2016-08-18

Functions of Argonaute Proteins in Self Versus Non-Self Recognition in the C. elegans Germline: A Dissertation

Meetu Seth

University of Massachusetts Medical School

FUNCTIONS OF ARGONAUTE PROTEINS IN SELF VERSUS NON-SELF RECOGNITION IN THE *C. ELEGANS* GERMLINE

A Dissertation Presented

By

Meetu Seth

Submitted to the Faculty of the

University of Massachusetts Graduate School of Biomedical Sciences, Worcester

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

August 18, 2016

Molecular Biology and Genetics

FUNCTIONS OF ARGONAUTE PROTEINS IN SELF VERSUS NON-SELF RECOGNITION IN THE *C. ELEGANS* GERMLINE

A Dissertation Presented By
MEETU SETH

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

Craig C. Mello, Ph.D., Thesis Advisor

Jeanne Lawrence, Ph.D., Member of Committee

Oliver Rando, Ph.D., Member of Committee

Sean P. Ryder, Ph.D., Member of Committee

Scott Kennedy, Ph.D., Member of Committee

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

Victor Ambros, Ph.D., Chair of Committee

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

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

Interdisciplinary Graduate Program
August 18, 2016

DEDICATION

I dedicate my thesis to my family

Utsav, Shaurya, and Sanjeev,

for their love and support

my parents and mother in law

for their blessings and encouragement.

ACKNOWLEDGMENTS

I extend my appreciation from the bottom of my heart to my parents who always taught me to give primary importance to education in life. They have been inspiration to what I have achieved till date. I also extend my thanks to my sisters for their continuous encouragement during my study. My son Utsav and husband Sanjeev have always given me uncompromised support during my Ph.D., with Utsav often spending weekends with me in the lab as I worked on my projects.

I am grateful to Dr. Craig Mello, my thesis advisor for his immense support, invaluable mentorship and continued guidance. As a child, I always dreamed about becoming a scientist and I am fortunate to work with one of the great mentors in scientific field. I am thankful to him for providing me an opportunity to work on exciting projects with novel ideas. Under his guidance, I have learned how to approach a scientific question, be a logical thinker and never give up. These years have been the best years of my life and I sincerely thank Craig for all his efforts in making me the scientist I am today.

I would like to express my gratitude to my committee chair Dr. Victor Ambros, and other committee members including Dr. Jeanne Lawrence, Dr. Oliver Rando and Dr. Sean Ryder for their guidance and useful suggestions during my TRAC meetings. They have been phenomenal in discussing ideas on my projects and helping me move forward to conclude my thesis work. I would also like to thank my external committee member Dr. Scott Kennedy for taking time to read my thesis and attending my defense presentation. I am also grateful

to Dr. Melissa Moore, Dr. Erik Sontheimer and Dr. Anastasia Khvorova for their excellent counseling and guidance for my career related queries.

I would like to thank my colleagues and friends at UMass for their support and help. My special thanks to Dr. Masaki Shirayama for all his collaborations on my projects from the time I joined the lab. I would also like to express my thanks to Dr. Darryl Conte and Dr. Gregoriy Dokshin for their discussions and critically reviewing my thesis. I also thank Dr. Wen Tang for all the scientific discussions on my project. I would like to thank Dr. Rita Sharma for helping me with RNAi screen. Thanks to Jean O'Connor, Deb Bordne and Laurie Cassella our lab managers who were always willing to answer any questions I had. I am grateful to my friends especially Ami Ashar-Patel for being there for me during all the good and tough times during my Ph.D. Thanks to Dr. Vahbiz Jokhi for her precious time and helpful suggestions on my thesis.

I am grateful to Dr. Weifeng Gu from University of California, Riverside for sharing all his reagents and protocols for small RNA cloning. I would like to thank our collaborator Dr. Shikui Tu from Dr. Zhiping Weng laboratory for helping me with the bioinformatics analysis on my project.

I acknowledge Howard Hughes Medical Institute for providing me with the International Student Research Fellowship to support the thesis work during my Ph.D.

Meetu, Nov 2016

ABSTRACT

Organisms employ sophisticated mechanisms to silence foreign nucleic acid, such as viruses and transposons. Evidence exists for pathways that sense copy number, unpaired DNA, or aberrant RNA (e.g., dsRNA), but the mechanisms that distinguish “self” from “non-self” are not well understood. Our studies on transgene silencing in *C. elegans* have uncovered an RNA surveillance system in which the PIWI protein, PRG-1, uses a vast repertoire of piRNAs to recognize foreign transcripts and to initiate epigenetic silencing. Partial base pairing by piRNAs is sufficient to guide PRG-1 targeting. PRG-1 in turn recruits RdRP to synthesize perfectly matching antisense siRNAs (22G-RNAs) that are loaded onto worm-specific Argonaute (WAGO) proteins. WAGOs collaborate with chromatin factors to maintain epigenetic silencing (RNAe). Since mismatches are allowed during piRNA targeting, piRNAs could—in theory—target any transcript expressed in the germline, but germline genes are not subject to silencing by RNAe. Moreover, some foreign sequences are expressed and appear to be adopted as “self.” How are “self” transcripts distinguished from foreign transcripts? We have found that another Argonaute, CSR-1, and its siRNAs—also synthesized by RdRP—protect endogenous genes from silencing by RNAe. We refer to this pathway as RNA-mediated gene activation (RNAa). Reducing CSR-1 or PRG-1 or increasing piRNA targeting can shift the balance towards expression or silencing, indicating that PRG-1 and CSR-1 compete for control over their targets. Thus worms have evolved a remarkable nucleic acids

immunity mechanism in which opposing Argonaute pathways generate and maintain epigenetic memories of self and non-self nucleotide sequences.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Copyright Information

Parts of this dissertation have been published or will be submitted for publication as:

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

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

Seth, M., Shirayama, M., Tang, W., Shen, E., Tu, S., and Mello, C.C. (2016). The *C. elegans* CSR-1 and PRG-1 Argonaute Pathways Compete with Each Other to Regulate Gene Expression (Manuscript in preparation)

ABBREVIATIONS

ADR-1: Adenosine Deaminases that act on RNA
AGO: Argonaute
ALG-1: Argonaute (plant)-Like Gene
AUB: AUBERGINE
C. elegans: Caenorhabditis elegans
CRISPR: Clustered Regularly Interspaced Short Palindromic Repeat
CSR-1: Chromosome Segregation and RNAi deficient
DCR-1: DiCer Related
dsRNA: Double stranded RNA
DNA: Deoxyribonucleic acid
ERI: Enhanced RNAi
ERGO-1: Endogenous-RNAi deficient arGOnaute
HPL-2: HP1 like heterochromatin protein
IP: Immunoprecipitation
MES: Maternal Effect Sterile
mRNA: messenger RNA
MUT-7: MUTator
MosSCI: Mos1-mediated single copy insertion
miRNAs: microRNAs
NT: nucleotide
NRDE-3: Nuclear RNAi-DEficient
piRNAs: Piwi-interacting RNAs
PIWI: P-element Induced WImpy testis
RNA: Ribonucleic acid
RNAi: RNA interference
ROL-6: (ROLler)
RDE: RNAi DEfective
RRF-1: RNA-dependent RNA polymerase Family
RdRP: RNA-directed RNA Polymerase
RNAa: RNA-induced epigenetic gene activation
RNAe: RNA-mediated epigenetic silencing
WAGOs: Worm ArGOnaute proteins

PUBLISHED WORKS

The following publication appear in whole or as a part in this thesis:

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

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

Seth, M., Shirayama, M., Tang, W., Shen, E., Tu, S., and Mello, C.C. (2016). The *C. elegans* Argonaute CSR-1 and PRG-1 Pathways Compete with Each Other to Regulate Gene Expression (Manuscript in preparation)

The following publication that was contributed but is not included in this thesis:

Shirayama, M., Stanney, W., 3rd, Gu, W., Seth, M., and Mello, C.C. (2014). The Vasa Homolog RDE-12 engages target mRNA and multiple argonaute proteins to promote RNAi in *C. elegans*. Curr Biol *24*, 845-851.

Kim, H., Ishidate, T., Ghanta, K.S., Seth, M., Conte, D., Jr., Shirayama, M., and Mello, C.C. (2014). A co-CRISPR strategy for efficient genome editing in *Caenorhabditis elegans*. Genetics *197*, 1069-1080.

CHAPTER I: INTRODUCTION

Self versus Non-self distinction in organisms

Genomes of living organisms are under constant attack by foreign nucleic acids such as transposons and viruses. If mobilized, transposons can disrupt protein-coding genes, alter transcriptional regulatory networks or lead to chromosomal breakage and genomic rearrangement (McClintock, 1951). In order to protect themselves from harmful effects of transposition and maintain their own genomic integrity, cells must actively silence transposable elements. Animals evolved a fascinating array of gene-silencing pathways to confront a constant onslaught of parasitic Ribonucleic acid (RNA) and Deoxyribonucleic acid (DNA) invaders. Since the germline serves as a reservoir of hereditary material, these defense pathways are specifically highly active in the germ cell lineage. Although evidence exists for several pathways that sense copy number, unpaired DNA or aberrant RNA, in many cases, the mechanism used to distinguish “self” from “non-self” is not well understood.

Logically, the concept of self is linked to the concept of biological self-identity. Organisms from bacteria to mammals possess different recognition systems to detect foreign invaders. This ability has evolved both as the function of immune system, which defends against attack of foreign organisms, and as consciousness of oneself as an individual, one of the most important functions of the brain that enables social life (Lopez-Larrea, 2012). Thus each of such system could have three parts to it that may include: 1) Scanning for wide variety of invaders. 2) Generating memory and 3) Selection to avoid accidental recognition

of self otherwise it could lead to deleterious effects. I will mention some of these systems in the examples below.

Adaptive immunity in animals

A classic example of non-self-recognition is the adaptive immune system, which is based on the immunological memory of B and T cells and is well studied in vertebrates. The adaptive immune system first arose approximately 500 million years ago in jawed fish (Flajnik and Kasahara, 2010). While jawed fish have almost all of the genes required for adaptive immunity, jawless fish have none (Pancer et al., 2004). The adaptive immune system in vertebrates can recognize a pathogen as foreign, eliminate it, and create an immunological memory that leads to an enhanced response to subsequent encounters with that pathogen. Identifying and eliminating cells that could be autoreactive is an essential feature of the adaptive immune system that avoids targeting “self.” If this system is impaired, it could lead to autoimmune diseases such as Psoriasis, Rheumatoid arthritis etc.

CRISPR defense in bacteria and archea

Another example of non-self-recognition is the sequence-directed genetic interference pathway in prokaryotes. In bacteria and archea, Clustered, Regularly Interspaced Short Palindromic Repeat (CRISPR) interference is an RNA-directed adaptive immune system protecting cells against foreign genetic elements such as plasmids and phages (Barrangou et al., 2007; Marraffini and Sontheimer, 2008, 2010). Short sequence tags of invading DNA are incorporated into the

CRISPR locus, which provides an adaptive, heritable record of past infections. During re-infection, these sequence tags are transcribed into a long precursor RNAs and subsequently processed into small RNAs by Cas proteins. The base-pairing potential of these small RNAs is used as guides for the Cas ribonuleoprotein effector complex for recognition and destruction of invasive nucleic acid (Marraffini and Sontheimer, 2010).

RNAi as a defense mechanism

In 1998, the discovery of RNA interference (RNAi) by Fire and Mello opened a new chapter of exploration for the sequence-directed immunity mechanism in organisms (Fire et al., 1998). RNAi involves the generation of anti-sense small RNAs that associate with Argonaute (AGO) proteins. These small RNAs guide AGOs to target messenger RNA (mRNAs) via base pair complementarity and induce their silencing by post-transcriptional repression. A number of AGO-interacting small RNA species have since been identified. They are classified according to their size, the proteins involved in their biogenesis and their mode of regulation (Ghildiyal and Zamore, 2009). RNAi is well conserved between plants, fungi, prokaryotes, and animals (Matzke et al., 2001; Waterhouse et al., 2001) and has been thoroughly studied as a genome surveillance system. Other systems analogous to RNAi include post-transcriptional gene silencing in plants (de Carvalho et al., 1992), co-suppression (Napoli et al., 1990), and quelling (Romano and Macino, 1992).

In *Caenorhabditis elegans* (*C. elegans*) RNAi pathway has been classified as an exogenous and endogenous RNAi pathway. The exogenous RNAi has been widely used as a tool to knock down genes to study their functions in various organisms. Efficient knock down in worms can be achieved by injection of double-stranded RNA (dsRNA) into the body cavity (Fire et al., 1998), by soaking worms in a high concentration of dsRNA growth medium (Tabara et al., 1998), by feeding them on *Escherichia coli* producing the desired dsRNA (Timmons et al., 2003), or by expressing inverted repeat transgenes (Tavernarakis et al., 2000). RNAi is systemic: the silencing agent can spread from the site of injection or from the intestine (Hinas et al., 2012) throughout the worm—including somatic and germline tissues—to silence target genes (Fire et al., 1998; Timmons et al., 2003). Neurons appear to be the exception, as RNAi targeting neuronal genes works poorly by feeding or injecting dsRNA (Timmons et al., 2003). Nevertheless neurons do support cell-autonomous RNAi when the dsRNA trigger is produced directly in the cells (Tavernarakis et al., 2000). These studies suggested that neurons lack the machinery to efficiently uptake dsRNA.

Exogenous RNAi effects can be inherited in the progeny (Fire et al., 1998) and can also lead to long term indefinite transcriptional silencing even in the absence of original trigger (Vastenhouw et al., 2006).

C. elegans as a model system for self and non-self distinction

C. elegans is an excellent model system to address how an organism differentiate between self and non-self nucleic acids. Key advantages of this

system include: the facility of microinjection and genetic transformation, and the ability of organism to elicit non-self-mediated gene silencing responses against transgenes.

Extrachromosomal array silencing in *C. elegans*

Judith Kimble initiated DNA transformation in *C. elegans* in 1982 by developing a procedure for microinjection into the adult gonad (Kimble et al., 1982). She injected suppressor tRNA and showed the suppression of an amber mutation in *C. elegans* (Kimble et al., 1982). Stinchcomb later showed that injected DNA plasmids form tandem extrachromosomal arrays containing several hundred copies of the plasmid. These arrays can be transmitted to progeny of the injected animal (Stinchcomb et al., 1985). Andrew Fire improved DNA transformation and demonstrated that integration of arrays could be achieved reproducibly by microinjection of DNA into the maturing oocyte (Fire, 1986).

In 1991, Mello et al. demonstrated the first efficient germline transformation of *C. elegans* by microinjection using a cocktail of plasmids, including a plasmid with a dominant allele of (ROLler) *rol-6* (Mello et al., 1991). Having single-stranded oligonucleotides in the injection mixture led to rare but reproducible integration of the co-injected DNA. These extrachromosomal arrays, can be stably integrated into worm genome using gamma or UV irradiation to stabilize their expression (Mello and Fire, 1995).

Germline expression of extrachromosomal DNA can also be achieved by increasing the complexity of the DNA mixture—e.g., by co-injecting a plasmid

with fragmented worm genomic DNA to form a complex array (Kelly et al., 1997). However these arrays are not always expressed in the germline. Consistent with studies in various organisms where tandem repeats lead to heterochromatin-mediated silencing, silencing of simple arrays in *C. elegans* germline requires heterochromatin factors and polycomb-related transcriptional repressors encoded by Maternal Effect Sterile (*mes*) genes (Holdeman et al., 1998; Kelly and Fire, 1998; Kelly et al., 2002; Korf et al., 1998). The *C. elegans* HP1 homolog, HP1 like heterochromatin protein (*hpl-2*), has also been shown to be involved in germline transgene silencing (Kelly et al., 2002). Thus, silencing of repetitive transgenic arrays in the germline is stable and epigenetically heritable due to change in chromatin state. This type of silencing has also been shown to be temperature dependent with stronger silencing observed at lower temperature (16^{o}C and 20^{o}C) compared to 25^{o}C (Kelly et al., 1997; Strome et al., 2001). Genetic analysis of the RNAi pathway in *C. elegans* indicated that RNAi *DE*fective (rde) *rde-2*, *rde-3* and (MUTator) *mut-7* were required for both transposon and transgene silencing. This suggested a possible connection between RNAi pathway and transgene silencing (Ketting et al., 1999; Tabara et al., 1999).

Another phenomenon related to repetitive array silencing is co-suppression whereby the introduction of high-copy transgenes in *C. elegans* can induce co-suppression of endogenous homologous genes (Dernburg et al., 2000; Ketting and Plasterk, 2000). Co-suppression in *C. elegans* has been described

for germline genes and has been shown to be dependent on *mut-7* and *rde-2* (Ketting and Plasterk, 2000), which also function in the exogenous RNAi pathway. Interestingly, although *rde-1* and *rde-4* are required for RNAi they are dispensable for transposon silencing and co-suppression (Dernburg et al., 2000; Ketting and Plasterk, 2000). This suggests that co-suppression and RNAi have overlapping but distinct genetic requirements.

Soma versus germline silencing of transgene

It is interesting to note that genes present in simple (i.e., repetitive) arrays can be ubiquitously expressed at high levels in the soma, while being stably silent in the germline. Thus, simple arrays appear to be regulated by different mechanisms in soma and the germline. In soma, silencing of simple arrays can be triggered by mutations in the ADAR-encoding genes (Adenosine deaminases that act on RNA), *adr-1* and *adr-2.* Evidence suggests that ADAR proteins modify dsRNAs produced from these simple arrays suppressing their ability to feed into the RNAi pathway (Knight and Bass, 2002; Ohta et al., 2008). Somatic silencing that is triggered by dsRNAi requires the RNAi pathway genes such as *rde-1, rde-4,* and RNA-dependent RNA polymerase Family (*rrf-1*) as well as *hpl-2* and Argonaute (plant)-Like Gene (*alg-1*). Somatic silencing is not heritable (Grishok et al., 2005). Another pathway that functions in silencing of transgenes in soma and the germline is the Enhanced RNAi (ERI) endogenous small RNA pathway (Duchaine et al., 2006; Kim et al., 2005; Simmer et al., 2002). In the *eri* mutants, RNAi effect is stronger than in the wild type worms. ERI pathway involves the

upstream AGO Endogenous-RNAi deficient arGOnaute (ERGO-1) that interacts with 26G-RNA in embryos (Conine et al., 2010; Duchaine et al., 2006; Vasale et al., 2010). The 26G-RNAs have bias to begin with 5' guanosine monophosphate and are predominantly 26 nts long (Duchaine et al., 2006; Kim et al., 2005; Simmer et al., 2002). A number of proteins like DiCer Related (DCR-1) and Worm ArGOnaute proteins (WAGOs) are shared by the ERI and by the exogenous RNAi pathway (Conine et al., 2010; Duchaine et al., 2006; Vasale et al., 2010). The enhanced RNAi phenotype associated with ERI mutants is thought to be a result of competition for the downstream machinery in both RNAi pathways. In these mutants the loss of the endogenous pathway frees more of the WAGO machinery to process exogenous RNAi pathway and leads to ERI phenotype (Conine et al., 2010; Duchaine et al., 2006; Kennedy et al., 2004; Kim et al., 2005; Vasale et al., 2010). Similar to germline transgene silencing (Duchaine et al., 2006; Yigit et al., 2006), somatic silencing also depends MES-4 protein (Kim et al., 2005). Since transgenes are expressed in soma but silenced in the germline, the possibility arises, that germline cells actively prevent expression of somatic genes by a pathway that is analogous to those involved in silencing of repetitive arrays.

Silencing of multicopy and single copy transgenes

Some aspects of transgene silencing in *C. elegans* can be circumvented by newer methods—e.g., microparticle bombardment and Mos1-mediated Single Copy Insertion (MosSCI)—that result in low copy number insertion of DNA

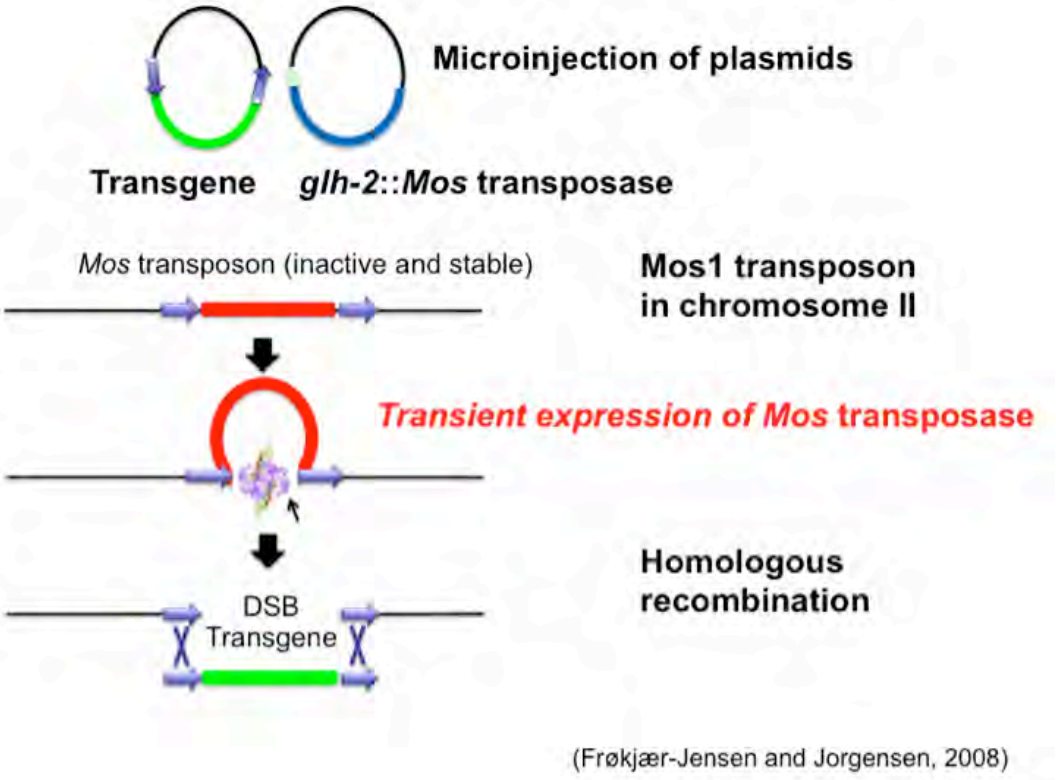
(Frokjaer-Jensen et al., 2008). During microparticle bombardment, DNA bound to the gold particle is shot into worms (Jackstadt et al., 1999; Praitis, 2006). Microparticle bombardment results in immediate integration of a low copy number transgene, but there is no control over the integration site. Furthermore the expression of a given transgene can vary depending on the site of integration and makes it difficult to study silencing mechanisms.

The MosSCI method overcomes the problems of variability in integration and expression of the transgenes. Briefly in this system, an extra-chromosomal array containing the transgene serves as a template for homologous recombination mediated repair of a double-strand DNA break. The breaks are induced by excision of a MOS1 transposon from a defined location in the genome, resulting in integration of the transgene in that specific site (Frokjaer-Jensen et al., 2008) (Figure 1.1). Using the MosSCI system, we found that the single-copy transgenes were stochastically silenced in the germline (see Chapter II). Transgenes generated by MosSCI were stably expressed or stably silenced for generations. When we crossed worms bearing expressed GFP to silent GFP, we found that silencing signal was trans-dominant, and the silent state was 100% penetrant in the subsequent generations. Further analysis of small RNAs from the silent and expressed transgenic strains revealed that small RNAs mediate silencing of the single-copy transgene, and we named this phenomenon as RNA-mediated epigenetic silencing (RNAe).

Figure 1.1. Single copy transgene insertion in *C. elegans* (MosSCI)

Schematic overview of MosSCI. A Mos1 element at a non-coding locus can be excised by transient transposase expression, resulting in a double-strand break in the chromosome. The break can then be repaired by synthesis-dependent strand annealing by homologous recombination resulting in insertion of the transgene at the desired locus.

Figure 1.1. Single copy transgene insertion in *C. elegans* (MosSCI)



Small RNA pathways in *C. elegans*

Beside the exogenous and endogenous pathways, small RNAs in *C. elegans* can be classified based on the length of small RNAs, mechanism of their biogenesis, and the AGO protein with which small RNAs associate. AGO proteins play a central role in RNA silencing process. *C. elegans* genome encodes 27 different *ago* genes, which associate with different classes of small non-coding RNAs including microRNAs (miRNAs), small interfering RNAs (siRNAs) and Piwi-interacting RNAs (piRNAs). Based on the AGO-small RNA association, they have been categorized into several pathways. These pathways can be further divided into two steps: the primary AGO targeting step and the secondary AGO maintenance step. For example, during exogenous RNAi pathway, the dsRNA trigger results in ribonuclease protein DCR-1 to processes dsRNA into siRNAs that are loaded onto the primary AGO protein RDE-1. Targeting of the cognate mRNA by RDE-1 leads to recruitment of an RNA directed RNA Polymerase (RdRPs), which synthesize secondary siRNAs—known as 22G-RNAs.

The secondary RNAi pathway: WAGO 22G RNA-mediated transcriptional and post-transcriptional gene silencing

The above-mentioned 22G RNAs are in turn loaded onto redundant group of secondary WAGOs (Gu et al., 2009; Pak and Fire, 2007; Shirayama et al., 2012; Sijen et al., 2001; Sijen et al., 2007; Yigit et al., 2006). WAGOs appear to lack the catalytic triad (DDH) required for the slicer activity of AGOs, but

nevertheless they function in various endogenous small RNA pathways to silence transposons, cryptic and aberrant genes, and foreign transgenes (Ashe et al., 2012; Gu et al., 2009; Guang et al., 2008; Shirayama et al., 2012; Yigit et al., 2006). WAGOs are required for both transcriptional silencing in the nucleus and posttranscriptional silencing in the cytoplasm (Guang et al., 2010; Guang et al., 2008; Shirayama et al., 2012). MUT-7 (3'-5'exonuclease) (Ketting et al., 1999) and RDE-3 (β nucleotidyltransferase) are required for WAGO-dependent silencing (Chen et al., 2005a) along with number of other factors.

Two of these WAGOs, WAGO-12/NRDE-3 and WAGO-9/HRDE-1, mediate somatic and germline nuclear RNAi which is also referred as Nuclear RNAi-DEficient NRDE-3 pathway (Guang et al., 2008). Like other WAGOs, WAGO-9 and -12 lack the catalytic triad (Yigit et al., 2006), but they are unique in that they contain a bipartite nuclear localization signal, and predominantly localizes to the nucleus (Ashe et al., 2012; Buckley et al., 2012; Guang et al., 2008). Consistent with their nuclear localization, these AGOs are involved in transcriptional gene silencing. WAGO 22G RNA binding triggers NRDE-3 and HRDE-1 to enter the nucleus and associate with nascent pre-mRNA targets. They then recruit NRDE-2, NRDE-1 and NRDE-4 to inhibit Pol II elongation and deposit the repressive H3K9me3 chromatin mark (Buckley et al., 2012; Burkhart et al., 2011; Burton et al., 2011; Gu et al., 2012a; Guang et al., 2010; Guang et al., 2008; Luteijn et al., 2012; Shirayama et al., 2012). Interestingly, GFP fused NRDE-3 is expressed in most of the somatic cells in 80-cell embryos (Guang et al., 2008) where as GFP

fused HRDE-1 is expressed in the male and female germ cells (Buckley et al., 2012; Shirayama et al., 2012). Both WAGOs appear to use common silencing mechanism and machinery (Ashe et al., 2012; Buckley et al., 2012; Burkhart et al., 2011; Luteijn et al., 2012). Thus NRDE-3 mediates somatic nuclear RNAi and HRDE-1 mediates germline nuclear RNAi.

Genome surveillance: PIWI and the piRNA pathway

P-element Induced WImpy testis (PIWI) protein was originally identified in a screen for factors involved in Germ line Stem Cell maintenance in *Drosophila melanogaster* (*D. melanogaster*) and is required for germline proliferation and fertility (Cox et al., 1998; Juliano et al., 2011). PIWI proteins are evolutionary conserved in animals and are most essential for fertility and transposon silencing (Cox et al., 1998; Lin and Spradling, 1997).

