poly(UG)-tailed RNAs and Transgenerational Epigenetic Inheritance in C. elegans

HARVARD UNIVERSITY
Graduate School of Arts and Sciences

DISSERTATION ACCEPTANCE CERTIFICATE

The undersigned, appointed by the Division of

Medical Sciences in the subject of Genetics and

Genomics, have examined a dissertation entitled

presented by Aditi Shukla

candidate for the degree of Doctor of Philosophy and hereby certify that it is worthy of acceptance.

Fred Winston

Signature: _____ Fred Winston (Dec 10, 2020 16:40 EST)

Typed Name: Dr. Fred Winston

Signature: _____ ERIK SONTHEIMER (Dec 10, 2020 16:39 EST)

Typed Name: Dr. Erik Sontheimer

Signature: _____

Typed Name: Dr. Maxwell Heiman

Gary Ruvkun

Signature: _____ Gary Ruvkun (Dec 10, 2020 17:22 EST)

Typed Name: Dr. Gary Ruvkun

Date: December 01, 2020

poly(UG)-tailed RNAs and Transgenerational Epigenetic Inheritance in *C. elegans*

A dissertation presented

by

Aditi Shukla

to

The Division of Medical Sciences

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Genetics and Genomics

Harvard University

Cambridge, Massachusetts

December 2020

Dissertation Advisor: Dr. Scott Kennedy Aditi Shukla

poly(UG)-tailed RNAs and Transgenerational Epigenetic Inheritance in *C. elegans*

Abstract

Small noncoding RNAs (small RNAs) such as Piwi-interacting RNAs (piRNAs), microRNA (miRNAs), and short-interfering RNAs (siRNAs) are important regulators of gene expression. Small RNAs are bound by Argonaute proteins and, together, this complex locates complementary target mRNAs and mediates their silencing. Known roles of siRNAs include protecting genomes from mobile genetic elements called transposons and mediating gene silencing in response to exogenous double-stranded RNAs (dsRNAs). RDE-3 is a ribonucleotidyltransferase required for the silencing of transposon RNAs and dsRNA-targeted mRNAs in *Caenorhabditis elegans (C. elegans)*. In heterologous expression systems, RDE-3 adds long stretches of alternating, non-templated uridine (U) and guanosine (G) ribonucleotides to the 3′ termini of RNAs (designated poly(UG) or pUG tails). In Chapter 2 of my dissertation, we show that, in its natural context in *C. elegans*, RDE-3 adds pUG tails to mRNA targets of exogenous dsRNA, as well as to transposon RNAs. RNA fragments attached to pUG tails with more than 16 perfectly alternating 3′ U and G nucleotides become potent mediators of gene silencing. pUG tails promote gene silencing by recruiting RNA-dependent RNA polymerases, which use pUG-tailed RNAs (pUG RNAs) as templates to synthesize gene-silencing siRNAs.

In *C. elegans*, siRNAs can promote gene silencing for several generations. This multigenerational silencing is a robust example of transgenerational epigenetic inheritance (TEI). Chapter 2 of my dissertation shows that cycles of pUG RNA-templated siRNA synthesis and siRNA-directed pUG RNA biogenesis (pUG RNA/siRNA cycling) underlie siRNA-directed TEI in

the *C. elegans* germline. How *C. elegans* ensures accuracy and fidelity to this feed-forward, heritable and, therefore, inherently dangerous gene regulatory system is not well understood. Chapter 3 of my dissertation demonstrates that *C. elegans* piRNAs are one system that prevents the wrong germline-expressed mRNAs from becoming permanently inactivated by pUG RNA/siRNA cycling. In the absence of piRNAs, a subset of germline-expressed genes that are not normally subjected to TEI enter a state of permanent silencing. Entry into the silenced state is irreversible, and genes thus silenced are paramutagenic. Taken together, this work reports a new 3' RNA modification that turns RNA fragments into potent mediators of gene silencing within and across generations and identifies one mechanism that *C. elegans* uses to ensure that this potentially dangerous modification is added to the correct RNAs.

Table of Contents

List of Figures and Tables

Chapter 1

Chapter 2

Chapter 3

Acknowledgments

A PhD is never a solo endeavor and I have many people to thank for supporting me through this journey. I would first like to thank my mentor, Scott Kennedy. Scott has always pushed me to be a better scientist, a better communicator and a better writer. One of the most important skills he has taught me during my time in his lab is how to think about the big picture and I know this skill will come in handy for the remainder of my scientific journey. Scott's pursuit of novel science is relentless, his imagination is boundless, and his willingness to play the odds is fearless. I especially benefited from his willingness to take a chance when he allowed me to put my first project on pause because I wanted to look for poly(UG)-tailed RNAs in *C. elegans*. Thank you, Scott, for putting up with me in my quest to always do the extra control and to find just the right words in our papers.

I am eternally grateful to have had the opportunity to collaborate with Marv Wickens and his lab. Their generosity and willingness to share with us unpublished data about the activity of RDE-3 set my research journey on an exciting new course. Marv's love and enthusiasm for science has been a constant inspiration over these past few years.

The members of my Dissertation Advisory Committee (DAC)—Danesh Moazed, Eric Greer and Fred Winston—have provided me with helpful guidance and advice throughout my graduate career and it has been an honor to present my work to these great scientists. I am also grateful to Fred, Max Heiman, Gary Ruvkun and Erik Sontheimer for being scientific inspirations and agreeing to serve on my defense examination committee and to evaluate my dissertation.

I would especially like to thank Fred, who has been a mentor to me since I was a first-year student and has been a part of every milestone in my graduate career. Fred always makes time to listen and has offered invaluable advice throughout the years. I will never forget his ability to captivate a room full of first-year students as he taught us about yeast genetics. In my time at HMS, I have met few people as calm, caring, honest, and funny as Fred. He has given an incredible amount to the Genetics community as an investigator, mentor, and teacher, something

that I will always carry with me and strive to emulate—even if I cannot condone his basketball allegiances.

Getting to this point in my graduate career would not have been possible without the amazing people who run the BBS program. Thank you to Susan Dymecki, Tom Bernhardt, Sam Reck-Peterson, Davie Van Vactor, Monica Colaiacovo, Kate Hodgins, Anne O'Shea and Danny Gonzalez for keeping us all on track and always being available when we need you.

The Genetics Department has been a wonderful place to do science for the past few years. I have learned a lot from my colleagues and have been inspired by their amazing work. I would specifically like to thank Sarah Elizabeth Merry, Vonda Shannon, Jene Etter, Lana Woods and Sabrina Shrestha for always having answers to all of my questions and for keeping our department running smoothly. In addition, I would like to thank Fred, Scott, Rachel Wright, Kale Hartmann and Emily Gleason for all of the hard work they have put into the Program in Genetics and Genomics (PGG). HMS and the BBS program, itself, are quite large and can sometimes feel a little overwhelming, but PGG has helped to provide a sense of community for students, like me, who have shared scientific interests.

I was extremely fortunate to have received a NSF Graduate Research Fellowship during graduate school and I would like to thank the NSF for their funding.

When experiments are not going well, it is helpful to be surrounded by people who motivate and inspire you to keep trying. I have to thank all of the past and present members of the Kennedy lab (Brandon Fields, Dan Pagano, Lakshya Bajaj, Anne Dodson, George Spracklin, Diveena Becker, Roberto Perales, Jenny Yan, Cindy Chang, Gang Wan, David Lowe, Bin Bao, Xiaoyun Yang and Yuhan Fei) for making the lab a fun place to do science. I don't think I've ever laughed more than at some of our lab lunches. Roberto Perales knew how to bring us all together for a good time over beers and generously provided the mutant strains that I characterized for a good chunk of my thesis. Jenny's work has been instrumental to helping us understand the function of poly(UG)-tailed RNAs in *C. elegans.* Cindy and Laksh have been particularly helpful

over the past few months as I try to wrap up my work in the lab. I have been fortunate that several of my relationships with my past labmates have extended to friendships outside of the lab. Brandon Fields and George Spracklin first welcomed me into the Kennedy lab when I was a second-year student and have remained great friends. I have cherished our endless conversations about science, pop culture, current events and television shows. Diveena and Karl Becker have provided great laughs and fun adventures outside of lab. Their annual Thanksgiving dinners have been particularly meaningful to those of us who couldn't make it home for the holiday. Finally, I would like to thank Anne Dodson. Anne has been a friend, mentor and inspiration for the past four years. Her attention to detail, thoroughness with her work and patience are all things I admire and strive towards as I grow as a scientist.

I have made some great friends on the second floor of the NRB. The members of the Winston lab have been friendly faces whom I could turn to for questions about science, but also just to chat. Their collective ability to convey science in a clear and engaging manner has inspired me to do the same. In particular, I would like to thank Katie Weiner and James Warner, for all of their helpful advice and shared enthusiasm for cooking, baking and drinking. Past and present members of the Wu lab, particularly Eric Joyce, Son Nguyen, Huy Nguyen, Guy Nir, Jelena Erceg and Talia Hatkevich, have been scientific role models, but have also made being at work more fun with their kindness and humor. Jon Seidman and Josh Gorham generously provided their time and resources to help us prepare libraries for RNA-sequencing. I am so fortunate to have met Heather Lorien de Rivera, Joe Zullo and Oliver Zullo, who have been a great support system and have provided many fun times outside of lab.

My BBS classmates have been a source of inspiration throughout the years. I am constantly blown away by their intelligence and creativity. In particular, I would like to thank Heather Drexler Landry, Patty Rohs, Emily Low, Kelsey Tyssowski, Colin Waters, Kelly Biette, Raja Gopalakrishnan and Brad Wierbowski, members of the G_{n+1} journal club that we formed

early on in graduate school to read interesting papers together and provide feedback on each other's talks and manuscripts.

Raja Gopalakrishnan has been a great friend since the start of graduate school. His optimism and enthusiasm make it always enjoyable to be around him, even when we are rooting for different tennis players.

I must thank Taylor Swift and Billie Eilish for making catchy music that got me through figure-making for both of my manuscripts and this dissertation.

I did not know what scientific research was until I was an undergraduate at Barnard College. It was at Barnard that I learned how to think like a scientist and where I discovered my passion for Molecular Biology. I was fortunate to have worked with and learned from a number of professors in the Biology Department, including Paul Hertz, John Glendinning, Elizabeth Bauer and Brian Morton. But most of all, I am grateful for the opportunity to have worked with Jennifer Mansfield. Jen was my mentor for my senior thesis research project. Her kindness, patience and encouragement were invaluable to me as a young scientist. Jen introduced me to the world of RNA and it was with her that I did my first ever RNA extraction, a skill that I have subsequently used countless times during graduate school. My project in Jen's lab was to try to get RNAi to work in the tobacco hornworm, *Manduca sexta,* and I guess fell in love with RNAi so much that I never stopped working on it.

There is absolutely no way I would have gotten to where I am today if my parents did not constantly push me to work harder and do better, both with their words and their actions. My parents left their friends and family in India to come to the United States and give me the opportunity for a better life. I realize now, more than ever, how fortunate I was that I grew up in a household where pursuing education to the highest degree was never questioned. I know that I can never repay my parents' sacrifice, but I hope that I can at least make them proud.

I have also been fortunate to have gained a second family during graduate school. The Wierbowskis have welcomed me into their lives with open arms and have provided constant support and encouragement over the past few years.

Last, but not least, I have to thank my husband, Brad Wierbowski, who has been my companion since day one of graduate school. Brad's excitement for science is infectious, his patience is admirable and his endless support means to the world to me. I could not imagine going on this wild journey with anyone else and am looking forward to our next adventure together.

My defense is something I looked forward to ever since I started graduate school. I never could have imagined that a worldwide pandemic would mean that I would be giving my defense virtually, rather than in room full of my friends, family and colleagues. However, I am thankful that my friends and family have so far been mostly unharmed during a time when many people's lives are being uprooted. I am thankful that science keeps going and has provided some stability during a time when things may seem a little unstable around us. I am thankful to everyone who has to make difficult and, often, unpopular decisions during this unprecedented time, but has done so to keep us all safe and healthy. I may not be defending under circumstances that I could have imagined, but the fact that I am defending at all is an honor and a privilege.

Chapter 1

Introduction

I. RNA-based gene regulation

Two of the founding tenets of modern molecular biology are: (1) DNA is the hereditary material; and (2) the Central Dogma, which states that there is sequential transfer of information from DNA to RNA to protein. Although first described in the 1950s, these two inaugural pillars of molecular biology still stand today, but have been built upon as a result of newer discoveries. Many of the new additions to these pillars concern the multifaceted roles of RNA in gene regulation and inheritance. Indeed, the discovery of noncoding RNAs has revealed that RNA is not simply an intermediary between DNA and proteins, but instead, RNA can regulate if, which and how much DNA (genes) actually gets expressed as protein. Further, while DNA is still known to be the primary mode of inheritance, in some cases, RNA molecules can also mediate the transfer of information between generations. For the past several years, my dissertation work has focused on expanding what we know about RNA-based gene regulatory mechanisms. I have studied how RNA molecules, as well as modifications added to RNA molecules, can affect gene expression and the mechanisms by which RNAs can mediate the transfer of gene expression states across generational boundaries. For this work, I have employed (or been employed by) the metazoan model organism *Caenorhabditis elegans* (*C. elegans*). These nematodes have a number of characteristics that have been invaluable to my research, including their genetic tractability, large brood sizes and short generation times.

This chapter of my dissertation provides a general overview, followed by a *C. elegans*-specific discussion, of some of what is known about the role of RNA molecules and RNA modifications in regulating gene expression and how gene expression states established by RNA-based gene regulatory mechanisms can sometimes be inherited. In Chapter 2 of my dissertation, I describe the discovery of a new RNA modification that gets added to mRNA fragments in *C. elegans* and discuss our work characterizing the function of this novel modification in mediating gene silencing within and across generations. In Chapter 3, I describe how the interaction

between noncoding RNA pathways can help to maintain proper gene expression states, by showing that, in the absence of one of these pathways, misregulation of gene expression can occur and be maintained across generations.

A. Overview of gene regulation

DNA is the immortal messenger that carries the blueprints for life across generations. However, organisms must pick and choose the DNA information that gets used at any given time in order to mediate processes such as growth and development, cell type specification and diversification, and response to environmental stresses. The mechanisms by which cells can induce or repress the expression of specific genes are collectively known as gene regulation.

The regulatory mechanisms that modulate gene expression are highly diverse. Indeed, gene expression can be regulated at many levels, including chromatin accessibility, transcription, mRNA processing, mRNA stability and translation. For instance, the packaging of DNA into chromatin in the nucleus can influence which genes are accessible to the transcription machinery. The most basic unit of chromatin is the nucleosome, which consists of 147bp of DNA wrapped around an octamer of the four core histone proteins (H3, H4, H2A and H2B) (Kouzarides 2007). A variety of modifications can be post-translationally added to the N-terminal tails of histone proteins, including acetylation, methylation (mono-, di-, and tri-methylation), phosphorylation, ADP-ribosylation, and ubiquitination (Stillman 2018). Histone modifications affect gene expression by influencing how tightly DNA is packaged (and, therefore, how accessible it is to the transcription machinery) and by recruiting proteins, such as chromatin remodelers, histone chaperones, DNA/histone-modifying enzymes and general transcription factors, that bind to specific modifications. For instance, methylation of lysine 9 on histone H3 (H3K9 methylation) is known to create a repressive chromatin state at pericentromeres (Kouzarides 2007; Suganuma and Workman 2011; Hyun *et al.* 2017).

DNA, itself, can also be modified and these modifications are associated with different gene expression states. Perhaps the best studied DNA modification is the methylation of the fifth carbon of cytosines (5-methylcytosine (5mC)), found at 5' cytidine-phosphate-guanosine-3' (CpG) dinucleotides. This covalent DNA modification is found in bacteria and some, but not all, eukaryotes (Zemach and Zilberman 2010; Raddatz *et al.* 2013; Greenberg and Bourc'his 2019). 5mC plays a particularly important role in gene silencing during mammalian fertilization and development, during X chromosome inactivation, in genomic imprinting (Smith and Meissner 2013; Greenberg and Bourc'his 2019) and in transposon silencing (Walsh *et al.* 1998).

Once a gene is transcribed, the resulting mRNA can be regulated post-transcriptionally by mechanisms that affect mRNA export, stability and decay. RNA modifications and noncoding RNAs have emerged as two important mechanisms that can accomplish this post-transcriptional gene regulation (Moazed 2009; Cech and Steitz 2014; Holoch and Moazed 2015a). As my dissertation work involved the study of noncoding RNAs and a novel RNA modification in *C. elegans,* I will focus the rest of this introduction on describing how RNA-based gene regulatory mechanisms can affect gene expression. Of note, gene regulatory mechanisms do not exist in isolation. In fact, as will be discussed below, noncoding RNAs may be best known for their cytoplasmic roles in mediating mRNA degradation and repressing translation, but these RNA molecules can help to coordinate different levels of gene regulation. For instance, noncoding RNAs can induce DNA methylation and histone modifications in the nucleus.

B. RNA modifications

Epitranscriptomics is the study of RNA modifications and the factors that interact with these modifications. As is the case for DNA and histone modifications, RNA modifications have writers (proteins that modify RNA), readers (proteins that recognize an RNA modification) and erasers (proteins that remove RNA modifications). Although ribosomal RNAs (rRNAs) (Sloan *et al.* 2017) and transfer RNAs (tRNAs) (Pan 2018) can be extensively modified, I will focus here

mostly on examples of modifications that are added to messenger RNAs (mRNAs). mRNA modifications include methylation of ribonucleotides (ex. N^6-methyladenosine or m^6A), acetylation of ribonucleotides (ex. N^4-acetylcytidine or ac^4C), addition of a 5' cap to mRNAs, as well as the addition of non-templated ribonucleotides to the 3' end of an RNA molecule (ex. polyadenylation or uridylation). Below I discuss some of these mRNA modifications in more detail.

1. N^6-methyladenosine

The N^6-methyladenosine (m^6A) modification in mRNAs was discovered in the 1970s (Perry and Kelley 1974) and since its discovery, it is now known to be the most abundant internal modification added to mRNAs (Zhao *et al.* 2017). Indeed, m^6A has been found to occur in organisms from all three domains of life (bacteria, archaea, eukaryotes) (Zhao *et al.* 2017). The m^6A modification is deposited by a multiprotein methyltransferase complex that contains METTL3 (methyltransferase-like 3), a protein that was shown to have methyltransferase activity and to bind S-adenosyl methionine (Bokar *et al.* 1994, 1997). Subsequent work identified the other components of the m^6A methyltransferase complex (METTL14, Wilms tumour 1-associated protein or WTAP and KIAA1429), as well as the m^6A readers (YT521-B homology or YTH proteins, HNRNPs) and erasers (fat mass and obesity-associated protein or FTO and alkB homologue 5 or ALKBH5) (Zhao *et al.* 2017).

mRNAs containing the m^6A modification have been identified by using an anti-m^6A antibody to immunoprecipitate m^6A-containing RNAs and then subjecting immunoprecipitated RNAs to high-throughput sequencing. These studies discovered more than 10,000 m^6A sites in mRNAs derived from ~7,000 genes. This work also revealed that m^6A sites: (1) are enriched in 3' UTRs, near stop codons and in long exons; (2) tend to be same between different tissues; and (3) exist as one site in most mRNAs, but some mRNAs can have up to 20 m^6A sites (Dominissini *et al.* 2012; Meyer *et al.* 2012).

The m^6A modification seems to have a variety of functions. For instance, m^6A can induce mRNA instability (Sommer *et al.* 1978) that is at least partially mediated by the binding of YTH domain-containing family protein 2 or YTHDF2 to the methyl mark (Wang *et al.* 2014a). m^6A is also involved in the posttranscriptional regulation of genes that control mammalian circadian rhythm. The so-called clock genes are modified with m^6A and in the absence of METTL3, two of these genes are retained in the nucleus and the circadian period is extended. m^6A is also involved in other steps of mRNA regulation, including splicing, nuclear export, translation and mRNA decay (Zhao *et al.* 2017). For instance, m^6A appears to promote the translation of mRNAs by recruiting the YTHDF1 reader protein, which then recruits translation initiation factor complex eukaryotic initiation factor 3 (Wang *et al.* 2015). In addition, m^6A seems to be enriched on mRNAs related to development and cell fate specification (Dominissini *et al.* 2012; Meyer *et al.* 2012; Geula *et al.* 2015). Indeed, m^6A is required for both the proper differentiation of mouse embryonic stem (ES) cells, possibly by promoting the degradation of transcripts required for pluripotency, and for the reprogramming of mouse epiblast cells to mouse ES cells (Geula *et al.* 2015). Finally, m^6A appears to be involved in stress responses, including the *Saccharomyces cerevisiae (S. cerevisiae)* response to nitrogen stress (Schwartz *et al.* 2013). Taken together, m^6A appears to have multifaceted functions related to the transcriptional and posttranscriptional regulation of mRNA.

2. 3' RNA tailing

RNA tailing is the addition of non-templated ribonucleotides to the 3' end of an RNA molecule. These tails are added by ribonucleotidyltransferase enzymes, which belong to the polymerase beta-like nucleotidyltransferase superfamily that includes DNA polymerase beta, as well as RNA modifying enzymes like poly(A) polymerases, poly(U) polymerases (also known as terminal uridylyl transferases), 2'-5'-oligoadenylate synthetases and CCA-adding enzymes (Martin and Keller 2007).

i. Polyadenylation

Perhaps the best known example of mRNA tailing is the post-transcriptional addition of non-templated adenosines (A) to the 3' ends of most mRNA molecules in the nucleus. Of note, one class of mRNAs that is not polyadenylated is the replication-dependent histone mRNAs. In eukaryotes, the poly(A) tail promotes the expression of mRNAs in a variety of ways, including by: (1) facilitating their export from the nucleus; (2) protecting their 3' ends from degradation; and (3) promoting their translation by recruiting cytoplasmic poly(A) binding proteins, which interact with the translation initiation factor eIF4G. mRNA polyadenylation occurs in two main steps that involve many proteins. In the first step, polyadenylation factors bind to sequence motifs within the mRNA, including to the polyadenylation signal (AAUAAA or AUUAAA). These polyadenylation factors then trigger the cleavage of the mRNA, usually 15-30 nucleotides downstream of the polyadenylation signal. Poly(A) polymerase then polymerizes the poly(A) tail, which can vary by length across organisms (Dreyfus and Régnier 2002; Proudfoot 2011; Lutz and Moreira 2011; Charlesworth *et al.* 2013; Norbury 2013; Yu and Kim 2020). Interestingly, polyadenylation can also occur in the cytoplasm. This cytoplasmic polyadenylation, which was first discovered in *Xenopus laevis* oocytes (Weill *et al.* 2012), normally extends short poly(A) tails on mRNAs that are being stored in the cytoplasm and is thought to allow for rapid changes in translation-competent mRNA populations and, therefore, protein levels, without requiring transcription of new mRNAs. Indeed, cytoplasmic polyadenylation has been implicated in promoting the mitotic to meiotic switch in the *C. elegans* germline and in the remodeling of synapses required for learning (Dreyfus and Régnier 2002; Proudfoot 2011; Lutz and Moreira 2011; Charlesworth *et al.* 2013; Norbury 2013).

ii. Uridylation

Adenosines are not the only non-templated ribonucleotides that can be added to the 3' ends of mRNA molecules. Indeed, mRNAs (and small RNAs) can also be modified with non-

templated uridines (poly(U) tails) via the action of poly(U) polymerases or PUPs, which are also known as terminal uridyl transferases or TUTs (Norbury 2013; Yu and Kim 2020). The Wickens lab first identified PUPs from four different species (*Arabidopsis thaliana, Schizosaccharomyces pombe, C. elegans* and humans) by expressing candidate ribonucleotidyltransferase (rNT) enzymes in *Xenopus* oocytes, tethering these rNTs to reporter mRNAs and then asking how the rNTs modified their reporter mRNAs (Kwak and Wickens 2007). Of note, below I will discuss a similar assay that the Wickens lab used to identify a poly(UG) polymerase in *C. elegans* (Preston *et al.* 2019). In the same year, another group also showed that the Caffeine-induced death suppressor 1 (Cid1), the PUP that the Wickens lab identified from fission yeast, can monouridylate the 3' ends of polyadenylated mRNAs, which promotes their decapping and leads to their subsequent degradation (Rissland *et al.* 2007). mRNAs derived from the replication-dependent histone genes, which are not polyadenylated, are also subject to 3' uridylation, which promotes their rapid degradation at the end of S phase (Mullen and Marzluff 2008). Uridylation also helps to promote genome stability. For instance, retrotransposon mRNAs, such as those derived from long interspersed element 1 (LINE-1), can be uridylated and this uridylation prevents their retrotransposition (Warkocki *et al.* 2018). Taken together, uridylation of mRNAs promotes mRNA decay and can be useful during processes that require rapid turnover of mRNA, including apoptosis, gametogenesis and after DNA replication. It should be noted that small noncoding RNAs, such as microRNAs and Piwi-interacting RNAs, can also be monouridylated or oligouridylated. For instance, the precursor *let-7* microRNA is oligouridylated in mammals and nematodes. This modification promotes its 3'→5' degradation during early larval stages to regulate developmental timing in the worm (Heo *et al.* 2009; Hagan *et al.* 2009; Lehrbach *et al.* 2009). In contrast, at least one report finds that the monouridylation of precursor *let-7* microRNAs might actually promote their processing and maturation (Heo *et al.* 2012). Additional work will show how

the number of Us added to RNAs might be read and interpreted differently, leading to different fates of these uridylated RNAs.

iii. CCA-addition to tRNAs

Finally, I will mention one more 3' RNA modification in this section, namely the post-transcriptional, non-templated addition of the trinucleotide cytidine-cytidine-adenosine (CCA) to the 3' ends of mature tRNAs. This discussion of a heteropolymeric 3' RNA modification is relevant to the work I will present in Chapters 2 and 3 concerning the identification of a new heteropolymeric 3' RNA modification in *C. elegans.* Prokaryotic and eukaryotic mature tRNAs are modified with a 3' CCA, which serves as the amino acid attachment site on the tRNA molecule and may protect mature tRNAs from endoribonucleolytic cleavage (Dutta *et al.* 2013). This 3' modification is particularly intriguing in that it, unlike polyadenylation and uridylation discussed above, involves the addition of two different ribonucleotides in an ordered arrangement. In most bacteria, archaea and eukaryotes, a single CCA-adding enzyme is able to add both two C residues and one A residue to the 3' end of mature tRNAs. Such an enzyme has a binding pocket that can accommodate either a CTP or an ATP. This activity is quite interesting in that the enzyme adds a very specific number of two different ribonucleotides in a very specific order. At least some of these CCA-adding enzymes accomplish this activity by employing phosphate groups found in the 3' end of the tRNA to interact with the nucleotide to be incorporated (Xiong and Steitz 2004; Vörtler and Mörl 2010). Interestingly, some thermophilic bacteria actually use two enzymes to modify tRNAs, one enzyme that adds two C residues and a second enzyme that then adds the A residue (Bralley *et al.* 2005; Neuenfeldt *et al.* 2008).

C. Noncoding RNAs

Noncoding RNAs, which can range in length from small to long, have emerged as major regulators of gene expression. These RNA molecules are not fated to be translated into proteins.

Instead, they influence the expression of protein-coding mRNAs at the transcriptional, post-transcriptional and even translational levels. Two breakthrough discoveries made in *C. elegans* broke open the floodgates for the study of noncoding RNAs. The first was the discovery of microRNAs. In 1993, the *lin-4* microRNA was shown to regulate the expression of the *lin-14* gene by base pairing with regions of partial complementarity in its 3' UTR (Lee *et al.* 1993; Wightman *et al.* 1993). The discovery of the second microRNA in *C. elegans*, *let-7* (Reinhart *et al.* 2000), and the demonstration that the *let-7* microRNA was not just a worm-specific RNA but, instead, was conserved in humans and other animals with bilateral symmetry (Pasquinelli *et al.* 2000), made microRNAs, and other noncoding RNAs, a topic of general interest. The second discovery that inspired the study of noncoding RNAs was the demonstration that double-stranded (dsRNAs) could induce the silencing of mRNAs in *C. elegans* (Fire et al. 1998). Both of these findings will be discussed in more detail below. Since these two seminal discoveries, the advent of high-throughput sequencing technologies has allowed the scientific community to really appreciate just how little of the transcriptome is composed of protein-coding mRNAs and how regions of the genome previously thought to transcriptionally silent actually give rise to long and small noncoding RNA molecules that regulate gene expression without coding for proteins. Here I will first provide a brief overview of long noncoding RNAs. I will then dive deeper into the functions of small noncoding RNAs, as most of my dissertation will be about these short, but powerful, RNA molecules.

1. Long noncoding RNAs

Historically, the study of the human genome has focused on the 20,000 protein-coding genes that were thought to be dispersed throughout junk DNA. However, high-throughput transcriptomics has shown us that while less than two percent of the human genome encodes functional proteins, seventy five percent of the genome is capable of being transcribed (Djebali *et al.* 2012). Long noncoding RNAs (lncRNAs) are one type of non-protein-coding transcript encoded

by eukaryotic genomes. lncRNAs are 200 or more nucleotide long, predominantly nuclear RNA molecules that are capped, spliced and polyadenylated, but are not translated into proteins (Ulitsky and Bartel 2013; Cech and Steitz 2014). lncRNAs are derived from RNA polymerase II-mediated intergenic transcription, transcription of enhancer DNA elements and antisense transcription that overlaps with protein-coding genes (Kopp and Mendell 2018).

There is much debate surrounding lncRNAs and their functions. Indeed, some of this debate stems from the difficulty in identifying lncRNAs and differentiating them from protein-coding mRNAs. One reason lncRNAs are hard to identify is because they often harbor the typical hallmarks of a protein-coding mRNAs, including a poly(A) tail and a 5' cap (Ulitsky and Bartel 2013). In addition, while lncRNAs have been identified in a diversity of organisms, including yeast, nematodes, zebrafish, mice and humans, there is often little primary sequence conservation between lncRNAs in different organisms (Kopp and Mendell 2018). Sometimes, even different human cell lines can also show varying lncRNA expression profiles (Djebali *et al.* 2012). Furthermore, even though lncRNAs do not encode function proteins, about half of all lncRNAs in mouse embryonic stem cells can associate with ribosomes (Ingolia *et al.* 2011). Indeed, some lncRNAs have been miscategorized and actually found to encode small proteins (Anderson *et al.* 2015). All of these difficulties have resulted in a very large range of potential lncRNA molecules in humans, from 20,000 (Harrow *et al.* 2012) to 100,000 (Zhao *et al.* 2016).

Despite the difficulties in identifying lncRNAs, the longstanding view that they represent transcriptional noise or junk RNA is continually challenged by the discovery of lncRNAs with important functions. How many lncRNAs will actually prove to have important functions is unknown, but will likely require analysis of specific lncRNAs on a case-by-case basis(Ulitsky and Bartel 2013; Cech and Steitz 2014). Below I will review just a few examples of lncRNAs that have been identified and shown to regulate gene expression through diverse mechanisms.

lncRNAs can broadly be characterized as those the act *in cis* by influencing the expression of nearby genes and those that act *in trans* by influencing gene expression anywhere in a cell.

One way in which lncRNAs can regulate gene expression *in cis* is by repressing the function of protein-coding genes that are in close proximity to the lncRNA gene by impeding the transcription of the protein-coding gene, a process known as transcription interference. For instance, *SRG1* is a lncRNA transcribed from intergenic DNA in the *S. cerevisiae* genome. The *SRG1* locus overlaps with the promoter of a downstream gene called *SER3.* The transcription of *SRG1* increases nucleosome density at the *SER3* promoter, a chromatin state that is unfavorable for transcription factor binding to the *SER3* promoter. As a result, transcription of *SRG1* results in the repression of *SER3* gene expression (Martens *et al.* 2004, 2005; Thebault *et al.* 2011).

Sometimes lncRNAs can act *in cis* in a manner that requires the lncRNA product and does not simply rely on transcription interference. The 17-kb long *Xist* (X-inactive-specific transcript) lncRNA, which is required for mammalian dosage compensation, is one well-known lncRNA that acts in this manner. During embryonic development in female mammals, *Xist* is transcribed from only one of the two X chromosomes. The *Xist* RNA then acts *in cis* by coating the X chromosome from which it was transcribed and mediating the transcriptional repression of this chromosome by directly or indirecting recruiting chromatin modifying proteins, such as PRC2 (Polycomb repressive complex 2), a complex that can deposit the repressive methylation on lysine 27 of histone H3 (Brockdorff *et al.* 1992; Brown *et al.* 1992; Penny *et al.* 1996). Interestingly, the *Xist* lncRNA is transcribed from a <500-kb region on the inactive X chromosome which contains at least seven genes that give rise to lncRNAs, several of which have been shown to be involved in X-chromosome inactivation (Lee 2009).

The two examples of lncRNA function discussed thus far showed how lncRNAs can act *in cis* to regulate nearby genes. However, lncRNAs can also act *in trans*. The highly conserved HOX transcription factors regulate positional identity and differentiation during development. These transcription factors are derived from four gene clusters (*HOXA, HOXB, HOXC* and *HOXD)* that are located on different chromosomes. *HOTAIR* (HOX antisense intergenic RNA), a 2.2-kb long lncRNA, is transcribed from the *HOXC* gene cluster and can bind the distantly located *HOXD*

cluster where it is thought to act as a scaffold to recruit chromatin-modifying complexes, such as PRC2, to repress the expression of the *HOXD* gene cluster (Rinn *et al.* 2007; Tsai *et al.* 2010; Chu *et al.* 2011). Of note, the ability of lncRNAs, such as Xist and *HOTAIR*, to recruit PRC2 is controversial and likely involves a complex cascade of events featuring several different proteins and RNAs (Holoch and Moazed 2015a).

The above discussion is by no means an exhaustive list of the all of the ways in which lncRNAs can act to regulate gene expression. It is clear from these examples, however, that lncRNAs can modulate gene expression in a variety of ways that only sometimes depend on the production of a functional lncRNA molecule. As a result, careful attention must be paid to how lncRNAs are studied. For instance, DNA regulatory elements, such as promoters, for neighboring genes can sometimes be embedded with a lncRNA locus, as is the case for the *SRG1* lncRNA discussed above. Thus deletion of the entire lncRNA locus can confound the interpretation of whether it is a lncRNA locus or the lncRNA molecule, itself, that influences gene expression. The advent of new technologies has made it easier to study lncRNAs in a genome-wide, but also targeted manner. For instance, CRISPR/Cas9-mediated genome-editing has made it possible to now screen for lncRNAs that are involved in a particular phenotype of interest and also to perform more targeted modifications to lncRNA loci to decipher their true functions (Zhu *et al.* 2016; John Liu *et al.* 2017; Joung *et al.* 2017).

2. Small noncoding RNAs

Small noncoding RNAs (small RNAs) are 20-35nt long RNA molecules that have emerged as key regulators of gene expression in eukaryotes, and even prokaryotes. Small RNAs can promote mRNA degradation and translational repression, as well as direct epigenetic DNA and chromatin modifications, which influence transcription. Small RNAs are bound by Argonaute proteins and, together, these ribonucleoprotein complexes target complementary mRNAs for silencing via RNA interference (RNAi)-related pathways (Filipowicz 2005; Holoch and Moazed

2015a). A number of different classes of small RNAs are produced in eukaryotic cells, including Piwi-interacting RNAs (piRNAs), microRNAs (miRNAs) and small interfering RNAs (siRNAs). In general, siRNAs and miRNAs are derived from longer dsRNA precursor molecules, whereas piRNAs are derived from single-stranded RNA precursors. Before I begin discussing these small RNAs in detail, I will first describe the discovery of RNAi in *C. elegans*, as this discovery helped to pave the way for the study of small RNAs and their functions. I will then introduce transposons as transposon silencing is a conserved function of small RNAs. Next, I will discuss Argonaute proteins, as these proteins serve as the essential binding partners that help to mediate the small RNA functions that I will describe. Finally, I will discuss the functions of siRNAs, miRNAs and piRNAs.

i. Discovery of RNAi

The field of genetics relies on the ability to disrupt gene expression to study gene function. Thus, tools that allow for the disruption of gene expression are invaluable to any geneticist. In the 1980s, several studies pointed to a role for antisense RNAs in transcriptional and translation repression in prokaryotes (Inouye 1988). As a result, antisense RNA-mediated gene disruption was attempted and shown to work (with varying efficiencies) in several prokaryotic and eukaryotic systems, including bacteria, flies, frogs, plants and mammalian cells (Takayama and Lnouye 1990). In 1991, antisense RNA was shown to disrupt expression of the *unc-22* and *unc-54* genes in *C. elegans* (Fire *et al.* 1991). It was proposed that the gene silencing effect of antisense RNA resulted from its ability to hybridize to a target sense mRNA, thereby blocking its transport or preventing the sense mRNA from being translated into protein (Takayama and Lnouye 1990; Fire *et al.* 1991). Interestingly, it was noted that sense *unc-22* RNA could also disrupt *unc-22* expression, but sense *unc-54* RNA did not have the same effect on *unc-54* gene expression. Indeed, soon after, it was shown that injection of both sense and antisense *par-1* RNA into the *C.*

elegans germline phenocopied *par-1* loss-of-function mutations, indicating that both sense and antisense RNA can induce knockdown of *par-1* gene expression (Guo and Kemphues 1995).

Andy Fire, Craig Mello and colleagues set out to explain the somewhat puzzling observations about the silencing effects of sense and antisense RNAs in *C. elegans* (Fire *et al.* 1998). They noted that previous studies that reported the silencing effects of sense and antisense RNAs had made use of high copy DNA transgene arrays or *in vitro* transcribed RNAs made with bacteriophage RNA polymerases, two methods that could result in contaminating RNA transcripts of the opposite strandedness, thus resulting in some double-stranded RNAs that might actually be the active gene interfering RNA molecules (Fire *et al.* 1991, 1998). To test this idea, they injected worms with purified sense and antisense mRNAs, as well as a mixture of the two. Indeed, the sense-antisense mixture was at least two orders of magnitude more effective at inducing silencing of the *unc-22* gene than either strand alone. Injection of dsRNA targeting the *unc-54, hlh-1* and *fem-1* genes also phenocopied loss-of-function mutations in these genes. Further, it was shown that dsRNA corresponding to only coding sequences could induce gene silencing. In addition, dsRNA-induced silencing corresponded with a decrease in endogenous mRNA levels, as shown by *in situ* hybridization (Fire *et al.* 1998).

Fire, Mello and colleagues made two other interesting observations about the silencing effects of injected dsRNA. First, the effects of the dsRNA appeared to be systemic. For instance, *unc-22* dsRNA injected in the gonad could result in silencing of somatic tissues, such as muscles. Second, the silencing effects of dsRNA could be observed in the progeny of injected animals. This was particularly surprising because the lowest dose of dsRNA that was effective at inducing silencing would quickly be diluted in the inheriting progeny to only a few molecules per cell. This result indicated that perhaps some sort of catalytic or amplification mechanism was involved in mediating gene silencing in response to dsRNA (Fire *et al.* 1998). I will discuss what this amplification mechanism is in greater detail below and report the discovery of a new noncoding RNA that helps to mediate this amplification in Chapter 2.

The discovery that dsRNA is a potent inducer of gene silencing in *C. elegans* (Fire *et al.* 1998) was quickly followed by a flurry of reports showing that dsRNAs can induce gene silencing in fungi, plants and animals (Mello and Conte 2004). Further, the endogenous functions of this dsRNA-induced gene silencing response started to become more clear. For instance, RNAi seemed to be related to two phenomena that had previously been described extensively in plants: post-transcriptional gene silencing and cosuppression. Previous work had shown that the introduction of foreign DNA (transgenes) into plants induced the silencing of the transgenes, as well as any endogenous genes with homologous sequences (Jorgensen *et al.* 1996; Que and Jorgensen 1998). This silencing was shown, in at least some cases, to be mediated by post-transcriptional silencing of transcripts (Ingelbrecht *et al.* 1994). Similarly, it was shown that plants infected with viral vectors showed viral resistance and also silenced endogenous genes that were homologous to viral sequences. Since viruses that could not replicate were defective in eliciting this silencing response, the trigger for this response was believed to be viral RNA (Angell and Baulcombe 1997). Indeed, the same year that dsRNA-induced gene silencing was reported in *C. elegans* (Fire *et al.* 1998), the simultaneous expression of sense and antisense RNA was shown to be sufficient to induce viral and transgene silencing in plants (Waterhouse *et al.* 1998).

Thus, although the term RNAi was initially used to describe the gene silencing response elicited by experimentally-provided dsRNAs, today the term encompasses a broad range of RNA-based gene silencing mechanisms that I will discuss more below. The commonality of these RNAi mechanisms are their two main molecular players: small RNAs and Argonaute proteins. As discussed below, these RNAi mechanisms have been implicated in a variety of functions, including transgene silencing, transposon silencing, silencing of repetitive sequences and even endogenous gene regulation.

ii. Argonaute proteins

Argonaute proteins are highly conserved small RNA-binding proteins that are found in nearly all eukaryotes (*S. cerevisiae* being a notable exception (Drinnenberg *et al.* 2009)) and even in some bacteria and archaea (Meister 2013). Eukaryotic Argonaute proteins can be divided into three clades: (1) AGO clade, which bind miRNA and siRNAs; (2) PIWI clade, which find piRNAs; and (3) worm-specific Argonautes or WAGOs. Argonaute proteins bound to their small RNA partners are often referred to as RNA-induced silencing complexes (RISC). Argonaute proteins have three domains, the N-terminal PAZ domain, the middle MID domain and the C-terminal PIWI domain, each with a distinct function in binding small RNAs (Meister 2013). The PAZ domain binds the 3' end of a small RNA (Jinek and Doudna 2009), while the MID domain binds the 5' end (Simon *et al.* 2011). The PIWI domain shares homology with RNAse H enzymes, allowing some Argonaute proteins to act as endonucleases that can cleave a target mRNA that has perfect complementarity to the small RNA to which they are bound (Höck and Meister 2008). Endonucleolytic activity of the PIWI domain is mediated by a catalytic triad of amino acids (aspartic acid - aspartic acid - histidine or aspartic acid). However, not all Argonautes with this conserved catalytic triad show cleavage activity (Höck and Meister 2008; Meister 2013).

iii. Transposons

Transposable elements or transposons are mobile genetic elements found in the genomes of eukaryotes, prokaryotes and even archaea. Active transposons can be detrimental to genome integrity as they can disrupt protein-coding genes, induce chromosomal breaks and cause genome rearrangements (Malone and Hannon 2009). Transposons were first described in maize by Barbara McClintock in 1951 (McClintock 1950). Earlier in her career, McClintock invented techniques to visualize the maize genome and show that maize has 10 chromosomes. She then used these techniques to describe how a locus she called *Dissociation* (*Ds)* could jump out of the maize chromosome 9. This jumping was dependent on a second locus called *Activator (Ac)*, which could, itself, jump out of the genome. *Ac* was later shown to be a transposon that encoded a

transposase protein, which could mediate the excision of *Ac* out of the genome (Fedoroff *et al.* 1983). *Ds* was shown to be an inactive transposon, which could be excised out of the genome by the transposase encoded by *Ac* (Fedoroff *et al.* 1983). At the time, the genome of an organism was thought to be a stable entity and McClintock's observations of mutable loci or unstable genes were not given the acknowledgment they deserved. However, the existence of transposons became more generally accepted in the 1970s when these mobile elements were discovered in bacteria and viruses.

Today, transposons are characterized into two main classes: retrotransposons (class I) and DNA transposons (class II), which vary by the mechanism they use to mobilize. The prevalence of each transposon class can vary greatly between organisms. For instance, about 45% of the human genome is composed of transposons and >90% of these are retrotransposons. In contrast, about 12% of the *C. elegans* genome is made up of transposons, and >80% of these are DNA transposons. Retrotransposons mobilize using a "copy and paste" mechanism. They first encode an RNA intermediate that then gets reverse transcribed into cDNA, which then gets integrated at another site in the genome. Retrotransposons are further divided into long terminal repeat (LTR) retrotransposons and non-LTR retrotransposons. LTR retrotransposons are thought to be derived from retroviruses. They encode an integrase protein that mediates the insertion of the retrotransposon cDNA elsewhere in the genome. non-LTR retrotransposons, which include long interspersed nucleotide elements (LINEs) and short interspersed nucleotide elements (SINEs) subclasses, couple integration to reverse transcription by using an endonuclease to perform target-primed reverse transcription at the insertion locus. SINEs are non-autonomous and depend on LINEs for their mobilization. In contrast, DNA transposons employ a "cut and paste" mechanism to mobilize, such that there is no duplication of the transposon. DNA transposons encode a transposase protein that recognizes the terminal inverted repeats (TIRs) that flank all DNA transposon sequences. The transposase protein then mediates the excision and reintegration of the DNA transposon at another location in the genome. Of note, there are

also DNA transposons that no longer encode transposase proteins. They can mobilize, however, through the action of a transposase protein encoded by a related DNA transposon. Subclasses of DNA transposons can further be divided into superfamilies that share a common genetic organization and a shared ancestry (Kazazian 2004; Malone and Hannon 2009; Bourque *et al.* 2018).

Given the potential harm that transposons can cause to gene function and genome stability, organisms have adapted ways of silencing these mobile genetic elements. These mechanisms of transposon silencing include small RNAs, chromatin remodeling, chromatin modifications, DNA methylation and Krüppel-associated box (KRAB) zinc-finger proteins (ZFPs) (Molaro and Malik 2016; Deniz *et al.* 2019). I will discuss the role of small RNAs/RNAi in transposon silencing below. It should be noted, however, that transposons can sometimes be beneficial. Transposons can promote the evolution of new gene regulatory pathways. After all, transposons contain regulatory sequences, such as promoters, that are required for their own expression. Further, transposons are recognized and targeted by a variety of host cell silencing machinery. Thus, the presence of transposons near host protein-coding genes can influence the expression of these genes, and can be coordinated to set up multi-gene regulatory networks (Rebollo *et al.* 2012; Bourque *et al.* 2018).

iv. siRNAs

Here I will briefly describe the discovery of siRNAs and their general functions, including their role in transcriptional gene silencing. Below, siRNAs will be discussed in additional detail when I introduce the RNAi pathways in *C. elegans*.

a. Discovery and Function

The discovery that dsRNA is a potent inducer of gene silencing in *C. elegans* (Fire *et al.* 1998) was quickly followed by the discovery of siRNAs (Hamilton and Baulcombe 1999;

Hammond *et al.* 2000; Zamore *et al.* 2000; Bernstein *et al.* 2001). These small RNAs range between 21-25nt in length and are the effector molecules that mediate the gene silencing effects of dsRNAs. Several studies that followed the first report of RNAi described the gene silencing effects of siRNAs. siRNAs were first described in plants, where 25nt long sense and antisense RNAs were detected during both transgene-induced post-transcriptional gene silencing and virus-induced post-transcriptional gene silencing (Hamilton and Baulcombe 1999). In *C. elegans,* dsRNA-derived siRNAs were shown to mediate transposon silencing (Sijen and Plasterk 2003). dsRNA-induced gene silencing was recapitulated *in vitro* using *Drosophila* embryo lysates and dsRNA was shown to be processed into 21-23nt long small RNAs in this system. The mRNA targeted by dsRNA in this system was shown to be cleaved at intervals that matched the length of the small RNAs, indicating that siRNAs are the effector molecules of dsRNA-induced silencing (Zamore *et al.* 2000). Transfection of 21nt siRNA duplexes into mammalian cells was shown to knock down the expression of endogenous genes (Elbashir *et al.* 2001). The detection of siRNAs made more sense at a molecular level once it was discovered that the conserved ribonuclease III (RNase III) endonuclease Dicer could cleave dsRNAs into siRNAs (Bernstein *et al.* 2001).

A few years later, the role of siRNAs expanded beyond silencing of foreign sequences, such as transposons and transgenes, to include endogenous gene regulation. For instance, in *Drosophila,* the Argonaute 2 (Ago2) protein was shown to bind 20-22nt long siRNAs. While the majority of these siRNAs corresponded to transposons and other repetitive elements in the fly genome (Kawamura *et al.* 2008), some of these siRNAs were derived from convergent transcription of endogenous genes or endogenous genes containing sequences that could form hairpin elements (Okamura *et al.* 2008a; b). Further, small RNA sequencing in *C. elegans* showed that siRNAs not only targeted transposons, pseudogenes and cryptic loci, but also about 50% of the protein-coding genes in the *C. elegans* genome. Related roles for siRNAs in endogenous gene regulation were also described in plants and mice (Malone and Hannon 2009). Thus, over time, the role of siRNAs has evolved from molecules that mediate silencing in response to

experimental dsRNAs to small RNAs that can silence foreign and parasitic nucleic acids, as well as regulate endogenous gene expression. Below, I will describe the multi-faceted role of siRNAs in *C. elegans* in more detail. Here, I will discuss how some siRNAs can modulate gene expression via transcriptional gene silencing (Holoch and Moazed 2015a).

b. siRNAs can mediate transcriptional gene silencing

siRNA-mediated transcriptional gene silencing has been observed in *Schizosaccharomyces pombe (S. pombe), C. elegans* and *Arabidopsis thaliana (A. thaliana)* (Holoch and Moazed 2015a). Here I will describe how this pathway works in *S. pombe* and later I will describe an analogous pathway in *C. elegans*. Of note, all three of these organisms has a special class of enzymes called RNA-dependent RNA polymerases, which are able to use RNA molecules as a template to either synthesize a dsRNA substrate for Dicer to cleave (Holoch and Moazed 2015a), or as is the case sometimes in *C. elegans,* to synthesize individual siRNAs (Pak and Fire 2007; Sijen *et al.* 2007; Aoki *et al.* 2007).

In *S. pombe,* siRNA-mediated transcriptional gene silencing mediates silencing of pericentromeric DNA repeat regions, as well as the deposition of repressive H3K9 methylation at these sites (Volpe *et al.* 2002). Pericentromeric repeat regions are transcribed by RNA polymerase II and become templates for the RdRP-containing complex RDRC. The activity of the *S. pombe* RdRP, Rdp1, generates long dsRNAs, which are processed by Dicer into siRNAs (Holoch and Moazed 2015a; Allshire and Ekwall 2015). These double-stranded siRNAs are then loaded onto the Argonaute small interfering RNA chaperone (ARC) complex, which includes the sole *S. pombe* Argonaute, Ago1. Two members of the ARC complex, Arb1 and Arb2, help to mediate the destruction of one of the two strands of the siRNA duplex. Ago1 and its bound single-stranded siRNA then switch binding partners and become part of the RNA-induced transcriptional silencing (RITS) complex (Buker *et al.* 2007) and the RITS complex is able to locate nascent transcripts derived from pericentromeric repeats via complementarity to the Ago1-bound siRNA.

The adapter protein, Stc1 (Bayne *et al.* 2010), then helps the RITS complex recruit the *S. pombe* H3K9 methyltransferase Clr4, which then deposits H3K9 methylation on nucleosomes at pericentromeric repeat loci. A positive-feedback loop is created as the HP1 protein Swi6 binds to methylated H3K9 and recruits RDRC, thereby starting the cycle over again. This positive-feedback loop is also aided by Chp1, a member of the RITS complex that harbors a chromodomain and, therefore, helps to anchor the RITS complex onto nucleosomes that contain H3K9 methylation. In plants, such as *A. thaliana,* an analogous transcriptional gene silencing mechanism exists whereby 24nt long siRNAs are able to mediate the deposition of DNA methylation at loci such as foreign transgenes and transposons. This pathway involves two plant-specific RNA polymerases (Pol IV and Pol V) and the RdRP RDR2 (Holoch and Moazed 2015a).

v. miRNAs

a. Discovery

miRNAs are genomically-encoded small RNAs that drive the post-transcriptional silencing of mRNAs in plants and animals. This discussion will focus on biogenesis and functions of animal miRNAs. miRNAs were discovered in *C. elegans* by the Ambros and Ruvkun labs, who were studying the heterochronic genes, which constitute a molecular clock that regulates cell differentiation at each larval stage of development in the worm. Mutations in heterochronic genes can lead to premature or delayed development (Chalfie *et al.* 1981; Ambros and Horvitz 1984, 1987; Ambros 1989). *lin-14* is a heterochronic gene that codes for the LIN-14 protein, which is expressed in embryos and first stage larval animals, but whose levels decline in second stage larvae (Ruvkun and Giusto 1989). *lin-14* loss-of-function mutants skip the first larval stage and precociously exhibit developmental programs of later larval stages. The *lin-4* heterochronic gene was known to be a negative regulator of the *lin-14* gene (Ambros and Horvitz 1987; Ambros 1989). In 1993, the Ambros lab cloned the *lin-4* gene from *C. elegans* and three other *Caenorhabditis* species and found that, even though the three other *lin-4* genes could rescue *lin-4* loss-of-function

in *C. elegans,* the four *lin-4* gene did not encoded a conserved protein sequence that began or ended with conventional start and stop codons. Together, this data suggested that the *lin-4* gene did not code for a protein. Instead, *lin-4* was found to code for two RNAs that were 22 and 61 nucleotides (nt) long (Lee *et al.* 1993). At the same time, the Ruvkun lab showed that the *lin-4* RNAs were complementary to multiple sites in the 3'-UTR of the *lin-14* gene. In addition, the *lin-4* RNAs were shown to trigger translational repression of *lin-14* (Wightman *et al.* 1993). Taken together, the two labs proposed that the *lin-4* RNA could regulate the expression of the *lin-14* mRNA during the progression from the first to second larval stage by binding to the *lin-14* 3'-UTR and blocking ability of the *lin-14* mRNA to be translated.

Today we know that the *lin-4* was the first ever miRNA to be discovered. However, at the time, the Ambros and Ruvkun lab's findings went somewhat ignored as they were thought to be a nematode-specific observation. However, seven years later the Ruvkun lab reported the discovery of another 22nt long RNA called *let-7* (Reinhart et al. 2000). This *let-7* RNA was found in humans, flies and eleven other bilaterally symmetrical animals (Pasquinelli *et al.* 2000). Subsequent studies have identified miRNAs across a diversity of organisms. In addition, unlike lncRNAs, miRNAs tend to be conserved, sometimes only between closely related species and sometimes, through evolution (Bartel 2004).

b. Biogenesis

Here I will describe the canonical miRNA biogenesis pathway that gives rise to the majority of miRNAs in animals. The majority of microRNAs are transcribed by RNA polymerase II from individual transcriptional units, while a small number are found in the introns or 3' UTRs of genes. Transcription of a miRNA gene produces a imperfectly-paired hairpin-containing primary miRNA (pri-miRNA), which is capped, spliced and polyadenylated like a protein-coding mRNA (Lee *et al.* 2002; Cai *et al.* 2004). This pri-miRNA then undergoes sequential processing steps to give rise to a mature miRNA. The first step occurs in the nucleus where the ribonuclease III (RNase III)

enzyme Drosha cleaves the base of the hairpin, removing the sequence flanking the stem-loop/hairpin of the pri-miRNA and producing a precursor miRNA (pre-miRNA) that gets exported out of the nucleus (Denli *et al.* 2004; Han *et al.* 2004). In the cytoplasm, Dicer cleaves off the terminal loop of the hairpin, resulting in a duplex mature miRNA (Bernstein *et al.* 2001; Grishok *et al.* 2001; Hutvágner *et al.* 2001; Ketting *et al.* 2001; Zhang *et al.* 2004). The two strands that comprise this duplex are referred to as the 5p strand, which arises from the 5′ end of the pre-miRNA hairpin, and the 3p strand, which arises from the 3' end of pre-miRNA hairpin. Ultimately, one strand of this duplex, called the guide strand, is loaded into an Argonaute protein, while the other strand, termed the passenger strand, is degraded. The guide strand tends to have weaker binding at its 5' end and tends to begin with 5' uracil (U) (Meijer *et al.* 2014). Interestingly, microRNA expression (Landgraf *et al.* 2007) and even guide strand choice (Meijer *et al.* 2014) show tissue specificity. Of note, other variations of this biogenesis pathway have been described, including Drosha-independent and Dicer-independent non-canonical pathways (Hammond 2015).

c. Function

An Argonaute (ex. AGO1-4 in humans, ALG-1 and ALG-2 in *C. elegans*) loaded with the guide strand of a miRNA duplex is called the miRNA-induced silencing complex (miRISC). miRISC then locates target mRNAs that are complementary to at least nucleotides 2-8 of the miRNA (also known as the seed region). These miRNA target sites are usually found in the 3' UTR of target mRNAs. Interestingly, one miRNA can interact with multiple target mRNAs and a target mRNA can have multiple miRNA-binding sites (Lim *et al.* 2003; Bartel 2009).

The majority of miRNAs act by inducing mRNA decay and translational repression of target mRNAs. Although miRNAs have also been reported to induce co-transcriptional gene silencing of actively transcribed mRNAs (Cernilogar *et al.* 2011), the role of miRNAs in the nucleus is still an active area of investigation and will not be discussed here. miRISC induces mRNA decay by promoting a cascade of events that lead to the deadenylation and decapping of the target

mRNA. miRISC recruits glycine-tryptophan protein of 182 kDa (GW182) proteins to a target mRNA. These GW182 proteins then recruit two poly(A)-deadenylase complexes, poly(A)-nuclease deadenylation complex subunit 2 (PAN2)-PAN3 and carbon catabolite repressor protein 4 (CCR4)-NOT, which, together, deadenylate the target mRNA. Deadenylation is followed by decapping of the target mRNA, a process mediated by the mRNA-decapping enzyme subunit 1 (DCP1)–DCP2 complex (Rehwinkel 2005; Behm-Ansmant *et al.* 2006; Chen *et al.* 2009; Fabian *et al.* 2011; Braun *et al.* 2011). The deadenylated and decapped RNA then undergoes 5'→3' degradation, carried out by exoribonuclease 1 (XRN1) (Braun *et al.* 2012). The role of miRISC in promoting translation inhibition is more controversial, but may involve inhibition of eukaryotic translation initiation factor binding to a target mRNA (Meijer *et al.* 2013; Fukao *et al.* 2014; Fukaya *et al.* 2014). While most miRNAs in animals do not show complementarity with their target mRNAs beyond the seed region (Grimson *et al.* 2007; Ellwanger *et al.* 2011), sometimes additional complementarity can exist. In humans, if complementarity between the miRNA and target mRNA exists beyond the seed region, then the Argonaute AGO2 is able to cleave the target mRNA directly (Liu *et al.* 2004; Jo *et al.* 2015).

vi. piRNAs

piRNAs are small noncoding RNAs that are best known for their role in mediating transposon silencing in animals (Siomi *et al.* 2011; Weick and Miska 2014; Ozata *et al.* 2019). Here, I will provide a brief overview of piRNA biogenesis and then discuss the function of piRNAs in flies and mice. I will discuss *C. elegans* piRNAs in more detail below.

a. Overview of piRNA biogenesis and function

piRNAs are an animal-specific class of genomically-encoded and germline-expressed small noncoding RNAs, which are bound by the PIWI clade of Argonaute proteins, and are

essential for germ cell function in most animals (Weick and Miska 2014; Ozata *et al.* 2019). These small RNAs were first isolated from mouse testes by four different groups in 2006 (Aravin *et al.* 2006; Grivna *et al.* 2006; Watanabe *et al.* 2006; Girard *et al.* 2006). Historically, piRNAs are best known for their role in transposon silencing (Siomi *et al.* 2011; Weick and Miska 2014; Ozata *et al.* 2019). However, as I will discuss below, transposon silencing is not the only function of piRNAs.

piRNA biogenesis can be divided into two steps: precursor piRNAs are first transcribed by RNA polymerase II in the nucleus and then exported to the cytoplasm where they undergo processing and are loaded onto PIWI proteins. In most animals (with *C. elegans* being a notable exception, see below), piRNAs are transcribed as long transcripts (either on one or on both strands) that are then cleaved to produce piRNAs. Once processed into their mature forms, piRNAs range in length from 21-35 nucleotides long, show a preference for a 5' uracil (U) and are 2'-O-methylated at their 3' ends. piRNA biogenesis is highly complex and remarkably diverse across eukaryotes (Weick and Miska 2014; Rojas-Ríos and Simonelig 2018; Ozata *et al.* 2019). As a result, I will focus this section on the function of piRNAs. Later, I will provide more detail about piRNA biogenesis in *C. elegans,* as this information will be relevant to the work I present in Chapter 3.

Before piRNAs were even isolated and sequenced, their protein binding partners, PIWI proteins, were known to be expressed in animal germlines and to be important for germ cell development. For instance, loss of Piwi, one of three PIWI proteins in *Drosophila,* was shown to impair germ cell regeneration and cause sterility (Lin and Spradling 1997; Cox *et al.* 1998). *C. elegans* hermaphrodites lacking PRG-1, the functional PIWI protein in *C. elegans*, were found to display reduced fertility at normal rearing temperatures and complete sterility at elevated temperatures (Wang and Reinke 2008; Batista *et al.* 2008; Das *et al.* 2008). Similarly, all three mouse PIWI proteins were shown to be essential for male fertility (Deng and Lin 2002; Kuramochi-Miyagawa 2004; Carmell *et al.* 2007). Studies in *Drosophila* hinted that some of these germline

defects may be due to the loss of transposon silencing. For example, mutations in the *Drosophila* PIWI proteins were shown to cause mobilization of the Gypsy family of transposons (Sarot *et al.* 2004). Since then, data has revealed that some, but not all, of the fertility defects observed in the absence of PIWI proteins are due to transposon upregulation.

b. piRNA function in flies

The *Drosophila melanogaster* genome encodes three Piwi proteins: (1) Piwi, which is expressed in germ cell nuclei and in somatic follicle cells of the ovary; (2) Aubergine (Aub), which localizes to the cytoplasm of germ cells; and (3) Ago3, which localizes to germ granules, called nuage, in the cytoplasm of germ cells (Brennecke *et al.* 2007). Transposon silencing is the main function of piRNAs in *D. melanogaster*. *D. melanogster* piRNAs are transcribed as long transcripts from large heterochromatic regions that are clustered in the genome and transcribed from either one or both strands. These piRNA clusters are thought of as transposon graveyards, as they are rich in inactive transposon sequence remnants, thought to serve as an archive of past transposon activity (Gunawardane *et al.* 2007; Brennecke *et al.* 2007). The process by which these piRNAs induce post-transcriptional silencing of transposons is called the Ping-Pong cycle and involves many proteins. I will briefly summarize how ping pong cycling works here. Before Ping-Pong can begin, mature (also known as primary) piRNAs must be generated. First, piRNA clusters are transcribed to generate long precursor piRNA transcripts from piRNA clusters. These precursor transcripts are then exported into the cytoplasm, where they are processed into piRNA intermediates by Zucchini, a protein localized on the mitochondrial outer membrane. How mitochondrial proteins were co-opted to participate in piRNA-mediated gene silencing is a mystery. The 5' ends of piRNA intermediates are bound by Piwi and the 3' ends are trimmed and 2'-O-methylated. Once mature piRNAs are generated, they can be bound by Ago3 and, together, this complex recognizes a transposon RNA and Ago3 cleaves the transposon RNA at the site that corresponds to the nucleotide 10 of the piRNA. This process then generates a new piRNA that is

bound by Aub. The Aub-piRNA complex then induces cleavage of a complementary transposon transcript, thereby generating a piRNA that is identical to the one that was bound by Ago3. Thus, in this way, the Ping-Pong cycle induces cleavage of transposon transcripts to generate sense and antisense piRNAs, thereby linking piRNA amplification to the post-transcriptional silencing of transposons in *Drosophila* (Luo and Lu 2017).

PIWI-bound piRNAs can also promote transcriptional gene silencing of transposons and protein-coding genes in flies, worms and mice by locating complementary nascent transcripts in the nucleus and triggering DNA methylation or histone modification of the genomic loci. Indeed, most euchromatic regions that contain H3K9me3 in *Drosophila* ovarian somatic cells actually correspond to transposon insertion sites. The maintenance of H3K9me3 at these sites depends on the nuclear-localized PIWI Argonaute Piwi (Sienski *et al.* 2012). Indeed, recruitment of Piwi to a specific locus using artificial piRNAs results in H3K9me3 deposition and exclusion of RNA polymerase II from the locus, further supporting that piRNAs can induce transcriptional silencing in flies (Huang *et al.* 2013; Le Thomas *et al.* 2013). Fascinatingly, H3K9 methylation of transposon loci in ovarian somatic cells can actually promote non-canonical transcription from these loci, generating precursor piRNAs that can be processed into mature piRNAs to carry out transposon silencing in the ovaries. In this way, somatic cells take on the burden and risk of expressing transposon loci to support transposon silencing in germ cells (Mohn *et al.* 2014; Le Thomas *et al.* 2014).

Although the majority of *Drosophila* piRNAs map to transposons, some fly piRNAs also have functions beyond transposon silencing. For instance, *Suppressor of Stellate (Su(Ste))* is piRNA-producing locus in flies that silences the multi-copy gene Stellate. In the absence of *Su(Ste)*, Stellate protein forms crystals or aggregates in spermatocytes, leading to sterility (Aravin *et al.* 2001, 2004). Work in flies has also suggested that piRNAs may play a role in promoting the degradation of maternally-deposited mRNAs. *Nanos (nos)* is a maternally deposited mRNA that gets degraded during the maternal-to-zygotic transition via deadenylation by the CCR4-NOT

deadenylation complex. Interestingly, this deadenylation event depends on the presence of a piRNA target site in the 3' UTR of the *nos* mRNA. Current models suggest that Aub and Ago3-bound piRNAs interact with the *nos* 3' UTR and either recruit or stabilize the CCR4-NOT deadenylation complex onto the *nos* transcript. In this case, although the piRNAs that interact with the *nos* 3' UTR were derived from retrotransposon piRNAs, they have been co-opted to regulate endogenous protein-coding mRNAs (Rouget *et al.* 2010). Another study suggests that some Aub-bound piRNAs might actually promote germ cell specification in embryos (before the maternal-to-zygotic transition) by binding to and trapping maternal mRNAs in the germ plasm (Vourekas *et al.* 2016). Thus, while a lot of *D. melanogaster* piRNA research has focused on the role of piRNAs in transposon silencing, novel functions of piRNAs are constantly being discovered. These new findings help to underscore why piRNAs are so essential for fertility.

c. piRNA function in mice

The mouse genome encodes three Piwi proteins (MIWI, MILI and MIWI2), all three of which are required for male fertility (Deng and Lin 2002; Kuramochi-Miyagawa 2004; Carmell *et al.* 2007). These Piwi proteins are expressed during different stages of development and bind different subsets of mouse piRNAs. Unlike *D. melanogaster* piRNAs, mouse piRNAs are not derived from heterochromatic clusters. Instead, they appear to be transcribed from euchromatic regions of the genome that are devoid of transposons and repetitive elements (Aravin *et al.* 2006; Grivna *et al.* 2006; Watanabe *et al.* 2006; Girard *et al.* 2006). piRNAs expressed in male mice can be divided into two types: pre-pachytene and pachytene piRNAs. Pre-pachytene piRNAs, which only represent about 5% of all piRNAs in adult mouse testes, are characterized by transposon sequences in the sense orientation, suggesting that they are derived from transposon transcripts (Ernst *et al.* 2017; Ozata *et al.* 2019). These piRNAs are bound by Mili and Miwi2 and, together, these proteins and their bound piRNAs can induce transcriptional and post-transcriptional silencing of transposons. Pre-pachytene piRNA biogenesis begins with the transcription of piRNA

precursor transcripts, which are then trimmed and processed to generate mature pre-pachytene piRNAs that bind to Mili. A Ping-Pong cycle then occurs and generates an amplified pool of sense and antisense secondary piRNAs that are bound by Mili and Miwi2. This cycling induces the cleavage of complementary mRNAs, including transposon transcripts. Miwi2 bound to piRNAs can translocate to the nucleus and induce DNA methylation of transposons, such as LINEs (Di Giacomo *et al.* 2013; Pezic *et al.* 2014; Itou *et al.* 2015). Loss of either Mili or Miwi2 results in arrest of spermatogenesis at the pachytene stage, thought to be caused by increased transposon mobilization, particularly of LINEs and intracisternal particle A transposons (Shoji *et al.* 2009; Reuter *et al.* 2009).

Pachytene piRNAs, which accumulate in spermatocytes in the pachytene stage of meiosis and associate with Miwi, comprise 95% of piRNAs in the mouse testes. Most pachytene piRNAs are not complementary to transposons (Aravin *et al.* 2006, 2008; Girard *et al.* 2006; Li *et al.* 2013). Instead, they are transcribed from intergenic regions and from the 3' UTRs of protein-coding genes and appear to be divergent across mammals (Robine *et al.* 2009; Vourekas *et al.* 2012). Thus, transposon silencing does not appear to be the only (or main) function of piRNAs in mice. Indeed, loss of Miwi does not cause transposon upregulation, but does result in male sterility, indicating that Miwi has an important function or functions unrelated to transposon silencing (Vasileva *et al.* 2009; Tanaka *et al.* 2011). However, the role of pachytene piRNAs has remained somewhat elusive. One study found that the Miwi protein binds to spermiogenic mRNAs directly, without using piRNAs as a guide. Further, instead of destabilizing these mRNAs, as small RNA/Argonaute complexes often do, Miwi seemed to protect these mRNAs from degradation in translationally repressed mRNA ribonucleoproteins complexes (mRNPs) until later stages in spermiogenesis, when translation of these mRNAs is initiated (Vourekas *et al.* 2012). Consistent with this finding, this study also found that pachytene piRNAs do not show clear complementary to the mRNAs with which Miwi interacts (Vourekas *et al.* 2012). Instead, this study hypothesized that pachytene piRNAs might just be degradation products generated during the degradation of

long noncoding RNAs during spermiogenesis. Another study reported a more active role for pachytene piRNAs during a different stage of spermatogenesis, namely during the elongating spermatid stage. This study reported that Miwi-bound piRNAs interact with mRNAs with imperfect base-pairing. Miwi then recruits CAF1, the catalytic subunit of the CCR-4-CAF-1-NOT deadenylase complex. Deadenylation of these mRNAs then leads to their degradation (Gou *et al.* 2015). This function of piRNAs in promoting mRNA deadenylation is reminiscent of a similar role for piRNAs described in flies and discussed above (Rouget *et al.* 2010). Taken together, the most abundant type of piRNAs in mice*,* pachytene piRNAs, do not seem to promote transposon silencing. What exactly the role of these piRNAs is is debated, but current work suggests the interesting possibilities that pachytene piRNAs may have changing roles during spermatogenesis.

Below, I will discuss the biogenesis and function of *C. elegans* piRNAs. These piRNAs are similar to mouse pachytene piRNAs in that their targets appear to be somewhat mysterious.

II. RNAi pathways in *C. elegans*

For the remainder of this chapter, I will focus on describing the role of RNAi/small RNA pathways in promoting gene silencing within and across generations in *C. elegans*, as this information will be relevant to Chapters 2 and 3 of this dissertation.

A. Exogenous dsRNA-mediated gene silencing in C. elegans

1. Overview of exogenous dsRNA-mediated gene silencing pathway in C. elegans

After dsRNA-induced gene silencing in *C. elegans* was first reported (Fire et al. 1998)*,* a number of labs performed forward genetic screens to identify factors that mediate this silencing in the worm (Collins *et al.* 1987; Ketting *et al.* 1999; Tabara *et al.* 1999; Winston *et al.* 2002, 2007; Vastenhouw *et al.* 2003; Tops *et al.* 2005; Chen *et al.* 2005). Here, I will summarize what is known

about some of the factors that were identified from these screens and, subsequently, characterized for their molecular functions in dsRNA-induced gene silencing.

