31(20), 4473-4485.

Shukla, A., Yan, J., Pagano, D. J., Dodson, A. E., Fei, Y., Gorham, J., ... & Kennedy, S. (2020). poly (UG)-tailed RNAs in genome protection and epigenetic inheritance. Nature, 582(7811), 283-288.

Sijen, Titia et al. (2001). On the Role of RNA Amplification in dsRNA-Triggered Gene Silencing. Cell 107(4): 465 – 476.

Slack, F. J., Basson, M., Liu, Z., Ambros, V., Horvitz, H. R., & Ruvkun, G. (2000). The lin-41 RBCC gene acts in the C. elegans heterochronic pathway between the let-7 regulatory RNA and the LIN-29 transcription factor. Molecular cell, 5(4), 659-669.

Strome, S. (2005) Specification of the germ line. WormBook, ed. The C. elegans Research Community, 10.1895/wormbook.1.9.1, www.wormbook.org.

Suen, K.M., Braukmann, F., Butler, R. et al. (2020). DEPS-1 is required for piRNA-dependent silencing and PIWI condensate organisation in Caenorhabditis elegans. Nat Commun 11, 4242.

Tang, W. W., Dietmann, S., Irie, N., Leitch, H. G., Floros, V. I., Bradshaw, C. R., ... & Surani, M. A. (2015). A unique gene regulatory network resets the human germline epigenome for development. Cell, 161(6), 1453-1467.

Timmons, L., Court, D. L., & Fire, A. (2001). Ingestion of bacterially expressed dsRNAs can produce specific and potent genetic interference in Caenorhabditis elegans. Gene, 263(1-2), 103-112.

Timmons, L., Fire, A. (1998). Specific interference by ingested dsRNA. Nature 395, 854.

van Wolfswinkel, JC., Claycomb JM., Batista PJ., Mello C., Berezikov E., Ketting R. (2009). CDE-1 Affects Chromosome Segregation through Uridylation of CSR-1-Bound siRNAs. Cell, 139(1): 135 – 148.

Watanabe, Y., & Yamamoto, M. (1994). S. pombe mei2+ encodes an RNA-binding protein essential for premeiotic DNA synthesis and meiosis I, which cooperates with a novel RNA species meiRNA. Cell, 78(3), 487-498.

Xu, F., Feng X., Chen X., Weng C., Yan Q., Xu T., Hong M., Guang S. (2018). A Cytoplasmic Argonaute Protein Promotes the Inheritance of RNAi. Cell Reports, 23(8): 2482 – 2494.

Yamashita A. (2019). meiRNA, A Polyvalent Player in Fission Yeast Meiosis. Non-coding RNA, 5(3): 45.

Yigit, E., Batista, P. J., Bei, Y., Pang, K. M., Chen, C. C. G., Tolia, N. H., ... & Mello, C. C. (2006). Analysis of the C. elegans Argonaute family reveals that distinct Argonautes act sequentially during RNAi. Cell, 127(4), 747-757.

Zhang D, Tu S, Stubna M, Wu WS, Huang WC, Weng Z, Lee HC. (2018). The piRNA targeting rules and the resistance to piRNA silencing in endogenous genes. Science. 359(6375):587-592.