Piwi proteins associate with a class of small non-coding RNAs called piRNAs (Aravin et al., 2007; Brennecke et al., 2007; Das et al., 2008; Lin, 2007; Ruby et al., 2006). piRNAs comprise the largest class of small non-coding RNAs. They are 21-35-nt long and have bias towards 5’ Uridine in most species. piRNA biogenesis is independent of DICER protein thus making them distinct from siRNAs and miRNAs (Batista et al., 2008; Das et al., 2008). In most species, piRNAs are derived from large genomic clusters, ranging from one kilobases (kbs) to hundreds of kbs in size. Such clusters may encode hundred to thousands of piRNAs (O'Donnell and Boeke, 2007). Although the biogenesis of

piRNA is not fully understood, several mechanisms had been proposed in various model organisms.

In *D. melanogaster* piRNA production involves the primary pathway and the secondary ping-pong pathways. In the primary pathway, piRNAs are transcribed from genomic regions called piRNA clusters, processed, and loaded onto Piwi or AUBERGINE (AUB). Silencing of transposons by these piRNA takes place both in cytoplasm and in the nucleus. Together with AGO-3, the AUB-piRNA complex serves as a trigger to start the ping-pong amplification cycle to produce secondary piRNAs (Brennecke et al., 2007; Gunawardane et al., 2007; O'Donnell and Boeke, 2007). The ping pong cycle silences the target transposon sequences and amplifies the piRNA sequence at the same time for silencing (Iwasaki et al., 2015). The ping-pong amplification in flies is analogous to the activity of RdRPs in worms and plants, since RdRPs are absent from the genome of flies and mammals (Brennecke et al., 2007). Recent work has uncovered an interesting interconnection between ping-pong amplification and the production of phased, primary piRNAs that are predominantly loaded into PIWI protein(Han et al., 2015; Wang et al., 2015). Phasing generates small RNA molecules further down the targeted transcript from the original targeting site. This mechanism allows the targeting of diverse sequences that lie in proximity to the original threat (Czech and Hannon, 2016).

In mice, piRNA are categorized as prepachytene and pachytene piRNAs depending on their expression during spermatogenesis (Aravin et al., 2006;

Girard et al., 2006; Grivna et al., 2006). Mice have three PIWI AGOs MIWI, MILI and MIWI2. Prepachytene piRNAs are derived mostly from transposable elements and are associated with MILI and MIWI2. Where as pachytene piRNAs comprises 95% of the total piRNAs and are derived from piRNA clusters located in various regions of the genome in adult mouse testis. These piRNAs bind to both MILI and MIWI AGO proteins (Iwasaki et al., 2015; Kawaoka et al., 2011; Weick and Miska, 2014). In mouse, A-MYB a transcriptional master regulator induces Polymerase II (POLII)-mediated transcription of both pachytene piRNAs and piRNA biogenesis machinery creating a feed-forward loop for the piRNA biogenesis (Aravin et al., 2006; Li et al., 2013). Pachytene piRNAs regulate their target genes by posttranscriptional gene silencing in cytoplasm (Iwasaki et al., 2015).

In *C. elegans* there are more than 30,000 piRNAs, also known as 21U-RNAs—21 nucleotides with a 5'-monophosphorylated Uridine (Batista et al., 2008; Das et al., 2008; Gu et al., 2012b; Ruby et al., 2006). Similar to mammalian pachytene piRNAs, 21U-RNAs are diverse in sequence, and majority of them lack perfect complementarity to their target RNAs (Bagijn et al., 2012; Lee et al., 2012; Ruby et al., 2006). *C. elegans* piRNAs are derived from 25- to 29-nucleotide capped small RNA precursors (Gu et al., 2012b; Ruby et al., 2006) transcribed by RNA POLII, Unlike mammalian piRNAs, which are derived from long precursor RNAs. About half of the piRNA loci—called type 1 piRNAs—map within two large clusters on Chromosome IV within intergenic regions. Type 1-

piRNA promoters have a conserved 8-nt (CTGTTTCA) A/T-rich “Ruby motif” (Gu et al., 2012b; Ruby et al., 2006). The expression of type 1-piRNAs appears to require Forkhead family of transcription factors (Cecere et al., 2012), the nuclear factor PRDE-1 (Weick et al., 2014) and the Myb-like DNA-binding protein SNPC-4 (Goh et al., 2014; Kasper et al., 2014) (Unpublished data from our lab). Binding of SNPC-4 to type 1-piRNA loci appears to require PRDE-1, which is required for transcription of type 1 precursors (Weick et al., 2014). The remaining piRNA loci—called type 2 piRNAs—map to transcription start sites of all other genes transcribed by RNA POL II and lack the 8-nt Ruby motif. With a few exceptions, type 2 piRNAs are much less abundant than type 1 piRNAs (Gu et al., 2012b).

Functional 21U piRNAs are generated by processing of the precursor piRNA which requires removal of the 5’ cap and the first 2 nucleotides by an unknown mechanism (Gu et al., 2012b) and 3’ trimming by the exoribonuclease PARN-1 (Tang et al., 2016). Using RNAi-based screening, the Hannon lab identified several factors required for *C. elegans* piRNA biogenesis. They named these factors as TOFUs, for twenty-one U fouled up, and suggested where they might function in piRNA biogenesis (Goh et al., 2014).

In worms, PIWI Argonaute PRG-1 associates with the piRNAs and has been linked primarily to the silencing of only one transposon family, Tc3 (Das et al., 2008). Interestingly, PRG-1 appears to target transposon by piRNAs and recruit RdRP and WAGOs/22G machinery to maintain Tc3 silencing. The exogenous RNAi share the downstream factors required for transposon silencing

although the initiation of silencing in each case requires a different trigger. An intriguing fact about the transposon silencing is that all other transposons do not require PRG-1 for silencing except Tc3. One explanation for this observation could be that other transposons except Tc3 are no longer dependent on PRG-1 for their silencing. It is very likely that millions of years ago, when these transposons were first acquired, they were initially subjected to piRNAs mediated PRG-1 silencing, as these sequences were foreign to worm genome. However during the course of time, these transposons were able to enter the WAGO dependent silencing where they no longer require PRG-1 targeting by piRNAs and hence became independent of PRG-1 mediated silencing. This might be very similar the maintenance phase of RNAe that we have discovered in case of transgene silencing which is no longer dependent on PRG-1.

The majority of piRNAs in worms maps uniquely to the genome and lack obvious targets including transposon (Batista et al., 2008; Das et al., 2008). As such, their function remains entirely unknown. However bioinformatics analysis from our group and Miska group revealed that piRNAs in *C. elegans* do not require a perfect match but base pair imperfectly to target mRNAs (Bagijn et al., 2012; Batista et al., 2008; Das et al., 2008; Lee et al., 2012). It has also been shown that the catalytic activity of PRG-1 is not required for piRNA-induced silencing (Bagijn et al., 2012; Lee et al., 2012).

piRNA-mediated transcriptional transgene silencing also requires NRDE-1, -2, and -4, suggesting that there are several downstream factors shared by both

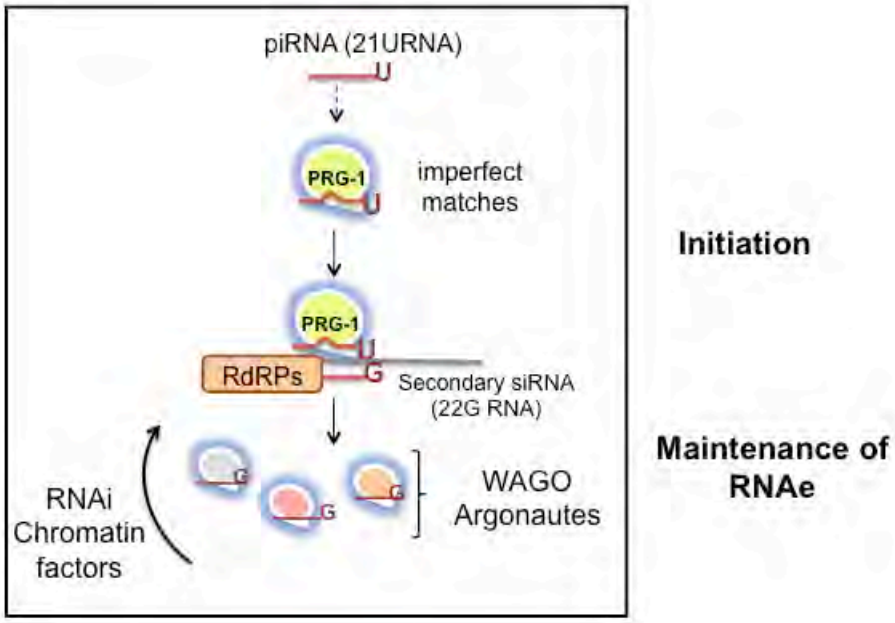
somatic and germline transgene silencing. Where as piRNAs play important role as a defense mechanism to silence transposons in flies and worms, it has different function in ciliates. For example in *Tetrahymena*, a set of PIWI protein-bound RNAs known as scanRNAs target heterochromatin modification to mark genome for elimination very similar to piRNA mediated silencing of metazoan (Liu et al., 2004b), where as in *Qxytricha trifallax* piRNAs serve as the opposite function of promoting retention of maternal genomic regions thus protecting the self sequences, while the non-self sequence are eliminated (Chalker and Yao, 2011; Fang et al., 2012; Mochizuki et al., 2002). Thus in one case piRNAs serve to eliminate non-self sequences while in other case piRNAs serve to protect the self-sequences.

Chapter II of my thesis describes how RNAe is initiated and maintained. Initiation of RNAe requires the primary AGO PRG-1 and maintenance is carried out by nuclear and cytoplasmic WAGOs. Thus WAGOs guide the memory of non-self gene expression. Using a candidate approach, we identified factors required for maintenance of epigenetic silencing, including genes previously shown to be required for silencing of extra-chromosomal arrays (Vastenhouw et al., 2006). Furthermore, methylation of histone H3 on lysine 9 (H3K9me3), the histone methylation associated with heterochromatin protein, was also enriched at the silent transgenes. Together our data support the model where WAGOs and their 22G-RNA cofactors maintain the heterochromatin state on epigenetically silenced transgenes (Figure 1.2).

Figure 1.2. Initiation and maintenance of RNAe pathway

The schematic figure shows that RNAe can be divided into two distinct stages. For initiation PRG-1 along with piRNAs scans using imperfect base pairing for target RNAs. On targeting the foreign RNA sequence, PRG-1 recruits RdRP to produce secondary anti-sense 22G RNAs that are loaded on to WAGOs. WAGOs maintain silencing and establish a memory of non-self. This memory is also mediated by chromatin factors, Polycomb, HP1 etc. This pathway is referred as RNA-mediated epigenetic silencing (RNAe).

Figure 1.2. Initiation and maintenance of RNAe pathway



Guardian of genome: CSR-1 and associated 22Gs

The Chromosome Segregation and RNAi deficient (CSR-1) derives its name because of its essential function in both Chromosome Segregation and RNAi (Yigit et al., 2006). Different from other AGOs, CSR-1 is an essential Argonaute required for fertility and development. Hermaphrodites lacking CSR-1 peotein exhibit sterile phenotype and the RNAi of *csr-1* results in embryonic lethal phenotype. The embryos also show defects in organization of chromosome at metaphase of each early embryonic cycle and the formation of anaphase DNA bridges (Yigit et al., 2006). Julie et al further characterized CSR-1 in 2009 and showed that CSR-1 is expressed in all developmental stages and localizes to the perinuclear P-granules in the germline (Claycomb et al., 2009). CSR-1 has two isoforms: a longer isoform CSR-1a and the shorter isoform CSR-1b. Interestingly, hermaphrodite lack the longer isoform where as males exhibit both isoforms (Ortiz et al., 2014).

CSR-1 also binds to a class of endogenous derived small RNAs that are 22 nts in length, has a 5' triphosphate and starts preferentially with a 5' Guanosine (Ambros et al., 2003; Claycomb et al., 2009; Gu et al., 2009). CSR-1 interacting 22G small RNAs are antisense to thousands of germline expressing genes. CSR-1 has catalytic activity (DDH) and *in vitro* CSR-1 has been shown to cleave complementary target mRNA when loaded with triphosphorylated 22G small RNAs (Aoki et al., 2007). This activity is lost in a D769A slider-dead mutant (Aoki et al., 2007), suggesting its role in the cleavage of mRNAs. Earlier studies on

CSR-1 have suggested that it does not seem to down regulate its target mRNA or protein expression (Claycomb et al., 2009). In contrast, Desai lab has recently demonstrated that CSR-1 and its slicer activity plays important role in down regulating the levels of maternally deposited mRNAs to fine-tune the expression of proteins with critical roles in embryonic cell division (Gerson-Gurwitz et al., 2016). However the down regulation of target genes were applicable only for a subset of 133 genes among the 3000 CSR-1 target genes (Gerson-Gurwitz et al., 2016).

Another fascinating role of CSR-1 associated small RNAs in germline is the activation of transgenes (Seth et al., 2013; Wedeles et al., 2013b). In chapter III of my thesis we provide evidence that in addition to an adaptive memory of silenced sequences, *C. elegans* can also develop an opposing adaptive memory of expressed/self-mRNAs (Figure 1.3). We named this mechanism that can prevent or reverse RNAe, as RNA-induced epigenetic gene activation (RNAa). We show that CSR-1 is required for RNAa and *C. elegans* can adaptively acquire and maintains a transgenerational memory that recognizes and protect self-mRNAs from piRNA induced silencing. This model was put forward in lie of the hypothesis that piRNA target by allowing mismatches (Figure 1.2). If two to three mismatches are allowed piRNA could target any foreign sequences (Bagijn et al., 2012; Lee et al., 2012) including the endogenous germline expressed mRNAs. Therefore piRNA surveillance system involves that self-mRNAs be protected from piRNA induced silencing (Seth et al., 2013; Wedeles et al., 2013a, b). This

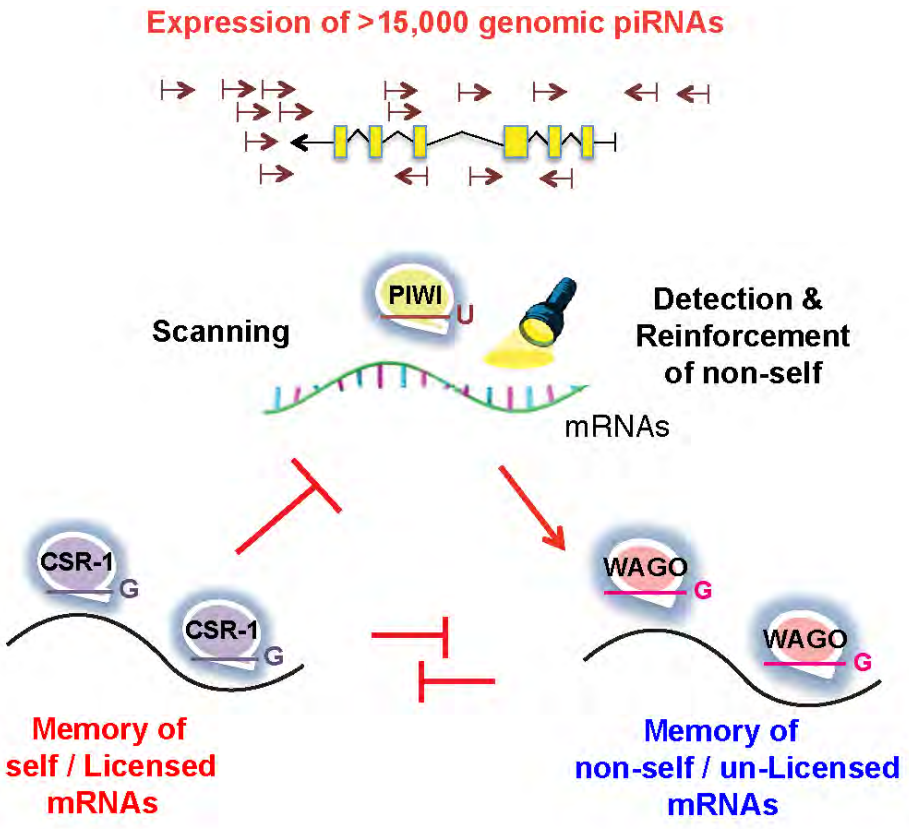
leads into the CSR-1 protection model where CSR-1 associated 22G small RNAs serves as a molecular marker of ‘self” and counteracts silencing by other small RNA pathways, including the piRNA pathway (Lee et al., 2012; Seth et al., 2013; Shirayama et al., 2012). In support of this model, Wedeles et al showed that tethering CSR-1 to a previously silent RNAe transgene is sufficient to activate expression at this transgene thus inducing the active chromatin modification by CSR-1 (Wedeles et al., 2013b). In addition, CSR-1 has been shown to positively regulate the expression of germline genes on a genome wide scale (Cecere et al., 2014). It is also particularly important during sperm development for promoting the expression of genes involved in sperm differentiation downstream of the ALG-3,-4 26G-RNA pathway (Conine et al., 2013) .

CSR-1 targets are highly enriched for histone modification associated with the active chromatin mark including mono-,di-,and tri-methylation at H3 lysine 4, and acetylation at histone H3 lysine 9, H4 lysine 8, and H4 lysine 16 suggestive of its role in maintenance of germline gene expression in the euchromatin region (Youngman and Claycomb, 2014).

Figure 1.3. Model: An RNA-mediated binary switch

PIWI Argonaute PRG-1 uses piRNAs for transcriptome-wide surveillance of germline transcripts. piRNAs can allow imperfect base-pairing to initiate silencing. Silencing is maintained by downstream WAGOs associated small RNAs that carries the memory of non-self transcripts whereas mRNAs targeted by CSR-1 Argonaute appears to protect transcript from silencing and carries the memory of self-transcripts.

Figure 1.3. Model: An RNA-mediated binary switch



Epigenetics as a regulator of self/non-self recognition

The self-identity of an individual can be defined by the specific genetic information carried in its own DNA. This information can be further modified epigenetically. Epigenetics is the study of heritable changes in gene expression (active versus inactive genes) that does not involve change in DNA sequence and can also be described as changes in phenotype without any change in genotype. Dr. Conrad Hal Waddington in 1942 first coined the term “epigenotype” that describes how genotype gives rise to phenotype during development (Bird, 2007; Waddington, 2012). Further studies in this area led to redefining epigenetics as the study of mitotically and meiotically heritable changes in gene expression that are not encoded in the DNA itself (Bird, 2007; E.A. Russo, 1996). Sustainable epigenetic inheritance involves changes in three systems including DNA methylation, histone modification, and non-coding RNA associated gene silencing (Egger et al., 2004).

Non-coding RNAs have been known to function as a major epigenetic modulator. There are two major groups of non-coding RNAs: small non-coding RNAs that are less than 30 nucleotides, e.g., these include microRNAs, siRNAs, piRNAs and long non-coding RNAs greater than 200 nucleotides. Small non-coding RNAs associate with AGO proteins and mediates transcriptional gene silencing, where as long non-coding RNAs bind to the chromatin modifying enzymes, recruit them to a specific site in the genome, thereby modulating chromatin status and regulating gene expression (Mercer and Mattick, 2013).

Both groups of non-coding RNAs play important roles in heterochromatin silencing, histone modification, DNA methylation and gene silencing.

In Summary in chapter II of my thesis, I will present the finding from a project led by Dr Masaki Shirayama, senior member in our lab. I worked closely on this project with Dr Shirayama and set the paradigm for self/non-self recognition in *C. elegans* germline using transgene as a tool. In this chapter we describe the phenomenon of RNA-mediated epigenetic silencing (RNAe). Mechanistically we show that PIWI AGO PRG-1 and its genomically encoded piRNAs cofactors initiate RNAe and maintenance of RNAe requires chromatin factors; nuclear and cytoplasmic WAGOs. We also set the stage for an opposing phenomenon of transgene activation in addition to transgene silencing in *C. elegans* germline.

In Chapter III of my thesis, we further provide insight into the phenomenon of RNA-mediated epigenetic gene activation (RNAa). We show that AGO CSR-1 and its associated small RNAs are required for this phenomenon. We show that in addition to the adaptive memory of non-self transcript, *C. elegans* can also develop an adaptive memory of self-transcripts. These memories of gene-expression are carried by small RNAs through sperm and oocyte to next generation.

In chapter IV of my thesis, we provide evidence for a competition between PRG-1 and CSR-1 Argonaute pathways. Eliminating PRG-1 or engineering new piRNA can shift the balance towards RNAa or RNAe. We also show that targeting by novel piRNAs triggers robust local synthesis of secondary siRNAs by RdRP and partial down-regulation of the target mRNA and protein levels, but

failed to trigger epigenetic silencing of endogenous germline targets. Thus worms have evolved a remarkable strategy in which multiple AGOs function together to generate and maintain an epigenetic memory of self and non-self gene expression.

Preface to Chapter II

This chapter describes the discovery of novel phenomenon of RNA-mediated epigenetic silencing triggered by piRNA in recognition to non-self RNAs.

My contributions to this chapter are: Figures 2.1A, D and E, Table 2.1, 2.3A-F, 2.4, 2.5A, 2.6 B

Masaki Shirayama contributed: Figures 2.1A-C, 2.2A-D, Table 2.1, 2.3G, I, J, 2.4, 2.6A

Heng-Chi Lee contributed: Figure 2.4

Weifeng Gu contributed: Figure 2.3H, J and 2.6A

Takao Ishidate contributed: Figure 2.5B

Darryl Conte Jr contributed: Table 2.1

Craig C. Mello contributed: Figure 2.6A and B

This chapter was published as:

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

CHAPTER II: piRNAs Initiate An Epigenetic Memory Of Non-self RNA In The *C. elegans* Germline

SUMMARY

Organisms employ a fascinating array of strategies to silence invasive nucleic acids such as transposons and viruses. Although evidence exists for several pathways that detect foreign sequences including pathways that sense copy-number, unpaired DNA, or aberrant RNA (e.g. dsRNA), in many cases the mechanisms used to distinguish "self" from "non-self" nucleic acids remain mysterious. Here we describe an RNA-induced epigenetic silencing pathway that permanently silences single copy transgenes. We show that the Piwi Argonaute PRG-1 and its genomically-encoded piRNA cofactors initiate permanent silencing, while maintenance depends on chromatin factors and the WAGO Argonaute pathway. Our findings support a model in which PRG-1 scans for foreign sequences, while two other Argonaute pathways serve as epigenetic memories of "self" and "non-self" RNAs. These findings suggest how organisms can utilize RNAi-related mechanisms to detect foreign sequences, not by any molecular signature, but by comparing the foreign sequence to a memory of previous gene expression.

INTRODUCTION

All organisms balance the need to maintain genetic variation against the danger of accumulating potentially deleterious genes or pathogenic sequences (Antonovics et al., 2011). The experimental introduction of DNA (transgenes) into the germline provides an opportunity to probe an organism’s response to foreign

DNA (Rulicke and Hubscher, 2000), and has revealed that organisms use a variety of mechanisms to silence transgenes in the germline (Birchler et al., 2003; Brodersen and Voinnet, 2006). Interestingly, some mutants that disrupt transgene silencing also de-silence endogenous genes, including self-replicating elements called transposons (Ketting et al., 1999; Tabara et al., 1999). Thus, the mechanisms involved in transgene silencing protect the genome from invasive DNA elements.

In many organisms transgene silencing has been linked to factors that are also required for the RNAi pathway (Bosher and Labouesse, 2000). RNAi was first identified as a sequence-specific response triggered by double-stranded (ds) RNA (Fire et al., 1998). During RNAi, dsRNA is processed by the RNase III-related protein, Dicer, into ~21 nucleotide (nt) short-interfering (si) RNAs (Bernstein et al., 2001; Carmell and Hannon, 2004; Zamore et al., 2000), which are loaded onto Argonaute (AGO) proteins to form the key effectors of RNA-induced silencing complexes (Hammond et al., 2001; Liu et al., 2004a; Meister et al., 2004). AGOs are RNase H-related proteins that use the base-pairing potential of small RNA cofactors to guide sequence-specific binding to target sequences (Song et al., 2004). In some cases, AGOs directly cleave their targets; in other cases, AGOs recruit co-factors that direct mRNA destruction or other modes of regulation.

Despite a clear overlap between the mechanisms that mediate RNAi and the silencing of transposons and transgenes, several findings point to distinct

triggering mechanisms. For example, the AGO protein RDE-1 is essential for the dsRNA response in *C. elegans*, but is not required for transposon or transgene silencing (Tabara et al., 1999). RDE-1 engages siRNAs produced by Dicer and mediates the initial search for target RNAs in the cell (Parrish and Fire, 2001; Yigit et al., 2006). RDE-1 is thought to recruit a cellular RNA-dependent RNA polymerase (RdRP), which then utilizes the target mRNA as a template for the production of secondary siRNAs, termed 22G-RNAs (Gu et al., 2009; Pak and Fire, 2007; Sijen et al., 2001; Sijen et al., 2007; Yigit et al., 2006). The 22G-RNAs are loaded onto members of an expanded, partially redundant, group of Worm-specific AGOs (WAGOs). WAGOs that localize to the cytoplasm are thought to mediate mRNA turnover, whereas WAGOs that localize to the nucleus mediate transcriptional silencing (Gu et al., 2009; Guang et al., 2008). Many components of the RNAi pathway that function downstream of RDE-1 are required for transposon and transgene silencing, including the RdRP system (Gu et al., 2009; Smardon et al., 2000), the polynucleotide polymerase RDE-3 (Chen et al., 2005b), the nuclease MUT-7 (Ketting et al., 1999), and the WAGO proteins (Yigit et al., 2006), among others (Robert et al., 2004). The fact that RDE-1 is not required for transposon and transgene silencing suggests that features unique to transposons and transgenes underlie the initial recruitment of RdRP to these targets and that dsRNA is unlikely to be the trigger.

In the germline, RdRPs not only produces 22G-RNAs that interact with WAGOs, but also produce 22G-RNAs that interact with a distinct AGO, CSR-1,

required for fertility and chromosome segregation (Claycomb et al., 2009; Yigit et al., 2006). However some factors, including RDE-3 and MUT-7, are only required for WAGO 22G-RNA accumulation (Gu et al., 2009), indicating that the CSR-1 and WAGO 22G pathways also involve distinct mechanisms. Indeed, the WAGO and CSR-1 22G pathways together target virtually all germline-expressed mRNAs, however their targets are largely non-overlapping (Gu et al., 2009). Furthermore, unlike the WAGO pathway, the CSR-1 22G pathway does not appear to silence its targets (Claycomb et al., 2009). Instead, the CSR-1 pathway may help to define and maintain euchromatic regions along the holocentric chromosomes in order to support the proper assembly of kinetochores.

In most animals, the Piwi-family AGOs are required for fertility and transposon silencing (Cox et al., 1998; Juliano et al., 2011). In *C. elegans*, however, the Piwi-related gene product PRG-1 has only been linked to the silencing of one transposon family, Tc3 (Batista et al., 2008; Das et al., 2008). Interestingly, PRG-1 appears to recruit RdRP and the WAGO 22G pathway to maintain Tc3 silencing. Piwi-interacting (pi) RNAs (21U-RNAs in *C. elegans*) are genomically encoded and appear to be expressed as Pol II transcripts whose single-stranded products are processed and loaded onto Piwi (Aravin and Hannon, 2008; Kim et al., 2009). More than 15,000 distinct piRNA species exist in *C. elegans*, while millions of species are expressed in the testes of mammals (Aravin et al., 2006; Batista et al., 2008; Das et al., 2008; Girard et al., 2006; Grivna et al., 2006; Lau et al., 2006). The majority of these piRNAs map uniquely

to the genome and lack obvious targets. As such, their function remains entirely unknown.

Here we use a homologous gene-targeting method, called “Mos1-mediated single copy insertion” (MosSCI) (Frokjaer-Jensen et al., 2008), to show that strains bearing identical single-copy transgenes inserted at the same chromosomal site can exhibit opposite and remarkably stable epigenetic fates, either expressed or silenced. Transgenes consisting of an endogenous germline-expressed gene fused to a relatively long foreign sequence (e.g. *gfp*) were prone to silencing. By contrast, otherwise identical transgenes fused to a short foreign sequence (e.g. *flag*) were always expressed. Our genetic and molecular analyses reveal that silencing is dependent on nuclear and cytoplasmic WAGOs and is correlated with the accumulation of 22G-RNAs targeting the foreign portion of the transgene. Importantly, PRG-1 is required to initiate, but not to maintain silencing. We propose that PRG-1 and its 21U-RNA co-factors scan for foreign RNA sequences and initiate WAGO-maintained gene silencing, while endogenous mRNAs are protected from silencing, perhaps by the CSR-1/22G-RNA pathway.