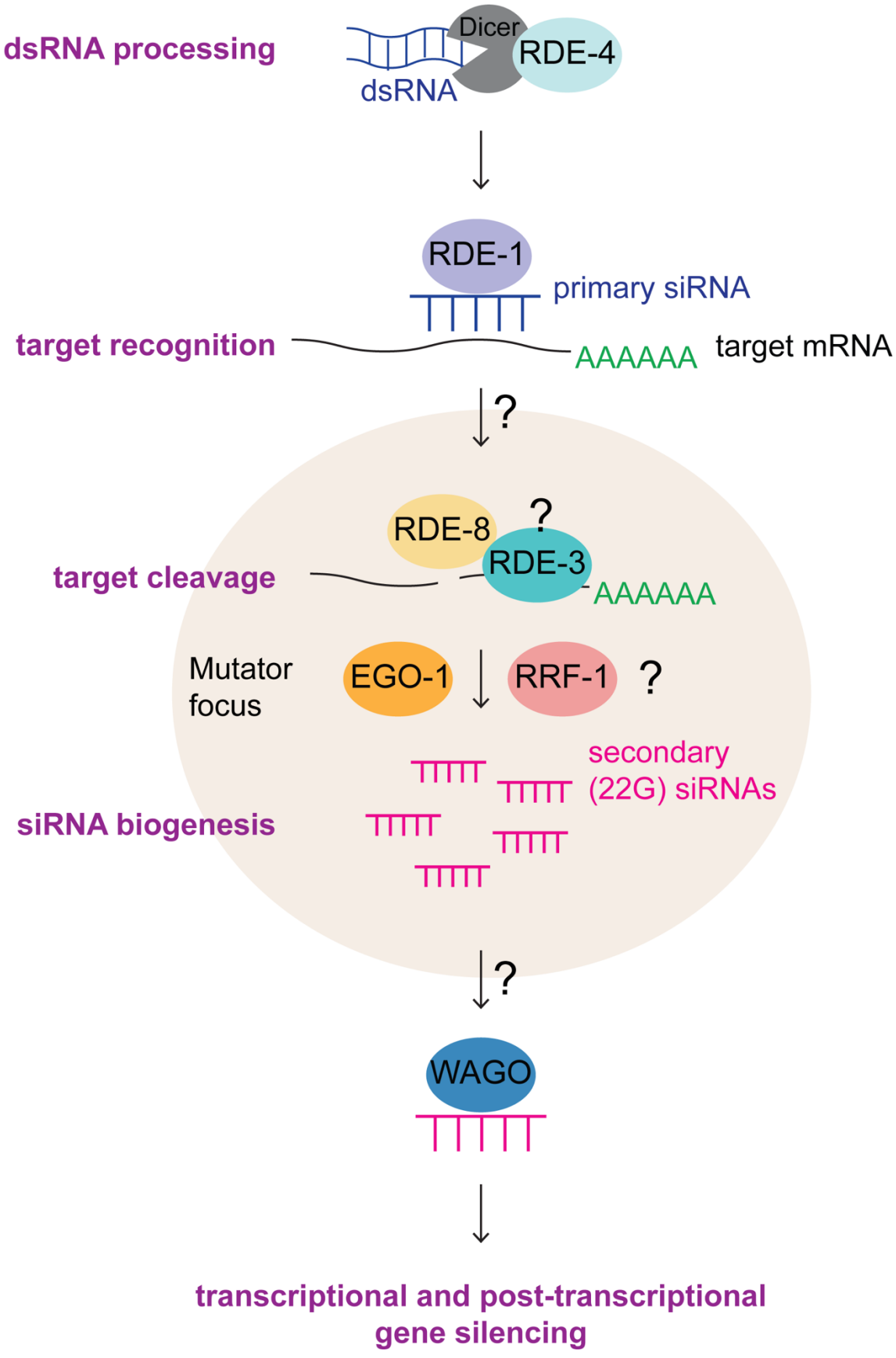
As noted by Andy Fire, Craig Mello and colleagues (Fire *et al.* 1998), dsRNA-induced gene silencing is systemic in *C. elegans*. In other words, when worms are fed dsRNA (i.e. dsRNA enters their gut) or injected with dsRNA in particular anatomical location, the silencing effects of the dsRNA are able to spread to other non-neuronal tissues, including the germline (Jose and Hunter 2007). SID-2 is a transmembrane protein expressed in the apical membrane of intestinal cells and is required for the uptake of dsRNA from the intestinal lumenal space when worms are fed dsRNA (Winston *et al.* 2007; McEwan *et al.* 2012). The transmembrane protein SID-1, which is expressed in all cells, except for neurons, is required for the uptake of dsRNA into the cytosol (Winston *et al.* 2002).

Once dsRNA is taken up into the cytosol, it is processed by Dicer (Figure 1.1) and the double-stranded RNA binding protein RNAi-defective 4 (RDE-4) into short dsRNAs (Parrish and Fire 2001; Tabara *et al.* 2002; Parker *et al.* 2006; Pak and Fire 2007). One strand of this short dsRNA (termed the primary or 1° siRNA) is then bound by the Argonaute RDE-1 (Tabara *et al.* 1999; Parrish and Fire 2001) and together, this siRNA/Argonaute complex recognizes a target mRNA via base complementarity to the siRNA. RDE-1 is phylogenetically related to Argonaute proteins and harbors an evolutionarily conserved DDH catalytic motif that confers endonuclease (Slicer) activity to some Argonaute and PIWI proteins, allowing these proteins to cleave a target mRNA (Meister 2013). However, RDE-1 does not cleave mRNAs and, instead, its catalytic activity is required to remove the passenger strand of the short dsRNA that is loaded into RDE-1 after Dicer processing (Steiner *et al.* 2009). After recognition by an RDE-1/siRNA complex, the target mRNA is then thought to enter *Mutator* foci, biomolecular condensates that localize perinuclear to germ cell nuclei in *C. elegans* (see below), where the endonuclease RDE-8 cleaves the target mRNA, thereby ensuring that it cannot encode a protein (Tsai *et al.* 2015). How exactly the target mRNA relocates from the cytoplasm into *Mutator* foci is not known. In the *C. elegans* germline,

two partially redundant RNA-dependent RNA polymerases or RdRPs (EGO-1 and RRF-1) are then thought to use the target mRNA as a template to produce secondary (2°) siRNAs (Gu *et al.* 2009; Vasale *et al.* 2010; Gent *et al.* 2010; Thivierge *et al.* 2012). These 2° siRNAs are antisense to target mRNAs, 22nt in length, 5' triphosphorylated and begin with a 5' guanosine. It is currently not known how these RdRPs get recruited to target mRNAs and whether they use full-length or cleaved mRNAs as templates for 22G-siRNA biogenesis. Of note, RDE-8 physically associates with the ribonucleotidyltranferase RDE-3 (see below), but the function of RDE-3 is not clear (Tsai *et al.* 2015). A previous study showed that, in animals lacking germlines, as well as the RdRP RRF-1, *sel-1* RNAi causes a small fraction of *sel-1* mRNA fragments to be uridylated in a largely RDE-3–dependent manner in the soma (Tsai *et al.* 2015). This data suggests that, in somatic tissues, RDE-3 may add non-templated Us to the 3' termini of mRNA fragments generated during RNAi. However, the work I present in Chapter 2 will show that we discovered a different mechanism that mediates RdRP recruitment to mRNAs targeted by RNAi in the *C. elegans* germline. RdRP-generated 2° siRNAs engage ≅12 worm-specific Argonautes (WAGOs) and together, this complex can induce transcriptional (see below) and post-transcriptional silencing of target mRNAs. WAGO proteins lack key catalytic residues found in Argonautes that have cleavage activity, suggesting that they mediate gene silencing through a different, currently unknown mechanism (Yigit *et al.* 2006; Guang *et al.* 2008; Gu *et al.* 2009; Shirayama *et al.* 2012). How and where 22G-siRNAs get loaded onto WAGO proteins, most of which do not seem to localize to *Mutator* foci, is currently not known.

Interestingly, although the RNAi pathway described above was initially studied in the context of exogenous dsRNA-mediated gene silencing, it soon became clear that other RNA triggers, such as piRNAs, can initiate related RNAi pathways. As will be discussed below, RdRPs and WAGO proteins are also key components of these other endogenous RNAi pathways in *C. elegans*.

Figure 1.1. dsRNA-mediated gene silencing in *C. elegans*. dsRNA is processed by endoribonuclease Dicer and the double-stranded RNA binding protein RDE-4 into short dsRNAs (Parrish and Fire 2001; Tabara *et al.* 2002; Parker *et al.* 2006; Pak and Fire 2007). One strand (primary siRNA) of this short dsRNA is then bound by the Argonaute RDE-1 (Tabara *et al.* 1999; Parrish and Fire 2001) and together, this siRNA/Argonaute complex recognizes a target mRNA via base complementarity. The target mRNA is then thought to enter *Mutator* foci, biomolecular condensates that localize perinuclear to germ cell nuclei, where the endonuclease RDE-8 cleaves the target mRNA (Tsai *et al.* 2015). How exactly the target mRNA relocates from the cytoplasm into *Mutator* foci is not known. EGO-1 and RRF-1, two partially redundant RNA-dependent RNA polymerases, then use the target mRNA as a template to produce secondary siRNAs that are 22nt in length, 5' triphosphorylated and begin with a 5' guanosine (e.g. 22G-siRNAs). It is currently not known how these RdRPs get recruited to target mRNAs and whether they use full-length or cleaved mRNAs as templates for 22G-siRNA biogenesis. RDE-8 is physically associated with the predicted nucleotidyltransferase RDE-3 (Tsai *et al.* 2015), but the activity of RDE-3 is not clear. 22G-siRNAs engage ≅12 worm-specific Argonautes (WAGOs) and, together, these WAGO/22G-siRNA complexes can induce transcriptional and post-transcriptional silencing of mRNAs. How and where 22G-siRNAs get loaded onto WAGO proteins, most of which do not seem to localize to *Mutator* foci, is currently not known.


dsRNA processing
Dicer
RDE-4
dsRNA
RDE-1
primary siRNA
target recognition
AAAAAA
target mRNA
?
RDE-8
?
RDE-3
target cleavage
AAAAAA
Mutator
focus
EGO-1
RRF-1
?
secondary
(22G) siRNAs
siRNA biogenesis
?
WAGO
transcriptional and post-transcriptional
gene silencing

2. RNAi in transposon silencing

As the *C. elegans* factors, specifically the RDE proteins, that mediate gene silencing in response to exogenous dsRNAs were being discovered, an interesting observation was made about some of them: several of these factors appeared to be required for DNA transposon silencing. The converse observation was also made: some mutants that had been identified a decade or more previously as being defective for transposon silencing were also defective in triggering gene silencing in response to exogenous dsRNAs (Collins *et al.* 1987; Ketting *et al.* 1999; Tabara *et al.* 1999; Vastenhouw *et al.* 2003; Tops *et al.* 2005; Chen *et al.* 2005). Together, these observations indicated that exogenous dsRNA targets and transposons might be silenced by a similar mechanism. Below I will discuss the discovery of DNA transposons in *C. elegans* and describe how the factors that silence these transposons were discovered.

i. Discovery of DNA transposons in C. elegans

DNA transposons were first described in *C. elegans* in 1983. Two different laboratory strains of *C. elegans*, called Bristol and Bergerac, were analyzed by Southern blotting after a restriction digest. While 37/40 restriction fragments were the same between the two strains, 3 were not. One of these fragments was about 1.7kb larger (e.g. 8.7kb vs. 7kb) in Bergerac than in Bristol. Furthermore, electron microscopy of the longer Bergerac fragment suggested that the 1.7kb segment was flanked by inverted repeats, as this 1.7kb fragment formed a loop due to annealing of the inverted repeats (Emmons *et al.* 1983). Interestingly, when the 8.7kb Bergerac fragment was used as a probe for Southern blot, it was found to hybridize to genomic DNA fragments from the Bristol strain to which a probe made from the 7.0kb Bristol fragment did not hybridize, suggesting that the 1.7kb segment was also found in the Bristol genome. Further analysis revealed that the Bristol genome contained about 20 repeats of this 1.7kb fragment, whereas the Bergerac genome contained 200 repeats. In addition, Southern blotting analysis using probes that flanked the 1.7kb insertion showed that, occasionally, the 1.7kb would precisely

excise out of the Bergerac genome but, interestingly, not out of the Bristol genome. Taken together, all of this data suggested that this 1.7kb fragment was a transposon that the authors named Tc1 (Emmons *et al.* 1983). A similar study characterized another restriction digest polymorphism between Bristol and Bergarac near the actin gene cluster and found that this polymorphism was also due to Tc1 (Liao *et al.* 1983).

Soon after, work from Phil Anderson's lab expanded on the observation that the Bergerac strain has more Tc1 transposons and that these transposons appear to be actively hopping. The Anderson lab, while working with Bergerac, noticed that 10 out of 18 spontaneous mutations that they observed in the *unc-54* gene resulted from the insertion of Tc1 into *unc-54* (Eide and Anderson 1985b). In contrast, none of the 65 spontaneous mutations in *unc-54* that they identified in the Bristol strain had a transposon insertion in *unc-54* (Eide and Anderson 1985a), suggesting that transposons are silenced in Bristol. These Tc1 insertions into *unc-54* in the Bergerac strain were also genetically unstable in that the Unc-54 phenotype would revert back to wild-type. Restriction mapping and Southern blotting showed that this reversion was, in fact, due to Tc1 hopping out of the *unc-54* gene, as the revertant animals always showed wild-type restriction patterns (Eide and Anderson 1985b). Of note, the genetic basis for why Tc1 is so active in the Bergerac strain is still not known.

Additional studies revealed that Tc1 is not the only transposon in the *C. elegans* genome. In fact, about 12% of the *C. elegans* genome is now thought to be derived from transposable elements. Most of these transposon-derived sequences are remnants of previously active DNA transposons and can no longer mobilize. The few DNA transposons that have been shown to retain activity under laboratory conditions and, in most cases, only in mutant backgrounds, include Tc1, Tc2, Tc3, Tc4, Tc5 and Tc7 (C. elegans Sequencing Consortium 1998; Bessereau 2006). Why or how retrotransposons have been kept out of the *C. elegans* genome remains a mystery.

ii. Genetic basis of transposon silencing in C. elegans

Forward genetic screens helped to reveal what factors maintain transposon silencing in the *C. elegans* Bristol strain. The earliest such screen was performed by Phil Anderson's lab (Collins *et al.* 1987). Bergerac animals containing a Tc1 insertion in the *unc-54* gene were outcrossed with the Bristol strain and propagated for several generations, thereby creating a hybrid Bristol/Bergerac strain that was presumably homozygous at all loci. This strain, referred to as *unc-54(r232::Tc1),* has a Tc1 excision rate of 1×10^{-5}. Animals with a Tc1 insertion in the *unc-54* gene display an Uncoordinated phenotype. *unc-54(r232::Tc1)* were mutagenized using EMS and mutants were screened clonally for populations that showed a higher frequency of wild-type animals, which indicated that Tc1 excised out of *unc-54*. The eight mutants identified in this screen were referred to as *mutator* mutants. Interestingly, all eight mutants also showed an increase in Tc1 transposition, defined as the insertion of Tc1 at a new location in the genome. This Tc1 transposition rate was measured by looking for Tc1 insertions into the *unc-22* gene, which result in a nicotine-induced twitching phenotype. One mutant identified from this screen, *mut-2(r459),* showed a 100-fold increase in Tc1 excision from the *unc-54* gene and a 40-fold increase in Tc1 insertion into the *unc-22* gene. Later work from the Anderson lab and others revealed that *mut-2* mutants also showed activation of other DNA transposons (Collins *et al.* 1989; Collins and Anderson 1994), suggesting that *mut-2* is part of a general mechanism of DNA transposon silencing in *C. elegans.* A large part of my dissertation research focused on characterizing MUT-2 (now known as RDE-3, see below) in *C. elegans* and more information about MUT-2/RDE-3 will be provided below. Of note, the increase in Tc1 reversion in *mut-2* mutants and other mutants identified from this screen were only observed in the germline, not the soma.

As screens for animals defective for dsRNA-induced silencing were being performed simultaneously with screens for animals defective for transposon silencing, it became obvious that there was a significant overlap in the factors identified from both types of screens, indicating that transposon silencing is one endogenous function of RNAi in *C. elegans* (Collins *et al.* 1987; Ketting *et al.* 1999; Tabara *et al.* 1999; Vastenhouw *et al.* 2003; Tops *et al.* 2005; Chen *et al.*

2005). For instance, the Plasterk lab looked for mutants in which Tc1 showed increased rates of mobilization in the germline. One mutant that they identified, *mut-7*, showed increased mobilization of Tc1, as well as Tc3, Tc4 and Tc5, which suggested that MUT-7, like MUT-2/RDE-3, functions in a pathway that regulates global transposon silencing in *C. elegans*. Interestingly, *mut-7* mutants were also resistant to RNAi (Ketting *et al.* 1999). On the flip side, Craig Mello's lab identified seven mutants, mapping to four complementation groups, that showed resistance to dsRNA. Two of these mutants, *rde-2* and *rde-3* (later shown to harbor mutations in the same gene as the *mut-2* mutants (Collins *et al.* 1987) discussed above) also showed increased transposon mobilization, while the other two, *rde-1* and *rde-4*, did not (Tabara *et al.* 1999). Further, an RNAi screen for genes required for transposon silencing conducted by the Plasterk lab identified the *mut-16* mutant, which also showed an RNAi resistant phenotype. Together, this work revealed that an RNAi-like pathway silences DNA transposons in *C. elegans*. Indeed, it was shown that the terminal inverted repeats that flank the Tc1 DNA transposon could base-pair and form dsRNA, which could potentially feed into dsRNA-triggered RNAi pathway described above (Sijen and Plasterk 2003). I will next describe how some MUT/RDE proteins not only have the same phenotypes, but they also localize to the same phase-separated biomolecular condensate in the *C. elegans* germline.

iii. RDE and MUT proteins localize to Mutator *foci*

Eukaryotic cells are compartmentalized into organelles to allow for the segregation and specialization of functions. Some of these organelles, such as the nucleus or endoplasmic reticulum, are membrane-bound, meaning that they are surrounded by a lipid bilayer. However, other organelles, such as germ granules, Cajal bodies and stress granules, although clearly distinct from their surrounding nucleoplasm or cytoplasm, respectively, are membraneless. Such phase-separated membraneless organelles or biomolecular condensates are composed of nucleic acids and proteins, which come together and separate from their surrounding nucleoplasm

or cytoplasm via a process called liquid-liquid phase separation (LLPS). LLPS is thought to occur via multivalent intra- and intermolecular interactions that promote the oligomerization or polymerization of the protein, DNA and RNA components of biomolecular condensates, thereby increasing their local concentration and causing them to become insoluble in their surrounding nucleoplasm or cytoplasm and, thus, promoting their phase separation (Banani *et al.* 2017; Shin and Brangwynne 2017). Proteins with multiple interaction domains and proteins with intrinsically disordered regions are particularly important to participating in the multivalent interactions that promote LLPS. In 2009, P granules, germ granules found in the *C. elegans* germline, became the first membraneless organelles shown to form via LLPS (Brangwynne *et al.* 2009).

In 2012, Carolyn Phillips in Gary Ruvkun's lab generated worms expressing transgenes consisting of *mut/rde* genes fused to either *gfp* or *mcherry* and used immunofluorescence and live imaging to ask where the following six MUT/RDE protein localized: the ribonucleotidyltransferase MUT-2/RDE-3 (Collins *et al.* 1987; Chen *et al.* 2005), the conserved 3' exonuclease MUT-7 (Ketting *et al.* 1999), the RNA helicase MUT-14 (Tijsterman *et al.* 2002), the glutamine/asparagine (Q/N)-rich MUT-16/RDE-6 (Vastenhouw *et al.* 2003) and two proteins lacking known functional domains, MUT-8/RDE-2 (Tops *et al.* 2005) and MUT-15/RDE-5 (Phillips *et al.* 2012). Interestingly, she found that all six proteins colocalized to a perinuclear membraneless organelle in the *C. elegans* germline, termed the *Mutator* focus. Interestingly, these *Mutator* foci were shown to lie adjacent to P granules, the first identified germ granules in *C. elegans* (Strome and Wood 1982). Furthermore, in the absence of MUT-16, the other five MUT proteins failed to localize to perinuclear foci, indicating that MUT-16 nucleates the formation of *Mutator* foci in *C. elegans* (Phillips *et al.* 2012; Uebel *et al.* 2018). More recently, the Phillips lab performed a more detailed analysis of *Mutator* foci and showed that *Mutator* foci are, in fact, phase-separated condensates that form via liquid-liquid phase separation. MUT-16, ~70% of which is intrinsically disordered, promotes the phase separation of *Mutator* foci (Uebel *et al.* 2018).

Below I will discuss the function of MUT-2/RDE-3, one protein that localizes to *Mutator* foci and whose activity was the major focus of the work I present in Chapters 2 and 3.

iv. RDE-3

MUT-2/RDE-3 is a member of the polymerase beta-like nucleotidyltransferase superfamily that includes poly(A) polymerases, poly(U) polymerases, 2'-5'-oligoadenylate synthetases and CCA-adding enzymes (Aravind and Koonin 1999; Martin and Keller 2007). As discussed above, the *mut-2* mutant was initially identified genetically as being required for transposon silencing in *C. elegans* (Collins *et al.* 1987). Later, Craig Mello's lab noticed that *mut-2* mutants had another interesting phenotype: they were resistant to the gene silencing effects of dsRNA (Tabara *et al.* 1999). At the same time, the Mello lab identified the *rde-3* mutant from a screen for factors required for dsRNA-induced gene silencing and showed that *rde-3* had an increased frequency of transposon mobilization (Tabara *et al.* 1999). Finally, in 2005, the Mello lab used genetic mapping and showed that *mut-2* and *rde-3* mutants harbor mutations in the same gene, which, at the time, was known as K04F10.6 (Chen *et al.* 2005). Because the *mut-2/rde-3* gene was cloned using its Rde phenotype, we now refer to this gene as *rde-3.*

At the time, the RDE-3 protein was known to contain the conserved nucleotidyltransferase 2 domain (Chen *et al.* 2005) found in the polymerase beta-like nucleotidyltransferase superfamily that was known to include poly(A) polymerases, 2'-5'-oligoadenylate synthetases and the *S. cerevisiae* Trf4P (Aravind and Koonin 1999; Martin and Keller 2007)). This superfamily now also includes poly(U) polymerases and CCA-adding enzymes (Aravind and Koonin 1999; Martin and Keller 2007). RDE-3 was also known to harbor the PAP/25A-associated domain (Chen *et al.* 2005) found in the poly(A) polymerases of this superfamily, including *C. elegans* GLD-2 and *S. cerevisiae* Trf4p (Aravind and Koonin 1999; Chen *et al.* 2005)). Thus, RDE-3 was predicted to be a poly(A) polymerase and the Mello lab proposed two models for its function in RNAi silencing (Chen *et al.* 2005). The first model posited that in the absence of the predicted poly(A) polymerase

RDE-3, mRNAs with short poly(A) tails accumulate and these aberrant transcripts trigger RNAi-mediated silencing, thus consuming RNAi factors such that they are no longer available to respond to exogenous dsRNAs. The second model, which was based on the absence of siRNAs that was observed in *rde-3* mutants, suggested that RDE-3 may be required for the amplification of the RNAi response. In this model, exogenous dsRNAs were thought to be processed into primary siRNAs, which were low in abundance. These primary siRNAs could then be bound by Argonaute proteins and trigger the cleavage of a target mRNA. RDE-3 was then proposed to polyadenylate the 5' product of this cleavage event, thereby stabilizing it and allowing RdRPs to use this cleavage product as a template to amplify the siRNA pool (Chen *et al.* 2005). Chapter 2 of my thesis will present our data showing that this second model, although inaccurate in some of the details, namely the nucleotidyltransferase activity of RDE-3, was correct in its proposed function of RDE-3 in promoting the amplification of siRNAs.

While genetic studies helped to identify some of the functions of RDE-3 in *C. elegans,* the molecular function of RDE-3 was a mystery until very recently. The Wickens lab developed tethered ribonucleotidyltransferase (rNTase) activity identified by high-throughput sequencing (TRAID-seq) to identify the activity of rNTases (Preston *et al.* 2019), enzymes that non-templated ribonucleotides to RNAs and other substrates (Aravind and Koonin 1999; Martin and Keller 2007). In this heterologous system, a rNTase is fused to MS2 coat protein and this fusion protein is expressed in *S. cerevisiae* cells that express a reporter RNA containing MS2 stem loops, which are high-affinity binding sites for the MS2 protein. When MS2 binds to the MS2 stem loops on the reporter RNA, the rNTase of interest is brought into close proximity of the 3' end of the reporter RNA. High-throughput sequencing can then be used to ask how the rNTase modified the 3' end of the reporter RNA. Using this system, the Wickens lab showed that RDE-3 added perfectly alternating uridine (U) and guanosine (G) repeats to the 3' ends of two different reporter RNAs, a tRNA and an RNase P RNA. Results were confirmed by using a similar system in *Xenopus laevis* oocytes.

TRAID-seq is subject to a few pitfalls. For example, it is possible that yeast may not express the proper co-factors or proteins that an rNTase needs for its activity. Further, since rNTases are fused to a reporter RNA and not their natural substrates, the observed activity may not be biologically relevant. For these and other reasons, the activity of an rNTase observed using TRAID-seq may not reflect its true activity. Despite these caveats, TRAID-seq provides a useful starting point for identifying the endogenous activity of an rNTase, particularly because the activities of these enzymes cannot be deciphered simply from their primary amino acid sequences (Preston *et al.* 2019). Chapter 2 of my thesis focuses on asking whether RDE-3 is, indeed, a poly(UG) polymerase in its natural context in *C. elegans* and if so, what the function of RNAs modified with poly(UG) repeats may be.

B. Overview of endogenous RNAi pathways in C. elegans

The study of experimental dsRNA-induced gene silencing, also known as exogenous RNAi, helped to reveal that *C. elegans* expresses a number of endogenous small RNAs, such as miRNAs, 21U-RNAs (piRNAs), 26G siRNAs and 22G-siRNAs, that can also influence gene expression via RNAi-like pathways. Endogenous RNAi-related pathways in *C. elegans* are very complex, as evidenced by the existence of an entire clade of Argonaute proteins, called the worm-specific Argonautes or WAGOs, that exists solely in nematodes (Almeida *et al.* 2019; Ketting and Cochella 2020). I will provide a brief overview of some of these endogenous RNAi pathways here and then dive deeper into the *C. elegans* piRNA pathway, as this pathway was a focus of the work that I present in Chapter 3 of this dissertation.

Endogenous RNAi pathways in *C. elegans* can be distinguished based on their Argonaute and small RNA participants. One of these pathways is the miRNA pathway. Like miRNAs in other animals, *C. elegans* miRNAs are genomically-encoded and transcribed by RNA polymerase II as long precursor transcripts. These long transcripts are sequentially processed by the endonucleases Drosha and Dicer and then loaded into the redundant Argonautes ALG-1 and

ALG-2 (Grishok *et al.* 2001; Tops *et al.* 2006). The mechanism by which miRNAs then induce the post-transcriptional silencing of their target mRNAs is highly conserved and was discussed above. The *C. elegans* genome encodes approximately 145 miRNAs, most of which show some degree of cell-type specific expression driven by the action of their promoters and enhancers (Martinez *et al.* 2008; Alberti *et al.* 2018). These microRNAs often function redundantly with other microRNAs that have shared seed sequences to regulate a variety of processes, including growth, behavior, cell fate determination and stress responses (Ambros and Ruvkun 2018; Fromm *et al.* 2020).

The remaining endogenous RNAi pathways, like the exogenous RNAi pathway described above, can be divided into an initiation phase, during which a target mRNA is recognized by a ribonucleoprotein complex composed of a so-called primary Argonaute and primary small RNA, followed by an amplification phase, during which the target mRNA is thought to become a template for RdRP-dependent secondary 22G-siRNA biogenesis. Secondary 22G-siRNAs are then loaded into WAGO proteins (Yigit *et al.* 2006; Guang *et al.* 2008; Gu *et al.* 2009; Shirayama *et al.* 2012). What happens after WAGO loading is somewhat mysterious, but in some cases, WAGO-bound siRNAs can induce co-transcriptional gene silencing in the nucleus (see below).

Primary Argonautes in *C. elegans* include RDE-1 (discussed above), PRG-1, ERGO-1, ALG-3 and ALG-4. PRG-1 is the functional PIWI protein in *C. elegans* and will be discussed in detail below (Wang and Reinke 2008; Batista *et al.* 2008; Das *et al.* 2008). The ERGO-1, ALG-3 and ALG-4 Argonautes bind to primary small RNAs referred to as 26G siRNAs because they are 26nt in length and tend to begin with a G. These 26G siRNAs are not genomically-encoded; instead, they are made by the ERI complex. This complex includes the RdRP RRF-3, which uses an RNA template to make a dsRNA intermediate that then gets processed by Dicer into 26G siRNAs (Vasale *et al.* 2010; Gent *et al.* 2010; Welker *et al.* 2011; Thivierge *et al.* 2012; Almeida *et al.* 2018). ERGO-1 binds to 26G-siRNAs that are expressed in the female germline and in embryos. These 26G-siRNAs (like *C. elegans* piRNAs, see below) are 2'-O-methylated by HENN-

1. The ERGO-1/26G siRNA complex then locates target mRNAs, which include pseudogenes, recent gene duplications and transcripts with defective splicing (Vasale *et al.* 2010; Gent *et al.* 2010; Fischer *et al.* 2011; Newman *et al.* 2018). ERGO-1/26G-siRNA binding to a target transcript triggers 22G-siRNA biogenesis, mostly mediated by the RdRP RRF-1, via a pathway that at least partially competes for protein resources with the *C. elegans* exogenous RNAi pathway (Kennedy *et al.* 2004; Duchaine *et al.* 2006; Gu *et al.* 2009). These 22G-siRNAs then engage WAGO proteins, such as the somatically-expressed, nuclear Argonaute NRDE-3 (Gent *et al.* 2010), which triggers co-transcriptional gene silencing in the soma (Guang *et al.* 2008). How exactly the 22G-siRNA products induced by a germline-expressed primary Argonaute are able to get into somatic tissues is a mystery. The paralogous Argonautes ALG-3 and ALG-4 (ALG-3/4) bind to 26G-siRNAs, which are not 2'-O-methylated, in the male germline. ALG-3/4 and their bound 26G-siRNAs also trigger 22G-siRNA biogenesis, but these 22G-siRNAs have been reported to repress expression some spermatogenesis-related genes, while promoting the expression of others (Conine *et al.* 2010, 2013).

The 22G-siRNAs I have discussed so far (i.e. those triggered by exogenous dsRNAs, ERGO-1/26G-siRNAs and PRG-1/piRNAs) all fall into the WAGO family of 22G-siRNAs. These 22G-siRNAs engage WAGO proteins and depend on the RdRPs EGO-1 and RRF-1 for their biogenesis (Gu *et al.* 2009). Interestingly, the biogenesis of the majority of these 22G-siRNAs is thought to occur in *Mutator* foci, biomolecular condensates in the *C. elegans* germline that were discussed above (Zhang *et al.* 2011; Phillips *et al.* 2012; Uebel *et al.* 2018). How different primary Argonaute/small RNA complexes can all hand-off their target mRNAs into *Mutator* foci is not well understood. Further, how 22G-siRNAs triggered by different primary Argonaute/small RNA complexes could all be synthesized in the same *Mutator* foci, but then engage different WAGO proteins is also a mystery (Yigit *et al.* 2006).

One class of 22G-siRNAs that does not have a clear primary Argonaute/small RNA complex that triggers its production is the CSR-1 class of 22G-siRNAs. These siRNAs are made

solely by the RdRP EGO-1 and bound by the WAGO protein CSR-1. Interestingly, both CSR-1 (but none of the other WAGO proteins) and EGO-1 are essential proteins (Claycomb *et al.* 2009; Maniar and Fire 2011). The role of CSR-1 and its bound 22G-siRNAs has been enigmatic and somewhat controversial. One prevailing model for CSR-1 function suggests that CSR-1 promotes the expression of germline-expressed mRNAs, perhaps as a way to counteract the silencing activity of PRG-1 and piRNAs (Shen *et al.*; Claycomb *et al.* 2009; Seth *et al.* 2013; Wedeles *et al.* 2013). However, several other findings contradict this model. First, CSR-1 can and does cleave target mRNAs that are complementary to 22G-siRNAs (Aoki *et al.* 2007). Further, CSR-1 can induce silencing of some mRNAs. For instance, in the absence of CSR-1, the *klp-7* mRNA becomes upregulated. The *klp-7* gene encodes the only MCAK-related kinesin-13 in *C. elegans* and the upregulation of *klp-7* is thought to drive the chromosome segregation defects observed in *csr-1* mutant animals (Gerson-Gurwitz *et al.* 2016). Further, newer studies (discussed in more detail below) suggest that PRG-1 and piRNAs, not CSR-1, might actually prevent germline-expressed mRNAs from aberrant 22G-siRNA-mediated silencing (de Albuquerque *et al.* 2015; Phillips *et al.* 2015; Barucci *et al.* 2020; Reed *et al.* 2020). Taken together, conflicting data has made it difficult to determine what the role of CSR-1 (and PRG-1 as I will discuss below) is in *C. elegans*. Perhaps some of these conflicting data arise from target-dependent anti-silencing and pro-expression functions of Argonaute protein. In addition, in an organism that has so many different endogenous small RNAs pathways, many of which converge on the same proteins and RNA molecules, one must consider how these different pathways interact with and affect one another and, therefore, how loss of one these pathways may have direct and indirect effects on other pathways that make the resulting gene expression changes difficult to interpret.

Below, I will discuss the *C. elegans* piRNA pathway in detail. The primary Argonaute associated with this pathway is the PIWI protein PRG-1. While transposon silencing is a conserved function of piRNAs in animals (Siomi *et al.* 2011; Weick and Miska 2014; Ozata *et al.*

2019), I will discuss how transposon silencing does not seem to be the main function of piRNAs in *C. elegans*.

C. piRNAs in C. elegans

1. Discovery of C. elegans *piRNAs*

piRNAs in *C. elegans* are known as 21U-RNAs because of their 21nt length and propensity for beginning with a 5' uridine (U). These small RNAs were first discovered in the worm in 2006, when David Bartel's lab used high-throughput pyrosequencing to sequence ~400,000 small RNAs from mixed-stage *C. elegans* (Ruby *et al.* 2006). The Bartel lab found ~5,000 unique 21U-RNAs, the majority of which mapped to two regions of chromosome IV. These 21U-RNAs were sensitive to alkaline hydrolysis and phosphatase treatment and could act as substrates for RNA ligase, indicating that they harbored a 5' monophosphate. Resistance to periodate treatment suggested that the 3' end of these 21U-RNAs was missing a *cis* diol and that either the 2' or 3' oxygen might be modified. In addition, the genomic loci encoding the 21U-RNAs shared two upstream sequence motifs, separated by an ~20nt spacer sequence. The large motif was 34bp long and had an 8nt long core consensus sequence (CTGTTTCA). At the time, these motifs were proposed to act as a promoter and each 21U-RNA was hypothesized to represent an independent transcriptional unit (Ruby *et al.* 2006). It was noted that these 21U-RNAs shared many features with piRNAs in other organisms, including their 5' uridine bias and tight clustering in the genome.

Soon after their discovery, 21U-RNAs were shown to be the piRNAs of *C. elegans*. The *C. elegans* genome encodes 27 Argonaute proteins, two of which, PRG-1 and PRG-2, belong to the PIWI clade that binds piRNAs in most animals (Meister 2013). In 2008, three groups showed that animals lacking PRG-1 failed to express 21U-RNAs. Further, these groups found that, even though PRG-1 and PRG-2 share 91% identity at the amino acid level, PRG-1 is likely the only functional PIWI protein in *C. elegans,* as 21U-RNA levels did not change in *prg-2* mutants (Wang and Reinke 2008; Batista *et al.* 2008; Das *et al.* 2008). Using Northern blotting and germline-

deficient mutant animals, all three groups showed that 21U-RNAs appear to be expressed only in the *C. elegans* germline. Indeed, PRG-1 was also shown to be expressed in the germline, localizing to P granules (Wang and Reinke 2008; Batista *et al.* 2008; Das *et al.* 2008), biomolecular condensates found perinuclear to germ cell nuclei in *C. elegans*. Immunoprecipitation of PRG-1 revealed that the protein physically interacts with 21U-RNAs (Batista *et al.* 2008). In addition, *prg-1* mutants were found to exhibit brood size defects at 20°C and almost complete loss of fertility at 25°C (Wang and Reinke 2008; Batista *et al.* 2008; Das *et al.* 2008), consistent with fertility defects observed for piRNA mutants in other organisms (Weick and Miska 2014; Ozata *et al.* 2019). *prg-2* mutants, on the other hand, did not show these fertility defects and *prg-1; prg-2* double mutants had the same phenotype as *prg-1* single mutants alone (Wang and Reinke 2008; Batista *et al.* 2008; Das *et al.* 2008). Together, this work indicated that PRG-1 is the functional PIWI Argonaute that binds to *C. elegans* piRNAs in the worm germline. The function of PRG-2 still remains elusive.

2. Biogenesis of C. elegans *piRNAs*

Subsequent work revealed more about how piRNAs in *C. elegans* are produced, processed and modified. While piRNAs in most organisms are produced by the transcription and processing a long precursor transcript (Ozata *et al.* 2019), *C. elegans* piRNAs are unique in that each piRNA is transcribed by RNA polymerase II as independent transcriptional units. The upstream sequence motif described by the Bartel lab (Ruby *et al.* 2006) does, in fact, act as a promoter for piRNA biogenesis (Cecere *et al.* 2012; Billi *et al.* 2013). The 8nt long core consensus sequence of this motif, now called the Ruby motif after the first author of the paper that described it (Ruby *et al.* 2006), is recognized by Forkhead transcription factors (Cecere *et al.* 2012) and transcription of individual piRNAs is initiated 2nt upstream of the 5' U found in mature piRNAs. While different groups have obtained different results as to the exact length of these precursor transcripts (Cecere *et al.* 2012; Gu *et al.* 2012), the most recent data suggests that each transcript

is 28-29nt in length and modified with a 5′ 7-methylguanylate cap (Goh *et al.* 2014; Weick *et al.* 2014). A forward genetic screen identified piRNA defective 1 (PRDE-1), a nuclear localized and germline-expressed protein that associates with the piRNA clusters found on chromosome IV. PRDE-1 is required for the biogenesis or stability of all piRNA precursor transcripts derived from Ruby motif-containing piRNA loci, but not for the expression of a small group of piRNA loci that do not contain a Ruby motif. TOFU-3, TOFU-4 and TOFU-5 are additional factors required for piRNA precursor biogenesis or stability and were identified from an RNAi screen performed by the Hannon lab, although their exact molecular functions are not yet known (Goh *et al.* 2014).

Once piRNA precursor transcripts are made, they are exported from the nucleus to the cytoplasm, where they undergo processing and are bound by PRG-1 once they are processed. The piRNA precursor transcripts are first cleaved at their 5' and 3' ends, leaving an uncapped 21nt piRNA that begins with a U. This cleavage process appears to require TOFU-1 and TOFU-2, as mutants lacking these factors do not express any mature piRNAs and have elevated levels of precursor transcripts (Goh *et al.* 2014). PID-1 (piRNA-induced silencing defective 1) is another factor that seems to be important for piRNA processing in the cytoplasm as mutants that lack PID-1 show upregulation of piRNA precursor transcripts (de Albuquerque *et al.* 2014). The RNA methyltransferase HENN-1 then 2'-O-methylates the 3' ends of piRNAs, resulting in mature piRNAs that get bound by PRG-1 (Montgomery *et al.* 2012; Kamminga *et al.* 2012). This last step of 2'-O-methylation is conserved in piRNAs from different species (Ozata *et al.* 2019) and was hinted at by the Bartel lab's observation that 21U-RNAs were resistant to periodate treatment. Binding of PRG-1 to mature piRNAs is thought to protect these RNAs from degradation, such that most piRNAs are not expressed in *prg-1* mutants (Wang and Reinke 2008; Batista *et al.* 2008; Das *et al.* 2008).

3. Functions of C. elegans *piRNAs*

i. piRNAs and transposon silencing

Ever since piRNAs were discovered in *C. elegans,* their function has been somewhat mysterious. One conserved function of piRNAs in animals is to promote transposon silencing (Siomi *et al.* 2011; Weick and Miska 2014; Ozata *et al.* 2019). Indeed, silencing of the Tc3 DNA transposon family depends on PRG-1 and piRNAs and there is a single piRNA that is complementary to all 22 copies of the Tc3 transposon in *C. elegans*. Interestingly, 22G-siRNAs targeting Tc3 were depleted in a *prg-1* mutant, raising the possibility that perhaps PRG-1/piRNA complexes induce the production of RdRP-dependent 22G-siRNAs (Batista *et al.* 2008) and these 22G-siRNAs might induce target mRNA silencing via WAGO proteins (Gu *et al.* 2009). As discussed below, this mechanism was later shown to be true (Ashe *et al.* 2012; Shirayama *et al.* 2012). However, while Tc3 silencing does depend on PRG-1 and piRNAs*,* piRNAs are not required for the silencing of most *C. elegans* transposable elements (Wang and Reinke 2008; Batista *et al.* 2008; Das *et al.* 2008; Bagijn *et al.* 2012; Barucci *et al.* 2020; Reed *et al.* 2020). Indeed, fewer than 1% of transposon families are upregulated in a *prg-1* mutant, which lacks all piRNAs (Barucci *et al.* 2020).

ii. piRNA-induced transgene silencing

In addition to mediating silencing of the Tc3 transposon, PRG-1 and piRNAs also promote transgene silencing in *C. elegans*. This piRNA-induced silencing is not dependent on the presence of a perfectly complementary piRNA binding site in the transgene sequence. Forward genetic screens have revealed how piRNAs are able to promote transgene silencing in the worm (Ashe *et al.* 2012; Shirayama *et al.* 2012). Current models posit that PRG-1/piRNA complexes bind target mRNAs and recruit RNA-dependent RNA Polymerases (RdRPs) (Lee *et al.* 2012; Bagijn *et al.* 2012), which use the target mRNA as a template to produce 22G-siRNAs (Gu *et al.* 2009). At least some of these RdRP-generated 22G-siRNAs are then thought to engage the primarily nuclear WAGO, Heritable RNAi-Defective 1 or HRDE-1 (Ashe *et al.* 2012; Shirayama *et al.* 2012; Buckley *et al.* 2012), and, together, this Argonaute/siRNA complex is thought to enter

the nucleus and locate a nascent transcript via complementary base pairing to the siRNA. The HRDE-1/22G-siRNA complex then mediates the assembly of the nuclear RNAi complex, which includes the proteins NRDE-1, NRDE-2 and NRDE-4 (Ashe *et al.* 2012; Buckley *et al.* 2012), and, together, this nuclear RNAi machinery mediates the: (1) deposition of repressive chromatin marks, such as H3K9me3 (Ashe *et al.* 2012; Buckley *et al.* 2012); and (2) pausing of RNA polymerase II at the transcription elongation phase at target loci (Guang *et al.* 2008, 2010; Burkhart *et al.* 2011). A related pathway that mediates dsRNA-induced heritable gene silencing will be discussed below.

iii. piRNAs and endogenous mRNA regulation

Recent studies have shown that PRG-1 and piRNAs also interact with endogenous protein-coding mRNAs in the *C. elegans* germline (Shen *et al.* 2018; Zhang *et al.* 2018). However, if and how PRG-1 and piRNAs influence the expression of these mRNAs is somewhat mysterious. One study used a crosslinking, ligation and sequencing of hybrids (CLASH) protocol to identify all of the piRNAs to which PRG-1 binds, as well as to identify all of the target mRNAs with which PRG-1 and its bound piRNAs interact. This work revealed that PRG-1 and its bound piRNAs interact with >16,000 mRNAs, the majority of which are expressed in the *C. elegans* germline (Shen *et al.* 2018). This result was somewhat surprising for many reasons (see below), one being that piRNAs match only ~1% of *C. elegans* mRNAs with perfect complementarity (Lee *et al.* 2012). However, the CLASH data revealed that the interaction between a piRNA and a target mRNA is able to tolerate some mismatches, except for where the target mRNA interacts with the second to eighth nucleotide of the piRNA (Shen *et al.* 2018). This region of perfect complementarity is reminiscent of the miRNA seed region (Bartel 2009).

Despite PRG-1 and piRNAs interacting with >16,000 mRNAs in *C. elegans,* fewer than 100 mRNAs undergo piRNA-dependent gene silencing (Barucci *et al.* 2020). Of note, the HRDE-1/22G-siRNA-based mechanism of piRNA-induced transgene silencing described above is

thought to mediate the silencing of these mRNAs. Meanwhile, the expression of the majority of mRNAs to which PRG-1 and piRNAs bind is unaffected in *prg-1* mutants, indicating that PRG-1 and piRNAs do not silence most of the mRNAs to which they are bound. Furthermore, several genes, most prominently the replication-dependent histones genes, undergo gene silencing in the absence of PRG-1 (Barucci *et al.* 2020; Reed *et al.* 2020). Silencing correlates with production of antisense 22G-siRNAs that are produced in the absence of PRG-1 and piRNAs (Barucci *et al.* 2020; Reed *et al.* 2020). Why some genes get aberrantly silenced in a *prg-1* mutant and how PRG-1 and piRNAs normally promote the expression of these genes remains a mystery. Furthermore, why PRG-1 and piRNAs interact with most of the *C. elegans* transcriptome is also not known.

iv. P granules restrict PRG-1 silencing activity

PRG-1 localizes to biomolecular condensates called P granules in the *C. elegans* germline (Wang and Reinke 2008; Batista *et al.* 2008) by interacting with the constitutive P granule protein DEPS-1 (Suen *et al.* 2020). Recent studies suggest that P granules help to restrict PRG-1/piRNA-induced gene silencing on the correct RNAs. Indeed, in *meg-3/meg-4* double mutants, which lack P granules during embryogenesis (Wang *et al.* 2014b), PRG-1 becomes mislocalized to the cytoplasm (Ouyang *et al.* 2019), where it promotes 22G-siRNA biogenesis from and silencing of the wrong mRNAs, including mRNAs derived from the *rde-11* and *sid-1* genes. The expression of these aberrant 22G-siRNAs is dependent on HRDE-1 (Ouyang *et al.*; Dodson and Kennedy 2019). *rde-11* (Zhang et al. 2012) and *sid-1* (Winston *et al.* 2002) are both required for exogenous RNAi in *C. elegans.* As a result of aberrant silencing of these two genes in *meg-3/meg-4* mutants, these mutants are unable to induce gene silencing in response to exogenous dsRNAs (Ouyang *et al.*; Dodson and Kennedy 2019). Aberrant *rde-11* and *sid-1* 22G-siRNA biogenesis and silencing is just one symptom of a global disorganization of 22G-siRNA levels in *meg-3/meg-4*

mutants (Ouyang *et al.*; Dodson and Kennedy 2019). However, it is not known whether all aberrant 22G-siRNAs produced in *meg-3/meg-4* mutants are PRG-1-dependent.

III. Transgenerational epigenetic inheritance (TEI) in *C. elegans*

Below, I will discuss how, sometimes, the small RNA and RNAi-based gene silencing mechanisms I discussed above can transmit gene silencing information across generations.

A. Overview of TEI

1. Epigenetics

The somatic cells of multicellular organisms carry the same DNA in their nuclei, yet these cells display obvious morphological and functional differences. Underlying these differences are vastly different transcriptomes and proteomes. The ability of these diverse cell types to emerge from a single-celled embryo and to maintain their cellular identities across many rounds of mitotic cell division is mediated by epigenetics.

The definition of the term epigenetics has undergone several transformations and has been somewhat controversial (Ptashne 2013; Deans and Maggert 2015) ever since the term was first introduced by Conrad Waddington in 1942 (WADDINGTON and H 1942). Waddington was a developmental biologist who used epigenetics to refer to the ability of a gene or genotype to give rise to multiple different phenotypes during development (WADDINGTON and H 1942). Robin Holliday moved beyond a definition of epigenetics that focused only on gene regulation and, instead, made heritability a necessary component of epigenetic phenomenon by referring to epigenetics as the study of changes in gene expression or function that arise independent of changes in DNA sequence and can be inherited through mitosis and/or meiosis (Holliday 1994). The requirement for heritability in Holliday's definition has been debated (Gibney and Nolan 2010;

Deans and Maggert 2015). I will use epigenetics to refer to changes in gene expression that are independent of changes in DNA sequence and are stable in the absence of initiating triggers (Holoch and Moazed 2015a; Horsthemke 2018). Mechanisms that mediate epigenetic gene expression changes include DNA modifications, histone modifications and noncoding RNAs.

2. TEI

Recent work has shown that epigenetic information can be inherited not only over mitotic cell divisions during development, but also transmitted through the germline across generations. The transmission of epigenetic information from one generation to the next is called intergenerational epigenetic inheritance (Heard and Martienssen 2014; Horsthemke 2018; Perez and Lehner 2019). For instance, in *D. melanogaster,* maternally inherited piRNAs can direct heritable silencing of transposable elements in progeny (Le Thomas *et al.* 2014). In addition, male mice fed low-protein can transmit tRNA fragments to their progeny via their sperm. These inherited tRNA fragments have been correlated with changes in transposon gene expression in offspring (Chen *et al.* 2016; Sharma *et al.* 2016).

Sometimes, epigenetic information can cross generational boundaries for multiple generations, a phenomenon called transgenerational epigenetic inheritance (TEI). For a phenomenon to be truly transgenerational, rather than intergenerational, the effects on an organism cannot be ascribed to a direct exposure to the trigger (Perez and Lehner 2019). For example, if a pregnant female mouse is exposed to a stimulus, any effects of that stimulus that are seen in her progeny or grand-progeny would not be considered transgenerational effects since her fetus, as well as the germline of her fetus (i.e. her grand-progeny) could have been exposed to the stimulus *in utero*. Examples of TEI include the siRNA-based inheritance of pericentromeric heterochromatin in *S. pombe,* protein-based inheritance in yeast (Shorter and Lindquist 2005; Harvey *et al.* 2020), dsRNA-mediated gene silencing in *C. elegans* (Vastenhouw *et al.* 2006; Ashe *et al.* 2012; Buckley *et al.* 2012), pathogen avoidance behavior in *C. elegans*

and paramutation in plants. The discovery of TEI was particularly surprising because germline reprogramming during germ cell differentiation and after fertilization was thought to erase any acquired chromatin marks or other epigenetic cues in order to ensure that the germline is a clean slate at the start of each generation (Heard and Martienssen 2014; Horsthemke 2018).

Studies of TEI have revealed that, although prion proteins, histone modifications and DNA methylation can mediate TEI, small noncoding RNAs are the major drivers of TEI across species (Heard and Martienssen 2014; Holoch and Moazed 2015a; Bošković and Rando 2018; Perez and Lehner 2019; Duempelmann *et al.* 2020). A well-studied example of RNA-based TEI is paramutation in plants. Defined by Alexander Brink in 1956 (Brink 1956), paramutation is defined as the interaction between alleles of a gene during which one allele can affect the gene expression state of another allele. An example of paramutation occurs at the *b1* locus (Coe 1966), which encodes a transcription factor that promotes the biosynthesis of purple anthocyanin pigments, in maize. Plants can have one of two alleles of the *b1* locus. Plants with the *B'* allele have light pigmentation as a result of low *b1* expression, whereas plants that have *B-I* allele are dark purple as a result of high *b1* expression. Interestingly, the sequences of these two alleles are identical. Further, in plants that are heterozygous for the two alleles, the *B'* allele can induce silencing of *B-I* allele and this *trans*-silencing is fully penetrant in future generations (Coe, 1996). *Trans*-silencing is mediated by an RNAi-like pathway, in which siRNAs are produced from tandem repeats located 100kb upstream of the *b1* transcription start site. These siRNAs induce repressive chromatin modifications, particularly DNA methylation, at the *b1* locus in *B'* plants and are able to act *in trans* to irreversibly silence the *B-I* allele in future generations (Arteaga-Vazquez and Chandler 2010). Paramutation is best-studied in plants, but examples of paramutation-like phenomenon, have been reported in worms (Ashe *et al.* 2012; Shirayama *et al.* 2012), flies (Hermant *et al.* 2015) and mice (Rassoulzadegan *et al.* 2006) and these other examples of paramutation have also been shown to be mediated by RNA.

Below I will review some examples of TEI reported in *C. elegans.* Interestingly, although the triggers that initiate the examples of TEI that I will discuss below are different (ex. dsRNA vs. pathogenic bacteria), many of these examples seem to converge on the same siRNA-based co-transcriptional gene silencing mechanism. The ability of the same mechanism to mediate TEI in response to a diversity of triggers begs the question of how diverse triggers are "read" and "translated" such that they can tap into the same siRNA/co-transcriptional gene silencing response.

B. Examples of TEI in C. elegans

1. RNAi inheritance

C. elegans emerged as a model system for the study of TEI when it was discovered that exogenously-introduced dsRNA could silence homologous genes and that this silencing was heritable for four or more generations, after which gene expression is restored (Fire *et al.* 1998; Vastenhouw *et al.* 2006; Ashe *et al.* 2012; Buckley *et al.* 2012; Spracklin *et al.* 2017). This heritable silencing is truly an epigenetic phenomenon as it is not accompanied by any changes in DNA sequence. The first report of the heritable nature of dsRNA-induced gene silencing (RNAi) dates back to the same paper that described RNAi for the first time (Fire *et al.* 1998). Since then, forward genetic screens have been instrumental to revealing how this dsRNA-induced inheritance (termed RNAi inheritance) occurs (Ashe *et al.* 2012; Buckley *et al.* 2012; Spracklin *et al.* 2017; Wan *et al.* 2018). Importantly, these screens have revealed that RNAi inheritance is not simply a passive phenomenon, but is instead mediated by an active process that amplifies and transmits an epigenetic silencing response across generations. Of note, transgenerational RNAi inheritance has been observed only for germline-expressed genes (Ashe *et al.* 2012; Buckley *et al.* 2012; Spracklin *et al.* 2017; Wan *et al.* 2018). Soma-expressed genes can be heritably silenced for one generation after dsRNA treatment, but this intergenerational inheritance will not be discussed in detail here (Burton *et al.* 2011).

RNAi inheritance is mediated by a co-transcriptional gene silencing phenomenon called nuclear RNAi (Figure 1.2) (Ashe *et al.* 2012; Buckley *et al.* 2012; Spracklin *et al.* 2017; Wan *et al.* 2018) that is reminiscent of the siRNA-based silencing of pericentromeric heterochromatin in *S. pombe* (Holoch and Moazed 2015a). As described above, RNAi begins in *C. elegans,* when dsRNA is processed by the endoribonuclease Dicer and the double-stranded RNA binding protein RDE-4 into short dsRNAs (Parrish and Fire 2001; Tabara *et al.* 2002; Parker *et al.* 2006). One strand of this short dsRNA (termed 1° siRNA) is then bound by the Argonaute (AGO) RDE-1 (Tabara *et al.* 1999; Parrish and Fire 2001) and together, this siRNA/AGO complex recognizes a target mRNA via Watson-Crick base-pairing to the siRNA and helps to assemble the downstream cytoplasmic RNAi machinery onto this target mRNA. RdRP proteins are recruited to the target mRNA and generate an amplified pool of siRNAs, called 22G-siRNAs, that are antisense to the target mRNA. These 22G-siRNAs are bound by worm-specific Argonaute proteins called WAGO proteins, which carry out gene silencing (Yigit *et al.* 2006; Gu *et al.* 2009).

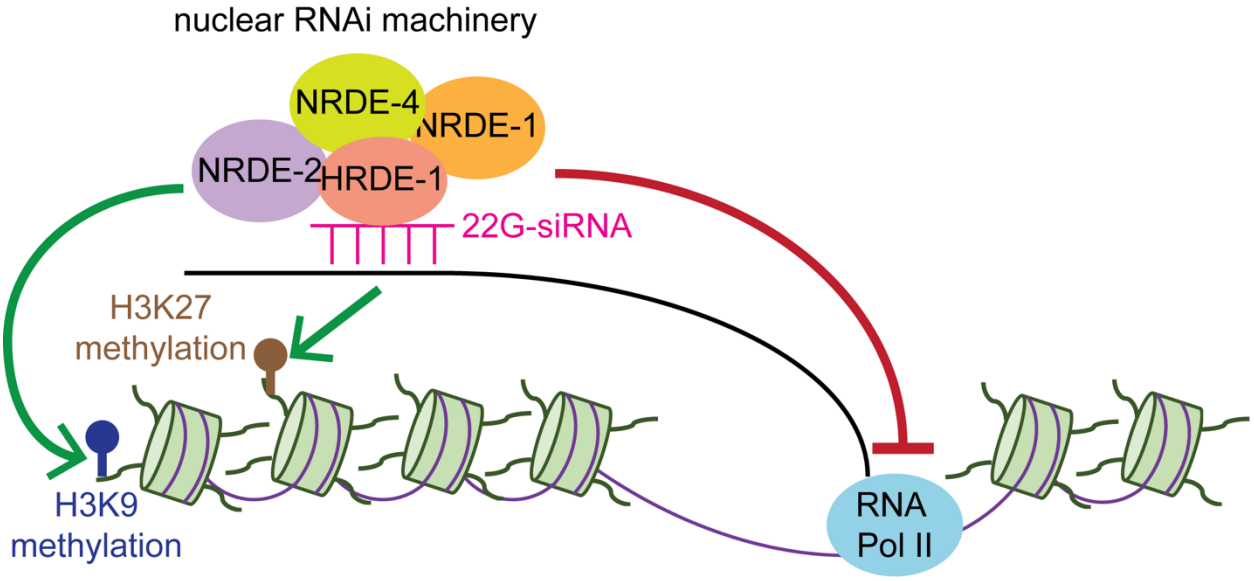


Figure 1.2. Nuclear RNAi in the *C. elegans* germline. The nuclear WAGO HRDE-1 binds to 22G-siRNAs and, together, this complex locates RNA Polymerase II (RNA Pol II)-dependent nascent transcripts based on complementarity to the bound 22G-siRNAs. Of note, it is not currently known whether HRDE-1 gets loaded with 22G-siRNAs in the cytoplasm or nucleus.

Figure 1.2 (Continued). HRDE-1 then recruits downstream RNAi factors, including NRDE-1, NRDE-2 and NRDE-4. Together, this complex mediates the: (1) deposition of repressive chromatin modifications, namely H3K9 methylation and H3K27 methylation; and (2) inhibition of RNA Pol II during the elongation step of transcription (Guang *et al.* 2008, 2010; Burkhart *et al.* 2011; Burton *et al.* 2011; Buckley *et al.* 2012; Mao *et al.* 2015)**.** This pathway is thought to operate downstream of dsRNA-induced gene silencing, piRNA-induced transgene silencing and piRNA-induced endogenous gene silencing in the *C. elegans* germline (Ashe *et al.* 2012; Shirayama *et al.* 2012; Buckley *et al.* 2012).

In the *C. elegans* germline, RdRP-generated 22G-siRNAs can be bound by HRDE-1 (heritable RNAi-defective 1), a WAGO that shows primarily nuclear localization (Buckley *et al.* 2012). Loading of HRDE-1 with dsRNA-induced siRNAs depends on HRDE-2, potentially by an analogous mechanism to how the Argonaute small interfering RNA chaperone (ARC) complex helps to load Ago1 with siRNAs in *S. pombe* (Holoch and Moazed 2015b). Together, the HRDE-1/22G-siRNA complex locates a nascent transcript via complementary base pairing to the siRNA and mediates the assembly of the nuclear RNAi complex, which includes the proteins NRDE-1, NRDE-2 and NRDE-4 (Ashe *et al.* 2012; Buckley *et al.* 2012). The nuclear RNAi machinery then mediates the: (1) deposition of repressive chromatin marks, such as H3K9me3 (Ashe *et al.* 2012; Buckley *et al.* 2012) and H3K27me3 (Mao *et al.* 2015), at the genomic locus being silenced by RNAi; and (2) pausing of RNA polymerase II at the transcription elongation phase (Guang *et al.* 2008, 2010; Burkhart *et al.* 2011). The deposition of H3K9me3 is at least partially mediated by SET-32, a histone methyltransferase in *C. elegans* (Spracklin *et al.* 2017). Of note, an analogous nuclear RNAi pathway exists in somatic cells where, instead of HRDE-1, the nuclear Argonaute NRDE-3 mediates transcriptional gene silencing (Guang *et al.* 2008, 2010; Burkhart *et al.* 2011). Indeed, this somatic nuclear RNAi pathway mediates the intergenerational inheritance of RNAi targeting soma-expressed genes mentioned earlier (Burton *et al.* 2011).

Heritable silencing during RNAi inheritance correlates with both the transgenerational expression of 22G-siRNAs antisense to the dsRNA-targeted mRNA and the persistence of H3K9me3 at the corresponding genomic locus for several generations, but both epigenetic marks

decline with each inheriting generation, concomitant with the loss of heritable silencing (Ashe *et al.* 2012; Buckley *et al.* 2012; Sapetschnig *et al.* 2015). How exactly these epigenetic marks can persist in the absence of any additional dsRNA silencing signal is a mystery. HRDE-1 is required in inheriting generations to maintain gene silencing after RNAi, suggesting that nuclear RNAi is part of the mechanism that maintains RNAi inheritance. Maintenance of 22G-siRNAs presumably involves RdRP-based amplification because any 22G-siRNAs generated from the initial dsRNA trigger would quickly be diluted with each inheriting generation since each hermaphrodite lays 250-350 progeny (Lev *et al.* 2017). Another mystery about RNAi inheritance is why, if amplification mechanisms exist to maintain silencing for a few generations, silencing eventually stops being maintained. Chapters 2 and 3 of my dissertation will address both of these questions.

Mutants lacking HRDE-1 and the other known HRDE proteins also display another transgenerational phenotype: progressive sterility at elevated temperatures (Buckley *et al.* 2012; Spracklin *et al.* 2017). This phenotype, known as Mrt (mortal germline) (Ahmed and Hodgkin 2000), results from the misregulation of the endogenous targets of HRDE-1 and the germline nuclear RNAi pathway. Indeed Mrt correlates with a loss of H3K9me3 at genomic loci corresponding to siRNAs to which HRDE-1 is normally bound, as well as upregulation of these HRDE-1 target genes (Buckley *et al.* 2012). Thus the germline nuclear RNAi machinery appears to promote germline immortality, presumably by helping to maintain the *C. elegans* epigenome at HRDE-1 target genes.

i. RNAi inheritance and germ granules

More recently, *C. elegans* germ granules have also been implicated in RNAi inheritance. Work from our lab showed that two proteins found in P granules (Gruidl *et al.* 1996; Spike *et al.* 2008a) are required for RNAi inheritance: the highly conserved Vasa homolog, GLH-1 (Spracklin *et al.* 2017), and the constitutive P granule protein, DEPS-1 (Wan *et al.* 2018). Additional forward genetic screens for factors required for RNAi inheritance in *C. elegans* also discovered two new

proteins as being required for RNAi inheritance: ZNFX-1, a highly conserved NFX1-type zinc finger-containing protein 1 homolog, and WAGO-4, one of the 27 Argonaute proteins found in *C. elegans*. *znfx-1* and *wago-4* mutants showed normal 22G-siRNA expression and gene silencing when fed RNAi. Instead, ZNFX-1 and WAGO-4 were specifically required for the expression of 22G-siRNAs, and gene silencing, in the inheriting generations after RNAi (Wan *et al.* 2018). Interestingly, ZNFX-1 and WAGO-4 were shown to localize to Z granules, a third liquid-like condensate in the *C. elegans* germline (Ishidate *et al.* 2018; Wan *et al.* 2018). However, Z granules display highly dynamic localization in the *C. elegans* germline. Confocal imaging revealed that ZNFX-1 and WAGO-4 colocalize with the P granule protein, PGL-1, during early embryonic cell divisions. However, sometime after the eight cell embryo stage, an interesting transition occurs where ZNFX-1 and WAGO-4 begin to separate from PGL-1. At around the 88-cell stage, the P_4 germline blastomere cell divides into the primordial germ cells, Z_2 and Z_3. At this point, ZNFX-1 and WAGO-4 were completely separated from, but directly adjacent to, P granules.

P (Strome and Wood 1982) and Z granules (Ishidate *et al.* 2018; Wan *et al.* 2018) segregate with the germ cell lineage during embryonic cell divisions and are expressed throughout the adult germline. The observations that these germ granules are required for RNAi inheritance are enticing in that these biomolecular condensates seem perfectly poised to concentrate RNAs and/or proteins required for RNAi inheritance and transmit them to the next generation. Indeed, the finding that WAGO-4, an Argonaute, is required for the heritable expression of 22G-siRNAs during RNAi inheritance (Wan *et al.* 2018) opens up the intriguing possibility that WAGO-4 may bind and transmit siRNAs across generations. Of note, because P and Z granules colocalize during early embryonic cell stages and GLH-1 (Spike *et al.* 2008b; Updike and Strome 2009) and DEPS-1 (Spike *et al.* 2008a) are required for P granule formation, it is possible that P granules are indirectly required for RNAi inheritance because, in the absence of P granules, Z granules formation or function is impaired. Indeed, *deps-1* mutants also show abnormal Z granule morphology (Wan *et al.* 2018).

2. piRNA-induced TEI

piRNAs are another RNA trigger that can induce TEI in *C. elegans.* Worms can be injected to introduce single-copy transgenes at ectopic loci (Frøkjaer-Jensen *et al.* 2008). However, as was discussed above, sometimes, these transgenes get silenced and this silencing depends on PRG-1 and piRNAs (Ashe *et al.* 2012; Shirayama *et al.* 2012). It is thought that PRG-1 is able to identify and initiate silencing of foreign sequences based on sequence complementarity to piRNAs to which it is bound (Shirayama *et al.* 2012; Zhang *et al.* 2018). Interestingly, once PRG-1-dependent transgene silencing has been initiated, PRG-1 and piRNAs become dispensable for maintenance of this silencing. Indeed, piRNA-induced transgene silencing can be maintained in the absence of piRNAs for several generations, and sometimes even permanently (Ashe *et al.* 2012; Shirayama *et al.* 2012). Like gene silencing in response to exogenous dsRNA, piRNA-induced TEI requires WAGO proteins, such as HRDE-1, and is associated with the heritable expression of 22G-siRNAs. In addition, transgene silencing depends on histone methyltransferases, like SET-25, and is correlated with H3K9 methylation deposited on transgene loci (Shirayama *et al.* 2012) . Taken together, it is believed that once piRNAs initiate silencing of transgenes, the maintenance of silencing becomes dependent on a very similar mechanism as the mechanism that mediates RNAi inheritance in response to an exogenous dsRNA trigger.

3. Environmentally-induced TEI

Environmental cues, such as pathogens, starvation and chemical stress, can also trigger TEI in *C. elegans* (Remy 2010; Rechavi *et al.* 2014; Schott *et al.* 2014; Jobson *et al.* 2015; Moore *et al.* 2019; Posner *et al.* 2019). Interestingly, many of these multigenerational responses to environmental signals converge on the same set of factors that mediate dsRNA-induced TEI and piRNA-induced TEI. For instance, *C. elegans* shows pathogen avoidance to *Pseudomonas aeruginosa* (*P. aeruginosa*) after an initial exposure (Zhang *et al.* 2005) and a *P. aeruginosa* exposed hermaphrodite is able to transmit this avoidance behavior to its progeny (Moore *et al.*

2019). Avoidance behavior can subsequently be inherited for up to four generations by progeny who were never exposed to *P. aeruginosa* (Moore *et al.* 2019). Inheritance of this behavior requires HRDE-1, the RdRP RRF-1 and the histone methyltransferase SET-25, suggesting that nuclear RNAi mediates the inheritance of *P. aeruginosa* avoidance behavior. Interestingly, inheritance of avoidance behavior also requires PRG-1 and piRNAs, although it is unclear what role the piRNA-pathway plays in this example of TEI.

C. Regulators of TEI in C. elegans

RNAi inheritance in *C. elegans* is a finite phenomenon, lasting for a few generations, but not forever. This is somewhat surprising as RNAi inheritance is mediated by the amplification of siRNAs, so why aren't siRNAs amplified indefinitely? Instead, siRNA and H3K9me3 levels decline with each generation during the RNAi inheritance (Ashe *et al.* 2012; Buckley *et al.* 2012; Lev *et al.* 2017). We and others have hypothesized that there may be mechanisms that prevent RNAi inheritance from lasting forever. This hypothesis implies that we should be able to isolate mutants in which the generational longevity of RNAi inheritance may be extended. Below I will summarize some findings that show that there are, in fact, mechanisms that limit the generational perdurdance of RNAi inheritance.

1. MET-2

MET-2 is one of several methyltransferases expressed in *C. elegans.* MET-2 is thought to add the first two methyl groups onto lysine 9 of histone H3 at most genomic loci bearing H3K9 methylation. SET-25 is another methyltransferase that is thought to add the third methyl group onto H3K9, after the activity of MET-2 (Towbin *et al.* 2012). *met-2* mutants showed enhanced RNAi inheritance, meaning that they are able to inherit exogenous dsRNA-induced gene silencing for more generations than are wild-type animals. The mechanism of enhanced RNAi inheritance in *met-2* mutants is thought to be indirect. *met-2* mutants show global reduction in endogenous

H3K9 methylation levels (Towbin *et al.* 2012) and global changes in siRNA levels (Lev *et al.* 2017). These changes in siRNAs levels are likely caused by the loss of siRNAs that have an interdependent relationship with H3K9 methylation via the nuclear RNAI pathway. Indeed, the levels of siRNAs normally bound to HRDE-1 decline in *met-2* mutants, which is thought to make HRDE-1 more available to silence mRNAs targeted by exogenous dsRNA, thereby allowing RNAi inheritance to extend for more generations than normal (Lev *et al.* 2017). Interestingly, H3K9me3 is still observed at exogenous dsRNA-targeted genes in the absence of MET-2. What protein mediates this H3K9 methylation was not explored, but SET-32 is a likely candidate as SET-32 has been shown to mediate H3K9me3 after RNAi (Spracklin *et al.* 2017).

2. HERI-1

Our lab employed forward genetics to look for *heritable enhancer of RNAi* (*heri*) mutations that extend the number of generations RNAi inheritance lasts at two different germline-expressed loci (Perales *et al.* 2018) that can normally be heritably silenced for 4-10 generations after dsRNA treatment (Ashe *et al.* 2012; Buckley *et al.* 2012; Spracklin *et al.* 2017). These two loci are: (1) oocyte maturation defective-1 (*oma-1)*, a gene that is important for oocyte maturation in *C. elegans* (Lin 2003); and (2) *pie-1::gfp::h2b,* a single-copy transgene that encodes a histone-GFP fusion protein in the germline of *C. elegans* (Ashe *et al.* 2012). More specifically, we employed animals that harbor *zu405,* a gain-of-function allele of *oma-1,* that results in embryonic arrest at 20°C. This embryonic arrest phenotype can be suppressed by feeding worms dsRNA targeting *oma-1*, which silences the toxic gain-of-function allele (Lin 2003). Since *oma-1* can be heritably silenced by dsRNA, *oma-1(zu405)* animals continue to lay viable progeny for several generations after RNAi (Buckley *et al.* 2012). Using this gain-of-function *oma-1* allele allowed for the selection of mutants that inherited *oma-1* RNAi, and, therefore, laid viable progeny for longer than wild-type animals. To ensure that we identified general regulators of RNAi inheritance and not simply

intragenic or extragenic suppressors of *oma-1(zu405),* we then screened these mutants to look for animals that were also still silencing *pie-1::gfp::h2b* expression (Perales *et al.* 2018)..

Our screen identified twenty independent mutants in which dsRNA-mediated heritable silencing of both *oma-1* and *pie-1::gfp::h2b* lasted seven or more generations longer than in non-mutagenized controls (Perales *et al.* 2018). Four of these mutants harbored mutations in the gene *heri-1*, which encodes a chromodomain-containing pseudokinase. Indeed, *heri-1* mutants showed *oma-1* and *gfp* RNAi inheritance for 20 generations longer than wild-type animals, indicating that the HERI-1 protein normally limits the generational perdurance of RNAi inheritance. The enhanced RNAi inheritance observed in *heri* mutants, like typical RNAi inheritance (Ashe *et al.* 2012; Buckley *et al.* 2012), was associated with elevated levels of antisense 22G-siRNAs and persistent H3K9me3 at dsRNA-targeted loci, suggesting that the normal mechanism that mediates RNAi inheritance lasts longer in *heri-1* mutants. The HERI-1 protein was shown to be directly recruited to genes undergoing RNAi inheritance by HRDE-1 and the nuclear RNAi machinery, suggesting that, unlike MET-2, HERI-1 may directly regulate dsRNA-directed TEI (Perales *et al.* 2018). Interestingly, this data indicated that nuclear RNAi, the mechanism of RNAi inheritance in *C. elegans*, normally recruits its own regulator to limit its own activity (Perales *et al.* 2018). As HERI-1 harbors a chromodomain, HERI-1 could be recruited to loci that are targeted by nuclear RNAi via its interaction with methylated histones. Taken together, these data suggest that there are active mechanisms that limit the generational duration of RNAi inheritance in *C. elegans*.

Chapter 3 of this dissertation will focus on my work characterizing two other *heri* mutants identified from our screen. These mutants were unique in that, five years after they were treated with dsRNA, they still have not restored expression of *oma-1* and *pie-1::gfp::h2b.*

IV. References

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

Alberti C., R. A. Manzenreither, I. Sowemimo, T. R. Burkard, J. Wang, *et al.*, 2018 Cell-type specific sequencing of microRNAs from complex animal tissues. Nat. Methods 15: 283–289.

Albuquerque B. F. M. de, M. J. Luteijn, R. J. Cordeiro Rodrigues, P. van Bergeijk, S. Waaijers, *et al.*, 2014 PID-1 is a novel factor that operates during 21U-RNA biogenesis in Caenorhabditis elegans. Genes Dev. 28: 683–688.

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

Allshire R. C., and K. Ekwall, 2015 Epigenetic Regulation of Chromatin States in Schizosaccharomyces pombe. Cold Spring Harb. Perspect. Biol. 7: a018770.

Almeida M. V., S. Dietz, S. Redl, E. Karaulanov, A. Hildebrandt, *et al.*, 2018 GTSF-1 is required for formation of a functional RNA-dependent RNA Polymerase complex in Caenorhabditis elegans. EMBO J. 37. https://doi.org/10.15252/embj.201899325

Almeida M. V., M. A. Andrade-Navarro, and R. F. Ketting, 2019 Function and Evolution of Nematode RNAi Pathways. Noncoding RNA 5. https://doi.org/10.3390/ncrna5010008

Ambros V., and H. R. Horvitz, 1984 Heterochronic mutants of the nematode Caenorhabditis elegans. Science 226: 409–416.

Ambros V., and H. R. Horvitz, 1987 The lin-14 locus of Caenorhabditis elegans controls the time of expression of specific postembryonic developmental events. Genes Dev. 1: 398–414.

Ambros V., 1989 A hierarchy of regulatory genes controls a larva-to-adult developmental switch in C. elegans. Cell 57: 49–57.

Ambros V., and G. Ruvkun, 2018 Recent Molecular Genetic Explorations of Caenorhabditis elegans MicroRNAs. Genetics 209: 651–673.

Anderson D. M., K. M. Anderson, C.-L. Chang, C. A. Makarewich, B. R. Nelson, *et al.*, 2015 A micropeptide encoded by a putative long noncoding RNA regulates muscle performance. Cell 160: 595–606.

Angell S. M., and D. C. Baulcombe, 1997 Consistent gene silencing in transgenic plants expressing a replicating potato virus X RNA. EMBO J. 16: 3675–3684.

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

Aravin A. A., N. M. Naumova, A. V. Tulin, V. V. Vagin, Y. M. Rozovsky, *et al.*, 2001 Double-stranded RNA-mediated silencing of genomic tandem repeats and transposable elements in the D. melanogaster germline. Current Biology 11: 1017–1027.

Aravin A. A., M. S. Klenov, V. V. Vagin, F. Bantignies, G. Cavalli, *et al.*, 2004 Dissection of a Natural RNA Silencing Process in the Drosophila melanogaster Germ Line. Molecular and Cellular Biology 24: 6742–6750.

Aravin A., D. Gaidatzis, S. Pfeffer, M. Lagos-Quintana, P. Landgraf, *et al.*, 2006 A novel class of small RNAs bind to MILI protein in mouse testes. Nature 442: 203–207.

Aravin A. A., R. Sachidanandam, D. Bourc'his, C. Schaefer, D. Pezic, *et al.*, 2008 A piRNA pathway primed by individual transposons is linked to de novo DNA methylation in mice.

Mol. Cell 31: 785–799.

Aravind L., and E. V. Koonin, 1999 DNA polymerase beta-like nucleotidyltransferase superfamily: identification of three new families, classification and evolutionary history. Nucleic Acids Res. 27: 1609–1618.

Arteaga-Vazquez M. A., and V. L. Chandler, 2010 Paramutation in maize: RNA mediated trans-generational gene silencing. Curr. Opin. Genet. Dev. 20: 156–163.

Ashe A., A. Sapetschnig, E.-M. Weick, J. Mitchell, M. P. Bagijn, *et al.*, 2012 piRNAs can trigger a multigenerational epigenetic memory in the germline of C. elegans. Cell 150: 88–99.

Bagijn M. P., L. D. Goldstein, A. Sapetschnig, E.-M. Weick, S. Bouasker, *et al.*, 2012 Function, targets, and evolution of Caenorhabditis elegans piRNAs. Science 337: 574–578.

Banani S. F., H. O. Lee, A. A. Hyman, and M. K. Rosen, 2017 Biomolecular condensates: organizers of cellular biochemistry. Nat. Rev. Mol. Cell Biol. 18: 285–298.

Bartel D. P., 2004 MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 116: 281–297.

Bartel D. P., 2009 MicroRNAs: target recognition and regulatory functions. Cell 136: 215–233.

Barucci G., E. Cornes, M. Singh, B. Li, M. Ugolini, *et al.*, 2020 Small-RNA-mediated transgenerational silencing of histone genes impairs fertility in piRNA mutants. Nat. Cell Biol. 22: 235–245.

Batista P. J., J. G. Ruby, J. M. Claycomb, R. Chiang, N. Fahlgren, *et al.*, 2008 PRG-1 and 21U-RNAs interact to form the piRNA complex required for fertility in C. elegans. Mol. Cell 31: 67–78.

Bayne E. H., S. A. White, A. Kagansky, D. A. Bijos, L. Sanchez-Pulido, *et al.*, 2010 Stc1: A Critical Link between RNAi and Chromatin Modification Required for Heterochromatin Integrity. Cell 140: 666–677.

Behm-Ansmant I., J. Rehwinkel, T. Doerks, A. Stark, P. Bork, *et al.*, 2006 mRNA degradation by miRNAs and GW182 requires both CCR4:NOT deadenylase and DCP1:DCP2 decapping complexes. Genes Dev. 20: 1885–1898.

Bernstein E., A. A. Caudy, S. M. Hammond, and G. J. Hannon, 2001 Role for a bidentate ribonuclease in the initiation step of RNA interference. Nature 409: 363–366.

Bessereau J.-L., 2006 Transposons in C. elegans. WormBook 1–13.

Billi A. C., M. A. Freeberg, A. M. Day, S. Y. Chun, V. Khivansara, *et al.*, 2013 A conserved upstream motif orchestrates autonomous, germline-enriched expression of Caenorhabditis elegans piRNAs. PLoS Genet. 9: e1003392.

Bokar J. A., M. E. Rath-Shambaugh, R. Ludwiczak, P. Narayan, and F. Rottman, 1994 Characterization and partial purification of mRNA N6-adenosine methyltransferase from HeLa cell nuclei. Internal mRNA methylation requires a multisubunit complex. J. Biol. Chem. 269: 17697–17704.

Bokar J. A., M. E. Shambaugh, D. Polayes, A. G. Matera, and F. M. Rottman, 1997 Purification and cDNA cloning of the AdoMet-binding subunit of the human mRNA (N6-adenosine)-methyltransferase. RNA 3: 1233–1247.

Bošković A., and O. J. Rando, 2018 Transgenerational Epigenetic Inheritance. Annu. Rev. Genet. 52: 21–41.

Bourque G., K. H. Burns, M. Gehring, V. Gorbunova, A. Seluanov, *et al.*, 2018 Ten things you

should know about transposable elements. Genome Biol. 19: 199.

Bralley P., S. A. Chang, and G. H. Jones, 2005 A phylogeny of bacterial RNA nucleotidyltransferases: Bacillus halodurans contains two tRNA nucleotidyltransferases. J. Bacteriol. 187: 5927–5936.

Brangwynne C. P., C. R. Eckmann, D. S. Courson, A. Rybarska, C. Hoege, *et al.*, 2009 Germline P granules are liquid droplets that localize by controlled dissolution/condensation. Science 324: 1729–1732.

Braun J. E., E. Huntzinger, M. Fauser, and E. Izaurralde, 2011 GW182 proteins directly recruit cytoplasmic deadenylase complexes to miRNA targets. Mol. Cell 44: 120–133.

Braun J. E., V. Truffault, A. Boland, E. Huntzinger, C.-T. Chang, *et al.*, 2012 A direct interaction between DCP1 and XRN1 couples mRNA decapping to 5′ exonucleolytic degradation. Nat. Struct. Mol. Biol. 19: 1324–1331.

Brennecke J., A. A. Aravin, A. Stark, M. Dus, M. Kellis, *et al.*, 2007 Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. Cell 128: 1089–1103.

Brink R. A., 1956 A Genetic Change Associated with the R Locus in Maize Which Is Directed and Potentially Reversible. Genetics 41: 872–889.

Brockdorff N., A. Ashworth, G. F. Kay, V. M. McCabe, D. P. Norris, *et al.*, 1992 The product of the mouse Xist gene is a 15 kb inactive X-specific transcript containing no conserved ORF and located in the nucleus. Cell 71: 515–526.

Brown C. J., B. D. Hendrich, J. L. Rupert, R. G. Lafrenière, Y. Xing, *et al.*, 1992 The human XIST gene: analysis of a 17 kb inactive X-specific RNA that contains conserved repeats

and is highly localized within the nucleus. Cell 71: 527–542.

Buckley B. A., K. B. Burkhart, S. G. Gu, G. Spracklin, A. Kershner, *et al.*, 2012 A nuclear Argonaute promotes multigenerational epigenetic inheritance and germline immortality. Nature 489: 447–451.

Buker S. M., T. Iida, M. Bühler, J. Villén, S. P. Gygi, *et al.*, 2007 Two different Argonaute complexes are required for siRNA generation and heterochromatin assembly in fission yeast. Nature Structural & Molecular Biology 14: 200–207.

Burkhart K. B., S. Guang, B. A. Buckley, L. Wong, A. F. Bochner, *et al.*, 2011 A Pre-mRNA–Associating Factor Links Endogenous siRNAs to Chromatin Regulation. PLoS Genetics 7: e1002249.

Burton N. O., K. B. Burkhart, and S. Kennedy, 2011 Nuclear RNAi maintains heritable gene silencing in Caenorhabditis elegans. Proceedings of the National Academy of Sciences 108: 19683–19688.

Cai X., C. H. Hagedorn, and B. R. Cullen, 2004 Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. RNA 10: 1957–1966.

Carmell M. A., A. Girard, H. J. G. van de Kant, D. Bourc'his, T. H. Bestor, *et al.*, 2007 MIWI2 Is Essential for Spermatogenesis and Repression of Transposons in the Mouse Male Germline. Developmental Cell 12: 503–514.

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

Cech T. R., and J. A. Steitz, 2014 The Noncoding RNA Revolution—Trashing Old Rules to Forge New Ones. Cell 157: 77–94.

C. elegans Sequencing Consortium, 1998 Genome sequence of the nematode C. elegans: a platform for investigating biology. Science 282: 2012–2018.

Cernilogar F. M., M. C. Onorati, G. O. Kothe, A. M. Burroughs, K. M. Parsi, *et al.*, 2011 Chromatin-associated RNA interference components contribute to transcriptional regulation in Drosophila. Nature 480: 391–395.

Chalfie M., H. R. Horvitz, and J. E. Sulston, 1981 Mutations that lead to reiterations in the cell lineages of C. elegans. Cell 24: 59–69.

Charlesworth A., H. A. Meijer, and C. H. de Moor, 2013 Specificity factors in cytoplasmic polyadenylation. Wiley Interdiscip. Rev. RNA 4: 437–461.

Chen C.-C. G., M. J. Simard, H. Tabara, D. R. Brownell, J. A. McCollough, *et al.*, 2005 A member of the polymerase beta nucleotidyltransferase superfamily is required for RNA interference in C. elegans. Curr. Biol. 15: 378–383.

Chen C.-Y. A., D. Zheng, Z. Xia, and A.-B. Shyu, 2009 Ago-TNRC6 triggers microRNA-mediated decay by promoting two deadenylation steps. Nat. Struct. Mol. Biol. 16: 1160–1166.

Chen Q., M. Yan, Z. Cao, X. Li, Y. Zhang, *et al.*, 2016 Sperm tsRNAs contribute to intergenerational inheritance of an acquired metabolic disorder. Science 351: 397–400.

Chu C., K. Qu, F. L. Zhong, S. E. Artandi, and H. Y. Chang, 2011 Genomic maps of long noncoding RNA occupancy reveal principles of RNA-chromatin interactions. Mol. Cell 44: 667–678.

Claycomb J. M., P. J. Batista, K. M. Pang, W. Gu, J. J. Vasale, *et al.*, 2009 The Argonaute CSR-1 and its 22G-RNA cofactors are required for holocentric chromosome segregation. Cell

139: 123–134.

Coe E. H., 1966 The properties, origin, and mechanism of conversion-type inheritance at the B locus in maize. Genetics 53: 1035–1063.

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

Collins J., E. Forbes, and P. Anderson, 1989 The Tc3 family of transposable genetic elements in Caenorhabditis elegans. Genetics 121: 47–55.

Collins J. J., and P. Anderson, 1994 The Tc5 family of transposable elements in Caenorhabditis elegans. Genetics 137: 771–781.

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

Conine C. C., J. J. Moresco, W. Gu, M. Shirayama, D. Conte Jr, *et al.*, 2013 Argonautes promote male fertility and provide a paternal memory of germline gene expression in C. elegans. Cell 155: 1532–1544.

Cox D. N., A. Chao, J. Baker, L. Chang, D. Qiao, *et al.*, 1998 A novel class of evolutionarily conserved genes defined by piwi are essential for stem cell self-renewal. Genes & Development 12: 3715–3727.

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

Deans C., and K. A. Maggert, 2015 What Do You Mean, “Epigenetic”? Genetics 199: 887–896.

Deng W., and H. Lin, 2002 miwi, a murine homolog of piwi, encodes a cytoplasmic protein essential for spermatogenesis. Dev. Cell 2: 819–830.

Deniz Ö., J. M. Frost, and M. R. Branco, 2019 Regulation of transposable elements by DNA modifications. Nature Reviews Genetics.

Denli A. M., B. B. J. Tops, R. H. A. Plasterk, R. F. Ketting, and G. J. Hannon, 2004 Processing of primary microRNAs by the Microprocessor complex. Nature 432: 231–235.

Di Giacomo M., S. Comazzetto, H. Saini, S. De Fazio, C. Carrieri, *et al.*, 2013 Multiple epigenetic mechanisms and the piRNA pathway enforce LINE1 silencing during adult spermatogenesis. Mol. Cell 50: 601–608.

Djebali S., C. A. Davis, A. Merkel, A. Dobin, T. Lassmann, *et al.*, 2012 Landscape of transcription in human cells. Nature 489: 101–108.

Dodson A. E., and S. Kennedy, 2019 Germ Granules Coordinate RNA-Based Epigenetic Inheritance Pathways. Dev. Cell 50: 704–715.e4.

Dominissini D., S. Moshitch-Moshkovitz, S. Schwartz, M. Salmon-Divon, L. Ungar, *et al.*, 2012 Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. Nature 485: 201–206.

Dreyfus M., and P. Régnier, 2002 The Poly(A) Tail of mRNAs. Cell 111: 611–613.

Drinnenberg I. A., D. E. Weinberg, K. T. Xie, J. P. Mower, K. H. Wolfe, *et al.*, 2009 RNAi in budding yeast. Science 326: 544–550.

Duchaine T. F., J. A. Wohlschlegel, S. Kennedy, Y. Bei, D. Conte Jr, *et al.*, 2006 Functional proteomics reveals the biochemical niche of C. elegans DCR-1 in multiple small-RNA-mediated pathways. Cell 124: 343–354.

Duempelmann L., M. Skribbe, and M. Bühler, 2020 Small RNAs in the Transgenerational Inheritance of Epigenetic Information. Trends Genet. 36: 203–214.

Dutta T., A. Malhotra, and Murray P. Deutscher, 2013 How a CCA Sequence Protects Mature tRNAs and tRNA Precursors from Action of the Processing Enzyme RNase BN/RNase Z. Journal of Biological Chemistry 288: 30636–30644.

Eide D., and P. Anderson, 1985a The gene structures of spontaneous mutations affecting a Caenorhabditis elegans myosin heavy chain gene. Genetics 109: 67–79.

Eide D., and P. Anderson, 1985b Transposition of Tc1 in the nematode Caenorhabditis elegans. Proc. Natl. Acad. Sci. U. S. A. 82: 1756–1760.

Elbashir S. M., J. Harborth, W. Lendeckel, A. Yalcin, K. Weber, *et al.*, 2001 Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. Nature 411: 494–498.

Ellwanger D. C., F. A. Büttner, H.-W. Mewes, and V. Stümpflen, 2011 The sufficient minimal set of miRNA seed types. Bioinformatics 27: 1346–1350.

Emmons S. W., L. Yesner, K.-S. Ruan, and D. Katzenberg, 1983 Evidence for a transposon in caenorhabditis elegans. Cell 32: 55–65.

Ernst C., D. T. Odom, and C. Kutter, 2017 The emergence of piRNAs against transposon invasion to preserve mammalian genome integrity. Nature Communications 8.

Fabian M. R., M. K. Cieplak, F. Frank, M. Morita, J. Green, *et al.*, 2011 miRNA-mediated deadenylation is orchestrated by GW182 through two conserved motifs that interact with CCR4–NOT. Nature Structural & Molecular Biology 18: 1211–1217.

Fedoroff N., S. Wessler, and M. Shure, 1983 Isolation of the transposable maize controlling

elements Ac and Ds. Cell 35: 235–242.

Filipowicz W., 2005 RNAi: the nuts and bolts of the RISC machine. Cell 122: 17–20.

Fire A., D. Albertson, S. W. Harrison, and D. G. Moerman, 1991 Production of antisense RNA leads to effective and specific inhibition of gene expression in C. elegans muscle. Development 113: 503–514.

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

Fischer S. E. J., T. A. Montgomery, C. Zhang, N. Fahlgren, P. C. Breen, *et al.*, 2011 The ERI-6/7 helicase acts at the first stage of an siRNA amplification pathway that targets recent gene duplications. PLoS Genet. 7: e1002369.

Frøkjaer-Jensen C., M. W. Davis, C. E. Hopkins, B. J. Newman, J. M. Thummel, *et al.*, 2008 Single-copy insertion of transgenes in Caenorhabditis elegans. Nat. Genet. 40: 1375–1383.

Fromm B., D. Domanska, E. Høye, V. Ovchinnikov, W. Kang, *et al.*, 2020 MirGeneDB 2.0: the metazoan microRNA complement. Nucleic Acids Res. 48: D1172.

Fukao A., Y. Mishima, N. Takizawa, S. Oka, H. Imataka, *et al.*, 2014 MicroRNAs trigger dissociation of eIF4AI and eIF4AII from target mRNAs in humans. Mol. Cell 56: 79–89.

Fukaya T., H.-O. Iwakawa, and Y. Tomari, 2014 MicroRNAs block assembly of eIF4F translation initiation complex in Drosophila. Mol. Cell 56: 67–78.

Gent J. I., A. T. Lamm, D. M. Pavelec, J. M. Maniar, P. Parameswaran, *et al.*, 2010 Distinct phases of siRNA synthesis in an endogenous RNAi pathway in C. elegans soma. Mol. Cell 37: 679–689.

Gerson-Gurwitz A., S. Wang, S. Sathe, R. Green, G. W. Yeo, *et al.*, 2016 A Small RNA-Catalytic Argonaute Pathway Tunes Germline Transcript Levels to Ensure Embryonic Divisions. Cell 165: 396–409.

Geula S., S. Moshitch-Moshkovitz, D. Dominissini, A. A. Mansour, N. Kol, *et al.*, 2015 Stem cells. m6A mRNA methylation facilitates resolution of naïve pluripotency toward differentiation. Science 347: 1002–1006.

Gibney E. R., and C. M. Nolan, 2010 Epigenetics and gene expression. Heredity 105: 4–13.

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

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

Gou L.-T., P. Dai, J.-H. Yang, Y. Xue, Y.-P. Hu, *et al.*, 2015 Pachytene piRNAs instruct massive mRNA elimination during late spermiogenesis. Cell Res. 25: 266.

Greenberg M. V. C., and D. Bourc'his, 2019 The diverse roles of DNA methylation in mammalian development and disease. Nat. Rev. Mol. Cell Biol. 20: 590–607.

Grimson A., K. K.-H. Farh, W. K. Johnston, P. Garrett-Engele, L. P. Lim, *et al.*, 2007 MicroRNA targeting specificity in mammals: determinants beyond seed pairing. Mol. Cell 27: 91–105.

Grishok A., A. E. Pasquinelli, D. Conte, N. Li, S. Parrish, *et al.*, 2001 Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control C. elegans developmental timing. Cell 106: 23–34.

Grivna S. T., E. Beyret, Z. Wang, and H. Lin, 2006 A novel class of small RNAs in mouse

spermatogenic cells. Genes & Development 20: 1709–1714.

Gruidl M. E., P. A. Smith, K. A. Kuznicki, J. S. McCrone, J. Kirchner, *et al.*, 1996 Multiple potential germ-line helicases are components of the germ-line-specific P granules of Caenorhabditis elegans. Proc. Natl. Acad. Sci. U. S. A. 93: 13837–13842.

Gu W., M. Shirayama, D. Conte Jr, J. Vasale, P. J. Batista, *et al.*, 2009 Distinct argonaute-mediated 22G-RNA pathways direct genome surveillance in the C. elegans germline. Mol. Cell 36: 231–244.

Gu W., H.-C. Lee, D. Chaves, E. M. Youngman, G. J. Pazour, *et al.*, 2012 CapSeq and CIP-TAP identify Pol II start sites and reveal capped small RNAs as C. elegans piRNA precursors. Cell 151: 1488–1500.

Guang S., A. F. Bochner, D. M. Pavelec, K. B. Burkhart, S. Harding, *et al.*, 2008 An Argonaute transports siRNAs from the cytoplasm to the nucleus. Science 321: 537–541.

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

Gunawardane L. S., K. Saito, K. M. Nishida, K. Miyoshi, Y. Kawamura, *et al.*, 2007 A slicer-mediated mechanism for repeat-associated siRNA 5’ end formation in Drosophila. Science 315: 1587–1590.

Guo S., and K. J. Kemphues, 1995 par-1, a gene required for establishing polarity in C. elegans embryos, encodes a putative Ser/Thr kinase that is asymmetrically distributed. Cell 81: 611–620.

Hagan J. P., E. Piskounova, and R. I. Gregory, 2009 Lin28 recruits the TUTase Zcchc11 to

inhibit let-7 maturation in mouse embryonic stem cells. Nat. Struct. Mol. Biol. 16: 1021–1025.

Hamilton A. J., and D. C. Baulcombe, 1999 A species of small antisense RNA in posttranscriptional gene silencing in plants. Science 286: 950–952.

Hammond S. M., E. Bernstein, D. Beach, and G. J. Hannon, 2000 An RNA-directed nuclease mediates post-transcriptional gene silencing in Drosophila cells. Nature 404: 293–296.

Hammond S. M., 2015 An overview of microRNAs. Adv. Drug Deliv. Rev. 87: 3–14.

Han J., Y. Lee, K.-H. Yeom, Y.-K. Kim, H. Jin, *et al.*, 2004 The Drosha-DGCR8 complex in primary microRNA processing. Genes Dev. 18: 3016–3027.

Harrow J., A. Frankish, J. M. Gonzalez, E. Tapanari, M. Diekhans, *et al.*, 2012 GENCODE: the reference human genome annotation for The ENCODE Project. Genome Res. 22: 1760–1774.

Harvey Z. H., A. K. Chakravarty, R. A. Futia, and D. F. Jarosz, 2020 A Prion Epigenetic Switch Establishes an Active Chromatin State. Cell 180: 928–940.e14.

Heard E., and R. A. Martienssen, 2014 Transgenerational epigenetic inheritance: myths and mechanisms. Cell 157: 95–109.

Heo I., C. Joo, Y.-K. Kim, M. Ha, M.-J. Yoon, *et al.*, 2009 TUT4 in concert with Lin28 suppresses microRNA biogenesis through pre-microRNA uridylation. Cell 138: 696–708.

Heo I., M. Ha, J. Lim, M.-J. Yoon, J.-E. Park, *et al.*, 2012 Mono-uridylation of pre-microRNA as a key step in the biogenesis of group II let-7 microRNAs. Cell 151: 521–532.

Hermant C., A. Boivin, L. Teysset, V. Delmarre, A. Asif-Laidin, *et al.*, 2015 Paramutation in

Drosophila Requires Both Nuclear and Cytoplasmic Actors of the piRNA Pathway and Induces Cis-spreading of piRNA Production. Genetics 201: 1381–1396.

Höck J., and G. Meister, 2008 The Argonaute protein family. Genome Biology 9: 210.

Holliday R., 1994 Epigenetics: an overview. Dev. Genet. 15: 453–457.

Holoch D., and D. Moazed, 2015a RNA-mediated epigenetic regulation of gene expression. Nat. Rev. Genet. 16: 71–84.

Holoch D., and D. Moazed, 2015b Small-RNA loading licenses Argonaute for assembly into a transcriptional silencing complex. Nat. Struct. Mol. Biol. 22: 328–335.

Horsthemke B., 2018 A critical view on transgenerational epigenetic inheritance in humans. Nat. Commun. 9: 2973.

Huang X. A., H. Yin, S. Sweeney, D. Raha, M. Snyder, *et al.*, 2013 A Major Epigenetic Programming Mechanism Guided by piRNAs. Developmental Cell 24: 502–516.

Hutvágner G., J. McLachlan, A. E. Pasquinelli, E. Bálint, T. Tuschl, *et al.*, 2001 A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. Science 293: 834–838.

Hyun K., J. Jeon, K. Park, and J. Kim, 2017 Writing, erasing and reading histone lysine methylations. Experimental & Molecular Medicine 49: e324–e324.

Ingelbrecht I., H. Van Houdt, M. Van Montagu, and A. Depicker, 1994 Posttranscriptional silencing of reporter transgenes in tobacco correlates with DNA methylation. Proceedings of the National Academy of Sciences 91: 10502–10506.

Ingolia N. T., L. F. Lareau, and J. S. Weissman, 2011 Ribosome profiling of mouse embryonic

stem cells reveals the complexity and dynamics of mammalian proteomes. Cell 147: 789–802.

Inouye M., 1988 Antisense RNA: its functions and applications in gene regulation — a review. Gene 72: 25–34.

Ishidate T., D. J. Durning, R. Sharma, E.-Z. Shen, H. Chen, *et al.*, 2018 ZNFX-1 Engages Argonaute Proteins to Promote the Stable Inheritance of Epigenetic States in C. Elegans

Itou D., Y. Shiromoto, Y. Shin-ya, C. Ishii, T. Nishimura, *et al.*, 2015 Induction of DNA Methylation by Artificial piRNA Production in Male Germ Cells. Current Biology 25: 901–906.

Jinek M., and J. A. Doudna, 2009 A three-dimensional view of the molecular machinery of RNA interference. Nature 457: 405–412.

Jo M. H., S. Shin, S.-R. Jung, E. Kim, J.-J. Song, *et al.*, 2015 Human Argonaute 2 Has Diverse Reaction Pathways on Target RNAs. Mol. Cell 59: 117–124.

Jobson M. A., J. M. Jordan, M. A. Sandrof, J. D. Hibshman, A. L. Lennox, *et al.*, 2015 Transgenerational Effects of Early Life Starvation on Growth, Reproduction, and Stress Resistance in Caenorhabditis elegans. Genetics 201: 201–212.

John Liu S., M. A. Horlbeck, S. W. Cho, H. S. Birk, M. Malatesta, *et al.*, 2017 CRISPRi-based genome-scale identification of functional long noncoding RNA loci in human cells. Science 355. https://doi.org/10.1126/science.aah7111

Jorgensen R. A., P. D. Cluster, J. English, Q. Que, and C. A. Napoli, 1996 Chalcone synthase cosuppression phenotypes in petunia flowers: comparison of sense vs. antisense constructs and single-copy vs. complex T-DNA sequences. Plant Mol. Biol. 31: 957–973.

Jose A. M., and C. P. Hunter, 2007 Transport of sequence-specific RNA interference information between cells. Annu. Rev. Genet. 41: 305–330.

Joung J., J. M. Engreitz, S. Konermann, O. O. Abudayyeh, V. K. Verdine, *et al.*, 2017 Genome-scale activation screen identifies a lncRNA locus regulating a gene neighbourhood. Nature 548: 343–346.

Kamminga L. M., J. C. van Wolfswinkel, M. J. Luteijn, L. J. T. Kaaij, M. P. Bagijn, *et al.*, 2012 Differential impact of the HEN1 homolog HENN-1 on 21U and 26G RNAs in the germline of Caenorhabditis elegans. PLoS Genet. 8: e1002702.

Kawamura Y., K. Saito, T. Kin, Y. Ono, K. Asai, *et al.*, 2008 Drosophila endogenous small RNAs bind to Argonaute 2 in somatic cells. Nature 453: 793–797.

Kazazian H. H., 2004 Mobile Elements: Drivers of Genome Evolution. Science 303: 1626–1632.

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

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

Ketting R. F., S. E. Fischer, E. Bernstein, T. Sijen, G. J. Hannon, *et al.*, 2001 Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in C. elegans. Genes Dev. 15: 2654–2659.

Ketting F. R., and L. Cochella, 2020 Concepts and functions of small RNA pathways in C. elegans

Kopp F., and J. T. Mendell, 2018 Functional Classification and Experimental Dissection of Long

Noncoding RNAs. Cell 172: 393–407.

Kouzarides T., 2007 Chromatin modifications and their function. Cell 128: 693–705.

Kuramochi-Miyagawa S., 2004 Mili, a mammalian member of piwi family gene, is essential for spermatogenesis. Development 131: 839–849.

Kwak J. E., and M. Wickens, 2007 A family of poly(U) polymerases. RNA 13: 860–867.

Landgraf P., M. Rusu, R. Sheridan, A. Sewer, N. Iovino, *et al.*, 2007 A mammalian microRNA expression atlas based on small RNA library sequencing. Cell 129: 1401–1414.

Lee R. C., R. L. Feinbaum, and V. Ambros, 1993 The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell 75: 843–854.

Lee Y., K. Jeon, J.-T. Lee, S. Kim, and V. N. Kim, 2002 MicroRNA maturation: stepwise processing and subcellular localization. EMBO J. 21: 4663–4670.

Lee J. T., 2009 Lessons from X-chromosome inactivation: long ncRNA as guides and tethers to the epigenome. Genes Dev. 23: 1831–1842.

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

Lehrbach N. J., J. Armisen, H. L. Lightfoot, K. J. Murfitt, A. Bugaut, *et al.*, 2009 LIN-28 and the poly(U) polymerase PUP-2 regulate let-7 microRNA processing in Caenorhabditis elegans. Nat. Struct. Mol. Biol. 16: 1016–1020.

Le Thomas A., A. K. Rogers, A. Webster, G. K. Marinov, S. E. Liao, *et al.*, 2013 Piwi induces piRNA-guided transcriptional silencing and establishment of a repressive chromatin state. Genes Dev. 27: 390–399.

Le Thomas A., E. Stuwe, S. Li, J. Du, G. Marinov, *et al.*, 2014 Transgenerationally inherited piRNAs trigger piRNA biogenesis by changing the chromatin of piRNA clusters and inducing precursor processing. Genes Dev. 28: 1667–1680.

Lev I., U. Seroussi, H. Gingold, R. Bril, S. Anava, *et al.*, 2017 MET-2-Dependent H3K9 Methylation Suppresses Transgenerational Small RNA Inheritance. Curr. Biol. 27: 1138–1147.

Li X. Z., C. K. Roy, X. Dong, E. Bolcun-Filas, J. Wang, *et al.*, 2013 An Ancient Transcription Factor Initiates the Burst of piRNA Production during Early Meiosis in Mouse Testes. Molecular Cell 50: 67–81.

Liao L. W., B. Rosenzweig, and D. Hirsh, 1983 Analysis of a transposable element in Caenorhabditis elegans. Proc. Natl. Acad. Sci. U. S. A. 80: 3585–3589.

Lim L. P., N. C. Lau, E. G. Weinstein, A. Abdelhakim, S. Yekta, *et al.*, 2003 The microRNAs of Caenorhabditis elegans. Genes Dev. 17: 991–1008.

Lin H., and A. C. Spradling, 1997 A novel group of pumilio mutations affects the asymmetric division of germline stem cells in the Drosophila ovary. Development 124: 2463–2476.

Lin R., 2003 A gain-of-function mutation in oma-1, a C. elegans gene required for oocyte maturation, results in delayed degradation of maternal proteins and embryonic lethality. Dev. Biol. 258: 226–239.

Liu J., M. A. Carmell, F. V. Rivas, C. G. Marsden, J. M. Thomson, *et al.*, 2004 Argonaute2 is the catalytic engine of mammalian RNAi. Science 305: 1437–1441.

Luo S., and J. Lu, 2017 Silencing of Transposable Elements by piRNAs in Drosophila: An Evolutionary Perspective. Genomics Proteomics Bioinformatics 15: 164–176.

Lutz C. S., and A. Moreira, 2011 Alternative mRNA polyadenylation in eukaryotes: an effective regulator of gene expression. Wiley Interdiscip. Rev. RNA 2: 22–31.

Malone C. D., and G. J. Hannon, 2009 Small RNAs as guardians of the genome. Cell 136: 656–668.

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

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

Martens J. A., L. Laprade, and F. Winston, 2004 Intergenic transcription is required to repress the Saccharomyces cerevisiae SER3 gene. Nature 429: 571–574.

Martens J. A., P.-Y. J. Wu, and F. Winston, 2005 Regulation of an intergenic transcript controls adjacent gene transcription in Saccharomyces cerevisiae. Genes Dev. 19: 2695–2704.

Martin G., and W. Keller, 2007 RNA-specific ribonucleotidyl transferases. RNA 13: 1834–1849.

Martinez N. J., M. C. Ow, J. S. Reece-Hoyes, M. I. Barrasa, V. R. Ambros, *et al.*, 2008 Genome-scale spatiotemporal analysis of Caenorhabditis elegans microRNA promoter activity. Genome Res. 18: 2005–2015.

McClintock B., 1950 The origin and behavior of mutable loci in maize. Proc. Natl. Acad. Sci. U. S. A. 36: 344–355.

McEwan D. L., A. S. Weisman, and C. P. Hunter, 2012 Uptake of extracellular double-stranded RNA by SID-2. Mol. Cell 47: 746–754.

Meijer H. A., Y. W. Kong, W. T. Lu, A. Wilczynska, R. V. Spriggs, *et al.*, 2013 Translational repression and eIF4A2 activity are critical for microRNA-mediated gene regulation. Science 340: 82–85.

Meijer H. A., E. M. Smith, and M. Bushell, 2014 Regulation of miRNA strand selection: follow the leader? Biochem. Soc. Trans. 42: 1135–1140.

Meister G., 2013 Argonaute proteins: functional insights and emerging roles. Nat. Rev. Genet. 14: 447–459.

Mello C. C., and D. Conte, 2004 Revealing the world of RNA interference. Nature 431: 338–342.

Meyer K. D., Y. Saletore, P. Zumbo, O. Elemento, C. E. Mason, *et al.*, 2012 Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. Cell 149: 1635–1646.

Moazed D., 2009 Small RNAs in transcriptional gene silencing and genome defence. Nature 457: 413–420.

Mohn F., G. Sienski, D. Handler, and J. Brennecke, 2014 The Rhino-Deadlock-Cutoff Complex Licenses Noncanonical Transcription of Dual-Strand piRNA Clusters in Drosophila. Cell 157: 1364–1379.

Molaro A., and H. S. Malik, 2016 Hide and seek: how chromatin-based pathways silence retroelements in the mammalian germline. Curr. Opin. Genet. Dev. 37: 51–58.

Montgomery T. A., Y.-S. Rim, C. Zhang, R. H. Dowen, C. M. Phillips, *et al.*, 2012 PIWI associated siRNAs and piRNAs specifically require the Caenorhabditis elegans HEN1 ortholog henn-1. PLoS Genet. 8: e1002616.

Moore R. S., R. Kaletsky, and C. T. Murphy, 2019 Piwi/PRG-1 Argonaute and TGF-β Mediate

Transgenerational Learned Pathogenic Avoidance. Cell 177: 1827–1841.e12.

Mullen T. E., and W. F. Marzluff, 2008 Degradation of histone mRNA requires oligouridylation followed by decapping and simultaneous degradation of the mRNA both 5′ to 3′ and 3′ to 5′. Genes Dev. 22: 50–65.

Neuenfeldt A., A. Just, H. Betat, and M. Mörl, 2008 Evolution of tRNA nucleotidyltransferases: a small deletion generated CC-adding enzymes. Proc. Natl. Acad. Sci. U. S. A. 105: 7953–7958.

Newman M. A., F. Ji, S. E. J. Fischer, A. Anselmo, R. I. Sadreyev, *et al.*, 2018 The surveillance of pre-mRNA splicing is an early step in C. elegans RNAi of endogenous genes. Genes Dev. 32: 670–681.

Norbury C. J., 2013 Cytoplasmic RNA: a case of the tail wagging the dog. Nat. Rev. Mol. Cell Biol. 14: 643–653.

Okamura K., S. Balla, R. Martin, N. Liu, and E. C. Lai, 2008a Two distinct mechanisms generate endogenous siRNAs from bidirectional transcription in Drosophila melanogaster. Nature Structural & Molecular Biology 15: 581–590.

Okamura K., W.-J. Chung, J. Graham Ruby, H. Guo, D. P. Bartel, *et al.*, 2008b The Drosophila hairpin RNA pathway generates endogenous short interfering RNAs. Nature 453: 803–806.

Ouyang J. P. T., A. Folkmann, L. Bernard, C.-Y. Lee, U. Seroussi, *et al.*, P granules protect RNA interference genes from silencing by piRNAs

Ozata D. M., I. Gainetdinov, A. Zoch, D. O'Carroll, and P. D. Zamore, 2019 PIWI-interacting RNAs: small RNAs with big functions. Nat. Rev. Genet. 20: 89–108.

Pak J., and A. Fire, 2007 Distinct Populations of Primary and Secondary Effectors During RNAi

in C. elegans. Science 315: 241–244.

Pan T., 2018 Modifications and functional genomics of human transfer RNA. Cell Res. 28: 395–404.

Parker G. S., D. M. Eckert, and B. L. Bass, 2006 RDE-4 preferentially binds long dsRNA and its dimerization is necessary for cleavage of dsRNA to siRNA. RNA 12: 807–818.

Parrish S., and A. Fire, 2001 Distinct roles for RDE-1 and RDE-4 during RNA interference in Caenorhabditis elegans. RNA 7: 1397–1402.

Pasquinelli A. E., B. J. Reinhart, F. Slack, M. Q. Martindale, M. I. Kuroda, *et al.*, 2000 Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. Nature 408: 86–89.

Penny G. D., G. F. Kay, S. A. Sheardown, S. Rastan, and N. Brockdorff, 1996 Requirement for Xist in X chromosome inactivation. Nature 379: 131–137.

Perales R., D. Pagano, G. Wan, B. D. Fields, A. L. Saltzman, *et al.*, 2018 Transgenerational Epigenetic Inheritance Is Negatively Regulated by the HERI-1 Chromodomain Protein. Genetics 210: 1287–1299.

Perez M. F., and B. Lehner, 2019 Intergenerational and transgenerational epigenetic inheritance in animals. Nat. Cell Biol. 21: 143–151.

Perry R. P., and D. E. Kelley, 1974 Existence of methylated messenger RNA in mouse L cells. Cell 1: 37–42.

Pezic D., S. A. Manakov, R. Sachidanandam, and A. A. Aravin, 2014 piRNA pathway targets active LINE1 elements to establish the repressive H3K9me3 mark in germ cells. Genes Dev. 28: 1410–1428.

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

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

Posner R., I. A. Toker, O. Antonova, E. Star, S. Anava, *et al.*, 2019 Neuronal Small RNAs Control Behavior Transgenerationally. Cell 177: 1814–1826.e15.

Preston M. A., D. F. Porter, F. Chen, N. Buter, C. P. Lapointe, *et al.*, 2019 Unbiased screen of RNA tailing activities reveals a poly(UG) polymerase. Nature Methods 16: 437–445.

Proudfoot N. J., 2011 Ending the message: poly(A) signals then and now. Genes & Development 25: 1770–1782.

Ptashne M., 2013 Epigenetics: core misconcept. Proc. Natl. Acad. Sci. U. S. A. 110: 7101–7103.

Que Q., and R. A. Jorgensen, 1998 Homology-based control of gene expression patterns in transgenic petunia flowers. Dev. Genet. 22: 100–109.

Raddatz G., P. M. Guzzardo, N. Olova, M. R. Fantappié, M. Rampp, *et al.*, 2013 Dnmt2-dependent methylomes lack defined DNA methylation patterns. Proc. Natl. Acad. Sci. U. S. A. 110: 8627–8631.

Rassoulzadegan M., V. Grandjean, P. Gounon, S. Vincent, I. Gillot, *et al.*, 2006 RNA-mediated non-mendelian inheritance of an epigenetic change in the mouse. Nature 441: 469–474.

Rebollo R., M. T. Romanish, and D. L. Mager, 2012 Transposable Elements: An Abundant and

Natural Source of Regulatory Sequences for Host Genes. Annual Review of Genetics 46: 21–42.

Rechavi O., L. Houri-Ze'evi, S. Anava, W. S. S. Goh, S. Y. Kerk, *et al.*, 2014 Starvation-induced transgenerational inheritance of small RNAs in C. elegans. Cell 158: 277–287.

Reed K. J., J. M. Svendsen, K. C. Brown, B. E. Montgomery, T. N. Marks, *et al.*, 2020 Widespread roles for piRNAs and WAGO-class siRNAs in shaping the germline transcriptome of Caenorhabditis elegans. Nucleic Acids Res. 48: 1811–1827.

Rehwinkel J., 2005 A crucial role for GW182 and the DCP1:DCP2 decapping complex in miRNA-mediated gene silencing. RNA 11: 1640–1647.

Reinhart B. J., F. J. Slack, M. Basson, A. E. Pasquinelli, J. C. Bettinger, *et al.*, 2000 The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. Nature 403: 901–906.

Remy J.-J., 2010 Stable inheritance of an acquired behavior in Caenorhabditis elegans. Curr. Biol. 20: R877–8.

Reuter M., S. Chuma, T. Tanaka, T. Franz, A. Stark, *et al.*, 2009 Loss of the Mili-interacting Tudor domain-containing protein-1 activates transposons and alters the Mili-associated small RNA profile. Nat. Struct. Mol. Biol. 16: 639–646.

Rinn J. L., M. Kertesz, J. K. Wang, S. L. Squazzo, X. Xu, *et al.*, 2007 Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. Cell 129: 1311–1323.

Rissland O. S., A. Mikulasova, and C. J. Norbury, 2007 Efficient RNA polyuridylation by noncanonical poly(A) polymerases. Mol. Cell. Biol. 27: 3612–3624.

Robine N., N. C. Lau, S. Balla, Z. Jin, K. Okamura, *et al.*, 2009 A Broadly Conserved Pathway Generates 3′UTR-Directed Primary piRNAs. Current Biology 19: 2066–2076.

Rojas-Ríos P., and M. Simonelig, 2018 piRNAs and PIWI proteins: regulators of gene expression in development and stem cells. Development 145. https://doi.org/10.1242/dev.161786

Rouget C., C. Papin, A. Boureux, A.-C. Meunier, B. Franco, *et al.*, 2010 Maternal mRNA deadenylation and decay by the piRNA pathway in the early Drosophila embryo. Nature 467: 1128–1132.

Ruby J. G., C. Jan, C. Player, M. J. Axtell, W. Lee, *et al.*, 2006 Large-scale sequencing reveals 21U-RNAs and additional microRNAs and endogenous siRNAs in C. elegans. Cell 127: 1193–1207.

Ruvkun G., and J. Giusto, 1989 The Caenorhabditis elegans heterochronic gene lin-14 encodes a nuclear protein that forms a temporal developmental switch. Nature 338: 313–319.

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

Sarot E., G. Payen-Groschêne, A. Bucheton, and A. Pélisson, 2004 Evidence for a piwi-Dependent RNA Silencing of the gypsy Endogenous Retrovirus by the Drosophila melanogaster flamenco Gene. Genetics 166: 1313–1321.

Schott D., I. Yanai, and C. P. Hunter, 2014 Natural RNA interference directs a heritable response to the environment. Sci. Rep. 4: 7387.

Schwartz S., S. D. Agarwala, M. R. Mumbach, M. Jovanovic, P. Mertins, *et al.*, 2013 High-resolution mapping reveals a conserved, widespread, dynamic mRNA methylation program

in yeast meiosis. Cell 155: 1409–1421.

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

Sharma U., C. C. Conine, J. M. Shea, A. Boskovic, A. G. Derr, *et al.*, 2016 Biogenesis and function of tRNA fragments during sperm maturation and fertilization in mammals. Science 351: 391–396.

Shen E.-Z., H. Chen, A. R. Ozturk, S. Tu, M. Shirayama, *et al.*, Identification of piRNA binding sites reveals the Argonaute regulatory landscape of the C. elegans germline

Shin Y., and C. P. Brangwynne, 2017 Liquid phase condensation in cell physiology and disease. Science 357. https://doi.org/10.1126/science.aaf4382

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

Shoji M., T. Tanaka, M. Hosokawa, M. Reuter, A. Stark, *et al.*, 2009 The TDRD9-MIWI2 Complex Is Essential for piRNA-Mediated Retrotransposon Silencing in the Mouse Male Germline. Developmental Cell 17: 775–787.

Shorter J., and S. Lindquist, 2005 Prions as adaptive conduits of memory and inheritance. Nat. Rev. Genet. 6: 435–450.

Sienski G., D. Dönertas, and J. Brennecke, 2012 Transcriptional Silencing of Transposons by Piwi and Maelstrom and Its Impact on Chromatin State and Gene Expression. Cell 151: 964–980.

Sijen T., and R. H. A. Plasterk, 2003 Transposon silencing in the Caenorhabditis elegans germ

line by natural RNAi. Nature 426: 310–314.

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

Simon B., J. P. Kirkpatrick, S. Eckhardt, M. Reuter, E. A. Rocha, *et al.*, 2011 Recognition of 2′-O-Methylated 3′-End of piRNA by the PAZ Domain of a Piwi Protein. Structure 19: 172–180.

Siomi M. C., K. Sato, D. Pezic, and A. A. Aravin, 2011 PIWI-interacting small RNAs: the vanguard of genome defence. Nat. Rev. Mol. Cell Biol. 12: 246–258.

Sloan K. E., A. S. Warda, S. Sharma, K.-D. Entian, D. L. J. Lafontaine, *et al.*, 2017 Tuning the ribosome: The influence of rRNA modification on eukaryotic ribosome biogenesis and function. RNA Biol. 14: 1138–1152.

Smith Z. D., and A. Meissner, 2013 DNA methylation: roles in mammalian development. Nature Reviews Genetics 14: 204–220.

Sommer S., U. Lavi, and J. E. Darnell Jr, 1978 The absolute frequency of labeled N-6-methyladenosine in HeLa cell messenger RNA decreases with label time. J. Mol. Biol. 124: 487–499.

Spike C. A., J. Bader, V. Reinke, and S. Strome, 2008a DEPS-1 promotes P-granule assembly and RNA interference in C. elegans germ cells. Development 135: 983–993.

Spike C., N. Meyer, E. Racen, A. Orsborn, J. Kirchner, *et al.*, 2008b Genetic analysis of the Caenorhabditis elegans GLH family of P-granule proteins. Genetics 178: 1973–1987.

Spracklin G., B. Fields, G. Wan, D. Becker, A. Wallig, *et al.*, 2017 The RNAi Inheritance Machinery of Caenorhabditis elegans. Genetics 206: 1403–1416.

Steiner F. A., K. L. Okihara, S. W. Hoogstrate, T. Sijen, and R. F. Ketting, 2009 RDE-1 slicer activity is required only for passenger-strand cleavage during RNAi in Caenorhabditis elegans. Nat. Struct. Mol. Biol. 16: 207–211.

Stillman B., 2018 Histone Modifications: Insights into Their Influence on Gene Expression. Cell 175: 6–9.

Strome S., and W. B. Wood, 1982 Immunofluorescence visualization of germ-line-specific cytoplasmic granules in embryos, larvae, and adults of Caenorhabditis elegans. Proc. Natl. Acad. Sci. U. S. A. 79: 1558–1562.

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

Suganuma T., and J. L. Workman, 2011 Signals and combinatorial functions of histone modifications. Annu. Rev. Biochem. 80: 473–499.

Tabara H., M. Sarkissian, W. G. Kelly, J. Fleenor, A. Grishok, *et al.*, 1999 The rde-1 gene, RNA interference, and transposon silencing in C. elegans. Cell 99: 123–132.

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

Takayama K. M., and M. Lnouye, 1990 Antisense RN. Crit. Rev. Biochem. Mol. Biol. 25: 155–184.

Tanaka T., M. Hosokawa, V. V. Vagin, M. Reuter, E. Hayashi, *et al.*, 2011 Tudor domain containing 7 (Tdrd7) is essential for dynamic ribonucleoprotein (RNP) remodeling of

chromatoid bodies during spermatogenesis. Proceedings of the National Academy of Sciences 108: 10579–10584.

Thebault P., G. Boutin, W. Bhat, A. Rufiange, J. Martens, *et al.*, 2011 Transcription regulation by the noncoding RNA SRG1 requires Spt2-dependent chromatin deposition in the wake of RNA polymerase II. Mol. Cell. Biol. 31: 1288–1300.

Thivierge C., N. Makil, M. Flamand, J. J. Vasale, C. C. Mello, *et al.*, 2012 Tudor domain ERI-5 tethers an RNA-dependent RNA polymerase to DCR-1 to potentiate endo-RNAi. Nature Structural & Molecular Biology 19: 90–97.

Tijsterman M., R. F. Ketting, K. L. Okihara, T. Sijen, and R. H. A. Plasterk, 2002 RNA helicase MUT-14-dependent gene silencing triggered in C. elegans by short antisense RNAs. Science 295: 694–697.

Tops B. B. J., H. Tabara, T. Sijen, F. Simmer, C. C. Mello, *et al.*, 2005 RDE-2 interacts with MUT-7 to mediate RNA interference in Caenorhabditis elegans. Nucleic Acids Res. 33: 347–355.

Tops B. B. J., R. H. A. Plasterk, and R. F. Ketting, 2006 The Caenorhabditis elegans Argonautes ALG-1 and ALG-2: almost identical yet different. Cold Spring Harb. Symp. Quant. Biol. 71: 189–194.

Towbin B. D., C. González-Aguilera, R. Sack, D. Gaidatzis, V. Kalck, *et al.*, 2012 Step-wise methylation of histone H3K9 positions heterochromatin at the nuclear periphery. Cell 150: 934–947.

Tsai M.-C., O. Manor, Y. Wan, N. Mosammaparast, J. K. Wang, *et al.*, 2010 Long noncoding RNA as modular scaffold of histone modification complexes. Science 329: 689–693.

Tsai H.-Y., C.-C. G. Chen, D. Conte Jr, J. J. Moresco, D. A. Chaves, *et al.*, 2015 A ribonuclease coordinates siRNA amplification and mRNA cleavage during RNAi. Cell 160: 407–419.

Uebel C. J., D. C. Anderson, L. M. Mandarino, K. I. Manage, S. Aynaszyan, *et al.*, 2018 Distinct regions of the intrinsically disordered protein MUT-16 mediate assembly of a small RNA amplification complex and promote phase separation of Mutator foci. PLOS Genetics 14: e1007542.

Ulitsky I., and D. P. Bartel, 2013 lincRNAs: genomics, evolution, and mechanisms. Cell 154: 26–46.

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

Vasale J. J., W. Gu, C. Thivierge, P. J. Batista, J. M. Claycomb, *et al.*, 2010 Sequential rounds of RNA-dependent RNA transcription drive endogenous small-RNA biogenesis in the ERGO-1/Argonaute pathway. Proc. Natl. Acad. Sci. U. S. A. 107: 3582–3587.

Vasileva A., D. Tiedau, A. Firooznia, T. Müller-Reichert, and R. Jessberger, 2009 Tdrd6 Is Required for Spermiogenesis, Chromatoid Body Architecture, and Regulation of miRNA Expression. Current Biology 19: 630–639.

Vastenhouw N. L., S. E. J. Fischer, V. J. P. Robert, K. L. Thijssen, A. G. Fraser, *et al.*, 2003 A genome-wide screen identifies 27 genes involved in transposon silencing in C. elegans. Curr. Biol. 13: 1311–1316.

Vastenhouw N. L., K. Brunschwig, K. L. Okihara, F. Müller, M. Tijsterman, *et al.*, 2006 Gene expression: long-term gene silencing by RNAi. Nature 442: 882.

Volpe T. A., C. Kidner, I. M. Hall, G. Teng, S. I. S. Grewal, *et al.*, 2002 Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. Science 297: 1833–1837.

Vörtler S., and M. Mörl, 2010 tRNA-nucleotidyltransferases: highly unusual RNA polymerases with vital functions. FEBS Lett. 584: 297–302.

Vourekas A., Q. Zheng, P. Alexiou, M. Maragkakis, Y. Kirino, *et al.*, 2012 Mili and Miwi target RNA repertoire reveals piRNA biogenesis and function of Miwi in spermiogenesis. Nature Structural & Molecular Biology 19: 773–781.

Vourekas A., P. Alexiou, N. Vrettos, M. Maragkakis, and Z. Mourelatos, 2016 Sequence-dependent but not sequence-specific piRNA adhesion traps mRNAs to the germ plasm. Nature 531: 390–394.

WADDINGTON, and C. H, 1942 The epigenotype. Endeavour 1: 18–20.

Walsh C. P., J. R. Chaillet, and T. H. Bestor, 1998 Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. Nat. Genet. 20: 116–117.

Wan G., B. D. Fields, G. Spracklin, A. Shukla, C. M. Phillips, *et al.*, 2018 Spatiotemporal regulation of liquid-like condensates in epigenetic inheritance. Nature 557: 679–683.

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

Wang X., Z. Lu, A. Gomez, G. C. Hon, Y. Yue, *et al.*, 2014a N6-methyladenosine-dependent regulation of messenger RNA stability. Nature 505: 117–120.

Wang J. T., J. Smith, B.-C. Chen, H. Schmidt, D. Rasoloson, *et al.*, 2014b Regulation of RNA granule dynamics by phosphorylation of serine-rich, intrinsically disordered proteins in C.

elegans. Elife 3: e04591.

Wang X., B. S. Zhao, I. A. Roundtree, Z. Lu, D. Han, *et al.*, 2015 N(6)-methyladenosine Modulates Messenger RNA Translation Efficiency. Cell 161: 1388–1399.

Warkocki Z., P. S. Krawczyk, D. Adamska, K. Bijata, J. L. Garcia-Perez, *et al.*, 2018 Uridylation by TUT4/7 Restricts Retrotransposition of Human LINE-1s. Cell 174: 1537–1548.e29.

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

Waterhouse P. M., M. W. Graham, and M. B. Wang, 1998 Virus resistance and gene silencing in plants can be induced by simultaneous expression of sense and antisense RNA. Proc. Natl. Acad. Sci. U. S. A. 95: 13959–13964.

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

Weick E.-M., P. Sarkies, N. Silva, R. A. Chen, S. M. M. Moss, *et al.*, 2014 PRDE-1 is a nuclear factor essential for the biogenesis of Ruby motif-dependent piRNAs in C. elegans. Genes Dev. 28: 783–796.

Weick E.-M., and E. A. Miska, 2014 piRNAs: from biogenesis to function. Development 141: 3458–3471.

Weill L., E. Belloc, F.-A. Bava, and R. Méndez, 2012 Translational control by changes in poly(A) tail length: recycling mRNAs. Nat. Struct. Mol. Biol. 19: 577–585.

Welker N. C., T. S. Maity, X. Ye, P. J. Aruscavage, A. A. Krauchuk, *et al.*, 2011 Dicer's helicase

domain discriminates dsRNA termini to promote an altered reaction mode. Mol. Cell 41: 589–599.

Wightman B., I. Ha, and G. Ruvkun, 1993 Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans. Cell 75: 855–862.

Winston W. M., C. Molodowitch, and C. P. Hunter, 2002 Systemic RNAi in C. elegans requires the putative transmembrane protein SID-1. Science 295: 2456–2459.

Winston W. M., M. Sutherlin, A. J. Wright, E. H. Feinberg, and C. P. Hunter, 2007 Caenorhabditis elegans SID-2 is required for environmental RNA interference. Proc. Natl. Acad. Sci. U. S. A. 104: 10565–10570.

Xiong Y., and T. A. Steitz, 2004 Mechanism of transfer RNA maturation by CCA-adding enzyme without using an oligonucleotide template. Nature 430: 640–645.

Yigit E., P. J. Batista, Y. Bei, K. M. Pang, C.-C. G. Chen, *et al.*, 2006 Analysis of the C. elegans Argonaute family reveals that distinct Argonautes act sequentially during RNAi. Cell 127: 747–757.

Yu S., and V. N. Kim, 2020 A tale of non-canonical tails: gene regulation by post-transcriptional RNA tailing. Nat. Rev. Mol. Cell Biol. 21: 542–556.

Zamore P. D., T. Tuschl, P. A. Sharp, and D. P. Bartel, 2000 RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. Cell 101: 25–33.

Zemach A., and D. Zilberman, 2010 Evolution of eukaryotic DNA methylation and the pursuit of safer sex. Curr. Biol. 20: R780–5.

Zhang H., F. A. Kolb, L. Jaskiewicz, E. Westhof, and W. Filipowicz, 2004 Single processing

center models for human Dicer and bacterial RNase III. Cell 118: 57–68.

Zhang Y., H. Lu, and C. I. Bargmann, 2005 Pathogenic bacteria induce aversive olfactory learning in Caenorhabditis elegans. Nature 438: 179–184.

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

Zhang C., T. A. Montgomery, S. E. J. Fischer, S. M. D. A. Garcia, C. G. Riedel, *et al.*, 2012 The Caenorhabditis elegans RDE-10/RDE-11 complex regulates RNAi by promoting secondary siRNA amplification. Curr. Biol. 22: 881–890.

Zhang D., S. Tu, M. Stubna, W.-S. Wu, W.-C. Huang, *et al.*, 2018 The piRNA targeting rules and the resistance to piRNA silencing in endogenous genes. Science 359: 587–592.

Zhao Y., H. Li, S. Fang, Y. Kang, W. Wu, *et al.*, 2016 NONCODE 2016: an informative and valuable data source of long non-coding RNAs. Nucleic Acids Res. 44: D203–8.

Zhao B. S., I. A. Roundtree, and C. He, 2017 Post-transcriptional gene regulation by mRNA modifications. Nat. Rev. Mol. Cell Biol. 18: 31.

Zhu S., W. Li, J. Liu, C.-H. Chen, Q. Liao, *et al.*, 2016 Genome-scale deletion screening of human long non-coding RNAs using a paired-guide RNA CRISPR-Cas9 library. Nat. Biotechnol. 34: 1279–1286.

Chapter 2

poly(UG)-tailed RNAs in Genome Defense and Transgenerational Epigenetic Inheritance

This chapter consists primarily of the published manuscript:

Shukla, A.*, Yan, J.*, Pagano, D.J., Dodson, A.E., Fei, Y., Gorham, J., Seidman, J.G., Wickens, M. & Kennedy, S. poly(UG)-tailed RNAs in genome protection and epigenetic inheritance. *Nature* 582, 283–288 (2020). * co-first authors

This manuscript was written by Scott Kennedy and Aditi Shukla. Experiments were performed by Aditi Shukla and Jenny Yan. All figures were made and assembled by Aditi Shukla. MiSeq and IP-MS/MS data analysis was performed by Anne E. Dodson. RNA-seq data analysis was performed by Daniel J. Pagano. Small RNA-seq analysis was performed by Yuhan Fei. Josh Gorham and Jon Seidman assisted with RNA-seq library preparation and sequencing. Marvin Wickens generously provided unpublished data and invaluable intellectual support.

I. Abstract

Mobile genetic elements threaten genome integrity in all organisms. MUT-2/RDE-3 is a ribonucleotidyltransferase required for transposon silencing and RNA interference (RNAi) in *C. elegans* (Collins *et al.* 1987; Ketting *et al.* 1999; Tabara *et al.* 1999; Chen *et al.* 2005). When tethered to RNAs in heterologous expression systems, RDE-3 can add long stretches of alternating non-templated uridine (U) and guanosine (G) ribonucleotides to the 3' termini of these RNAs (poly(UG) or pUG tails) (Preston *et al.* 2019). Here we show that, in its natural context in *C. elegans*, RDE-3 adds pUG tails to targets of RNAi, as well as to transposon RNAs. pUG tails with more than 16 perfectly alternating 3' U and G nucleotides convert RNA fragments into agents of gene silencing. pUG tails promote gene silencing by recruiting RNA-dependent RNA polymerases (RdRPs), which use pUG-tailed RNAs (pUG RNAs) as templates to synthesize small interfering RNAs (siRNAs). Our results show that cycles of pUG RNA–templated siRNA synthesis and siRNA-directed mRNA pUGylation underlie dsRNA-directed transgenerational epigenetic inheritance in the *C. elegans* germline. We speculate that this pUG RNA/siRNA silencing loop allows parents to inoculate progeny against the expression of unwanted or parasitic genetic elements.

II. Introduction

Transposable elements are mobile parasitic genetic elements present in all genomes. Transposons threaten genome integrity, and can cause disease by disrupting genes or inducing non-allelic recombination. RNA interference (RNAi) is a conserved gene silencing mechanism initiated by double-stranded RNA (dsRNA) (Fire *et al.* 1998). Forward genetic screens to identify factors required for either transposon silencing or RNAi have been conducted in the model metazoan *C. elegans* (Collins *et al.* 1987; Ketting *et al.* 1999; Tabara *et al.* 1999). These screens identified an overlapping set of genes, indicating that an RNAi-related process silences transposons (Collins *et al.* 1987; Ketting *et al.* 1999; Tabara *et al.* 1999). One gene required for both efficient transposon silencing and RNAi in *C. elegans* is *mut-2/rde-3*, which encodes a protein with homology to ribonucleotidyltransferases (rNTs) (Collins *et al.* 1987; Ketting *et al.* 1999; Tabara *et al.* 1999; Chen *et al.* 2005). rNTs add non-templated ribonucleotides to RNAs and other substrates (Aravind and Koonin 1999; Martin and Keller 2007). Recently, *C. elegans* MUT-2/RDE-3 (henceforth, RDE-3) was shown to add perfectly alternating U and G ribonucleotides to the 3' termini of RNAs (termed polyUG or pUG tails) to which it was tethered either in *S. cerevisiae* or in *X. laevis* oocytes (Preston *et al.* 2019). Taken together, these data prompted the proposal that RDE-3 may append non-templated pUG tails to the 3' termini of RNAs during transposon silencing and/or RNAi in *C. elegans* (Preston *et al.* 2019).

III. Results

A. RNAi directs RDE-3–dependent mRNA pUGylation

We first asked whether pUG tails are added to RNAs targeted by RNAi in *C. elegans*. We used an $(AC)_9$ oligo to reverse transcribe (RT) total RNA extracted from animals exposed to dsRNA targeting the germline-expressed gene *oma-1* (Detwiler *et al.* 2001), and then performed nested

PCR to try to detect *oma-1* RNAs modified with 3' pUG repeats (Figure 2.1a). This approach (termed pUG PCR) detected PCR products that were dependent on *oma-1* dsRNA (Figure 2.1b), as well as on components of the RNAi machinery including RDE-4, which promotes dsRNA processing into siRNAs (Parrish and Fire 2001; Tabara *et al.* 2002); the siRNA-binding Argonaute (AGO) protein RDE-1 (Parrish and Fire 2001; Tabara *et al.* 2002); and RDE-8, which is an endonuclease that cleaves mRNAs exhibiting homology to siRNAs (Tsai *et al.* 2015) (Figure 2.1c).

Sanger and Illumina sequencing revealed that most (>89%) pUG PCR products were derived from hybrid RNAs consisting of stretches of nearly perfectly alternating (error rate <2%, Table 1) U and G repeats appended to the 3' termini of sense and spliced *oma-1* mRNA fragments (Figure 2.1d, Figure 2.2a). pUGylation sites were non-randomly distributed along the *oma-1* mRNA (Figure 2.2). Critically, most (64%) pUG tails were longer (range=19-75nt) than the $(AC)_9$ oligo used for RT (Figure 2.1d), indicating that the detected pUG RNAs were not the result of priming from genomically encoded UG-rich sequences. RDE-3 was required for addition of pUG repeats to mRNA fragments (termed pUGylation): *rde-3* mutants, including *rde-3(ne3370)* animals, which harbor a deletion that removes residues required for catalysis within the rNT domain of RDE-3 (henceforth *rde-3(-)*) (Preston *et al.* 2019), failed to produce *oma-1* pUG RNAs in response to *oma-1* dsRNA (Figure 2.1b). pUGylation defects in *rde-3* mutants were rescued by introducing a wild-type copy of *rde-3* into *rde-3(-)* animals or by CRISPR/Cas9-mediated reversion of a missense allele (*ne298)* of *rde-3* to wild-type (Figure 2.1b). Furthermore, RNA pUGylation was a general response to RNAi: animals exposed to dsRNA targeting a germline-expressed *gfp::h2b* transgene or the hypodermis-expressed *dpy-11* gene (Ko and Chow 2002) produced RDE-3–dependent *gfp* and *dpy-11* pUG RNAs, respectively (Figure 2.3a,b). Finally, pUGylation was sequence-specific, since *dpy-11* dsRNA did not induce *oma-1* pUG RNA biogenesis and vice versa (Figure 2.3b). Together, these data indicate that RDE-3 adds pUG tails to mRNAs targeted for silencing by RNAi.

Figure 2.1. pUG tails are added to mRNA fragments *in vivo*. a, Assay to detect gene-specific pUG RNAs. Note: $(AC)_9$ RT oligo can anneal anywhere along the pUG tail. **b**, *oma-1* pUG PCR on total RNA isolated from animals of indicated genotypes, +/- *oma-1* dsRNA (RNAi). *rde-3* mutants were rescued as described in Results and Methods sections. *gsa-1*, which has an 18nt long genomically encoded pUG repeat in its 3'UTR, is a loading control. Wild-type (WT) vs. *rde-3(ne3370)* and WT vs. *rde-3(ne298)* data is representative of >10 and 2 biologically independent experiments, respectively. **c,** *oma-1* pUG PCR on total RNA from animals of indicated genotypes, +/- *oma-1* dsRNA. Data is representative of 3 biologically independent experiments. **d,** Sanger sequencing chromatogram (red=T, black=G, blue=C, green=A) of *oma-1* pUG PCR product.

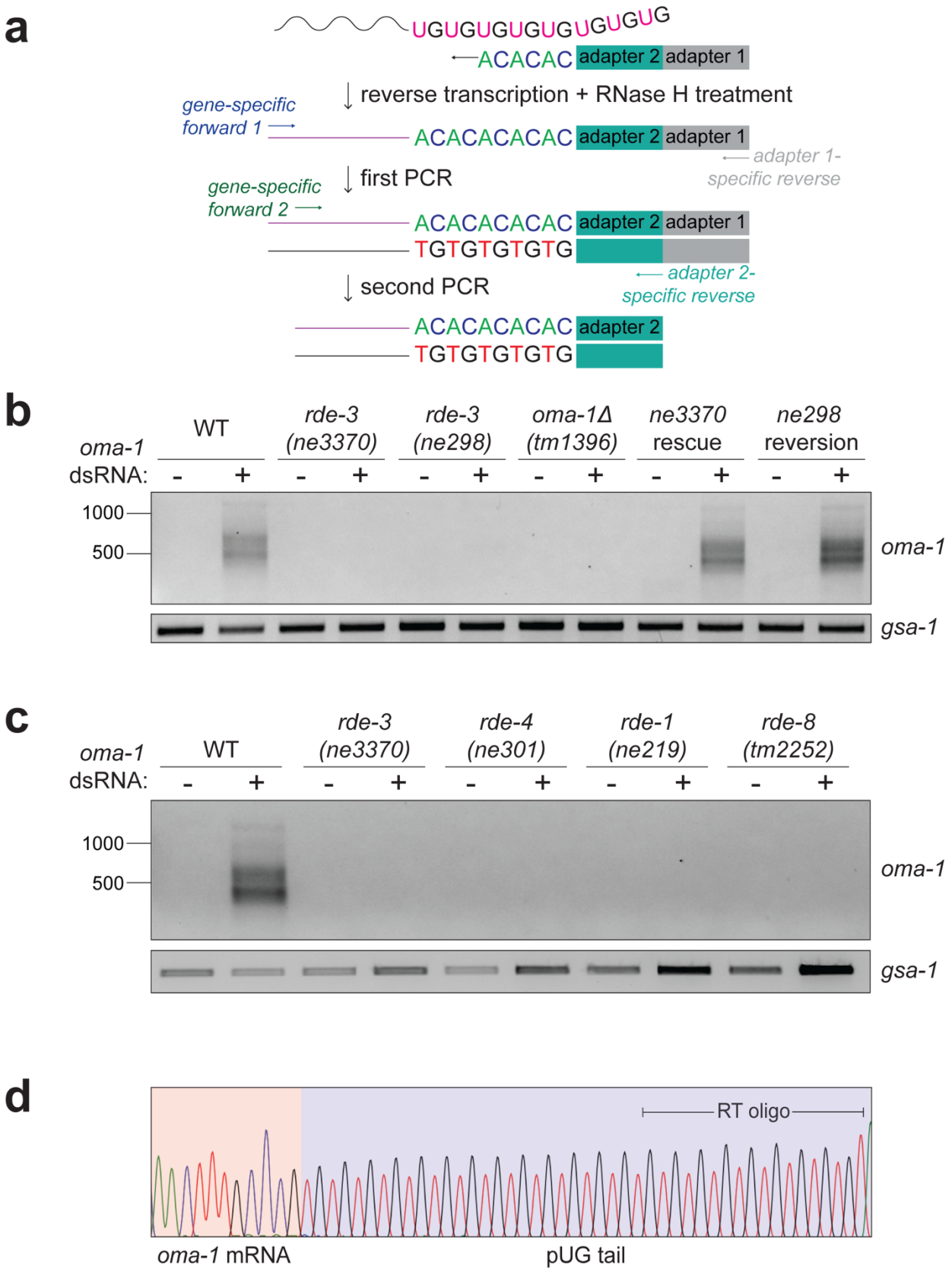

a
UGUGUGUGUGUGUGUG
ACACAC
adapter 2
adapter 1
reverse transcription + RNase H treatment
gene-specific forward 1
ACACACACAC
adapter 1-specific reverse
first PCR
gene-specific forward 2
TGTGTGTGTG
adapter 2-specific reverse
second PCR
b
oma-1
WT
rde-3 (ne3370)
rde-3 (ne298)
oma-1Δ (tm1396)
ne3370 rescue
ne298 reversion
dsRNA:
1000
500
oma-1
gsa-1
c
rde-4 (ne301)
rde-1 (ne219)
rde-8 (tm2252)
d
RT oligo
oma-1 mRNA
pUG tail

Table 2.1. RDE-3 fidelity calculations. >18nt long pUG tails were randomly chosen from *oma-1* pUG RNAs that were sequenced using Sanger and Illumina sequencing. The accuracy of RDE-3 was calculated as the [(length of the pUG tail without the 18nt RT oligo - the # of interruptions)/length of pUG tail] * 100. Of note, errors could have arisen as a result erroneous RDE-3 activity, during PCR or during sequencing.

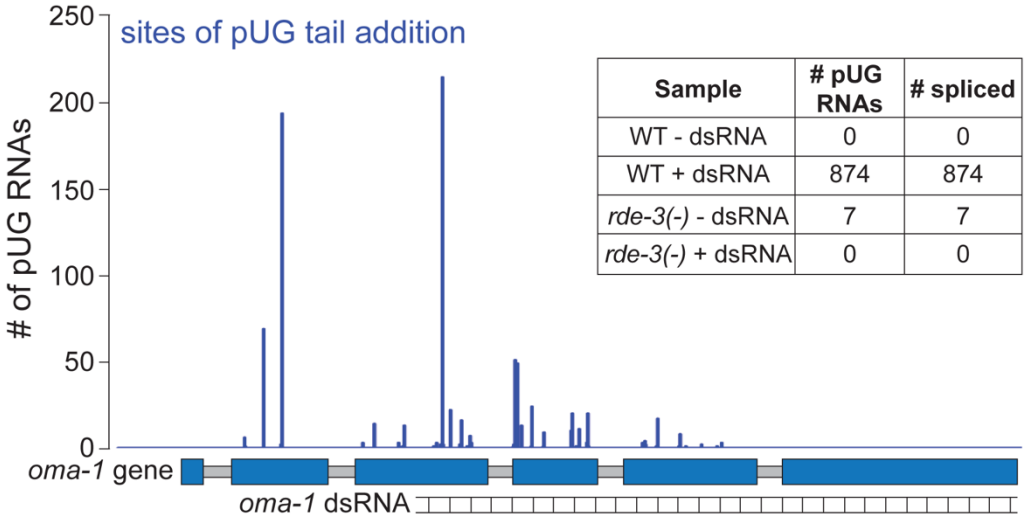
Sanger Sequencing							
pUG RNA #	**Total length of pUG tail**	**Total length - 18nt oligo**	**Trimmed pUG tail (Sequenced tail - 18nt oligo)**	**Pairs**	**# errors**	**Type**	**Accuracy**
1	41	23	TGTGTGTGTGTGTGTGTG TGTGT	11.5	0	NA	100
2	25	7	TGTGTGT	3.5	0	NA	100
3	30	12	TGTGTGTGTTGT	6	1	insertion of T	91.66666667
4	50	32	GTGTGTGTGTGTGTGTGT GTGTGCGTGTGTGT	16	1	T --> C transition	96.875
5	33	15	TGTGGGTGTGTGTGT	7.5	1	T --> G transition	93.33333333
6	45	27	TGTGTGTGTGTGTGTGTG TGTGTGTGT	13.5	0	NA	100
7	25	7	TGTGTGT	3.5	0	NA	100
8	33	15	TGTGTGTGTGTGTGT	7.5	0	NA	100
9	22	4	GTGT	2	0	NA	100
10	25	7	TGTGTGT	3.5	0	NA	100
11	27	9	GTGTATGT	4.5	1	G --> transition	88.88888889
12	35	17	TGTGTGTGTGTGTGTGT	8.5	0	NA	100
13	23	5	TGTGT	2.5	0	NA	100
14	32	14	TGTGTGTGTGTGTT	7	1	insertion of T	92.85714286
15	74	56	GTGTGTGTGTGTGTGTGT GTGTGTGTGTGTGTGTGT GTGTGTGTGTGTGTGTGT GT	28	0	NA	100
Illumina MiSeq							
Read #	**Total length of pUG tail**	**Total length - 18nt oligo**	**Trimmed pUG tail (Sequenced tail - 18nt oligo)**	**Pairs**	**# errors**	**Type**	**Accuracy**
1	32	14	TGTGTGTGTTGTGT	7	1	insertion of T	92.85714286

2	50	32	TGTGTGTGTTGTGTGTGTGTGTGTGTGTGT	16	1	insertion of T	96.875
3	23	5	TGTGT	2.5	0	NA	100
4	31	13	TGTGTGTGTGTGT	6.5	0	NA	100
5	57	39	TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	19.5	0	NA	100
6	37	19	TGTGTGTGTGTGTGTGTGT	9.5	0	NA	100
7	33	15	TGTGTGTGTGTGTGT	7.5	0	NA	100
8	44	26	TGTGTGTGTTGTGTGTGTGTGTGTGT	13	1	insertion of T	96.15384615
9	35	17	TGTGTGTGTGTGTGTGT	8.5	0	NA	100
10	45	27	TGTGTGTGTGTGTGTGTGTGTGTGT	13.5	0	NA	100
11	47	29	TGTGTGTGTGTGTGTGTGTGTGTGT	14.5	0	NA	100
12	53	35	TGTGTGTGTGTGTGTGTGTGTGTGTGTGT	17.5	0	NA	100
13	31	13	TGTGTGTGTGTGT	6.5	0	NA	100

Figure 2.2. Analysis of *oma-1* pUGylation sites. a, Illumina MiSeq was performed (n=1 biological experiment) on *oma-1* pUG PCR products derived from WT and *rde-3(-)* animals, +/- *oma-1* dsRNA. # of sequenced pUG RNAs (y-axis) mapping to each pUGylation site (x-axis) is shown. Inset: total number of sequenced *oma-1* pUG RNAs from indicated samples and total number of these sequenced pUG RNAs in which the *oma-1* sequence was spliced. **b,** MiSeq-sequenced *oma-1* pUG RNAs were sorted into four groups based on the nucleotide (nt) at the last templated position (-1) of the *oma-1* mRNA. The % of *oma-1* pUG RNAs (MiSeq reads) with each nt in the -1 position is shown. Logo analysis was then performed on each of the four groups to determine the probability of finding each nucleotide at the first position of the pUG tail (+1), as well as at the second-to-last templated nucleotide of *oma-1* (-2). This analysis showed that if the last templated nucleotide of the *oma-1* mRNA fragment was an A or a C, then RDE-3 was equally likely to add a U or a G as the first nucleotide of an elongating pUG tail. If, however, the last templated nt was a U or G, then RDE-3 preferentially added a G or U, respectively, as the first nt in an elongating pUG tail. *Note: To perform the analyses in this figure, we assumed that if a U or G could have been genomically encoded, then it was. If, instead, RDE-3 added the U or G shown in the -1 position as the first nucleotide of the pUG tail, then these data show that the second nucleotide that RDE-3 prefers to add is a G after a U or a U after a G.