RESULTS

Heritable and dominant silencing of single-copy transgenes

Single-copy insertions can overcome barriers to transgene expression in the germline (Rieckher et al., 2009). Indeed, the single-copy insertion of transgenes at a defined chromosomal locus via the recently developed MosSCI

approach reproducibly achieves germline expression (Frokjaer-Jensen et al., 2008). However, while using MosSCI, we were surprised to find that not all single-copy transgenes were expressed in the germline (Figure 2.1A-C). The failure to express was only common for transgene fusions to lengthy foreign sequences, *gfp* (Figure 2.1A); transgenes with the *flag* epitope sequences were nearly always fully expressed (Figure 2.1A). Furthermore, we observed that transgenes where *gfp* was inserted at the 5′ (rather than 3′) end of the construct were much less likely to be expressed (Figure 2.1A). PCR and sequence analyses indicated that non-expressed transgenes are structurally identical to expressed transgenes, suggesting that the former are actively silenced.

We next crossed a silent line to an expressing line to see which phenotype dominates. Strikingly, we found that 100 % of the F1 cross-progeny (n=12) and F2 self-progeny (n=24) failed to express *gfp* in the germline (Figure 2.1D). Identical results were obtained even when the silent and active alleles were inserted on separate chromosomes (Figure 2.1E), suggesting that chromosomal pairing is not required for transfer of the silent state. Although transgenes with 3′ *gfp* insertions were less prone to silencing during transgene formation, they were fully silenced when crossed to a silent line (Figure 2.3J and data not shown).

We found that either parent could contribute the dominant silencing signal. However, when the silent allele was male-derived, it took more than one generation to completely silence the active allele. For example, silencing was observed in 67 % (n=15) of F1 progeny when the silent allele was paternally

derived, while 100 % (n=12) of F1 progeny were silenced when maternally derived. Nevertheless, regardless of the parent of origin, in the F3 and subsequent generations, 100 % of the descendants were GFP negative (n>100). The silent phenotype was fully penetrant with no evidence of expression or reversion even after the formerly active allele was re-segregated as a homozygote (Figure 2.1E). These results clearly indicate that the failure to express these single-copy transgenes represents an active silencing process that involves a dominant trans-acting silencing signal. We first observed this dominant silencing activity in crosses with *gfp::csr-1*, which raised a concern since CSR-1 is an Argonaute potentially involved in silencing mechanisms. However, identical results were obtained in crosses with *cdk-1* transgenes (data not shown), indicating that there is nothing unusual about the *csr-1* transgenic lines.

We refer to this phenomenon as RNA-induced epigenetic silencing (RNAe), because the silent state is stable indefinitely (without evidence of reversion), and (as shown below) maintenance of silencing involves a small-RNA silencing signal that is epigenetically programmed (not genomically encoded). We identify transgenes exhibiting this type of silencing by including the term “(RNAe)” after the transgene name (e.g. *neSi11 gfp-1::cdk-1*(*RNAe*)). For clarity, active versions of the same alleles are referred to using (+), (e.g. *neSi11 gfp-1::cdk-1(+)*).

High-copy transgenes in *C. elegans* can induce co-suppression of endogenous homologous genes (Dernburg et al., 2000; Ketting and Plasterk, 2000). Several of the transgenes we analyzed are fusion constructs with

essential genes (e.g. *gfp*::*cdk-1*) and should result in obvious visible phenotypes if the corresponding endogenous locus was co-suppressed. However, no phenotypic evidence of co-suppression was observed in the silent lines analyzed (data not shown), suggesting that despite the dominant nature of the silencing signal, silencing does not spread to the endogenous locus. To ask if there is a partial suppression of the endogenous locus, we performed Western blot analysis to determine the relative expression of the transgene and endogenous protein products in both active and silent lines. Consistent with the lack of phenotypic evidence for co-suppression, we observed identical levels of endogenous protein expression in both the active and silent transgenic lines (Figure 2.2A).

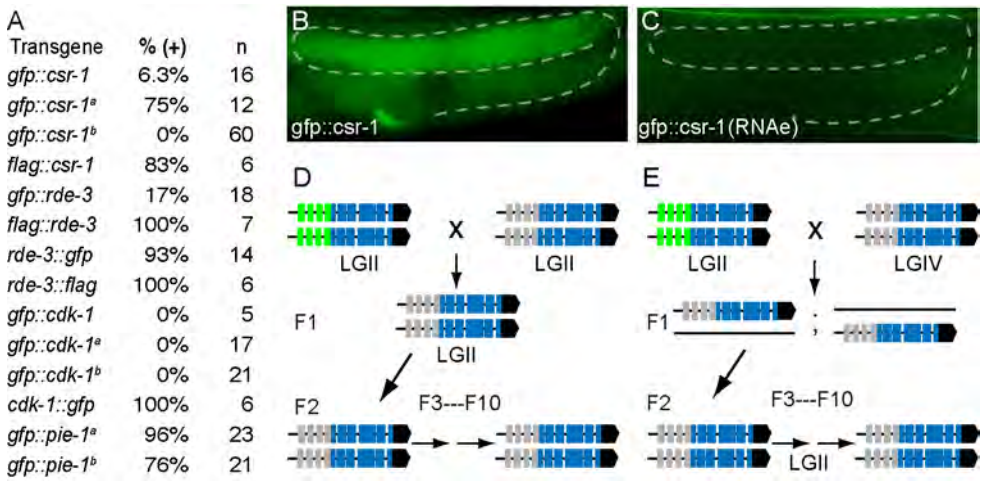
Figure 2.1. Heritable and dominant silencing of single-copy transgenes

(A) Transgenic lines created by MosSCI. MosSCI injection mixture made with 1 ng/µl (a) or 50 µg/ml (b) target plasmid for heat shock method.

(B, C) Fluorescence micrographs of adult hermaphrodite germ lines from (B) GFP positive *neSi9 gfp*::*csr-1(+)* and (C) GFP negative *neSi8 gfp*::*csr-1*(*RNAe*) transgenic lines. GFP::CSR-1 is expressed prominently in the peri-nuclear P-granules in the syncytial germ line (dashed outline) and is also visible in the cytoplasm of maturing oocytes.

(D, E) Schematic diagrams illustrating the results of genetic crosses between expressed and silenced (gray) *gfp::csr-1* transgenic lines (>100 animals scored per generation after F2). In (D) *neSi8 gfp*::*csr-1*(*RNAe*) hermaphrodites were mated with *neSi9 gfp*::*csr-1(+)* males. In (E) *neSi10 gfp*::*csr-1*(*RNAe*) hermaphrodites, integrated on chromosome IV (LGIV), were mated to *neSi9 gfp*::*csr-1(+)* males, integrated on LGII (LGII). In the F2 generation, the *neSi9 gfp*::*csr-1(+)* allele was segregated away from *neSi10* and propagated for 8 more generations.

Figure 2.1. Heritable and dominant silencing of single-copy transgenes



RNAe requires chromatin factors and correlates with H3K9me3

To ask if silencing is regulated transcriptionally or post-transcriptionally, we isolated total RNA from otherwise identical silent and active *gfp*::*csr-1* strains and measured the abundance of pre-mRNAs and mRNAs by real-time quantitative PCR (qPCR). We found that both the pre-mRNA and mRNA levels were significantly reduced in the silent line compared to the active line (Figure 2.2B and D). Moreover, although a reduction at the pre-mRNA level appeared to account for the majority of silencing, a further reduction was evident at the mRNA level, suggesting that silencing is achieved at both transcriptional and post-transcriptional levels (Figure 2.2B and 2.2D).

Previous work has shown that the methylation of lysine 9 on histone H3 (H3K9me), a histone modification associated with silent chromatin, is enriched on high-copy number transgenes in the germline (Bessler et al., 2010; Kelly et al., 2002). Furthermore, germline silencing of high-copy transgenes is dependent on a number of chromatin-associated factors, including the Polycomb-Group complex (MES-2/-3/-6), Trithorax-related (MES-4) and the heterochromatin proteins (HPL-1 and -2) (Couteau et al., 2002; Grishok et al., 2005; Kelly and Fire, 1998; Kelly et al., 1997). Consistent with these previous findings, we found that transgene sequences from a silent MosSCI allele, but not an active MosSCI allele, were enriched in Chromatin Immunoprecipitation (ChIP) experiments using antibodies specific for H3K9me3 (Figure 2.2C and 2.2D). The lysates used were from whole worms, therefore only a portion of the chromatin present in the total

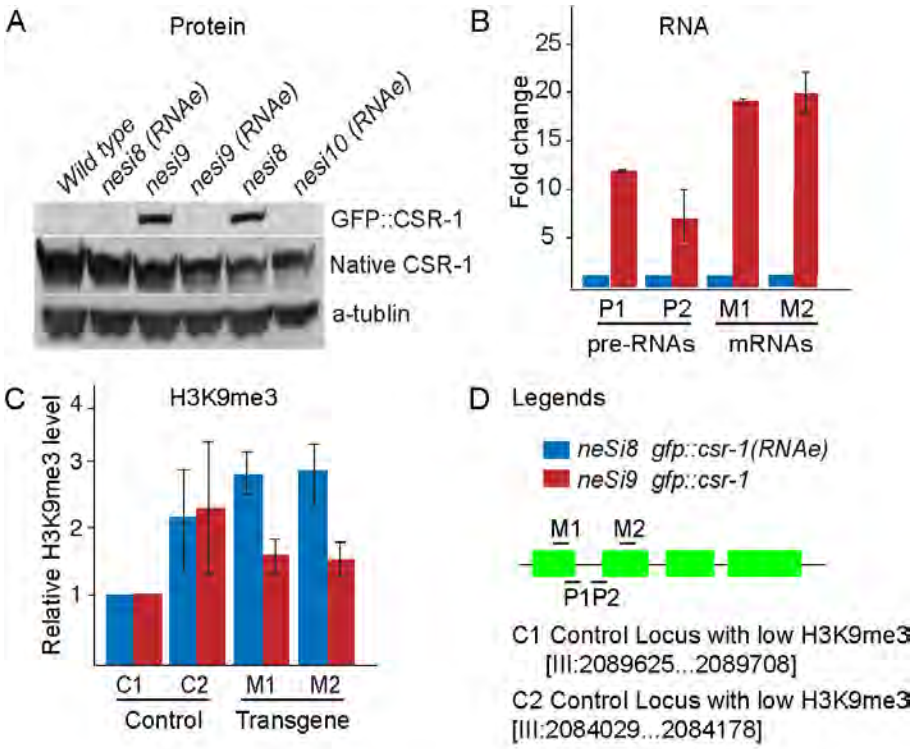
lysate corresponds to germline chromatin, perhaps accounting for the relatively weak, 2-fold enrichment observed. Finally, we found that *mes-3*, *mes-4*, and *hpl-2* mutants all de-silenced the *gfp*::*csr-1* and *gfp*::*cdk-1* transgenes (Table 2.1). These findings suggest that the maintenance of single-copy transgene silencing involves a chromatin component.

Figure 2.2. *RNAe* alleles exhibit evidence of transcriptional silencing

(A) Analysis of protein expression in wild-type and transgenic strains (as indicated). The blot was probed with anti-GFP (GFP::CSR-1), anti-CSR-1 (Native CSR-1) and anti-α-tubulin (α-tubulin) antibodies (as indicated). The *neSi9 gfp*::*csr-1*(*RNAe*) strain was generated by crossing *neSi9 gfp*::*csr-1(+)* to *neSi10 gfp*::*csr-1*(*RNAe*). The *neSi8 gfp*::*csr-1(+)* strain was generated by crossing *neSi8 gfp*::*csr-1*(*RNAe*) to *rde-3*.

(B, C) qPCR analysis of *gfp::csr-1* mRNA, pre-mRNA, and H3K9me3 levels in silent (blue) and expressed (red) transgenic lines. The strains and probes used are indicated in (D). In (B) *gfp*::*csr-1* expression was normalized to the *clp-3* mRNA. The data is shown as fold-change between the expressed and silent *gfp*::*csr-1* alleles. Error bars represent the standard deviation for two experimental replicates. In (C), error bars indicate the standard deviation for three experimental replicates.

Figure 2.2. *RNAe* alleles exhibit evidence of transcriptional silencing



Maintenance of silencing requires RNAi-related factors

The trans-acting nature of the silencing phenomenon suggested the possible involvement of an RNAi-related small RNA pathway. To explore this possibility we crossed a silent transgenic strain into strains bearing mutations in RNAi components. Two downstream factors in the exo-RNAi pathway, *rde-3* and *mut-7*, which encode a beta-nucleotidyl transferase and a 3'-5' exonuclease respectively (Chen et al., 2005b; Ketting et al., 1999), are known to be required for the maintenance of transposon silencing and have been implicated in co-suppression (Dernburg et al., 2000; Ketting and Plasterk, 2000) and high-copy number transgene silencing (Tabara et al., 1999). Consistent with the involvement of these factors in the maintenance of RNAe we found that crossing a silent transgene into these mutant strains resulted in fully restored transgene expression (Table 2.1).

We also examined the consequences of crossing strains de-silenced in the *rde-3* mutant background back into a wild-type *rde-3(+)* background. We found that for a *gfp*::*csr-1* transgene de-silenced by *rde-3,* 27 % of *rde-3(+)* segregants (n=15) retained expression after outcross (Figure 2.2A). However, in contrast, strains bearing the *gfp*::*cdk-1* transgene, also desilenced by *rde-3,* were always rapidly and fully re-silenced by reintroducing *rde-3(+)* (n>20).

Table 2.1. Genetic test for maintenance of gene silencing

Gene(allele)	Gene function	Transgene expression	
		gfp::csr-1	*gfp::cdk-1*
rde-1(*ne300*)	Argonaute in RNAi	-	-
prg-1(*tm872*)	Piwi homolog	-	-
rde-3(*ne3370*)	Poly(A) polymerase	+	+
mut-7(*ne4255*)	3' to 5' exonuclease	+	+
hpl-1(*tm1624*)	HP1 homolog	-	-
hpl-2(*tm1489*)	HP1 homolog	+[c]	+[b]
hpl-1(*tm1624*) *hpl-2*(*tm1489*)	HP1 homolog HP1 homolog	+	+
met-1(n4337) *met-2(n4256)*	Methyltransferase Methyltransferase	-	NA
mes-3(*bn35*)[a]	Polycomb complex	+[b]	+[b]
mes-4(*bn23*)[a]	Trithorax complex	+[b]	+[b]
wago-1(*tm1414*)	*Cytoplasmic WAGO*	-	+[b]
nrde-3(*tm1116*)	*Nuclear WAGO*	-	NA
wago-9(*tm1200*)	*Nuclear WAGO*	+	+[b]
wago-1(*tm1414*) *wago-9*(*tm1200*)	*Cytoplasmic WAGO* *Nuclear WAGO*	NA	+
wago-9(*tm1414*) *wago-10*(*tm1186*)	*Nuclear WAGO* *Nuclear WAGO*	NA	+
wago-9(*tm1414*) *wago-10*(*tm1186*) *nrde-3*(*tm1116*)	*Nuclear WAGO* *Nuclear WAGO* *Nuclear WAGO*	NA	+
wago-9(*tm1414*) *wago-10*(*tm1186*) *wago-11*(*tm1127*) *nrde-3*(*tm1116*)	*Nuclear WAGO* *Nuclear WAGO* *Nuclear WAGO* *Nuclear WAGO*	NA	+

[a] Scored in sterile M-Z- mutants

[b] GFP is partially desilenced (GFP signal is weak in each worm)

[c] GFP is desilenced in fraction of germline in the same worm

Nuclear and Cytoplasmic WAGOs are required for silencing maintenance

Because RDE-3 and MUT-7 are required for the accumulation of RdRP-derived 22G-RNAs that engage WAGOs (Gu et al., 2009), we asked whether WAGOs are required for the maintenance of single-copy transgene silencing by crossing silent lines with several different *wago* mutant strains. We found that a mutation in the predominantly cytoplasmic germline WAGO, *wago-1*(*tm1414*) (Gu et al., 2009) partially de-silenced a *gfp*::*cdk-1* transgene but did not de-silence a *gfp*::*csr-1* transgene (Table 2.1 and Figure 2.3A and 2.3C).

The finding that *wago-1* mutants failed to de-silence *gfp*::*csr-1* and only partially de-silenced *gfp*::*cdk-1* suggested that additional WAGOs contribute to RNAe (Figure 2.3I). Furthermore, because RNAe involves a chromatin component, we suspected that nuclear WAGOs might be important for RNAe. The nuclear WAGO, NRDE-3/WAGO-12, is required for nuclear RNAi and transcriptional silencing in somatic tissues (Burton et al., 2011; Guang et al., 2008), and *nrde-3* mutants failed to de-silence a *gfp*::*csr-1* transgene in the germline (Table 2.1). However, within the WAGO sub-clade that includes NRDE-3 (Figure 2.3I), we identified WAGO-9 (HRDE-1/C16C10.3) as a nuclear WAGO that is restricted to the germline (Figure 2.3G). Furthermore, we found that *wago-9* (*tm1200*) mutants fully de-silenced a *gfp*::*csr-1* transgene and partially de-silenced a *gfp*::*cdk-1* transgene (Figure 2.3B and 2.3D), the converse of the relationship between *wago-1*(*tm1414*) and these RNAe lines. The de-silencing of *gfp*::*cdk-1* was increased in a *wago-1*; *wago-9* double mutant (Figure 2.3E). The

wago-9 locus was also identified by two other groups (Ashe et al., 2012) as a gene required for heritable RNAi (hence its other name, heritable RNAi-defective, *hrde-1).*

Because *gfp*::*cdk-1* was not completely desilenced by these *wago* mutant combinations, we asked if additional members of the nuclear WAGO sub-clade play a role in *gfp*::*cdk-1* silencing. Indeed, *gfp*::*cdk-1* was strongly de-silenced in a *wago-9*; *wago-10* (*t22h9.3*); *wago-11*(*f49f6a.1*); *nrde-3* quadruple mutant, as well as in a *wago-9*; *wago-10* double mutants (Table 2.1 and Figure 2.3F). Taken together, these findings indicate that cytoplasmic and nuclear WAGOs contribute to RNAe in parallel and that the input from cytoplasmic and nuclear WAGOs varies between individual RNAe lines.

The small RNAs that associate with WAGO-1 were previously identified by immunoprecipitation (IP) of FLAG::WAGO-1 followed by deep sequencing of associated small RNAs (Gu et al., 2009). We performed similar studies using a *flag::wago-9* transgene. We found that the targets of WAGO-9 largely overlap with those of WAGO-1 (Figure 2.3H). These observations suggest that nuclear and cytoplasmic WAGOs share targets and are likely to share a common 22G biogenesis pathway.

Figure 2.3. Genetic requirements for maintenance of *RNAe*

(A-F) Fluorescence microscopy of transgene desilencing in *wago* mutant backgrounds. The transgenes used were *neSi8 gfp*::*csr-1*(*RNAe*), which localizes to P-granules when expressed (indicated by arrow in A and B), and *neSi11 gfp*::*cdk-1*(*RNAe*), which is most prominent in oocyte nuclei (indicated by arrowheads in C-F).

(G) WAGO-9 is a germline expressed nuclear Argonaute. Fluorescence micrograph of GFP::WAGO-9 in the adult hermaphrodite germline. The dashed lines in the micrograph indicate the position of the syncytial germline.

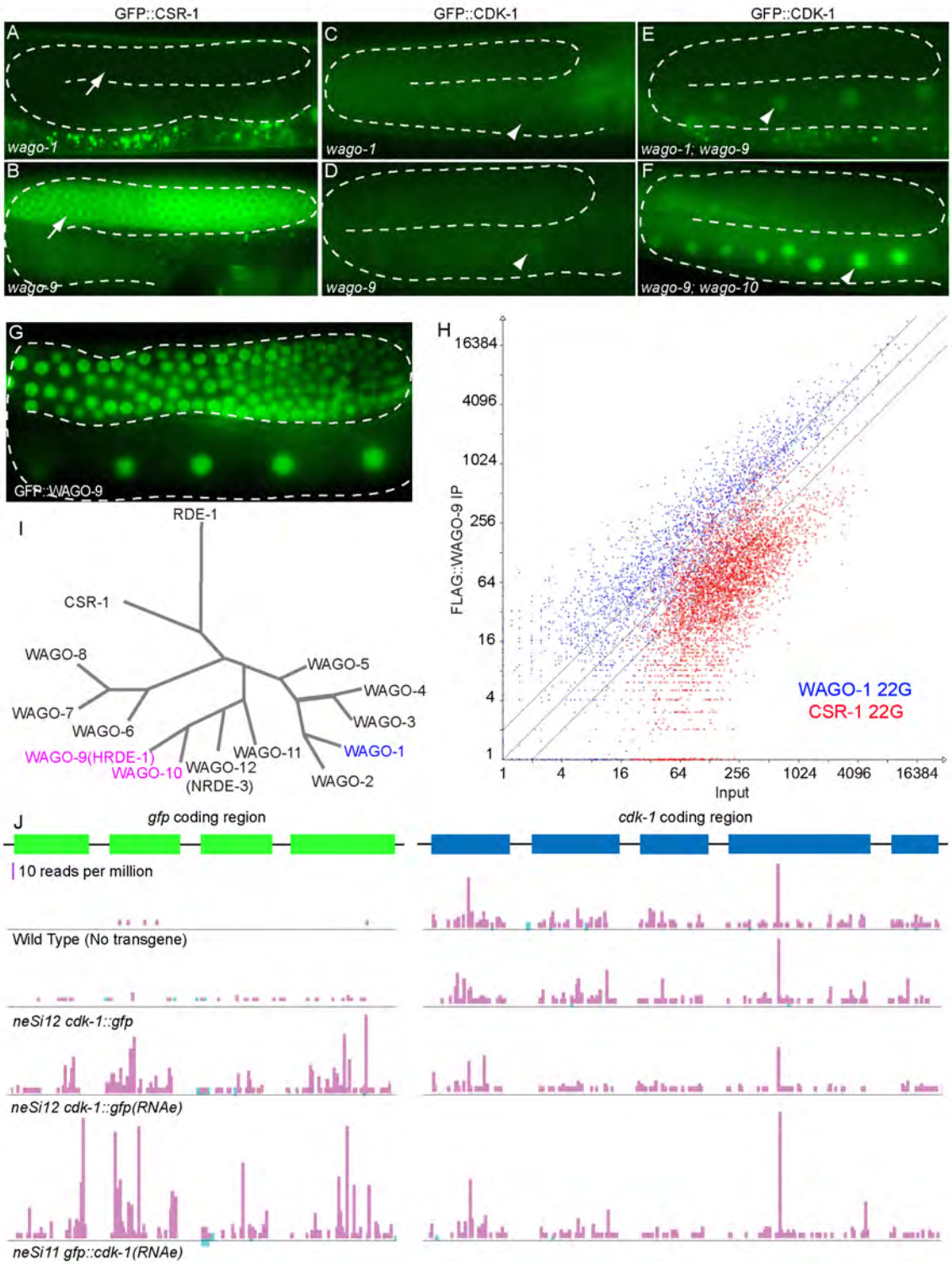
(H) WAGO-9-associated small RNAs overlap extensively with WAGO-1 small RNAs. The plot shows the enrichment of 22G-RNAs in FLAG::WAGO-9 IP relative to input. Each point in the graph corresponds to previously identified WAGO-1 (blue) and CSR-1 (red) target genes. The x- and y-axes represent the number of 22Gs ($\log_2$ scale) targeting each gene in the Input and WAGO-9 IP samples, respectively. The diagonal lines signify 2-fold enrichment (upper), identity (middle), and 2-fold depletion of 22G-RNAs in the WAGO-9 IP.

(I) Phylogenetic tree of WAGOs, CSR-1 and RDE-1. Adapted from (Gu et al., 2009).

(J) Small RNA density along the *gfp* and *cdk-1* coding regions of wild-type and indicated transgenic lines. Vertical bars represent the 5′ nt of a small RNA, and the height of each bar indicates the number of reads that start at that position. The strand is represented by color; sense (light blue) and antisense (pink). Scale

bar indicates 10 reads per million. Strain *neSi12 cdk-1::gfp(RNAe)* was generated by crossing *neSi12 cdk-1::gfp(+)* to *neSi11 gfp::cdk-1(RNAe)*.

Figure 2.3. Genetic requirements for maintenance of *RNAe*



Silencing correlates with accumulation of 22Gs targeting GFP

To examine the small RNA profile associated with germline silencing, we dissected gonads from different transgenic lines, including active, silent, and converted lines (e.g. active to silent and silent to active lines), and prepared small RNA libraries for deep sequencing (Figure 2.3J). Strikingly, each silenced line exhibited a marked accumulation of 22G-RNAs that were restricted to the *gfp* portion of the transgene sequence (Figure 2.3J). Consistent with the idea that these 22Gs are WAGO-pathway dependent, we found that 22G-RNA levels targeting *gfp* were significantly reduced in lines converted from silent to active by crossing through an *rde-3* mutant background.

Native germline-expressed genes are recognized by low levels of 22G-RNAs that engage CSR-1 (CSR-1-22Gs) (Claycomb et al., 2009). We found that the transgene sequences corresponding to endogenous germline-expressed mRNA sequences always exhibited low 22G-RNA levels similar to those observed for the endogenous sequences in wild-type non-transgenic animals (Figure 2.3J). These findings suggest that the WAGO-mediated silencing signal only targets the foreign sequences of the transgene.

Initiation of silencing requires the Piwi Argonaute PRG-1

Despite interacting with distinct small RNA species, both PRG-1 and RDE-1 function as primary AGOs upstream of WAGO-22G mediated silencing (Batista et al., 2008; Das et al., 2008; Pak and Fire, 2007; Sijen et al., 2007; Yigit et al., 2006). However, we found that neither *prg-1* nor *rde-1* mutants could activate an

already established silent transgene (Table 2.1). To explore the possibility that either PRG-1 or RDE-1 is involved in the initiation of RNAe, we generated new transgenic lines by directly injecting into *prg-1* and *rde-1* mutants. We chose to inject the *gfp::cdk-1* construct, because 100 % of MosSCI lines were silent when established in the wild-type background (n=21) (Figure 2.1A). In an *rde-1*(*ne300*) mutant strain, we found that the *gfp::cdk-1* transgene was silenced in all three newly isolated lines. Strikingly, however, when we repeated the same experiments with *prg-1*(*tm872*) mutants, the *gfp::cdk-1* transgene was fully active in all five independently generated transgenic lines (Figure 2.4). Taken together, these findings suggest that PRG-1 and piRNAs are involved in the initiation of transgene silencing, whereas dsRNA (e.g. from bi-directional transcription of the transgene) is not involved.

When established in the wild-type background, the epigenetic state of a transgene, whether active or silent, is stably maintained over many generations. If PRG-1 is only required for the initiation of silencing, then we expected that active transgenes established in a *prg-1* mutant background would remain active even after outcrossing to a wild-type strain. We found that *gfp*::*cdk-1* was expressed in 96 % (n=24) of the heterozygous F1 progeny. However, by the F3 generation, the *gfp*::*cdk-1* transgene was only expressed in 9 % (n=66) of animals heterozygous or homozygous for a wild-type allele of *prg-1*, and by the F4 *gfp*::*cdk-1* was silent in all wild-type descendants (Figure 2.4). Conversely, among the F3 animals that were once again homozygous for the *prg-1* mutation,

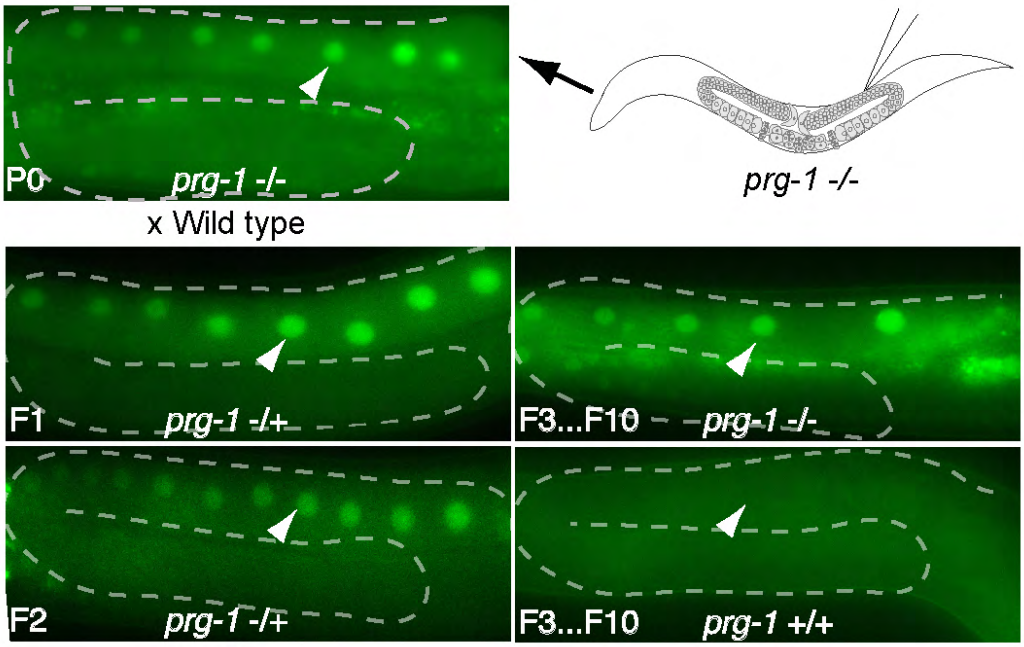
77 % (n=30) maintained expression of the gfp::cdk-1 transgene (Figure 2.4). These findings support the idea that PRG-1 is involved in the initiation of gene silencing.