CCA-adding rNT enzymes modify the 3' termini of tRNAs with non-templated CCA nts. The mechanism by which these enzymes add non-templated nonhomopolymeric stretches of nts is thought to involve allosteric regulation of the nt binding pocket by the 3' nt of a substrate tRNA (Xiong and Steitz 2004). A similar mechanism may explain how RDE-3 can add pUG tails to its mRNA substrates. For instance, when the 3' nt of an RDE-3 substrate is a U, the rNTP binding pocket of RDE-3 might adopt a structure that preferentially binds G and *vice versa* when 3' nt of an RDE-3 substrate is a G. Such a model could explain how a single rNT enzyme adds perfectly alternating U and G nts to RNA substrates. There are also alternative models for how RDE-3 might add pUG tails to an RNA. These include: 1) the existence a poly(AC) nucleic acid template used by RDE-3 during pUG tail synthesis, 2) the existence of one or more rNTs that cooperate with RDE-3 to produce pUG tails, or 3) the possibility that RDE-3 binds and incorporates UG or GU dinucleotides. We disfavor the first two possibilities as these models are difficult to reconcile with the observation that RDE-3 adds UG repeats to tethered RNAs in yeast or in *Xenopus* oocytes (Preston *et al.* 2019). The third proposed model may be true, but because our sequencing shows that pUG tails can initiate with either a U or G (this figure, Tables 1 and 2), then RDE-3 would need to be able to bind both UG and GU dinucleotides. Determining the mechanism by which RDE-3 adds pUG tails will likely involve structural studies and/or *in vitro* pUGylation assays using recombinant RDE-3 protein.

a



Sample	# pUG RNAs	# spliced
WT - dsRNA	0	0
WT + dsRNA	874	874
rde-3(-) - dsRNA	7	7
rde-3(-) + dsRNA	0	0

b

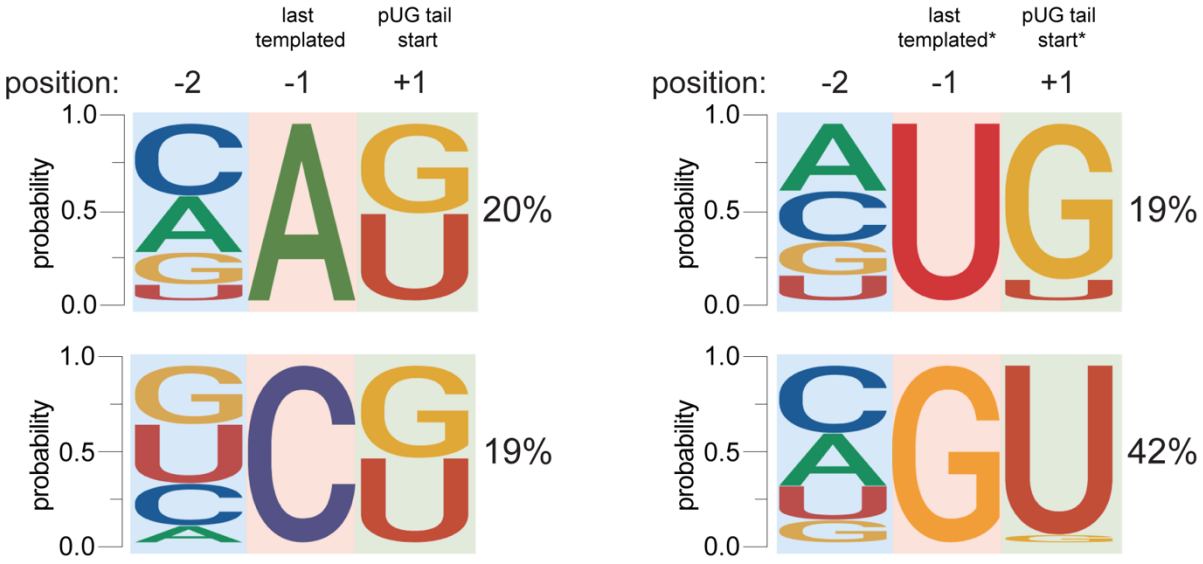
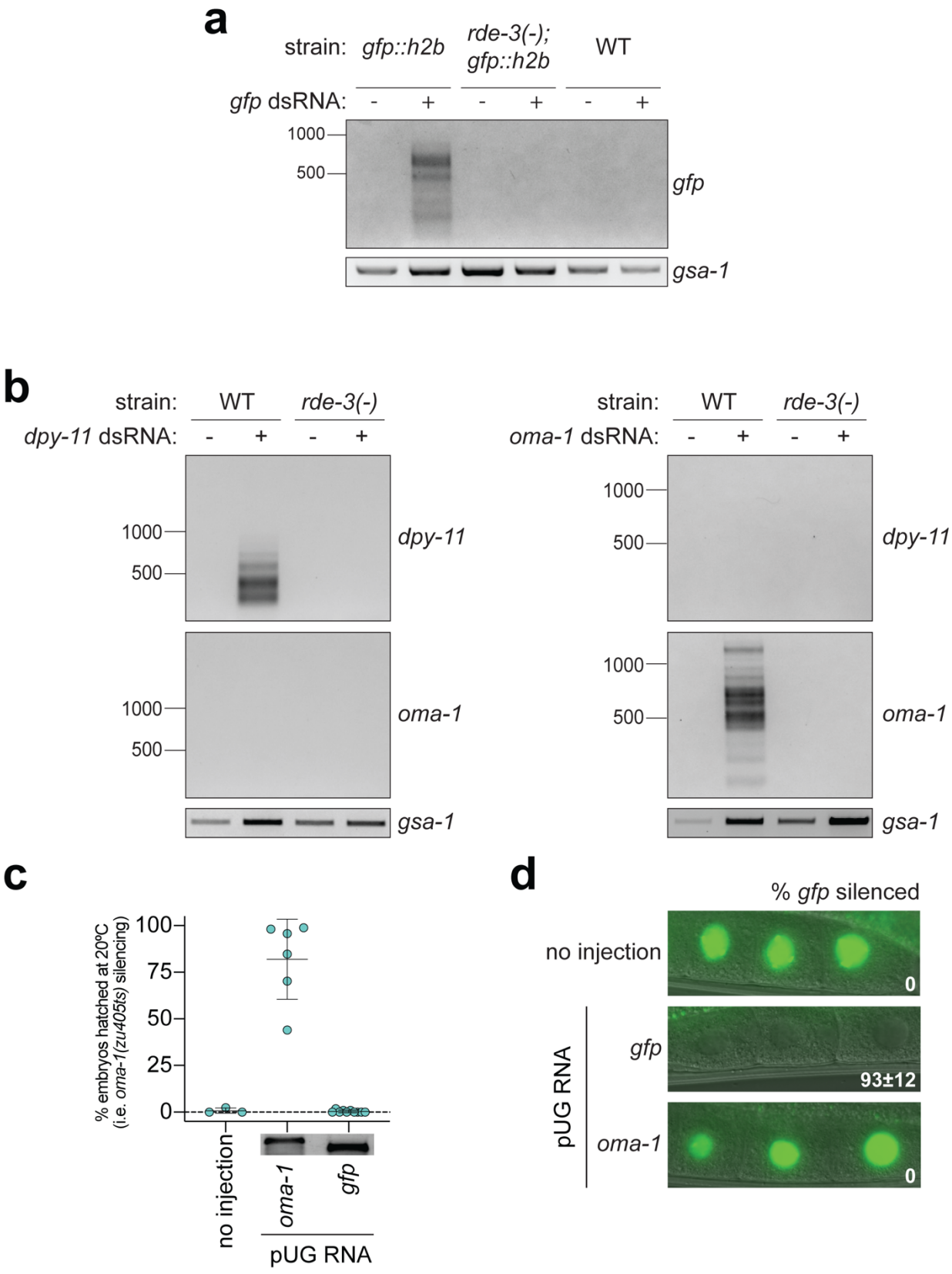


Figure 2.3. RNAi-triggered pUGylation and pUG RNA-directed gene silencing are general and sequence-specific. a, *gfp::h2b, rde-3(-); gfp::h2b* and WT (no *gfp::h2b*) animals were fed *E.coli* expressing either empty vector control or *gfp* dsRNA. **b,** WT and *rde-3(-)* animals were fed *E.coli* expressing empty vector control and either *oma-1* or *dpy-11* dsRNA. For **a** and **b**, *gfp*, *dpy-11* and *oma-1* pUG RNAs were detected using the assay outlined in Figure 2.1a. Data is representative of 3 biologically independent experiments. **c,** *rde-1(ne219); oma-1(zu405ts)* animals were injected with either an *oma-1* (n=6) or *gfp* (n=10) pUG RNA. n=3 for no injection. % embryos hatched was scored for the progeny of injected animals. Inset: injected RNAs run on a 2% agarose gel to assess RNA integrity. Error bars: s.d. of the mean. **d,** *rde-1(ne219); gfp::h2b* animals were injected with either an *oma-1* or *gfp* pUG RNA (n=10 for both, 3 for no injection). Mean % progeny with *gfp::h2b* silenced is indicated ± standard deviation (s.d.). For **c** and **d**, all pUG tails were 36nt in length.


a
strain:
gfp::h2b
rde-3(-); gfp::h2b
WT
gfp dsRNA:
- + - + - +
1000
500
gfp
gsa-1
b
strain:
WT
rde-3(-)
dpy-11 dsRNA:
- + - +
1000
500
dpy-11
oma-1
gsa-1
oma-1 dsRNA:
c
% embryos hatched at 20ºC
(i.e. oma-1(zu405ts) silencing)
100
75
50
25
0
no injection
oma-1
gfp
pUG RNA
d
% gfp silenced
no injection
0
pUG RNA
gfp
93±12
oma-1
0

B. pUG RNAs drive gene silencing

pUG tails could either mark mRNA fragments for degradation or convert mRNA fragments into active mediators of gene silencing. To differentiate these possibilities, we asked whether *in vitro* transcribed pUG RNAs possess gene silencing activity. Indeed, injection of a *gfp* pUG RNA (i.e. 18 3'-terminal pUG repeats appended to the first 369nt of the *gfp* mRNA) into animals expressing a germline-expressed *gfp::h2b* transgene was sufficient to silence *gfp::h2b* expression (Figure 2.4a). The same *gfp* mRNA fragment without a 3' tail or with 18 3'-terminal pAU, pGC or pAC repeats lacked gene silencing activity (Figure 2.4a). Note: to control for potential dsRNA contamination in our *in vitro* transcription reactions, all RNAs were injected into *rde-1(ne219)* mutants, which cannot respond to dsRNA (Figure 2.4a,b) (Tabara *et al.* 1999). The ability of a pUG tail to confer gene silencing activity on an mRNA fragment was both general and sequence-specific. *oma-1(zu405ts)* animals lay arrested embryos at 20°C unless *oma-1(zu405ts)* is silenced (Lin 2003). An *in vitro* transcribed 541nt long *oma-1* mRNA fragment with 18 3' pUG repeats (hereafter, *oma-1* pUG RNA)—but not 18 3' pAU, pGC or pAC repeats—was capable of silencing *oma-1(zu405ts)* (Figure 2.4b). Additionally, an *oma-1* pUG RNA injection did not silence *gfp::h2b* and *vice versa* (Figure 2.3c,d). Finally, while RDE-3 was required for efficient *oma-1* RNAi (Figure 2.5a), this requirement could be bypassed by an *oma-1* pUG RNA injection (Figure 2.5b), establishing that RDE-3-mediated pUGylation is necessary for RNAi. We conclude that pUG tails convert otherwise inert mRNA fragments into agents of gene silencing.

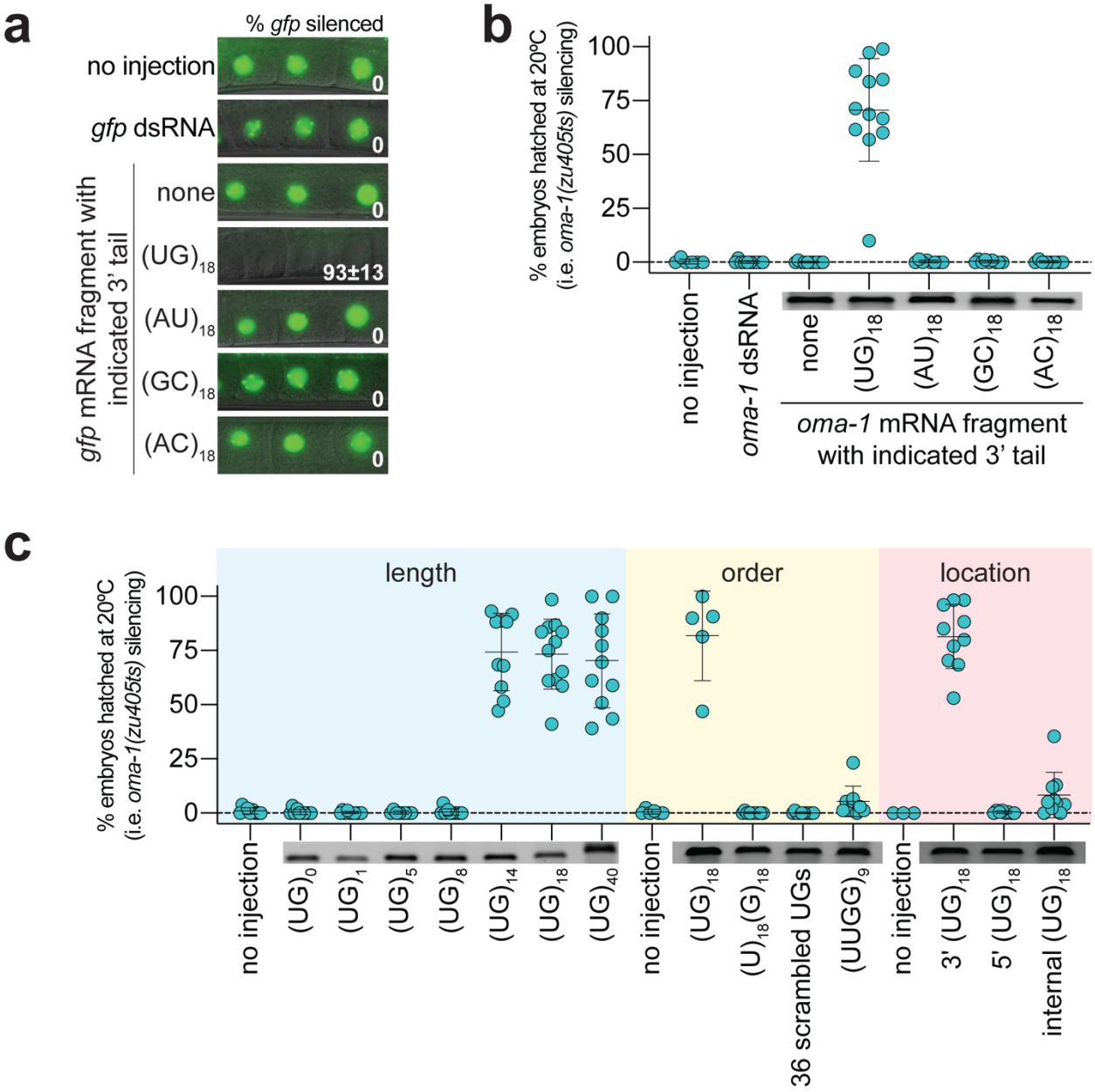


Figure 2.4. pUG tails convert inert RNA fragments into agents of gene silencing. a-c, To control for potential dsRNA contamination in *in vitro* transcription reactions, RNAs were injected into *rde-1(ne219)* mutants, which cannot respond to dsRNA(Tabara *et al.* 1999). **a,** Fluorescence micrographs showing -1 to -3 oocytes of adult progeny of *rde-1(ne219); gfp::h2b* animals injected in the germline with *in vitro* transcribed RNAs consisting of the first 369nt of *gfp* mRNA with the indicated 3' terminal repeats. Mean % progeny with *gfp::h2b* silenced is indicated ± standard deviation (s.d.). # of injected animals (n) = 3 (no injection); 9 [*gfp* dsRNA, $(AU)_{18}$]; 10 [no tail, $(GC)_{18}$, $(AC)_{18}$]; and 16 [$(UG)_{18}$]. **b-c,** *oma-1(zu405ts)* animals lay arrested embryos at 20°C unless *oma-1(zu405ts)* is silenced (Lin 2003). Adult *rde-1(ne219); oma-1(zu405ts)* animals were injected in the germline with *in vitro* transcribed RNAs consisting of the first 541nt of *oma-1* mRNA with **b,** indicated 3' terminal repeats or **c,** varying 3' pUG tail lengths; different 3' UG repeat sequences; or with $(UG)_{18}$ on the 3' end, 5' end or in the middle of the *oma-1* mRNA. For all *oma-1* pUG RNA injection data, each point represents % hatched embryos laid by 5 progeny derived

Figure 2.4 (Continued). from one injected animal at 20°C (see Methods). Error bars: s.d. of the mean. Insets: injected RNAs run on 2% agarose gel to assess RNA integrity. For **b,** n=6 (no injection); 10 (*oma-1* dsRNA, no tail); 12 [$(UG)_{18}$]; 9 [$(GC)_{18}$]; and 8 [$(AU)_{18}$, $(AC)_{18}$]. For **c,** n=9 [no injection #1, ($U_{18}G_{18}$), $(UUGG)_9$]; 10 [$(UG)_{0,\ 1,\ 8,\ 14}$, scrambled UGs, 3' $(UG)_{18}$, internal $(UG)_{18}$]; 8 [$(UG)_5$, 5' $(UG)_{18}$]; 12 [$(UG)_{18}$ #1]; 5 [$(UG)_{18}$ #2]; 11 [$(UG)_{40}$]; 6 (no injection #2) and 3 (no injection #3).

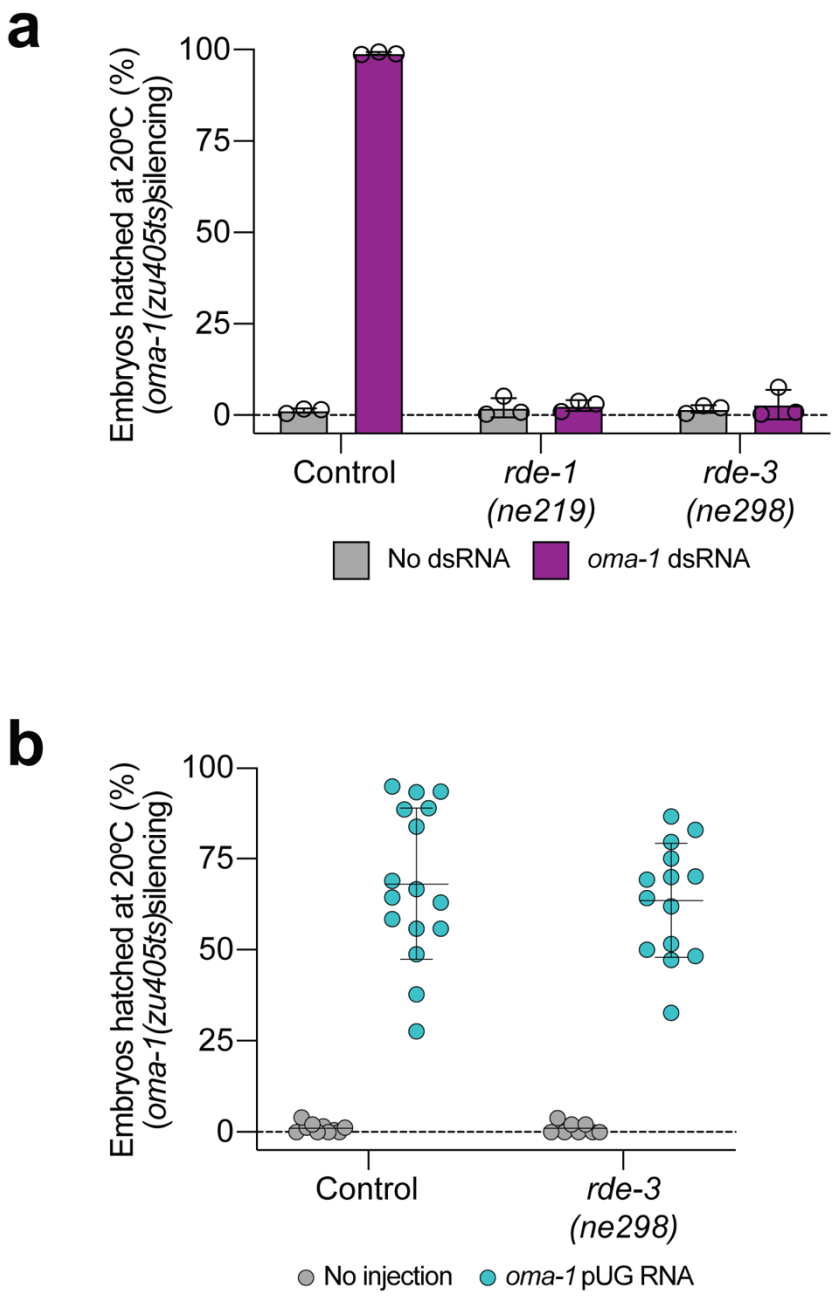


Figure 2.5. RDE-3-mediated pUGylation is necessary for RNAi. a, Animals of the indicated genotypes (all harboring the *oma-1(zu405ts)* mutation) were treated +/- *oma-1* dsRNA. For each experiment, % embryos hatched was scored at 20°C and averaged for 6 individual animals per treatment for each genotype. *rde-1(ne219)* mutants, which cannot respond to dsRNA (Tabara *et al.* 1999), serve as a control for this experiment. Error bars: s.d. of the mean for 3 biologically independent experiments. **b,** Control or *rde-3(ne298)* animals (all *rde-1(ne219); oma-1(zu405ts)* background) were injected with *oma-1* pUG RNAs and % embryos hatched was scored at 20°C. n=10 noninjected and 16 injected animals for control. n=8 noninjected and 14 injected animals for *rde-3(ne298)*. Error bars: s.d. of the mean.

We used our pUG RNA injection assay to define the features of pUG RNAs required for biological activity. We injected animals with the same *oma-1* mRNA fragment harboring varying numbers of 3' UG repeats and found that *oma-1* pUG RNAs with 14, 18, or 40—but not 1, 5, or 8—UG repeats were capable of triggering *oma-1* gene silencing (Figure 2.4c). We also found that while perfectly alternating 3' U and G repeats conferred silencing activity on an mRNA fragment, 3' tails with scrambled UG sequence or other combinations of Us and Gs did not (Figure 2.4c). Moreover, while an *oma-1* mRNA fragment with a 3' pUG tail triggered *oma-1(zu405ts)* silencing, *oma-1* mRNA fragments with 5' or internal UG repeats did not (Figure 2.4c). Finally, the *oma-1* segment of an *oma-1* pUG RNA had to possess the sense coding sequence (Figure 2.6a) and be >50nt in length for pUG RNA functionality (Figure 2.6b). Together, these data show that a pUG RNA must consist of >8 3' UG repeats appended to >50nt of sense RNA in order to trigger gene silencing.

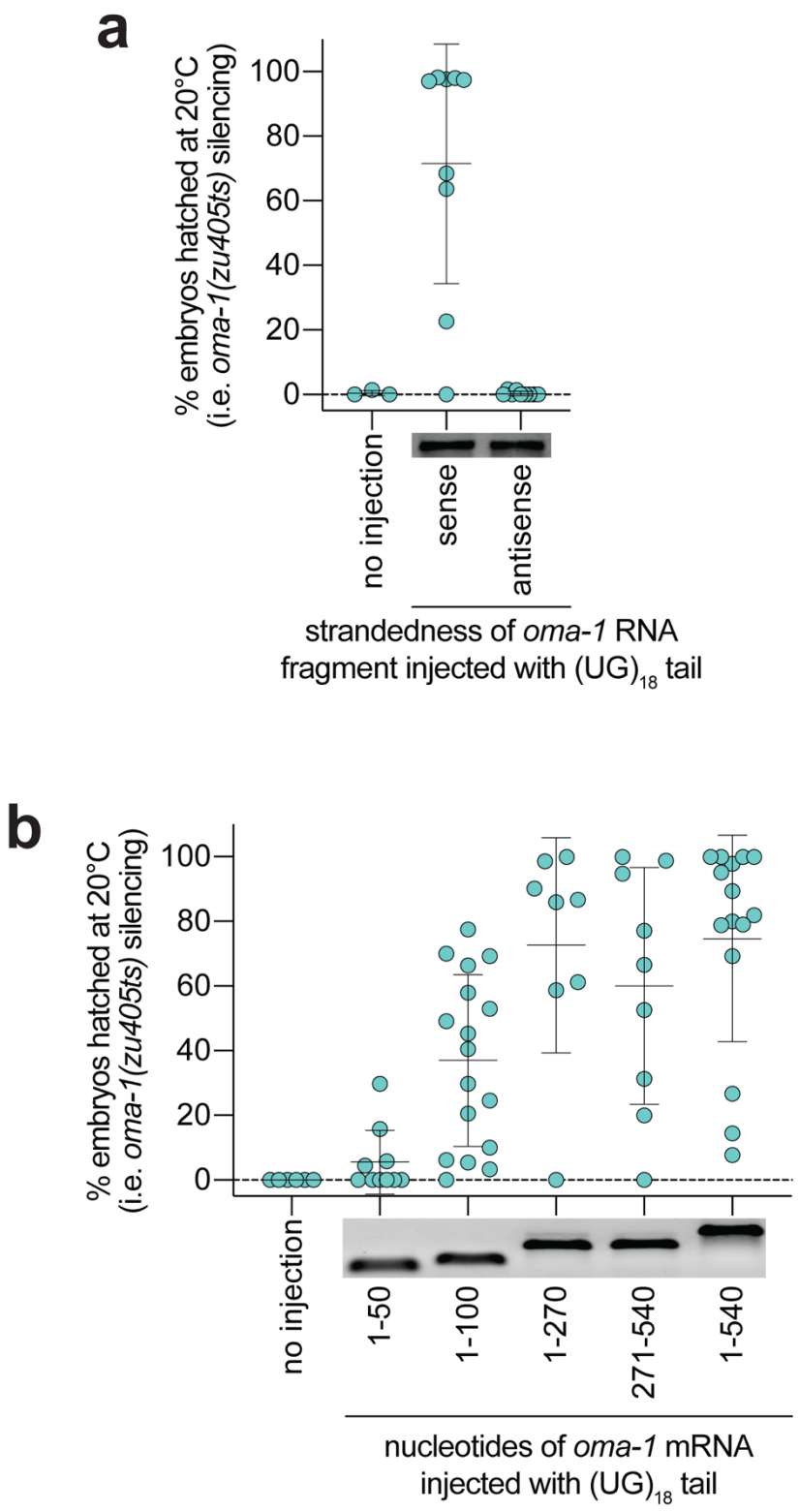


Figure 2.6. pUG tails must be appended to sense RNAs of >50nt for functionality. *rde-1(ne219); oma-1(zu405ts)* animals were injected with: **a,** an *oma-1* pUG RNA consisting of the sense or antisense strand of the same 541nt long *oma-1* mRNA fragment (beginning at the *aug)*

Figure 2.6 (Continued). with a 36nt 3' pUG tail (n=9 for both; n=3 for no injection). **b,** *oma-1* pUG RNAs consisting of *oma-1* mRNA fragments of varying lengths (with position 1 starting at the *aug* of the *oma-1* mRNA sequence) all appended to a 36nt pUG tail. n=6 (no injection), 10 (1-50), 17 (1-100), 8 (1-270), 9 (271-541) and 15 (1-541). For **a** and **b**, % embryonic arrest was scored at 20°C. Error bars: s.d. of the mean.

C. RDE-3 pUGylates germline-expressed RNAs

We next asked whether endogenous mRNAs are pUGylated in *C. elegans*. Tc1 is the most common DNA transposon in the *C. elegans* genome (Fischer *et al.* 2003). In the absence of RDE-3, Tc1 transposase RNA is upregulated and Tc1 mobilizes (Collins *et al.* 1987), suggesting that Tc1 RNA might be pUGylated in wild-type animals. Indeed, using a Tc1-specific pUG PCR assay (Figure 2.1a), we observed RDE-3–dependent pUG tails appended to Tc1 RNA fragments (Figure 2.7a, Table 2). In addition, Tc1 mobilization caused by *rde-3* mutation was suppressed by injection of a Tc1 pUG RNA (Figure 2.7b). We conclude that RDE-3–based pUGylation silences the Tc1 transposon in *C. elegans.*

To identify additional targets of pUGylation, we conducted mRNA-seq on wild-type and *rde-3(-)* animals and identified 346 RNAs that were upregulated in *rde-3(-)* animals (Table 3, adjusted p value <0.05 and log2 fold change >1.5), including Tc1 RNA (Collins *et al.* 1987), as well as six other DNA transposons (Tc1A, Tc4, Tc5, MIRAGE1, CEMUDR1, Chapaev-2), several LTR retrotransposons (Cer3, Cer9, Cer13), and 294 predicted protein-coding RNAs (Table 3, Figure 2.8a). Directed pUG PCR analyses confirmed that Tc4v, Tc5, Cer3, and four of five genes tested from amongst our list of the top 25 most RDE-3–regulated mRNAs were pUGylated in an RDE-3–dependent manner (Figure 2.8b,c). pUG tails were not detected on RNAs whose expression is unchanged in *rde-3* mutants, including *oma-1, gfp, dpy-11* (Figure 2.3a,b), and two additional genes selected at random (Figure 2.8d). We conclude that RDE-3 adds pUG tails to endogenous RNAs in *C. elegans*, which include, but are not limited to, transposon RNAs.

Figure 2.7. Endogenous pUG RNAs exist and localize to germline *Mutator* foci. a, Tc1 pUG PCR (Figure 2.1a) on total RNA from two replicates of indicated genotypes (rescue/reversion as in Figure 2.1b). WT vs. *rde-3(ne3370)* and WT vs. *rde-3(ne298)* is representative of >5 and 2 biologically independent experiments, respectively. **b,** 18 and 22 *rde-3(-); unc-22::tc1* animals were injected with Tc1 pUG RNA + co-injection marker or co-injection marker alone, respectively. Each data point (n) represents # of mobile progeny (indicating Tc1 mobilized from *unc-22*) laid by 25 randomly pooled co-injection marker–expressing progeny derived from injected animals (see Methods). n=9 for co-injection marker only, 6 for Tc1 pUG RNA + co-injection marker, 6 for noninjected *rde-3(+); unc-22::tc1*. Error bars: s.d. of the mean. **c-d,** Fluorescence micrographs of adult pachytene stage germ cell nuclei. DNA stained with 4',6-diamidino-2-phenylindole (DAPI, blue). Data is representative of 3 biologically independent experiments. **c,** RNA FISH to detect pUG RNAs on germlines dissected from WT or *rde-3(-)* animals using $(AC)_9$ DNA oligo conjugated to Alexa 647 (magenta). Positive control: RNA FISH to detect *ama-1* mRNA (green). **d,** pUG RNA FISH (magenta) combined with immunofluorescence to detect GFP::degron::RDE-3 (green). **e,** Tc1 pUG PCR on total RNA isolated from replicates of *glp-1(q224/ts)* animals grown at 15°C (germ cells present) or 25°C (<99% of germ cells). Data is representative of 2 biologically independent experiments. **f,** Control, *rde-8(tm2252)* or *mut-16(pk710)* animals (all *rde-1(ne219); oma-1(zu405ts)* background) were injected with *oma-1* pUG RNAs (n=9, 12 and 8, respectively) and % embryos hatched scored at 20°C. n=3 for all no injection. Error bars: s.d. of the mean.

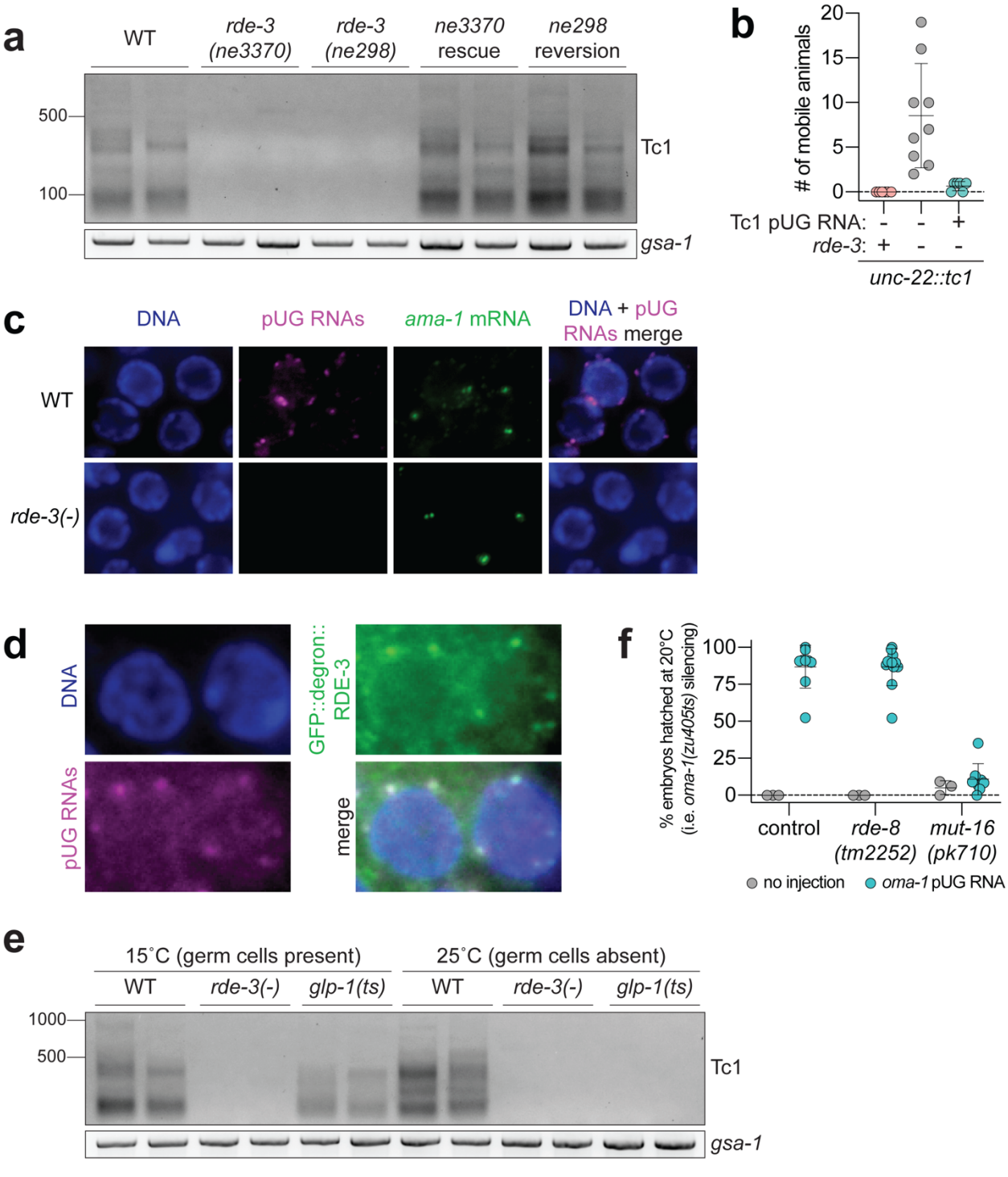

a
WT
rde-3 (ne3370)
rde-3 (ne298)
ne3370 rescue
ne298 reversion
500
100
Tc1
gsa-1
b
# of mobile animals
20
15
10
5
0
Tc1 pUG RNA: - - +
rde-3: + - -
unc-22::tc1
c
DNA
pUG RNAs
ama-1 mRNA
DNA + pUG RNAs merge
WT
rde-3(-)
d
DNA
pUG RNAs
GFP::degron::RDE-3
merge
f
% embryos hatched at 20°C (i.e. oma-1(zu405ts) silencing)
100
75
50
25
0
control
rde-8 (tm2252)
mut-16 (pk710)
no injection
oma-1 pUG RNA
e
15˚C (germ cells present)
25˚C (germ cells absent)
WT
rde-3(-)
glp-1(ts)
WT
rde-3(-)
glp-1(ts)
1000
500
Tc1
gsa-1

Table 2.2. Tc1 pUG RNAs sequenced using Sanger sequencing. Tc1 pUG RNAs were detected (Figure 2.7) in WT animals using the assay in Figure 2.1a and then Sanger sequenced. Sequencing results are reported for Tc1 pUG RNAs that had a pUG tail >18nt (i.e. the length of the RT oligo).

Tc1 pUG RNA #	Sequence
1	CGTTCTCCGTCGACTATTTGGAAAGTAATCAAGAAGTACCAAACTGAGGT GAGTTCGAAAAATATTATTTTTTAATAATAAATGTTTAGTGTGTGTGTGTGT GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTTGTGTGTGT
2	CGTTCTCCGTCGACTATTTGGAAAGTAATCAAGAAGTACCAAACTGAGAAA TCCGTCGCTTTGAGTGAGTGTGTGTGTGTGTGT
3	CGTTCTCCGTCGACTATTTGGAAAGTAATCAAGAAGTACCAAACTGAGAAA TCCGTCGCTTTGAGAATCTCGTGTGTGTGTGTGTGTTGTGTGTGT
4	CGTTCTCCGTCGACTATTTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT GTGTGTGTGTGTGTGTGTGTGTGTGT
5	CGTTCTCCGTCGACTATTTGGAAAGTAATCAAGAAGTACCAAACTGAGAAA TCCGTCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT GTGTGT
6	CGTTCTCCGTCGACTATTTGGAAAGTAATCAAGAAGTACCAAACTGAGTGT GGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT
7	CGTTCTCCGTCGACTATTTGGAAAGTAATCAAGAAGTACCAAACTGAGAAA TCCGTCGCTTTGAGAATCTCGCCCGGCAGGCCTCGAGTGACAACCCATA GGATGGATCGCAACATCCTCCGATCAGCAAGAGAAGATCCGCATAGGAC CGCCACGGAGATGTGTGTGTGTGTGTGTGTGT
8	CGTTCTCCGTCGACTATTTGGAAAGTAATCAAGAAGTACCAAACTGAGAAA TCCGTCGCTTTGAGAATCTCGCCCGGCAGGCCTCGAGTGACAATGTGTG TGTGTGTGTGTGT
9	CGTTCTCCGTCGACTATTTGGAAAGTAATCAAGAAGTACCAAACTGAGGT GAGTTCGTGTGTGTGTGTGTGTGTGTGT
10	CGTTCTCCGTCGACTATTTGGAAAGTAATCAAGAAGTATCAAACTGAGGT GAGTTCGAAAAATATTATTTTTTAATAATAAATGTTTAGAAATCCGTCGCTT TGAGAATCTCGCCCGGCAGGCCTCGAGTGACAACCCATAGGATGGATCG CAACATCCTCCGATCAGCAAGAGAAGATCCGCATAGGACCGCCACGTGT GTGTGTGTGTGTGTGTGTGTGT
11	CGTTCTCCGTCGACTATTTGGAAAGTAATCAAGAAGTACCAAACTGAGAAA TCCGTCGCTTTGAGAATCTCGCCCGGCAGGCCTCGAGTGACAACCCATA GGATGGATCGCAACATCCTCCGATCAGCAAGAGAAGATCCGCATAGGAC CGCCACAGATATTCAAATGATTATAATGTGTGTGTGTGTGTGTGT
12	CGTTCTCCGTCGACTATTTGGAAAGTAATCATGTGTGTGTGTGTGTGTGT GTGTGTGTGTGTGTGTGTGT

Table 2.3. Genes upregulated in *rde-3(-)* mutants. List of mRNAs upregulated in *rde-3(-)* mutants (adjusted p value <0.05 and log2fold change >1.5). Data is based on mRNA-seq performed on two replicates each of WT vs. *rde-3(-)* animals.

Gene	log2FoldChange	Fold Change	padj
F11D11.3	11.004630	2054.583329	0.000000
C38D9.2	9.579959	765.341134	0.000000
21ur-10687	8.762330	434.234280	0.000000
21ur-10492	8.707421	418.017880	0.000000
Y47H10A.5	8.644678	400.227849	0.000000
F15D4.5	8.532555	370.301097	0.000000
F49F1.17	8.464473	353.232208	0.000002
F55B11.6	8.301006	315.392879	0.000000
K08D10.11	8.105936	275.505278	0.000009
K08D10.15	7.993551	254.858254	0.000015
ZC15.3	7.872015	234.267744	0.000000
cyp-35A1	7.394325	168.233919	0.000197
F11D11.4	6.890244	118.623359	0.001195
F07B7.1	6.717298	105.222409	0.000000
K09H11.11	6.710726	104.744166	0.000000
F40G12.7	6.595150	96.680278	0.003510
srg-31	6.494310	90.153397	0.006330
sdz-25	6.427758	86.089038	0.006302
abu-8	6.420851	85.677915	0.005803
cyp-13A8	6.329242	80.406587	0.008667
nhx-6	6.316558	79.702793	0.000000
clec-168	6.227828	74.948490	0.001056
C38D9.13	6.187207	72.867660	0.001135
phat-5	6.126339	69.857280	0.015744
W09B7.1	6.067499	67.065505	0.000000
fbxb-44	6.002675	64.118755	0.019368
F49F1.8	5.993906	63.730216	0.000000
C08B6.4	5.944474	61.583576	0.002490
Y41C4A.32	5.913574	60.278614	0.000000
F55C9.5	5.889758	59.291686	0.027822
F33H12.7	5.830795	56.917299	0.000000
F53H2.1	5.811589	56.164590	0.003959
C08F11.7	5.790694	55.356998	0.000000

col-165	5.776923	54.831136	0.000000
ZK262.9	5.705727	52.190923	0.000000
ins-31	5.684206	51.418155	0.043985
F15B9.6	5.667191	50.815306	0.000000
ZC15.1	5.593766	48.293794	0.006974
cut-3	5.544692	46.678694	0.000000
F15D4.6	5.282491	38.921392	0.000000
C18H7.1	5.281804	38.902865	0.000683
C08E8.4	5.081604	33.862200	0.000000
mtl-1	5.060825	33.377975	0.005567
F47B8.4	5.031002	32.695081	0.029601
hch-1	4.999135	31.980825	0.000007
col-74	4.972545	31.396790	0.000003
tbb-6	4.859603	29.032618	0.000000
ZC15.4	4.727083	26.484629	0.000002
cyp-13A7	4.593912	24.149340	0.000000
ZK262.8	4.562640	23.631517	0.000000
F58H7.5	4.480639	22.325781	0.000000
fbxa-163	4.427819	21.523175	0.018092
CER3-LTR_CE:Gypsy:LTR	4.359420	20.526565	0.018610
fbxb-97	4.351254	20.410700	0.000000
T24E12.5	4.349305	20.383144	0.000001
dpy-14	4.340054	20.252867	0.000000
Y37H9A.2	4.234259	18.820845	0.000088
phg-1	4.204443	18.435866	0.026945
T26H5.9	4.163461	17.919536	0.000000
C25F9.11	4.153501	17.796243	0.000366
hil-7	4.149487	17.746796	0.042054
Y61A9LA.12	4.089781	17.027336	0.000000
F08G2.5	4.065528	16.743485	0.000000
asp-17	4.015041	16.167678	0.000000
T08B6.5	3.992850	15.920905	0.016596
clec-219	3.986530	15.851312	0.046371
pqn-73	3.964612	15.612308	0.005495
cdr-4	3.949260	15.447055	0.000000
F30H5.3	3.945270	15.404395	0.000719
R09E10.5	3.933795	15.282359	0.006578
E04D5.4	3.745942	13.416550	0.000011

F19B10.14	3.714450	13.126859	0.000001
Y61A9LA.5	3.657123	12.615477	0.000000
srw-85	3.617988	12.277870	0.006157
R02F11.1	3.606017	12.176411	0.009512
lpr-4	3.599340	12.120190	0.000043
npp-26	3.564763	11.833154	0.005804
CER3-I_CE:Gypsy:LTR	3.562396	11.813757	0.000000
ZK896.1	3.546181	11.681725	0.000000
Y25C1A.6	3.503706	11.342811	0.003319
cyp-13A6	3.473748	11.109699	0.039265
K04H4.2	3.473022	11.104113	0.000061
sams-2	3.422187	10.719661	0.000354
bath-13	3.408033	10.615003	0.000062
C32B5.6	3.402852	10.576949	0.007459
T19D12.4	3.376438	10.385065	0.000000
fbxa-88	3.373428	10.363417	0.030176
tsp-2	3.356647	10.243569	0.000033
C49G7.12	3.355344	10.234325	0.000000
clec-71	3.337145	10.106034	0.031671
F28F8.7	3.324013	10.014463	0.000000
linc-45	3.320965	9.993328	0.043119
C18H2.1	3.298360	9.837966	0.000000
zip-10	3.296879	9.827873	0.000000
C41D11.6	3.283218	9.735252	0.000000
F49F1.7	3.281479	9.723519	0.000000
arf-1.1	3.277306	9.695438	0.000000
Y47D3B.1	3.272426	9.662700	0.010918
tsp-1	3.245472	9.483842	0.000011
C46G7.111	3.230185	9.383884	0.006662
hpo-2	3.229671	9.380538	0.023181
abu-14	3.213699	9.277262	0.001722
AC8.3	3.187580	9.110817	0.000928
C49G7.7	3.167423	8.984404	0.003716
CER9-I_CE:Pao:LTR	3.155929	8.913110	0.000000
Y48G1BM.8	3.151502	8.885802	0.000000
fbxa-162	3.125136	8.724885	0.006796
acd-1	3.114159	8.658753	0.037769
W09H1.4	3.094846	8.543612	0.000000
rnh-1.3	3.071562	8.406833	0.000000

Y46G5A.6	3.071265	8.405098	0.000000
K01A11.3	3.058378	8.330354	0.000043
T23G5.10	3.053022	8.299488	0.000052
gst-38	3.019840	8.110778	0.000000
C46A5.4	3.010156	8.056518	0.010633
dpy-17	2.949794	7.726389	0.000000
dpy-7	2.939940	7.673792	0.000402
swt-6	2.889336	7.409296	0.000000
Y54G2A.21	2.881957	7.371493	0.000000
F40G9.20	2.881191	7.367580	0.000707
Y54G2A.36	2.880202	7.362532	0.000000
C09G5.7	2.878442	7.353553	0.010903
W09B7.2	2.863314	7.276849	0.000004
grl-7	2.859286	7.256560	0.020234
egl-46	2.851156	7.215786	0.008479
ptr-10	2.851037	7.215186	0.003108
clec-266	2.840712	7.163736	0.000000
Y46H3C.6	2.821313	7.068052	0.000500
F59B2.12	2.819088	7.057161	0.000000
C46G7.5	2.813060	7.027738	0.000002
igcm-1	2.812350	7.024279	0.003971
cdr-2	2.805722	6.992083	0.000000
T19D12.5	2.804440	6.985871	0.000000
arrd-1	2.796316	6.946645	0.000487
TC5:TcMar-Tc4:DNA	2.793723	6.934170	0.000000
strm-1	2.772454	6.832691	0.000549
ZK430.5	2.772184	6.831411	0.000884
T20F5.4	2.768153	6.812351	0.002343
clec-78	2.762211	6.784353	0.000004
MIRAGE1:CMC-Mirage:DNA	2.756012	6.755261	0.000000
Y67D8C.23	2.752907	6.740740	0.000005
F27C1.18	2.739675	6.679198	0.041629
C34H4.2	2.738183	6.672293	0.000000
C08F11.6	2.726997	6.620763	0.000748
F41G4.7	2.724488	6.609257	0.005965
egl-1	2.719917	6.588349	0.034359
dod-20	2.706592	6.527779	0.001940
T28H10.3	2.705124	6.521139	0.000000

fbxa-11	2.693623	6.469360	0.000000
agmo-1	2.665745	6.345547	0.023388
T01D3.3	2.657826	6.310811	0.000001
vet-6	2.657114	6.307698	0.000000
saeg-2	2.649579	6.274843	0.000000
C30G12.1	2.644909	6.254561	0.000226
tts-1	2.629962	6.190097	0.008901
F56D6.16	2.617864	6.138407	0.000000
F29G9.1	2.600949	6.066857	0.024337
T28D6.15	2.597090	6.050648	0.042190
Y38F2AL.12	2.595460	6.043818	0.040078
ptr-4	2.573508	5.952551	0.000023
dod-19	2.558824	5.892270	0.000000
timm-17B.2	2.533041	5.787903	0.000000
fbxa-24	2.518936	5.731593	0.000126
T24H10.5	2.516404	5.721541	0.000964
Y24F12A.3	2.506992	5.684335	0.000068
tts-2	2.504681	5.675237	0.001909
plg-1	2.490226	5.618660	0.000000
B0462.5	2.479055	5.575322	0.000152
Y65B4BL.1	2.474421	5.557441	0.000106
dct-17	2.468628	5.535172	0.000000
cdr-1	2.438822	5.421989	0.032058
dpy-2	2.433454	5.401852	0.000119
F08F3.8	2.428970	5.385089	0.000863
F46E10.2	2.413318	5.326982	0.001648
mul-1	2.413287	5.326867	0.000141
F07B7.2	2.404694	5.295234	0.000055
Tc1:TcMar-Tc1:DNA	2.397966	5.270594	0.000000
Y46G5A.43	2.397633	5.269378	0.000000
F43D9.1	2.395709	5.262356	0.000416
T12B5.9	2.387164	5.231279	0.044456
pqn-32	2.374891	5.186965	0.000043
cut-2	2.371967	5.176466	0.007154
Y53F4B.6	2.346285	5.085130	0.014895
C06C3.5	2.346012	5.084170	0.000000
lec-11	2.336431	5.050516	0.000000
srd-64	2.327448	5.019169	0.013190
T05H10.3	2.324983	5.010599	0.002243

R03H10.6	2.318646	4.988638	0.000772
cyp-13A5	2.317515	4.984730	0.000000
F58E6.13	2.311960	4.965573	0.044408
ZK402.3	2.310253	4.959699	0.006287
sdz-26	2.309087	4.955694	0.027034
mlt-11	2.309043	4.955543	0.000000
Y41D4B.18	2.289897	4.890211	0.000151
F23H12.5	2.283918	4.869987	0.013225
H42K12.3	2.282254	4.864374	0.019479
C03A7.2	2.277905	4.849732	0.000433
B0334.6	2.269271	4.820794	0.011206
H14E04.3	2.269149	4.820387	0.001223
F53B2.8	2.266754	4.812391	0.000068
M163.16	2.264488	4.804838	0.000018
irg-2	2.260546	4.791729	0.000117
grl-5	2.255076	4.773593	0.037253
glf-1	2.253156	4.767245	0.000188
wrt-2	2.249126	4.753947	0.019887
fbxc-51	2.247245	4.747754	0.000002
TC4:TcMar-Tc4:DNA	2.246492	4.745277	0.000000
B0024.4	2.239530	4.722433	0.000000
Y38H8A.8	2.233258	4.701947	0.000008
dao-2	2.231859	4.697389	0.005106
K06B9.2	2.210300	4.627716	0.000044
his-24	2.208196	4.620970	0.000000
arrd-2	2.205732	4.613087	0.014332
Y110A2AL.4	2.187985	4.556687	0.019808
lea-1	2.187661	4.555664	0.000000
pho-9	2.168313	4.494976	0.000000
Y37E3.30	2.157676	4.461957	0.004418
C50F7.5	2.153421	4.448814	0.000065
F41B5.1	2.148172	4.432658	0.000004
C09F9.2	2.142670	4.415786	0.000008
sepa-1	2.134980	4.392309	0.000001
fbxa-182	2.108245	4.311666	0.000000
skr-4	2.105608	4.303791	0.000000
lpr-3	2.102801	4.295424	0.004186
B0391.8	2.094849	4.271815	0.000000
F33D4.6	2.084832	4.242258	0.015383

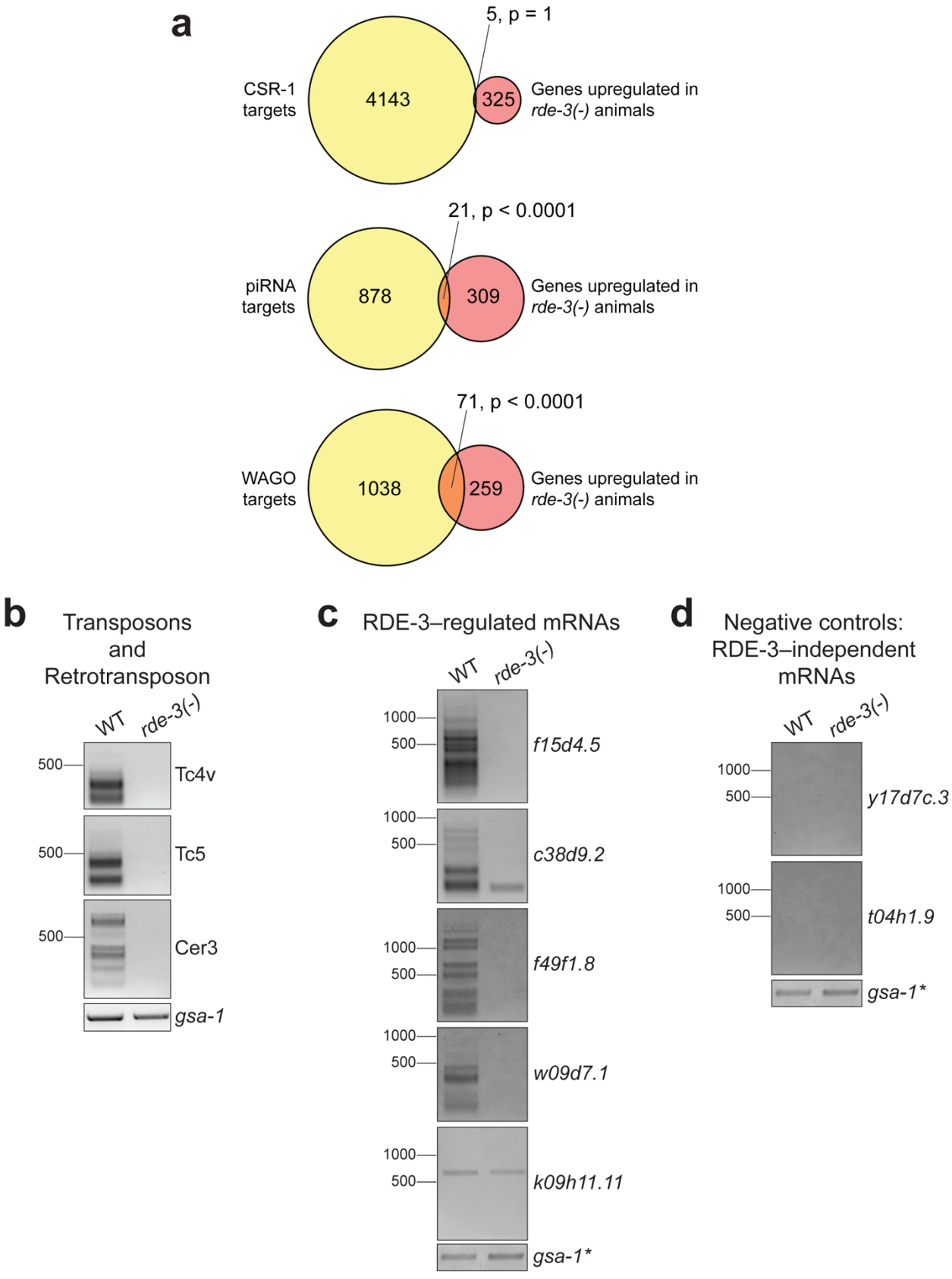
parn-1	2.084298	4.240687	0.000000
sqt-3	2.081944	4.233773	0.000000
tba-7	2.077381	4.220403	0.000000
C48D1.5	2.077229	4.219960	0.001184
F53B1.4	2.073615	4.209401	0.002709
H43E16.1	2.067568	4.191794	0.000000
Y41D4B.17	2.052917	4.149442	0.044964
ZK1055.7	2.040567	4.114071	0.000000
Chapaev-2_CE:CMC-Chapaev:DNA	2.040490	4.113853	0.000000
Y94H6A.10	2.028985	4.081176	0.000000
CEMUDR1:MULE-MuDR:DNA	2.023402	4.065412	0.000000
enu-3.1	2.020330	4.056765	0.000000
ZC123.1	2.018942	4.052865	0.032584
ZC239.14	2.015146	4.042215	0.010266
mlt-8	2.014158	4.039447	0.013092
Y69A2AL.2	2.004571	4.012694	0.000000
hil-3	1.993884	3.983080	0.000001
mlt-9	1.990122	3.972705	0.026865
lys-7	1.976495	3.935358	0.014670
K08D8.4	1.955054	3.877303	0.000000
comt-3	1.951048	3.866553	0.000000
ceh-32	1.946678	3.854857	0.000448
T21C12.8	1.942707	3.844263	0.034781
lys-1	1.941587	3.841280	0.000000
oac-14	1.940129	3.837401	0.002814
Y22D7AR.7	1.936073	3.826627	0.025483
clec-72	1.911538	3.762100	0.000719
igcm-4	1.902375	3.738281	0.000000
pgp-9	1.898403	3.728004	0.000000
M01G12.9	1.896202	3.722319	0.000000
mth-1	1.893001	3.714070	0.000000
wrt-1	1.884184	3.691441	0.006902
cld-9	1.879455	3.679360	0.000000
tbb-4	1.872586	3.661883	0.034709
ugt-62	1.863517	3.638937	0.000000
F37H8.2	1.862728	3.636947	0.000084
anr-35	1.850002	3.605006	0.006993

hsp-17	1.842491	3.586288	0.000000
CER13-I_CE:Pao:LTR	1.839358	3.578508	0.018236
E02C12.6	1.838569	3.576552	0.043798
K03A11.1	1.832727	3.562098	0.000274
valv-1	1.832709	3.562053	0.000442
bath-20	1.827719	3.549754	0.022927
cht-1	1.824882	3.542781	0.000000
clec-41	1.823543	3.539493	0.000000
Tc1A:TcMar-Tc1:DNA	1.816054	3.521169	0.000000
C10C5.2	1.814567	3.517542	0.000069
rrn-2.1	1.813654	3.515316	0.048495
K10D3.4	1.811521	3.510122	0.000231
ceh-20	1.803033	3.489531	0.000000
eva-1	1.797482	3.476130	0.003641
cfz-2	1.795807	3.472097	0.046135
pgp-1	1.786986	3.450931	0.000000
T12D8.5	1.764736	3.398119	0.002525
spe-41	1.763584	3.395406	0.000004
K06B9.4	1.754980	3.375215	0.000000
Y54G11A.14	1.747616	3.358032	0.000006
Y58A7A.4	1.740913	3.342465	0.038594
ugt-19	1.739389	3.338937	0.000000
T02G5.11	1.735085	3.328992	0.001140
cpr-3	1.734979	3.328746	0.000000
ZK6.11	1.727820	3.312270	0.000000
tbx-11	1.725042	3.305897	0.003944
cyp-35A4	1.725000	3.305801	0.047282
Y53F4B.45	1.724599	3.304883	0.000000
dhs-26	1.703384	3.256638	0.000479
EEED8.4	1.703338	3.256536	0.020659
pals-23	1.700221	3.249507	0.000001
lec-8	1.688582	3.223397	0.000000
B0250.7	1.685814	3.217219	0.000000
F12E12.11	1.678468	3.200878	0.016104
C49G7.10	1.676631	3.196805	0.000007
bath-21	1.676520	3.196560	0.018723
Y37H2A.14	1.676051	3.195521	0.002583
enu-3.5	1.672910	3.188571	0.000000
ctl-3	1.672844	3.188424	0.000000

21ur-14806	1.665717	3.172712	0.016454
C40A11.8	1.663948	3.168825	0.015561
clec-180	1.659335	3.158708	0.047894
cgt-1	1.655458	3.150232	0.000685
F26H11.4	1.653446	3.145841	0.001623
ife-5	1.642048	3.121085	0.000000
clec-67	1.633662	3.102997	0.000000
ncam-1	1.633115	3.101820	0.002949
Y97E10AR.1	1.622652	3.079406	0.030518
Y82E9BR.17	1.621113	3.076122	0.021133
clec-62	1.615904	3.065035	0.000000
fbxa-60	1.611315	3.055301	0.000000
C01B10.6	1.609369	3.051184	0.000000
Y39A3A.4	1.606370	3.044848	0.025483
cpt-4	1.605936	3.043931	0.026922
K06G5.1	1.604766	3.041463	0.000000
Y57G11C.51	1.600070	3.031581	0.000000
sid-2	1.598854	3.029025	0.000000
ZK973.8	1.585386	3.000881	0.011123
C25F9.5	1.579854	2.989395	0.000020
ham-1	1.579142	2.987921	0.023319
bcmo-2	1.577406	2.984328	0.000000
R08E3.1	1.576254	2.981945	0.000000
nnt-1	1.576112	2.981652	0.006188
B0205.13	1.574684	2.978703	0.012506
sqst-1	1.562269	2.953179	0.000000
alr-1	1.561742	2.952102	0.006541
cec-8	1.558560	2.945598	0.005085
R09A1.3	1.552950	2.934165	0.001923
col-95	1.552532	2.933315	0.000526
Y22D7AR.2	1.550633	2.929456	0.000095
C44B11.6	1.545497	2.919046	0.049593
T25G12.6	1.538477	2.904876	0.000033
drd-50	1.537954	2.903824	0.000108
cav-1	1.530893	2.889647	0.000000
bam-2	1.530791	2.889443	0.000000
T24B8.3	1.529005	2.885867	0.000000
K10D11.5	1.523299	2.874477	0.000003
gst-39	1.523084	2.874047	0.000270

thn-1	1.514209	2.856422	0.000004
F41E7.1	1.509942	2.847985	0.000631
cyp-33C8	1.509528	2.847168	0.000000
ugt-2	1.506345	2.840894	0.000001
dot-1.5	1.504564	2.837390	0.003026

Figure 2.8. Endogenous targets of pUGylation in *C. elegans*. a, mRNAs upregulated in *rde-3(-)* mutants (Table 3) were compared to published lists of: (1) RNAs targeted by CSR-1–bound endo-siRNAs (Claycomb *et al.* 2009), (2) piRNA-targeted mRNAs (based on predictive and experimental approaches) (Wu *et al.* 2019), and (3) WAGO-class mRNAs (Gu *et al.* 2009). *p*-values were generated using a one-sided Fisher's exact test. This analysis showed statistically significant overlap between the mRNAs upregulated in *rde-3(-)* mutants and both piRNAs targets and WAGO-class mRNAs. **b-d**, Total RNA was extracted from WT or *rde-3(-)* animals. The assay outlined in Figure 2.1a was used to detect pUG RNAs for **b,** two DNA transposons (Tc4v and Tc5) and a retrotransposon (Cer3) that were significantly upregulated in *rde-3(-)* animals; **c,** predicted protein-coding mRNAs that were significantly upregulated in *rde-3(-)* animals; and **d,** two randomly selected mRNAs whose expression does not change in *rde-3(-)* mutants. Data is representative of 3 biologically independent experiments. *Note: the same RT samples were used for panels **c** and **d** and, therefore, the *gsa-1* loading control is the same for both panels.


a
5, p = 1
CSR-1 targets
4143
325
Genes upregulated in rde-3(-) animals
21, p < 0.0001
piRNA targets
878
309
Genes upregulated in rde-3(-) animals
71, p < 0.0001
WAGO targets
1038
259
Genes upregulated in rde-3(-) animals
b
Transposons and Retrotransposon
WT
rde-3(-)
500
Tc4v
Tc5
Cer3
gsa-1
c
RDE-3–regulated mRNAs
1000
500
f15d4.5
c38d9.2
f49f1.8
w09d7.1
k09h11.11
gsa-1*
d
Negative controls: RDE-3–independent mRNAs
y17d7c.3
t04h1.9

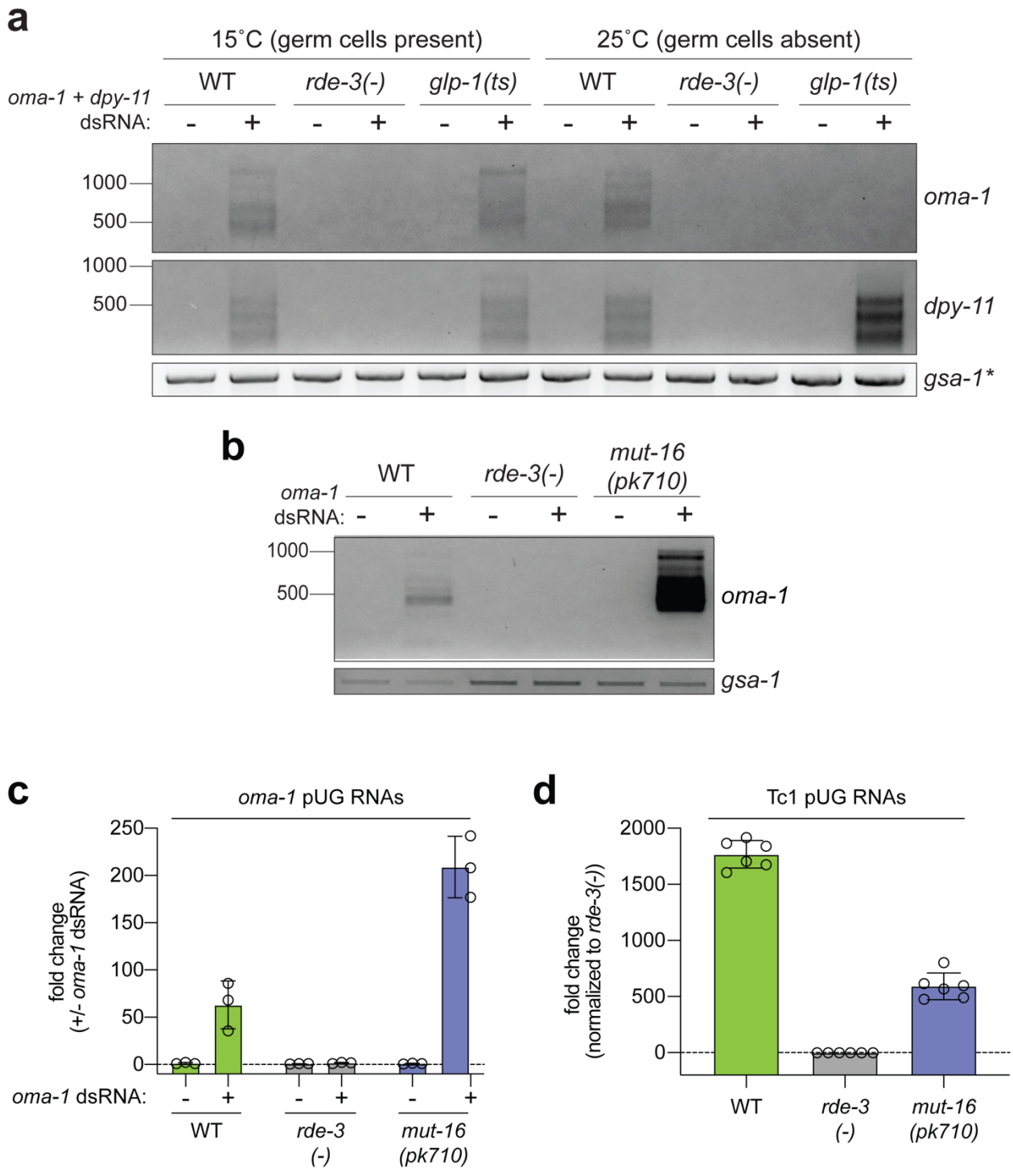
D. pUG RNAs localize to germ granules

Germ granules are liquid-like condensates that form near the outer nuclear membrane in most animal germ cells and likely promote germ cell totipotency by concentrating germline determinants, including maternal RNAs and proteins, into developing germline blastomeres (Voronina *et al.* 2011). *C. elegans* RDE-3 localizes to perinuclear germ granules termed *Mutator* foci (Phillips *et al.* 2012). RNA fluorescence *in situ* hybridization (RNA FISH) using a fluorescently labeled $p(AC)_9$ probe to detect pUG RNAs (pUG FISH) showed that pUG RNAs localized to perinuclear puncta in germ cells of wild-type, but not *rde-3(-),* animals (Figure 2.7c). pUG FISH coupled with immunofluorescence (IF) to detect a GFP- and degron-tagged RDE-3 showed that pUG RNA foci co-localized with RDE-3 and, therefore, *Mutator* foci (Figure 2.7d). These data suggest that pUG RNAs are produced, function, and/or are stored in *Mutator* foci in the *C. elegans* germline. Indeed, *glp-1(q224)* animals, which lack ≅99% of their germ cells when grown at 25°C (hereafter, *glp-1(ts)*) (Austin and Kimble 1987), failed to produce detectable Tc1 pUG RNAs (Figure 2.7e) or *oma-1* dsRNA-induced *oma-1* pUG RNAs (Figure 2.9a) when grown at 25°C, confirming that pUG RNAs are produced or stored in germ cells. Incidentally, when *glp-1(ts)* animals were treated with dsRNA targeting the hypodermis-expressed *dpy-11* gene (Ko and Chow 2002), *dpy-11* pUG RNAs were detected in somatic cells (Figure 2.9a), consistent with previous reports showing that RDE-3 promotes RNAi within and between cells in the *C. elegans* soma (Jose *et al.* 2011; Phillips *et al.* 2012). Hereafter, this work focuses on the biogenesis and function of pUG RNAs in the germline.

To explore further how germline pUG RNAs and *Mutator* foci might relate, we asked if the glutamine/asparagine (Q/N) motif–rich protein MUT-16, which is required for *Mutator* focus assembly in germ cells (Phillips *et al.* 2012), was needed for pUG RNA biogenesis or function. *mut-16(pk710)* animals, which harbor a nonsense mutation in *mut-16,* produced elevated levels of *oma-1* pUG RNAs in response to *oma-1* dsRNA (Figure 2.9b,c) and decreased levels of Tc1

pUG RNAs (Figure 2.9d), suggesting that *Mutator* foci help to coordinate pUG RNA biogenesis. *mut-16(pk710)* animals were completely defective for silencing *oma-1* after an *oma-1* pUG RNA injection (Figure 2.7f), indicating that *Mutator foci* are required for pUG RNA–based gene silencing, downstream of pUG RNA biogenesis.

Figure 2.9. *Mutator* foci likely coordinate pUG RNA biogenesis within germ cells. a, *dpy-11* and *oma-1* pUG PCR (Figure 2.1a) were performed on total RNA from *glp-1(q224/ts)* animals grown at 15°C (germ cells present) or 25°C (<99% of germ cells), +/- *oma-1* and *dpy-11* dsRNA. Data is representative of 2 biologically independent experiments. *Note: the samples in **a** are the same as those used in Figure 2.7e and, therefore, the *gsa-1* loading control is the same. **b,** *oma-1* pUG PCR was performed on total RNA extracted from wild-type, *rde-3(-)*, and *mut-16(pk710)* animals, +/- *oma-1* dsRNA. Data is representative of 4 biologically independent experiments. **c,** qRT-PCR was used to quantify levels of *oma-1* pUG RNAs in wild-type, *rde-3(-)*, and *mut-16(pk710)* animals, +/- *oma-1* dsRNA. Data is represented as fold change in the levels of *oma-1* pUG RNAs +/- *oma-1* dsRNA (y-axis) for each strain (x-axis). n=3 biologically independent samples per treatment for each strain. Error bars: s.d. of the mean. **d,** qRT-PCR was used to quantify levels of Tc1 pUG RNAs in wild-type, *rde-3(-)*, and *mut-16(pk710)* animals. Note: the RNA samples used for **d** are the same as those used in **c**, except that the data for +/- *oma-1* dsRNA samples were pooled for each strain. n=6 biologically independent samples for each strain. Error bars: s.d. of the mean. The analyses in **c** and **d** showed that *mut-16* mutant animals produced more *oma-1*, but fewer Tc1, pUG RNAs, than wild-type animals. The increased levels of *oma-1* pUG RNAs in *mut-16(pk710)* animals was also suggested by the gel in **b**. Together, these data suggest that *Mutator* foci likely have an important role in coordinating pUG RNA biogenesis in germ cells, as pUG RNA levels become misregulated in *mut-16(pk710)* mutants.


a
15˚C (germ cells present)
25˚C (germ cells absent)
WT
rde-3(-)
glp-1(ts)
WT
rde-3(-)
glp-1(ts)
oma-1 + dpy-11
dsRNA:
- + - + - + - + - + - +
1000
500
oma-1
1000
500
dpy-11
gsa-1*
b
oma-1
WT
rde-3(-)
mut-16 (pk710)
dsRNA: - + - + - +
1000
500
oma-1
gsa-1
c
oma-1 pUG RNAs
fold change (+/- oma-1 dsRNA)
250
200
150
100
50
0
oma-1 dsRNA: - + - + - +
WT
rde-3 (-)
mut-16 (pk710)
d
Tc1 pUG RNAs
fold change (normalized to rde-3(-))
2000
1500
1000
500
0
WT
rde-3 (-)
mut-16 (pk710)

E. pUG RNAs are templates for RdRPs

To understand how pUG tails might convert RNAs into agents of gene silencing, we sought to identify pUG tail–binding proteins. We conjugated 5' biotinylated RNA oligonucleotides (oligos) consisting of 18 UG repeats, which conferred gene silencing activity onto *oma-1* and *gfp* mRNA fragments (Figure 2.4), to streptavidin beads. Beads were incubated with wild-type *C. elegans* extracts and bound proteins were analyzed with liquid chromatography–tandem mass spectrometry (LC-MS/MS). Beads conjugated to oligos consisting of 5 UG repeats and 36 scrambled UGs, neither of which conferred gene silencing activity (Figure 2.4c), served as controls. This analysis identified 54 proteins that were enriched ≥2-fold in our $(UG)_{18}$ RNA pull-down versus both control pull-downs (Figure 2.10a). Amongst these proteins were TDP-1, the *C. elegans* ortholog of the mammalian UG-binding protein TDP-43 (Buratti and Baralle 2001; Kuo *et al.* 2009), as well as the RNA-dependent RNA polymerases (RdRPs) EGO-1 and RRF-1 (Figure 2.10a). Current models posit that, during RNAi, RdRPs: 1) are recruited to mRNAs by siRNAs generated from dsRNA, and 2) use these mRNAs as templates to synthesize additional siRNAs, termed secondary (2°) siRNAs, which carry out gene silencing (Sijen *et al.* 2001, 2007; Ambros *et al.* 2003; Pak and Fire 2007). Interestingly, RRF-1, one of four *C. elegans* RdRPs, localizes to *Mutator* foci (Phillips *et al.* 2012). Thus, pUG tails may promote gene silencing by recruiting RdRPs, such as RRF-1, to pUG RNAs, which could then act as templates for siRNA synthesis.

To confirm and expand upon potential pUG tail and RRF-1 interactions, we incubated beads conjugated to RNA oligos consisting of 5, 8, 14, or 18 UG repeats; 18 GC repeats; or 36 scrambled UGs with extracts from animals expressing HA- and TagRFP-tagged RRF-1. α-HA immunoblotting showed that HA::TagRFP::RRF-1 interacted with $(UG)_{18}$, but not scrambled UG or $(GC)_{18}$, RNA oligos (Figure 2.10b). Additionally, HA::Tag::RFP::RRF-1 precipitated strongly with $(UG)_{14}$ and $(UG)_{18}$ RNAs, weakly with a $(UG)_8$ RNA, but not with a $(UG)_5$ RNA (Figure 2.10c). Together, these data show that the RdRP RRF-1 interacts with UG repeat RNAs and that the

sequence determinants of this interaction largely mirror those required for pUG tail–mediated gene silencing *in vivo* (Figure 2.4c).

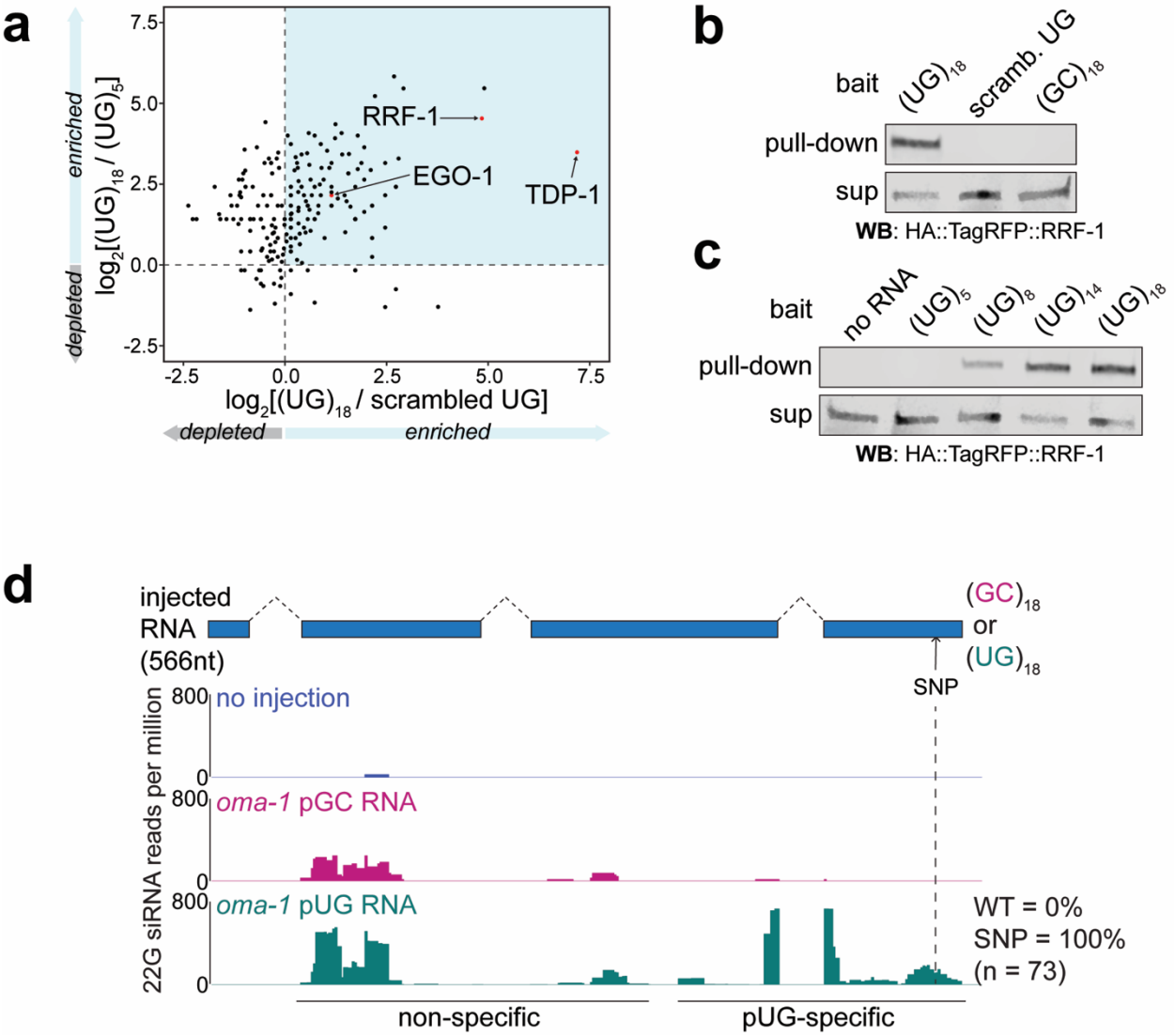


Figure 2.10. pUG RNAs are templates for RdRPs. a, LC-MS/MS was performed on proteins that bound to 5' biotinylated RNA oligos [$(UG)_5$, $(UG)_{18}$ or 36 scrambled UGs] conjugated to streptavidin beads. Shown is a scatter plot of log_2-transformed fold enrichment in $(UG)_{18}$ vs. scrambled UG pull-down (x-axis) and $(UG)_{18}$ vs. $(UG)_5$ pull-down (y-axis). Proteins enriched ≥2-fold in $(UG)_{18}$ vs. beads-only pull-down are plotted. **b-c,** Indicated 5' biotinylated RNA oligos were conjugated to streptavidin beads and incubated with extracts from animals expressing HA::tagRFP::RRF-1. Bead-bound material (pull-down) and supernatant (sup) were subjected to α-HA immunoblotting. Data is representative of 2 biologically independent experiments. **d,** *rde-1(ne219); oma-1(zu405ts)* animals were injected with SNP-containing (dotted line) *oma-1* (*oma-1(SNP)*) pUG or pGC-tailed RNAs and collected 1-4 hours later; small RNAs (20-30nts) were sequenced. Distribution of 22G siRNAs mapping antisense to *oma-1* is shown, with 22G siRNA

Figure 2.10 (Continued). reads normalized to reads per million total reads. pUG-specific: 22G siRNAs observed only after *oma-1(SNP)* pUG RNA injection; non-specific: 22G siRNAs observed after *oma-1(SNP)* pUG and pGC RNA injections. See Figure 2.11 for biological replicate and details about sequenced small RNAs.

To determine whether pUG RNAs act as templates for RdRPs *in vivo*, we sequenced small (20-30nt) RNAs from animals injected with either an *oma-1* pGC or pUG RNA engineered to contain a single-nucleotide polymorphism (SNP) not present in the genomic copy of *oma-1* (Figure 2.10d, termed *oma-1(SNP)* RNAs). This SNP enabled differentiation of siRNAs templated from genomically encoded *oma-1* mRNAs versus those templated from injected *oma-1(SNP)* pGC or pUG RNAs. In *C. elegans,* RdRP-derived (2°) siRNAs are also known as 22G siRNAs as they are typically antisense, 22nt in length and begin with a guanosine (Gu *et al.* 2009). Small RNA sequencing showed that injection of the *oma-1(SNP)* pUG RNA, but not the *oma-1(SNP)* pGC RNA, triggered the synthesis of *oma-1* 22G siRNAs mapping near (≅100bp upstream) the site where the pUG tail was appended (Figure 2.10d, Figure 2.11a) (Billi *et al.* 2014). For unknown reasons, both *oma-1(SNP)* pUG and pGC RNAs triggered non-specific siRNA synthesis ≅0.4kb upstream of where the tails were appended. Importantly, most (90-100%) pUG-specific 22G siRNAs antisense to the region of *oma-1* containing the engineered SNP encoded the complement of the SNP (Figure 2.10d, Figure 2.11a), indicating that these siRNAs were templated from the injected *oma-1(SNP)* pUG RNA. We conclude that one function of a pUG tail is to convert RNAs into templates for RdRPs.

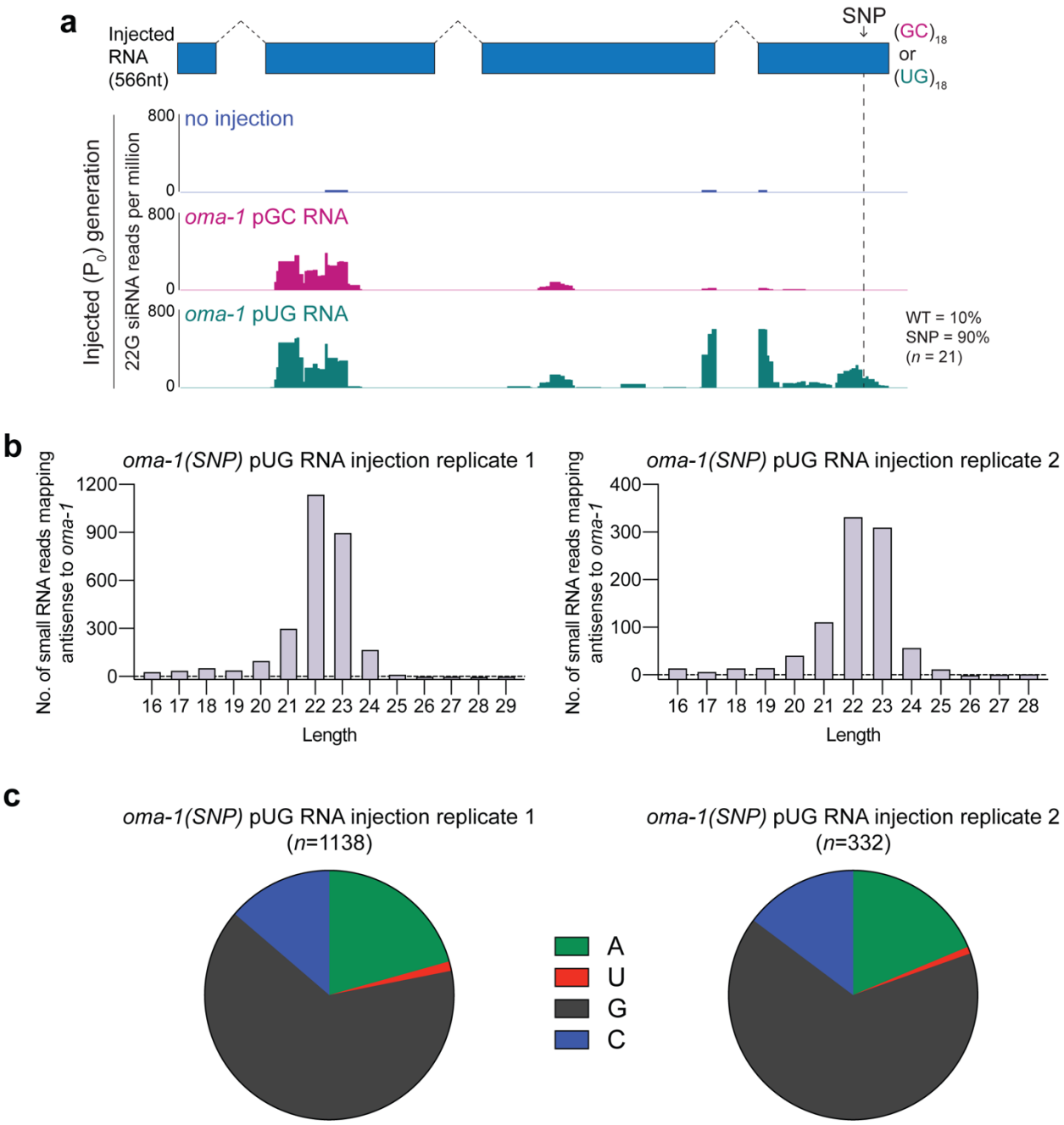


Figure 2.11. pUG RNAs are templates for 22G siRNA biogenesis. a, A biological replicate of the experiment shown in Figure 2.10d was performed. *oma-1(SNP)* pUG or pGC RNAs were injected into *rde-1(ne219); oma-1(zu405ts)* germlines. SNP location is indicated with the dotted line. Injected animals were collected 1-4 hours after injection, total RNA was isolated and small RNAs (20-30nt) were sequenced. Distribution of 22G siRNAs mapping antisense to *oma-1* is shown, with 22G siRNA reads normalized to reads per million total reads. *oma-1* pUG (but not pGC) RNA injection triggered 22G siRNA production near the site of the pUG tail ("pUG-specific" 22G siRNAs). For unknown reasons, both pUG and pGC RNA injections triggered small RNA production ≅400bp 5' of either tail. **b,** Length distribution of small RNA reads mapping antisense to *oma-1* is shown for small RNAs sequenced after *oma-1(SNP)* pUG RNA injections (Figures

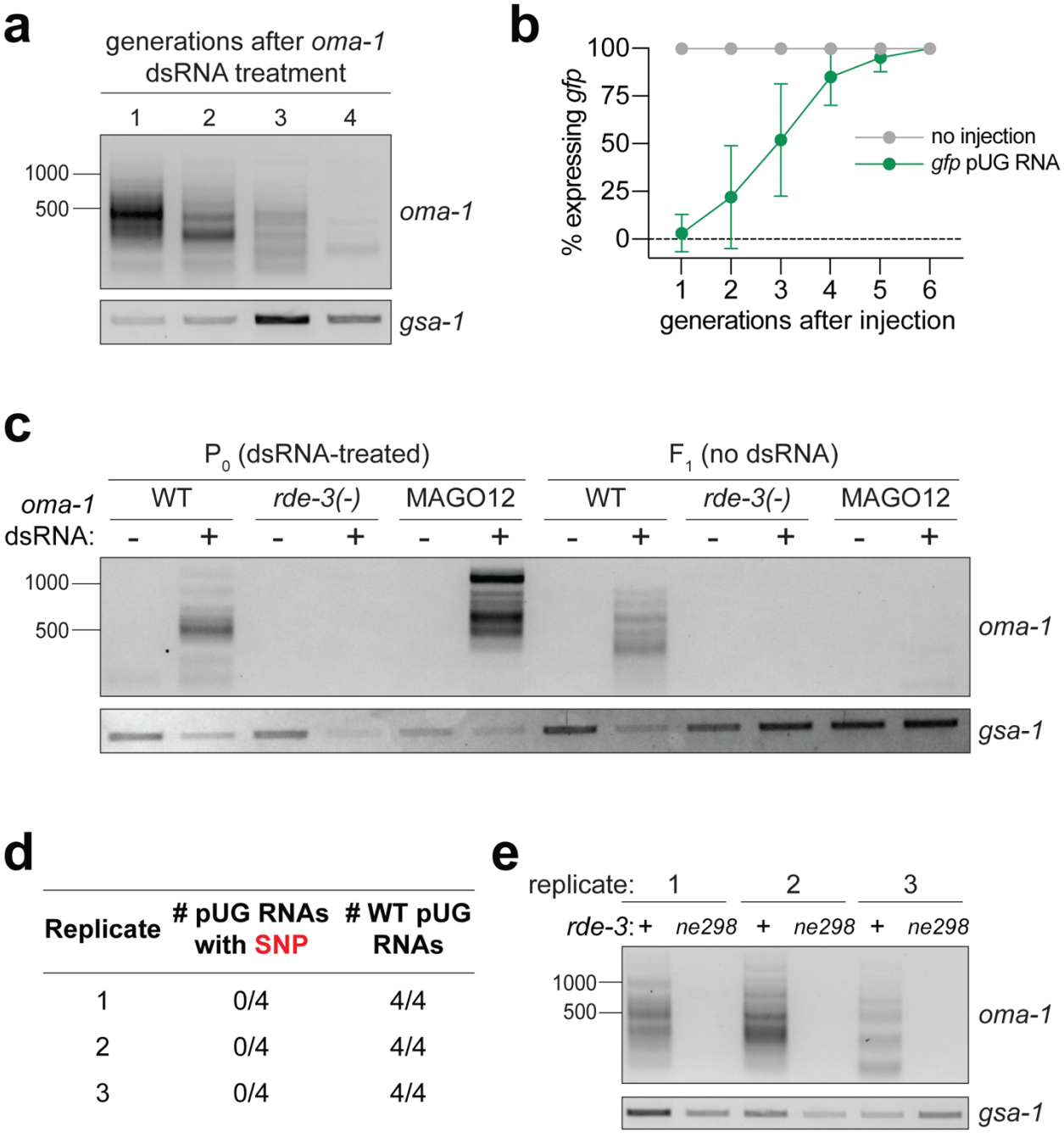
Figure 2.11 (Continued). 2.10d and **a**). **c,** Proportion of 22nt long small RNAs mapping antisense to *oma-1* containing 5' adenine, uracil, guanine, or cytosine is shown.

F. pUG RNAs are vectors for TEI

RNAi-triggered gene silencing can be inherited for multiple generations in *C. elegans,* making RNAi inheritance a robust and dramatic example of transgenerational epigenetic inheritance (TEI) (Vastenhouw *et al.* 2006; Ashe *et al.* 2012; Shirayama *et al.* 2012; Luteijn *et al.* 2012; Buckley *et al.* 2012; Sapetschnig *et al.* 2015). Interestingly, a one-time exposure of animals to *oma-1* dsRNA not only initiated the production of *oma-1* pUG RNAs (Figure 2.1b), but also caused *oma-1* pUG RNAs to be expressed for four additional generations (Figure 2.12a), concomitant with *oma-1* gene silencing (Figure 2.13a), suggesting that pUG RNAs may contribute to TEI. To test this idea, we injected animals with *gfp* or *oma-1* pUG RNAs and monitored *gfp* or *oma-1* silencing over generations. *gfp* or *oma-1* pUG RNAs were sufficient to silence *gfp* (Figure 2.12b) or *oma-1* (Figure 2.13b), respectively, for multiple generations. We conclude that pUG RNAs are sufficient to induce TEI.

How might pUG RNAs drive TEI? We speculated that if pUG RNA–templated siRNAs (Figure 2.10d, Figure 2.11a) could direct *de novo* mRNA pUGylation, then generationally repeated cycles of pUG RNA–templated siRNA synthesis and siRNA-directed pUG RNA biogenesis could be maintained in the absence of initiating dsRNA triggers and, thus, propagate gene silencing across generations. Three lines of evidence support this "pUG/siRNA cycling" model for RNAi-directed TEI. First, RdRP-derived 2° siRNAs in *C. elegans* can engage twelve AGO proteins (termed WAGOs) to mediate gene silencing (Yigit *et al.* 2006; Gu *et al.* 2009). MAGO12 animals, which harbor deletions in all twelve WAGOs, produced *oma-1* pUG RNAs after *oma-1* RNAi (Figure 2.12c). Progeny of RNAi-treated MAGO12 animals, however, failed to produce *oma-1* pUG RNAs (Figure 2.12c). Thus, the 2° siRNA system is needed to maintain pUG RNAs specifically during the inheriting generations of TEI, consistent with a pUG/siRNA cycling model for TEI. Interestingly, pUG RNAs derived from endogenous pUGylation targets *c38d9.2* and Tc1

were also dependent upon the WAGOs (Figure 2.13c), suggesting that the endogenous targets of RDE-3 also undergo heritable silencing via pUG/siRNA cycling.



Replicate	# pUG RNAs with SNP	# WT pUG RNAs
1	0/4	4/4
2	0/4	4/4
3	0/4	4/4

Figure 2.12. pUG RNA/siRNA cycles drive heritable gene silencing. a, *oma-1* pUG PCR performed on total RNA from descendants of *oma-1* dsRNA-treated animals. **b,** *rde-1(ne219); gfp::h2b* animals were injected with *gfp* pUG RNA, and *gfp* expression was monitored for six generations. n=3 (no injection), 9 (*gfp* pUG RNA). Error bars: s.d. of the mean. **c,** *oma-1* pUG PCR performed on total RNA from *oma-1* dsRNA-treated (P_0) animals of indicated genotypes and their progeny (F_1). Note: pUG RNAs appear longer in MAGO12 animals (see Figure 2.14). **d,** pUG

Figure 2.12 (Continued). RNAs were Sanger sequenced from F_2 progeny of *rde-1(ne219); oma-1(zu405ts)* animals injected with *oma-1(SNP)* pUG RNA. **e,** *oma-1(zu405ts)* hermaphrodites were fed *oma-1* dsRNA and crossed to *rde-3(ne298); oma-1(zu405ts)* males (3 biologically independent crosses). *oma-1* pUG PCR was performed on total RNA from *rde-3(+)* or *rde-3(ne298)* F_3 progeny. **a, c, d, e.** Data is representative of 3 biologically independent experiments.

Second, when we injected animals with an *oma-1(SNP)* pUG RNA, *oma-1* pUG RNAs were detectable in subsequent generations (Figure 2.13d), but did not contain the engineered SNP (Figure 2.12d). Similarly, <1% of siRNAs sequenced from progeny of *oma-1(SNP)* pUG RNA injected animals possessed the SNP complement (Figure 2.13e). Combined, these data show *de novo* pUGylation events occur during the inheriting generations of RNAi-directed TEI and these newly derived pUG RNAs become templates for further siRNA synthesis, supporting the idea that repeated pUG/siRNA cycling mediates TEI**.**

Third, we conducted a genetic analysis that showed that these *de novo* pUGylation events in inheriting generations are required for TEI. We crossed *oma-1* RNAi–treated wild-type hermaphrodites with *rde-3(ne298)* males, isolated *rde-3(+)* and *rde-3(ne298)* F_2 progeny, and then assayed the F_3 generation of this cross (Figure 2.12e) for *oma-1* pUG RNA expression and *oma-1* gene silencing. *rde-3(ne298)* animals lacked *oma-1* pUG RNAs (Figure 2.12e) and failed to silence the *oma-1* locus (Figure 2.13f) during the inheriting generations of TEI, supporting the idea that pUG RNA biogenesis and, therefore, pUG/siRNA cycling in progeny is necessary for TEI maintenance. We conclude that pUG tails convert otherwise inert RNA fragments into drivers of an RNA-based memory system, which is likely propagated across generations via iterative cycles of sense pUG RNA and antisense siRNA biogenesis.

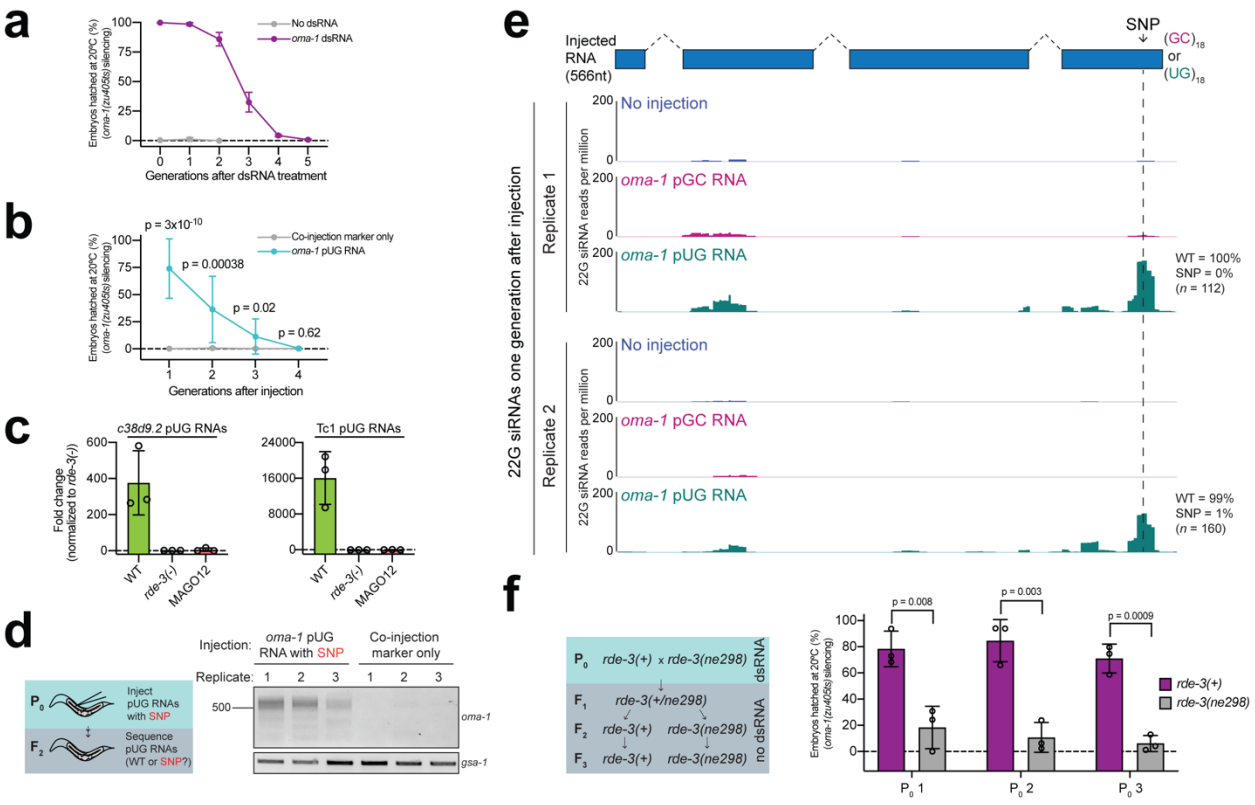


Figure 2.13. *de novo* pUGylation events in progeny are required for TEI. a, *oma-1(zu405ts)* animals were fed bacteria expressing empty vector control or *oma-1* dsRNA and % embryos hatched at 20°C was scored for 6 generations. Error bars represent s.d. of the mean of three biologically independent experiments. For each experiment, % embryos hatched at 20°C was averaged for 6 individual animals per treatment for each genotype. **b,** *rde-1(ne219); oma-1(zu405ts)* animals were injected with co-injection marker alone (n=12) or co-injection marker + *oma-1* pUG RNA (n=19) and % embryos hatched at 20°C was scored for four generations in lineages of animals established from injected parents (see Methods for details of experimental setup). Error bars: s.d. of the mean. *p*-values: two-tailed unpaired Student's *t*-test. **c,** *c38d9.2* and Tc1 pUG RNA expression quantified in embryos harvested from wild-type, *rde-3(-)* or MAGO12 animals using qRT-PCR. Fold change normalized to *rde-3(-).* Each point (n) represents a biologically independent replicate, n=3 independent replicates/strain. Error bars: s.d. of the mean. **d,** Same experiment as Figure 2.12d. *rde-1(ne219); oma-1(zu405ts)* animals were injected with an *oma-1(SNP)* pUG RNA or with co-injection marker only. Co-injection marker-expressing F_1 progeny were picked and allowed to lay their F_2 broods. *oma-1* pUG PCR was performed on total RNA from F_2 progeny. Shown is data from three biological replicates. **e,** Two biological replicates of small RNAs sequenced from the progeny of *rde-1(ne219); oma-1(zu405ts)* animals injected with *oma-1(SNP)* pUG or pGC RNAs are shown. Dotted line indicates the location of the SNP incorporated into *oma-1*. Distribution of 22G siRNAs mapping antisense to *oma-1* is shown, with 22G siRNA reads normalized to reads per million total reads. In Figure 2.10d and Figure 2.11a, small RNAs were sequenced 1-4 hours after injection and 100% of 22G siRNAs antisense to the region of the engineered SNP in *oma-1* were found to encode the complement of the SNP. Shown here, <1% of 22G siRNAs from progeny of injected animals encoded the SNP complement. Note: siRNAs mapping near the pUG tail were observed only after *oma-1(SNP)* pUG RNA injection (pUG-specific siRNAs). For unknown reasons, both *oma-1(SNP)* pUG and pGC RNAs triggered small RNA production 5' of the pUG-specific siRNAs. It is possible that these siRNAs were

Figure 2.13 (Continued). triggered by systems that respond to foreign RNAs, such as the piRNA system. Further work will be needed to ascertain the etiology of these siRNAs. **f,** Same experiment as Figure 2.12e. *oma-1(zu405ts)* hermaphrodites were fed *oma-1* dsRNA and crossed to *rde-3(ne298); oma-1(zu405ts)* males. F_2 progeny from this cross were genotyped for *rde-3(ne298)*. WT and *rde-3(ne298)* homozygous F_3 progeny were phenotyped for % embryonic arrest at 20°C. 3 biologically independent crosses (P_0 1-3) were performed. Error bars: s.d. of the mean. *p*-values: two-tailed unpaired Student's *t*-tests.

IV. Discussion

Here we show that RDE-3 adds pUG tails to germline- and soma-expressed RNAs in *C. elegans* and also demonstrate a role for this modification in transposon silencing and TEI. We find that RdRPs are recruited, either directly or indirectly, to pUG tails and use pUG RNAs as templates for siRNA synthesis. Assemblage of RDE-3 (Phillips *et al.* 2012) and other proteins, like the endonuclease RDE-8 (Tsai *et al.* 2015) and the RdRP RRF-1 (Phillips *et al.* 2012), into germline condensates termed *Mutator* foci likely coordinates RNA target recognition, cleavage, pUGylation, and siRNA amplification (Figure 2.14). We find that functional pUG tails consist of more than eight pairs of perfect or near-perfect 3' UG repeats. These precise length and sequence requirements for pUG tail function hint that long pUG tails may impart stability upon mRNA fragments and/or form a structure which helps to recruit, and possibly prime, RdRPs, similar to the proposed role for poly(U)-tailing in small RNA–based gene silencing in *Tetrahymena* (Talsky and Collins 2010). Additionally, our data show that proteins other than RdRPs, such as TDP-1, the *C. elegans* ortholog of the mammalian UG-binding protein TDP-43 (Buratti and Baralle 2001; Kuo *et al.* 2009), also interact with pUG repeats. We speculate that these other proteins may regulate the localization, stability, or function of pUG-tailed RNAs. Note: Figure 2.14 relates our findings to a previous report suggesting RDE-3 may uridylate targets of RNAi (Tsai *et al.* 2015).

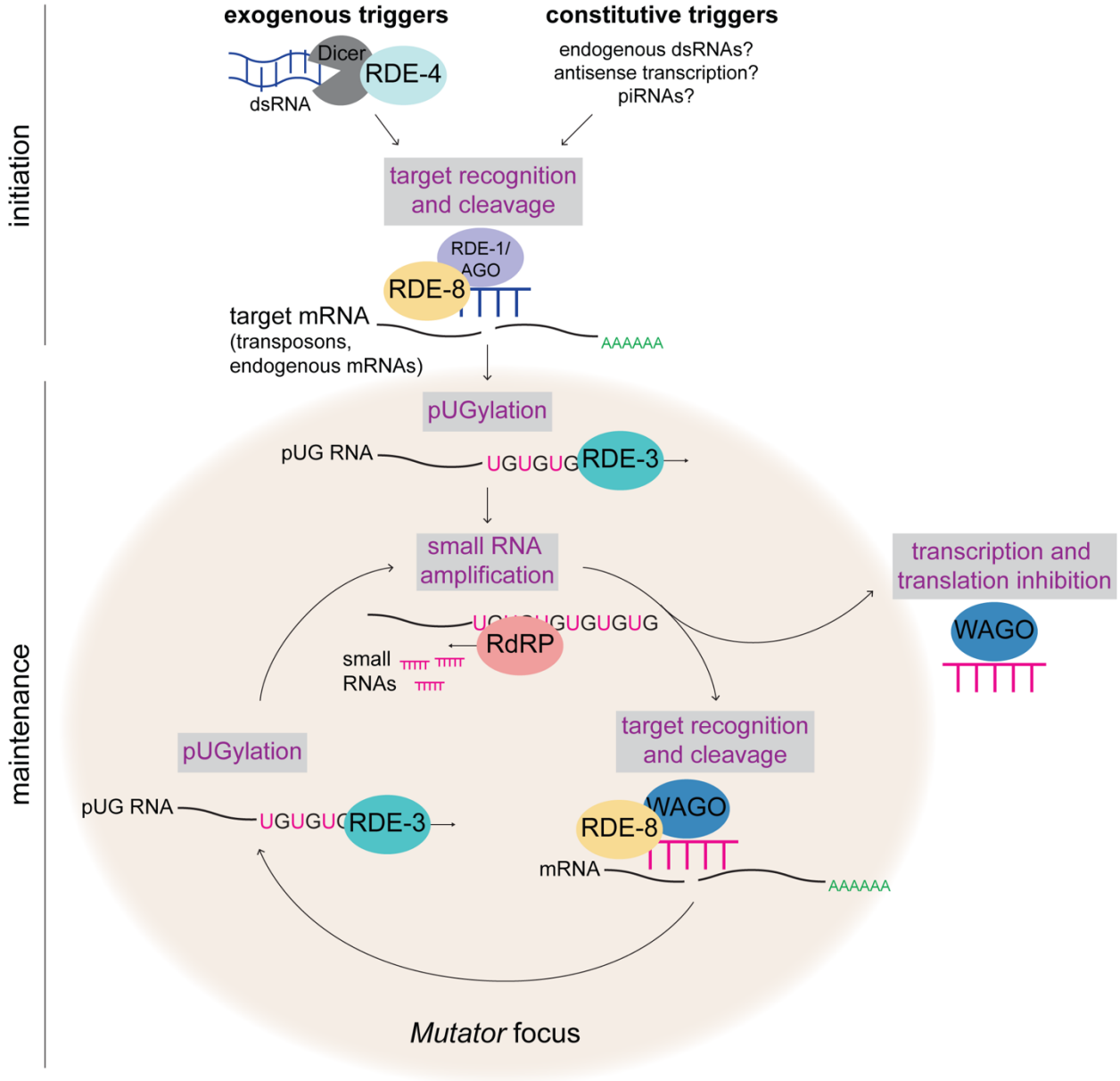


Figure 2.14. Working model for pUG RNA/siRNA cycling during RNAi. *Initiation*: exogenous and constitutive (i.e. genomically encoded such as dsRNA, piRNAs) triggers direct RDE-3 to pUGylate RNAs previously fragmented by factors in the RNAi pathway. *Maintenance*: pUG RNAs are templates for 2° siRNA synthesis by RdRPs. Argonaute proteins (termed WAGOs) bind 2° siRNAs and: 1) target homologous RNAs for transcriptional and translational silencing (Yigit *et al.* 2006; Guang *et al.* 2008, 2010; Buckley *et al.* 2012), as well as 2) direct the cleavage and *de novo* pUGylation of additional mRNAs. In this way, cycles of pUG RNA-based siRNA production and siRNA-directed mRNA pUGylation maintain silencing over time and across generations. This model shows germline perinuclear condensates termed *Mutator* foci as the likely sites of pUG RNA biogenesis in germ cells for several reasons. RDE-3 localizes to *Mutator* foci (Phillips *et al.* 2012) and we show in Figure 2.7d that endogenous pUG RNAs localize to *Mutator* foci. The fact that enzyme and enzyme product both localize to *Mutator* foci suggests that *Mutator* foci may be sites of RNA pUGylation. In addition, while pUG RNAs are still made in *mut-16* mutants (Figure 2.9b-d), which lack *Mutator* foci, the levels of both dsRNA-triggered and endogenous pUG RNAs

Figure 2.14 (Continued). are misregulated. Thus, while RDE-3 still has enzymatic activity in the absence of *Mutator* foci, these perinuclear condensates are likely coordinating target recognition and pUGylation in wild-type animals. Indeed, both the endonuclease RDE-8, which cleaves mRNAs targeted by dsRNA (Tsai *et al.* 2015), and the RdRP RRF-1 (Phillips *et al.* 2012) also localize to *Mutator* foci, further suggesting that pUG RNA/siRNA cycling occurs in *Mutator* foci.

Previous studies have shown that animals lacking RDE-3 still produce some 22G endo-siRNAs, including 22G siRNAs that associate with the Argonaute CSR-1 and whose biogenesis depends upon the RdRP EGO-1 (Gu *et al.* 2009; Zhang *et al.* 2011). Thus, EGO-1 may also produce some 22G siRNAs via a pUG RNA-independent mechanism.

A previous study showed that, in *rrf-1* mutants that lack germlines, *sel-1* RNAi causes a small fraction of *sel-1* mRNA fragments to be uridylated in a largely RDE-3–dependent manner in the soma (Tsai *et al.* 2015). This data suggests that, in somatic tissues, RDE-3 may add non-templated Us to the 3' termini of mRNA fragments generated during RNAi. It was proposed that this uridylation may be important for turnover or decay of RNAi targets (Tsai *et al.* 2015). Our work, combined with this earlier data about RDE-3–dependent uridylation (Tsai *et al.* 2015), suggests two models. First, RDE-3 may possess two distinct catalytic activities: uridylation and pUGylation. According to this model, RDE-3 might add Us or UGs depending on context (e.g. cell/tissue-type or developmental timing). Alternatively, the mRNA uridylation observed in the soma could depend upon RDE-3 and the pUGylation system, but may be mediated by another, currently unknown, poly(U) polymerase.

Further, our data show that pUG RNAs act as informational vectors for TEI when they engage in feed-forward amplification cycles with RdRP-generated 2° siRNAs (Figure 2.14). These pUG/siRNA cycles, we speculate, allow *C. elegans* to remember past gene silencing events and inoculate progeny against expressing unwanted and/or dangerous genetic elements. Experimental RNAi-initiated pUG/siRNA cycles perdure for several generations (Figure 2.12a), but are not permanent, suggesting that *C. elegans* possesses systems to prevent pUG/siRNA cycles from propagating in perpetuity. Interestingly, we find that RNAi-initiated pUG RNAs shorten progressively during TEI, suggesting that pUG RNA shortening, which may be an inevitable consequence of RdRP-based 2° siRNA synthesis (Figure 2.15), could function as one such brake on TEI. In contrast, the natural targets of pUGylation, such as transposons, are constitutively silenced by the pUG/siRNA system, suggesting that genetic systems, such as genomically encoded PIWI-interacting RNAs or endogenous dsRNAs, likely reinforce and refocus epigenetic pUG/siRNA silencing at these loci each generation (Figure 2.14).

The logic of sense pUG RNA/antisense siRNA cycling resembles that of fly and mammalian piRNA "ping-pong" systems in which iterative base-pairing between genomically

encoded sense/antisense transposon RNAs, and piRNAs derived from these RNAs, mediates stable transposon silencing (Czech *et al.* 2018). We speculate that related sense/antisense RNA systems could contribute to other biological processes during which long-term memories of past expression states are needed, such as antiviral immunity, development, or inheritance of environmentally acquired traits. Finally, our data show that long non-templated and non-homopolymeric tracts of ribonucleotides can be appended to, and confer novel functions to, RNAs in *C. elegans*. It will be of obvious interest to ask whether pUG-tailed RNAs, or RNAs bearing other unexpected tails, are restricted to *C. elegans* or are, instead, emissaries of a new class of eukaryotic RNA.

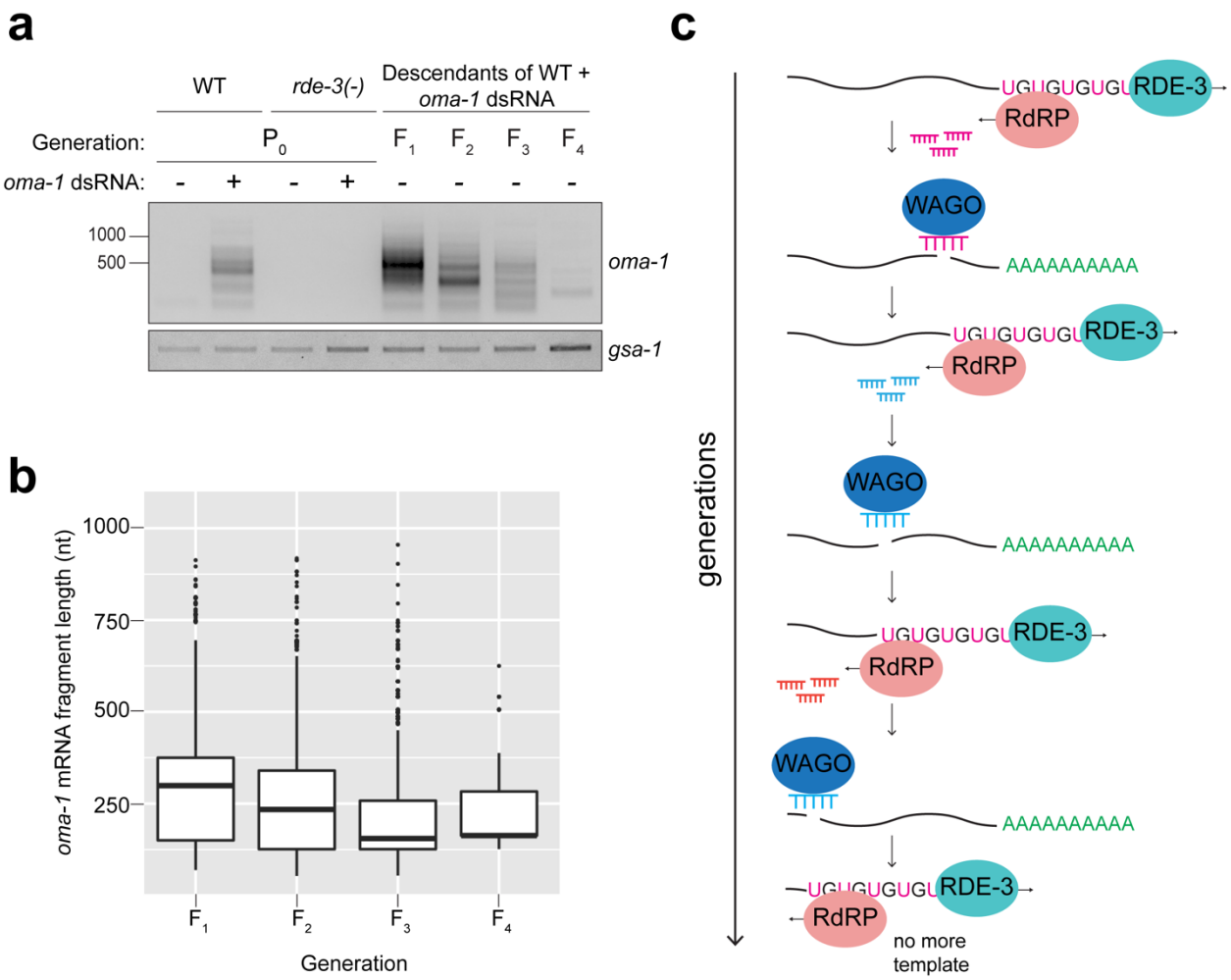


Figure 2.15. pUG RNA shortening may act as a brake on TEI. **a,** The gel shown is the same as in Figure 2.12a, except that *oma-1* pUG RNAs from the P_0 generation are included for WT and *rde-3(-)* animals. Data is representative of 3 biologically independent experiments. **b,** *oma-1* pUG RNA reads from MiSeq (n=1 biological experiment) were mapped to *oma-1* and the length of the *oma-1* mRNA portion of each pUG RNA read was determined (y-axis). Shown is a Box and Whisker plot representing the interquartile range (IQR, box) and median (line in the box) of lengths at the indicated generations after dsRNA treatment. The y-axis starts at the *aug* of the *oma-1* mRNA. The whiskers extend to values below and above 1.5*IQR from the first and third quartiles, respectively. Data beyond the end of the whiskers are outliers and plotted as points. These data support the gel in **a**, showing that pUG RNAs get shorter in each generation during RNAi-triggered TEI. **c,** A "ratchet" model to explain pUG RNA shortening. pUG RNA shortening may be due to the 3'→5' directionality of RdRPs, which, during the maintenance phase of pUG/siRNA cycling (see model in Figure 2.14), causes each turn of the pUG/siRNA cycle to trigger cleavage and pUGylation of target mRNAs at sites more 5' than in the previous cycle. Eventually, pUG RNAs are too short to act as RdRP templates, cycling cannot be maintained and silencing ends. Additional support for the ratchet model comes from Figure 2.12c, which shows that RNAi-triggered pUG RNAs are longer in MAGO12 mutant animals than in wild-type animals. Note: the P_0 generation animals in Figure 2.12c were exposed to dsRNA continuously from embryos to adulthood, when they were harvested. These longer pUG RNAs are likely due to continued

Figure 2.15 (Continued). initiation of pUGylation triggered by the exogenously provided dsRNA without downstream pUG/siRNA cycling. In the absence of this cycling, pUG RNA shortening does not occur. Finally, a number of recent studies in *C. elegans* have reported transgenerational inheritance of acquired traits, which lasts 3-4 generations (Remy 2010; Rechavi *et al.* 2014; Schott *et al.* 2014; Jobson *et al.* 2015; Moore *et al.* 2019; Posner *et al.* 2019). As shown in **a,** the expression of *oma-1* RNAi–directed pUG RNAs also perdures for 3-4 generations. These shared generational timescales of inheritance hint that the inheritance of acquired traits in *C. elegans* may be mediated by pUG RNAs whose generational "half-life" is limited to 3-4 generations due to the built-in brake on TEI provided by pUG RNA shortening.

V. Methods

A. Genetics

C. elegans culture and genetics were performed as described previously (Brenner 1974). Unless otherwise noted, all *C. elegans* strains (Table 4) were maintained at 20°C on NGM growth media and fed OP50 *E. coli* bacteria.

Table 2.4. *C. elegans* strains used in this study.

Strain Name	Genotype	Source
N2	wild-type	CGC
YY1449	*rde-3(ne3370) I*	2X outcross of WM286 with N2
YY1505	*rde-3(ne298) I*	2X outcross of WM30 with N2
YY1711	*rde-3(ne3370) I; rde-3(ggSi19) II*	RDE-3 + 2kb upstream and downstream sequence integrated using CRISPR into the LGII MosSCI site (*ttTi5605*. See wormbuilder.org for more information on MosSCI.
YY1670	*rde-3(gg679)*	Reversion of *rde-3(ne298)* mutation using oligo-mediated CRISPR.
tm1396	*oma-1(tm1396) IV*	Mitani Lab
YY1569	*rde-3(ne298) I; oma-1(zu405) IV*	
YY1635	*rde-3(ne298) I; oma-1(zu405) IV; rde-1(ne219) V*	
SX461	*mjIs31[pie-1::gfp::h2b] II*	Miska Lab
YY1332	*rde-3(ne3370) I; mjIs31[pie-1::gfp::h2b] II*	
NL1810	*mut-16(pk710) I*	CGC
WM27	*rde-1(ne219) V*	CGC

WM49	*rde-4(ne301) III*	CGC
	rde-8(tm2252) IV	Mello Lab
JK4605	*glp-1(q224) I*	CGC
YY1808	*rde-8(tm2252) oma-1(zu405) IV; rde-1(ne219) V*	introduced *zu405* allele into *oma-1* locus using oligo-mediated CRISPR in an *rde-8(tm2252); rde-1(ne219)* double mutant
WM191	*sago-2(tm894) ppw-1(tm914) ppw-2(tm1120) wago-2(tm2686) wago-1(tm1414) I; wago-11(tm1127) wago-5(tm1113) wago-4(tm1019) II; hrde-1(tm1200) sago-1(tm1195) III; wago-10(tm1186) V; nrde-3(tm1116) X*	CGC
YY1818	*unc-22(st136) IV*	derived from >8X outcross of NL3643
YY1671	*rde-3(ne3370) I; oxTi956 III; unc-22(st136) IV*	
YY1471	*oma-1(zu405) IV; rde-1(ne219) V*	
OD56	*unc-119(ed3) III; ltIs37 [(pAA64) pie-1p::mCherry::his-58 + unc-119(+)] IV*	CGC
YY1920	*unc-119(ed3) III?; ltIs37 IV; c38d9.2(gg746) V*	*unc-119(ed3)* status unknown
YY1921	*rde-3(ne298) I; unc-119(ed3) III?; ltIs37 IV; c38d9.2(gg746) V*	*unc-119(ed3)* status unknown
YY1702	*rde-3(gg693) I*	*gfp::degron::rde-3* at endogenous *rde-3* locus
YY1568	*mjIs31[pie-1::gfp::h2b] II; rde-1 (ne219) V*	
YY1635	*rde-3(ne298) I; oma-1(zu405) IV; rde-1(ne219) V*	
YY1664	*mut-16(pk710) I ; oma-1(zu405) IV; rde-1(ne219) V*	
YY1978	*rrf-1(gg794) I*	*ha::TagRFP* tag on C-terminus of endogenous *rrf-1* locus

B. RNAi

To perform RNAi experiments, embryos were obtained via hypochlorite treatment (egg prep) of gravid adult hermaphrodites and dropped onto RNAi plates (standard NGM plates with 1mM isopropyl β-D-1-thiogalactopyranoside and 25μg/ml carbenicillin) seeded with HT115 *E. coli*

bacteria expressing either L4440 (Addgene, #1654) empty vector control or L4440 carrying inserts to trigger the production of dsRNA against a gene of interest. To perform pUG RNA analysis after RNAi treatment, gravid adults were washed off plates after 3-4 days using M9 + Triton X-100 buffer, collected in TRIzol, flash frozen in liquid nitrogen and stored at -80°C until total RNA extraction (see below). To look at *oma-1* pUG RNAs across generations after RNAi, embryos were dropped onto plates seeded with HT115 bacteria expressing either empty vector control or dsRNA of interest. Some gravid adults were collected for the P_0 generation sample and the remaining were egg prepped onto plates without dsRNA every generation, for the indicated number of generations. Each generation, some adult animals were collected while some were egg prepped to obtain the next generation. pUG RNAs were then detected as described below. To measure % embryos hatched after *oma-1* RNAi, embryos obtained from animals harboring the *oma-1(zu405ts)* allele (Lin 2003) were dropped onto plates seeded with HT115 bacteria expressing either empty vector control or *oma-1* RNAi and grown at 20°C. 6 adults were then singled per treatment for each strain/genotype and allowed to lay embryos overnight. The total number of embryos laid was counted, and then embryos were allowed to hatch for 24 hours, after which the total number of embryos that hatched was counted. For transgenerational RNAi experiments, empty vector control-treated and *oma-1* RNAi–treated adults were egg prepped onto plates without dsRNA every generation and % embryos hatched was counted as just described until embryos no longer hatched. The *dpy-11* RNAi clone and the *oma-1* RNAi clone used, unless noted below, came from the *C. elegans* RNAi collection (Dr. Julie Ahringer & Source BioScience). The second *oma-1* RNAi clone (referred to as pAS74 and used for Figures 2.2b, 2.12a, 2.13a, 2.15a, 2.15b) was a custom clone made to target exon 6 of *oma-1*. The *gfp* RNAi clone was obtained from the Fire lab.

C. pUG PCRs and qRT-PCRs

Total RNA was extracted using TRIzol Reagent (Life Technologies, 15596018). 5ug of total RNA and 1pmol of reverse transcription (RT) oligo was used to generate first-strand cDNA using the Superscript III First-Strand Synthesis System (Invitrogen, 18080051). Note: total RNA was heated with dNTPs and RT oligo to 65°C for 5 mins and immediately chilled on ice before proceeding with remaining cDNA synthesis steps. 1ul of cDNA was used for the first PCR (20ul volume) performed with Taq DNA polymerase (New England BioLabs, M0273) and primers listed in Table 5. First PCR reactions were diluted 1:100, and then 1ul was used for a second PCR (50ul volume) using primers listed in Table 5. *gsa-1,* which has an 18nt long genomically encoded pUG repeat in its 3'UTR, served as a control for all pUG PCR analyses. PCR reactions were then run on agarose gels. For Sanger sequencing, lanes of interest were cut out from agarose gels and gel extracted using a QIAquick Gel Extraction Kit (Qiagen, 28706). 3ul of gel extracted PCR product was used for TA cloning with the pGEM-T Easy Vector System (Promega, A1360) according to manufacturer's instructions. Ligation reactions were incubated overnight at 4°C. Transformations were performed with 5-alpha Competent *E. coli* cells (NEB, C2987H) and plated on LB/ampicillin/IPTG/X-gal plates (according to pGEM-T Easy Vector System manufacturer's instructions). On the next day, white colonies were selected and inoculated, and then liquid cultures were miniprepped using QIAprep Spin Miniprep Kit (Qiagen, 27106). Plasmids were Sanger sequenced using a universal SP6 primer (5'-CATACGATTTAGGTGACACTATAG-3') (Dana-Farber/Harvard Cancer Center DNA Resource Core, Harvard Medical School). qRT-PCRs were performed using 2ul of 1:100 diluted first PCRs as a template with qPCR primers (Table 5) and iTaq Universal SYBR Green Supermix (Bio-Rad) according to manufacturer's instructions.

Table 2.5. Oligonucleotides used in this study.

pUG PCR and qPCR primers	
Name	**Sequence (5'-3')**
adapter 1 specific reverse	GCT ATG GCT GTT CTC ATG GC
adapter 2 specific reverse	GGC GTC GCC ATA TTC TAC TT
oma-1 pUG PCR forward 1	ATG AAC GTT AAC GGT GAA AAC
oma-1 pUG PCR forward 2	AAC AAC GAG AAG ATC GAT GA
oma-1 pUG PCR miSeq forward	AAT GAT ACG GCG ACC ACC GAG ATC TAC ACT CTT TCC CTA CAC GAC GCT CTT CCG ATC TNN NAA CAA CGA GAA GAT CGA TGA
oma-1 pUG PCR miSeq reverse	CAA GCA GAA GAC GGC ATA CGA GAT (6nt index) GTG ACT GGA GTT CAG ACG TGT GCT CTT CCG ATC TNN NGG CGT CGC CAT ATT CTA CTT
oma-1 pUG qPCR forward	AGG AAC TTC GTC CAA CAT TCG
oma-1 pUG qPCR reverse	TTC CCG TAT GGG CAG AGT
Tc1 pUG PCR forward 1	ATG GTA AAA TCT GTT GGG TG
Tc1 pUG PCR forward 2	CGTTCTCCGTCGACTATTTGG
Tc1 pUG qPCR forward	CAA GAA GTA CCA AAC TGA GAA ATC C
Tc1 pUG qPCR reverse	GGA TGT TGC GAT CCA TCC TAT
gsa-1 pUG PCR forward 1	GAG TTC TAC GAT CAC ATT CT
gsa-1 pUG PCR forward 2	CAC TTG CTG GAA AGA CAA GG
dpy-11 pUG PCR forward 1	ATG CTG CTC CGA TTG CTC GC
dpy-11 pUG PCR forward 2	CGT GCT CGG ACT TTT CGC CG
gfp pUG PCR forward 1	ATG AGT AAA GGA GAA GAA CT
gfp pUG PCR forward 2	TTTCACTGGAGTTGTCCCAA
c38d9.2 pUG PCR forward 1	AAC AAC GAG AAA AGG CCG GA
c38d9.2 pUG PCR forward 2	GAG CCA ACT CAA GAA GCT GG
c38d9.2 pUG qPCR forward	TTA CCA ACG AAA GGA GGA TGA A
c38d9.2 pUG qPCR reverse	GGG TCT TCT CTT GTC TCT TTC TC
f15d4.5 pUG PCR forward 1	AAC AAC GAG AAA AGG CCG GA
f15d4.5 pUG PCR forward 2	AAA CCG ACT CAA GAA TCT GG
f49f1.8 pUG PCR forward 1	ATG GCT GAC AAA GAA GAG CC
f49f1.8 pUG PCR forward 2	GCC ACC ACC CGG TAC TCT GC
w09d7.1 pUG PCR forward 1	ATG GGA GGC AGA AAG TCC AA
w09d7.1 pUG PCR forward 2	CAA AGC AAA AAA TCA ACA AC
k09h11.11 pUG PCR forward 1	ATG AAT TCG GGA TAC GAA AT
k09h11.11 pUG PCR forward 2	ACG AAA TGC CAA ATA ATT TG
y17d7c.3 pUG PCR forward 1	ATG GCA ATT GTC CGG GAA TG

y17d7c.3 pUG PCR forward 2	CAA ATT GTA CAC ATT CAC AC
t04h1.9 pUG PCR forward 1	ATG AAA GAA ATT ATT AAC GT
t04h1.9 pUG PCR forward 2	TCA AGT TGG ACA ATG TGG AA
tc4v pUG PCR forward 1	TGTCGTCATTCGCAACATTC
tc4v pUG PCR forward 2	CGCCATGGAAGAGCTTGA
tc5 pUG PCR forward 1	CCAAAGTGGGTTCACGAAAG
tc5 pUG PCR forward 2	GTGGTTCTGCAGGAGAAAGG
cer3 pUG PCR forward 1	ATG TCC CGC AAC CCG CAG TT
cer3 pUG PCR forward 2	CAA TAC ACC CCC ACC CAA TG

oligos to amplify DNA templates for T7 reactions	
Name	**Sequence (5'-3')**
Tc1 pUG RNA T7 template forward	TAA TAC GAC TCA CTA TAG GGA GAA TGG TAA AAT CTG TTG GGT G
Tc1 pUG RNA T7 template reverse	CAC ACA CAC ACA CAC ACA CAC ACA CAC ACA CAC ACA TGC TGT AAA CGT CGA CGA AC
*gfp*pUG RNA T7 template forward	TAATACGACTCACTATAGGGAGACCAC ATGAGTAAAGGAGAAGAACT
*gfp*pUG RNA T7 template reverse	CACACACACACACACACACACACACACAC ACACA GATTCTATTAACAAGGGTAT
*gfp*pAU RNA T7 template reverse	ATATATATATATATATATATATATATATAT AT GATTCTATTAACAAGGGTAT
*gfp*pGC RNA T7 template reverse	CGCGCGCGCGCGCGCGCGCGCGCGCGC GCGCGCG GATTCTATTAACAAGGGTAT
*gfp*pCA RNA T7 template reverse	GTGTGTGTGTGTGTGTGTGTGTGTGTG TGTGT GATTCTATTAACAAGGGTAT
oma-1 pUG RNA T7 template forward	TAATACGACTCACTATAGGGAGACCAC ATGAACGTTAACGGTGAAAAC
oma-1 pUG RNA T7 template reverse	CACACACACACACACACACACACACACAC ACACA CGGGATGAATGAAAAGGCAC
oma-1(SNP) pUG RNA T7 template reverse	CACACACACACACACACACACACACACAC ACACA CGAATGTATGCGTTTGGTCCATGAACGGGAT GAATGAAAAGGCACCG
oma-1 pAU RNA T7 template reverse	ATATATATATATATATATATATATATATAT AT CGGGATGAATGAAAAGGCAC
oma-1 pGC RNA T7 template reverse	CGCGCGCGCGCGCGCGCGCGCGCGCGC GCGCGCG CGGGATGAATGAAAAGGCAC

oma-1(SNP) pGC RNA T7 template reverse	CGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG CGAATGTATGCGTTTGGTCCATGAACGGGAT GAATGAAAAGGCACCG
oma-1 pCA RNA T7 template reverse	GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT CGGGATGAATGAAAAGGCAC
oma-1 $p(UG)_1$ RNA T7 template reverse	CA CGGGATGAATGAAAAGGCAC
oma-1 $p(UG)_5$ RNA T7 template reverse	CACACACACA CGGGATGAATGAAAAGGCAC
oma-1 $p(UG)_8$ RNA T7 template reverse	CACACACACACACACA CGGGATGAATGAAAAGGCAC
oma-1 $p(UG)_{14}$ RNA T7 template reverse	CACACACACACACACACACACACACACA CGGGATGAATGAAAAGGCAC
oma-1 $p(UG)_{40}$ RNA T7 template reverse	CACACACACACACACACACACACACACACACAC ACACACACACACACACACACACACACACACA CACACACACACACACACA CGGGATGAATGAAAAGGCAC
oma-1 $p(U)_{18}(G)_{18}$ RNA T7 template reverse	CCCCCCCCCCCCCCCCCCAAAAAAAAAAAAAA AAAAA CGGGATGAATGAAAAGGCACCG
oma-1 scrambled UG RNA T7 template reverse	CCCAAAAAACAAACCCACCCAAACCCAACCC CAAAC CGGGATGAATGAAAAGGCACCG
oma-1 $p(UUGG)_9$ RNA T7 template reverse	CCAACCAACCAACCAACCAACCAACCAACCA ACCAA CGGGATGAATGAAAAGGCACCG
oma-1 5' pUG tail T7 template forward	TAATACGACTCACTATAGGGAGACCAC GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG TGTGT ATGAACGTTAACGGTGAAAAC
oma-1 internal pUG T7 template forward	GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG TGTGT CACCTCCATTTCAACAGTTTG
oma-1 internal pUG T7 template reverse	CACACACACACACACACACACACACACACACAC ACACA CAACCTGGTTGTTCTGGGGAA
antisense *oma-1* pUG RNA T7 template forward	TAATACGACTCACTATAGGGAGACCAC CGGGATGAATGAAAAGGCAC
antisense *oma-1* pUG RNA T7 template reverse	GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG TGTGT ATGAACGTTAACGGTGAAAAC
oma-1(1-50) pUG RNA T7 template reverse	CACACACACACACACACACACACACACACACAC ACACA TCGAGATGATGCTCATCGAT
oma-1(1-100) pUG RNA T7 template reverse	CACACACACACACACACACACACACACACACAC ACACA CGTGATCGAGTGGTGAAACGG
oma-1(271-541) pUG RNA T7 template forward	TAATACGACTCACTATAGGGAGACCAC GCACCTCCATTTCAACAGTTTG

D. MiSeq

oma-1 pUG PCRs were sequenced on an Illumina MiSeq from animals fed HT115 bacteria expressing L4440 empty vector control plasmid (2 biological replicates), *oma-1* RNAi clone from the Ahringer RNAi library or our custom *oma-1* RNAi clone (pAS74). F_1 to F_4 descendants from pAS74-fed animals were obtained as described above and also sequenced (1 replicate each generation). A first round of PCR was performed with the same primers as described above (Table 5). Primers were modified for the second PCR to contain Illumina p5 and p7 sequences, read 1 and 2 sequencing primers, a unique index (reverse primer only) for multiplexing and unique molecular identifiers (NNN) (Table 5). PCR reactions were then pooled, run on an agarose gel and gel purified as described above. Gel purified DNA was sequenced on an Illumina MiSeq (Biopolymers Facility, Harvard Medical School) to obtain paired-end reads (67bp for Read 1, 248bp for Read 2).

E. MiSeq sequencing analysis.

First, unique molecular identifiers (UMIs) were removed from each read pair and appended to the end of the read name using UMI-tools v.1.0.0 (Smith *et al.* 2017). Then, Cutadapt v2.5 was used for the following: 1) low-quality bases (quality score < 20) were trimmed from the 3' ends of reads; 2) read pairs containing the inline portion of the 5' adapter (AACAACGAGAAGATCGATGA) in Read 1 were selected for and then trimmed; 3) Read pairs containing the inline portion of the 3' adapter (GGCGTCGCCATATTCTACTTACACACACACACACACAC) in Read 2 were selected for and trimmed; and 4) If the 5' adapter was present in any Read 2 sequences, the adapter was trimmed from those sequences (Martin 2011). After adapter trimming, Read 2 sequences were screened for additional pACs at the 5' end: reads that did not contain additional pACs (and therefore did not have a pUG tail longer than the adapter) were discarded; reads that did contain additional

pACs were retained, and the pACs were trimmed using Cutadapt v2.5 (pAC and pCA sequences were provided as non-internal 5' adapters) (Martin 2011). After pAC trimming, Read 2 sequences shorter than 5 nucleotides were discarded. The remaining Read 2 sequences were aligned to the *C. elegans* genome (WormBase release WS260) using STAR v2.7.0f (Dobin *et al.* 2013b). SAM and BED files of unique alignments were generated using SAMtools v1.9 and BEDtools v2.27.1 and then imported into R for subsequent analyses(Li *et al.* 2009; Quinlan and Hall 2010; RStudio Team 2016). Alignments were deduplicated based on the combination of the UMI and end coordinate. Alignments that mapped to the "+" strand and/or to coordinates outside of the *oma-1* gene were discarded.

To systematically define the "*oma-1*" and "pUG" portions of each read, the pre-pAC-trimmed version of the read was reverse-complemented and then split as follows. By default, the aligned portion of the read was designated as "*oma-1*", and any sequence downstream of the aligned portion was designated as the "pUG." Then, the "*oma-1*" portion was matched to an *oma-1* reference sequence (spliced + UTRs) using Biostrings v2.50.2 (Pagès *et al.* 2017). If the first 1-6 nucleotides that occurred 3' of the match were the same in the *oma-1* reference as they were in the read prior to pAC trimming (and therefore had the potential to be templated), then those nucleotides were reassigned to the "*oma-1*" portion of the read. End coordinates of the alignments were adjusted accordingly. A small portion of reads (<15%) were misannotated with the above approach, largely due to soft-clipping at the 3' end during alignment. To systematically filter out such reads, reads for which the annotated "pUG" started with a base other than "U" or "G" and/or contained 2 or more bases other than "U" or "G" within the "pUG" sequence were discarded. The abundance of each pUGylation site (Figure 2.2a) was plotted in R using Sushi v1.20.0 (Phanstiel *et al.* 2014) for pUG RNAs sequenced from wild-type animals fed the *oma-1* RNAi clone from the RNAi collection. To generate the pUG site logos shown in Figure 2.2b, a list of unique pUG sites was combined for pUG RNAs sequenced from wild-type animals fed the *oma-1* RNAi clone from the RNAi collection and pAS74 (our custom *oma-1* RNAi clone). This combined list was sorted by

the last nucleotide of the "*oma-1*" portion and then plotted in R using ggseqlogo v0.1 (Wagih 2017).

F. pUG RNA injections

A list of all pUG RNAs injected in this study is provided in Table 6. For *gfp* and *oma-1* pUG RNA injections, pUG RNAs were synthesized *in vitro* using MEGAscript T7 Transcription Kit (Invitrogen, AM1334). DNA templates for *in vitro* transcription reactions were gel purified PCR products (150ng per *in vitro* transcription reaction) amplified using primers listed in Table 5. Reactions were incubated overnight at 37°C. *in vitro* transcribed RNA was purified using TRIzol Reagent (Life Technologies, 15596018) and stored at -80°C. Injection mix consisted of 0.5pmol/ul *in vitro* transcribed RNA and 2.5ng/ul co-injection marker (*pmyo-2::mCherry::unc-54 3'UTR*) plasmid pCFJ90 (Addgene, plasmid #19327), dissolved in water. Animals expressing the co-injection marker show *mCherry* expression in the pharynx. For *gfp* pUG RNA injections, *mjIs31 (gfp::h2b); rde-1(ne219)* animals were injected in the germline and allowed to lay a brood at 20°C. Progeny of injected animals (F_1 generation) were washed off plates using M9 + Triton X-100 buffer, mounted onto slides and 12-20 *mCherry*-expressing progeny per injected animal were scored for *gfp* expression using the Plan-Apochromat 20 × /0.8 M27 objective on an Axio Observer.Z1 fluorescent microscope (Zeiss). Images were taken with the Plan-Apochromat 63 × /1.4 Oil DIC M27 objective. For transgenerational inheritance experiments, 5 F_1 progeny expressing *mCherry* per injected animal were picked under an Axis Zoom.V16 fluorescent dissecting microscope using a PlanNeoFluar Z 1x/0.25 FWD 56mm objective to a new plate to lay an F_2 brood. Five random F_2 animals were picked to a new plate to lay F_3 progeny, while the remaining F_2 adults (~40/injected animal) were scored for *gfp* expression as described above, but without regard for *mCherry* expression. This process was continued for several generations until 100% of animals expressed *gfp*. For all experiments, no injection control animals or animals injected with other RNA species were scored as described for pUG RNA injected animals. For

oma-1 pUG RNA injections, *oma-1(zu405ts); rde-1(ne219*) were injected in the germline and allowed to recover at 15°C for two days. *oma-1(zu405ts*) is a gain-of-function temperature-sensitive allele of *oma-1*. *oma-1(zu405ts)* animals lay arrested embryos at 20°C, unless *oma-1(zu405ts)* is silenced (Alcazar *et al.* 2008). Two days after injections, injected animals were shifted to 20°C. To measure *oma-1(zu405ts)* silencing, 5 adult *mCherry*-expressing progeny (F_1 generation) per injected animal were picked as described above and transferred to a new plate. Animals were removed after laying 50-100 embryos, and *oma-1(zu405ts)* silencing was measured as percentage of embryos hatched (# hatched embryos/total number of embryos laid). For transgenerational inheritance experiments, F_2 progeny that hatched were picked to new plates and allowed to grow to adults at 20°C. Five adult F_2 animals per injected animal were picked to a new plate and % embryos hatched was scored as described above. This process was continued for several generations until 100% of embryos failed to hatch at 20°C. No injection control animals were maintained at 15°C. Every generation, 5 animals were shifted to 20°C before reaching adulthood and % embryos hatched was scored once they were adults. For Tc1 pUG RNA injections, T7 *in vitro* transcription was performed as described above to synthesize a Tc1 pUG RNA consisting of a 36nt pUG tail appended to a 338nt long fragment of Tc1 RNA (see Table 5 for primers used). 18 and 22 *rde-3(-); unc-22::tc1* animals were injected with a Tc1 pUG RNA + co-injection marker or co-injection marker alone, respectively. *unc-22(st136)* animals have a Tc1 DNA transposon insertion in the *unc-22* gene, resulting in paralysis. *mCherry*-expressing progeny of Tc1 pUG RNA + co-injection marker or co-injection marker only injected animals were picked at the L4 stage and randomly pooled (25 animals per pool) onto 10cM NGM plates and allowed to lay a brood. The number of mobile adult progeny in each pool was counted 6-7 days later.