However the finding that the transgene becomes silent after outcross to wild-type indicates that the active state for this transgene does not become epigenetically stable when propagated in the *prg-1* mutant background. This observation raises the possibility that PRG-1 is upstream of competing epigenetic pathways; one that initiates silencing and one that initiates anti-silencing (see below and DISCUSSION).

Figure 2.4. PRG-1 is required for the initiation of *RNAe*

prg-1(*tm872*) mutant worms injected with the *gfp*::*cdk-1* construct (top right) give rise to MosSCI lines that express GFP::CDK-1 (P0, top left). The micrographs show the expression status of GFP::CDK-1 in oocyte nuclei (arrowheads) before (P0) and after outcrossing to wild type (F1 and F2 panels), and after segregating homozygous *prg-1(+)* and *prg-1(-)* strains for several generations (F3-F10 panels). More than 10 worms were examined per generation. Results are detailed in the text.

Figure 2.4. PRG-1 is required for the initiation of *RNAe*



A trans-acting anti-silencing signal

The findings described above indicate that extremely stable silencing associated with single-copy transgenes is initiated by piRNAs and requires the same downstream factors that are required for RDE-1-dependent dsRNA-induced silencing. However unlike the silencing described here, to our knowledge, dsRNA-induced silencing (even when transmitted for numerous generations) has not been observed to become stable. Instead, all previous descriptions of inherited RNAi described reversion frequencies in the range of 80 % per generation (Alcazar et al., 2008; Vastenhouw et al., 2006).

We therefore wondered if PRG-1 somehow initiates a more stable mode of silencing than that initiated through RDE-1. To test this idea, we used *gfp* dsRNA to initiate silencing of active GFP(+) transgenes and monitored expression for multiple generations after removal of the dsRNA trigger. In each generation, we scored 10 animals from each of 10 independent lines for a total of 100 worms per generation. For the *gfp*::*csr-1* transgene, we found that, as expected, 100 % of the animals were silenced in the F1 generation. Remarkably, however, 100 % of *gfp*::*csr-1* worms remained silent in all ten lines for greater than 10 generations with no evidence of reversion. Similar results were obtained for the *cdk-1*::*gfp* transgene. This transgene, which was less prone to silencing during initial transgenesis, remained completely silent in 6 of 10 lines, whereas 4 lines recovered expression. Thus, the susceptibility of these active transgene lines to piRNA induced silencing mirrors their susceptibility to dsRNA-induced permanent

silencing.

The above data suggest that the MosSCI transgenes studied here are more sensitive than endogenous genes to permanent silencing by RNAi. To ask if this is generally true of transgenes, we asked whether exposure to *gfp* (*RNAi*) could permanently silence low-copy transgenes generated several years ago by different methods. For this analysis we chose two different transgenes generated by different approaches, *gfp*::*wrm-1* (Nakamura et al., 2005), which was produced by injecting an engineered yeast artificial chromosome, and *oma-1*::*gfp* (Lin, 2003), which was generated by biolistic gold-particle mediated transformation (Praitis, 2006). We found that both transgenes were efficiently silenced by RNAi in the F1 (100 %, n=100), but expression always fully recovered after removal of the dsRNA trigger (100 % GFP+ by the F3 generation).

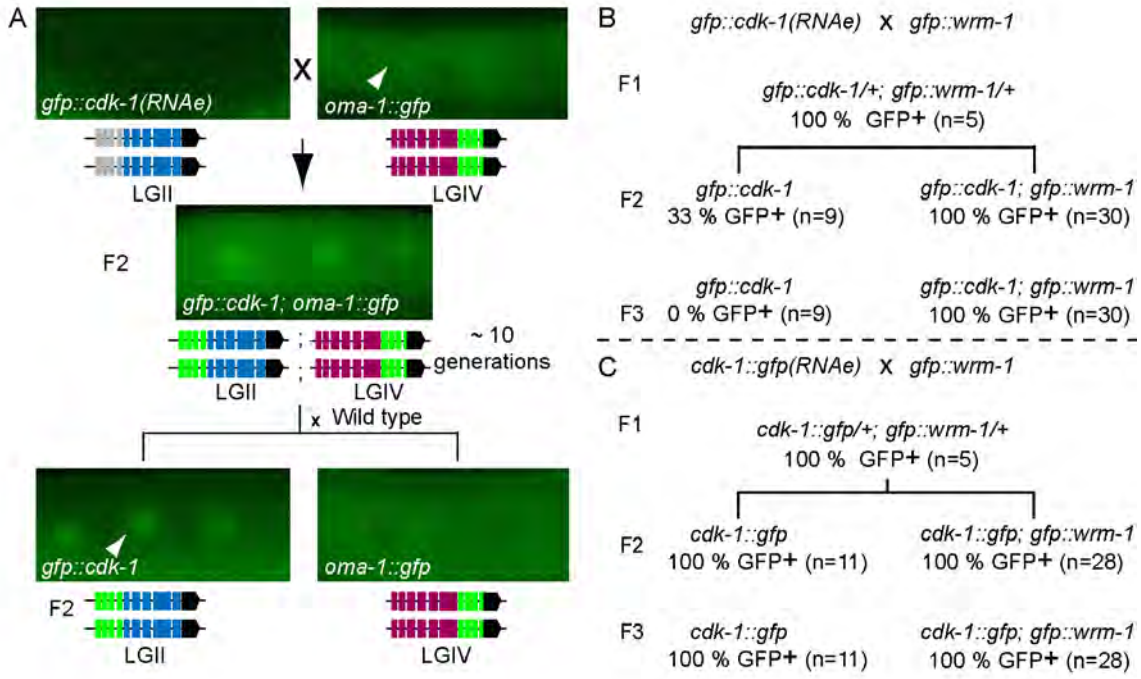
Considering the resistance of *gfp*::*wrm-1* and *oma-1*::*gfp* to permanent silencing by dsRNA, we wondered if they might also be resistant to trans-silencing in crosses with silent transgenes. Surprisingly, not only were both *gfp*::*wrm-1* and *oma-1*::*gfp* resistant to trans-silencing, we found that both transgenes could dominantly activate the expression of a silent transgene in the F1 cross progeny (Figure 2.5A-2.5C). Expression was initially low in the F1 and F2, but, when propagated along with *gfp*::*wrm-1* or *oma-1*::*gfp* transgenes, the trans-activated transgene alleles became fully expressed by the third generation (Figure 2.5A-2.5C). Finally, after propagating the activated transgene lines in the

presence of *gfp*::*wrm-1* or *oma-1*::*gfp* for a few generations, we segregated the transgenes away from each other. We found that *gfp::cdk-1* returned to its silent state (Figure 2.5B), while *cdk-1*::*gfp* remained stably expressed after exposure to the active transgene (Figure 2.5C). Although we need to test more transgenic lines, these findings indicate that a trans-acting dominant mechanism can activate a silent transgene and suggests that activating and silencing signals compete with each other for dominance when transgene alleles interact.

Figure 2.5. Evidence for a trans-acting anti-silencing activity

(A) Schematic illustrating the cross between *neSi11 gfp::cdk-1*(*RNAe*) and *teIs1 oma-1*::*gfp(+)*. The micrographs show the expression status of GFP::CDK-1 in oocyte nuclei (arrowhead) when expressed and OMA-1::GFP in the oocyte cytoplasm. The dashed circles (top left) show the position of GFP-negative oocyte nuclei in the *neSi11 gfp::cdk-1*(*RNAe*) strain. The cartoon below each micrograph indicates whether the transgene is expressed or silent (gray).
(B-C) Schematics illustrating crosses between *neIs2 gfp::wrm-1(+)* males and (B) *neSi11 gfp*::*cdk-1*(*RNAe*) or (C) *neSi12 cdk-1*::*gfp*(*RNAe*) hermaphrodites. After each cross the two transgenes were either maintained together or allowed to segregate away from each other. The GFP::WRM-1 signal is very weak and was scored periodically during the analysis. The percentage of GFP+ worms indicates the expression of the CDK-1 fusion proteins.

Figure 2.5. Evidence for a trans-acting anti-silencing activity



DISCUSSION

Recognition of self and non-self nucleic acids

Organisms employ an array of mechanisms that afford some control over the expression of foreign sequences (Hornung and Latz, 2010; Murray, 2002). In *Drosophila*, for example, piRNAs have been shown to mediate transposon silencing in the germline (Malone and Hannon, 2009). In this remarkable system transposons are thought to move freely at first, until a spontaneous insertion into a genomic piRNA generating locus results in the expression of piRNAs perfectly complementary to the new transposon (Khurana and Theurkauf, 2010). The stable genomic integration of the transposon within the piRNA-generating locus, initiates silencing and provides a genetic (rather than epigenetic) memory of the invasive sequence. Maternally inherited piRNAs function to prime production of piRNAs but cannot function without a genetic reservoir of transposon sequence in the maternal genome (Brennecke et al., 2008). Even defective transposon remnants embedded in piRNA-producing loci are sufficient to maintain piRNA production in the absence of a functional transposon (Grentzinger et al., 2012). Here we have shown that *C. elegans* employs piRNAs in a very different mechanism that recognizes even single-copy foreign sequences, and initiates a remarkably stable epigenetic memory of silencing. Rather than recognition based on the site of integration or on an aberrant feature of the transgene DNA or RNA product, our findings suggest that initiation of silencing involves the comparison of the foreign sequence to an epigenetic memory of previously expressed

sequences. Thus, genetically identical individuals in *C. elegans* can exhibit remarkably stable but opposite patterns of expression.

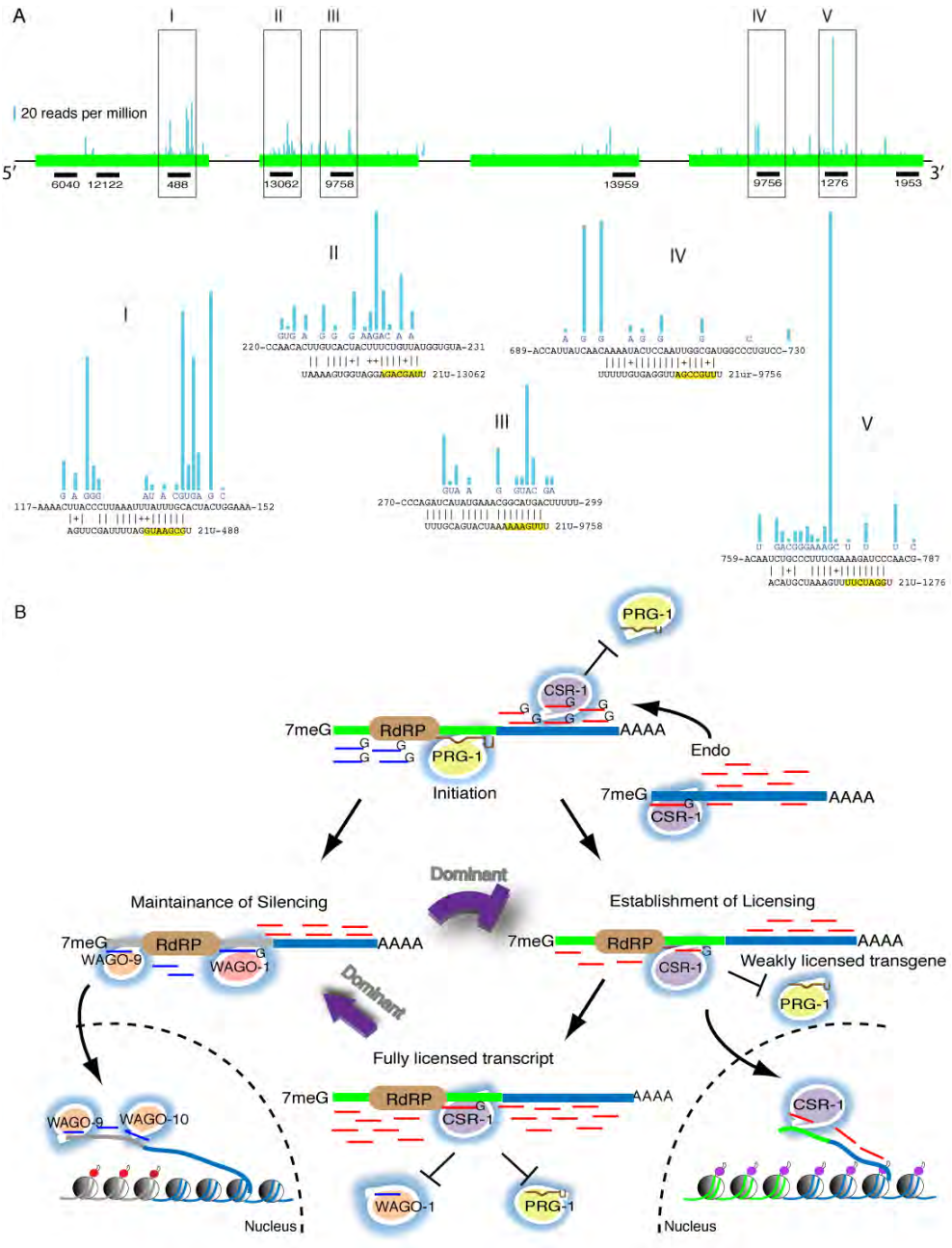
We propose a model in which three AGO pathways function together in a system that maintains an inventory of expressed mRNAs while constantly scanning for foreign sequences (Figure 2.6B). In this system, PRG-1 uses genomically encoded piRNA co-factors to scan, via imperfect base pairing interactions, for foreign RNAs expressed in the germline. Upon targeting, PRG-1 recruits RdRP to produce anti-sense 22G-RNAs, which are loaded onto WAGO Argonautes. In turn, WAGOs mediate silencing and establish a memory of non-self RNA. A third, as yet unidentified pathway provides a memory of self and is capable of acting as an anti-silencing signal. Although our studies have not yet identified the anti-silencing (self-recognition) mechanism, the CSR-1 22G-RNA pathway provides an attractive candidate for this activity (See further discussion below). We propose that the self-recognition pathway can prevent PRG-1 from recruiting the WAGO pathway, providing a function that helps expressed transgenes to maintain their expression and helps endogenous genes to recover from WAGO-mediated silencing induced by RNAi. The initial decision to silence or express the transgene represents a stochastic outcome of competition between establishments of these epigenetic self- or non-self memories.

Figure 2.6. Model: Self non-self RNA recognition in *C.elegans*

(A) Schematic showing the density of 22G-RNAs targeting GFP in *neSi8 gfp::csr-1(RNAe)* worms, as described in the legend of Figure 2.3J. Scale bar indicates 20 reads per million. The positions of several 21U-RNAs that could base pair with mismatches to the *gfp* sequence are indicated below the gene diagram. Five major 22G hotspots (numbered boxes) are enlarged to show the base pairing between the candidate 21U-RNA and *gfp*, as well as the density of 22G-RNAs at single-nucleotide resolution. Each 21U-RNA has at most two G:U pairs within the seed region (nts 2-8, yellow highlight), and at most 3 non-seed mismatches (nts 9-21).

(B) Model for the allelic interactions between transgenes observed in this study.

Figure 2.6. Model: Self non-self RNA recognition in *C.elegans*



Repetitive and single copy transgenes exhibit distinct but overlapping silencing mechanisms

The silencing of high-copy and single-copy transgenes share several features including chromatin-related and WAGO 22G pathway requirements. Furthermore, both high-copy (Praitis, 2006) and single-copy silencing (the present study) occur independently of RDE-1 and thus are unlikely to be initiated by dsRNA. However, several observations suggest that high-copy transgenes are subject to distinct modes of recognition and silencing. First, high-copy transgenes were at best only partially de-silenced in WAGO-pathway mutant contexts, such as *rde-3* and *mut-7* (Tabara et al., 1999) and (data not shown) whereas single copy transgenes were fully desilenced and in some cases even maintained their expression after outcrossing to wild type. Second, high-copy transgenes were fully and rapidly silenced in the germline of *prg-1* mutant animals (data not shown), indicating that a distinct initiation step is involved in high-copy number silencing. Third, high-copy number silencing was observed even when only the native germline gene sequences were present in the transgene (data not shown), whereas silencing of the single-copy transgene was correlated with the presence of foreign sequences within the germline-expressed portion of the transgene construct. Finally, unlike the single-copy silencing described here, where trans-silencing remains focused on foreign sequences, high-copy transgenes were found to elicit co-suppression of the endogenous gene (Dernburg et al., 2000; Ketting and Plasterk, 2000). Taken together these

observations are consistent with the existence of at least two distinct modes of silencing that act on transgenes, one that depends on high-copy number and can spread throughout the transgene, and a second that requires PRG-1 and is restricted to portions of the transgene composed of foreign sequences.

21U-RNAs complementary to *gfp* are correlated with 22G biogenesis

Our findings suggest that transgene silencing is initiated by PRG-1 and depends on the presence of foreign *gfp* sequences in the transgene. In a parallel study, PRG-1 was shown to initiate silencing of synthetic reporters containing sites perfectly complementary to 21U-RNAs (Bagijn et al., 2012; Lee et al., 2012). Mismatched pairing was also correlated with silencing both on transgenes (Bagijn et al., 2012) and on presumptive endogenous targets (Bagijn et al., 2012; Lee et al., 2012). We have not identified 21U-RNAs that are perfectly complementary to *gfp*; however, there are dozens of potential high-affinity 21U-RNA-GFP target sites (data not shown). Our recent studies (Lee et al., 2012) suggest that PRG-1/21U-RNA targeting initiates 22G-RNA biogenesis within a +/- 40 nt window around the site of 21U-RNA complementarity on the target RNA. We found 8 regions in *gfp* where 22G-RNAs were detected at greater than 75 reads per million in a silent strain (Figure 2.6A). We identified potential high-affinity 21U-RNA interactions in all 8 regions. The potential base-pairing interactions and the proximal 22G-RNAs found in a silent transgenic strain are shown at single-nucleotide resolution in Figure 2.6A (also see EXPERIMENTAL PROCEDURES). Validation of these candidate 21U-RNA target sites and the

general rules that govern piRNA targeting remain to be elucidated.

CSR-1 as an anti-silencing Argonaute

At least three mechanisms must work together to explain the all-or-none nature (expressed or silent) of the epigenetic states observed, and the stable heritability of these states once established (Figure 2.6B). The genetic studies, thus far, have implicated PRG-1 in the initiation of silencing and the WAGO pathway in the maintenance of silencing. The third pathway required is a "maintenance of expression" or "anti-silencing pathway". Such a pathway is necessary to explain why, once established, active transgenes are stably transmitted from one generation to the next without undergoing spontaneous silencing. An anti-silencing pathway could also explain how certain active transgenes are able to dominantly activate silent transgenes (Figure 2.6B).

The CSR-1 22G pathway targets endogenous germline-expressed mRNAs (Claycomb et al., 2009), and is an ideal candidate for an anti-silencing pathway. In vitro, CSR-1 is catalytically active and capable of cleaving a target (Aoki et al., 2007), whereas the all WAGOs lack key catalytic residues (Yigit et al., 2006). Perhaps CSR-1 can compete by selectively destroying RNAs on which RdRP is bound, thus preventing or attenuating the production of WAGO 22G-RNAs. It is not known how CSR-1 targeting is first established. However, all of the transgenes that we analyzed contain endogenous germline expressed sequences known to be targeted by CSR-1 22Gs. Perhaps CSR-1 22Gs can spread *in trans* along a target transcript as has been shown for the transitive

RNAi mediated by WAGOs after dsRNA targeting (Pak and Fire, 2007; Sijen et al., 2007; Yigit et al., 2006)**.** If so, then stable expression of a transgene may reflect the spread of CSR-1 targeting to the foreign portion of the transgene prior to PRG-1 recognition.

Interestingly, although the anti-silencing signal initially appears to be sufficient to prevent PRG-1 driven silencing, it is not sufficient to prevent silencing initiated in crosses with a silent transgene or when dsRNA is used to stimulate gene silencing. If CSR-1 22G-RNAs represent the anti-silencing signal, then it will be interesting to explore whether the levels of CSR-1 22G-RNAs build up over generations. If so then, the older transgenes, which were able to activate a silent transgene, may show relatively high levels of CSR-1 22G-RNAs targeting *gfp* when compared to newly established lines. However, it is also possible that as yet unknown features of the chromatin environments of the different transgenes drives their different sensitivity to trans-silencing and their differing abilities to trans-activate or to recover from silencing spontaneously.

Finally, it is worth noting that PRG-1 may function upstream of RdRP recruitment for both the CSR-1 and WAGO pathways. If so, then the decision to express or silence a new transgene may represent the result of a competition between the CSR-1 and WAGO pathways for RdRP loading, downstream of this initial recruitment. An expectation for such a model would be that both the maintenance-of-silencing (non-self) and maintenance-of-expression (self) pathways should fail to initiate when PRG-1 is absent. To further explore this

question it will be important to analyze the behavior of additional transgenes established in the *prg-1* mutant background.

RNA-induced epigenetic inheritance

Here we have described a remarkably stable form of epigenetic inheritance (RNAe) that is initiated by *C. elegans* piRNAs. While RNAe likely serves as a defense against transposons and other invasive sequences, it is also possible that it could have a more general role with significant potential impact on evolution. For example, RNAe could accelerate evolutionary change by heritably modulating the expression of unpaired parental alleles to allow the phenotypic expression of recessive traits among F1 progeny. Consistent with this idea a recent report has shown that a paternally derived allele with no homolog in the hermaphrodite genome is subject to dominant silencing, and that silencing was prevented by injecting single stranded RNAs matching the coding region of the absent gene into hermaphrodite gonads prior to the cross (Johnson and Spence, 2011). These observations are consistent with a mechanism for the licensing of gene expression by maternal RNA and, along with the present study, support the existence of an epigenetic switch that is sensitive to prior expression of a gene. These phenomena are also similar to a form of allelic interaction known as paramutation that has been described in organisms ranging from mice to corn (Erhard and Hollick, 2011). Thus, it appears likely that diverse organisms can both track and respond epigenetically to the history of gene expression. In *C. elegans*, this process overlaps mechanistically with RNAi, but involves a distinct

triggering mechanism that requires the genomically-encoded piRNAs. Mammalian genomes encode abundant piRNA species that are analogous to *C. elegans* 21U-RNAs. Our findings together with those of Ashe et al (this issue) raise the intriguing possibility that these so-called meiotic piRNAs of mammals function in epigenetic programming.

EXPERIMENTAL PROCEDURES

Genetics

All *C. elegans* strains were derived from the Bristol N2 strain and cultured as described (Brenner, 1974). The strains used in this study are listed in Supplemental Information.

MosSCI by direct injection

MosSCI lines were generated by the direct insertion method using strain EG4322 and EG5003 as described (Frokjaer-Jensen et al., 2008). Targeting vectors are described in Supplemental Information.

MosSCI by heat-shock and ivermectin selection

Strain WM186 was injected with a DNA mixture containing 50 ng/ml each of pRF4::*rol-6(su1006)*, pCCM416::*Pmyo-2*::*avr-15*, and pJL44::*Phsp-16.48::MosTase::glh-2utr* (Frokjaer-Jensen et al., 2008), and either 1 ng/ml or 50 ng/ml of targeting vector. MosSCI was performed using the heat-shock method (Frokjaer-Jensen et al., 2008) and single-copy insertion lines were selected on ivermectin to select against animals carrying the extrachromosomal array. Additional details are provided in Supplemental Information.

Small RNA cloning from isolated germlines

Ten gonads from each strain were dissected in 1x PBS containing 0.1 mM EDTA, 1 mM Aurin tricarboxylate, 0.1% Tween 20, and 0.2 mM levamisole (Wang et al., 2009). Total RNAs were extracted with 5 volumes of TRI Reagent (MRC). Small RNAs were gel-purified and cloned as described (Gu et al., 2009). *gfp::csr-1* small RNAs were pre-treated with Tobacco Acid Phosphatase (TAP, Epicenter Biotechnologies). *gfp::cdk-1* and *cdk-1::gfp* small RNAs were pre-treated with CIP/PNK (NEB). Libraries were sequenced in the UMass Deep Sequencing Core using an Illumina GAII instrument.

Small RNA cloning from FLAG::WAGO-9 immune complexes

Synchronous adult *flag::wago-9* worms were dounced in a stainless steel homogenizer. FLAG::WAGO-9 was immunoprecipitated from 20 mg of lysate essentially as described (Gu et al., 2009). Small RNAs were extracted from WAGO-9 immune complexes as well as a portion of the input lysate, gel-purified, pre-treated with TAP, cloned and sequenced as above.

Computational analysis of small RNAs

Deep sequencing data were processed and analyzed using custom Perl scripts (Gu et al., 2009). Definition of WAGO and CSR-1 22Gs are described in (Gu et al., 2009). Candidate 21U-RNAs that target *gfp* were identified by searching for seed sequences (nts 2-8) that base-pair with at most two G:U wobbles, and allowing at most 3 unpaired non-seed residues (nts 9-21). Additional details are provided in Supplemental Information. Perl scripts are available on request.

Chromatin Immunoprecipitation

ChIP was performed essentially as described (Claycomb et al., 2009) except that synchronized adult *neSi8 gfp*::*csr-1* (*RNAe*) and *neSi9 gfp*::*csr-1(+)* worms were dounced in a stainless steel homogenizer (30 strokes) prior to cross-linking with 2.6 % formaldehyde. Immunoprecipitations were performed in a total volume of 1 mL (5 mg) with 10 mg of anti-Histone H3 (ab1791, Abcam) or anti-H3K9me3 (ab8898, Abcam) antibodies. Immune complexes were recovered with 50 mL of Protein A Dynabeads (Invitrogen). Three independent ChIP experiments were performed and analyzed by quantitative PCR.

Quantitative PCR

Quantitative PCR was performed as described (Claycomb et al., 2009) using an ABI 7500 Fast Real-Time PCR instrument. For RNA analysis, cDNA was generated from 1 mg of total RNA using random hexamers and Superscript III Reverse Transcriptase (Invitrogen). *gfp*::*csr-1* expression was measured relative to *clp-3* mRNA levels. H3K9me3 ChIP was first normalized to Histone H3 ChIP, and fold enrichment was then determined relative to an H3K9me3 negative control locus. Primer sequences are provided in Supplemental Information.

Transgenerational RNAi phenotype

A single *neSi9 gfp::csr-1(+)*, *neSi12 cdk-1::gfp(+), tsIs1 oma-1::gfp(+)* or *neIs2 gfp::wrm-1(+)* adult worm was placed onto each of 10 plates seeded with *gfp(RNAi)* food. A single F1 worm from each plate was transferred to OP50 (control) or *gfp(RNAi)* food and each line was maintained for 10 generations by

transferring a single worm from each plate to the corresponding food source, OP50 or *gfp(RNAi)*. In each generation, 10 progeny from each plate were scored for *gfp* expression (100 total for each condition).

Western blot analysis

Antibodies used for Western blotting are anti-CSR-1 (Claycomb et al., 2009), anti-GFP (A01704, Genscript) and anti-a-Tubulin (MCA78A, Serotec) antibodies.