Table 2.6. pUG RNAs injected in this study.

pUG RNA Name	Sequence (5' to 3')
gfp $p(UG)_{18}$	AUGAGUAAAGGAGAAGAACUUUUCACUGGAGUUGUCCCAAUUCUUGUUGAAUUAGAUGGU GAUGUUAAUGGGCACAAAUUUUCUGUCAGUGGAGAGGGUGAAGGUGAUGCAACAUACGGA AAACUUACCCUUAAAUUUAUUUGCACUACUGGAAAACUACCUGUUCCAUGGCCAACACUUG UCACUACUUUCUGUUAUGGUGUUCAAUGCUUCUCGAGAUACCCAGAUCAUAUGAAACGGC AUGACUUUUUCAAGAGUGCCAUGCCCGAAGGUUAUGUACAGGAAAGAACUAUAUUUUUCA AAGAUGACGGGAACUACAAGACACGUGCUGAAGUCAAGUUUGAAGGUGAUACCCUUGUUA AUAGAAUC UGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUG
gfp RNA fragment with no tail (also the sense strand of the *gfp* dsRNA)	AUGAGUAAAGGAGAAGAACUUUUCACUGGAGUUGUCCCAAUUCUUGUUGAAUUAGAUGGU GAUGUUAAUGGGCACAAAUUUUCUGUCAGUGGAGAGGGUGAAGGUGAUGCAACAUACGGA AAACUUACCCUUAAAUUUAUUUGCACUACUGGAAAACUACCUGUUCCAUGGCCAACACUUG UCACUACUUUCUGUUAUGGUGUUCAAUGCUUCUCGAGAUACCCAGAUCAUAUGAAACGGC AUGACUUUUUCAAGAGUGCCAUGCCCGAAGGUUAUGUACAGGAAAGAACUAUAUUUUUCA AAGAUGACGGGAACUACAAGACACGUGCUGAAGUCAAGUUUGAAGGUGAUACCCUUGUUA AUAGAAUC
antisense strand of *gfp* dsRNA	GAUUCUAUUAACAAGGGUAUCACCUUCAAACUUGACUUCAGCACGUGUCUUGUAGUUCCC GUCAUCUUUGAAAAAUAUAGUUCUUUCCUGUACAUAACCUUCGGGCAUGGCACUCUUGAA AAAGUCAUGCCGUUUCAUAUGAUCUGGGUAUCUCGAGAAGCAUUGAACACCAUAACAGAAA GUAGUGACAAGUGUUGGCCAUGGAACAGGUAGUUUUCCAGUAGUGCAAAUAAAUUUAAGG GUAAGUUUUCCGUAUGUUGCAUCACCUUCACCCUCUCCACUGACAGAAAAUUUGUGCCCA UUAACAUCACCAUCUAAUUCAACAAGAAUUGGGACAACUCCAGUGAAAAGUUCUUCUCCUU UACUCAU
gfp $p(AU)_{18}$	AUGAGUAAAGGAGAAGAACUUUUCACUGGAGUUGUCCCAAUUCUUGUUGAAUUAGAUGGU GAUGUUAAUGGGCACAAAUUUUCUGUCAGUGGAGAGGGUGAAGGUGAUGCAACAUACGGA AAACUUACCCUUAAAUUUAUUUGCACUACUGGAAAACUACCUGUUCCAUGGCCAACACUUG UCACUACUUUCUGUUAUGGUGUUCAAUGCUUCUCGAGAUACCCAGAUCAUAUGAAACGGC AUGACUUUUUCAAGAGUGCCAUGCCCGAAGGUUAUGUACAGGAAAGAACUAUAUUUUUCA AAGAUGACGGGAACUACAAGACACGUGCUGAAGUCAAGUUUGAAGGUGAUACCCUUGUUA AUAGAAUC AUAUAUAUAUAUAUAUAUAUAUAUAUAUAUAU
gfp $p(GC)_{18}$	AUGAGUAAAGGAGAAGAACUUUUCACUGGAGUUGUCCCAAUUCUUGUUGAAUUAGAUGGU GAUGUUAAUGGGCACAAAUUUUCUGUCAGUGGAGAGGGUGAAGGUGAUGCAACAUACGGA AAACUUACCCUUAAAUUUAUUUGCACUACUGGAAAACUACCUGUUCCAUGGCCAACACUUG UCACUACUUUCUGUUAUGGUGUUCAAUGCUUCUCGAGAUACCCAGAUCAUAUGAAACGGC AUGACUUUUUCAAGAGUGCCAUGCCCGAAGGUUAUGUACAGGAAAGAACUAUAUUUUUCA AAGAUGACGGGAACUACAAGACACGUGCUGAAGUCAAGUUUGAAGGUGAUACCCUUGUUA AUAGAAUC GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
gfp $p(CA)_{18}$	AUGAGUAAAGGAGAAGAACUUUUCACUGGAGUUGUCCCAAUUCUUGUUGAAUUAGAUGGU GAUGUUAAUGGGCACAAAUUUUCUGUCAGUGGAGAGGGUGAAGGUGAUGCAACAUACGGA AAACUUACCCUUAAAUUUAUUUGCACUACUGGAAAACUACCUGUUCCAUGGCCAACACUUG UCACUACUUUCUGUUAUGGUGUUCAAUGCUUCUCGAGAUACCCAGAUCAUAUGAAACGGC AUGACUUUUUCAAGAGUGCCAUGCCCGAAGGUUAUGUACAGGAAAGAACUAUAUUUUUCA AAGAUGACGGGAACUACAAGACACGUGCUGAAGUCAAGUUUGAAGGUGAUACCCUUGUUA AUAGAAUC CACACACACACACACACACACACACACACACACACA

oma-1 p(UG)$_{18}$	AUGAACGUUAACGGUGAAAACAACGAGAAGAUCGAUGAGCAUCAUCUCGAGAGCUCCCUC GCUGGCGUUCCAACUUUGCCCGUUUCACCACUCGAUCACGCUAAGGACUUGUCUCAAACU AAUCCGAAUGCCCAGAUUGGAGAUUUGGUUACUCAAACUGCUAACUUGAUUGCUAUCAAAA AGCAGUUGCUUGAAGAUAUUGCAUUCAACCAACAUAUUCAAUCGAUGCAAGUUCGCGCGA UUCAGAGUUUUCCCCAGAACAACCAGGUUGCACCUCCAUUUCAACAGUUUGAUCCUAGAC GCCGCGGAUUAGCCCGAAUGCAGAAACCAGAAUCAUACAAGACUGUGAUCUGCCAAGCUU GGCUCGAAUCUAAGACUUGUUCUUUCGCAGACAACUGCCGGUUUGCACAUGGAGAGGAGG AACUUCGUCCAACAUUCGUCGAGCCGCUGCAGAACAAUAAGUAUAAAACGAAACUCUGUGA CAAAUAUACGACUACAGGACUCUGCCCAUACGGGAAACGGUGCCUUUUCAUUCAUCCCG UGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUG
oma-1 RNA fragment with no tail (also the sense strand of the *oma-1* dsRNA)	AUGAACGUUAACGGUGAAAACAACGAGAAGAUCGAUGAGCAUCAUCUCGAGAGCUCCCUC GCUGGCGUUCCAACUUUGCCCGUUUCACCACUCGAUCACGCUAAGGACUUGUCUCAAACU AAUCCGAAUGCCCAGAUUGGAGAUUUGGUUACUCAAACUGCUAACUUGAUUGCUAUCAAAA AGCAGUUGCUUGAAGAUAUUGCAUUCAACCAACAUAUUCAAUCGAUGCAAGUUCGCGCGA UUCAGAGUUUUCCCCAGAACAACCAGGUUGCACCUCCAUUUCAACAGUUUGAUCCUAGAC GCCGCGGAUUAGCCCGAAUGCAGAAACCAGAAUCAUACAAGACUGUGAUCUGCCAAGCUU GGCUCGAAUCUAAGACUUGUUCUUUCGCAGACAACUGCCGGUUUGCACAUGGAGAGGAGG AACUUCGUCCAACAUUCGUCGAGCCGCUGCAGAACAAUAAGUAUAAAACGAAACUCUGUGA CAAAUAUACGACUACAGGACUCUGCCCAUACGGGAAACGGUGCCUUUUCAUUCAUCCCG
antisense strand of *oma-1* dsRNA	CGGGAUGAAUGAAAAGGCACCGUUUCCCGUAUGGGCAGAGUCCUGUAGUCGUAUAUUUGU CACAGAGUUUCGUUUUAUACUUAUUGUUCUGCAGCGGCUCGACGAAUGUUGGACGAAGUU CCUCCUCUCCAUGUGCAAACCGGCAGUUGUCUGCGAAAGAACAAGUCUUAGAUUCGAGCC AAGCUUGGCAGAUCACAGUCUUGUAUGAUUCUGGUUUCUGCAUUCGGGCUAAUCCGCGG CGUCUAGGAUCAAACUGUUGAAAUGGAGGUGCAACCUGGUUGUUCUGGGGAAAACUCUGA AUCGCGCGAACUUGCAUCGAUUGAAUAUGUUGGUUGAAUGCAAUAUCUUCAAGCAACUGC UUUUUGAUAGCAAUCAAGUUAGCAGUUUGAGUAACCAAAUCUCCAAUCUGGGCAUUCGGA UUAGUUUGAGACAAGUCCUUAGCGUGAUCGAGUGGUGAAACGGGCAAAGUUGGAACGCCA GCGAGGGAGCUCUCGAGAUGAUGCUCAUCGAUCUUCUCGUUGUUUUCACCGUUAACGUU CAU
oma-1 p(AU)$_{18}$	AUGAACGUUAACGGUGAAAACAACGAGAAGAUCGAUGAGCAUCAUCUCGAGAGCUCCCUC GCUGGCGUUCCAACUUUGCCCGUUUCACCACUCGAUCACGCUAAGGACUUGUCUCAAACU AAUCCGAAUGCCCAGAUUGGAGAUUUGGUUACUCAAACUGCUAACUUGAUUGCUAUCAAAA AGCAGUUGCUUGAAGAUAUUGCAUUCAACCAACAUAUUCAAUCGAUGCAAGUUCGCGCGA UUCAGAGUUUUCCCCAGAACAACCAGGUUGCACCUCCAUUUCAACAGUUUGAUCCUAGAC GCCGCGGAUUAGCCCGAAUGCAGAAACCAGAAUCAUACAAGACUGUGAUCUGCCAAGCUU GGCUCGAAUCUAAGACUUGUUCUUUCGCAGACAACUGCCGGUUUGCACAUGGAGAGGAGG AACUUCGUCCAACAUUCGUCGAGCCGCUGCAGAACAAUAAGUAUAAAACGAAACUCUGUGA CAAAUAUACGACUACAGGACUCUGCCCAUACGGGAAACGGUGCCUUUUCAUUCAUCCCG AUAUAUAUAUAUAUAUAUAUAUAUAUAUAUAU
oma-1 p(GC)$_{18}$	AUGAACGUUAACGGUGAAAACAACGAGAAGAUCGAUGAGCAUCAUCUCGAGAGCUCCCUC GCUGGCGUUCCAACUUUGCCCGUUUCACCACUCGAUCACGCUAAGGACUUGUCUCAAACU AAUCCGAAUGCCCAGAUUGGAGAUUUGGUUACUCAAACUGCUAACUUGAUUGCUAUCAAAA AGCAGUUGCUUGAAGAUAUUGCAUUCAACCAACAUAUUCAAUCGAUGCAAGUUCGCGCGA UUCAGAGUUUUCCCCAGAACAACCAGGUUGCACCUCCAUUUCAACAGUUUGAUCCUAGAC GCCGCGGAUUAGCCCGAAUGCAGAAACCAGAAUCAUACAAGACUGUGAUCUGCCAAGCUU GGCUCGAAUCUAAGACUUGUUCUUUCGCAGACAACUGCCGGUUUGCACAUGGAGAGGAGG AACUUCGUCCAACAUUCGUCGAGCCGCUGCAGAACAAUAAGUAUAAAACGAAACUCUGUGA CAAAUAUACGACUACAGGACUCUGCCCAUACGGGAAACGGUGCCUUUUCAUUCAUCCCG GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC

oma-1 p(CA)$_{18}$	AUGAACGUUAACGGUGAAAACAACGAGAAGAUCGAUGAGCAUCAUCUCGAGAGCUCCCUC GCUGGCGUUCCAACUUUGCCCGUUUCACCACUCGAUCACGCUAAGGACUUGUCUCAAACU AAUCCGAAUGCCCAGAUUGGAGAUUUGGUUACUCAAACUGCUAACUUGAUUGCUAUCAAAA AGCAGUUGCUUGAAGAUAUUGCAUUCAACCAACAUAUUCAAUCGAUGCAAGUUCGCGCGA UUCAGAGUUUUCCCCAGAACAACCAGGUUGCACCUCCAUUUCAACAGUUUGAUCCUAGAC GCCGCGGAUUAGCCCGAAUGCAGAAACCAGAAUCAUACAAGACUGUGAUCUGCCAAGCUU GGCUCGAAUCUAAGACUUGUUCUUUCGCAGACAACUGCCGGUUUGCACAUGGAGAGGAGG AACUUCGUCCAACAUUCGUCGAGCCGCUGCAGAACAAUAAGUAUAAAACGAAACUCUGUGA CAAAUAUACGACUACAGGACUCUGCCCAUACGGGAAACGGUGCCUUUUCAUUCAUCCCG CACACACACACACACACACACACACACACACACACA
oma-1 p(UG)$_{1}$	AUGAACGUUAACGGUGAAAACAACGAGAAGAUCGAUGAGCAUCAUCUCGAGAGCUCCCUC GCUGGCGUUCCAACUUUGCCCGUUUCACCACUCGAUCACGCUAAGGACUUGUCUCAAACU AAUCCGAAUGCCCAGAUUGGAGAUUUGGUUACUCAAACUGCUAACUUGAUUGCUAUCAAAA AGCAGUUGCUUGAAGAUAUUGCAUUCAACCAACAUAUUCAAUCGAUGCAAGUUCGCGCGA UUCAGAGUUUUCCCCAGAACAACCAGGUUGCACCUCCAUUUCAACAGUUUGAUCCUAGAC GCCGCGGAUUAGCCCGAAUGCAGAAACCAGAAUCAUACAAGACUGUGAUCUGCCAAGCUU GGCUCGAAUCUAAGACUUGUUCUUUCGCAGACAACUGCCGGUUUGCACAUGGAGAGGAGG AACUUCGUCCAACAUUCGUCGAGCCGCUGCAGAACAAUAAGUAUAAAACGAAACUCUGUGA CAAAUAUACGACUACAGGACUCUGCCCAUACGGGAAACGGUGCCUUUUCAUUCAUCCCG UG
oma-1 p(UG)$_{5}$	AUGAACGUUAACGGUGAAAACAACGAGAAGAUCGAUGAGCAUCAUCUCGAGAGCUCCCUC GCUGGCGUUCCAACUUUGCCCGUUUCACCACUCGAUCACGCUAAGGACUUGUCUCAAACU AAUCCGAAUGCCCAGAUUGGAGAUUUGGUUACUCAAACUGCUAACUUGAUUGCUAUCAAAA AGCAGUUGCUUGAAGAUAUUGCAUUCAACCAACAUAUUCAAUCGAUGCAAGUUCGCGCGA UUCAGAGUUUUCCCCAGAACAACCAGGUUGCACCUCCAUUUCAACAGUUUGAUCCUAGAC GCCGCGGAUUAGCCCGAAUGCAGAAACCAGAAUCAUACAAGACUGUGAUCUGCCAAGCUU GGCUCGAAUCUAAGACUUGUUCUUUCGCAGACAACUGCCGGUUUGCACAUGGAGAGGAGG AACUUCGUCCAACAUUCGUCGAGCCGCUGCAGAACAAUAAGUAUAAAACGAAACUCUGUGA CAAAUAUACGACUACAGGACUCUGCCCAUACGGGAAACGGUGCCUUUUCAUUCAUCCCG UGUGUGUGUG
oma-1 p(UG)$_{8}$	AUGAACGUUAACGGUGAAAACAACGAGAAGAUCGAUGAGCAUCAUCUCGAGAGCUCCCUC GCUGGCGUUCCAACUUUGCCCGUUUCACCACUCGAUCACGCUAAGGACUUGUCUCAAACU AAUCCGAAUGCCCAGAUUGGAGAUUUGGUUACUCAAACUGCUAACUUGAUUGCUAUCAAAA AGCAGUUGCUUGAAGAUAUUGCAUUCAACCAACAUAUUCAAUCGAUGCAAGUUCGCGCGA UUCAGAGUUUUCCCCAGAACAACCAGGUUGCACCUCCAUUUCAACAGUUUGAUCCUAGAC GCCGCGGAUUAGCCCGAAUGCAGAAACCAGAAUCAUACAAGACUGUGAUCUGCCAAGCUU GGCUCGAAUCUAAGACUUGUUCUUUCGCAGACAACUGCCGGUUUGCACAUGGAGAGGAGG AACUUCGUCCAACAUUCGUCGAGCCGCUGCAGAACAAUAAGUAUAAAACGAAACUCUGUGA CAAAUAUACGACUACAGGACUCUGCCCAUACGGGAAACGGUGCCUUUUCAUUCAUCCCG UGUGUGUGUGUGUG
oma-1 p(UG)$_{14}$	AUGAACGUUAACGGUGAAAACAACGAGAAGAUCGAUGAGCAUCAUCUCGAGAGCUCCCUC GCUGGCGUUCCAACUUUGCCCGUUUCACCACUCGAUCACGCUAAGGACUUGUCUCAAACU AAUCCGAAUGCCCAGAUUGGAGAUUUGGUUACUCAAACUGCUAACUUGAUUGCUAUCAAAA AGCAGUUGCUUGAAGAUAUUGCAUUCAACCAACAUAUUCAAUCGAUGCAAGUUCGCGCGA UUCAGAGUUUUCCCCAGAACAACCAGGUUGCACCUCCAUUUCAACAGUUUGAUCCUAGAC GCCGCGGAUUAGCCCGAAUGCAGAAACCAGAAUCAUACAAGACUGUGAUCUGCCAAGCUU GGCUCGAAUCUAAGACUUGUUCUUUCGCAGACAACUGCCGGUUUGCACAUGGAGAGGAGG AACUUCGUCCAACAUUCGUCGAGCCGCUGCAGAACAAUAAGUAUAAAACGAAACUCUGUGA CAAAUAUACGACUACAGGACUCUGCCCAUACGGGAAACGGUGCCUUUUCAUUCAUCCCG UGUGUGUGUGUGUGUGUGUGUGUGUG

oma-1 p(UG)$_{40}$	AUGAACGUUAACGGUGAAAACAACGAGAAGAUCGAUGAGCAUCAUCUCGAGAGCUCCCUC GCUGGCGUUCCAACUUUGCCCGUUUCACCACUCGAUCACGCUAAGGACUUGUCUCAAACU AAUCCGAAUGCCCAGAUUGGAGAUUUGGUUACUCAAACUGCUAACUUGAUUGCUAUCAAAA AGCAGUUGCUUGAAGAUAUUGCAUUCAACCAACAUAUUCAAUCGAUGCAAGUUCGCGCGA UUCAGAGUUUUCCCCAGAACAACCAGGUUGCACCUCCAUUUCAACAGUUUGAUCCUAGAC GCCGCGGAUUAGCCCGAAUGCAGAAACCAGAAUCAUACAAGACUGUGAUCUGCCAAGCUU GGCUCGAAUCUAAGACUUGUUCUUUCGCAGACAACUGCCGGUUUGCACAUGGAGAGGAGG AACUUCGUCCAACAUUCGUCGAGCCGCUGCAGAACAAUAAGUAUAAAACGAAACUCUGUGA CAAAUAUACGACUACAGGACUCUGCCCAUACGGGAAACGGUGCCUUUUCAUUCAUCCCG UGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUG UGUGUGUGUGUGUGUGUGUG
oma-1 (U)$_{18}$(G)$_{18}$	AUGAACGUUAACGGUGAAAACAACGAGAAGAUCGAUGAGCAUCAUCUCGAGAGCUCCCUC GCUGGCGUUCCAACUUUGCCCGUUUCACCACUCGAUCACGCUAAGGACUUGUCUCAAACU AAUCCGAAUGCCCAGAUUGGAGAUUUGGUUACUCAAACUGCUAACUUGAUUGCUAUCAAAA AGCAGUUGCUUGAAGAUAUUGCAUUCAACCAACAUAUUCAAUCGAUGCAAGUUCGCGCGA UUCAGAGUUUUCCCCAGAACAACCAGGUUGCACCUCCAUUUCAACAGUUUGAUCCUAGAC GCCGCGGAUUAGCCCGAAUGCAGAAACCAGAAUCAUACAAGACUGUGAUCUGCCAAGCUU GGCUCGAAUCUAAGACUUGUUCUUUCGCAGACAACUGCCGGUUUGCACAUGGAGAGGAGG AACUUCGUCCAACAUUCGUCGAGCCGCUGCAGAACAAUAAGUAUAAAACGAAACUCUGUGA CAAAUAUACGACUACAGGACUCUGCCCAUACGGGAAACGGUGCCUUUUCAUUCAUCCCG UUUUUUUUUUUUUUUUUUGGGGGGGGGGGGGGGGGG
oma-1 UG scrambled	AUGAACGUUAACGGUGAAAACAACGAGAAGAUCGAUGAGCAUCAUCUCGAGAGCUCCCUC GCUGGCGUUCCAACUUUGCCCGUUUCACCACUCGAUCACGCUAAGGACUUGUCUCAAACU AAUCCGAAUGCCCAGAUUGGAGAUUUGGUUACUCAAACUGCUAACUUGAUUGCUAUCAAAA AGCAGUUGCUUGAAGAUAUUGCAUUCAACCAACAUAUUCAAUCGAUGCAAGUUCGCGCGA UUCAGAGUUUUCCCCAGAACAACCAGGUUGCACCUCCAUUUCAACAGUUUGAUCCUAGAC GCCGCGGAUUAGCCCGAAUGCAGAAACCAGAAUCAUACAAGACUGUGAUCUGCCAAGCUU GGCUCGAAUCUAAGACUUGUUCUUUCGCAGACAACUGCCGGUUUGCACAUGGAGAGGAGG AACUUCGUCCAACAUUCGUCGAGCCGCUGCAGAACAAUAAGUAUAAAACGAAACUCUGUGA CAAAUAUACGACUACAGGACUCUGCCCAUACGGGAAACGGUGCCUUUUCAUUCAUCCCG GUUUGGGGUUGGGUUUGGGUGGGUUUGUUUUUUGGG
oma-1 (UUGG)$_{9}$	AUGAACGUUAACGGUGAAAACAACGAGAAGAUCGAUGAGCAUCAUCUCGAGAGCUCCCUC GCUGGCGUUCCAACUUUGCCCGUUUCACCACUCGAUCACGCUAAGGACUUGUCUCAAACU AAUCCGAAUGCCCAGAUUGGAGAUUUGGUUACUCAAACUGCUAACUUGAUUGCUAUCAAAA AGCAGUUGCUUGAAGAUAUUGCAUUCAACCAACAUAUUCAAUCGAUGCAAGUUCGCGCGA UUCAGAGUUUUCCCCAGAACAACCAGGUUGCACCUCCAUUUCAACAGUUUGAUCCUAGAC GCCGCGGAUUAGCCCGAAUGCAGAAACCAGAAUCAUACAAGACUGUGAUCUGCCAAGCUU GGCUCGAAUCUAAGACUUGUUCUUUCGCAGACAACUGCCGGUUUGCACAUGGAGAGGAGG AACUUCGUCCAACAUUCGUCGAGCCGCUGCAGAACAAUAAGUAUAAAACGAAACUCUGUGA CAAAUAUACGACUACAGGACUCUGCCCAUACGGGAAACGGUGCCUUUUCAUUCAUCCCG UUGGUUGGUUGGUUGGUUGGUUGGUUGGUUGG
5' p(UG)$_{18}$ + sense *oma-1*	GUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGU AUGAACGUUAACGGUGAAAACAACGAGAAGAUCGAUGAGCAUCAUCUCGAGAGCUCCCUC GCUGGCGUUCCAACUUUGCCCGUUUCACCACUCGAUCACGCUAAGGACUUGUCUCAAACU AAUCCGAAUGCCCAGAUUGGAGAUUUGGUUACUCAAACUGCUAACUUGAUUGCUAUCAAAA AGCAGUUGCUUGAAGAUAUUGCAUUCAACCAACAUAUUCAAUCGAUGCAAGUUCGCGCGA UUCAGAGUUUUCCCCAGAACAACCAGGUUGCACCUCCAUUUCAACAGUUUGAUCCUAGAC GCCGCGGAUUAGCCCGAAUGCAGAAACCAGAAUCAUACAAGACUGUGAUCUGCCAAGCUU GGCUCGAAUCUAAGACUUGUUCUUUCGCAGACAACUGCCGGUUUGCACAUGGAGAGGAGG AACUUCGUCCAACAUUCGUCGAGCCGCUGCAGAACAAUAAGUAUAAAACGAAACUCUGUGA CAAAUAUACGACUACAGGACUCUGCCCAUACGGGAAACGGUGCCUUUUCAUUCAUCCCG

internal p(UG)$_{18}$ in *oma-1* sense RNA fragment	AUGAACGUUAACGGUGAAAACAACGAGAAGAUCGAUGAGCAUCAUCUCGAGAGCUCCCUC GCUGGCGUUCCAACUUUGCCCGUUUCACCACUCGAUCACGCUAAGGACUUGUCUCAAACU AAUCCGAAUGCCCAGAUUGGAGAUUUGGUUACUCAAACUGCUAACUUGAUUGCUAUCAAAA AGCAGUUGCUUGAAGAUAUUGCAUUCAACCAACAUAUUCAAUCGAUGCAAGUUCGCGCGA UUCAGAGUUUUCCCCAGAACAACCAGGUU GUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGU GCACCUCCAUUUCAACAGUUUGAUCCUAGACGCCGCGGAUUAGCCCGAAUGCAGAAACCA GAAUCAUACAAGACUGUGAUCUGCCAAGCUUGGCUCGAAUCUAAGACUUGUUCUUUCGCA GACAACUGCCGGUUUGCACAUGGAGAGGAGGAACUUCGUCCAACAUUCGUCGAGCCGCUG CAGAACAAUAAGUAUAAAACGAAACUCUGUGACAAAUAUACGACUACAGGACUCUGCCCAU ACGGGAAACGGUGCCUUUUCAUUCAUCCCG
oma-1 p(UG)18 with SNP	AUGAACGUUAACGGUGAAAACAACGAGAAGAUCGAUGAGCAUCAUCUCGAGAGCUCCCUC GCUGGCGUUCCAACUUUGCCCGUUUCACCACUCGAUCACGCUAAGGACUUGUCUCAAACU AAUCCGAAUGCCCAGAUUGGAGAUUUGGUUACUCAAACUGCUAACUUGAUUGCUAUCAAAA AGCAGUUGCUUGAAGAUAUUGCAUUCAACCAACAUAUUCAAUCGAUGCAAGUUCGCGCGA UUCAGAGUUUUCCCCAGAACAACCAGGUUGCACCUCCAUUUCAACAGUUUGAUCCUAGAC GCCGCGGAUUAGCCCGAAUGCAGAAACCAGAAUCAUACAAGACUGUGAUCUGCCAAGCUU GGCUCGAAUCUAAGACUUGUUCUUUCGCAGACAACUGCCGGUUUGCACAUGGAGAGGAGG AACUUCGUCCAACAUUCGUCGAGCCGCUGCAGAACAAUAAGUAUAAAACGAAACUCUGUGA CAAAUAUACGACUACAGGACUCUGCCCAUACGGGAAACGGUGCCUUUUCAUUCAUCCCGU UCAUGGACCAAACGCAUACAUUCG GUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUG
oma-1 (pGC)$_{18}$ with SNP	AUGAACGUUAACGGUGAAAACAACGAGAAGAUCGAUGAGCAUCAUCUCGAGAGCUCCCUC GCUGGCGUUCCAACUUUGCCCGUUUCACCACUCGAUCACGCUAAGGACUUGUCUCAAACU AAUCCGAAUGCCCAGAUUGGAGAUUUGGUUACUCAAACUGCUAACUUGAUUGCUAUCAAAA AGCAGUUGCUUGAAGAUAUUGCAUUCAACCAACAUAUUCAAUCGAUGCAAGUUCGCGCGA UUCAGAGUUUUCCCCAGAACAACCAGGUUGCACCUCCAUUUCAACAGUUUGAUCCUAGAC GCCGCGGAUUAGCCCGAAUGCAGAAACCAGAAUCAUACAAGACUGUGAUCUGCCAAGCUU GGCUCGAAUCUAAGACUUGUUCUUUCGCAGACAACUGCCGGUUUGCACAUGGAGAGGAGG AACUUCGUCCAACAUUCGUCGAGCCGCUGCAGAACAAUAAGUAUAAAACGAAACUCUGUGA CAAAUAUACGACUACAGGACUCUGCCCAUACGGGAAACGGUGCCUUUUCAUUCAUCCCGU UCAUGGACCAAACGCAUACAUUCG GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC
antisense *oma-1* p(UG)$_{18}$	CGGGAUGAAUGAAAAGGCACCGUUUCCCGUAUGGGCAGAGUCCUGUAGUCGUAUAUUUGU CACAGAGUUUCGUUUUAUACUUAUUGUUCUGCAGCGGCUCGACGAAUGUUGGACGAAGUU CCUCCUCUCCAUGUGCAAACCGGCAGUUGUCUGCGAAAGAACAAGUCUUAGAUUCGAGCC AAGCUUGGCAGAUCACAGUCUUGUAUGAUUCUGGUUUCUGCAUUCGGGCUAAUCCGCGG CGUCUAGGAUCAAACUGUUGAAAUGGAGGUGCAACCUGGUUGUUCUGGGGAAAACUCUGA AUCGCGCGAACUUGCAUCGAUUGAAUAUGUUGGUUGAAUGCAAUAUCUUCAAGCAACUGC UUUUUGAUAGCAAUCAAGUUAGCAGUUUGAGUAACCAAAUCUCCAAUCUGGGCAUUCGGA UUAGUUUGAGACAAGUCCUUAGCGUGAUCGAGUGGUGAAACGGGCAAAGUUGGAACGCCA GCGAGGGAGCUCUCGAGAUGAUGCUCAUCGAUCUUCUCGUUGUUUUCACCGUUAACGUU CAU UGUGUGUGUGUGUGUGUGUGUGUGUGUGUG
oma-1(1-50) p(UG)$_{18}$	AUGAACGUUAACGGUGAAAACAACGAGAAGAUCGAUGAGCAUCAUCUCGA UGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUG
oma-1(1-100) p(UG)$_{18}$	AUGAACGUUAACGGUGAAAACAACGAGAAGAUCGAUGAGCAUCAUCUCGAGAGCUCCCUC GCUGGCGUUCCAACUUUGCCCGUUUCACCACUCGAUCACG UGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUG
oma-1(1-270) p(UG)$_{18}$	AUGAACGUUAACGGUGAAAACAACGAGAAGAUCGAUGAGCAUCAUCUCGAGAGCUCCCUC GCUGGCGUUCCAACUUUGCCCGUUUCACCACUCGAUCACGCUAAGGACUUGUCUCAAACU AAUCCGAAUGCCCAGAUUGGAGAUUUGGUUACUCAAACUGCUAACUUGAUUGCUAUCAAAA AGCAGUUGCUUGAAGAUAUUGCAUUCAACCAACAUAUUCAAUCGAUGCAAGUUCGCGCGA UUCAGAGUUUUCCCCAGAACAACCAGGUU UGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUG

oma-1(271-541) p(UG)$_{18}$	GCACCUCCAUUUCAACAGUUUGAUCCUAGACGCCGCGGAUUAGCCCGAAUGCAGAAACCA GAAUCAUACAAGACUGUGAUCUGCCAAGCUUGGCUCGAAUCUAAGACUUGUUCUUUCGCA GACAACUGCCGGUUUGCACAUGGAGAGGAGGAACUUCGUCCAACAUUCGUCGAGCCGCUG CAGAACAAUAAGUAUAAAACGAAACUCUGUGACAAAUAUACGACUACAGGACUCUGCCCAU ACGGGAAACGGUGCCUUUUCAUUCAUCCCG UGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUG
Tc1 p(UG)$_{18}$	AUGGUAAAAUCUGUUGGGUGUAAAAAUCUUUCCUUGGACGUCAAGAAAGCCAUUGUAGCU GGCUUCGAACAAGGAAUACCCACGAAAAUGCUCGCGCUGCAAAUUCAACGUUCUCCGUCG ACUAUUUGGAAAGUAAUCAAGAAGUACCAAACUGAGAAAUCCGUCGCUUUGAGAAUCUCGC CCGGCAGGCCUCGAGUGACAACCCAUAGGAUGGAUCGCAACAUCCUCCGAUCAGCAAGAG AAGAUCCGCAUAGGACCGCCACGGAUAUUCAAAUGAUUAUAAGUUCUCCAAAUGAACCUGU ACCAAGUAAACGAACUGUUCGUCGACGUUUACAGCAUGUGUGUGUGUGUGUGUGUGU GUGUGUGUGUGUG

G. RNA FISH + Immunofluorescence

Approximately 30 animals were dissected in 15 µl of 1X egg buffer (25 mM HEPES (pH 7.3), 118 mM $NaCl_2$, 48 mM KCl, 2 mM $CaCl_2$, 2 mM $MgCl_2$) on SuperFrost Plus Adhesion slides (VWR, 631-0448) to isolate gonads. A coverslip was placed on top of dissected tissue, excess buffer was soaked up using a Kimwipe and slides were placed onto a metal block pre-chilling on dry ice for 10 min. Coverslips were popped off and slides were submerged in methanol at −20°C for 10 min. Slides were then washed twice, 5 min per wash in 1X PBS + 0.1% Tween-20 (PBSTW). Samples were then fixed with 4% paraformaldehyde solution in 1X PBS for 20 min, followed by two 5 min washes in PBSTW. Samples were then incubated at 37°C for 6 hours in a humid chamber with a 1:50 dilution of fluorescent RNA FISH probe in hybridization buffer (10% formamide, 2X SSC, 10% dextran sulfate (w/v)). The RNA FISH probe to detect pUG RNAs (/5Alexa647/CACACACACACACACACACA) was ordered from Integrated DNA Technologies (IDT) and stored at a stock concentration of 100uM at -20°C. The RNA FISH probe to detect *ama-1* mRNA was ordered from Stellaris (SMF-6011-1). After 6 hours, slides were washed twice, 10 min per wash, in FISH Wash Buffer (2X SSC, 10% formamide, 0.1% Tween-20). Samples were then washed for 5 min in 2X SSC. Slides were sealed using 15ul of VECTASHIELD Antifade Mounting Medium (H-1000) with 4′,6-diamidino-2-phenylindole (DAPI). For experiments in which

RNA FISH and immunofluorescence were combined, RNA FISH was first performed as above. After the final 2X SSC wash, slides were washed once with PBST for 5 min, samples were incubated overnight at room temperature in a humid chamber with a 1:1000 dilution of GFP antibody (Abcam, ab290) in PBSTW. Slides were then washed three times, 10 min per wash, in PBSTW and incubated in a 1:100 dilution (in PBSTW) of Goat anti-Rabbit secondary antibody, Alexa Fluor 555 (Invitrogen, A-21429) for 2 hours at room temperature in a humid chamber. Slides were next washed three times, 10 min per wash, in PBSTW and then sealed with 15ul of VECTASHIELD Antifade Mounting Medium (H-1000) with DAPI. All imaging was performed on an Axio Observer.Z1 fluorescent microscope (Zeiss) using the Plan-Apochromat 63 × /1.4 Oil DIC M27 objective. All image processing was done using Fiji (Schindelin *et al.* 2012).

H. RNA-seq

Total RNA was extracted using TRIzol Reagent (Life Technologies, 15596018). RNA quality (RIN) and quantity were assessed on the TapeStation 2200 (Agilent). Two rounds of mRNA purification were performed on 1ug total RNA using the Dynabeads mRNA DIRECT Kit (Invitrogen, 61011). First-strand cDNA was generated using the Superscript III First-Strand Synthesis System (Invitrogen, 18080051), followed by second-strand synthesis using DNA polymerase I (Invitrogen, 11917010). cDNA libraries were prepared using the Nextera XT DNA Library Preparation Kit (Illumina, FC-131-1024). Libraries were sequenced on the Illumina NextSeq500 platform (Biopolymers Facility, Harvard Medical School) and 75bp paired-end reads were obtained.

I. RNA-seq analysis

Reads were trimmed to remove sequencing adapters and low-quality bases using Trim Galore version 0.4.4_dev (https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/).

Trimmed reads were then aligned to the *C. elegans* genome (UCSC ce11/WBcel235) using STAR version 2.7.0a (Dobin *et al.* 2013a). Differential expression analysis of genes and repeat elements was performed using the TEtranscripts package in TEToolkit version 2.0.3 (Jin *et al.* 2015). Gene annotations were obtained from Ensembl (WormBase release WS260) (Cunningham *et al.* 2019). Repeat annotations were obtained from UCSC by downloading the RepeatMasker (rmsk) table in the Table Browser program. The table was reformatted to a GTF file using the Perl script makeTEgtf.pl (http://labshare.cshl.edu/shares/mhammelllab/www-data/TEToolkit/TE_GTF/). Features with an adjusted p value of < 0.05 and a log2 fold change > 1.5 were reported. Overlap (Figure 2.8a) was determined between mRNAs upregulated in *rde-3(-)* animals and published lists of: (1) RNAs targeted by CSR-1–bound endo-siRNAs (Claycomb *et al.* 2009), (2) piRNA-targeted mRNAs (http://cosbi6.ee.ncku.edu.tw/piRTarBase/, Stringent and CLASH list) (Wu *et al.* 2019), and (3) WAGO-class mRNAs (Gu *et al.* 2009).

J. CRISPR

The CRISPR strategy described previously (Arribere *et al.* 2014; Farboud and Meyer 2015; Wan *et al.* 2018) was used to revert the missense mutation in *rde-3(ne298)* animals to wild-type and to tag the N-terminus of *rrf-1* with *ha::tagRFP*. SapTrap cloning (Schwartz and Jorgensen 2016; Dickinson *et al.* 2018) and the selection-based CRISPR strategy described previously (Dickinson *et al.* 2015) was used to tag *rde-3* at the N-terminus with *gfp::degron* and to introduce *3xflag::rde-3* (with 2kb upstream of the ATG and 2kb downstream of the stop codon) at the LGII MosSCI site *ttTi5605 (Frøkjaer-Jensen et al. 2008)* into *rde-3(ne3370)* animals. All guide RNAs were designed using the guide RNA selection tool CRISPOR (Haeussler *et al.* 2016).

K. Small RNA sequencing

rde-1(ne219); oma-1(zu405) animals were injected with an *oma-1* pUG or pGC RNA (*oma-1* mRNA fragment with $(UG)_{18}$ or $(GC)_{18}$ tail) in which the *oma-1* sequence (the first 566nt of *oma-1* mRNA) was modified to contain a SNP in exon 4 (ATTCATCCCG A>T TCATGGACCA). Injection mix was prepared as described above. For P_0 analysis (Figure 2.10d and Figure 2.11a), ~100 *rde-1(ne219)* animals were injected per experiment. After recovering for 1-4 hours at room temperature, injected animals were collected for total RNA extraction. For F_1 analysis (Figure 2.13e), ~20 *rde-1(ne219)* animals were injected per experiment. Injected animals recovered at 15°C for two days and were returned to room temperature. ~500 adult co-injection marker-expressing progeny of injected animals were collected for total RNA extraction. Small RNAs were size-selected, cloned and sequenced as described previously (Dodson and Kennedy 2019). Libraries were sequenced on the Illumina NextSeq500 platform (Biopolymers Facility, Harvard Medical School) to obtain 50bp single-end reads. Note: the same SNP-containing *oma-1* pUG RNA was injected for the experiment described in Figure 2.12d and Figure 2.13d. For this experiment, *rde-1(ne219); oma-1(zu405ts)* animals were injected with either co-injection marker only or the *oma-1(SNP)* pUG RNA with co-injection marker. Injected animals were allowed to recover at 15°C for ~4 days, after which 5-8 co-injection marker-expressing progeny from 5 injected animals were picked (as described above), pooled onto a 10cM plate for each replicate (3 replicates total) and allowed to lay a brood, which was then collected in TRIzol and *oma-1* pUG PCR analysis was performed as described above.

L. Small RNA sequencing analysis

A custom Python script was used to select reads starting with the last 4 nucleotides of the 5' adaptor (either AGCG or CGTC). Cutadapt 1.14 (Martin 2011) was then used to trim the 3' adaptor (CTGTAGGCACCATCAATAGATCGGAAGAGCAC) and the in-line portion of the 5' adaptor (AGCG and CGTC) (both with a minimum phred score = 20), allowing only sequences >= 16nt after trimming to pass (cutadapt -q 20 -m 16 -u 4 -a

CTGTAGGCACCATCAATAGATCGGAAGAGCAC --discard-untrimmed). The quality of the trimming was assessed with FastQC 0.11.5 (Andrews and Others 2010). For downstream analysis, custom Python scripts were used to select reads that were 22nt in length and began with a G (22G siRNA reads). Tophat 2.1.1 (Trapnell *et al.* 2009) was then used to map 22G siRNA reads to the *C. elegans* genome (WBcel235). Gene annotations were obtained from Ensembl (Cunningham *et al.* 2019) (WormBase release WS269) and custom shell scripts were used to select protein-coding genes only. One mismatch was allowed to identify 22G siRNAs with SNPs. Using Samtools v0.1.19, only uniquely mapping sequences were retained. 22G siRNA pileup figures were generated as follows: first, bam files generated from Tophat v2.1.1 (Trapnell *et al.* 2009) were normalized by DeepTools v3.0.2 (Ramírez *et al.* 2014) based on counts per million and only antisense reads were kept for further analysis (bamCoverage -bs 2 --normalizeUsing CPM -samFlagExclide 16). Then, the normalized antisense 22G small RNA sequences (bedGraph files) were visualized using Sushi 1.20.0 (Phanstiel *et al.* 2014) in R. The number of reads mapping antisense to each gene was calculated by featureCounts 1.6.0 (Liao *et al.* 2014) (featureCounts -s 2 -a *.gtf -t exon -g gene_name). All custom scripts used in this section are available at: https://github.com/Yuhan-Fei/pUG-analysis.

M. pUG RNA chromatography

Wild-type adult animals (~1-2 full 10cm plates per experiment) were frozen in liquid nitrogen as small droplets and ground into powder with a mortar and pestle. Powder was dissolved in lysis buffer (5mM HEPES-NaOH(pH7.5), 50mM NaCl, 5mM $MgCl_2$, 0.5mM EDTA (pH8.0), 5% glycerol, 0.25% Triton X-100, 0.5mM DTT, 1mM PMSF, 1 tablet of cOmplete protease inhibitor (Roche, 11697498001)) and rotated for 30 min at 4°C. The resulting lysate was centrifuged at top speed for 10 min at 4°C. Supernatant was distributed evenly among experiments, and RNaseOUT recombinant ribonuclease inhibitor (Invitrogen, 10777019) was added to lysate (1ul per 100ul lysate). For each experiment, 160pmol of biotinylated RNA was conjugated to 400ug Dynabeads

MyOne Streptavidin beads (Invitrogen, 65001) as per manufacturer's instructions. Beads were added to lysates and rotated at room temperature for 1 hour. Beads were separated from supernatant on a magnetic rack, and the supernatant was collected and saved ("sup" fraction). Beads were washed 3 times with lysis buffer and rotated for 5 min at 4°C in lysis buffer. To perform liquid chromatography with tandem mass spectrometry (LC/MS/MS), beads were incubated in 500mM NH_4OH and shaken at 37°C for 20 min after the last wash with lysis buffer. Beads were then pelleted using a magnetic rack and the supernatant was removed and vacuum dried until NH_4OH had completely evaporated. 100ul of 100% TCA and 400ul of pre-chilled water was added to each sample, followed by a 15 min incubation on ice. Samples were spun at top speed at 4°C for 20 min. The supernatant was removed and the previous step was repeated using 1mL of 10% TCA solution. Samples were then washed two times with acetone and centrifuged for 10 min at 4°C after each wash. Acetone was removed and samples were dried in a speed vacuum to remove all residual acetone. Samples were analyzed using liquid chromatography-tandem mass spectrometry (LC-MS/MS, Taplin Mass Spectrometry Facility, Harvard Medical School). Note: a beads-only pull-down served as a control for this experiment. To analyze the LC-MS/MS data, 1 peptide count was assigned to all proteins with 0 peptide counts and the peptide counts for each pull-down sample were then normalized by the total number of peptides identified for that sample. Only proteins with normalized peptide counts that were ≥2-fold more enriched in the $(UG)_{18}$ pull-down versus the beads-only control pull-down were kept for further analysis (Figure 2.10a). For gel electrophoresis and Western blotting experiments, lysates were made from animals harboring *rrf-1* tagged with *ha* and *TagRFP* at the endogenous locus. pUG RNA chromatography was performed as above. After the last wash with lysis buffer, beads ("pull-down" fraction) and "sup" were dissolved in 2x Laemmli sample buffer (Biorad, 1610737, final concentration=1x) with 5% 2-mercaptoethanol and heated for 5 min at 95°C and chilled on ice. "Pull-down" and "sup" fractions

were loaded into 4–15% Mini-PROTEAN TGX Precast protein gels (Biorad, 4561086) and run in Tris-glycine running buffer (25mM Tris, 192mM glycine, 0.1% SDS). Proteins were then transferred to nitrocellulose membrane (BioRad) at 100V for 1 hour in electrotransfer buffer (50mM Tris, 40mM glycine, 9% methanol, 0.2% SDS). Blotted membranes were blocked with 5% milk in PBST (phosphate-buffered saline, 1.0% Tween-20) for 1 hour at room temperature and probed with primary antibody (1:1000 HA-Tag Rabbit mAb, Cell Signaling, #3724, in 5% milk) overnight at 4°C. After washing with PBST 3 times, membrane was probed with secondary antibody (1:10000 IRDye 800CW Goat anti-Rabbit IgG, LI-COR, 926-32211, in 5% milk) for 1 hour at room temperature. Membrane was washed with PBST 3 times before imaging using Odyssey Fc Dual-Mode Imaging System (LI-COR).

N. Heterozygous Experiment

To perform the experiment in Figure 2.12e and Figure 2.13f, embryos were obtained via hypochlorite treatment of *oma-1(zu405ts)* gravid adult hermaphrodites and dropped onto RNAi plates seeded with HT115 bacteria expressing dsRNA targeting *oma-1.* L_4 hermaphrodites were then transferred, along with *rde-3(ne298); oma-1(zu405ts)* males, onto RNAi plates seeded with 25ul (small area of food to encourage mating) of *oma-1* dsRNA-expressing bacteria. Once hermaphrodites were adults, they were singled onto NGM plates seeded with OP50 and allowed to lay F_1 progeny. 12-15 F_1s were singled from 3 independently mated hermaphrodites and genotyped to ensure that they were heterozygous for *rde-3(ne298).* To obtain F_3 animals, 12-15 F_2s per F_1 (verified to be heterozygous for *rde-3(ne298)*) were singled to 15°C (so as to avoid embryonic arrest due to temperature) and allowed to lay a brood. F_2s were then single worm genotyped to identify *rde-3(+)* and *rde(ne298)* homozygous animals. Then, % embryonic arrest was calculated by pooling 5 L_4 stage F_3 animals per F_2 at 20°C until they had laid a brood of 50-200 progeny and counting the # of embryos that were laid vs. hatched on the following day. *rde-*

3(+) and *rde(ne298)* homozygous F_3 broods were pooled for all plates that were derived from the same P_0 and *oma-1* pUG PCR was performed as described above.

O. Code availability

A description of custom scripts used to analyze MiSeq and RNA-seq data is provided in the Methods section and scripts are available upon request. Custom Python scripts used to analyze small RNA sequencing data are deposited at: https://github.com/Yuhan-Fei/pUG-analysis.

P. Data availability

All DNA and RNA sequencing data discussed in this publication have been deposited in NCBI's Gene Expression Omnibus (Edgar *et al.* 2002) and are accessible through GEO Series accession number GSE148134.

VI. Author Contributions

Aditi Shukla contributed to Figures 2.1, 2.2, 2.3a-b, 2.5a, 2.7a-d, 2.7e, 2.8, 2.9, 2.12a, 2.12c-e, 2.13a, 2.13c-d, 2.13f, 2.14, 2.15 and Tables 1, 2, 3. Jenny Yan contributed to Figures 2.3c-d, 2.4, 2.5b, 2.6, 2.7f, 2.10, 2.11, 2.12b, 2.13b, 2.13e. Daniel J. Pagano, Josh Gorham and Jon G. Seidman contributed to Figure 2.8 and Table 3. Anne E. Dodson contributed to Figures 2.2, 2.8a, 2.10 and Table 1. Yuhan Fei contributed to Figures 2.10d, 2.11 and 2.13e.

VII. Acknowledgements

We thank past and present members of the Kennedy, Butcher and Wickens labs for helpful discussions; the Biopolymers Facility at HMS for Illumina sequencing; the Dana-Farber/Harvard Cancer Center DNA Resource Core for Sanger sequencing; and Taplin Mass Spectrometry

Facility in the Cell Biology Department at HMS for performing LC-MS/MS. Some strains were provided by the *Caenorhabditis* Genetics Center (CGC), which is funded by the NIH Office of Research Infrastructure Programs (P40 OD010440). Some strains were provided by the Mitani laboratory through the National BioResource Project (Tokyo, Japan), which is part of the International *C. elegans* Gene Knockout Consortium. A.S. (DGE1144152, DGE1745303) and J.Y. (DGE1745303) were supported by NSF Graduate Research Fellowships. A.E.D. was a Damon Runyon Fellow supported by the Damon Runyon Cancer Research Foundation (DRG-2304-17). D.J.P. was supported by a Ruth L. Kirschstein National Research Service Award (1F32GM125345-01).

VIII. References

Alcazar R. M., R. Lin, and A. Z. Fire, 2008 Transmission dynamics of heritable silencing induced by double-stranded RNA in Caenorhabditis elegans. Genetics 180: 1275–1288.

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

Andrews S., and Others, 2010 FastQC: a quality control tool for high throughput sequence data

Aravind L., and E. V. Koonin, 1999 DNA polymerase beta-like nucleotidyltransferase superfamily: identification of three new families, classification and evolutionary history. Nucleic Acids Res. 27: 1609–1618.

Arribere J. A., R. T. Bell, B. X. H. Fu, K. L. Artiles, P. S. Hartman, *et al.*, 2014 Efficient marker-free recovery of custom genetic modifications with CRISPR/Cas9 in Caenorhabditis elegans. Genetics 198: 837–846.

Ashe A., A. Sapetschnig, E.-M. Weick, J. Mitchell, M. P. Bagijn, *et al.*, 2012 piRNAs can trigger

a multigenerational epigenetic memory in the germline of C. elegans. Cell 150: 88–99.

Austin J., and J. Kimble, 1987 glp-1 is required in the germ line for regulation of the decision between mitosis and meiosis in C. elegans. Cell 51: 589–599.

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

Brenner S., 1974 The genetics of Caenorhabditis elegans. Genetics 77: 71–94.

Buckley B. A., K. B. Burkhart, S. G. Gu, G. Spracklin, A. Kershner, *et al.*, 2012 A nuclear Argonaute promotes multigenerational epigenetic inheritance and germline immortality. Nature 489: 447–451.

Buratti E., and F. E. Baralle, 2001 Characterization and functional implications of the RNA binding properties of nuclear factor TDP-43, a novel splicing regulator of CFTR exon 9. J. Biol. Chem. 276: 36337–36343.

Chen C.-C. G., M. J. Simard, H. Tabara, D. R. Brownell, J. A. McCollough, *et al.*, 2005 A member of the polymerase beta nucleotidyltransferase superfamily is required for RNA interference in C. elegans. Curr. Biol. 15: 378–383.

Claycomb J. M., P. J. Batista, K. M. Pang, W. Gu, J. J. Vasale, *et al.*, 2009 The Argonaute CSR-1 and its 22G-RNA cofactors are required for holocentric chromosome segregation. Cell 139: 123–134.

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

Cunningham F., P. Achuthan, W. Akanni, J. Allen, M. R. Amode, *et al.*, 2019 Ensembl 2019. Nucleic Acids Res. 47: D745–D751.

Czech B., M. Munafò, F. Ciabrelli, E. L. Eastwood, M. H. Fabry, *et al.*, 2018 piRNA-Guided Genome Defense: From Biogenesis to Silencing. Annu. Rev. Genet. 52: 131–157.

Detwiler M. R., M. Reuben, X. Li, E. Rogers, and R. Lin, 2001 Two zinc finger proteins, OMA-1 and OMA-2, are redundantly required for oocyte maturation in C. elegans. Dev. Cell 1: 187–199.

Dickinson D. J., A. M. Pani, J. K. Heppert, C. D. Higgins, and B. Goldstein, 2015 Streamlined Genome Engineering with a Self-Excising Drug Selection Cassette. Genetics 200: 1035–1049.

Dickinson D. J., M. M. Slabodnick, A. H. Chen, and B. Goldstein, 2018 SapTrap assembly of repair templates for Cas9-triggered homologous recombination with a self-excising cassette. MicroPublication Biol. . Dataset 10: W2KT0N.

Dobin A., C. A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski, *et al.*, 2013b STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29: 15–21.

Dodson A. E., and S. Kennedy, 2019 Germ Granules Coordinate RNA-Based Epigenetic Inheritance Pathways. Dev. Cell 50: 704–715.e4.

Edgar R., M. Domrachev, and A. E. Lash, 2002 Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res. 30: 207–210.

Farboud B., and B. J. Meyer, 2015 Dramatic enhancement of genome editing by CRISPR/Cas9 through improved guide RNA design. Genetics 199: 959–971.

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

Fischer S. E. J., E. Wienholds, and R. H. A. Plasterk, 2003 Continuous exchange of sequence information between dispersed Tc1 transposons in the Caenorhabditis elegans genome. Genetics 164: 127–134.

Frøkjaer-Jensen C., M. W. Davis, C. E. Hopkins, B. J. Newman, J. M. Thummel, *et al.*, 2008 Single-copy insertion of transgenes in Caenorhabditis elegans. Nat. Genet. 40: 1375–1383.

Gu W., M. Shirayama, D. Conte Jr, J. Vasale, P. J. Batista, *et al.*, 2009 Distinct argonaute-mediated 22G-RNA pathways direct genome surveillance in the C. elegans germline. Mol. Cell 36: 231–244.

Guang S., A. F. Bochner, D. M. Pavelec, K. B. Burkhart, S. Harding, *et al.*, 2008 An Argonaute transports siRNAs from the cytoplasm to the nucleus. Science 321: 537–541.

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

Haeussler M., K. Schönig, H. Eckert, A. Eschstruth, J. Mianné, *et al.*, 2016 Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR. Genome Biol. 17: 148.

Jin Y., O. H. Tam, E. Paniagua, and M. Hammell, 2015 TEtranscripts: a package for including transposable elements in differential expression analysis of RNA-seq datasets. Bioinformatics 31: 3593–3599.

Jobson M. A., J. M. Jordan, M. A. Sandrof, J. D. Hibshman, A. L. Lennox, *et al.*, 2015 Transgenerational Effects of Early Life Starvation on Growth, Reproduction, and Stress Resistance in Caenorhabditis elegans. Genetics 201: 201–212.

Jose A. M., G. A. Garcia, and C. P. Hunter, 2011 Two classes of silencing RNAs move between Caenorhabditis elegans tissues. Nat. Struct. Mol. Biol. 18: 1184–1188.

Ketting R. F., T. H. A. Haverkamp, H. G. A. M. van Luenen, and R. H. A. Plasterk, 1999 mut-7 of C. elegans, Required for Transposon Silencing and RNA Interference, Is a Homolog of Werner Syndrome Helicase and RNaseD. Cell 99: 133–141.

Ko F. C. F., and K. L. Chow, 2002 A novel thioredoxin-like protein encoded by the C. elegans dpy-11 gene is required for body and sensory organ morphogenesis. Development 129: 1185–1194.

Kuo P.-H., L. G. Doudeva, Y.-T. Wang, C.-K. J. Shen, and H. S. Yuan, 2009 Structural insights into TDP-43 in nucleic-acid binding and domain interactions. Nucleic Acids Res. 37: 1799–1808.

Li H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, *et al.*, 2009 The Sequence Alignment/Map format and SAMtools. Bioinformatics 25: 2078–2079.

Liao Y., G. K. Smyth, and W. Shi, 2014 featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 30: 923–930.

Lin R., 2003 A gain-of-function mutation in oma-1, a C. elegans gene required for oocyte maturation, results in delayed degradation of maternal proteins and embryonic lethality. Dev. Biol. 258: 226–239.

Luteijn M. J., P. van Bergeijk, L. J. T. Kaaij, M. V. Almeida, E. F. Roovers, *et al.*, 2012 Extremely stable Piwi-induced gene silencing in Caenorhabditis elegans. EMBO J. 31: 3422–3430.

Martin G., and W. Keller, 2007 RNA-specific ribonucleotidyl transferases. RNA 13: 1834–1849.

Martin M., 2011 Cutadapt removes adapter sequences from high-throughput sequencing reads.

EMBnet.journal 17: 10–12.

Moore R. S., R. Kaletsky, and C. T. Murphy, 2019 Piwi/PRG-1 Argonaute and TGF-β Mediate Transgenerational Learned Pathogenic Avoidance. Cell 177: 1827–1841.e12.

Pagès H., P. Aboyoun, R. Gentleman, and S. DebRoy, 2017 Biostrings: Efficient manipulation of biological strings. R package version 20.50.2 2.

Pak J., and A. Fire, 2007 Distinct Populations of Primary and Secondary Effectors During RNAi in C. elegans. Science 315: 241–244.

Parrish S., and A. Fire, 2001 Distinct roles for RDE-1 and RDE-4 during RNA interference in Caenorhabditis elegans. RNA 7: 1397–1402.

Phanstiel D. H., A. P. Boyle, C. L. Araya, and M. P. Snyder, 2014 Sushi.R: flexible, quantitative and integrative genomic visualizations for publication-quality multi-panel figures. Bioinformatics 30: 2808–2810.

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

Posner R., I. A. Toker, O. Antonova, E. Star, S. Anava, *et al.*, 2019 Neuronal Small RNAs Control Behavior Transgenerationally. Cell 177: 1814–1826.e15.

Preston M. A., D. F. Porter, F. Chen, N. Buter, C. P. Lapointe, *et al.*, 2019 Unbiased screen of RNA tailing activities reveals a poly(UG) polymerase. Nat. Methods 16: 437–445.

Quinlan A. R., and I. M. Hall, 2010 BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26: 841–842.

Ramírez F., F. Dündar, S. Diehl, B. A. Grüning, and T. Manke, 2014 deepTools: a flexible platform for exploring deep-sequencing data. Nucleic Acids Res. 42: W187–91.

Rechavi O., L. Houri-Ze'evi, S. Anava, W. S. S. Goh, S. Y. Kerk, *et al.*, 2014 Starvation-induced transgenerational inheritance of small RNAs in C. elegans. Cell 158: 277–287.

Remy J.-J., 2010 Stable inheritance of an acquired behavior in Caenorhabditis elegans. Curr. Biol. 20: R877–8.

RStudio Team, 2016 *RStudio: Integrated Development for R*.

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

Schindelin J., I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, *et al.*, 2012 Fiji: an open-source platform for biological-image analysis. Nat. Methods 9: 676–682.

Schott D., I. Yanai, and C. P. Hunter, 2014 Natural RNA interference directs a heritable response to the environment. Sci. Rep. 4: 7387.

Schwartz M. L., and E. M. Jorgensen, 2016 SapTrap, a Toolkit for High-Throughput CRISPR/Cas9 Gene Modification in Caenorhabditis elegans. Genetics 202: 1277–1288.

Shirayama M., M. Seth, H.-C. Lee, W. Gu, T. Ishidate, *et al.*, 2012 piRNAs Initiate an Epigenetic Memory of Nonself RNA in the C. elegans Germline. Cell 150: 65–77.

Sijen T., J. Fleenor, F. Simmer, K. L. Thijssen, S. Parrish, *et al.*, 2001 On the role of RNA amplification in dsRNA-triggered gene silencing. Cell 107: 465–476.

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

Smith T., A. Heger, and I. Sudbery, 2017 UMI-tools: modeling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy. Genome Res. 27: 491–499.

Tabara H., M. Sarkissian, W. G. Kelly, J. Fleenor, A. Grishok, *et al.*, 1999 The rde-1 gene, RNA interference, and transposon silencing in C. elegans. Cell 99: 123–132.

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

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

Trapnell C., L. Pachter, and S. L. Salzberg, 2009 TopHat: discovering splice junctions with RNA-Seq. Bioinformatics 25: 1105–1111.

Tsai H.-Y., C.-C. G. Chen, D. Conte Jr, J. J. Moresco, D. A. Chaves, *et al.*, 2015 A ribonuclease coordinates siRNA amplification and mRNA cleavage during RNAi. Cell 160: 407–419.

Vastenhouw N. L., K. Brunschwig, K. L. Okihara, F. Müller, M. Tijsterman, *et al.*, 2006 Gene expression: long-term gene silencing by RNAi. Nature 442: 882.

Voronina E., G. Seydoux, P. Sassone-Corsi, and I. Nagamori, 2011 RNA granules in germ cells. Cold Spring Harb. Perspect. Biol. 3. https://doi.org/10.1101/cshperspect.a002774

Wagih O., 2017 ggseqlogo: a versatile R package for drawing sequence logos. Bioinformatics 33: 3645–3647.

Wan G., B. D. Fields, G. Spracklin, A. Shukla, C. M. Phillips, *et al.*, 2018 Spatiotemporal regulation of liquid-like condensates in epigenetic inheritance. Nature 557: 679–683.

Wu W.-S., J. S. Brown, T.-T. Chen, Y.-H. Chu, W.-C. Huang, *et al.*, 2019 piRTarBase: a database of piRNA targeting sites and their roles in gene regulation. Nucleic Acids Res. 47: D181–D187.

Xiong Y., and T. A. Steitz, 2004 Mechanism of transfer RNA maturation by CCA-adding enzyme without using an oligonucleotide template. Nature 430: 640–645.

Yigit E., P. J. Batista, Y. Bei, K. M. Pang, C.-C. G. Chen, *et al.*, 2006 Analysis of the C. elegans Argonaute family reveals that distinct Argonautes act sequentially during RNAi. Cell 127: 747–757.

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

Chapter 3

---

piRNAs coordinate poly(UG) tailing to prevent aberrant and permanent silencing of germline RNAs

This chapter consists primarily of unpublished work that we are preparing to submit as the manuscript:

Shukla, A., Perales, R. & Kennedy, S. piRNAs coordinate poly(UG) tailing to prevent aberrant and permanent silencing of germline RNAs.

This manuscript was written by Scott Kennedy and Aditi Shukla. Experiments in this manuscript were performed by Aditi Shukla. All figures were made and assembled by Aditi Shukla. Mutant strains characterized in this study were generously provided by Roberto Perales.

I. Abstract

Small noncoding RNAs such as piRNAs, miRNAs, and siRNAs are important regulators of gene expression. In some cases, small RNA-based gene regulation can be inherited via the germline for multiple generations (termed transgenerational epigenetic inheritance or TEI). A feed-forward cascade of siRNA synthesis from poly(UG)-tailed (pUG-tailed) mRNAs followed by siRNA-directed pUG RNA biogenesis (henceforth pUG RNA/siRNA loops) mediates TEI in the *C. elegans* germline. How *C. elegans* ensures accuracy and fidelity of this feed-forward, heritable and, therefore, inherently dangerous gene regulatory system is unknown. Here we show that the *C. elegans* piRNA system prevents germline-expressed genes from becoming permanently inactivated by the pUG RNA/siRNA TEI pathway. In the absence of the *C. elegans* piRNA system, a subset of germline-expressed genes, which are not normally subjected to TEI, enter a state of permanent epigenetic silencing. Entry into the silenced state is irreversible, and genes thus silenced are paramutagenic. Permanent silencing and paramutation is mediated by perpetual activation of pUG RNA/siRNA loops. Our results show that one function of *C. elegans* piRNAs is to insulate germline-expressed genes from aberrant and runaway inactivation by the pUG RNA/siRNA epigenetic inheritance system.

II. Introduction

Small noncoding RNAs, such as PIWI-interacting RNAs (piRNAs), microRNAs, and small interfering RNAs (siRNAs) are key regulators of gene expression in eukaryotes. Small RNAs are bound by Argonaute proteins and, together, these ribonucleoprotein complexes target complementary mRNAs for silencing (Filipowicz 2005; Meister 2013). piRNAs are an animal-specific class of genomically encoded and germline-expressed small noncoding RNAs, which are bound by the PIWI clade of Argonaute proteins and are essential for germ cell function in most animals. One widely conserved function of piRNAs is the silencing of mobile genetic elements termed transposons (Siomi *et al.* 2011; Ozata *et al.* 2019).

The function of some piRNAs, however, remains elusive. For instance, most mammalian piRNAs do not exhibit complementarity to transposable elements (Aravin *et al.* 2006, 2008; Girard *et al.* 2006; Li *et al.* 2013). Additionally, loss of Miwi, one of three PIWI proteins encoded by the mouse genome, does not cause transposon upregulation, but does result in male sterility, indicating that Miwi and, therefore, piRNAs have important functions unrelated to transposon silencing (Vourekas *et al.* 2012). The biological functions of the *C. elegans* piRNAs (also known as 21U-RNAs (Ruby *et al.* 2006)) are also somewhat mysterious. *C. elegans* piRNAs are transcribed as 28-29nt long capped transcripts that are processed into 21nt uncapped piRNAs that are 2'-O-methylated, initiate with uracil, and are bound and stabilized by PRG-1, one of two PIWI clade Argonaute proteins in *C. elegans. C. elegans* piRNAs are not required for the silencing of most *C. elegans* transposable elements (Wang and Reinke 2008; Batista *et al.* 2008; Das *et al.* 2008; Bagijn *et al.* 2012; Barucci *et al.* 2020; Reed *et al.* 2020). Indeed, fewer than 1% of transposon families are upregulated in a *prg-1* mutant, which lacks all piRNAs (Barucci *et al.* 2020). PRG-1 and piRNAs are required for the silencing of some germline-expressed mRNAs, such as *bath-45* (Bagijn *et al.* 2012), as well as for transgene silencing (Shirayama et al., 2012; Ashe et al., 2012). Current models posit that PRG-1/piRNA complexes bind target mRNAs and

recruit RNA-dependent RNA Polymerases (RdRPs) (Lee *et al.* 2012; Bagijn *et al.* 2012), which use the target mRNA as a template to produce siRNAs, which are antisense to mRNA templates, 22nt in length, and begin with a guanosine (termed 22G-siRNAs or secondary (2°) siRNAs) (Gu *et al.* 2009). piRNA-dependent 22G-siRNAs are bound by a worm-specific clade of Argonaute proteins, called WAGOs and, together, this complex can mediate target mRNA silencing (Gu *et al.* 2009). Somewhat surprisingly, PRG-1 and its bound piRNAs were recently shown to interact with >16,000 mostly germline-expressed mRNAs (Shen *et al.* 2018; Zhang *et al.* 2018). However, fewer than 100 of these mRNAs undergo piRNA-dependent gene silencing (Barucci *et al.* 2020). Instead, the expression of the majority of these mRNAs is unaffected in *prg-1* mutants, indicating that PRG-1 and piRNAs do not silence most of the mRNAs to which they are bound. Indeed, recent studies show that *C. elegans* piRNAs actually protect some mRNAs from aberrant siRNA-mediated gene silencing (de Albuquerque *et al.* 2015; Phillips *et al.* 2015; Barucci *et al.* 2020; Reed *et al.* 2020) such that, in the absence of piRNAs, some germline-expressed genes, most prominently the replication-dependent histones genes, undergo aberrant gene silencing (Barucci *et al.* 2020; Reed *et al.* 2020). Why some genes get aberrantly silenced in a *prg-1* mutant and how PRG-1 and piRNAs normally promote expression of these genes remains a mystery.

The PRG-1/piRNA-induced silencing of some loci, such as transgenes, is heritable (Ashe *et al.* 2012; Shirayama *et al.* 2012). Transgene silencing initially requires PRG-1 and piRNAs. However, once silencing is established, it can be maintained in the absence of PRG-1 and piRNAs for many generations. Such heritable gene silencing, in the absence of initiating triggers, is an example of transgenerational epigenetic inheritance (TEI), which in *C. elegans* is also known as RNA-induced epigenetic silencing (RNAe) (Ashe *et al.* 2012; Shirayama *et al.* 2012). piRNA-induced transgenerational silencing correlates with heritable expression of 22G-siRNAs and histone 3 lysine 9 trimethylation (H3K9me3) of the genomic locus and depends on nuclear factors, such as the nuclear WAGO HRDE-1, the HP1 homolog HPL-2, and histone methyltransferases.

Double-stranded RNAs (dsRNAs) can also initiate TEI in *C. elegans* (Vastenhouw *et al.* 2006; Ashe *et al.* 2012; Shirayama *et al.* 2012; Buckley *et al.* 2012; Luteijn and Ketting 2013; Sapetschnig *et al.* 2015). dsRNA initiates TEI via a conserved gene silencing program termed RNA interference (RNAi) (Fire *et al.* 1998). In *C. elegans*, RNAi begins when dsRNA is processed by the endoribonuclease Dicer and the double-stranded RNA binding protein RDE-4 into short dsRNAs (Parrish and Fire 2001; Tabara *et al.* 2002; Parker *et al.* 2006). One strand of this short dsRNA (termed 1° siRNA) is then bound by the Argonaute (AGO) RDE-1 (Tabara *et al.* 1999; Parrish and Fire 2001) and together, this siRNA/AGO complex recognizes a target mRNA via Watson-Crick base-pairing and the target mRNA is cleaved by the ribonuclease RDE-8 (Tsai *et al.* 2015). RdRPs then generate 22G-siRNAs that are antisense to the target mRNA and these 22 siRNAs are bound by WAGO proteins, which carry out gene silencing (Yigit *et al.* 2006; Gu *et al.* 2009). RNAi-directed silencing of germline-expressed mRNAs can be inherited (termed RNAi inheritance) and inheritance is correlated with the inheritance of 22G-siRNAs, some of which bind the nuclear AGO HRDE-1, as well as the inheritance of repressive chromatin modifications, such as H3K9me3 (Ashe *et al.* 2012; Shirayama *et al.* 2012; Buckley *et al.* 2012) and H3K27me3 (Mao *et al.* 2015), at loci undergoing dsRNA-directed TEI. The shared requirements for RdRPs and WAGOs, such as HRDE-1, between RNAi-directed and piRNA-directed TEI suggests that these two TEI pathways converge on a common set of gene silencing effector proteins. Other factors required for RNAi inheritance, such as the Vasa homolog GLH-1 (Spracklin *et al.* 2017) and the NFX1-type zinc finger-containing protein 1 homolog, ZNFX-1 (Wan *et al.* 2018), localize to germline biomolecular condensates (Brangwynne *et al.* 2009; Wan *et al.* 2018) termed P granules (Gruidl *et al.* 1996) and Z granules (Wan *et al.* 2018), respectively, suggesting that these condensates play important roles in RNAi inheritance.

The maintenance of siRNA expression over generations in *C. elegans* during dsRNA-triggered TEI depends upon a recently discovered noncoding RNA modification (Shukla *et al.* 2020). mRNAs targeted by RNAi and cleaved by the ribonuclease RDE-8 (Tsai *et al.* 2015) are

modified with perfectly alternating 3' uridine (U) and guanosine (G) repeats (termed poly(UG) or pUG tails) (Shukla *et al.* 2020). RNAs modified with pUG tails (pUG RNAs) serve as templates for RdRPs, which use pUG RNAs to generate 22G-siRNAs. Further, we recently showed that generationally repeated rounds of pUG RNA-templated 22G-siRNA synthesis and 22G-siRNA-directed mRNA pUGylation (termed pUG RNA/siRNA cycling) drive RNAi inheritance in *C. elegans* (Shukla *et al*. 2020). Although not yet tested, piRNA-directed TEI may also depend upon pUG/siRNA cycling, as the poly(UG) polymerase RDE-3 is known to be required for piRNA-directed TEI (Shirayama *et al.* 2012). The feed-forward and theoretically irreversible nature of the pUG/siRNA cycling pathway hints that systems likely exist to regulate this potentially dangerous pathway to prevent mistargeting of essential genes for generationally stable silencing. In support of this idea, RNAi inheritance in *C. elegans* is usually finite, typically only lasting a few generations (Ashe *et al.* 2012; Buckley *et al.* 2012; Lev *et al.* 2017; Spracklin *et al.* 2017; Perales *et al.* 2018). Additionally, recent studies have shown that the methyltransferase MET-2 (Lev *et al.* 2017) and the chromodomain protein HERI-1 (Perales *et al.* 2018) limit the generational perdurance of RNAi inheritance in *C. elegans*. Interestingly, HERI-1 is physically recruited to the chromatin of genes undergoing RNAi inheritance, suggesting that HERI-1 may play a direct role in limiting TEI (Perales *et al.* 2018). Little else is known about how *C. elegans* regulates and focuses its potent, feed-forward, and potentially dangerous TEI pathways.

Here we show that piRNAs limit the duration of pUG/siRNA-based transgenerational gene silencing in the *C. elegans* germline. In the absence of piRNAs, a number of germline-expressed genes are inappropriately targeted for poly(UG) tailing, which causes these mRNAs to become permanent targets of self-perpetuating pUG/siRNA silencing loops. We conclude that one function of *C. elegans* piRNAs is to protect germline-expressed genes from aberrant and runway gene silencing by the pUG/siRNA silencing pathway.

III. Results

A. Identification of mutations that cause RNAi to become essentially permanent

We previously conducted a forward genetic screen to identify factors that normally limit the duration of dsRNA-directed TEI in *C. elegans* (Perales *et al.* 2018). For this screen, we mutagenized *oma-1(zu405); pie-1::gfp::h2b* animals, which harbor: (1) *zu405* (hereafter referred to as *zu405ts*), a temperature-sensitive, gain-of-function mutation in the germline-expressed gene o*ma-1* that results in embryonic arrest, unless *oma-1* is silenced by RNAi (Lin 2003); and (2) the *pie-1::gfp::h2b* transgene, which encodes a GFP::H2B fusion protein driven by a germline-expressed promoter (hereafter referred to as *gfp*). In wild-type animals, dsRNA-induced silencing of *oma-1* or *gfp* lasts for four to ten generations after initiating dsRNA triggers are removed (Ashe *et al.* 2012; Buckley *et al.* 2012; Spracklin *et al.* 2017; Perales *et al.* 2018). Our screen identified 20 mutations that extended *oma-1* and *gfp* RNAi inheritance for seven or more generations longer than in non-mutagenized controls (Perales *et al.* 2018). For 18 of the 20 identified mutations, o*ma-1* and *gfp* expression was restored in 20 or fewer generations after RNAi (Perales *et al.* 2018). However, two mutants, *gg531* and *gg540*, were unique in that, more than five years, and hundreds of generations, after *oma-1* and *gfp* were initially targeted for silencing by RNAi, 100% of animals in these lineages continued to silence both reporter genes (Figure 3.1). Data presented below will show that the silencing observed in these mutants is epigenetic. We conclude that genetically-encoded systems exist that prevent dsRNA-initiated epigenetic inheritance from becoming essentially permanent. Henceforth, we refer to this remarkably stable epigenetic inheritance as "permanent silencing" and genes undergoing silencing in these lineages as "permanently silenced alleles."