Microscopy

Transgenic worms were mounted in dH_2O on RITE-ON glass slides (Beckton Dickinson). Epi-fluorescence and differential interference contrast (DIC) microscopy were performed using an Axioplan2 Microscope (Zeiss). Images were captured with an ORCA-ER digital camera (Hamamatsu) and AxioVision (Zeiss) software.

Preface to Chapter III

This chapter provides insight into the phenomenon of RNA-induced epigenetic activation mediated by Argonaute CSR-1/22G in recognition of self-RNA.

My contributions to this chapter are: Figures 3.1, 3.2, 3.3, and 3.4

Masaki Shirayama contributed: Figures 3.2B-F

Weifeng Gu contributed: Figure 3.2B-F

Craig C. Mello contributed: Figure 3.4

This chapter was published as:

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

CHAPTER III: The *C. elegans* CSR-1 Argonaute Pathway Counteracts Epigenetic Silencing To Promote Germline Gene Expression

SUMMARY

Organisms can develop adaptive sequence-specific immunity by re-expressing pathogen-specific small RNAs that guide gene silencing. For example, the *C. elegans* PIWI-Argonaute/piRNA pathway recruits RNA-dependent RNA polymerase RdRP to foreign sequences to amplify a trans-generational small RNA-induced epigenetic silencing signal (termed RNAe). Here we provide evidence that in addition to an adaptive memory of silenced sequences, *C. elegans* can also develop an opposing adaptive memory of expressed/self mRNAs. We refer to this mechanism, which can prevent or reverse RNAe as RNA-induced epigenetic gene activation (RNAa). We show that CSR-1, which engages RdRP-amplified small RNAs complementary to germline-expressed mRNAs, is required for RNAa. We show that a transgene with RNAa activity also exhibits accumulation of cognate CSR-1 small RNAs. Our findings suggest that *C. elegans* adaptively acquires and maintains a trans-generational CSR-1 memory that recognizes and protects self mRNAs, allowing piRNAs to recognize foreign sequences innately, without need for prior exposure.

INTRODUCTION

Epigenetics is often defined as the stable transmission of gene expression programs through mitotic or meiotic cell division without alteration in the DNA sequence (Bird, 2007). In eukaryotic cells epigenetic inheritance can be driven by covalent modifications to chromatin, often referred to as chromatin marks or

simply epigenetic marks (Grewal and Elgin, 2007; Henderson and Jacobsen, 2007; Lippman and Martienssen, 2004; Strome and Lehmann, 2007).

An emerging theme in epigenetic regulation is the frequent involvement of non-coding RNAs (Daxinger and Whitelaw, 2012; Grewal and Elgin, 2007; Henderson and Jacobsen, 2007; Lessing and Lee, 2013; Lim and Brunet, 2013). In many organisms, epigenetic silencing has been linked to RNAi-related mechanisms, which involve small non-coding RNAs termed short-interfering (si) RNAs (Ghildiyal and Zamore, 2009). Interestingly, the best-studied examples of RNAi-related epigenetic silencing also involve chromatin marks and their associated enzymatic mediators (Grewal and Elgin, 2007; Lippman and Martienssen, 2004), suggesting that RNAi and chromatin-modifying mechanisms reinforce and synergize with each other. Whereas the propagation of chromatin marks occurs in cis, RNAi can propagate in trans, allowing coordinate regulation of alleles on sister chromatids or of whole gene families such as transposons dispersed throughout the genome.

The core effectors of all RNAi-related pathways are Argonaute proteins. Argonautes present their guide RNAs for base pairing with target sequences and, upon binding, can cleave the target RNA and/or recruit cofactors that mediate post-transcriptional or transcriptional silencing (Ghildiyal and Zamore, 2009; Kuhn and Joshua-Tor, 2013). Although much less common, there are several examples of small-RNA pathways that appear to activate gene expression. For example, studies in human cultured cells have implicated small RNAs and/or

Argonautes in gene activation, a phenomenon referred to as RNAa (Janowski et al., 2007; Li et al., 2006; Place et al., 2008). In these examples, targeting is thought to occur within the promoter region of the gene, perhaps acting on nascent promoter-derived transcripts, and is correlated with the induction of chromatin marks characteristic of gene activation. In plants small dsRNAs have been implicated in the activation of the Petunia pMADS3 homeotic gene and are thought to act by promoting DNA-methylation at a CpG site within an intronic cis-promoter element (Shibuya et al., 2009a).

Two major groups of Argonaute proteins, the AGO proteins and the PIWI proteins, are encoded by animal genomes. PIWI Argonautes are expressed abundantly in the germline where they engage small-RNA species termed piwi-interacting (pi) RNAs (for review, see Juliano et al., 2011). In *C. elegans*, the PIWI Argonaute PRG-1 engages over 30,000 distinct genomically-encoded piRNA species (Batista et al., 2008; Das et al., 2008; Gu et al., 2012b). Recent studies have shown that PRG-1 initiates silencing of transgenes containing foreign, non-*C. elegans* sequences (Shirayama et al., 2012) and suggest that it does so while allowing imperfect base pairing with target sequences (Bagijn et al., 2012; Lee et al., 2012; Shirayama et al., 2012). Upon recognition of foreign sequences PRG-1 is thought to recruit a cellular RdRP, which in turn amplifies the silencing signal by producing antisense siRNAs perfectly complementary to the foreign sequences. These amplified siRNAs are loaded onto members of an expanded clade of WAGOs, which are implicated in both cytoplasmic and

nuclear gene silencing (Buckley et al., 2012; Gu et al., 2009; Guang et al., 2008; Yigit et al., 2006). The result is a remarkably stable mode of epigenetic silencing, termed RNA-induced epigenetic silencing (RNAe) (Shirayama et al., 2012). Alleles that are silenced by RNAe send trans-acting Argonaute-small-RNA signals that act in a sequence-specific manner to induce the permanent trans-generational silencing of their targets (Shirayama et al., 2012). The maintenance of RNAe requires chromatin factors, including heterochromatin protein 1 (HP1) and multiple histone methyltransferases (Ashe et al., 2012; Luteijn et al., 2012; Shirayama et al., 2012). Given the high numbers and the sequence diversity of *C. elegans* piRNAs, the allowance of two or three mismatches during target recognition should suffice, in principle, for piRNA to bind virtually any foreign RNA sequence. However, piRNAs should also recognize endogenous RNAs and therefore the piRNA surveillance model requires that “self” RNA be protected from RNAe (Shirayama et al., 2012).

The CSR-1 Argonaute engages antisense siRNAs complementary to the majority (perhaps all) endogenous germline-expressed genes (Claycomb et al., 2009; Gu et al., 2009). This finding, and the fact that its targets do not appear to exhibit CSR-1-dependent silencing, make this Argonaute a candidate for a self-RNA recognition factor. Paradoxically, however, CSR-1 protein has been shown to exhibit slicer activity in vitro (Aoki et al., 2007), and *csr-1* mutants are partially deficient in dsRNA induced silencing (Claycomb et al., 2009; Yigit et al., 2006). The siRNAs that engage CSR-1, like those that engage WAGO Argonautes, are

RdRP products. *C. elegans* RdRP products are often referred to as 22G-RNAs because they exhibit a predominant length of 22 nucleotides and a strong bias for a 5’ guanosine.

Evidence for a trans-activating signal that can counteract RNAe was discovered in crosses between an RNAe transgene and homologous actively expressed transgenes (Shirayama et al., 2012). Because this process involves the epigenetically transmitted, RNA-induced trans-activation of a silent allele (see below), we refer to the phenomenon as “RNAa” for RNA-induced epigenetic gene activation. Transgene alleles that are capable of sending the activating signal are designated as RNAa alleles; for example, *oma-1::gfp(RNAa)*.

Here we show that CSR-1 is required for RNAa and that the ability of a foreign sequence to direct transactivation is correlated with acquisition of CSR-1-associated small RNAs antisense to the foreign sequence. In contrast to previously studied RNAa phenomena, the CSR-1-associated activating small-RNAs target sequences present in the mature mRNA rather than promoter or intron sequences. We show that propagation of an RNAe and an RNAa allele together for multiple generations results in a gradual transfer of a stable, expressed state to the formerly silent transgene. Finally, consistent with the idea that RNAa counteracts PRG-1 recognition, we show that re-silencing of a trans-activated RNAe allele depends on PRG-1 activity. Our findings suggest that CSR-1 small RNAs constitute a memory of previous germline-gene expression that protects endogenous genes from piRNA recognition. This self-memory

system allows foreign sequences to be recognized innately without the need for prior exposure. Taken together, these findings and previous work on RNAe suggest that the *C. elegans* germline employs Argonaute-small-RNA complexes as trans-generational binary signals that program and reinforce the ON/OFF expression state for thousands of germline genes.

RESULTS

CSR-1 is required for RNAa

As a first test of whether trans-activation depends on CSR-1 activity, we crossed *oma-1::gfp(RNAa)* to *gfp::cdk-1(RNAe)* and exposed newly hatched F1 cross progeny to either *csr-1(RNAi)* by feeding, or to a control RNAi. Since OMA-1::GFP is expressed uniformly in oocyte cytoplasm (Lin, 2003), transactivation in this assay is evidenced by accumulation of the nuclear GFP::CDK-1 gene product (as shown in Figure 3.1A-3.1C). When cross progeny were exposed to a control RNAi directed against *sel-1,* an abundant germline gene with a function unrelated to small RNA pathways, we found that 100% (n=66) of the F1s exhibited trans-activation of *gfp::cdk-1(RNAe)* (Figure S3.1A). In contrast, we found that 0% (n=80) of F1s exposed to *csr-1(RNAi)* exhibited GFP::CDK-1 nuclear expression (Figure S3.1A). These findings suggest that CSR-1 activity is required in the zygote for transactivation of an RNAe allele.

We next wished to explore the consequences of reducing the dose of *csr-1* activity. To do this we conducted the transactivation assay using heterozygous *csr-1(tm892)* null mutant animals (Figure 3.1D), which exhibit wild-type fertility.

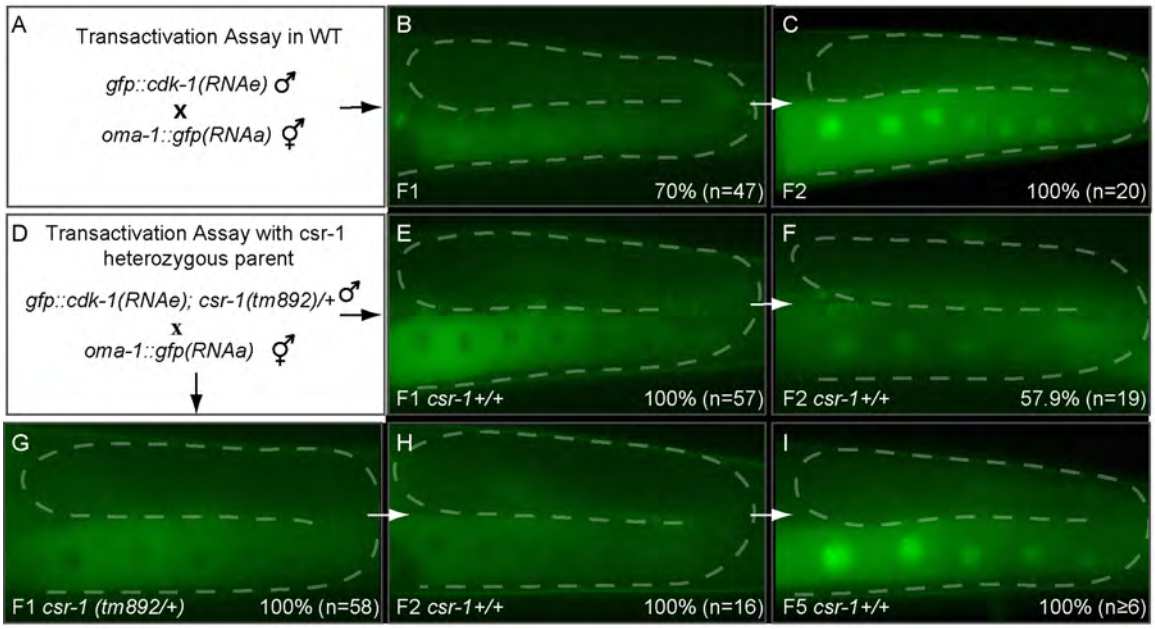
Interestingly, we found that trans-activation failed to occur when either transgenic parent was heterozygous for *csr-1(tm892)* (Figure 3.1E and 3.1G). We found that 100% of the F1 cross progeny failed to activate *gfp::cdk-1(RNAe)* when the *csr-1(tm892)* mutant was introduced from the father (*n=115)* or from the mother (*n*=15). This parental effect indicates that zygotic expression of CSR-1, although necessary, as suggested by the RNAi studies above, is not sufficient for trans-activation: Even F1 progeny homozygous for wild-type *csr-1(+)* activity failed to exhibit trans-activation if either parent was heterozygous for *csr-1(tm892)* (Figure 3.1E). As expected, when F1 wild-type *csr-1(+)* hermaphrodites were allowed to self cross, we observed trans-activation in the germlines of their F2 progeny (57.9%, n=19; Figure 3.1F). In contrast, heterozygous *csr-1(tm892)* hermaphrodites produced self progeny that failed to exhibit trans-activation (0%, n=16; Figure 3.1H) and transactivation was only restored among their wild-type progeny in subsequent generations (100%, n≥6; Figure 3.1I).

Figure 3.1. CSR-1 is required for RNAa.

(A and D) Schematic diagrams of crosses between silenced (RNAe) and licensed (RNAa) GFP transgenic strains as indicated.

(B, C, E-I) Epifluorescence images of representative germlines (outlined with dashes) in first (F1) and subsequent (F2, F3, F5) generations. The cytoplasmic fluorescence signal is OMA-1::GFP; the nuclear signal is GFP::CDK-1. The percentages indicate the number of animals exhibiting the shown phenotype in this and the subsequent figures.

Figure 3.1. CSR-1 is required for RNAa.



RNAa activity correlates with the accumulation of CSR-1 22G-RNAs

A previous study indicated that 22G-RNAs targeting *cdk-1::gfp*, a neutral transgene that is expressed but sensitive to silencing via RNAe, are present at very low levels, much lower for example than the level of CSR-1 22G-RNAs targeting the endogenous *cdk-1* portion of the transgene (Shirayama et al., 2012). The genetic analysis of RNAa described above suggest that transactivation of an RNAe allele is acutely sensitive to the dose of CSR-1 activity. We therefore wondered if small RNAs targeting GFP in the *oma-1::gfp(RNAa)* strain might be enriched to levels similar to an endogenous germline-expressed gene and whether they depend on CSR-1 activity. To explore this possibility, we first analyzed total small-RNA levels targeting *oma-1::gfp* in wild-type animals and in mutants defective in RNAa, *csr-1(tm892),* or defective in RNAe, *rde-3(ne3370).* In wild-type and *rde-3* mutant animals, we found that 22G-RNAs targeting *gfp* exhibited levels similar to 22G-RNAs targeting *oma-1* itself (Figure 3.2B and 3.2C). Conversely, and consistent with the idea that these *gfp*-targeted 22G-RNAs are in the CSR-1 pathway, we found that small RNAs targeting *gfp* were reduced by 73% in *csr-1(tm892)* mutants, a reduction similar to that observed for small RNAs targeting *oma-1* and other germline-expressed RNAs (Figure 3.2D and data not shown).

We next examined the physical association of *gfp*-directed 22G-RNAs by sequencing RNAs recovered in Argonaute protein immunoprecipitation (IP) complexes. To do this we conducted IP assays using epitope-tagged Argonaute

proteins, FLAG::CSR-1 and FLAG::WAGO-9/HRDE-1(Ashe et al., 2012; Buckley et al., 2012; Shirayama et al., 2012). Consistent with their genetic dependence on *csr-1,* we found that 22G-RNAs antisense to *gfp* were enriched (3.14-fold) in the FLAG::CSR-1 IP from *oma-1::gfp* transgenic animals (Figure 3.2E and F), and were not enriched in the FLAG::WAGO-9/HRDE-1 IP (Figure S3.2A and B). For comparison we also performed IP studies in a *gfp::cdk-1(RNAe)* strain. As expected, we found a reciprocal relationship in this silent strain; 22G-RNAs targeting *gfp* were depleted (3.35-fold) in the FLAG::CSR-1 IP relative to input (Figure S3.2C and E), and were enriched (1.75-fold) in the FLAG::WAGO-9/HRDE-1 IP (Figure S3.2D).

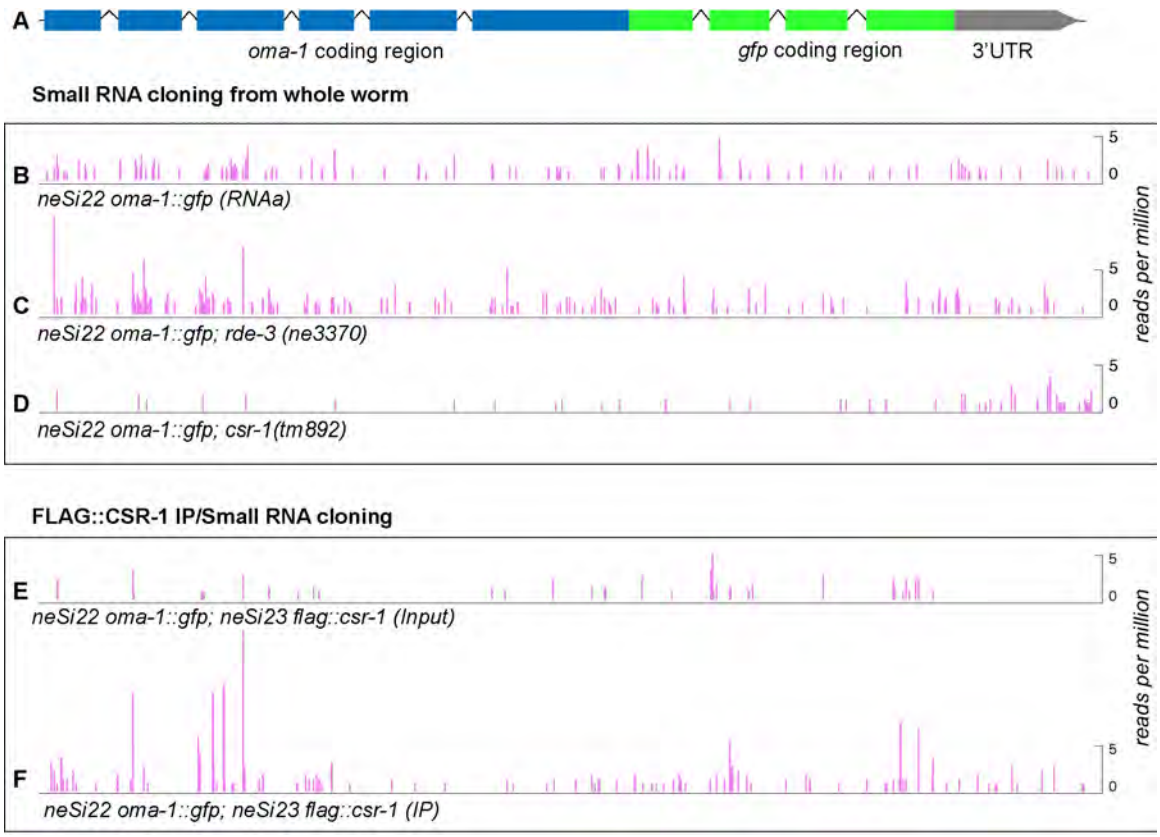
Thus we have shown that in three small-RNA Seq libraries independently prepared from *csr-1(+)* animals, 22G-RNAs targeting *gfp* were present at levels similar to CSR-1 22G-RNAs targeting the *oma-1*-derived portion of the RNAa transgene. Furthermore, we have shown that these *gfp* 22G-RNAs were depleted in *csr-1* mutants and were enriched in the CSR-1 IP. In contrast, an RNAe transgenic strain exhibited *gfp* 22G-RNAs that were enriched in the WAGO-9 IP and were depleted in the CSR-1 IP. Finally, a strain with a neutral transgene (sensitive to RNAe) exhibited very low levels of *gfp* 22G-RNAs relative to the levels of CSR-1 22G-RNA targeting the endogenously-derived portion of the transgene (Shirayama et al., 2012). Taken together, these findings indicate that the RNAa activity of *oma-1::gfp* correlates with the accumulation CSR-1 22G-RNAs targeting the foreign, *gfp*, sequences of the transgene.

Figure 3.2. CSR-1-associated small RNAs targeting GFP in *neSi22 oma-1::gfp(RNAa)*.

(A) Schematic of *oma-1::gfp* transgene. The exon-intron structure is indicated with boxes and lines, respectively.

(B-F) Plots showing the density of antisense small RNAs mapping along *oma-1::gfp* in wild-type (B) and mutant strains *rde-3* (C) and *csr-1*(D). In (E and F) the histograms show read densities of small RNAs obtained from the same lysate before (Input) and after FLAG::CSR-1 Immunoprecipitation (IP). The height of each peak corresponds to the number of RNA reads that begin at that position per million total reads.

Figure 3.2. CSR-1-associated small RNAs targeting GFP in *neSi22 oma-1::gfp (RNAa)*.



Multi-generational exposure to RNAa can gradually license an RNAe allele

The above findings indicate that *C. elegans* transgenes can adopt at least three different states: i) a dominant-acting trans-silencing state (RNAe); ii) a neutral, expressed state that is sensitive to trans-silencing; and (iii) a dominant trans-activating state (RNAa). Previous studies have shown that an RNAe allele can transfer the silent state to a neutral allele. We therefore wished to know whether transient exposure to an RNAa allele could stably activate (or license) the expression of an RNAe allele. To explore this possibility, we set up a series of crosses between an RNAa transgene and a number of distinct RNAe transgenes. After establishing the double transgenic lines, we outcrossed the strains to wild-type to separate the two transgenes again and then monitored expression and RNAa or RNAe status. We found that different transgenes behaved differently in these crosses. For example, *gfp::cdk-1(RNAe)* was activated in the presence of *oma-1::gfp(RNAa)* (Figure 3.3A), but was immediately silenced after crossing away the RNAa transgene (Figure 3.3B) (Shirayama et al., 2012). In contrast, a *cdk-1::gfp(RNAe)* allele remained stably expressed after transient exposure to the RNAa transgene (Shirayama et al., 2012). Finally, a *gfp::csr-1(RNAe)* transgene was never activated upon exposure to *oma-1::gfp(RNAa)*. Instead, each allele maintained its expression status in the double homozygote – silent *gfp::csr-1(RNAe)* and active *oma-1::gfp(RNAa)* (data not shown).

We next wanted to explore whether prolonged exposure to RNAa could influence the tendency of *gfp::cdk-1* to revert back to an RNAe status. Consistent with this idea, after propagating the *oma::gfp; gfp::cdk-1* double transgenic strain for 10 generations and then outcrossing to wild-type to separate the two transgenes, we found that the *gfp::cdk-1(+)* transgene remained expressed for one full generation after separation before re-silencing. Interestingly, the period of sustained expression increased to nearly 10 generations when *gfp::cdk-1* and *oma-1::gfp(RNAa)* were separated after 30 generations of co-propagation (Figure 3.3C and D). However we found that, even though expression of the formerly RNAe transgene was stabilized by long-term exposure to RNAa, the RNAa status was not transferred. Instead, the activated transgene remained sensitive to silencing when exposed through a genetic cross to *gfp::csr-1(RNAe)* (100%, n=24). Taken together these findings suggest that an RNAa transgene can, over time, influence the epigenetic stability of an RNAe allele. However, the transfer of RNAa status is either very slow or depends on other factors that remain to be identified.

Figure 3.3. RNAa counteracts Piwi-dependent silencing and acts over multiple generations to establish an active epigenetic gene-expression state.

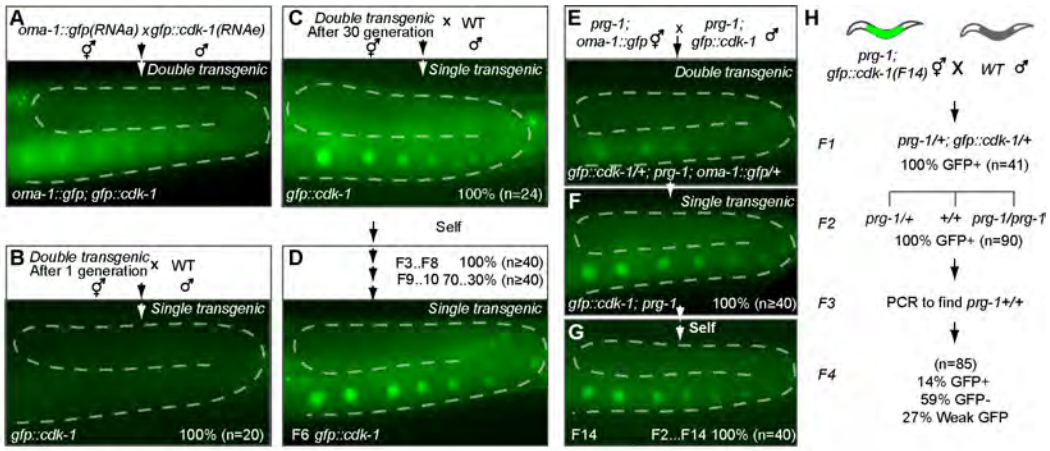
(A-H) Genetic crosses with corresponding epifluorescence images showing representative germlines of resulting progeny. The percentages of animals expressing *gfp::cdk-1* (nuclear GFP signal) at each generation and the number of animals scored 'n' are indicated.

(A-D) Analysis of RNAa exposure on the durability of gene activation in wild-type animals. Newly trans-activated F2 double transgenic animals (A), were outcrossed to wild-type (WT), either immediately or after propagating as a double transgenic strain for 30 generations (F30), to obtain *gfp::cdk-1* "single transgenic" animals shown in B and C, respectively. Siblings of animals shown in (C) were allowed to produce self progeny (D) for multiple generations, and GFP fluorescence was scored in each generation as indicated.

(E-H) Analysis of the genetic influence of Piwi (*prg-1*) on transactivation. RNAa and RNAe transgenes established in a wild-type background were crossed into *prg-1* prior to conducting the trans-activation assay shown in (E). After one generation, *oma-1::gfp* was segregated away to yield *gfp::cdk-1* single transgenic animals assayed in (F). Siblings of animals shown in (F) were allowed to produce self progeny for multiple generations and GFP fluorescence was scored in each generation (G) as indicated. In (H) *gfp::cdk-1*was outcrossed from the *prg-*

1(tm872) mutant background and the animals were scored for GFP in subsequent generations as indicated.

Figure 3.3. RNAa counteracts Piwi-dependent silencing and acts over multiple generations to establish an active epigenetic gene-expression state.



RNAa counteracts PRG-1-dependent silencing

The PIWI Argonaute PRG-1 is required for the initiation of RNAe, but not for the maintenance of silencing (Shirayama et al., 2012). We therefore wondered whether PRG-1 activity is required to re-initiate silencing of an RNAe transgene after trans-activation. To test this possibility, we first crossed the *gfp::cdk-1(RNAe)* and *oma-1:gfp(RNAa)* transgenes into the *prg-1(tm872)* mutant background. As expected, we found that each transgene, singly, maintained its silent or active expression state in the *prg-1* mutant background. We then repeated the trans-activation crosses by mating these *prg-1* mutant strains (Figure 3.3E). As observed in the wild-type *prg-1(+)* background, the *gfp::cdk-1(RNAe)* transgene was activated in the F1 cross progeny (Figure 3.3E). We then allowed the two transgenes to segregate from one another. Strikingly, we found that 100% of the F2 through F14 *gfp::cdk-1* transgenic animals examined maintained expression in the absence of the *oma-1::gfp(RNAa)* transgene (Figure 3.3F and 3.3G). Thus in the absence of *prg-1* activity the RNAa allele is not required to maintain the activated status of the formerly RNAe transgene. We next crossed these actively expressing *prg-1* mutant transgenic animals to wild-type to restore *prg-1* activity. We found that, once homozygous for *prg-1(+)* activity, 85% of the animals examined (n=85) exhibited re-silencing of the transgene by the F4 generation (Figure 3.3H). These findings indicate that *prg-1* is required to re-initiate silencing on an RNAe transgene, and that RNAa opposes this PRG-1 silencing activity.