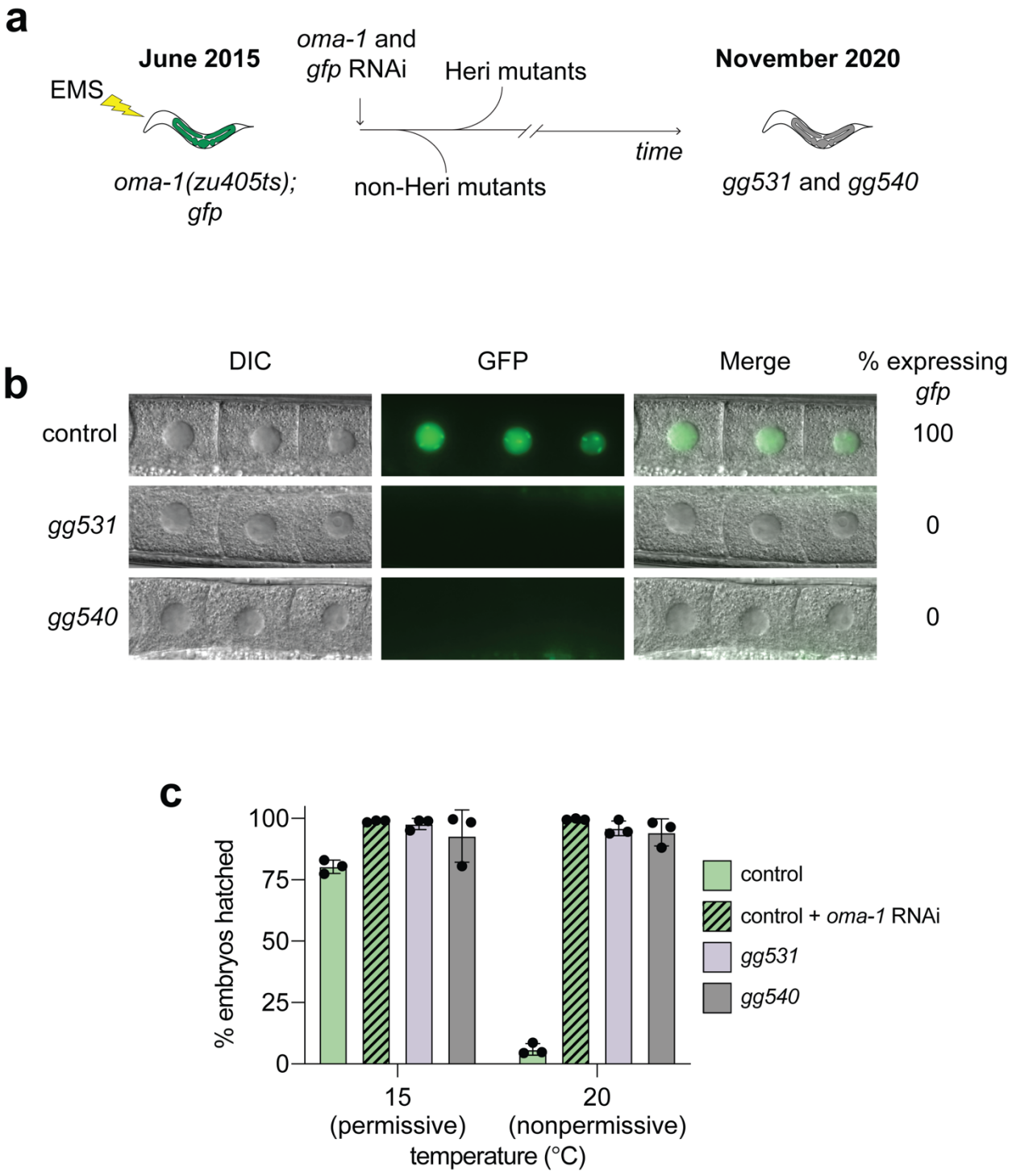


Figure 3.1. RNAi inheritance can become permanent in *C. elegans*. a, *oma-1(zu405ts); pie-1::gfp::h2b* (hereafter, *gfp*) animals were mutagenized with EMS to look for mutations that extended *oma-1* and *gfp* RNAi inheritance by seven or more generations (Perales *et al.* 2018). *oma-1(zu405ts)* is a temperature-sensitive, gain-of-function allele of *oma-1* that results in 100% embryonic arrest unless *oma-1* is silenced with RNAi (Lin 2003). We employed *oma-1(zu405ts)*

Figure 3.1 (Continued). so that we could select for animals that continued to silence *oma-1* after non-mutagenized controls had reset *oma-1* expression and, therefore, stopped laying viable progeny. We then scored these surviving mutants for *gfp* expression. For 18 of the 20 Heritable enhancers of RNAi (Heri) mutants that we isolated, o*ma-1* and *gfp* expression returned to normal in 20 or fewer generations. Two mutants, *gg531* and *gg540*, however, continue to silence *oma-1* and *gfp*, more than five years after they were treated with *oma-1* and *gfp* RNAi. **b,** Fluorescence micrographs showing *gfp* expression in the oocyte nuclei of control (e.g. non-mutagenized *oma-1(zu405ts); gfp*) animals vs. *gg531* and *gg540* mutants. **c,** *oma-1(zu405ts)* animals can lay viable progeny at 15°C (permissive temperature), but lay arrested embryos at 20°C (nonpermissive temperature) unless *oma-1* is silenced (Lin 2003). % embryos hatched (e.g. # of hatched embryos / total # embryos laid) can be used as a proxy for *oma-1* silencing. % embryonic arrest was measured for control (e.g. non-mutagenized *oma-1(zu405ts); gfp*) animals +/- *oma-1* RNAi and for *gg531* and *gg540* mutants. Each point represents a biological replicate. Each biological replicate consisted of % embryonic arrest scored for 6 individual animals. Error bars represent standard deviation (s.d.) of the mean.

B. PRG-1 limits RNAi inheritance

Whole-genome sequencing of *gg531* and *gg540* animals identified independent nonsense mutations in the gene *prg-1* (Figure 3.2a), which encodes the major *C. elegans* PIWI clade Argonaute protein (Wang and Reinke 2008; Batista *et al.* 2008; Das *et al.* 2008). PRG-1 binds to piRNAs in *C. elegans* and initiates transgenerational silencing of transgenes via a nuclear RNAi mechanism involving piRNA-directed RdRP-based synthesis of siRNAs (Ashe *et al.* 2012; Shirayama *et al.* 2012). Our data suggests that, surprisingly, PRG-1 and piRNAs can also limit transgenerational gene silencing initiated by exogenous dsRNA triggers. To test this idea further, we asked if animals harboring an independently isolated deletion allele, *tm872*, of *prg-1* exhibited an enhanced RNAi inheritance phenotype. *tm872* is a 640bp deletion (Figure 3.2a) in *prg-1* that removes most of the MID domain and part of the PIWI domain, two conserved domains found in Argonaute proteins (Meister 2013), and is thought to represent a null allele of *prg-1* (Wang and Reinke 2008). RNAi inheritance was enhanced in *prg-1(tm872)* animals (Figure 3.2b, Figure 3.3a). Finally, we outcrossed *prg-1(gg531)* from one of our two permanently silenced mutants, regenerated *prg-1(gg531)*; *gfp* animals, which expressed *gfp,* and then tested these animals for *gfp* RNAi inheritance. Indeed, *prg-1(gg531)* animals displayed enhanced *gfp* RNAi inheritance

(Figure 3.3b). These data confirm that the loss of PRG-1 can enhance the generational perdurance of RNAi-induced gene silencing.

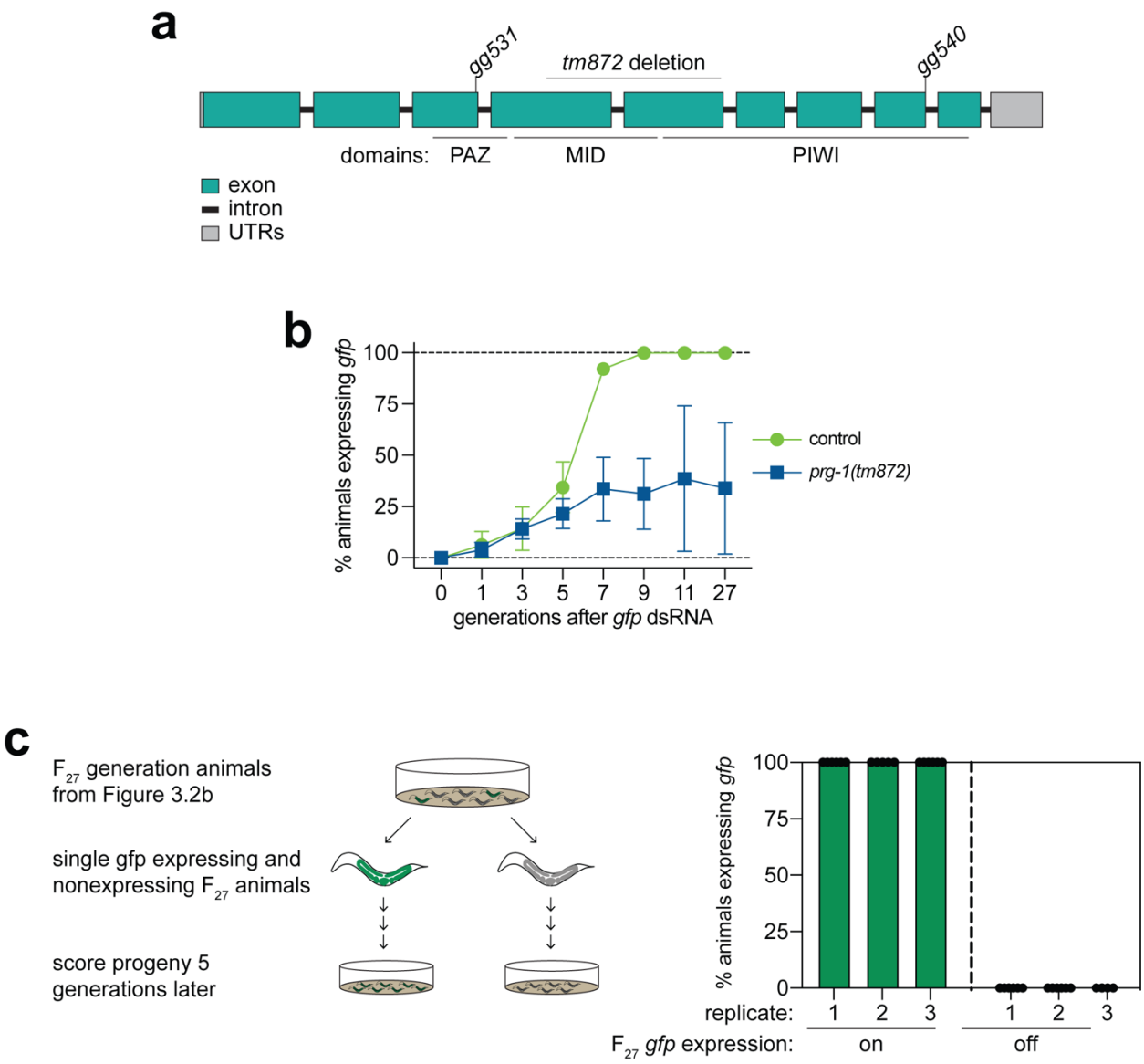


Figure 3.2. PRG-1 prevents permanent RNAi inheritance. a, The *prg-1* gene encodes a PIWI Argonaute protein. A schematic of the *prg-1* gene is shown, with the locations of the conserved PAZ, MID and PIWI domains (Meister 2013). *prg-1(gg531)* and *prg-1(gg540)* are two nonsense alleles and *prg-1(tm872)* encodes a partial deletion in *prg-1.* Gene features and coordinates were obtained from WormBase (Harris *et al.* 2020). **b,** *gfp* RNAi inheritance assay was performed on *gfp* expressing control vs. *prg-1(tm872)* animals. Animals were fed *gfp* dsRNA and the percentage of animals expressing *gfp* was scored for the indicated number of generations after dsRNA treatment. Three biological replicates of this experiment were performed, with >50 animals

Figure 3.2 (Continued). counted per replicate for each genotype every generation. Error bars represent s.d. of the mean. **c,** 5-6 *gfp* expressing and nonexpressing *prg-1(tm872)* animals (represented by each dot) were singled 27 generations after *gfp* dsRNA treatment from each biological replicate in the experiment shown in Figure 3.2b. Lineages established from these animals were allowed to grow for 5 generations and then *gfp* expression was scored.

In Figure 3.2b, some, but not all, *prg-1(tm872)* animals continued to inherit *gfp* silencing 27 generations after dsRNA treatment. Given that the *prg-1* mutants we identified in our genetic screen showed 100% penetrant silencing, we wondered if, following *gfp* RNAi, a subset of *prg-1(tm872)* animals enter a state of permanent silencing in which they, and all their progeny, exhibit fully penetrant gene silencing. To test this idea, we isolated individual *prg-1(tm872); gfp* animals, 27 generations after *gfp* RNAi, in which *gfp* was either expressed or silenced and then monitored *gfp* expression in lineages established from these individuals for an additional five generations. This analysis showed that 27 generations after RNAi, individuals had entered one of two epigenetic states: either they and all of their progeny expressed *gfp* or they and all their progeny did not (Figure 3.2c, Figure 3.3c). For unknown reasons, the percentage of *prg-1* animals entering the permanently silenced state can vary in different experiments (Figure 3.2b and Figure 3.3a,b). We conclude that, in the absence of PRG-1, dsRNA triggers one of two epigenetic states: finite gene silencing, in which gene expression is eventually restored after several generations, or permanent gene silencing, in which targets of dsRNA are silenced forever.

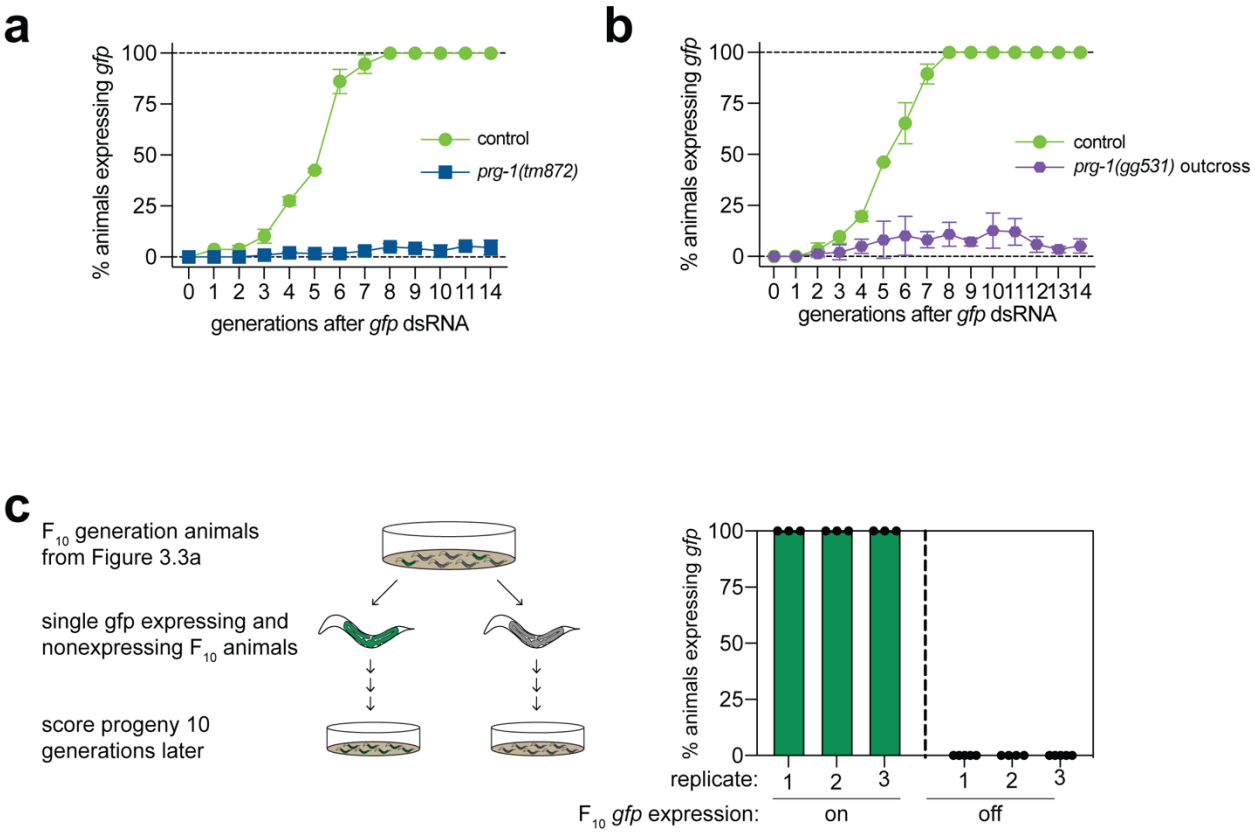


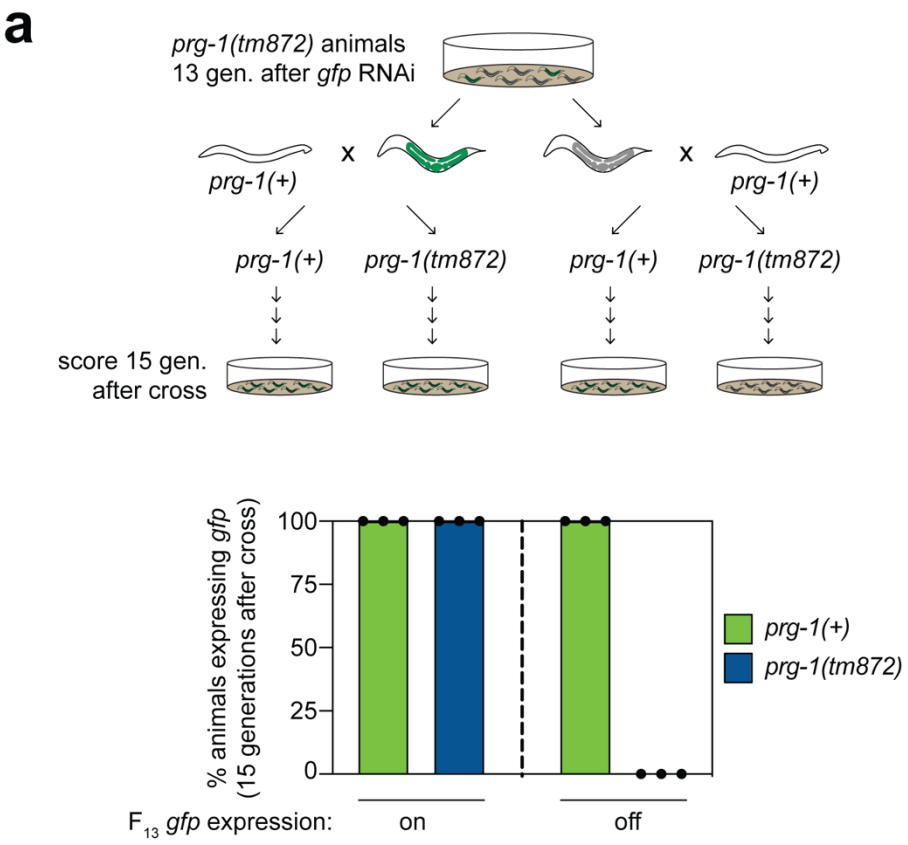
Figure 3.3. PRG-1 limits the generational perdurance of *gfp* RNAi inheritance. a, *gfp* RNAi inheritance assay was performed on *gfp* expressing control vs. *prg-1(tm872)* animals. Animals were fed *gfp* dsRNA and the percentage of animals expressing *gfp* was scored at the indicated number of generations after dsRNA treatment. Three biological replicates of this experiment were performed, with >50 animals counted per replicate for each genotype every generation. Note: this inheritance assay was performed two years before the inheritance assay shown in Figure 3.2b. This experiment employed a *prg-1(tm872)* strain that was nearing 100% sterility, so it was outcrossed with *gfp* expressing animals to generate the strain used for Figure 3.2b. **b,** *prg-1(gg531)* mutant animals were outcrossed to remove the silent *gfp* allele and to reintroduce an expressed *gfp* allele. *gfp* inheritance assay was then performed with *prg-1(gg531)* animals harboring this expressed *gfp* allele and control (*gfp* expressing) animals. **c,** 3-5 *gfp* expressing and nonexpressing *prg-1(tm872)* animals (represented by each dot) were singled 10 generations after *gfp* dsRNA treatment from each biological replicate in the experiment shown in Figure 3.3a. Lineages established from these animals were allowed to grow for 10 additional generations and then *gfp* expression was scored.

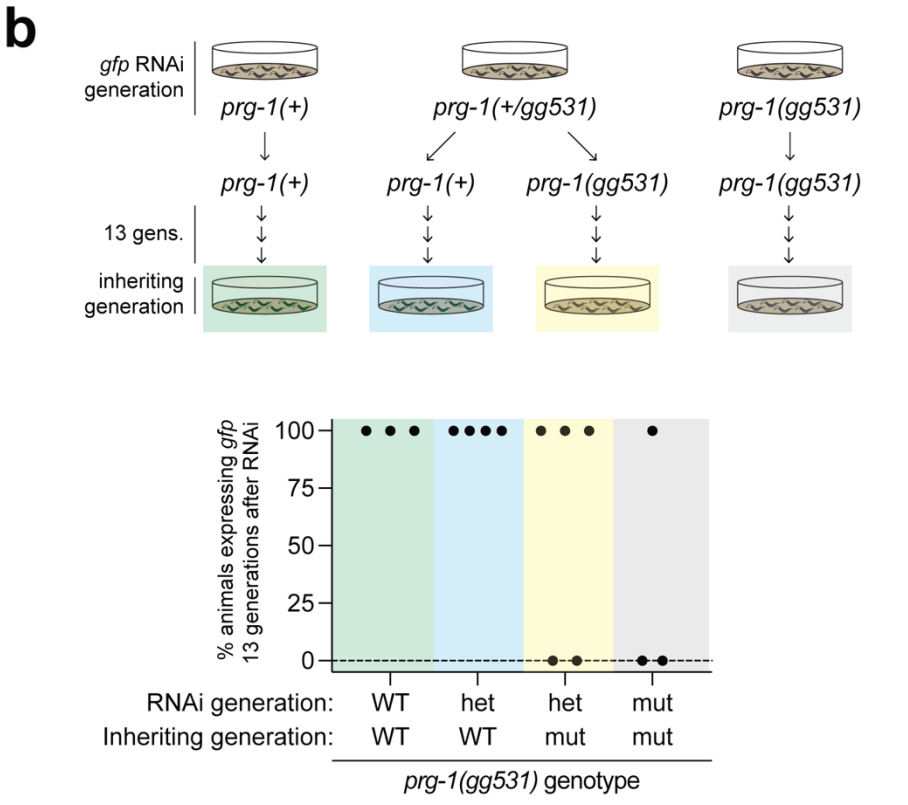
We next asked if PRG-1 normally inhibits the initiation or maintenance of permanent RNAi silencing. We crossed *gfp* nonexpressing *prg-1(tm872)* animals from the 13th generation of a *gfp* inheritance assay to wild-type animals. We then monitored *gfp* expression in *prg-1(+)* and *prg-*

1(tm872) progeny of this cross, which were homozygous for *gfp.* Only *prg-1(tm872)* progeny continued to silence *gfp* ~15 generations after this cross (Figure 3.4a). Thus, PRG-1 inhibits maintenance of permanent RNAi-induced silencing. To ask if PRG-1 inhibits the initiation of permanent silencing, we treated *prg-1(tm872/+)* heterozygous animals, which are wild-type for PRG-1 activity (Ashe *et al.* 2012), with *gfp* RNAi and scored the inheritance of *gfp* silencing in *prg-1*(*+)* or *prg-1*(*tm872)* homozygous progeny of this cross. After ten generations, most *prg-1(tm872)* animals were still inheriting *gfp* silencing and no *prg-1(+)* progeny maintained *gfp* silencing (Figure 3.4b). Taken together, these results: (1) show that PRG-1 does not inhibit the initiation of permanent silencing, and (2) confirm that PRG-1 inhibits the maintenance of *gfp* silencing after *gfp* RNAi. Thus, PRG-1 acts in inheriting generations to limit the long-term maintenance of permanent silencing.

C. elegans PRG-1 possesses endonuclease (Slicer) activity *in vitro* (Bagijn *et al.* 2012), which is dependent on an evolutionarily conserved DDH motif (catalytic triad) (Bagijn *et al.* 2012; Meister 2013). The biological function of PRG-1 Slicer activity is not known. CRISPR/Cas9 mediated mutation of the catalytic triad, in a manner that disrupted PRG-1 Slicer activity *in vitro*, did not result in permanent gene silencing after RNAi (Figure 3.5). Thus, Slicer activity is not required for PRG-1 to limit RNAi inheritance and, therefore, the purpose of PRG-1 slicing remains enigmatic. Further, whereas our data show that PRG-1 suppresses permanent silencing triggered by exogenous dsRNA, the other *C. elegans* PIWI clade Argonaute, PRG-2, whose function is not yet known (Wang and Reinke 2008; Batista *et al.* 2008; Das *et al.* 2008), did not have a role in limiting RNAi inheritance (Figure 3.6).

Figure 3.4. PRG-1 antagonizes the maintenance of heritable silencing. a, *prg-1(tm872)* animals were fed *gfp* dsRNA. 3 *gfp* expressing and nonexpressing animals were singled 13 generations after dsRNA treatment and crossed to WT animals. Lineages (represented by each dot) were established from one *prg-1(+)* and one *prg(tm872)* F_2 animal (all homozygous for *gfp*) from each cross and allowed to grow for 15 generations, after which *gfp* expression was scored. **b,** *prg-1(+), prg-1(gg531)* or *prg-1(gg531/+)* animals were fed *gfp* dsRNA. Lineages were then established by singling dsRNA-fed animals and allowing them to lay a brood. # of animals singled per genotype is represented by each dot. Progeny of *prg-1(gg531/+)* were genotyped for *prg-1(gg531)* after they had been singled and laid their progeny. *gfp* expression was then scored in each lineage 13 generations after dsRNA treatment.


a
prg-1(tm872) animals
13 gen. after gfp RNAi
x
x
prg-1(+)
prg-1(+)
prg-1(+)
prg-1(tm872)
prg-1(+)
prg-1(tm872)
score 15 gen.
after cross
% animals expressing gfp
(15 generations after cross)
100
75
50
25
0
prg-1(+)
prg-1(tm872)
F13 gfp expression:
on
off


b
gfp RNAi
generation
prg-1(+)
prg-1(+/gg531)
prg-1(gg531)
prg-1(+)
prg-1(+)
prg-1(gg531)
prg-1(gg531)
13 gens.
inheriting
generation
% animals expressing gfp
13 generations after RNAi
100
75
50
25
0
RNAi generation:
WT
het
het
mut
Inheriting generation:
WT
WT
mut
mut
prg-1(gg531) genotype

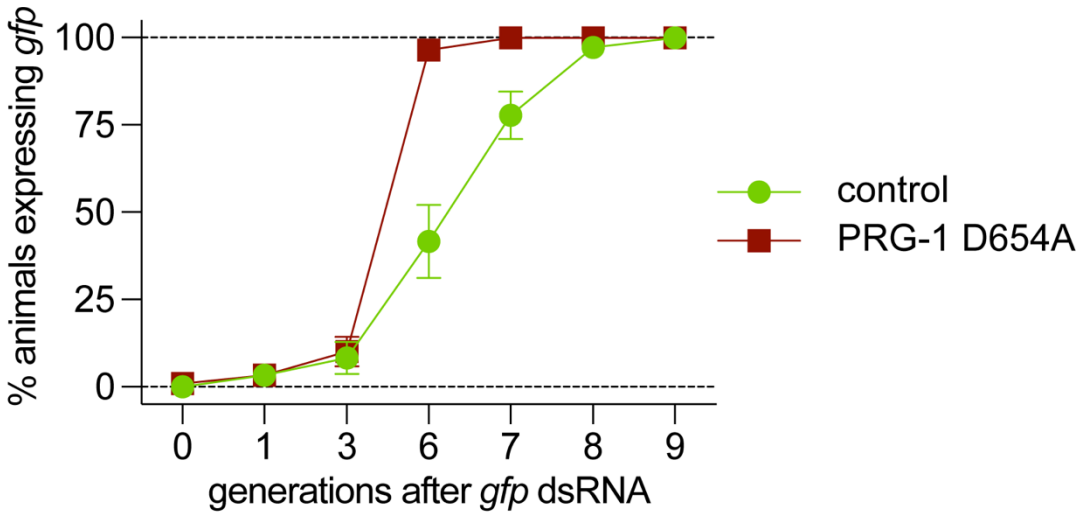


Figure 3.5. PRG-1 prevents permanent RNAi inheritance independent of its Slicer activity. An evolutionarily conserved DDH motif (catalytic triad) mediates the Slicer activity of Argonaute proteins (Bagijn *et al.* 2012; Meister 2013). CRISPR/Cas9 was used to mutate the DDH catalytic triad in PRG-1 to DAH, a mutation previously shown to abolish PRG-1 Slicer activity *in vitro* (Bagijn *et al.* 2012). A *gfp* RNAi inheritance assay was then performed on *gfp* expressing control animals and animals harboring this DAH mutation. >50 animals were counted each generation for each genotype. Error bars represent s.d. of the mean of three independent biological replicates.

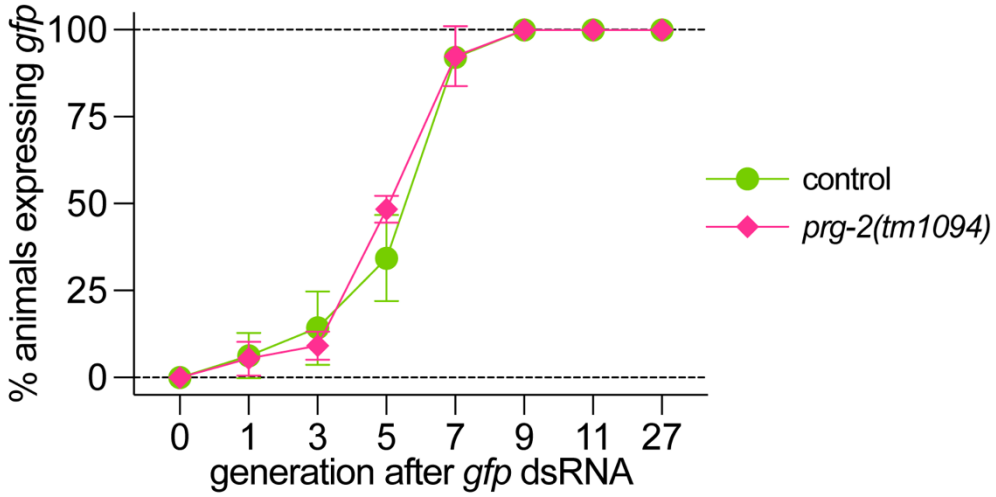


Figure 3.6. PRG-2 does not limit RNAi inheritance. PRG-2 is the second PIWI Argonaute expressed in the *C. elegans* genome, but does not seem to bind *C. elegans* piRNAs (Wang and Reinke 2008; Batista *et al.* 2008; Das *et al.* 2008). *gfp* RNAi inheritance assay was performed on *gfp* expressing control vs. *prg-2(tm1094)* animals. >50 animals were counted each generation for each genotype. Error bars represent s.d. of the mean of three independent biological replicates.

C. Permanently silenced alleles are paramutagenic

Paramutation is an epigenetic gene silencing phenomenon whereby the epigenetic state of one allele is transmitted to another allele of the same gene (Chandler 2010). First documented in plants (Chandler 2010), a related process occurs in the *C. elegans* germline, where it is also known as RNA epigenetic (RNAe) (Shirayama *et al.* 2012). While conducting genetic crosses with *gg531* and *gg540* animals, we noticed that the permanently silenced *oma-1* and *gfp* alleles in *gg531* and *gg540* animals exhibited paramutagenic properties. For instance, when we crossed *prg-1(gg531)* animals harboring permanently silenced *oma-1* or *gfp* alleles to animals harboring expressed alleles of *oma-1* or *gfp* (identified by linked mutations in *dpy-20* or *dpy-10*, respectively), expressed alleles were converted into the silenced state (Figure 3.7a,b). This *trans*-silencing was highly penetrant, specifically when transmitted via the female germline (Figure 3.7a), and was permanently maintained in lineages lacking PRG-1 (Figure 3.7b). We conclude that genes undergoing permanent silencing in *prg-1* mutant animals are paramutagenic, emphasizing the irreversible and permanent nature of RNAi silencing that occurs in the absence of PRG-1.

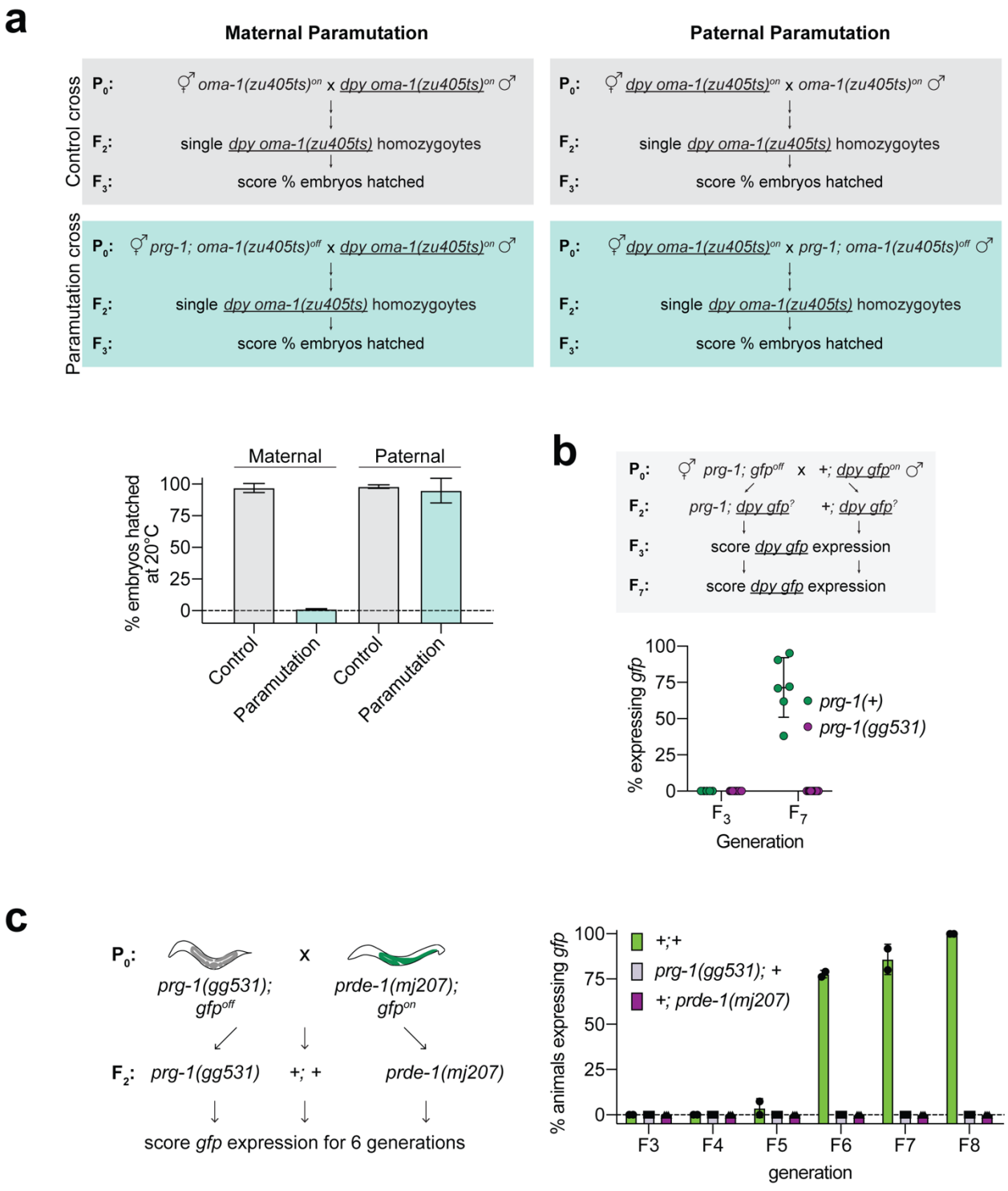


Figure 3.7. Permanently silenced alleles are paramutagenic. a, *prg-1(gg531)* hermaphrodites (maternal paramutation) or males (paternal paramutation) harboring a silent *oma-1(zu405ts)* allele (represented as *prg-1; oma-1(zu405ts)*^{off}) were crossed to animals harboring an expressed *oma-1* allele that was marked by a tightly-linked *dpy-20(e1282)* mutation (represented as *dpy oma-1(zu405ts)*^{on}). F_2 animals homozygous for *dpy oma-1(zu405ts)* were singled and allowed to

Figure 3.7 (Continued). lay an F_3 brood. % hatched embryos laid by F_3 animals was counted. *oma-1(zu405ts)*-expressing hermaphrodites and males, represented as *oma-1(zu405ts)*^{on}, were crossed to *dpy oma-1(zu405ts)*^{on} animals as a control. For each cross, four independent P_0 hermaphrodites were mated, 20 *dpy oma-1(zu405ts)* F_2 animals were singled (from 5 different F_1 animals), and % embryos hatched was counted for pools of 5 F_3 animals derived from each F_2 animal. % embryos hatched was averaged for all pools derived from the same P_0 and is shown here. Error bars represent s.d. of the mean. **b,** *prg-1(gg531)* animals were outcrossed with *gfp* expressing animals to remove the *oma-1(zu405ts)* allele, allowing this experiment to be performed at 20°C. Note: because the silenced *gfp* allele is paramutagenic in *prg-1(gg531)* animals, the *gfp* allele remained silent after outcross. Two hermaphrodites, represented as *prg-1; gfp*^{off}, were then mated with animals homozygous for the expressed *dpy-10(e128) gfp* allele, represented as *dpy gfp*^{on}. Lineages were established from homozygous *dpy gfp* F_2 animals, which were genotyped for *prg-1(gg531)* once they had laid a brood. Gravid adults were egg prepped every generation to maintain these lineages. *gfp* expression was scored in the F_3 and F_7 generations. At least 50 animals were counted for each strain. Each dot represents % animals expressing *gfp* in lineages derived from each F_2 *prg-1(+)* or *prg-1(tm872)* animal. **c,** *prg-1(gg531)* hermaphrodites were crossed to *gfp* expressing *prde-1(mj207)* males. F_2 animals were singled and genotyped for *prg-1(gg531)* and *prde-1(mj207).* Lineages were established from animals of the indicated genotypes and *gfp* expression was scored for the indicated number of generations. Each dot represents the number of lineages established for each genotype, each from a single F_2 homozygote of the indicated genotype. Error bars represent s.d. of the mean.

D. piRNAs limit RNAi inheritance

PRG-1 is a PIWI protein, which binds piRNAs in the *C. elegans* germline (Wang and Reinke 2008; Batista *et al.* 2008; Das *et al.* 2008). Therefore, the permanent silencing and paramutation we observe in animals lacking PRG-1 is likely to be due to loss of piRNA function in the germline. In support of this idea, one of the twenty mutant strains identified by our genetic screen harbored a nonsense mutation, *gg530,* in the *prde-1* gene, which encodes a nuclear-localized protein required for the biogenesis and/or stability of piRNA precursor transcripts (Figure 3.8a) (Weick *et al.* 2014). To further test the idea that PRDE-1, and, therefore, piRNAs limit maintenance of permanent silencing, we asked if animals harboring an independently isolated null allele of *prde-1(mj207)* also exhibited enhanced RNAi inheritance (Weick *et al.* 2014). We conducted *gfp* RNAi on *gfp* and *prde-1(mj207); gfp* animals and detected a complex and subtle, yet statistically significant, enhancement of *gfp* RNAi inheritance in *prde-1(mj207)* animals 11

generations after dsRNA exposure (Figure 3.8b). Further, some *prde-1(mj207)* animals, but no control animals, continued to silence *gfp* 27 generations after RNAi (Figure 3.8b). A related analysis that used paramutation to initiate gene silencing revealed a more dramatic and clearer role for PRDE-1 in the maintenance of permanent silencing. We crossed *prg-1(gg531)* animals harboring a permanently silenced *gfp* allele to *prde-1(mj207)* animals that expressed *gfp* and isolated *prg-1(+); prde-1(+)*, *prg-1(+); prde-1(mj207)* or *prg-1(gg531); prde-1(mj207)* progeny, which were all homozygous for *gfp* (Figure 3.7c)**.** We monitored GFP expression over generations in these lineages and, as expected based upon previous results (Figure 3.7b), we observed a loss of *gfp* silencing ≅6 generations after the cross in *prg-1(+); prde-1(+)* progeny, but no loss of *gfp* silencing in *prg-1(gg531); prde-1(+)* lineages (Figure 3.7c). *prde-1(mj207)* lineages behaved like *prg-1(gg531)* lineages in that they maintained *gfp* silencing for far more generations than wild-type (Figure 3.7c). We conclude that, like PRG-1, PRDE-1 limits the generational perdurance of RNAi inheritance, arguing that piRNAs prevent permanent gene silencing in the *C. elegans* germline.

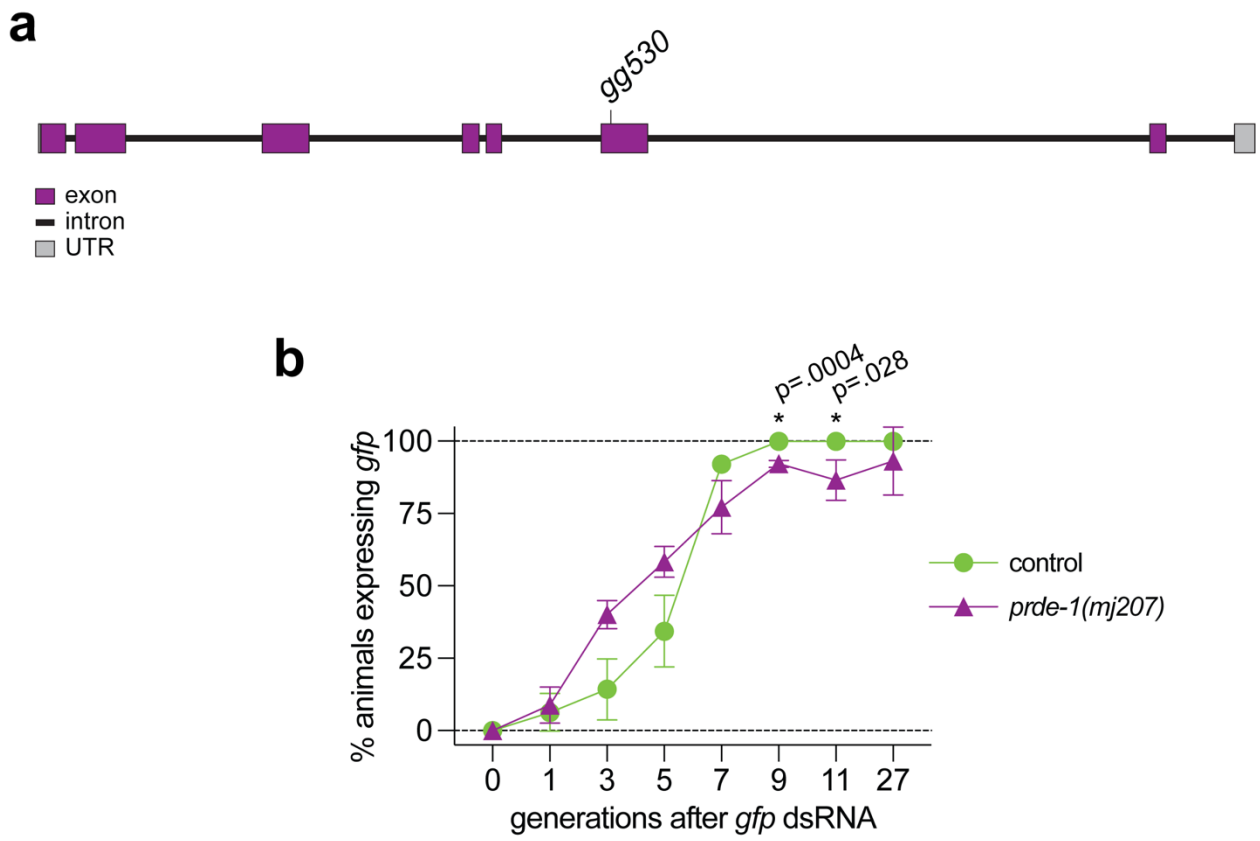


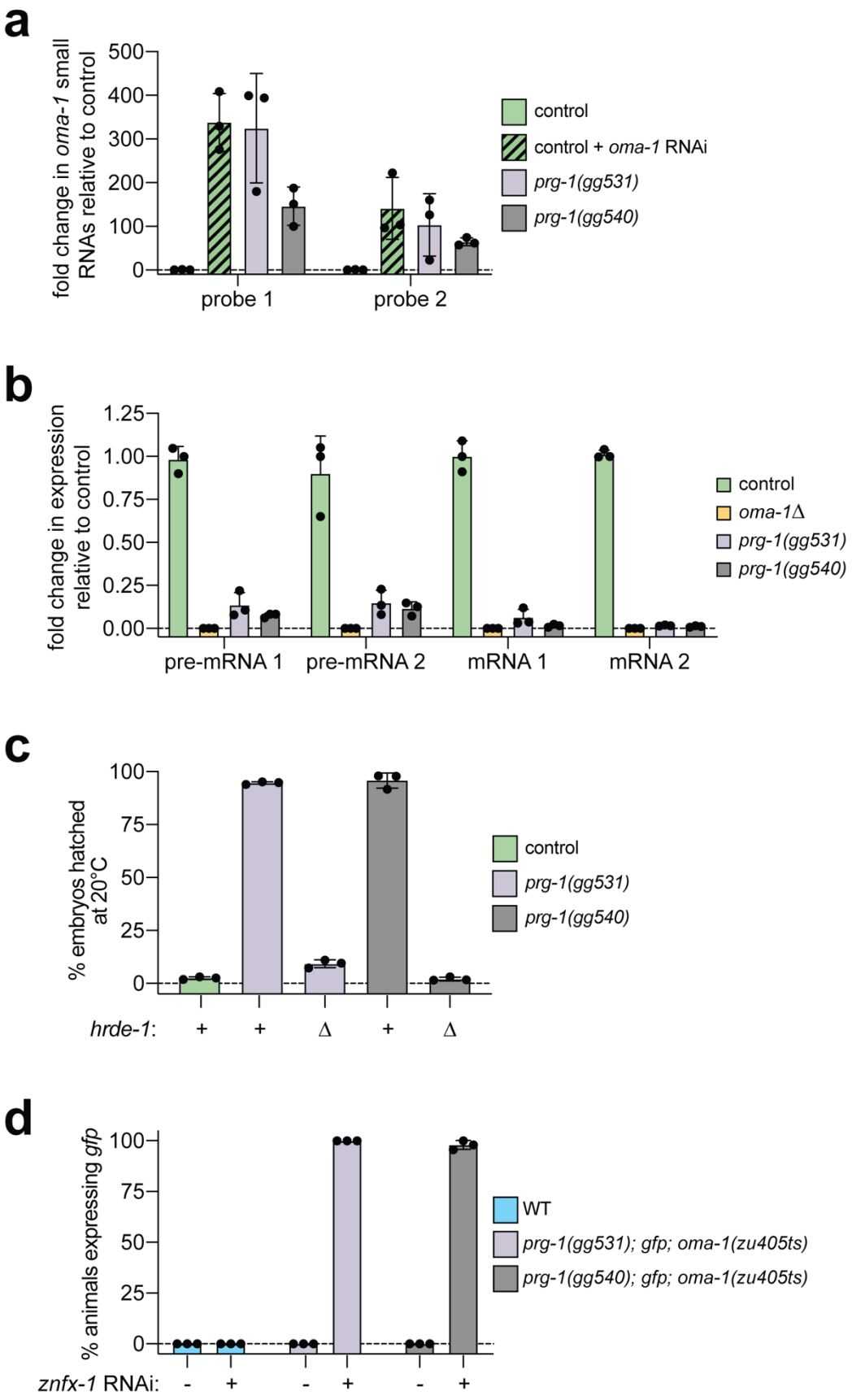
Figure 3.8. PRDE-1 may also prevent permanent RNAi inheritance. **a,** *gg530* mutants were identified in our Heri screen as showing enhanced RNAi inheritance of *oma-1* and *gfp* RNAi compared to unmutagenized animals. These mutants have the indicated nonsense mutations in the *prde-1* gene. **b,** *gfp* RNAi inheritance assay was performed on *gfp* expressing control or *prde-1(mj207)* animals. >50 animals were counted each generation for each genotype. Error bars represent s.d. of the mean of three independent biological replicates. Of note, *prde-1(mj207)* animals were tested for *gfp* RNAi inheritance alongside the *prg-1(tm872)* animals shown in Figure 3.2b. Thus, the control data for these two figures is the same.

E. Permanent silencing is driven by continuous siRNA production and continuous co-transcriptional gene silencing

The data in this section show that the permanent silencing we see in *prg-1(-)* animals is caused by permanent activation of silencing pathways that are normally induced, but finite, after RNAi in wild-type animals. First, 22G-siRNAs that are antisense to genes undergoing TEI are heritably expressed during dsRNA-triggered TEI (Ashe *et al.* 2012; Buckley *et al.* 2012). TaqMan-based siRNA-quantification showed that *oma-1* and *gfp* siRNAs were expressed hundreds of

generations after *oma-1* and *gfp* RNAi in *prg-1(gg531)* and *prg-1(gg540)* animals (Figure 3.9a and Figure 3.10a,b). Second, the *C. elegans* nuclear RNAi machinery is required for RNAi inheritance (Ashe *et al.* 2012; Buckley *et al.* 2012), likely because co-transcriptional gene silencing is a necessary component of gene silencing occurring during RNAi inheritance. qRT-PCR analyses measuring nascent pre-mRNAs showed that the *oma-1* and *gfp* loci remained co-transcriptionally silenced in *prg-1(gg531)* and *prg-1(gg540)* animals, hundreds of generations after these loci were targeted by RNAi (Figure 3.9b). Third, the genetic requirements for permanent silencing in *prg-1(gg531)* and *prg-1(gg540)* animals resembled the genetic requirements for RNAi inheritance in *prg-1(+)* animals. For instance, the germline-expressed and nuclear-localized Argonaute HRDE-1, which is required for RNAi inheritance in wild-type animals (Ashe *et al.* 2012; Buckley *et al.* 2012), was required for permanent silencing in *prg-1(-)* animals. CRISPR/Cas9 was used to create a 2,885bp deletion in the *hrde-1* gene (Figure 3.11) in both *prg-1(gg531)* and *prg-1(gg540)* mutants and this deletion restored *oma-1* (Figure 3.9c) and *gfp* (Figure 3.12) expression to wild-type levels. In addition, the conserved RNA helicase ZNFX-1 is also required for RNAi inheritance in *C. elegans* (Wan *et al.* 2018). RNAi-based depletion of ZNFX-1 was sufficient to restore *gfp* expression to permanently silenced *prg-1(-)* animals (Figure 3.9d). Finally, DEPS-1, which localizes to germ granules in the *C. elegans* germline and contributes to RNAi inheritance in wild-type animals (Wan *et al.* 2018), was also required for permanent silencing in *prg-1(-)* animals (Figure 3.12). Together, these data show that the normally finite mechanism underlying RNAi inheritance in the *C. elegans* germline becomes permanent in the absence of PRG-1 and piRNAs.

Figure 3.9. Permanent silencing is co-transcriptional and depends on known HRDE factors. **a,** Taqman-based qRT-PCR was used to quantify the expression of two different *oma-1* siRNAs (probe 1 and 2) in control (*oma-1(zu405ts); gfp*) animals +/- *oma-1* RNAi and in *prg-1(gg531)* and *prg-1(gg540)* animals. **b,** qRT-PCR was used to measure *oma-1* pre-mRNA and mRNA levels (2 primer pairs each) in control (e.g. *oma-1(zu405ts); gfp*) animals or in *prg-1(gg531)* and *prg(gg540)* animals. Animals harboring a deletion in the *oma-1* gene served as negative controls for this experiment. Results were normalized to the germline-expressed gene *nos-3.* **c,** CRISPR/Cas9 was used to create a 2,885bp deletion in the *hrde-1* gene (Figure 3.11) in *prg-1(gg531)* and *prg-1(gg540)* animals. Silencing of *oma-1(zu405ts)* was then quantified by measuring % embryos hatched (as described above) for control (*oma-1(zu405ts); gfp*) animals vs. *prg-1(gg531)* and *prg-1(gg540)* animals with or without the *hrde-1* deletion. **d,** WT, *prg-1(gg531)* and *prg-1(gg540)* animals were fed bacteria expressing empty vector control or dsRNA targeting the gene *znfx-1* for two generations and *gfp* expression was scored. Three biological replicates of this experiment were performed, with 50 animals counted per condition for each genotype. Each dot represents % animals expressing *gfp* for each replicate.


a
fold change in oma-1 small RNAs relative to control
500
400
300
200
100
0
probe 1
probe 2
control
control + oma-1 RNAi
prg-1(gg531)
prg-1(gg540)
b
fold change in expression relative to control
1.25
1.00
0.75
0.50
0.25
0.00
pre-mRNA 1
pre-mRNA 2
mRNA 1
mRNA 2
control
oma-1Δ
prg-1(gg531)
prg-1(gg540)
c
% embryos hatched at 20°C
100
75
50
25
0
hrde-1: + + Δ + Δ
control
prg-1(gg531)
prg-1(gg540)
d
% animals expressing gfp
100
75
50
25
0
znfx-1 RNAi: - + - + - +
WT
prg-1(gg531); gfp; oma-1(zu405ts)
prg-1(gg540); gfp; oma-1(zu405ts)

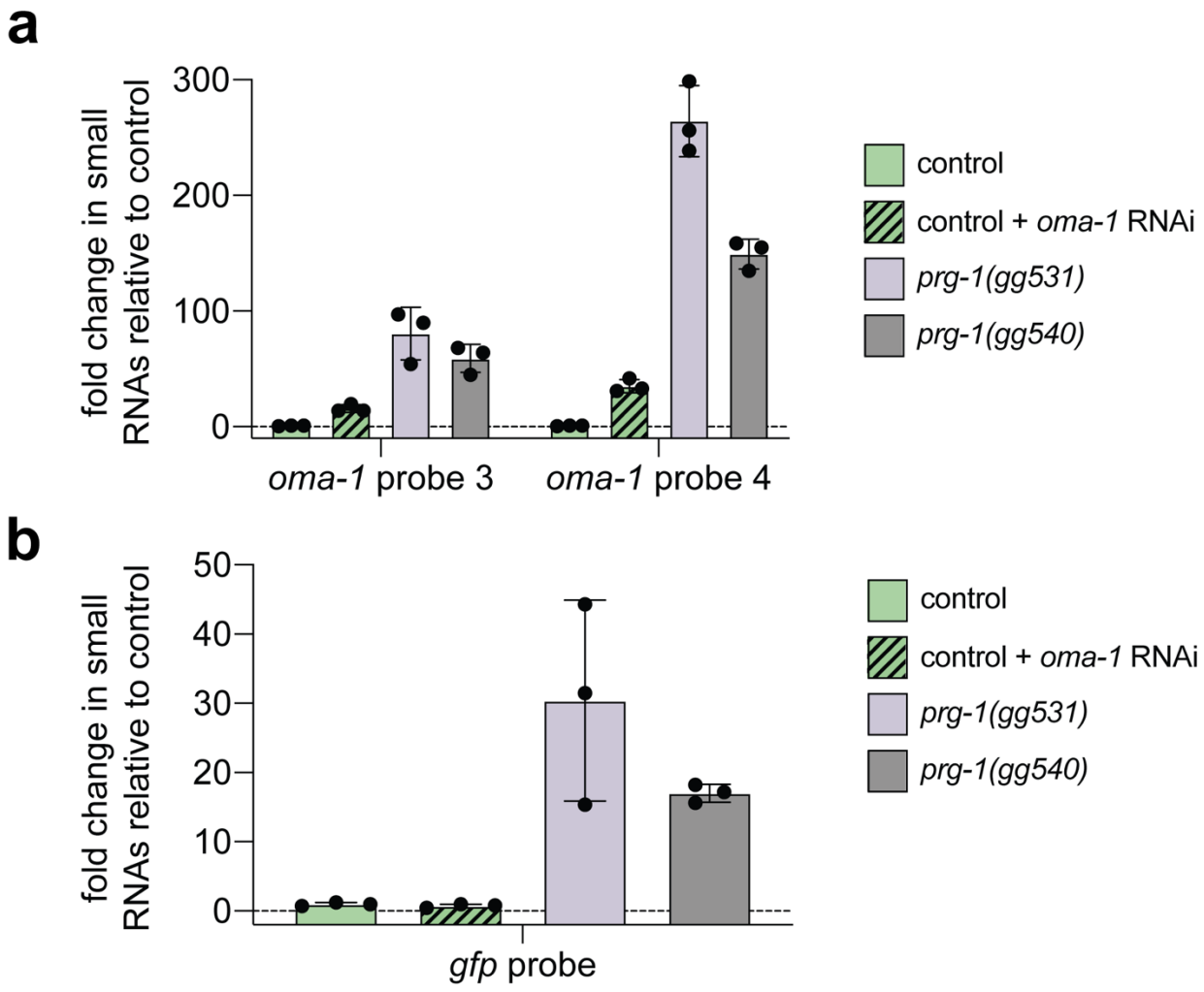


Figure 3.10. *prg-1(gg531)* and *prg-1(gg540)* animals show elevated *oma-1* and *gfp* small RNA levels. Taqman-based qRT-PCR was used to quantify the expression of two additional (see Figure 3.9a) *oma-1* siRNAs (probes 3 and 4) and a *gfp* siRNA in control (*oma-1(zu405ts); gfp*) animals +/- *oma-1* RNAi vs. *prg-1(gg531)* and *prg-1(gg540)* animals.

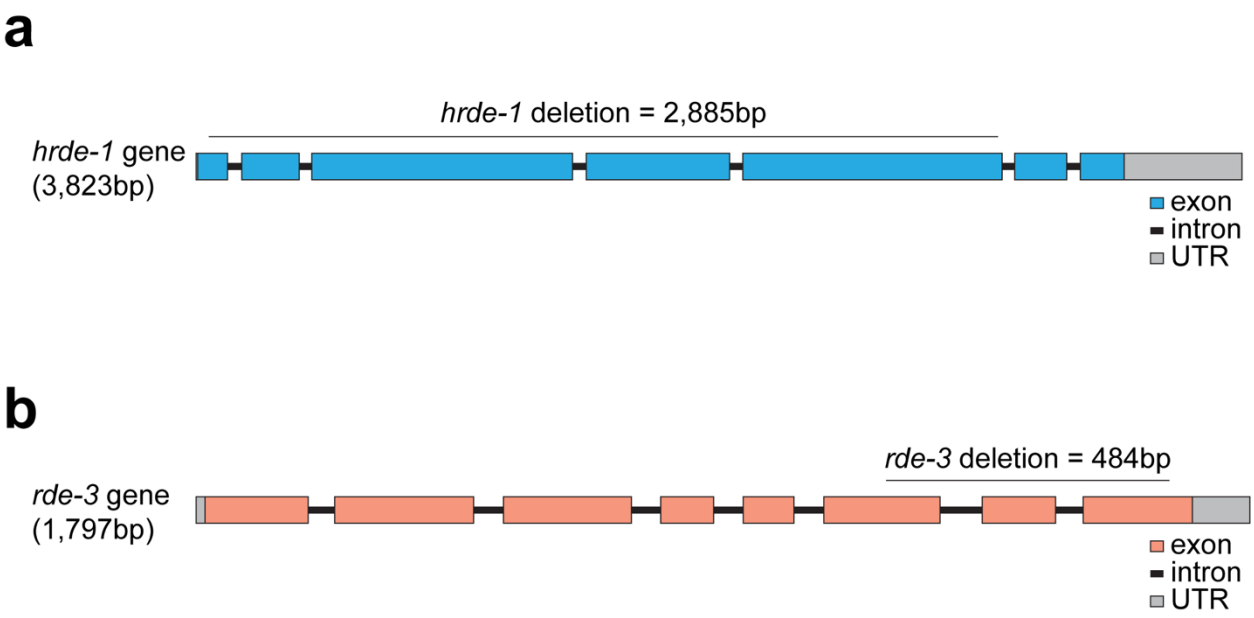
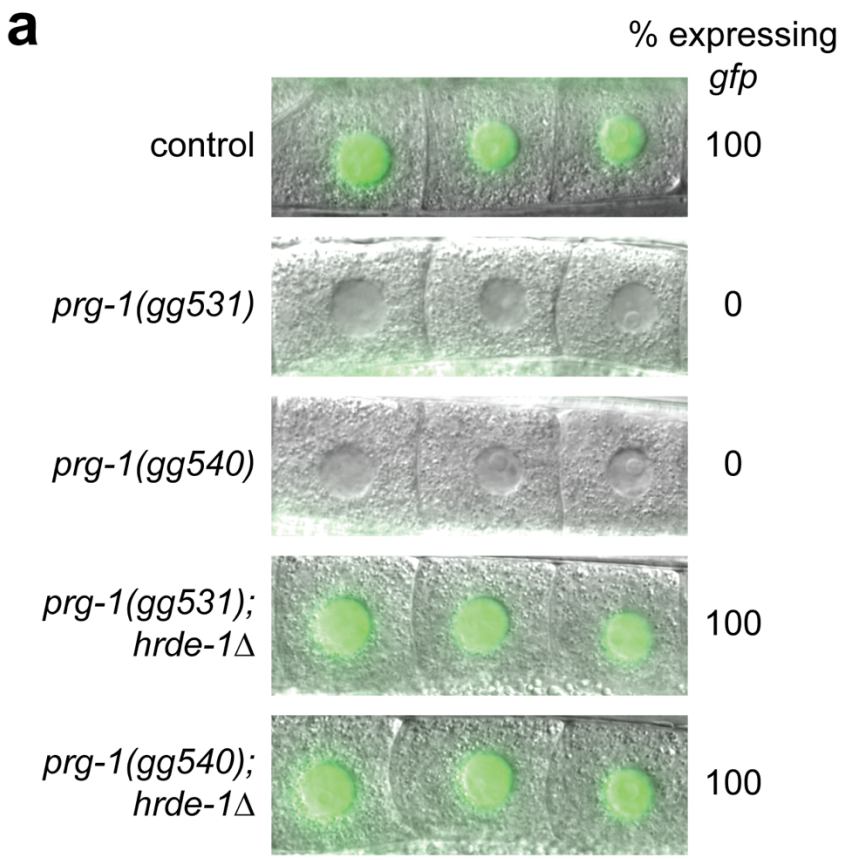
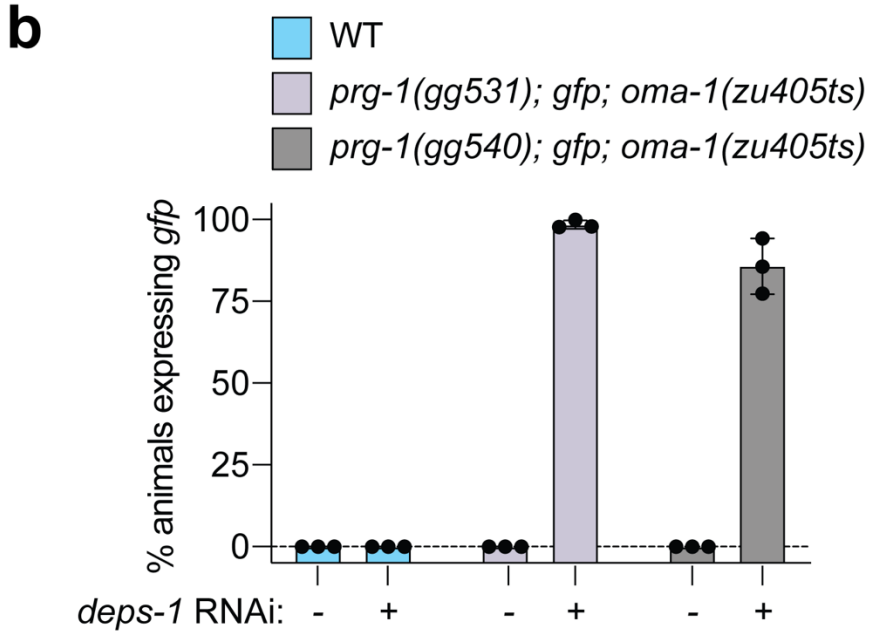


Figure 3.11. CRISPR/Cas9-induced deletions in *hrde-1* and *rde-3* genes. a, Schematic of the *hrde-1* gene, annotated with the 2,885 deletion created in *prg-1(gg531)* and *prg-1(gg540)* animals using CRISPR/Cas9. **b,** Schematic of the *rde-3* gene, annotated with the 484bp deletion created in *prg-1(gg531)* animals using CRISPR/Cas9. Gene features and coordinates were obtained from WormBase (Harris *et al.* 2020).

Figure 3.12. HRDE-1 and DEPS-1 are required for permanent RNAi inheritance in *prg-1(gg531)* and *prg-1(gg540)* animals. a, *prg-1(gg531)* and *prg-1(gg540)* animals with and without the *hrde-1* deletion shown in Figure 3.11 were scored for *gfp* expression, alongside control (*oma-1(zu405ts); gfp*) animals. >100 animals were scored for each genotype. **b,** Control (*oma-1(zu405ts); gfp*), *prg-1(gg531)* and *prg-1(gg540)* animals were fed bacteria expressing empty vector control or dsRNA targeting the gene *deps-1* for two generations and *gfp* expression was scored.

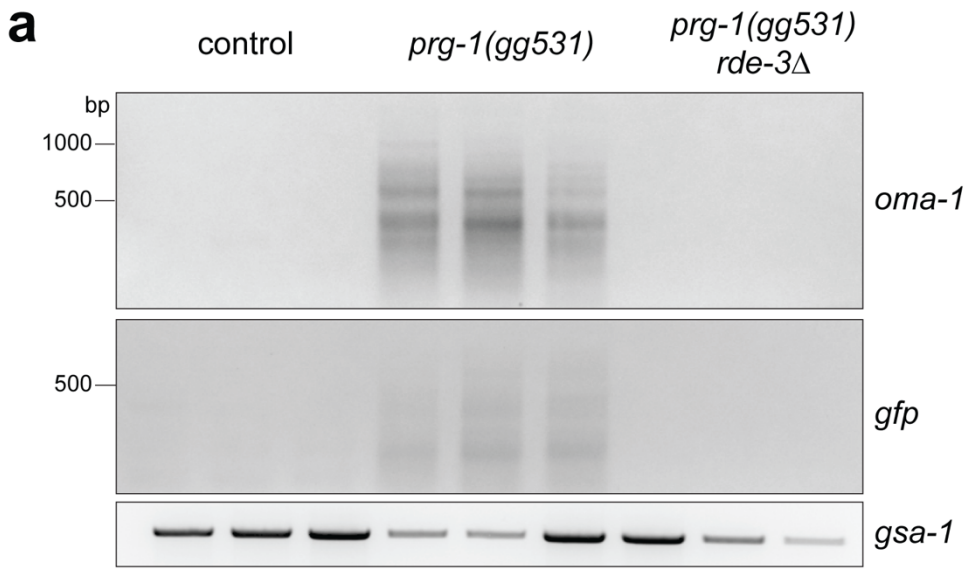

a
% expressing gfp
control
100
prg-1(gg531)
0
prg-1(gg540)
0
prg-1(gg531); hrde-1Δ
100
prg-1(gg540); hrde-1Δ
100

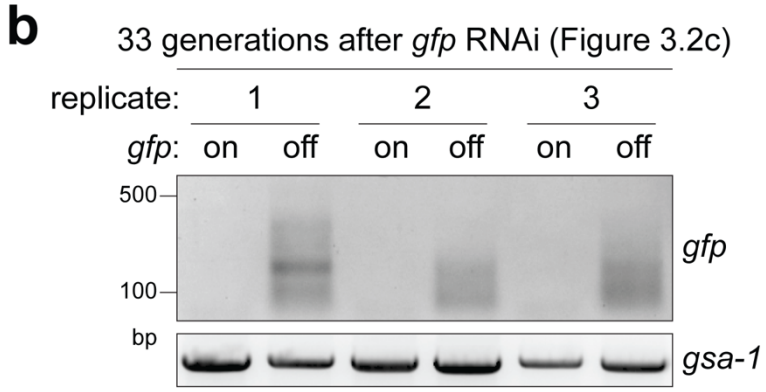

b
WT
prg-1(gg531); gfp; oma-1(zu405ts)
prg-1(gg540); gfp; oma-1(zu405ts)
% animals expressing gfp
100
75
50
25
0
deps-1 RNAi: - + - + - +

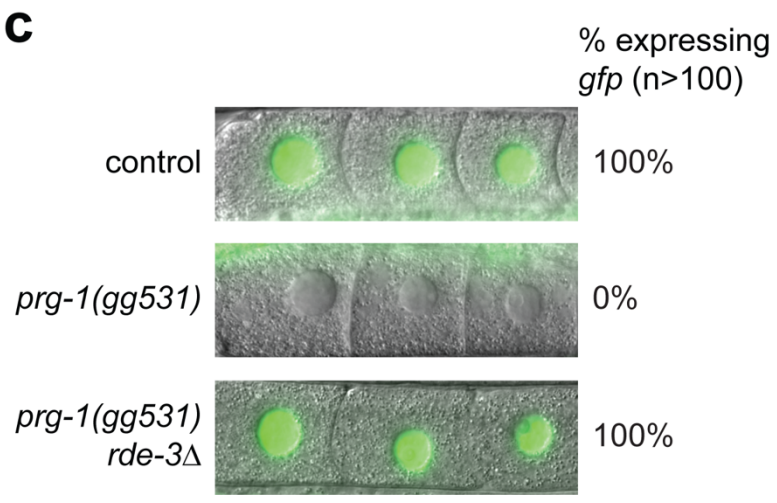
F. Perpetual pUG RNA/siRNA cycling causes permanent silencing

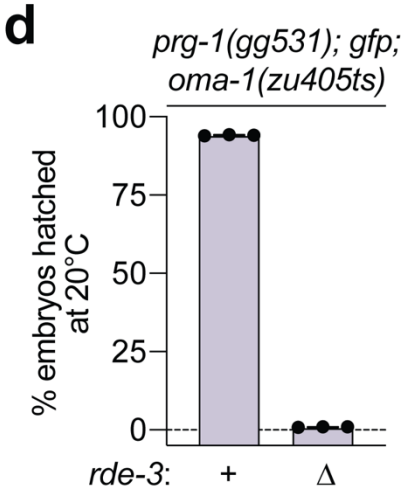
RDE-3 is a ribonucleotidyltransferase (rNT) required for RNAi (Chen *et al.* 2005; Phillips *et al.* 2012) and RNAi inheritance in *C. elegans* (Shukla *et al.* 2020). RDE-3 adds poly(UG) or pUG tails to mRNAs targeted for silencing by RNAi (Shukla *et al.* 2020). RdRPs, such as RRF-1, are recruited to pUG tails and use pUG RNAs as templates for antisense siRNA synthesis (Shukla *et al.* 2020). We recently showed that generationally repeated cycles of pUG RNA-mediated siRNA biogenesis coupled with siRNA-directed mRNA pUGylation drive RNAi inheritance in the *C. elegans* germline (Shukla *et al.* 2020). The following data show that the permanent silencing of *oma-1* and *gfp* we observe in *prg-1* mutants is due to perpetual activation of these pUG RNA/siRNA cycles. First, pUG RNAs are expressed concomitantly with gene silencing during RNAi inheritance (Shukla *et al.* 2020). We previously described a PCR-based assay to detect gene-specific pUG RNAs (pUG PCR). Using this assay, we detected *oma-1* and *gfp* pUG RNAs in both *prg-1(gg531)* and *prg-1(gg540)* animals, hundreds of generations after dsRNA exposure (Figure 3.13a). Non-templated pUG tails were added to *oma-1* and *gfp* mRNA fragments (Figure 3.14), which is similar to what was observed for pUG RNA production after RNAi in wild-type animals (Shukla *et al.* 2020). Second, pUG RNA expression correlated with permanent gene silencing. For instance, 3/3 *prg-1(-)* lineages that continued to silence *gfp* 33 generations after *gfp* RNAi expressed *gfp* pUG RNAs, while 3/3 lineages no longer silencing *gfp* did not (Figure 3.13b). Third, the *C. elegans* poly(UG) polymerase RDE-3 was required for permanent silencing. CRISPR/Cas9 was used to introduce a 484bp deletion in the *rde-3* gene in *prg-1(gg531)* animals, which were undergoing permanent *gfp* and *oma-1* silencing. This deletion abrogated *gfp* and *oma-1* pUG RNA expression (Figure 3.13a) and halted permanent *gfp* and *oma-1* silencing (Figure 3.13c,d). We conclude that permanent RNAi-initiated gene silencing in animals lacking a functional piRNA system is driven by perpetual cycling of the pUG RNA/siRNA pathway.

Figure 3.13. Permanent silencing of *oma-1* and *gfp* depends on permanent pUG RNA/siRNA cycles. a, *oma-1* and *gfp* pUG PCR was performed on total RNA isolated from control (*oma-1(zu405ts); gfp*) vs. *prg-1(gg531)* animals with or without a 484bp deletion in the *rde-3* gene created using CRISPR/Cas9 (Figure 3.11). *gsa-1*, which has an 18nt long genomically encoded pUG repeat in its 3'UTR, is a loading control. Three biological replicates of this experiment are shown. **b,** *gfp* pUG PCR was performed on total RNA isolated from *gfp* expressing and nonexpressing lineages that were established in Figure 3.2c. Animals were collected 33 generations after *gfp* dsRNA treatment. **c,** Fluorescence micrographs showing *gfp* expression in the oocytes of control (e.g. *oma-1(zu405ts); gfp*) animals vs. *gg531* with or without an *rde-3* deletion. >100 animals of each genotype were counted. **d**, *oma-1(zu405ts)* silencing was measured by quantifying the % embryos hatched at 20°C for *prg-1(gg531)* animals with or without an *rde-3* deletion.


a
control
prg-1(gg531)
prg-1(gg531)
rde-3Δ
bp
1000
500
oma-1
500
gfp
gsa-1


b
33 generations after gfp RNAi (Figure 3.2c)
replicate: 1 2 3
gfp: on off on off on off
500
100
bp
gfp
gsa-1


c
% expressing
gfp (n>100)
control
100%
prg-1(gg531)
0%
prg-1(gg531)
rde-3Δ
100%


d
prg-1(gg531); gfp;
oma-1(zu405ts)
% embryos hatched
at 20°C
100
75
50
25
0
rde-3: + Δ

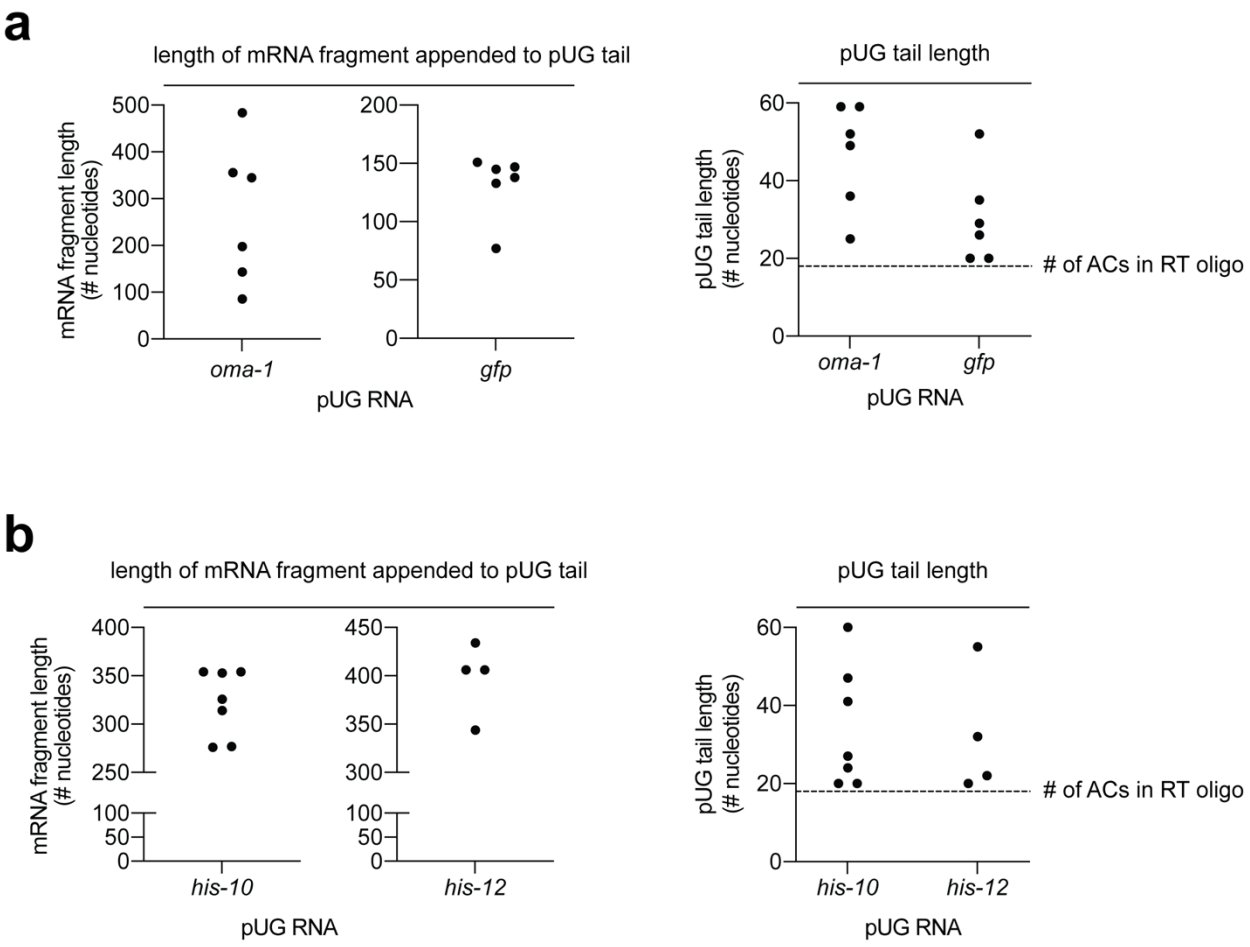


Figure 3.14. *oma-1, gfp* and histone pUG RNAs in *prg-1(gg531)* animals. Sanger sequencing (see Methods) was performed on **a,** *oma-1* and *gfp* pUG RNAs; **b,** histone pUG RNAs (Figures 6a and 7b, respectively) detected in *prg-1(gg531)* animals (see Methods). As shown previously (Shukla *et al.* 2020), these pUG RNAs corresponded to spliced mRNA fragments appended to long non-templated pUG tails. Shown here is the length of the spliced mRNA fragment in each pUG RNA that was sequenced, as well as the range of pUG tail lengths observed. Only pUG RNAs harboring >18nt long pUG tails (i.e. the length of the RT oligo used to detect them) are shown here.

G. piRNAs prevent pUG RNA-based permanent silencing of germline-expressed genes

Recent studies find that one function of the *C. elegans* piRNAs is to coordinate endogenous 22G-siRNA systems in the germline and that, in the absence of this coordination, some essential genes, most prominently the replication-dependent histones genes, undergo 22G-

siRNA-dependent aberrant gene silencing (de Albuquerque *et al.* 2015; Phillips *et al.* 2015; Barucci *et al.* 2020; Reed *et al.* 2020). Above, we showed that genes undergoing permanent RNAi silencing in *prg-1* mutants are silenced by perpetual pUG/siRNA cycling. We wondered if previously documented cases of aberrant gene silencing observed in *prg-1* mutants (Barucci *et al.* 2020; Reed *et al.* 2020) might also be explained by gain or loss of endogenous pUG/siRNA cycles. *his-10/14/26, his-11/15/44* and *his-12/16/43* are three sets of three nearly identical replication-dependent histones genes subjected to aberrant silencing in *prg-1* mutants (Barucci *et al.* 2020; Reed *et al.* 2020). We first confirmed that the mRNAs encoded by these genes were aberrantly silenced in our *prg-1(gg531)* mutants (Figure 3.15a). We next asked if the aberrant silencing of histone mRNAs in *prg-1(gg531)* animals was associated with aberrant histone pUG RNA production. Indeed, we detected *his-10/14/26, his-11/15/44* and *his-12/16/43* pUG RNAs in *prg-1(gg531)* animals but not in control animals (Figure 3.15b). Like pUG RNAs produced in response to dsRNA, histone pUG RNAs consisted of 5' fragments of histone mRNAs modified with non-templated pUG tails (Figure 3.14). Deletion of *rde-3* restored histone gene expression to near wild-type levels in *prg-1(gg531)* animals (Figure 3.15a) and abolished aberrant histone pUG RNA expression (Figure 3.15b). RNA-Seq studies have identified a number of genes, in addition to the replication-dependent histones, whose expression is downregulated in *prg-1* mutants (Barucci *et al.* 2020; Reed *et al.* 2020). We wondered if the aberrant silencing of these loci might also be explained by aberrant mRNA pUGylation. Indeed, we found that the predicted protein-coding gene *r03d7.2*, which is not normally pUGylated in *prg-1(+)* animals and becomes downregulated in *prg-1* mutants (Figure 3.15c) (Barucci *et al.* 2020; Reed *et al.* 2020), was pUGylated in an RDE-3-dependent manner in *prg-1(gg531)* animals (Figure 3.15d). Interestingly, during the course of analyzing predicted protein-coding genes for aberrant pUGylation, we identified as least one gene, *ztf-28*, which was previously reported to be downregulated in *prg-1* animals, but was pUGylated in *prg-1(gg540)* animals, but not *prg-1(gg531)* animals (Figure 3.16a). Aberrant pUGylation correlated with silencing of *ztf-28* in only *prg-1(gg540)* animals

(Figure 3.16b), indicating that while some genes, like the replication-dependent histone genes, are highly predisposed to become silenced in the absence of piRNAs, other genes, like *ztf-28,* enter the permanently silenced state more stochastically. We conclude that the decreased expression of at least some mRNAs in *prg-1* mutants, including the replicative histone mRNAs, is mediated by aberrant and perpetual pUG RNA/siRNA cycling.

Finally, we wondered if misrouting of some mRNAs into the pUG RNA/siRNA silencing loop was an indicator of global disorganization of the mRNA targets of pUGylation. Indeed, as previously reported (Barucci *et al.* 2020; Reed *et al.* 2020), we confirmed loss of *bath-13* silencing in the absence of PRG-1 and piRNAs (Figure 3.15e) and found that *bath-13* mRNAs, which are normally pUGylated in *prg-1(+)* animals, were not pUGylated in *prg-1(gg531)* animals (Figure 3.15f). Taken together, we conclude that one function of PRG-1 and piRNAs is to coordinate gene expression programs in the *C. elegans* germline, specifically by focusing the activity of the poly(UG) polymerase RDE-3 activity to the appropriate mRNAs.

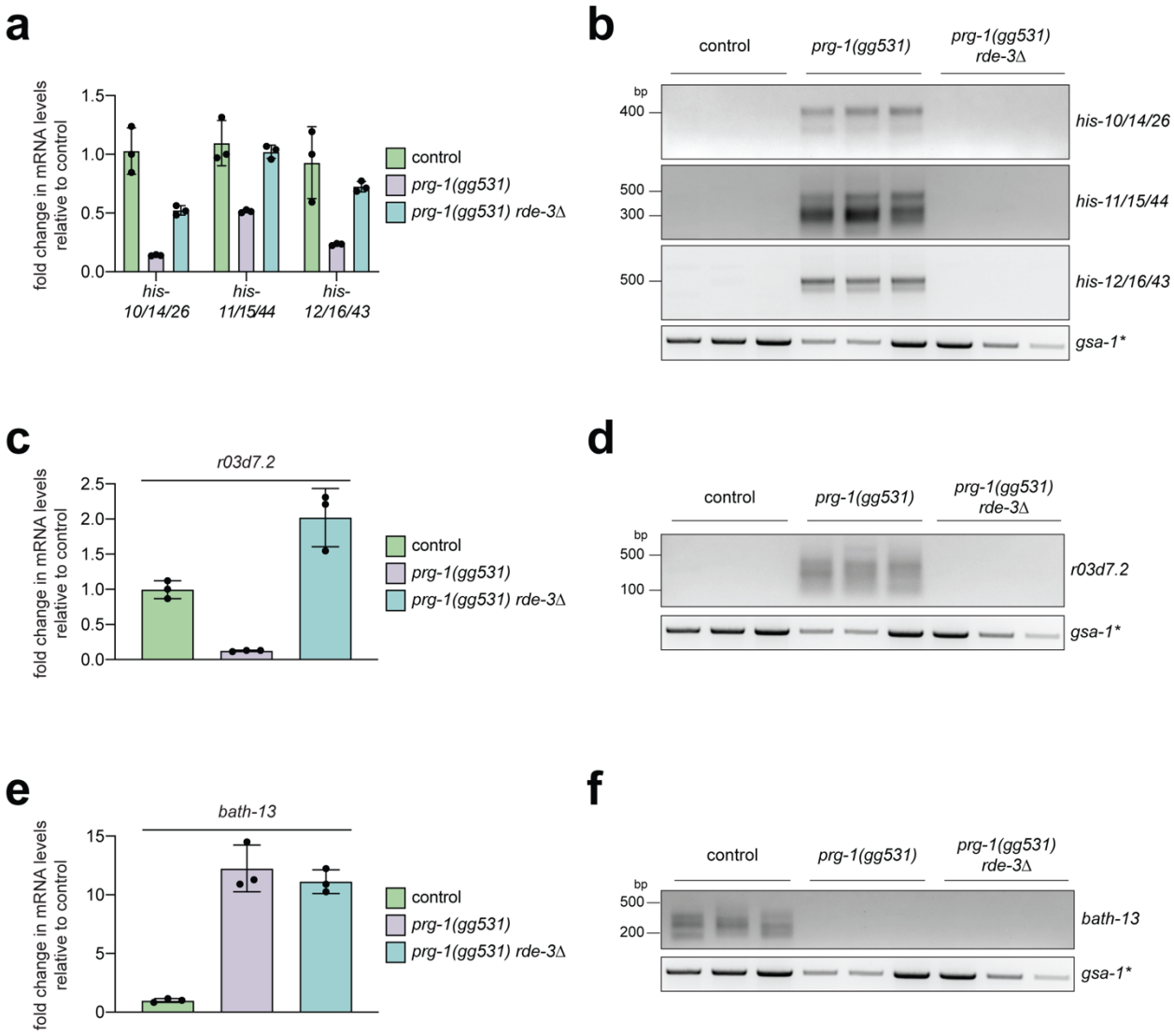


Figure 3.15. Aberrant pUG RNA/siRNA cycles drive silencing of germline-expressed mRNAs. For all panels, three biological replicates of the following strains were used: control (*oma-1(zu405ts); gfp*) animals or *prg-1(gg531)* animals with or without an *rde-3* deletion (Figure 3.11). **a, c, e,** qRT-PCR was used to quantify mRNA expression of the indicated genes. Each dot represents an independent biological replicate. Error bars represent s.d. of the mean. Results were normalized to the germline-expressed gene *nos-3.* **b, d, f,** Gene-specific pUG PCR assays were used to detect pUG RNAs corresponding to the indicated genes. *gsa-1*, which has an 18nt long genomically encoded pUG repeat in its 3'UTR, is a loading control. The same RNA samples were used for all three panels, so the *gsa-1* control is the same.

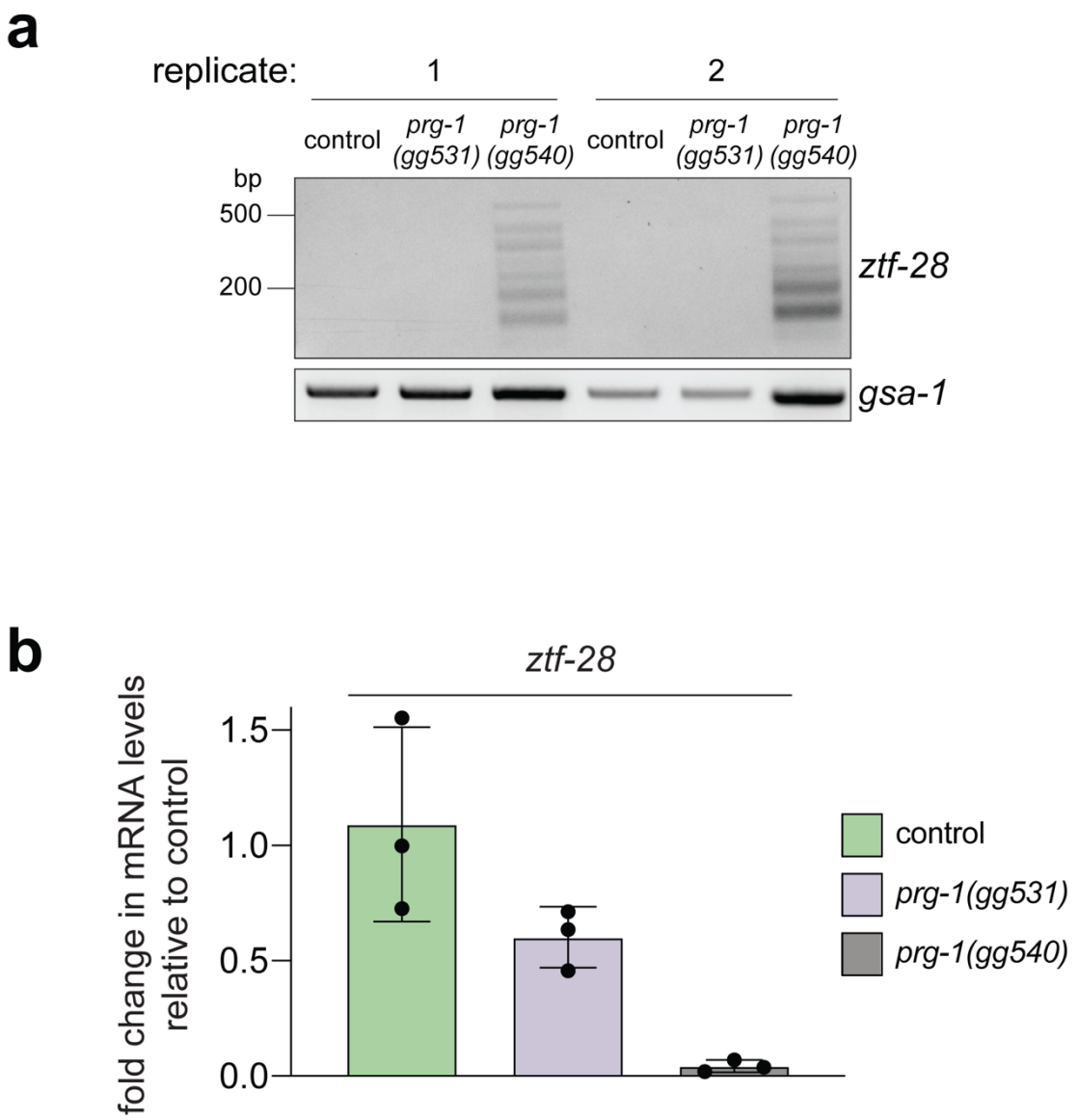


Figure 3.16. *ztf-28* is pUGylated and downregulated in only *prg-1(gg540)* animals. a, *ztf-28* pUG PCR assay was used to detect *ztf-28* pUG RNAs in control (*oma-1(zu405ts); gfp*) animals or in *prg-1(gg531)* and *prg-1(gg540)* animals. Two biological replicates of this experiment are shown. **b,** qRT-PCR was used to quantify *ztf-28* mRNA expression in three biological replicates of control (*oma-1(zu405ts); gfp*) animals or in *prg-1(gg531)* and *prg-1(gg540)* animals. Each dot represents an independent biological replicate. Error bars represent s.d. of the mean. Results were normalized to the germline-expressed gene *nos-3.*

IV. Discussion

Here we show that in the absence of PRG-1 and piRNAs, dsRNA-triggered TEI, which is normally confined to a finite number of generations, becomes essentially permanent. This

permanent silencing relies on nuclear RNAi machinery, germ granules, and RDE-3-mediated poly(UG) tailing of mRNAs targeted by dsRNA. We also find that a subset of germline-expressed mRNAs, including the replication-dependent histone RNAs, become novel substrates for RDE-3 in the absence of piRNAs. Together, our results show that piRNAs prevent germline-expressed mRNAs from entering self-perpetuating pUG RNA/siRNA silencing loops, thus protecting these mRNAs from undergoing permanent transgenerational silencing.

Previous reports have shown that one function of piRNAs is to regulate gene expression in *C. elegans* by coordinating germline small RNA pathways (de Albuquerque *et al.* 2015; Phillips *et al.* 2015; Barucci *et al.* 2020; Reed *et al.* 2020). Our data show that piRNAs do so, at least in part, by regulating RDE-3-dependent pUG/siRNA cycling. Whereas the trigger for permanent RNAi inheritance is exogenous dsRNA, the molecular triggers initiating aberrant silencing of germline-expressed transcripts, such as the histone mRNAs, are not yet known. It has been suggested that lack of a poly(A) tail, which is characteristic of eukaryotic histone mRNAs, might predispose histone mRNAs to aberrant silencing in piRNA mutants (Reed *et al.* 2020). However, we find that other mRNAs, such as *r03d7.2*, which are thought to be polyadenylated, are also subjected to irreversible pUGylation (Figure 3.15d). Thus, the molecular signals initiating aberrant silencing in the absence of piRNAs are likely to be more complex than previously thought. Adding to this complexity is our finding that there appear to be gene-specific differences in intrinsic susceptibility to runaway silencing. For instance, the *ztf-28* mRNA, which was previously reported to be aberrantly silenced in a piRNA mutant, was subjected to aberrant pUG RNA/siRNA-based silencing in just one of our two *prg-1(-)* strains. Similarly, 27 generations after RNAi, some *prg-1(-)* lineages showed permanent silencing of *gfp,* while others did not (Figure 3.2b). Together, our results hint at a molecular threshold that must be reached for an mRNA to enter the pUG RNA/siRNA silencing loop and to become irreversibly silenced. Some mRNAs, like the histone mRNAs, are more highly predisposed to reaching this threshold and, therefore, runaway histone silencing is highly penetrant in a population of *prg-1* mutants. Other mRNAs, like *ztf-28*, are less

likely to reach this threshold and, therefore, runaway silencing of these mRNAs is less penetrant. However, while genes may different in their propensity to enter pUG RNA/siRNA cycling, because pUG RNA/siRNA cycling is feed-forward, once the threshold is reached, gene silencing is complete and irreversible. It is not yet known if all mRNAs are susceptible to permanent silencing in the absence of PRG-1 and piRNAs or if sompUG RNA/siRNA silencing loop.

While initiating signals for aberrant silencing are not yet known, the mechanism by which silencing is maintained is via perpetual pUG/siRNA cycling (Figures 3.13 and 3.15). Our genetic analyses showed that it is maintenance, not initiation, of pUG/siRNA cycling that is inhibited by piRNAs (Figure 3.4). Therefore, the function of piRNAs is likely rescue of genes that have mistakenly entered aberrant pUG/siRNA cycling. Here we consider two models to explain how piRNAs might rescue genes from irreversible silencing. The piRNA complex, consisting of PRG-1 and its bound piRNAs, interacts with thousands of mRNAs in the *C. elegans* germline and most of these mRNAs are not silenced by this piRNA targeting (Shen *et al.* 2018). Many of these interactions likely occur in P granules, biomolecular condensates found in the *C. elegans* germline, to which PRG-1 localizes (Wang and Reinke 2008; Batista *et al.* 2008). The poly(UG) polymerase RDE-3 (Phillips *et al.* 2012), as well as pUG RNAs (Shukla *et al.* 2020), localize to a distinct germline biomolecular condensate, the *Mutator* focus. Therefore, it is possible that interactions with piRNA complexes sequester mRNAs in P granules, preventing them from interacting with RDE-3 in *Mutator* foci, such that in the absence of piRNAs, mRNAs might become mislocalized and come into contact with RDE-3. mRNAs targeted by dsRNA are normally pUGylated by RDE-3, but levels of these pUG RNAs decline over generations, which is reflective of the transient nature of RNAi inheritance in wild-type animals (Shukla *et al.* 2020). The loss of pUG RNA expression over generations during RNAi inheritance in wild-type animals could result from a tug-of-war between the piRNA complex pulling mRNAs into P granules and the RNAi machinery pulling them into *Mutator* foci. Interestingly, one function of piRNAs in *Drosophila* is to promote germline specification by binding to and, thereby, entrapping maternal mRNAs in the

germ plasm (Vourekas *et al.* 2016). Thus, promoting gene expression via regulated mRNA localization may be a conserved function of the animal piRNAs. Alternatively, a "competition" model may explain how piRNAs limit aberrant and irreversible gene silencing. Current models posit that *C. elegans* small RNA pathways compete with each other for limiting downstream silencing factors. Indeed, a number of studies show that piRNAs promote gene silencing by utilizing the same downstream silencing machinery (RdRPs and WAGOs) utilized to silence mRNAs targeted by exogenous RNAi (Smardon *et al.* 2000; Sijen *et al.* 2001; Yigit *et al.* 2006; Ashe *et al.* 2012). Therefore, in the absence of piRNAs, and the 22G-siRNAs normally produced from piRNA-targeted mRNAs, RdRPs and WAGOs might become more available for maintaining aberrant pUG/siRNA cycles in the germline. Related competition models have been proposed to explain why mutations in the ERI/DICER complex enhance exo-RNAi (Kennedy *et al.* 2004; Duchaine *et al.* 2006; Pavelec *et al.* 2009), why endo-RNAi mutants show enhanced miRNA silencing (Zhuang and Hunter 2012) and why animals lacking the MET-2 methyltransferase show prolonged RNAi inheritance (Lev *et al.* 2017).

The germline pUG/siRNA gene silencing pathway is feed-forward and, therefore, potentially dangerous. If, for example, any germline mRNA was to mistakenly enter the pathway, it might never exit. Therefore, dedicated systems, such as PRG-1 and piRNAs, exist to limit and/or regulate pUG/siRNA cycling. *prg-1* mutants exhibit fertility defects, including reduced brood size, that are exacerbated at elevated temperatures (Wang and Reinke 2008; Batista *et al.* 2008; Das *et al.* 2008). Indeed, the aberrant silencing of essential genes that occurs in the absence of piRNAs has been linked to sterility in *C. elegans* (de Albuquerque *et al.* 2015; Phillips *et al.* 2015). Additionally, genetic perturbations that restored expression of aberrantly silenced genes, such as the replication-dependent histones, in *prg-1* mutants correlated with partial rescue of the brood size defects in these animals (Barucci *et al.* 2020). These results, combined with our data showing that at least some of the aberrant silencing in *prg-1* mutants is driven by pUG/siRNA cycling, suggests that some of the germline defects associated with *prg-1* are due to the disorganization

of RDE-3-dependent pUGylation. We conclude that inhibition of aberrant pUG/siRNA cycling is one mechanism by which piRNAs coordinate germline gene expression programs and, therefore, promote germ cell function.

V. Methods

A. Genetics

C. elegans culture and genetics were performed as described previously (Brenner 1974). Most strains were maintained at 20°C on Nematode Growth Medium (NGM) plates and fed OP50 *E. coli* bacteria, unless otherwise noted. Of note, animals expressing *zu405,* a gain-of-function, temperature-sensitive mutation in the gene *oma-1* were maintained at 15°C. A list of strains used in this study are provided in Table 1.

Table 3.1. *C. elegans* strains used in this study.

Strain Name	Genotype	Source
N2	wild-type	CGC
SX461	*mjIs31[pie-1::gfp::h2b][pie-1::gfp::h2b] II*	Miska lab
YY565	*mjIs31[pie-1::gfp::h2b][pie-1::gfp::h2b] II; oma-1(zu405) IV*	
YY1165	*prg-1(tm872) I; mjIs31[pie-1::gfp::h2b][pie-1::gfp::h2b] II*	Was nearing 100% sterility, so was outcrossed twice with SX461 to generate YY1810.
YY1810	*prg-1(tm872) I; mjIs31[pie-1::gfp::h2b][pie-1::gfp::h2b] II*	
YY1356	*prg-2(tm1094); mjIs31[pie-1::gfp::h2b]*	
YY1415	*prg-1(tm872) I; mjIs31[pie-1::gfp::h2b] II; prg-2(tm1094) IV*	
YY1166	*oma-1(zu405) IV*	Outcross of strain TX20 from the CGC.

YY1099	*dpy-20(e1282) oma-1(zu405) IV*	
tm1396	*oma-1(tm1396) IV*	Mitani Lab
YY944	*prg-1(gg531) I; mjIs31[pie-1::gfp::h2b] II; oma-1(zu405) IV*	EMS screen - *pie-1::gfp::h2b* and *oma-1(zu405)* are both "permanently" silenced in this strain
YY963	*prg-1(gg540) I; mjIs31[pie-1::gfp::h2b] II; oma-1(zu405) IV*	EMS screen - *pie-1::gfp::h2b* and *oma-1(zu405)* are both "permanently" silenced in this strain
YY1999	*prg-1(gg531) rde-3(gg798) I; mjIs31[pie-1::gfp::h2b] II; oma-1(zu405) IV*	CRISPR on YY944; note that *pie-1::g2b::h2b* and *oma-1(zu405)* expression was restored in these strains after *rde-3* deletion
YY2001	*prg-1(gg531) I; mjIs31[pie-1::gfp::h2b] II; hrde-1(gg800) III; oma-1(zu405) IV*	CRISPR on YY944; note that *pie-1::g2b::h2b* and *oma-1(zu405)* expression was restored in these strains after *hrde-1* deletion
YY2005	*prg-1(gg540) I; mjIs31[pie-1::gfp::h2b] II; hrde-1(gg800) III; oma-1(zu405) IV*	CRISPR on YY963; note that *pie-1::g2b::h2b* and *oma-1(zu405)* expression was restored in these strains after *hrde-1* deletion
YY1451	*prg-1(gg630) I; mjIs31[pie-1::gfp::h2b] II*	Mutated DDH catalytic triad to DAH in N2 and then crossed to SX461
YY1256	*prg-1(gg531); mjIs31[pie-1::gfp::h2b] II; oma-1(WT) IV*	Outcrossed silenced *oma-1(zu405)* allele from YY944 using SX461 males. Due the paramutagenic nature of the silenced *oma-1* allele, the WT *oma-1* allele is silenced in this strain. The *pie-1::gfp::h2b* allele is also silenced in this strain.
YY1095	*dpy-10(e128) mjIs31[pie-1::gfp::h2b] II*	
YY1254	*prde-1(mj207) I; mjIs31[pie-1::gfp::h2b] II*	Cross between strains SX2499 and SX461 from the Miska lab

B. Heri Screen

The forward genetic screen described in this study was previously reported in (Perales *et al.* 2018).

C. RNAi and RNAi inheritance assays

To perform RNAi experiments, embryos were isolated via hypochlorite treatment (egg prep) of gravid adult hermaphrodites and dropped onto RNAi plates (standard NGM plates with

1mM Isopropyl β-D-1-thiogalactopyranoside and 25ug/ml carbenicillin) seeded with HT115 *E. coli* bacteria expressing either L4440 (Addgene, #1654) empty vector control or L4440 carrying inserts to trigger the production of dsRNA (RNAi) against a gene of interest. The *dpy-11* RNAi clone and the *oma-1* RNAi clone used, unless noted below, came from the *C. elegans* RNAi collection (Ahringer lab). The *gfp* RNAi clone was obtained from the Fire lab. The *znfx-1* RNAi clone was a custom clone made for this study.

For *gfp* RNAi inheritance assays, embryos were dropped onto RNAi plates seeded with bacteria expressing either the L4440 empty vector control or *gfp* dsRNA. Once animals were gravid adults, they were washed off of plates using M9 + Triton X-100 buffer. 50-150 animals were placed onto a microscope slide and *gfp* expression was scored using the Plan-Apochromat 20 × /0.8 M27 objective on an Axio Observer.Z1 fluorescent microscope (Zeiss). Images were taken with the Plan-Apochromat 63 × /1.4 Oil DIC M27 objective. All image processing was done using Fiji (Schindelin *et al.* 2012). The remaining animals were egg prepped and embryos were dropped onto NGM plates seeded with OP50 *E. coli* bacteria. This process was repeated every generation for the indicated number of generations.

D. Quantification of oma-1(zu405ts) *silencing*

To measure *oma-1(zu405ts)* silencing, embryos were isolated via hypochlorite treatment of gravid adult hermaphrodites and dropped onto plates that were stored at 20°C. After 2-3 days, 6 larval stage 4 (L4) or adult animals were singled for each strain/genotype and allowed to lay embryos overnight at 20°C. Adults were removed from plates on the next day and the total number of embryos laid was counted. On the following day, the total number of embryos that hatched was counted. % embryos hatched was then calculated as the (# of hatched embryos / # embryos laid) x 100.

E. Paramutation crosses

Expressed *oma-1(zu405ts)* and *gfp* alleles were marked with tightly-linked *dpy-20(e1282)* and *dpy-10(e128)*, respectively, mutations to allow for differentiation of expressed/naive alleles from silent *oma-1(zu405ts)* and *gfp* alleles in *prg-1(gg531)* and *prg-1(gg540)* mutants. More specifically, *oma-1(zu405ts)* and *gfp* animals were crossed with *dpy-20(e1282)* and *dpy-10(e128)* animals, respectively, and 200-300 F_2s were singled and genotyped to identify rare double mutants resulting from a crossover. Paramutation crosses for *oma-1(zu405ts)* were performed at 15°C using the *prg-1(gg540)* mutant. Four hermaphrodites were mated per cross (see Figure 3.7a) and 5 F_1 heterozygotes were singled for each mated hermaphrodite. Four F_2s animals homozygous for the naive *dpy-20(e1282) oma-1(zu405ts)* allele were singled for each F_1 animal and allowed to lay a brood. Four pools of 5 F_3 animals for each F_2 were tested for *oma-1(zu405ts)* silencing based on the embryonic arrest assay described above. For the *gfp* paramutation test in Figure 3.7b, *prg-1(gg531)* mutants were outcrossed with *gfp* expressing animals to remove the *oma-1(zu405ts)* allele, allowing this experiment to be performed at 20°C. Note: because the silenced *oma-1(zu405ts)* and *gfp* alleles are paramutagenic in *prg-1(gg531)* mutants, the wild-type *oma-1* allele in the new strain (YY1256) is also silent, as is the *gfp* allele. Two hermaphrodites were then mated with animals homozygous for the expressed <u>*dpy-10(e128) gfp*</u> allele. Lineages were established from homozygous <u>*dpy-10(e128) gfp*</u> F_2 animals, which were genotyped for *prg-1(gg531)* once they had laid a brood. Gravid adults were egg prepped every generation to maintain these lineages. *gfp* expression was scored in the F3 and F7 generations using the Plan-Apochromat 20 × /0.8 M27 objective on an Axio Observer.Z1 fluorescent microscope (Zeiss). At least 50 animals were counted for each strain.

F. CRISPR/Cas9-mediated genome editing

The CRISPR/Cas9 strategy described previously (Arribere *et al.* 2014; Farboud and Meyer 2015; Wan *et al.* 2018) was used to generate deletions of *rde-3* and *hrde-1*, as well as to introduce the DDH → DAH mutation in *prg-1.* Guide RNA plasmids and repair template DNA were prepared as described previously (Wan *et al.* 2018). All guide RNAs were designed using the guide RNA selection tool CRISPOR (Haeussler *et al.* 2016).

G. Gene expression quantification using qRT-PCR

Total RNA was extracted from gravid adult animals using TRIzol Reagent (Life Technologies, 15596018). 2ug of total RNA was reverse-transcribed to generate first-strand cDNA using the Superscript III First-Strand Synthesis System (Invitrogen, 18080051) and random hexamers. Note: total RNA was heated with dNTPs and random hexamers to 65°C for 5 mins and immediately chilled on ice before proceeding with remaining cDNA synthesis steps. First-strand cDNA was then treated with RNAse H at 37°C for 20 mins. cDNA was then 1:100 (for *oma-1* and histone RNA quantification) or 1:8 (for all other quantifications) and 2ul was used to quantify gene expression using iTaq Universal SYBR Green Supermix (Bio-Rad) according to manufacturer's instructions. qRT-PCRs were performed using the CFX Connect machine (Bio-Rad) and semi-skirted PCR plates (Bio-Rad, 2239441). All qRT-PCR data was normalized to quantification of *nos-3*, a germline-expressed gene. qRT-PCR primers are listed in Table 2.

Table 3.2. Oligonucleotides used in this study.

pUG PCR primers	
Name	**Sequence (5'-3')**
pUG PCR RT oligo	GCT ATG GCT GTT CTC ATG GCG GCG TCG CCA TAT TCT ACT TAC ACA CAC ACA CAC ACA C
reverse primer for 1st PCR	GCT ATG GCT GTT CTC ATG GC
reverse primer for 2nd PCR	GGC GTC GCC ATA TTC TAC TT

oma-1 1st PCR	ATG AAC GTT AAC GGT GAA AAC
oma-1 2nd PCR	AAC AAC GAG AAG ATC GAT GA
gfp 1st PCR	ATG AGT AAA GGA GAA GAA CT
gfp 2nd PCR	TTTCACTGGAGTTGTCCCAA
his-10/14/26 1st PCR	CTT GGA AAA GGA GGA GCC AA
his-10/14/26 2nd PCR	GCG TCA TCG CAA AGT GCT TC
his-11/15/44 1st PCR	GTT ACG AAG CCA AAG GAC GG
his-11/15/44 2nd PCR	GTA AGG AAT CAT ACT CCG TC
his-12/16/43 1st PCR	AAG ACC GGA GGA AAG GCC AA
his-12/16/43 2nd PCR	GTC CCG CTC ATC AAG AGC CG
r03d7.2 1st PCR	GGT CAA ACT TCC TTA TAC TT
r03d7.2 2nd PCR	CTA TTC CAA AAG AGA TGA CT
bath-13 1st PCR	ATG GTC CTC CCA AAG AAG GA
bath-13 2nd PCR	ATT CGT GCT TTC GCA CAC TT
ztf-28 1st PCR	ATG GAT ACT TCC GAT CAC AA
ztf-28 2nd PCR	GCC CCT TAT CGA CCG CTT CG
qRT-PCR primers	
oma-1 pre-mRNA forward 1	GCG TTG GCT AAT TTC CTG AAT ATC
oma-1 pre-mRNA reverse 1	CGG CGT CTA GGA TCA AAC TG
oma-1 pre-mRNA forward 2	GCT AGT TGC AAT TAA TGA CCT CTG
oma-1 pre-mRNA reverse 2	TCC AAG CGT CTA GAG CAA AC
oma-1 mRNA forward 1	ACA ACG AGA AGA TCG ATG AGC
oma-1 mRNA reverse 1	TCT GGG CAT TCG GAT TAG TTT
oma-1 mRNA forward 2	GCT TGA AGA TAT TGC ATT CAA CCA
oma-1 mRNA reverse 2	ACT GTT GAA ATG GAG GTG CAA

his-10/14/26 forward	CAT CCA AGG TAT CAC CAA GCC G
his-10/14/26 reverse	GTA TGT GAC GGC ATC ACG GAT C
his-11/15/44 forward	GAA GCC AAA GGA CGG AAA GA
his-11/15/44 reverse	GAA ACT CCA GTG TCT GGA TGA A
his-12/16/43 forward	CCC AAG ACA TCT TCA ACT TGC C
his-12/16/43 reverse	CTC CTC CTT GAG CGA TTG TG
r03d7.2 forward	GAT CAA GTC ATC GCT CAG GAA G
r03d7.2 reverse	CGG TTG GAT CCG AAG ATG TG
bath-13 forward	TGA AGA AAG CTT GTC TCC TGA A
bath-13 reverse	CAC GAC TGT TGG ATC CAT GT
ztf-28 forward	CGG ATC TCA GAG GCG ATT TC
ztf-28 reverse	CGT CGC TAT GCA CCA GTT TA
nos-3 forward	GCC TCA GCA ACA ACA TCA AC
nos-3 reverse	CTC CTT GTG GCA TTC CGT AT
CRISPR guide RNAs (gRNAs)	
rde-3 deletion gRNA 1	GGA CTT CGC TGA TCT TTA TT
rde-3 deletion gRNA 2	GCC ACA GAA GTG AAT CGA GT
hrde-1 deletion gRNA 1	GTC TCG TTT GGG AAT CCT GT
hrde-1 deletion gRNA 2	GAA CGG ACA TCT CAT CAA CAA
prg-1 DAH mutation gRNA	GCG CCA GCA CCA TCT CTG TAG

H. Taqman-based small RNA quantification

1ug of Trizol-extracted RNA from gravid adult animals (see above) was used for Taqman assays. Small RNAs were reverse transcribed into cDNA using the Taqman MicroRNA Reverse Transcription Kit (Applied Biosystems, 4366596). *oma-1* and *gfp* small RNAs were then quantified

by qRT-PCR using TaqMan Universal Master Mix II, no UNG (Applied Biosystems, 4440040) and custom TaqMan small RNA assays from Applied Biosystems (assay IDs: *gfp:* CSLJH0V, *oma-1* probe 1: CSKAJ9W, *oma-1* probe 2: CSLJIF4, *oma-1* probe 3: CSMSGMC, *oma-1* probe 4: CSN1ESK). qRT-PCRs were performed using the CFX Connect machine (Bio-Rad) and semi-skirted PCR plates (Bio-Rad, 2239441).

I. pUG RNA detection using pUG PCR

pUG RNAs were detected using pUG PCR, a PCR-based assay described previously (Shukla *et al.* 2020). Briefly, total RNA was extracted using TRIzol Reagent (Life Technologies, 15596018). 5ug of total RNA and 1pmol of reverse transcription (RT) oligo (see Table 2) was used to generate first-strand cDNA using the Superscript III First-Strand Synthesis System (Invitrogen, 18080051). Total RNA, dNTPs and RT oligo were mixed and heated to 65°C for 5 mins and immediately chilled on ice before proceeding with remaining cDNA synthesis steps. 1ul of cDNA was used to perform a first round of PCR (20ul volume) with Taq DNA polymerase (New England BioLabs, M0273) for 20-25 cycles and primers listed in Table 2. These PCRs were then diluted 1:100 and 1ul was used to perform a second round of PCR (50ul volume) for 25-30 cycles using primers listed in Table 2. *gsa-1,* which has an 18nt long genomically UG repeat in its 3'UTR, served as a control for all pUG PCR analyses. PCR reactions were then run on 1.5-2% agarose gels. Images were acquired using a ChemiDoc MP Imaging System (Bio-Rad). All image processing was done using Fiji (Schindelin *et al.* 2012). All pUG PCR reactions were sequenced by cutting out lanes of interest from agarose gels and gel extracting the DNA using QIAquick Gel Extraction Kit (Qiagen, 28706). 3ul of gel extracted PCR product was used for TA cloning with the pGEM-T Easy Vector System (Promega, A1360) according to manufacturer's instructions. Ligation reactions were incubated overnight at 4°C. Transformations were performed with 5-alpha Competent *E. coli* cells (NEB, C2987H) and plated on LB/ampicillin/IPTG/X-gal plates (according to pGEM-T Easy Vector System manufacturer's instructions). White colonies were selected on

the day next and inoculated in Luria Broth overnight. Liquid cultures were then miniprepped using QIAprep Spin Miniprep Kit (Qiagen, 27106) and plasmid DNA was Sanger sequenced using a universal SP6 primer (5'-CATACGATTTAGGTGACACTATAG-3; Dana-Farber/Harvard Cancer Center DNA Resource Core, Harvard Medical School) or a universal M13 primer (5'-TGTAAAACGACGGCCAGT-3'; Quintarabio, Cambridge, MA).

VI. Acknowledgements

We thank past and present members of the Kennedy lab helpful discussions. Some strains were provided by the *Caenorhabditis* Genetics Center (CGC), which is funded by the NIH Office of Research Infrastructure Programs (P40 OD010440). Some strains were provided by the Mitani laboratory through the National BioResource Project (Tokyo, Japan), which is part of the International *C. elegans* Gene Knockout Consortium. A.S. (DGE1144152, DGE1745303) was supported by NSF Graduate Research Fellowship.

VII. References

Arribere J. A., R. T. Bell, B. X. H. Fu, K. L. Artiles, P. S. Hartman, *et al.*, 2014 Efficient marker-free recovery of custom genetic modifications with CRISPR/Cas9 in Caenorhabditis elegans. Genetics 198: 837–846.

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

Aravin A., D. Gaidatzis, S. Pfeffer, M. Lagos-Quintana, P. Landgraf, *et al.*, 2006 A novel class of small RNAs bind to MILI protein in mouse testes. Nature 442: 203–207.

Aravin A. A., R. Sachidanandam, D. Bourc'his, C. Schaefer, D. Pezic, *et al.*, 2008 A piRNA pathway primed by individual transposons is linked to de novo DNA methylation in mice. Mol. Cell 31: 785–799.

Ashe A., A. Sapetschnig, E.-M. Weick, J. Mitchell, M. P. Bagijn, *et al.*, 2012 piRNAs can trigger a multigenerational epigenetic memory in the germline of C. elegans. Cell 150: 88–99.

Bagijn M. P., L. D. Goldstein, A. Sapetschnig, E.-M. Weick, S. Bouasker, *et al.*, 2012 Function, targets, and evolution of Caenorhabditis elegans piRNAs. Science 337: 574–578.

Barucci G., E. Cornes, M. Singh, B. Li, M. Ugolini, *et al.*, 2020 Small-RNA-mediated transgenerational silencing of histone genes impairs fertility in piRNA mutants. Nat. Cell Biol. 22: 235–245.

Batista P. J., J. G. Ruby, J. M. Claycomb, R. Chiang, N. Fahlgren, *et al.*, 2008 PRG-1 and 21U-RNAs interact to form the piRNA complex required for fertility in C. elegans. Mol. Cell 31: 67–78.

Brangwynne C. P., C. R. Eckmann, D. S. Courson, A. Rybarska, C. Hoege, *et al.*, 2009 Germline P granules are liquid droplets that localize by controlled dissolution/condensation. Science 324: 1729–1732.

Brenner S., 1974 The genetics of Caenorhabditis elegans. Genetics 77: 71–94.

Buckley B. A., K. B. Burkhart, S. G. Gu, G. Spracklin, A. Kershner, *et al.*, 2012 A nuclear Argonaute promotes multigenerational epigenetic inheritance and germline immortality. Nature 489: 447–451.

Chandler V. L., 2010 Paramutation's properties and puzzles. Science 330: 628–629.

Chen C.-C. G., M. J. Simard, H. Tabara, D. R. Brownell, J. A. McCollough, *et al.*, 2005 A

member of the polymerase beta nucleotidyltransferase superfamily is required for RNA interference in C. elegans. Curr. Biol. 15: 378–383.

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

Duchaine T. F., J. A. Wohlschlegel, S. Kennedy, Y. Bei, D. Conte Jr, *et al.*, 2006 Functional proteomics reveals the biochemical niche of C. elegans DCR-1 in multiple small-RNA-mediated pathways. Cell 124: 343–354.

Farboud B., and B. J. Meyer, 2015 Dramatic enhancement of genome editing by CRISPR/Cas9 through improved guide RNA design. Genetics 199: 959–971.

Filipowicz W., 2005 RNAi: the nuts and bolts of the RISC machine. Cell 122: 17–20.

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

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

Gruidl M. E., P. A. Smith, K. A. Kuznicki, J. S. McCrone, J. Kirchner, *et al.*, 1996 Multiple potential germ-line helicases are components of the germ-line-specific P granules of Caenorhabditis elegans. Proc. Natl. Acad. Sci. U. S. A. 93: 13837–13842.

Gu W., M. Shirayama, D. Conte Jr, J. Vasale, P. J. Batista, *et al.*, 2009 Distinct argonaute-mediated 22G-RNA pathways direct genome surveillance in the C. elegans germline. Mol. Cell 36: 231–244.

Haeussler M., K. Schönig, H. Eckert, A. Eschstruth, J. Mianné, *et al.*, 2016 Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR. Genome Biol. 17: 148.

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

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

Lev I., U. Seroussi, H. Gingold, R. Bril, S. Anava, *et al.*, 2017 MET-2-Dependent H3K9 Methylation Suppresses Transgenerational Small RNA Inheritance. Curr. Biol. 27: 1138–1147.

Li X. Z., C. K. Roy, X. Dong, E. Bolcun-Filas, J. Wang, *et al.*, 2013 An Ancient Transcription Factor Initiates the Burst of piRNA Production during Early Meiosis in Mouse Testes. Molecular Cell 50: 67–81.

Lin R., 2003 A gain-of-function mutation in oma-1, a C. elegans gene required for oocyte maturation, results in delayed degradation of maternal proteins and embryonic lethality. Dev. Biol. 258: 226–239.

Luteijn M. J., and R. F. Ketting, 2013 PIWI-interacting RNAs: from generation to transgenerational epigenetics. Nat. Rev. Genet. 14: 523–534.

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

Meister G., 2013 Argonaute proteins: functional insights and emerging roles. Nat. Rev. Genet.

14: 447–459.

Ozata D. M., I. Gainetdinov, A. Zoch, D. O'Carroll, and P. D. Zamore, 2019 PIWI-interacting RNAs: small RNAs with big functions. Nat. Rev. Genet. 20: 89–108.

Parker G. S., D. M. Eckert, and B. L. Bass, 2006 RDE-4 preferentially binds long dsRNA and its dimerization is necessary for cleavage of dsRNA to siRNA. RNA 12: 807–818.

Parrish S., and A. Fire, 2001 Distinct roles for RDE-1 and RDE-4 during RNA interference in Caenorhabditis elegans. RNA 7: 1397–1402.

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

Perales R., D. Pagano, G. Wan, B. D. Fields, A. L. Saltzman, *et al.*, 2018 Transgenerational Epigenetic Inheritance Is Negatively Regulated by the HERI-1 Chromodomain Protein. Genetics 210: 1287–1299.

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

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

Reed K. J., J. M. Svendsen, K. C. Brown, B. E. Montgomery, T. N. Marks, *et al.*, 2020 Widespread roles for piRNAs and WAGO-class siRNAs in shaping the germline transcriptome of Caenorhabditis elegans. Nucleic Acids Res. 48: 1811–1827.

Ruby J. G., C. Jan, C. Player, M. J. Axtell, W. Lee, *et al.*, 2006 Large-scale sequencing reveals 21U-RNAs and additional microRNAs and endogenous siRNAs in C. elegans. Cell 127: 1193–1207.

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

Schindelin J., I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, *et al.*, 2012 Fiji: an open-source platform for biological-image analysis. Nat. Methods 9: 676–682.

Shen E.-Z., H. Chen, A. R. Ozturk, S. Tu, M. Shirayama, *et al.*, Identification of piRNA binding sites reveals the Argonaute regulatory landscape of the C. elegans germline

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

Shukla A., J. Yan, D. J. Pagano, A. E. Dodson, Y. Fei, *et al.*, 2020 poly(UG)-tailed RNAs in genome protection and epigenetic inheritance. Nature 582: 283–288.

Siomi M. C., K. Sato, D. Pezic, and A. A. Aravin, 2011 PIWI-interacting small RNAs: the vanguard of genome defence. Nat. Rev. Mol. Cell Biol. 12: 246–258.

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

Spracklin G., B. Fields, G. Wan, D. Becker, A. Wallig, *et al.*, 2017 The RNAi Inheritance Machinery of Caenorhabditis elegans. Genetics 206: 1403–1416.

Tabara H., M. Sarkissian, W. G. Kelly, J. Fleenor, A. Grishok, *et al.*, 1999 The rde-1 gene, RNA interference, and transposon silencing in C. elegans. Cell 99: 123–132.

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

Tsai H.-Y., C.-C. G. Chen, D. Conte Jr, J. J. Moresco, D. A. Chaves, *et al.*, 2015 A ribonuclease coordinates siRNA amplification and mRNA cleavage during RNAi. Cell 160: 407–419.

Vastenhouw N. L., K. Brunschwig, K. L. Okihara, F. Müller, M. Tijsterman, *et al.*, 2006 Gene expression: long-term gene silencing by RNAi. Nature 442: 882.

Vourekas A., Q. Zheng, P. Alexiou, M. Maragkakis, Y. Kirino, *et al.*, 2012 Mili and Miwi target RNA repertoire reveals piRNA biogenesis and function of Miwi in spermiogenesis. Nature Structural & Molecular Biology 19: 773–781.

Vourekas A., P. Alexiou, N. Vrettos, M. Maragkakis, and Z. Mourelatos, 2016 Sequence-dependent but not sequence-specific piRNA adhesion traps mRNAs to the germ plasm. Nature 531: 390–394.

Wan G., B. D. Fields, G. Spracklin, A. Shukla, C. M. Phillips, *et al.*, 2018 Spatiotemporal regulation of liquid-like condensates in epigenetic inheritance. Nature 557: 679–683.

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

Weick E.-M., P. Sarkies, N. Silva, R. A. Chen, S. M. M. Moss, *et al.*, 2014 PRDE-1 is a nuclear factor essential for the biogenesis of Ruby motif-dependent piRNAs in C. elegans. Genes Dev. 28: 783–796.

Yigit E., P. J. Batista, Y. Bei, K. M. Pang, C.-C. G. Chen, *et al.*, 2006 Analysis of the C. elegans Argonaute family reveals that distinct Argonautes act sequentially during RNAi. Cell 127:

747–757.

Zhang D., S. Tu, M. Stubna, W.-S. Wu, W.-C. Huang, *et al.*, 2018 The piRNA targeting rules and the resistance to piRNA silencing in endogenous genes. Science 359: 587–592.

Zhuang J. J., and C. P. Hunter, 2012 The Influence of Competition Among C. elegans Small RNA Pathways on Development. Genes 3. https://doi.org/10.3390/genes3040671

Chapter 4

Summary and Perspectives

The work presented in this dissertation reports the discovery of a new 3' RNA modification that amplifies RNAi-mediated gene silencing in *C. elegans* and identifies one mechanism that *C. elegans* uses to ensure that this modification is added to the correct RNAs. Below I will summarize some of the key findings presented in Chapters 2 and 3, discuss some questions that arise from this work and propose some ideas for how some of these questions may be tested.

I. Identification of poly(UG)-tailed RNAs in *C. elegans*

A paradox exists within the RNAi pathway in *C. elegans*, namely that the function of RNAi is to silence mRNAs, but at the same time, RdRPs need to use mRNAs as templates to generate amplified pools of 22G-siRNAs that carry out gene silencing. In Chapter 2 of this dissertation, we show that fragments of mRNAs that are targeted for silencing by RNAi pathways, such as mRNAs targeted by experimental dsRNA and transposon RNAs, can be modified with pUG tails in *C. elegans*. pUG tails then serve to recruit RdRPs to pUG RNAs and RdRPs use the mRNA fragments appended to pUG tails as templates for siRNA biogenesis, thereby amplifying RNAi-mediated gene silencing. Thus, by using pUG RNAs, rather than full-length mRNAs, as templates for RdRPs, the RNAi machinery can silence (i.e. fragment) mRNAs, but still store the sequence information contained in mRNAs in the form of pUG RNAs. In other words, pUG RNAs can serve as a memory of RNAi silencing, without serving as a template for the translation machinery.

Why the worm uses this highly ordered repeat to recruit proteins, like RdRPs, is still a mystery. Two somewhat inseparable possibilities are: (1) a poly(UG) tail forms a unique structure that recruits RNA-binding proteins, or (2) a poly(UG) tail protects mRNA fragments from degradation, thereby allowing them to serve as templates for RdRPs. These possibilities are somewhat inseparable because it is possible that a unique structure is what could make a pUG RNA more stable and potentially resistant to exonucleases. The G-rich nature of a pUG tail could promote G-quadruplex formation (Varshney *et al.* 2020) or the unique properties of U and G to

form wobble base-pairs (Varani and McClain 2000) could promote the formation of a secondary structure, either of which could be recognized by pUG tail-binding proteins. Whether or not pUG RNAs are stabilized by pUG tails could be tested by incubating radiolabeled mRNA fragments with different 3' tails, including 3' pUG tails, with worm lysates and monitoring the rate at which these mRNA fragments are degraded.

While exogenous dsRNA is a clear trigger that initiates pUGylation and silencing of an mRNA, how the proper RNAs (e.g. transposon RNA fragments and the other mRNA fragments that we show are pUGylated in wild-type animals), but not essential mRNAs, are targeted for pUG RNA-dependent silencing is a mystery. This mystery can be broken down into a few parts: (1) how does the RNAi machinery know to target some mRNAs for silencing; (2) how does an mRNA that is recognized by the RNAi machinery enter *Mutator* foci; and (3) how does RDE-3 recognize the mRNA fragments that it pUGylates. Our data shows that pUGylation occurs downstream of RDE-4, RDE-1 and RDE-8, factors that recognize and mediate the cleavage of a target mRNA (Parrish and Fire 2001; Tabara *et al.* 2002; Tsai *et al.* 2015). RDE-8 and RDE-3 have been shown to physically interact in *Mutator* foci (Tsai *et al.* 2015), suggesting that RDE-8 might recruit RDE-3 to transcripts that it has cleaved, thereby providing RDE-3 with a free 3' end to pUGylate. However, how mRNAs are recognized by the RNAi machinery and brought into *Mutator* foci is still not known. Past work has suggested that the terminal inverted repeats that flank DNA transposon sequences can base-pair, thereby forming dsRNAs that trigger RNAi (Sijen and Plasterk 2003). While this model helps to explain how mRNAs derived from DNA transposons might be recognized by the RNAi machinery, and handed off to RDE-3 for pUGylation, it does not explain how the RNAi machinery recognizes the other mRNAs that it silences. Forward genetic screens that look for shared factors required for silencing of endogenously pUGylated mRNAs might help to reveal the proteins that help to identify mRNAs that are fated to be silenced by the RNAi machinery and help to drive mRNAs into *Mutator foci* for RDE-8-dependent cleavage and RDE-3-dependent pUGylation.

The localization of pUG RNAs to *Mutator* foci is intriguing for a number of reasons. First, a number of *Mutator* proteins, such as MUT-7 and MUT-8/RDE-2, are known to be required for dsRNA-induced gene silencing and transposon silencing (Ketting *et al.* 1999; Sijen and Plasterk 2003; Phillips *et al.* 2012), but their molecular activities have yet to be identified. However, we can now ask whether and how these uncharacterized MUT/RDE proteins influence pUG RNA/siRNA cycling. Second, *Mutator* foci are phase separated biomolecular condensates in the *C. elegans* germline. The interactions between proteins and RNA molecules are known to play a critical role in promoting liquid-liquid phase separation (Molliex *et al.* 2015). *In vitro,* RNA molecules have been shown to increase the liquidity of biomolecular condensates by decreasing their viscosity (Elbaum-Garfinkle *et al.* 2015). Further, RNA molecules alone are capable of phase separating (Jain and Vale 2017). Together, these and other observations noting the importance of RNA in phase separation beg the question of whether pUG RNAs play a role in the phase separation of *Mutator* foci. Indeed, we have shown that proteins are capable of interacting with pUG repeats, but whether pUG RNA-protein interactions, or even pUG RNA-pUG RNA interactions (mediated by quadruplex formation or GU wobble base pairing) can promote or help to maintain the liquid-like character of *Mutator* foci remains to be studied. While it is known that MUT-16 still forms perinuclear foci in the absence of RDE-3 (Phillips *et al.* 2012), it is not known whether the protein and RNA constituents of these foci are changed (aside from the obvious lack of RDE-3 and pUG RNAs and, consequently, potentially RRF-1) or if the biophysical properties of these foci are changed. Measuring the liquid-like character of *Mutator* foci (with techniques such as fluorescence recovery after photobleaching (Brangwynne *et al.* 2009)), in wild-type animals versus *rde-3* mutants, could be informative in getting at this question, as could an immunoprecipitation, followed by mass spectrometry, of MUT-16-interacting proteins in wild-type animals versus *rde-3* mutants.

The work presented in Chapter 2 also helps to explain an observation that was made in the first paper that described dsRNA-mediated gene silencing (Fire *et al.* 1998), namely that

dsRNA-triggered gene silencing can be propagated across generations in the absence of any additional silencing triggers. We find that generationally repeated rounds of pUG RNA-templated 22G-siRNA synthesis and 22G-siRNA-directed mRNA pUGylation (pUG RNA/siRNA cycling) drive RNAi inheritance in *C. elegans.* The self-sufficient relationship between pUG RNAs and 22G-siRNAs shows why dsRNA triggers are no longer needed to maintain gene silencing in inheriting generations. Future work will no doubt explore whether pUG RNAs or siRNAs are the inherited molecule that is transmitted from one generation to the next, but our work shows that either molecule is sufficient to kick off the pUG RNA/siRNA cycling that drives RNAi silencing. Indeed, the localization of pUG RNAs (Shukla *et al.* 2020), as well as RDE-3 and RRF-1 (Phillips *et al.* 2012), to *Mutator* foci, germline-localized biomolecular condensates, means that the pUG RNA/siRNA silencing loop seems perfectly poised to transmit silencing information from one generation to the next. Based on our data showing the shortening of *oma-1* pUG RNAs across generations during *oma-1* RNAi inheritance, we also propose a model to explain the finite nature of RNAi inheritance whereby the directionality of RdRPs results in pUG RNA shortening such that, eventually, there is not enough pUG RNA template left for siRNA biogenesis and cycling ceases. Interestingly, a number of recent studies have shown that environmental triggers, such as starvation or pathogenic bacteria, can trigger TEI that lasts for 3-4 generations and is also dependent on siRNAs and RNAi factors, such as HRDE-1 (Remy 2010; Rechavi *et al.* 2014; Schott *et al.* 2014; Jobson *et al.* 2015; Moore *et al.* 2019; Posner *et al.* 2019). Future work may reveal that environmentally-triggered pUGylation of one or more mRNAs promotes these examples of TEI. While it is still not known why the worm has a system that is able to amplify and transgenerationally transmit epigenetic information, recent work from *C. elegans* showing the transgenerational transmission of pathogenic avoidance behavior (Moore *et al.* 2019; Kaletsky *et al.* 2020) suggests that, perhaps, TEI can be adaptive.

The work presented in Chapter 2 also opens up two exciting areas of future research. The first is the search for whether pUG RNAs exist in other organisms. This question is made more

difficult by the fact that the primary sequence of a nucleotidyltransferase is not sufficient to reveal its biological activity so one cannot simply look for poly(UG) polymerases in other organisms using a homology search. Thus, nucleotidyltransferase activity must be tested on a case-by-case basis. Because pUG RNAs do not show up in normal RNA-seq libraries, potentially because they are lowly expressed or difficult substrates for adapter ligation, sequencing approaches developed to specifically enrich for pUG-tailed RNAs (pUG-seq) could get at this question. One pitfall with this approach is the lack of a negative control (i.e. a polymerase mutant) to help reveal sequencing artifacts in organisms other than *C. elegans.* An obvious place to start this search would be in organisms that have RdRPs, such as *S. pombe* and plants (Holoch and Moazed 2015), in the event that the function of pUG RNAs serving as templates for siRNA biogenesis is conserved. However, pUG RNAs need not just serve as templates for RdRPs. Indeed, if pUG tails do form a unique structure that can recruit proteins that recognize this structure, then long UG repeats, whether genomically-encoded or post-transcriptionally added, could serve to recruit proteins during any process that uses information stored in RNA. Our work also inspires the search for other 3' post-transcriptional modifications that might exist in *C. elegans* and other organisms. Indeed, any sequencing approach to search for pUG RNA could be adapted to search for other 3' modifications.

II. piRNAs protect mRNAs from aberrant poly(UG) tailing

Chapter 2 of this dissertation prompted the question of how RDE-3 identifies its targets. Chapter 3 of this dissertation reveals that *C. elegans* piRNAs are one mechanism that helps to restrict RDE-3 activity to the correct mRNAs. Indeed, we find that, in the absence of the PIWI Argonaute PRG-1 and piRNAs, RDE-3 pUGylates the wrong mRNAs, thereby driving the aberrant silencing of germline-expressed mRNAs that has previously been reported to occur in piRNA mutants (de Albuquerque *et al.* 2014; Phillips *et al.* 2015; Barucci *et al.* 2020; Reed *et al.* 2020).

Thus, while piRNAs are perhaps best known for their conserved role in mediating gene silencing (Ozata *et al.* 2019), our work helps to reveal that piRNAs can also have (direct or indirect) pro-expression functions. An interesting consequence of misdirected RDE-3 activity in the absence of piRNAs is that dsRNA-induced heritable gene silencing can become permanent in *prg-1* mutants. This permanent silencing is driven by the permanent activation of pUG RNA/siRNA cycling, which is normally confined to a finite number of generations after RNAi in wild-type animals. Thus, piRNAs normally prevent targets of dsRNA from being silenced forever. Taken together, our data indicates that piRNAs help to organize the *C. elegans* pUGylome (e.g. all of the pUG RNAs in the worm) and also points to a delicate balance that exists between small RNA pathways in *C. elegans.*

It is currently not known what factors determine which mRNAs undergo aberrant pUGylation in the absence of piRNAs. Indeed, how RDE-3 normally identifies its targets in wild-type animals is not known, but perhaps studying the triggers of aberrant pUGylation in *prg-1* mutants could help get at what factors normally regulate RDE-3 activity. It is clear that some mRNAs, like the histone mRNAs, are more highly predisposed to silencing in the absence of piRNAs, whereas other mRNAs have a lower probability of becoming silenced. One outstanding question, thus, is whether every mRNA has some probability of undergoing aberrant pUGylation and silencing in the absence of piRNAs or if only a subset of mRNAs are at risk of being aberrantly pUGylated and silenced. One could try to address this question by establishing lineages of *prg-1* mutants that each arose from a single animal and then using RNA-seq (or pUG-seq) to look for the variation in silenced mRNAs. If only some, but not all, mRNAs can become improperly pUGylated in *prg-1* mutants, then a study of this subset of mRNAs (e.g. localization, number of PRG-1/piRNA binding sites, number of 22G-siRNAs normally targeting this mRNAs, etc.) could help reveal the aberrant pUGylation trigger.

What causes the disorganization of the *C. elegans* pUGylome in the absence of piRNAs? One model that may explain how piRNAs normally protect mRNAs from RDE-3-driven silencing

comes from several studies that have explored the function of piRNAs in *C. elegans* and *Drosophila*, recently. PRG-1 and piRNAs in *C. elegans* were found to interact with most of the germline transcriptome (Shen *et al.* 2018), yet very few of the mRNAs with which PRG-1 and piRNAs interact are normally silenced as a result of this interaction (Barucci *et al.* 2020; Reed *et al.* 2020). Therefore, it is likely that the binding of PRG-1 and piRNAs to mRNAs has some other function. One such function might be that the presence of PRG-1 and piRNAs on mRNAs helps to protect these mRNAs from being silenced by other pathways. This protective function of PRG-1 and piRNAs could be due to this complex physically blocking other proteins from binding to mRNAs or by this complex potentially sequestering mRNAs to P granules, where PRG-1 is normally localized. A combination of RNA FISH to look at mRNAs that undergo silencing in *prg-1* mutants and immunofluorescence to look at PRG-1 and P granules might help to reveal whether the mRNAs that undergo gene silencing in *prg-1* mutants are, indeed, normally sequestered to P granules and become mislocalized in *prg-1* mutants (potentially to *Mutator* foci). Indeed, Piwi-bound piRNAs in *Drosophila* were shown to bind to and sequester maternal transcripts in the germ plasm, thereby promoting germ cell specification in embryos (Vourekas *et al.* 2016). A more indirect model of why mRNAs undergo aberrant gene silencing in the absence of PRG-1 and piRNAs is that, without piRNAs and piRNA-directed 22G-siRNAs, the RNAi machinery that induces gene silencing downstream of piRNAs, such as HRDE-1, becomes more available and begins to misdirect its silencing activity to the wrong mRNAs.

While the aberrant pUGylation and silencing of mRNAs in *prg-1* mutants is interesting in that it reveals a novel pro-expression function of piRNAs, an enticing question that arises from this work is whether natural or stochastic variation in siRNA levels or other epigenetic cues could drive mRNAs into the pUG RNA/siRNA silencing loop in wild-type animals. In other words, could spontaneous pUGylation-induced epigenetic silencing occur in wild-type animals and provide epigenetic variation that could be selected for (or against)? The low probability of such an event that is based on molecular thresholds would ensure that its penetrance would be low in a

population, but if populations of isogenic animals could be established by singling many wild-type animals, then one could probe the natural variation in the pUGylome of wild-type animals. While sequencing efforts could identify this natural variation if it exists, the harder question to answer becomes whether this natural variation in the pUGylome is advantageous.

III. References

Albuquerque B. F. M. de, M. J. Luteijn, R. J. Cordeiro Rodrigues, P. van Bergeijk, S. Waaijers, *et al.*, 2014 PID-1 is a novel factor that operates during 21U-RNA biogenesis in Caenorhabditis elegans. Genes Dev. 28: 683–688.

Barucci G., E. Cornes, M. Singh, B. Li, M. Ugolini, *et al.*, 2020 Small-RNA-mediated transgenerational silencing of histone genes impairs fertility in piRNA mutants. Nat. Cell Biol. 22: 235–245.

Brangwynne C. P., C. R. Eckmann, D. S. Courson, A. Rybarska, C. Hoege, *et al.*, 2009 Germline P granules are liquid droplets that localize by controlled dissolution/condensation. Science 324: 1729–1732.

Elbaum-Garfinkle S., Y. Kim, K. Szczepaniak, C. C.-H. Chen, C. R. Eckmann, *et al.*, 2015 The disordered P granule protein LAF-1 drives phase separation into droplets with tunable viscosity and dynamics. Proc. Natl. Acad. Sci. U. S. A. 112: 7189–7194.

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

Holoch D., and D. Moazed, 2015 RNA-mediated epigenetic regulation of gene expression. Nat. Rev. Genet. 16: 71–84.

Jain A., and R. D. Vale, 2017 RNA phase transitions in repeat expansion disorders. Nature 546: 243–247.

Jobson M. A., J. M. Jordan, M. A. Sandrof, J. D. Hibshman, A. L. Lennox, *et al.*, 2015 Transgenerational Effects of Early Life Starvation on Growth, Reproduction, and Stress Resistance in Caenorhabditis elegans. Genetics 201: 201–212.

Kaletsky R., R. S. Moore, G. D. Vrla, L. R. Parsons, Z. Gitai, *et al.*, 2020 C. elegans interprets bacterial non-coding RNAs to learn pathogenic avoidance. Nature 586: 445–451.

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

Molliex A., J. Temirov, J. Lee, M. Coughlin, A. P. Kanagaraj, *et al.*, 2015 Phase separation by low complexity domains promotes stress granule assembly and drives pathological fibrillization. Cell 163: 123–133.

Moore R. S., R. Kaletsky, and C. T. Murphy, 2019 Piwi/PRG-1 Argonaute and TGF-β Mediate Transgenerational Learned Pathogenic Avoidance. Cell 177: 1827–1841.e12.

Ozata D. M., I. Gainetdinov, A. Zoch, D. O'Carroll, and P. D. Zamore, 2019 PIWI-interacting RNAs: small RNAs with big functions. Nat. Rev. Genet. 20: 89–108.

Parrish S., and A. Fire, 2001 Distinct roles for RDE-1 and RDE-4 during RNA interference in Caenorhabditis elegans. RNA 7: 1397–1402.

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

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

Posner R., I. A. Toker, O. Antonova, E. Star, S. Anava, *et al.*, 2019 Neuronal Small RNAs Control Behavior Transgenerationally. Cell 177: 1814–1826.e15.

Rechavi O., L. Houri-Ze'evi, S. Anava, W. S. S. Goh, S. Y. Kerk, *et al.*, 2014 Starvation-induced transgenerational inheritance of small RNAs in C. elegans. Cell 158: 277–287.

Reed K. J., J. M. Svendsen, K. C. Brown, B. E. Montgomery, T. N. Marks, *et al.*, 2020 Widespread roles for piRNAs and WAGO-class siRNAs in shaping the germline transcriptome of Caenorhabditis elegans. Nucleic Acids Res. 48: 1811–1827.

Remy J.-J., 2010 Stable inheritance of an acquired behavior in Caenorhabditis elegans. Curr. Biol. 20: R877–8.

Schott D., I. Yanai, and C. P. Hunter, 2014 Natural RNA interference directs a heritable response to the environment. Sci. Rep. 4: 7387.

Shen E.-Z., H. Chen, A. R. Ozturk, S. Tu, M. Shirayama, *et al.*, Identification of piRNA binding sites reveals the Argonaute regulatory landscape of the C. elegans germline

Shukla A., J. Yan, D. J. Pagano, A. E. Dodson, Y. Fei, *et al.*, 2020 poly(UG)-tailed RNAs in genome protection and epigenetic inheritance. Nature 582: 283–288.

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

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

871.

Tsai H.-Y., C.-C. G. Chen, D. Conte Jr, J. J. Moresco, D. A. Chaves, *et al.*, 2015 A ribonuclease coordinates siRNA amplification and mRNA cleavage during RNAi. Cell 160: 407–419.

Varani G., and W. H. McClain, 2000 The G x U wobble base pair. A fundamental building block of RNA structure crucial to RNA function in diverse biological systems. EMBO Rep. 1: 18–23.

Varshney D., J. Spiegel, K. Zyner, D. Tannahill, and S. Balasubramanian, 2020 The regulation and functions of DNA and RNA G-quadruplexes. Nat. Rev. Mol. Cell Biol. 21: 459–474.

Vourekas A., P. Alexiou, N. Vrettos, M. Maragkakis, and Z. Mourelatos, 2016 Sequence-dependent but not sequence-specific piRNA adhesion traps mRNAs to the germ plasm. Nature 531: 390–394.