DISCUSSION

A genome-wide mechanism for the epigenetic adaptation of gene expression

The term epigenetics is used to describe many diverse types of biological events, ranging from the activity of prions (Halfmann and Lindquist, 2010), to the transmission of heritable membrane structures (Harold, 2005), and extending even to cellular differentiation events (Goldberg et al., 2007). In a recent review, Adrian Bird (Bird, 2007) suggested a compelling definition for chromatin-focused epigenetic events as “the structural adaptation of chromosomal regions so as to register, signal or perpetuate altered activity states”. A key element of this definition is that epigenetic chromatin marks are seen as responsive and adaptive: they help to canalize and buffer gene expression programs that may have more direct upstream triggers. Our findings are consistent with this adaptive view of epigenetic programming. They suggest how Argonaute small RNA pathways can work in concert with chromatin pathways to create heritable binary signals that communicate a memory of germline gene-expression from one generation to the next. In this system, small RNAs can both perpetuate expression states in cis and signal adapted gene-expression states to dispersed alleles of a gene.

In this work we focused on the role of Argonaute-small-RNA pathways in the control of transgene expression states. Yet these Argonaute pathways also act globally in the germline to target expressed (CSR-1-targeted) and silenced

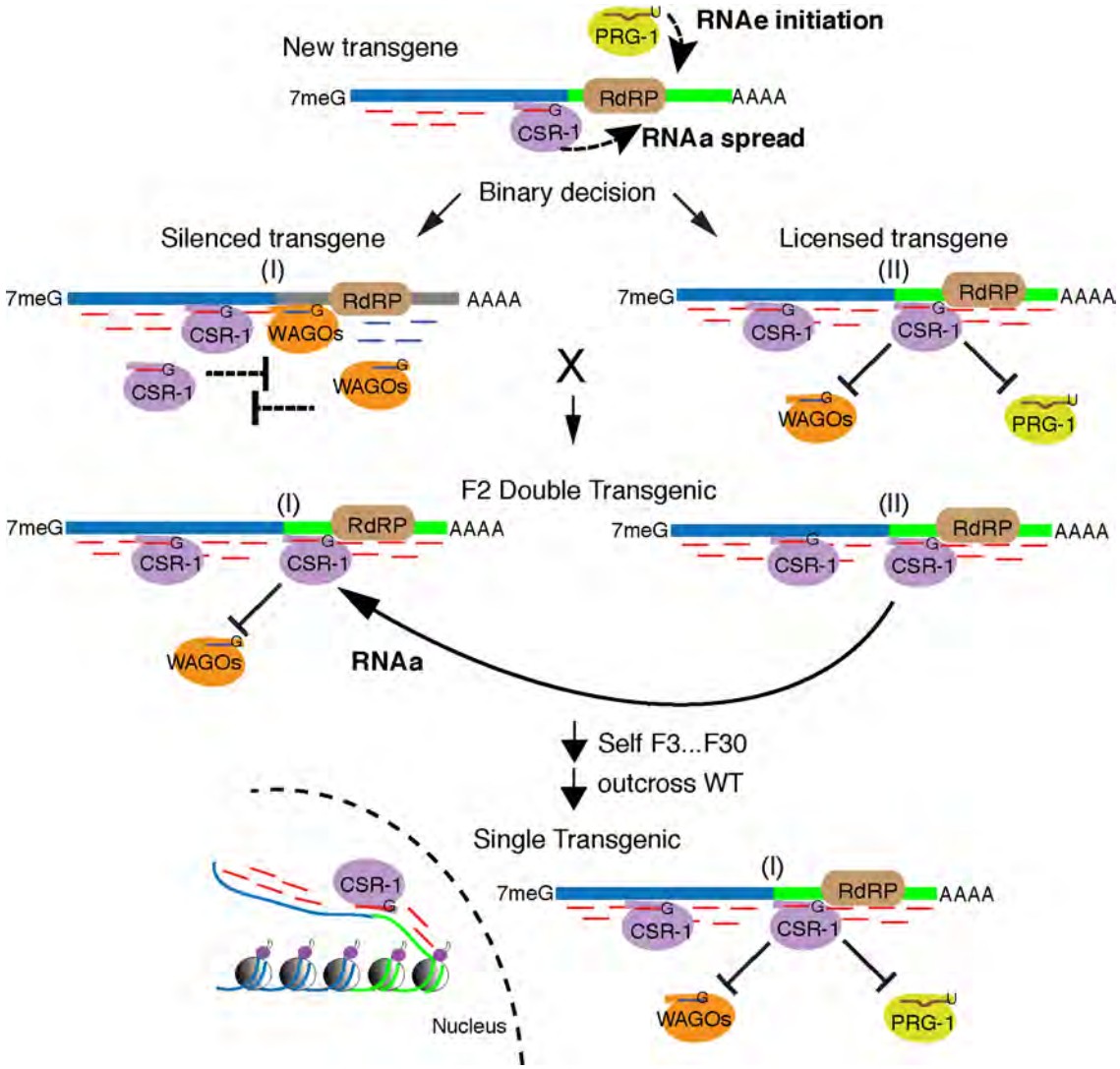
(WAGO-targeted) genes genome wide. A parallel paper by (Conine et al., 2013) shows that CSR-1 is required to promote the expression of many male-specific germline genes. In the absence of paternal CSR-1 activity, males are initially fertile, but progressively become sterile over a period of 5 to 6 generations. This “germline mortal” phenotype is consistent with previous work on the loss of specific Argonaute-silencing pathways (Buckley et al., 2012) and may reflect a gradual loss of the “adapted” epigenetic state after the reinforcing activities of the small RNA pathways are lost.

Studies on *prg-1* mutants suggest that the default state for transgene expression is “ON”. Therefore a simple model for the CSR-1 pathway is that it prevents the incursion of silencing signals within its targeted sequence domain (Model, Figure 3.4). It is possible that CSR-1 prevents PRG-1 and WAGO silencing by using its slicer activity to destroy template RNAs engaged in RdRP transcription and WAGO loading. Understanding the mechanistic details of RNAa will require further exploration of how chromatin and small RNA pathways change as alleles switch from a silenced to expressed status, and will also require new tools for directly intervening in the feed-forward Argonaute and chromatin pathways. A recent study describes one such tool, a tethering system that recruits CSR-1 to target sequences through direct RNA binding, thus activating an RNAe allele without the need for a transactivating allele and its cognate small RNAs (Wedeles et al., 2013b).

Figure 3.4. Model for transactivation by CSR-1 Argonaute.

See Discussion for details.

Figure 3.4. Model for transactivation by CSR-1.



An innate sequence-specific genome-defense mechanism

The findings described here support a model for genome defense that employs a truly surprising strategy – one that permits a rapid "innate" and yet sequence-specific response without the need for prior exposure to a pathogenic sequence or for structural triggers of pathogen-specific activity such as the expression of long dsRNA. Instead, our findings suggest that the recognition of foreign sequences in *C. elegans* depends directly on the Piwi pathway, which scans for foreign sequences (Ashe et al., 2012; Lee et al., 2012; Shirayama et al., 2012), and indirectly on the CSR-1 pathway, which protects endogenous germline-expressed genes from piRNA-mediated recognition. Thus sequence specificity is achieved, not by capturing and remembering foreign sequences as in some systems (Khurana et al, 2011; Sorek et al., 2008), but rather by remembering all self sequences, thereby permitting the innate recognition of foreign sequences (Model, Figure 3.4).

Under some circumstances foreign sequences appear to be adopted as self. One possible model for this adoption process is that CSR-1 recognition can spread, in cis, from fused endogenous sequences within a transgene (Model, Figure 3.4). Targeting by CSR-1 within the endogenous sequences could promote the local recruitment of RdRP, leading to the de novo synthesis of CSR-1 22G-RNAs within the adjacent foreign sequences. Molecular spreading of this type has been observed in gene silencing in both plants and animals (Axtell et al., 2006; Pak and Fire, 2007; Sijen et al., 2001; Sijen et al., 2007). The decision

to silence or license a newly introduced transgene would then be determined through a competition between cis-spreading of CSR-1 recognition and initial recognition by the PRG-1/21U-RNA pathway (Model, Figure 3.4). For some transgenes, such as *oma-1::gfp(RNAa)*, this process leads to the "adoption" of the foreign sequences (through acquisition of CSR-1 targeting) permitting these *gfp* sequences to trans-activate homologous transgenes (Model, Figure 3.4).

CONCLUSION

Epigenetic pathways are diverse and can differ widely from organism to organism. This is particularly true with Argonaute pathways, which exhibit evidence of extensive gene duplication and pathway diversification in both plants and animals (Cerutti and Casas-Mollano, 2006; Ghildiyal and Zamore, 2009). The rapid evolution of these pathways could reflect selective pressure exerted in response to their targets, which in most organisms include a striking genomic load of transposons. While the details may differ from one system to another, the concepts revealed in one organism will likely be relevant in other systems. For example, it is now clear that a dynamic interplay between Argonaute/small RNA pathways and chromatin modifiers is involved both in the silencing of repetitive gene families and in essential chromosome functions such as kinetochore assembly and chromosome segregation in organisms as diverse as fungi, plants and animals (Grewal and Elgin, 2007).

Here we have shown that *C. elegans* employs Argonautes to protect expressed genes from silencing. Interestingly, while the interaction between an

RNAa allele and an RNAe allele resulted in a rapid reversal of the silenced state, conversion of the formerly silent allele to a state permissive of independent sustained gene expression required dozens of generations of continuous exposure to RNAa. This slow conversion of the RNAe allele is consistent with the adaptive definition of an epigenetic process (Bird, 2007) and could reflect a gradual elimination of either small RNAs or of chromatin marks that can stimulate re-silencing (or possibly a slow accumulation of chromatin marks that enforce expression). CSR-1 localizes on chromatin and immunoprecipitates with target DNA sequences (Claycomb et al., 2009). Thus CSR-1 could influence chromatin directly perhaps by engaging nascent transcripts at target genes. It will be interesting in the future to determine whether CSR-1 actively recruits chromatin modifiers to promote gene expression. Furthermore, CSR-1 and members of the WAGO family are abundantly expressed in both oocytes and mature sperm (Claycomb et al., 2009; Conine, 2013; Conine et al., 2010; Gu et al., 2012b; Shirayama et al., 2012). Germline transmission of these Argonautes and their associated small RNAs may thus have genome-wide effects on epigenetic inheritance with potentially significant evolutionary implications.

EXPERIMENTAL PROCEDURES

Genetics

The *C. elegans* strains used in this study were derived from the Bristol N2 strain and cultured as described (Brenner, 1974). Strain WM288 contains a single-copy *oma-1::gfp* transgene that was created using the MosSCI heat shock protocol combined with ivermectin selection as described previously (Shirayama et al., 2012) .

Small RNA cloning and deep sequencing

Total RNA was extracted from 10 young adult worms (Shirayama et al., 2012). Small RNAs (18 – 40 nucleotides) were gel-purified, treated with TAP to generate monophosphate 5' ends, ligated to 5' and 3' linkers and converted to cDNA (Gu et al., 2009; Shirayama et al., 2012). Illumina adapters were added by PCR (Gu et al., 2009; Shirayama et al., 2012). To clone CSR-1 associated small RNAs, M2 FLAG antibody (Sigma) was used to immunoprecipitate FLAG::CSR-1 from 20 mg of lysate from synchronous adult worms homogenized in a stainless steel dounce (Gu et al., 2009). Small RNAs were extracted from FLAG::CSR-1 immune complexes and processed for deep sequencing as described above. Libraries were sequenced in the UMass Medical School Deep Sequencing Core using an Illumina GAII instrument.

For AGO IP studies the relative enrichment was measured by calculating the (# of antisense GFP reads)/(total # of genome matching antisense reads) in the Input and the IP, and then dividing the two numbers.

Computational analysis

Deep sequencing data were processed and analyzed using Bowtie (version 0.12.7) (Langmead et al., 2009) and custom Perl scripts (Gu et al., 2009; Shirayama et al., 2012). Small RNA reads were mapped to WormBase WS215 and normalized to non-structural RNA reads or to the total number of small RNAs that map antisense to protein coding genes. CSR-1 small RNA targets were defined previously (Claycomb et al., 2009; Gu et al., 2009). All scripts are available upon request.

Microscopy

Transgenic worms expressing GFP were mounted on RITE-ON glass slides (Beckton Dickinson) in the presence of 0.2 mM levamisole. Epi-fluorescence and differential interference contrast (DIC) microscopy were performed using an Axioplan2 Microscope (Zeiss). Images were captured with an ORCA-ER digital camera (Hamamatsu) and Axiovision (Zeiss) software.

Preface to Chapter IV

This chapter provides evidence for the competition between PRG-1 and CSR-1 AGO pathways.

My contributions to this chapter are: Figures 4.1, 4.2, 4.3, 4.4, and 5

Masaki Shirayama contributed: Figures 4.3A-B, F-G

Wen Tang contributed: Figure 4.2G, 4.3F-G, 4.4G

Enzhi Shen contributed: Figure 4.2F, 4.4F

This chapter is a manuscript in preparation:

Seth, M., Shirayama, M., Tang, W., Shen, E., Tu, S., and Mello, C.C. (2013). The *C. elegans* CSR-1 and PRG-1 Argonaute Pathways Compete with Each Other to Regulate Gene Expression (Manuscript in preparation)

CHAPTER IV: The *C. elegans* CSR-1 and PRG-1 Argonaute Pathways Compete With Each Other To Regulate Gene Expression

SUMMARY

PIWI proteins function with PIWI-interacting (pi) RNAs to promote fertility and silence transposons in diverse animals. However, little is known about whether and how piRNAs regulate other germline-expressed mRNAs. Here we use genome editing to create *C. elegans* piRNAs targeting endogenous germline mRNAs. Targeting by these novel piRNAs triggered robust local synthesis of secondary siRNAs by RdRP and partial down-regulation of the target mRNA and protein levels, but failed to trigger epigenetic silencing of endogenous germline targets. Resistance to silencing correlated with the expression level of CSR-1, which engages RdRP-derived small RNAs that target most self-mRNAs. Increasing piRNA targeting or decreasing PRG-1 protein levels had opposing effects on piRNA-initiated silencing. These findings suggest that the opposing activities of Piwi and CSR-1 within the *C. elegans* germline permit mRNAs to sample a broad spectrum of post-transcriptional control.

INTRODUCTION

PIWI-interacting RNAs (piRNAs) are Dicer-independent small RNAs, which specifically associate with the germline Argonaute PIWI (P-element Induced WImpy testis) (Aravin et al., 2006; Girard et al., 2006; Grivna et al., 2006; Lau et al., 2006; Ruby et al., 2006). These are the largest class of small non-coding RNAs, 21-35 nucleotides long, have bias towards 5’ Uridine and are evolutionary conserved in metazoans. Very often piRNAs arise from intergenic repetitive elements in the genome referred as piRNA clusters (Batista et al.,

2008; Brennecke et al., 2007; Das et al., 2008), which contain hundreds and thousands of piRNAs with diverse sequences.

Piwi-piRNA pathway has diverse biological and molecular functions including germline specification, gametogenesis, stem cell maintenance, epigenetic programming, transposon silencing, genome protection, and post-transcriptional regulation of mRNAs (Juliano et al., 2011). The most important functions of piRNAs are to maintain germline fertility and transposon silencing (Juliano et al., 2011). Loss of functional mutation in PIWI protein and their interacting co-factors derepress transposons, resulting in their random insertion into the genome (Kalmykova et al., 2005). In *Drosophila melanogaster* this results in activation of DNA damage proteins resulting in defects in germline, which often lead to infertility (Khurana and Theurkauf, 2010). Nevertheless, not all piRNAs target transposons. In *Tetrahymena* Piwi protein is required to complete its sexual life cycle where a set of Piwi-bound RNAs known as scanRNAs (scnRNAs) target heterochromatin modification to mark genome for elimination very similar to piRNA mediated silencing of metazoan (Chalker and Yao, 2011; Liu et al., 2004b; Mochizuki et al., 2002). In contrast, in *Oxytricha trifallax* piRNAs are remarkable as they serve the opposite function to silencing by preserving maternally expressed sequences rather than eliminating foreign sequences (Fang et al., 2012).

Recent studies indicate that piRNAs could silence transcripts other than transposons. One such study by Goh et al provide evidence for piRNA-directed

cleavage of meiotically expressed protein-coding genes in mouse (Goh et al., 2015). In mouse, it has also been shown that pachytene piRNAs are involved in massive RNA elimination in elongating spermatids (Gou et al., 2014). Surprisingly results from transgene studies in worms have advanced this field and such studies have shown that piRNA targets can be stably silenced over generations (Ashe et al., 2012; Bagijn et al., 2012; Lee et al., 2012; Shirayama et al., 2012). We have previously described a model in which PRG-1 with its genomically encoded piRNAs scan for foreign sequences (e.g., *gfp*) while allowing mismatch pairing with the target mRNA (Shirayama et al., 2012). Upon recognition, PRG-1 recruits RdRP to amplify secondary siRNAs that are loaded on to WAGOs, which maintain and propagate epigenetic silencing (RNAe).

How and when such heritable silencing is initiated is not well understood. Given the diversity of piRNAs, it becomes more challenging to understand piRNA target spectrum in a particular system where piRNA are not perfectly complementary to transposable elements. As piRNAs target while allowing mismatches, and as these sequences are diverse and numerous, they could, in principle, target any sequences (Bagijn et al., 2012; Lee et al., 2012). However this promiscuity of piRNA targeting raises the problem of how self-mRNA avoids silencing from piRNAs. In *C. elegans*, this problem of self-non-self recognition appears to be solved by self-protecting Argonaute CSR-1 (Seth et al., 2013; Wedeles et al., 2013b). This leads into the CSR-1 protection model where CSR-1—an essential Argonaute that is required for fertility and development—uses its

small RNAs as a molecular markers of “self” to counteract silencing by the piRNA pathway (Seth et al., 2013; Shirayama et al., 2012).

Here we show that CSR-1 and PRG-1 Argonaute pathways compete with each other to determine gene expression. We use genome editing to create *C. elegans* piRNAs with new specificities. We show that piRNA can induce localized secondary 22Gs. Interestingly, endogenous germline mRNAs were resistant to piRNA-directed silencing however was partially reduced at mRNA and protein level. Artificially increasing piRNA targeting or artificially decreasing PRG-1 protein levels had opposing effects on gene expression. These findings suggest that opposing activities of PRG-1 and CSR-1 within the *C. elegans* germline permit mRNAs to sample a broad spectrum of post-transcriptional control.

RESULTS

Balanced silencing and activation signals reveal that PRG-1 opposes CSR-1 to maintain transgene silencing.

Transgenes inserted in different chromosomal regions experience distinct local effects that determine their expression status (Frokjaer-Jensen et al., 2016). We therefore explored how silent transgenes inserted in different locations respond to an activating transgene. Interestingly, we found that two silent transgenes *gfp::csr-1* and *gfp::cdk-1* when inserted on LGII at location, *ttTi5605*, were activated in crosses with *oma-1::gfp*, but were not activated when inserted at location *cxTi10882* on LGIV. For example, there was no transactivation of *gfp::csr-1* on LGIV in the presence of *oma-1::gfp* on LGII in the F2 cross progeny

(n>24) (Figure 4.1A) and each of the transgenes maintained their status, OFF and ON, respectively.

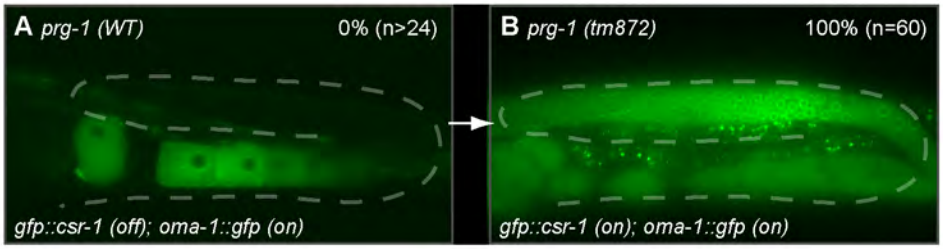
The balanced state observed in the above crosses was surprising as it suggested that *gfp-targeted* WAGO 22G-RNAs were sufficient to silence the *gfp::csr-1* mRNA but were unable to silence *gfp* target sequences in the *oma-1::gfp* mRNA within the same cells. Conversely, *gfp-targeted* CSR-1 22G-RNAs from *oma-1::gfp* could apparently protect *oma-1::gfp* mRNA but could not activate *gfp::csr-1* mRNAs. To further explore this phenomenon we first wished to confirm that WAGO small-RNAs were still required for maintaining *gfp::csr-1* silencing in this strain. To do this we crossed in an *rde-3* mutant, deficient in the WAGO 22G-RNA pathway. Consistent with the idea that WAGO silencing is required to maintain *gfp::csr-1* silencing in this strain we found that GFP::CSR-1 expression was fully restored in the *rde-3* mutant (n>20). We previously showed that PRG-1 activity is required to initiate, but not to maintain, transgene silencing (Shirayama et al., 2012). However, when a transgene was de-silenced by exposure to an RNAa transgene such as *oma-1::gfp*, re-silencing of the transgene required PRG-1(+) activity (Seth et al., 2013). We therefore wished to ask if the balanced state of the *gfp::csr-1; oma-1::gfp double* transgenic strain might reflect a constant re-silencing of *gfp::csr-1* via the PRG-1 piRNA pathway. To explore this possibility we first confirmed that *gfp::csr-1* single transgenic strain was not activated by crossing them into a *prg-1(tm872)* mutant background (data not shown). We next crossed the OMA-1::GFP (ON); GFP::CSR-1 (OFF)

strain into *prg-1(tm872).* Consistent with the idea that PRG-1 activity prevents the *oma-1::gfp*-dependent transactivation of *gfp::csr-1* we observed robust GFP::CSR-1 expression in 100% of the double transgenic in *prg-1(tm872)* homozygotes analyzed (n=60) (Figure 4.1B). This finding suggests that PRG-1 activity can reinforce WAGO dependent silencing to prevent CSR-1-dependent transactivation.

Figure 4.1. Balanced silencing and activation signals

(A, B) Epifluorescence image of representative germline (outlined with dashes) in F2 and F3 generations. The cytoplasmic fluorescence signal is OMA-1::GFP; the P-granule signal is GFP::CSR-1. The percentages indicate the number of animals that exhibited the expression of GFP::CSR-1 in the wild-type and *prg-1* mutant.

Figure 4.1. Balanced silencing and activation signals



Transgenes differ in their resistance to piRNA targeting

The above findings suggest that piRNAs contribute to the silencing of *gfp::csr-1,* but either do not target or cannot silence *oma-1::gfp*. Previous studies have shown that engineering piRNA-complementary sequences into a transgene 3' UTR can result in a local induction of 22G-RNA and silencing of the transgene (Bagijn et al., 2012; Lee et al., 2012). As an alternative approach we chose to use CRISPR homologous recombination (Kim et al., 2014) to replace several abundantly expressed piRNA sequences with antisense sequences targeting *gfp* named as *21ur-antigfpx1*. We then crossed worms bearing gfp-antisense piRNAs with animals expressing either *cdk-1::gfp* or *oma-1::gfp* transgene. As expected, we found that *cdk-1::gfp* was rapidly silenced after crossing to the *21ur-antigfpX1* strain (n>20). However, interestingly, the *oma-1::gfp* transgene remained expressed even with the *21ur-antigfpX1* gene was homozygous in the strain (n>20). Next, we crossed the two homozygous *21ur-antigfpX1* piRNA transgenic strains together to generate a *cdk-1::gfp, oma-1::gfp* double transgenic strain that is also homozygous for the *21ur-antigfpx1* piRNA. Strikingly, we found that this double transgenic strain not only failed to exhibit the transactivation of CDK-1::GFP, but also exhibit silencing of the OMA-1::GFP (Figure 4.2A). These findings suggest that the addition of a single *21ur-antigfpx1* piRNA, although it cannot silence *oma-1::gfp* by itself, can shift the balance of small RNA signals in the germline sufficiently to abolish the RNAa activity associated with *oma-1::gfp* and render the transgene sensitive to RNAe.

We then compared small RNAs profile generated from *gfp* targeting piRNAs on *oma-1::gfp* and *cdk-1::gfp* transgenic worms (Figure 4.2 B and C). We found that there were 22G generated next to *21ur-antigfpX1* piRNA sequence in both transgenic strains, however they were much more abundant in *cdk-1::gfp* transgene as compared to *oma-1::gfp* transgene (Figure 4.2B and C). Strikingly, there was obvious spreading of 22Gs in the *gfp* region of the *cdk-1::gfp* transgene where as there was little to no spreading of 22Gs in *gfp* region of the *oma-1::gfp* transgenic worms (Figure 4.2B and 4.2C). Single nucleotide resolution of 22G peaks near the *21ur-antigfpX1* piRNA revealed that there are many more and increased number of 22Gs targeting *cdk-1::gfp* as compared to *oma-1::gfp* (Figure 4.2 D and 4.2E). This again suggests that for some reason *oma-1::gfp* is adopted as a self transgene and is resistant piRNA targeting. When we compared GFP/*gfp* protein and mRNA levels in these piRNA-targeted strains, we found that the protein and mRNA levels of these targets were reduced (Figure 4.2 F and 4.2G). These finding suggests that although piRNA targeting *oma-1::gfp* does not silence the transgene, it still reduce *oma-1::gfp* transcripts and protein levels.

Figure 4.2. Transgenes differ in their resistance to piRNA targeting

(A) Analysis of the genetic influence of piRNA (*21ur-antigfpX1*) targeting. CDK-1::GFP (OFF) and OMA-1::GFP (ON) with the *21ur-antigfpX1* were crossed and the GFP fluorescence of cytoplasmic OMA-1::GFP was scored in the homozygous siblings (n ≥ 20). The worm cartoon represents the Germline GFP off or on for their respective transgenes.

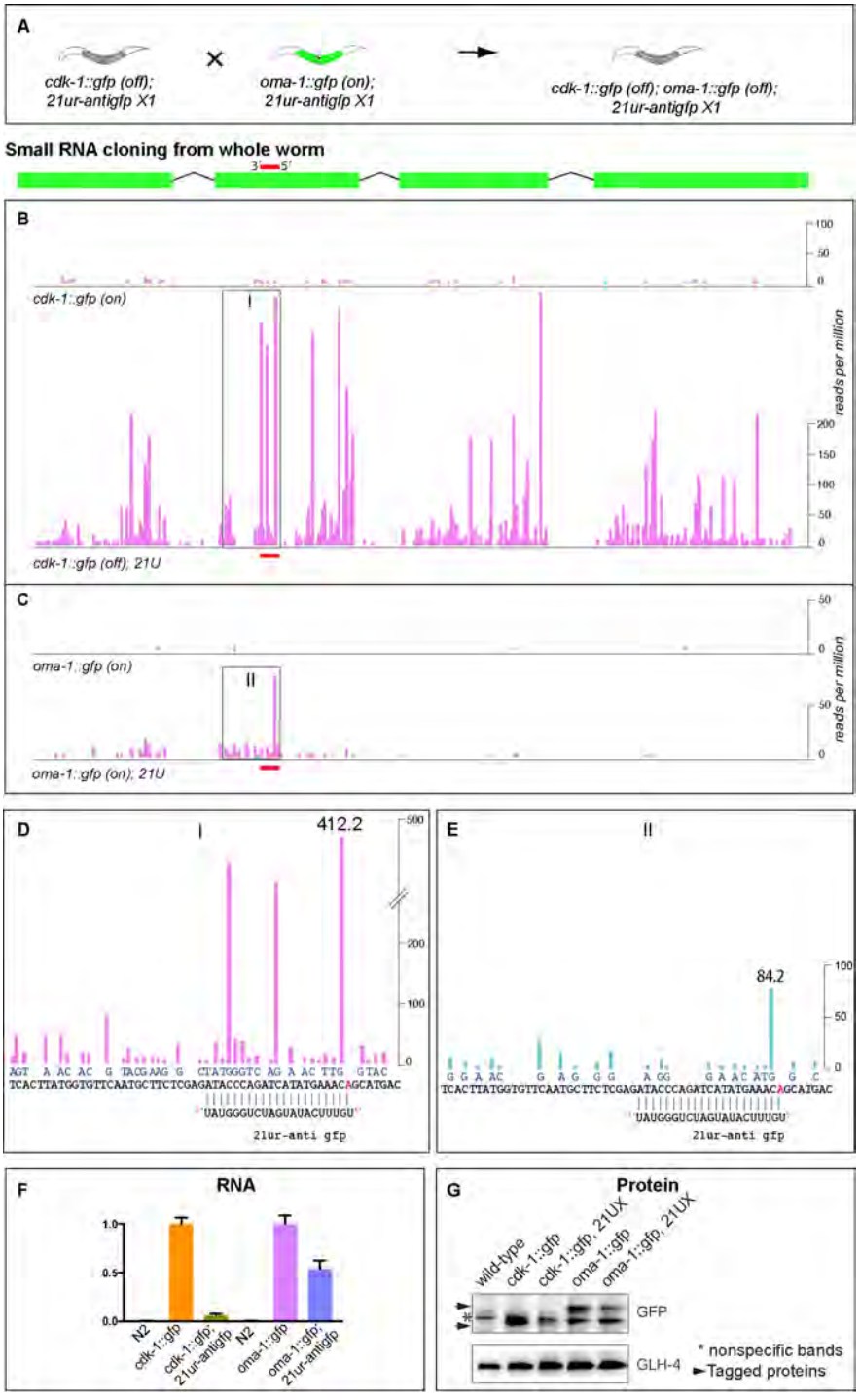
(B-C) Schematic showing density of 22G-RNAs targeting *gfp* in *cdk-1::gfp* and *oma-1::gfp*. Plots showing the density of antisense small RNAs mapping along *gfp* in CDK-1::GFP (ON) in wild-type and CDK-1::GFP in the presence of (*21ur-antigfpX1)* (B) and OMA-1::GFP (ON) in wild-type and OMA-1::GFP (ON) in the presence of (*21ur-antigfpX1)* (C). The height of each peak corresponds to the number of RNA reads that begin at that position per million total reads.

(D-E) Single nucleotide resolution of antisense small RNAs. The position of *21ur-antigfpX1* RNA that base pair to the *gfp* sequence are indicated in the gene diagram as well as the density of 22G-RNAs at single-nucleotide resolution.

(F) qRT-PCR analysis of *cdk-1::gfp*-RNA and *oma-1::gfp*-RNA from total RNA prepared from different transgenic strains (as indicated). Error bars represent the standard deviation for three replicates in one experiment.

(G) Analysis of protein expression in wild-type and transgenic strains (as indicated). The blot was probed with anti-GFP and anti-GLH-4 antibodies (as indicated).

Figure 4.2. Transgenes differ in their resistance to piRNA targeting



Oma-1 coding sequence confer RNAa activity

We next wished to explore why the *oma-1* transgene is more resistant to silencing than the *cdk-1* gene. The above findings suggest that one reason could be less targeting by piRNAs within the *oma-1* mRNA. However, a recent study suggests that the context of certain AT-rich non-coding sequences within and around a gene (Frokjaer-Jensen et al., 2016) may render the gene resistant to silencing. To explore the possibility that associated non-coding sequences explain the resistance of *oma-1* transgenes to silencing we created an *oma-1* construct and replaced the promoter and 3'UTR of *oma-1::gfp* transgene with sequences from the *cdk-1* locus (Figure 4.3A). We also created a reciprocal transgene in which the *oma-1* promoter and 3'UTR were appended to the *cdk-1::gfp* or *gfp::cdk-1* transgene. We then introduced these transgenes to monitor their expression. We found that the transgene containing *gfp::cdk-1* with the *oma-1* promoter and 3'UTR was rapidly silenced upon single-copy integration, n≥5, and thus behaved no differently than the same construct with cdk-1 promoter and 3'UTR. Conversely, the transgenes containing the *oma-1::gfp* with the *cdk-1* promoter and 3'UTR was expressed in all n≥5 integrated lines analyzed. Furthermore, in crosses with a silent *gfp::cdk-1* transgene these oma-1 open reading frame were sufficient to drive transactivation of GFP::CDK-1 (n=24). In addition, we also made *oma-1* transgene without intron sequence and then introduced it *de novo* to test it for RNAa activity of *oma-1::gfp (cDNA).* We show that coding sequence of *oma-1* was expressed under its promoter and was

sufficient to drive transactivation (Figure 4.3B). All the above finding suggests that the expression of *oma-1::gfp* and its transactivation was independent of AT-rich mer coding sequence surrounding the transgene (Frokjaer-Jensen et al., 2016).

To further explore what aspects of the *oma-1* coding region were responsible for its ability to license GFP expression we performed two experiments. First, we altered the codons within the open reading frame to maintain the protein sequence while maximizing differences in the nucleic acids sequence. Secondly, we frame-shifted the open reading frame and introduced 11 single nucleotide substitutions to remove stop codons from the second reading frame, thus generating an entirely novel protein that nevertheless maintained a nucleotide sequence nearly identical to *oma-1* (Figure 4.3B). These transgenes were then introduced and monitored for their expression. We found that neither of these transgenes generated a visible GFP signal. We next crossed these transgenes with a strain expressing a nuclear CDK-1::GFP. We found that the codon altered transgene exhibited RNAe activity, inducing the silencing of the CDK-1::GFP (Figure 4.3C). However the frame-shifted transgene did not cause silencing. We reasoned that the non-optimal codon bias of the frame-shifted transgene might prevent stable protein production, and so we monitored mRNA levels by Northern blotting. This analysis revealed that the frame-shifted mRNA was expressed, while, as expected the codon-altered mRNA (Figure 4.3D) was not expressed. Moreover, and consistent with its RNAe status, the expression of

the codon altered mRNA was restored upon CRISPR *rde-3* mutant which is defective in WAGO- 22G-RNA expression required for silencing. We next asked if the frame-shifted mRNA was sufficient to confer transactivation on a silent *gfp::cdk-1* transgene. Strikingly, all the F1 cross progeny analyzed restored GFP::CDK-1 expression in this assay (Figure 4.3E). Further small RNA seq analysis from the codon altered *oma-1::gfp* in the WT and *rde-3* mutant transgenic worms demonstrate that the silent version of codon altered *oma-1::gfp* has accumulation of novel and many more WAGOs 22Gs as compared to one in the *rde-3* mutant background (Figure 4.3F and G). These findings suggest that the coding region but not the reading frame or coding potential of the *oma-1* gene is sufficient for trans-activation. In addition, for the first time we were able to show that a “self” protein encoded by “non-self” DNA is recognized as “non-self”.

Figure 4.3. *Oma-1* coding sequence confer RNAa activity

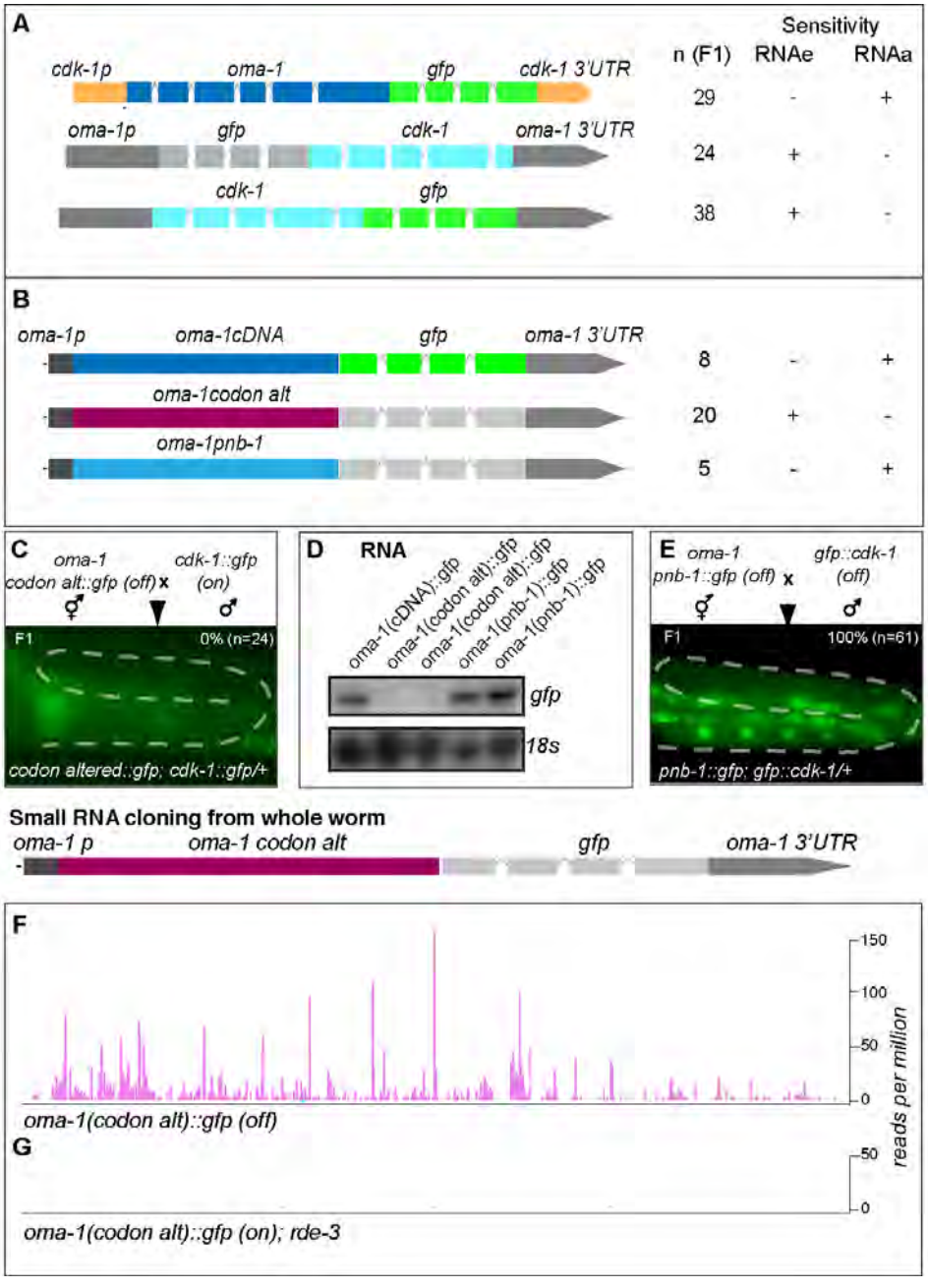
(A and B) Schematic of the respective transgenes (as indicated). The exon-intron structure is indicated with boxes and lines, respectively. n represents the number of F1 cross progeny scored in the transactivation assay while crossing it to off or on transgene. Sensitivity to RNAe represent if the transgene was able to silence or was silenced by other RNAe transgene when scored in the transactivation assay. Sensitivity to RNAa represent if the transgene was able to transactivate another silent transgene in the transactivation assay. *pnb-1* refers to frame shifted *oma-1*.

(C and E) Epifluorescence image of representative germline (outlined with dashes) in F1 cross progeny. The nuclear fluorescence signal is GFP::CDK-1 that is OFF in (C) and ON in (F). The percentages indicate the number of animals that exhibited the expression of GFP::CDK-1 in the F1 cross progeny.

(D) Northern blot analysis of *oma-1::gfp*-RNA using *gfp* probe 168 base pair in the first exon of *gfp.* Total RNA was prepared from different transgenic strains (as indicated).

(F and G) Schematic showing density of 22G-RNAs targeting *oma-1(codon alt)::gfp* in wild-type and *rde-3* mutant. Plots show the density of antisense small RNAs mapping along the transgene (as indicated). The height of each peak corresponds to the number of small RNA reads that begin at that position per million total reads.

Figure 4.3. *Oma-1* coding sequence confer RNAa activity



Increasing the piRNA targeting induces *oma-1::gfp* to silence

The above findings suggest that the coding region of *oma-1* is sufficient to protect a *gfp* transgene from silencing. We wondered if this might reflect a relatively lower level of piRNA targeting within this gene than for example within the open-reading frame of *cdk-1*. Although the rules for piRNA targeting are still not known with any precision, our informatics analysis did not provide support for this idea. The frequency of highly matched piRNAs targeting *oma-1* sequences appeared no lower than was observed for *cdk-1*. Because we saw no evidence for a natural lack of piRNA targeting within *oma-1* we decided to use CRISPR to artificially increase piRNA targeting on *oma-1*. To do this we altered the sequences of two abundantly expressed type-1 piRNAs to target the *oma-1* coding sequences. Surprisingly, upon introduction of *oma-1::gfp* into this double *21ur-antioma-1IV* strain, we still failed to see silencing of the *oma-1::gfp* transgene (Figure 4.4A). However, when we crossed in the *21ur-antigfpX1* reporter strain, we found that *oma-1::gfp* was finally silenced upon homozygosing all three perfectly matched piRNA loci over generations (Figure 4.4A). These findings suggest that the open-reading frame of *oma-1* confers resistance to piRNA silencing but that this resistance can be overcome by increasing the number of piRNAs with high complementarity (in this case perfectly complementary to *oma-1* mRNA) sequences. We then looked at the small RNAs associated with these transgenic strains and found that although there were increased number of 22G generated in the *oma-1::gfp* strain that was targeted by

multiple perfect complementary piRNAs as compared to the strain that had no piRNA targeting (Figure 4.4C and 4.4D), however there was no spreading of 22G *in cis* as in case of *cdk-1::gfp* (Figure 4.2B). Interestingly the triple piRNA strain with *oma-1::gfp* was desilenced in the *prg-1(tm872)* mutant background, however *cdk-1::gfp* with *21ur-antigfpX1* was not (data not shown). This suggests that in response to piRNA targeting in *oma-1::gfp*, WAGO 22Gs was generated locally however did not undergo the amplification process to spread throughout the transgene.

We also compared the OMA-1 protein levels blotting with GFP antibody and found that OMA-1::GFP level was reduced in the presence of piRNAs targeting *oma-1* (Figure 4.4F) consistent with our model of PRG-1 targeting the endogenous genes by virtue of piRNAs.

Figure 4.4. Increasing the piRNA targeting induces *oma-1::gfp* to silence

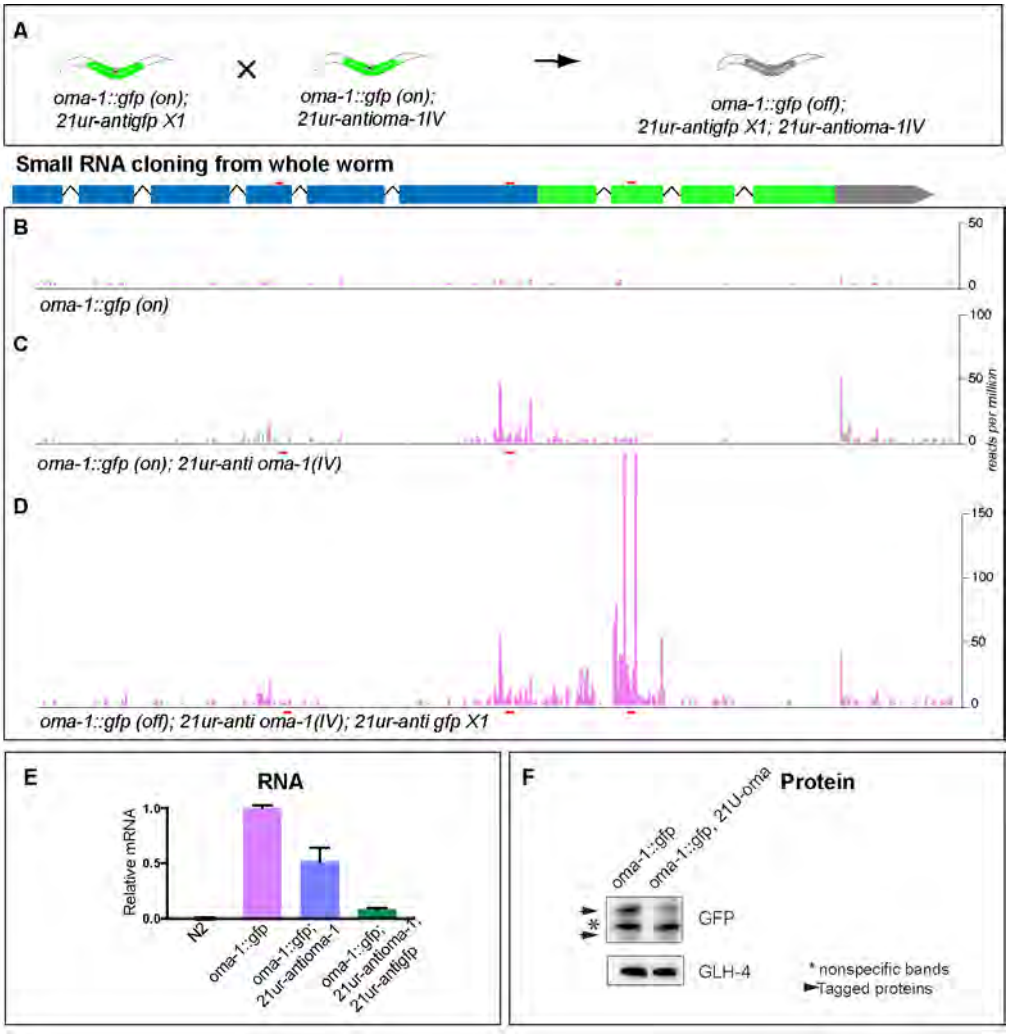
(A) Analysis of the genetic influence of multiple piRNA targeting. Worms bearing OMA-1::GFP (ON); *21ur-antigfpX1* and OMA-1::GFP (ON)*;21ur-antioma-1IV* transgenes were crossed and the GFP fluorescence of cytoplasmic OMA-1::GFP was scored in the homozygous siblings (n ≥ 20). The worm cartoon represents the Germline GFP OFF or ON with their respective genetic background.

(B-D) Schematic showing density of 22G-RNAs targeting *oma-1::gfp*. Plots showing the density of antisense small RNAs mapping along transgene in *oma-1::gfp* in wild-type (B) with *21ur-antigfpX1* in (C) and *21ur-antigfpX1; 21ur-antioma-1IV* in (D). The height of each peak corresponds to the number of small RNA reads that begin at that position per million total reads.

(E) qRT-PCR analysis of *oma-1::gfp*-RNA from total RNA prepared from different transgenic strains (as indicated). Error bars represent the standard deviation for three experimental replicates.

(F) Analysis of protein expression in wild type and transgenic strains (as indicated). The blot was probed with anti-GFP and anti-GLH-4 antibodies (as indicated).

Figure 4.4. Increasing the piRNA targeting induces *oma-1::gfp* to silence



Regulation of endogenous targets by loss of *21ur-x1* piRNA

In order to understand the targeting of piRNAs on endogenous genes, we looked at 22G changes at endogenous genes by the loss of *21ur-x1* piRNA. There were at least 16 genes that had reduced 22Gs levels greater or equal to1.5 fold (Figure 4.5A). Interestingly, all of these target genes were WAGO targets as enriched in the WAGO-1 IP data sets. Since there was reduction in WAGO-22G on these target genes in the absence of *21ur-x1* piRNA (Figure 4.5B), one could hypothesize that this reduction would result in increased mRNA levels. Therefore we designed qRT-PCR primers for two of these targets (*xol-1* and *rde-11)* and compared their expression levels in the wild type and (*21ur-x1* piRNA) mutant background (Figure 4.5C). As indicated, the small RNA profile show reduced 22Gs levels (Figure 4.5B) and the qRT-PCR showed 1.2-2 Fold increase in mRNA levels (Figure 4.5C), thus demonstrating that piRNA can silence the target gene by allowing mismatches and control the endogenous gene expression.

Figure 4.5. Regulation of endogenous targets by loss of *21ur-x1* piRNA

(A) List of genes showing reduced levels of 22G from 1.5-2 fold.

(B) Schematic showing density of 22G-RNAs targeting *xol-1 and rde-11*. Plots showing small RNAs density mapping along these target gene loci. The height of each peak corresponds to the number of small RNA reads that begin at that position per million total reads.

(C) qRT-PCR analysis of *xol-1 and rde-11* from total RNA prepared from WT and *21ur-x1* piRNA mutant strain. Each bar represents the relative mRNA transcript compared to WT and normalized to *csr-1* transcripts. Error bars represent the standard deviation for three experimental replicates.

Figure 4.5. Regulation of endogenous targets by loss of *21ur-x1* piRNA

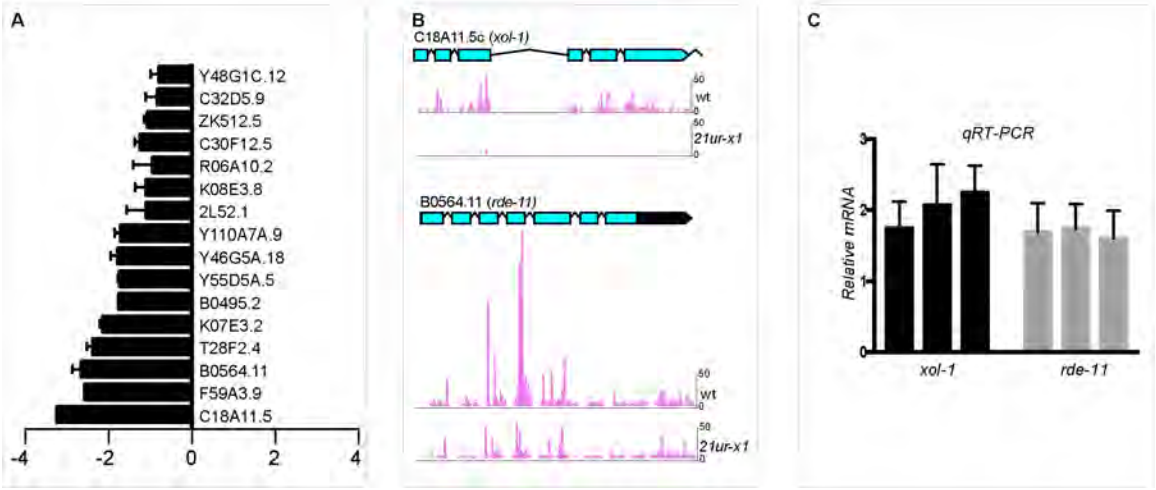
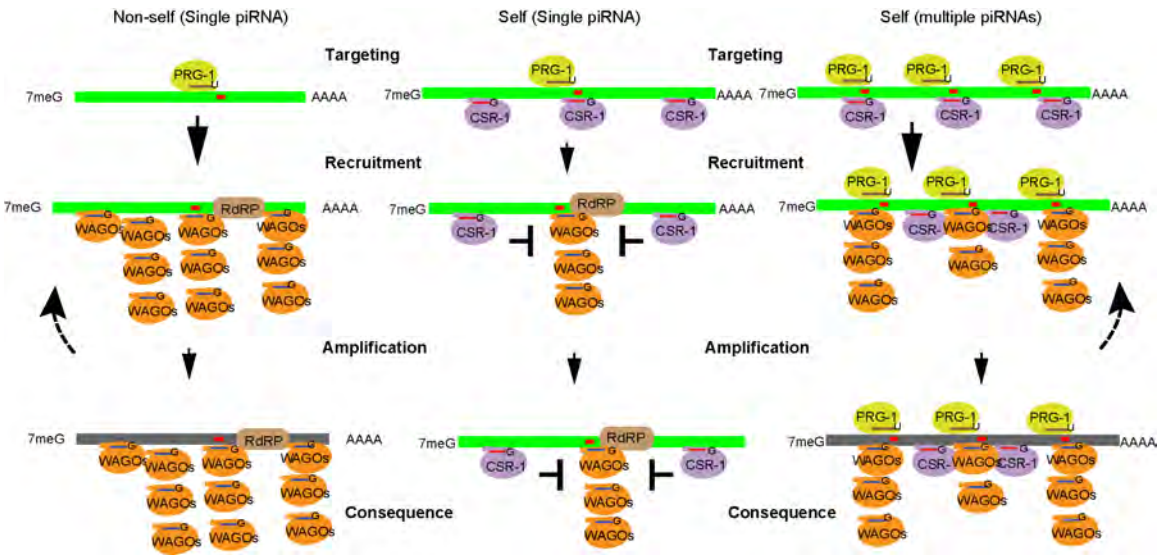


Figure 4.6. Model: PRG-1 regulate endogenous gene by virtue of piRNAs

See Discussion for details.

Figure 4.6. Model: PRG-1 regulate endogenous gene by virtue of piRNAs



DISCUSSION

We have shown that although *cdk-1::gfp* is silenced by a single piRNA that is perfectly complementary to *gfp*, *oma-1::gfp* is not. One reason for this discrepancy might be that *oma-1::gfp* is protected by CSR-1/22G, where as *cdk-1::gfp* is not. Although PRG-1-associated *21ur-antigfpX1* piRNA targets *oma-1::gfp* and recruits RdRP to generate WAGO-22Gs next to the piRNA target site, WAGO-22G are not amplified to cause spreading and silencing of *oma-1::gfp* (Model, Figure 4.6). We can however induce PRG-1-dependent silencing of *oma-1::gfp* by targeting with multiple artificial piRNAs. Nevertheless, silencing is dependent on the presence of the artificial piRNAs, and WAGO-22Gs fail to spread on *oma-1::gfp*. Conversely piRNAs trigger epigenetic silencing of *cdk-1::gfp,* which is maintained by WAGO-22G (Figure 4.2B) and becomes independent of PRG-1 (Shirayama et al., 2012). These findings suggest that piRNAs can repeatedly initiate silencing of CSR-1 target, but cannot maintain silencing by WAGO-22Gs. Perhaps because the catalytic activity of CSR-1 cleaves any of its target mRNA that are recognized by piRNAs and prevents spreading of silencing by WAGO-22Gs.

We have shown that codon altered version for *oma-1* is epigenetically silenced presumably because it is recognized as foreign. Interestingly, the WAGO-22Gs targeting *oma-1(codon alt)::gfp* peak at regular intervals throughout the transcript. It may be possible that the transcript is targeted by piRNAs at regular intervals. It is also possible that the transcript is marked at regular

intervals by a protein(s) that assists PRG-1-dependent recruitment of RDRP to synthesize anti-sense 22G that are loaded onto WAGOs. VASA-related helicases would be good candidates to test because they are known to interact with Argonaute proteins (Shirayama et al., 2014). For example, it would be interesting to see if these peaks disappear in the helicase mutant background or the *prg-1* mutant background.

In the absence of RDE-3—i.e., no memory of foreign sequence—the *oma-1(codon alt)::gfp* appears to be licensed by CSR-1 and adopted as self, acquiring RNAa activity. Although our data suggest that the oma-1 3' UTR is insufficient for RNAa in a wild-type background (Figure 4.3 A), it remains possible that CSR-1 22Gs from the *oma-1* 3'UTR are sufficient to license the *oma-1(codon alt)::gfp* transgene in the absence of RDE-3. If so, we might expect the *oma-1* promoter and 3'UTR could potentially license *gfp::cdk-1* (normally RNAe) in the absence of RDE-3. Conversely, we might also expect that *cdk-1* promoter and 3'UTR will fail to license *oma-1(codon alt)::gfp* because they do not license *gfp::cdk-1* in the absence of RDE-3 (Shirayama et al., 2012).

Replacing the *21ur-X1* piRNA with the *gfp* piRNA resulted in the loss of WAGO-22Gs on ten *21ur-X1* piRNA endogenous targets. Notably *xol-1* is one of them that are targeted by *21ur-X1* and fifteen other piRNA (allowing up to 4 mismatches). *xol-1*, which promotes male development and prevents dosage compensation (Miller et al., 1988; Rhind et al., 1995), is normally repressed in hermaphrodites, but ectopic expression of *xol-1* in hermaphrodites causes low

penetrance lethality (Carmi et al., 1998; Nicoll et al., 1997). There are two questions related to this observation: Does the loss of 21ur-X1 or *prg-1* increase the expression of *xol-1*? Are the phenotypes of *prg-1* mutants related to elevated *xol-1* expression? If so, these findings would indicate a role for piRNAs in sex determination and dosage compensation.

EXPERIMENTAL PROCEDURES

Genetics

The *C. elegans* strains used in this study (see Supplementary Information) were derived from the Bristol N2 strain and cultured as described (Brenner, 1974). Strain WM288 contains a single-copy *oma-1::gfp* transgene that was created using the MosSCI heat shock protocol combined with ivermectin selection as described previously .

Small RNA cloning and deep sequencing

Total RNA was extracted from 40,000 adult worms (Shirayama et al., 2012). Small RNAs (18 – 40 nucleotides) were gel-purified, treated with TAP to generate monophosphate 5’ ends, ligated to 5’ and 3’ linkers and converted to cDNA (Shirayama et al., 2012). Illumina adapters were added by PCR (Gu et al., 2009; Shirayama et al., 2012). Small RNAs were extracted from total RNA and processed for deep sequencing as described above. Libraries were sequenced in the UMass Medical School Deep Sequencing Core using an Illumina GAII instrument.

Computational analysis

Deep sequencing data were processed and analyzed using Bowtie (version 0.12.7) (Gu et al., 2009) and custom Perl scripts (Langmead et al., 2009). Small RNA reads were mapped to WormBase WS215 and normalized to non-structural RNA reads or to the total number of small RNAs that map antisense to protein coding genes. CSR-1 small RNA targets were defined previously (Gu et al., 2009; Shirayama et al., 2012). All scripts are available upon request.

Statistical analyses

Crosses were performed at least 3 times. For multigenerational experiments, at least 10 progeny from each of three or four independent lines were analyzed per generation. Two-tailed P values were calculated using Fisher's exact test with 2x2 contingency table.

Microscopy

Transgenic worms expressing GFP were mounted on RITE-ON glass slides (Beckton Dickinson) in the presence of 0.2 mM levamisole. Epi-fluorescence and differential interference contrast (DIC) microscopy were performed using an Axioplan2 Microscope (Zeiss). Images were captured with an ORCA-ER digital camera (Hamamatsu) and Axiovision (Zeiss) software.

CHAPTER V: Discussion and Future Work

Is RNAa phenomenon evolutionary conserved?

Our findings on RNAe and RNAa phenomenon provide insight into a remarkable multi-Argonaute system that scans the entire transcriptome to distinguish self from non-self nucleic acids and transmit memories of gene expression from one generation to the next via sperm and oocyte. If pathways exist to silence nucleic acids, there ought to be some ways by which an organism can protect self-genes from being wrongly targeted. A key question remains: Is RNAa evolutionary conserved in other organisms and what is the implication of such conservation? For example in *Oxytricha trifallax* piRNAs serve as the function of retention of the maternal genomic regions thus protecting the self-sequences and eliminating the non-self sequences (Fang et al., 2012). Other studies in human cultured cells have implicated small RNAs and/or Argonautes in gene activation, a phenomenon referred to as RNAa (Janowski et al., 2007; Li et al., 2006; Place et al., 2008). In these examples, targeting is thought to occur within the promoter region of the gene, perhaps acting on nascent promoter-derived transcripts, and is correlated with the induction of chromatin marks characteristic of gene activation. Similar studies in human cancer cells have indicated that AGO1 interacts with RNA polymerase II and binds to promoters of actively transcribed genes (Huang et al., 2013). Does this mean that cancer cells can employ RNAa to sustain active expression of genes necessary for growth and survival? Although less common, there are several other examples of small-RNA pathways that appear to activate gene expression. In plants small dsRNAs

have been implicated in the activation of the Petunia pMADS3 homeotic gene and are thought to act by promoting DNA-methylation at a CpG site within an intronic cis-promoter element (Shibuya et al., 2009b). This study indicates that RNA-directed DNA methylation induces transcriptional activation in plants as well.

Mechanism of licensing by CSR-1

In chapter three of my thesis, we have shown that CSR-1 is required for RNAa phenomenon. However the exact mechanism of RNAa remains largely unknown. It is possible that CSR-1 prevents silencing by PRG-1 and WAGOs by using its slicer activity to destroy template mRNAs engaged in RdRP transcription and WAGO loading. Or the RNAa function of CSR-1 is independent of its slicer activity? Understanding the mechanistic details of RNAa will require further exploration of how chromatin and small RNA pathways change as alleles switch from a silenced to expressed status. In other words, does RNAa have initiation and maintenance steps just like RNAe and if yes what are the genes involved in various steps of this process?

Why and how *oma-1::gfp* is adopted as self?

An important question that we are yet to find an answer for is how under some circumstances foreign sequences, such as *oma-1::gfp*, are adopted as self? One possible model for this adoption process is that CSR-1 recognition can spread, *in trans*, from fused endogenous sequences within a transgene (Model, Figure 5.1A). Targeting by CSR-1 within the endogenous sequences could

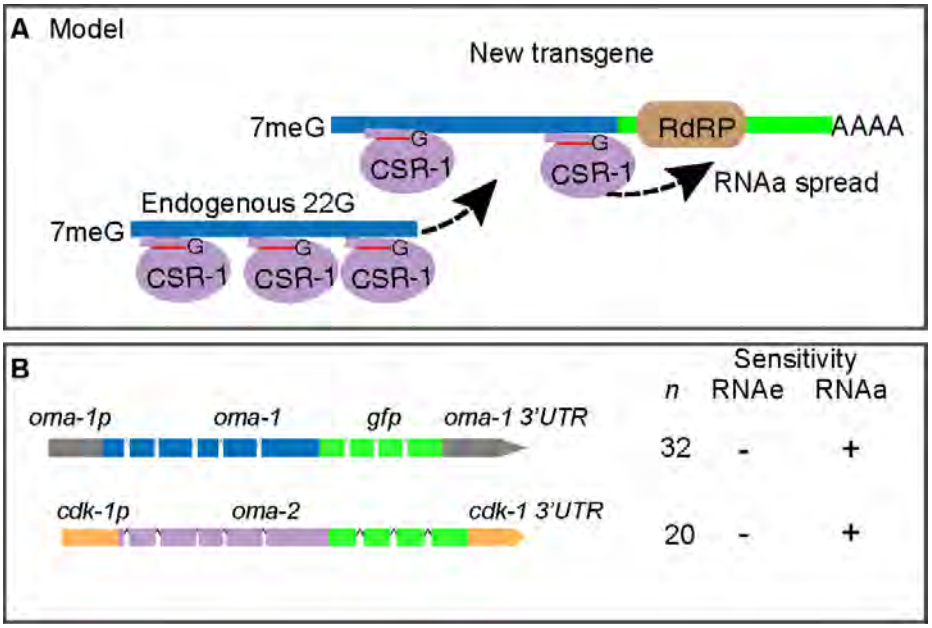
promote the local recruitment of RdRP, leading to the *de novo* synthesis of CSR-1 22G-RNAs within the adjacent foreign sequences. Molecular spreading of this type has been observed in gene silencing in both plants and animals (Axtell et al., 2006; Pak and Fire, 2007; Sijen et al., 2001; Sijen et al., 2007). To test this hypothesis, we used CRISPR gene editing to delete *oma-1* from the endogenous locus and reintroduced *oma-1::gfp de novo* in the deletion background (n=2). When this transgene was tested for transactivation, it was able to activate GFP::CDK-1 (Figure 5.1B). This raises the possibility if *oma-1::gfp* licensing was coming from CSR-1/22Gs associated from some other locus in the genome perhaps *oma-2*. To test this hypothesis we first drove *oma-2* expression by a *cdk*-1promoter and 3'UTR and found that like *oma-1::gfp*, *cdk-1p:oma-2::gfp:cdk-13'UTR* (n=7) was able to transactivate a silent transgene (n=20). Next to rule out any CSR-1/22G coming from either of the endogenous locus, it is necessary to construct the double knockout of *oma-1* and *oma-2*. However, oma-1/2 double knockout is lethal and therefore it becomes difficult to test this hypothesis. From the above experiment we can only conclude that just like *oma-1*, *oma-2* open reading frame was sufficient for its RNAa activity. Further experiments need to be done to determine if CSR-1/22Gs from *oma-2* can license *oma-1* to confer it transactivation property.

Figure 5.1. Endogenous *oma-1* associated CSR-1/22G is not required for RNAa

(A) Model for RNAa activity of *oma-1::gfp*. Refer to text for details.

(B) Schematic of the respective transgene (as indicated). The exon-intron structure is indicated with boxes and lines, respectively. *n* represents the number of F1 cross progeny scored in the transactivation assay while crossing it to off or on transgene. Sensitivity to RNAe represent if the transgene was able to silence or was silenced by other RNAe transgene when scored in the transactivation assay. Sensitivity to RNAa represent if the transgene was able to transactivate another silent transgene in the transactivation assay.

Figure 5.1. Endogenous *oma-1* associated CSR-1/22G is not required for RNAa



Because of our curiosity to understand if the *oma-1* coding sequence or the coding potential was required for transactivation, we generated STOP codon transgene of *oma-1* te33 (R_{106}STOP) (Lin, 2003) (n>3) and found that although not expressed at the protein level (under the fluorescence microscope), it was able to activate a silent transgene (Figure 5.2A). This observation suggests that it is the *oma-1* coding sequence and not OMA-1 protein is required for transactivation. In another approach to test similar hypothesis, I first made *oma-1* tagged with *gfp* at the 5' end (n=12) (Figure 5.2A). I noticed that all *gfp::oma-1* transgenic lines were expressed and when tested in a cross, it was able to transactivate a silent transgene. I further mutated an amino acid to a STOP codon in the 2nd exon of the *gfp* of *oma-1::gfp* transgene and introduce this transgene as a single copy MosSCI (n=5). Although *gfp* (STOP)*::oma-1* transgene was silent at protein level (as detected by GFP fluorescence), it was able to transactivate a silent transgene (Figure 5.2A). This experiment again suggests that it is the *oma-1* mRNA, not the OMA-1 protein that is required for transactivation. It further suggests that perhaps the small RNAs are generated in the P-granules before the translation starts in the cytoplasm. Similarly, we also showed that *gfp::cdk-1 (RNAe)* was still RNAe when an amino acid was mutated to the STOP codon in the second exon of *gfp* (n>3) (Figure 5.2A). The above STOP codon alleles for various transgenic strains further conclude that small RNAs are generated in the P-granule before the NMD (Non-sense Mediated Decay) surveillance happens.

If the hypothesis of *oma-2* providing transactivation property to *oma-1* is not true, it could be possible that there is some feature of the *oma-1* sequence *in cis* that is providing it transactivation activity. In order to test this hypothesis, we made various truncations of *oma-1* coding sequence in the *oma-1::gfp* transgene as shown in (Figure 5.2B). Interestingly, we found that each half of *oma-1*, tagged with *gfp* is expressed, however is not sufficient for transactivation activity of the transgene (n> 3) (Figure 5.2B). This made me think if there is some specific sequence in the junction of the two halves responsible for the transactivation. So I initially took the middle 800 *bp* that included both halves of *oma-1*, tagged with *gfp* (n=10) and then introduced as a single copy transgene. Surprisingly this transgene was able to transactivate another silent transgene (Figure 5.2B). Further, we reduced this 800 *bp* to150 *bp* at the junction of two halves with *gfp* tagged at 3' end (n>3). Interestingly, this transgene was also able to transactivate (Figure 5.2B) another silent transgene. Since these truncations were still able to exhibit transactivation, we hypothesize that there could be two alternate possibilities for this phenomenon. In the first scenario, it is possible that this region is a lot more similar to *oma-2*, thus endogenous small RNAs from *oma-2* is providing it RNAa activity. When we looked into the overlap of sequences between *oma-1* and *oma-2,* we found that there were three stretches of 22 *nts* that could be possibly driving 22Gs from *oma-2* and thus licensing the *150 bp* of *oma-1::gfp* bearing transgene. In the second scenario, it could be possible that there is some transcription factor binding to this 150 *bp in cis* that is

providing *oma-1* its RNAa activity. And to test this hypothesis we could design future experiments to nail down the factors binding to this 150 *bp* of RNA.

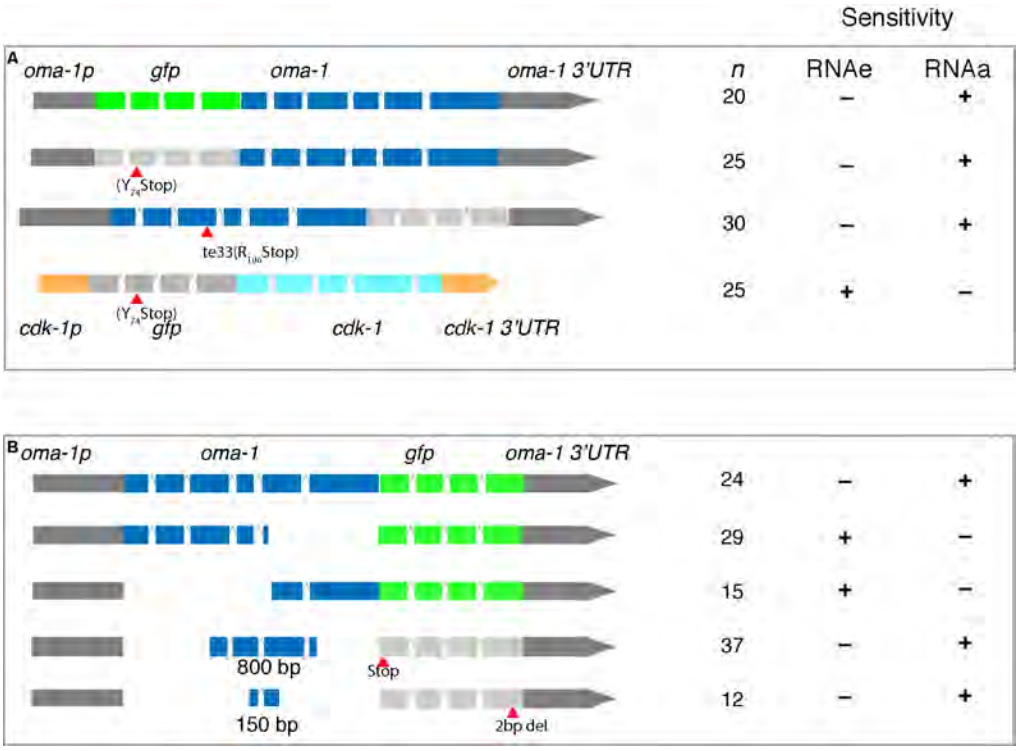
Figure 5.2. Dissecting *oma-1* coding sequence to study its ability to transactivate a silent transgene.

(A) RNAa or RNAe happens before NMD surveillance.

Schematic diagram showing the respective transgenes (as indicated). The exon-intron structure is indicated with boxes and lines, respectively. *n* represents the number of F1 cross progeny scored in the transactivation assay while crossing it to off or on transgene. Sensitivity to RNAe represent if the transgene was able to silence or was silenced by other RNAe transgene when scored in the transactivation assay. Sensitivity to RNAa represent if the transgene was able to transactivate another silent transgene in the transactivation assay.

(B) Schematic of the respective transgene (as indicated). The exon-intron structure is indicated with boxes and lines, respectively. *n* represents the number of F1 cross progeny scored in the transactivation assay while crossing it to off or on transgene. Sensitivity to RNAe represent if the transgene was able to silence or was silenced by other RNAe transgene when scored in the transactivation assay. Sensitivity to RNAa represent if the transgene was able to transactivate another silent transgene in the transactivation assay.

Figure 5.2. Dissecting *oma-1* coding sequence to study its ability to transactivate a silent transgene.



Reverse genetic screen to identify factors required in PRG-1 pathway

As we know that PRG-1 is only required for the initiation of RNAe (Shirayama et al., 2012). This makes it very challenging to screen for factors required in the PRG-1 pathway because it is not possible to generate new transgene in thousands of mutant background as a screening strategy. However we found that PRG-1 is required to re-silence a transgene that is activated by RNAa whereas in *prg-1(+)* activity, this transgene was silent Chapter IV (Figure 4.1A). Using a double transgenic strain OMA-1::GFP (ON) LGII, GFP::CSR-1 (OFF) LGIV*,* I have designed and executed a reverse genetic screen and identified numerous factors involved in both the PIWI-Argonaute pathway, which scans for foreign nucleic acids, and the CSR-1 pathway, which protects self-mRNAs from PIWI silencing (Figure 5.3). Since the screening involves worms to be scored individually in the Normaski microscope and is very labor intensive we started screening with a smaller set of embryonic lethal library. Along with Rita Sharma, a former technician in our lab, we have screened approximately 1152 RNAi clones and have discovered many interesting candidates involved in the piRNA biogenesis (Table 5.1). We validated the involvement of subset of candidate genes in the piRNA biogenesis by performing small RNA cloning on the RNAi mutant and quantified their piRNA levels (Figure 5.4A and 5.4B). We also looked at PRG-1 protein levels by western by probing it with PRG-1 antibody in mutant background (Figure 5.4C and 5.4D). However these factors need to be

further characterized for their mechanistic role in various steps of the piRNA biogenesis pathway (Figure 5.5).

Figure 5.3. Screening Method: Reverse genetic screen for PRG-1 pathway genes

(A) Epifluorescence image of representative germline (outlined with dashes) of the genotype as indicated. The cytoplasmic fluorescence signal is OMA-1::GFP; the P-granule signal is GFP::CSR-1.

Method: Fifty L1 transgenic worms bearing ON and OFF transgene as indicated were put on both candidate and control RNAi food in duplicates. Two days later ten adults worms from each plate was scored under the Normanski Microscope in a circled slide with 10 µl of 0.02mM levamisole for expression of GFP::CSR-1 in the P-granule.

Figure 5.3. Screening Method: Reverse genetic screen for PRG-1 pathway genes

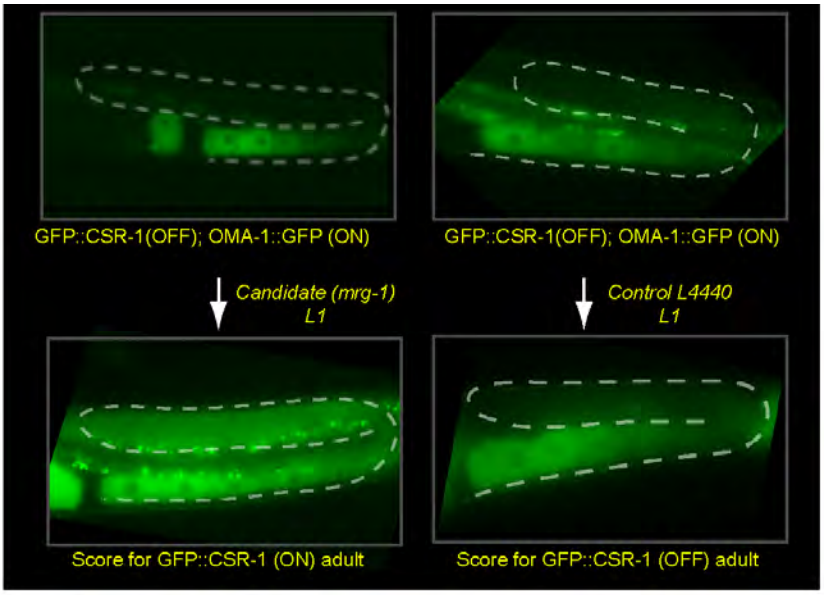


Table 5.1. Candidate genes identified in PRG-1 pathway

The first column represents the name of the candidate gene identified in the screen. *n* represents the total number of worms scored on that particular candidate RNAi food in 2 or more than two experiments. Percentage represents the number of worms that were desilenced for GFP::CSR-1(ON) in the double transgenic strain (OMA-1::GFP (ON); GFP::CSR-1(OFF) on the candidate RNAi food.

Table 5.1. Candidate genes identified in PRG-1 pathway

Candidate Gene	Number of de-silenced *gfp::csr-1* in the germline of (*oma-1::gfp, gfp::csr-1 double*)
snpc-4 (SNAPc (Small Nuclear RNA Activating Complex) homolog)	100% n > 100
phi-16 (Ortholog of human SUPT-16H, FACT Complex (facilitate chromatin transcription)	70-100% n = 30
use-1 (unconventional SNARE in endoplasmic reticulum)	100% n = 63
C53H9.2 (GTP binding protein)	80-100% n = 50
prp-17 (yeast PRP (splicing factor) related)	90-100% n = 60
T08G11.4 (Ortholog of human TGS-1) Trimethylguanosine synthase 1	88-100% n = 43
F54H12.1 (Aconitase-2)	94-100% n = 33
hrde-1 (Heritable RNAi Deficient)	100% n = 13
xpo-2 (eXPortin (nuclear export receptor)	90-100% n = 50
rsp-20 (Ribosomal subunit protein)	60-100% n = 30
npp-6 (Nuclear pore complex)	60-100% n = 37
npp-7 (Nuclear pore complex)	50-90% n = 40
npp-8 (Nuclear pore complex)	86% n = 36
mog-5 (DEAH helicase, orthologous to PRP22)	80-86% n = 25
mrg-1 (human MRG (Mortality factor-Related Gene) related),	69% n = 29

Figure 5.4. Relative piRNAs and PRG-1 levels in mutant defective for PRG-1 pathway

(A and B) The expression profile for the bulk population of piRNAs as determined by small-RNA sequencing. Plotted for each library is the percent of reads that represented piRNAs after normalized to total (non-structural) RNAs. Each library was made from total RNA prepared from the adult N2 worms collected from the respective RNAi food.

(C and D) Analysis of PRG-1 protein levels in different RNAi strains (as indicated). Plotted for mutant is the relative PRG-1 level normalized to the MRG-1 level. Total protein was extracted from the adult N2 worms collected from the respective RNAi food. The blot was probed with anti-PRG-1 and anti-MRG-1 antibodies (as indicated).

Figure 5.4. Relative piRNAs and PRG-1 levels in mutant defective for PRG-1 pathway

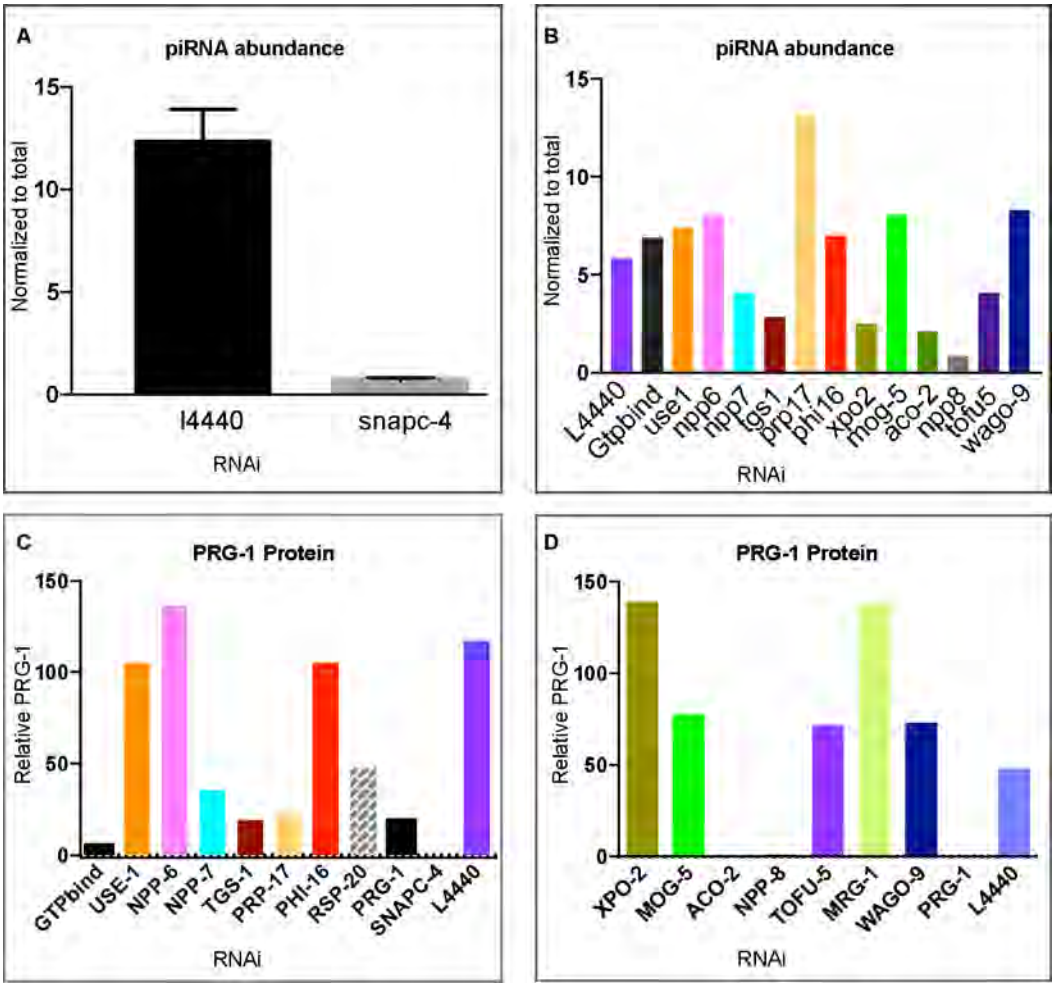
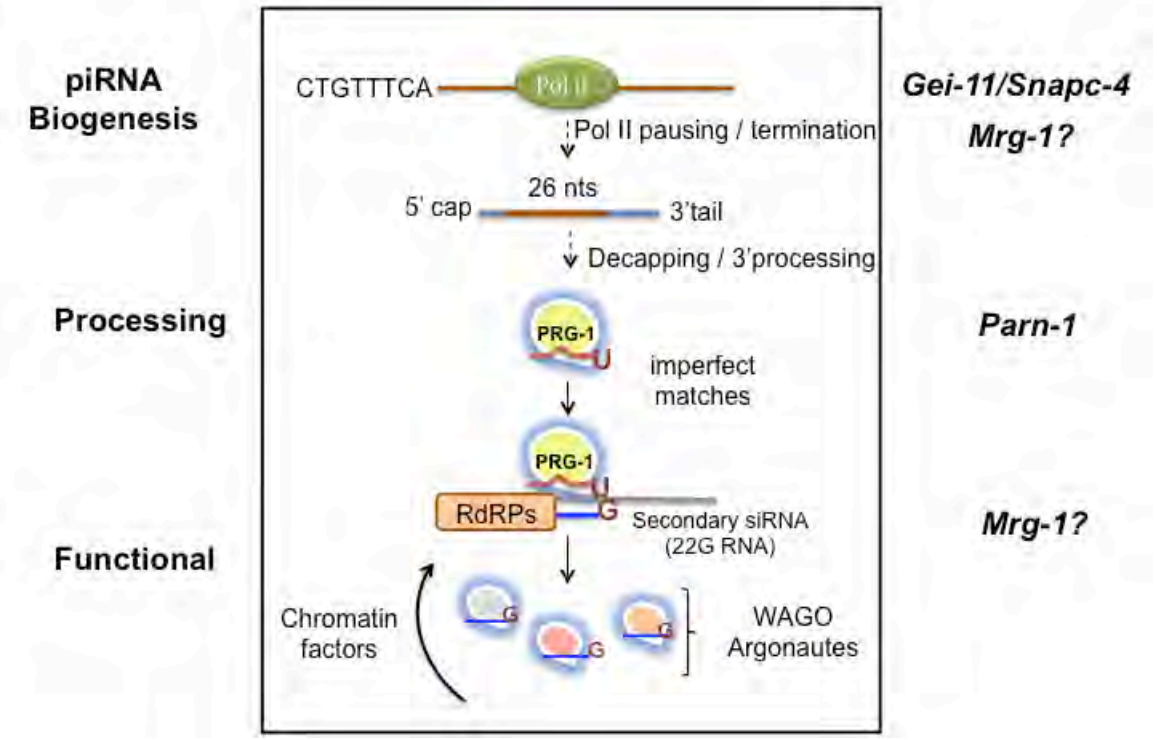


Figure 5.5. Model for categorizing candidate genes into different steps in piRNA/PRG-1 pathway

See Discussion for details.

Figure 5.5. Model for categorizing candidate genes into different steps in piRNA/PRG-1 pathway



Super silencing is PRG-1 independent

We have demonstrated in Chapter II of my thesis that PRG-1 is required for the initiation but not the maintenance of RNAe (Figure 2.4). Surprisingly few single copy transgenes with completely foreign sequences like *Cas-9 (n=3)* or *oma-1 (Codon alt)::gfp (n=11)* as described in Chapter IV, were never expressed in the germline when injected in the *prg-1(tm872)* mutant worms. Despite being independent of PRG-1, the so-called “super-silenced” transgenes require RDE-3 and secondary Argonautes for their silencing.

How is super-silencing initiated? Multiple primary Argonautes can trigger WAGO-mediated silencing in worms, so perhaps a different primary Argonaute triggers super-silencing. RDE-1—which triggers WAGO-mediated silencing in response to exogenous double-stranded RNA—might be a good candidate to test. When plasmids are injected into the germline they likely form extrachromosomal arrays from which the donor molecule is copied into the recipient MosSCI site in the genome. Extrachromosomal arrays can trigger RDE-1–dependent silencing of transgenes in somatic tissues of ERI pathway mutants (Kennedy et al., 2004; Simmer et al., 2002). The ERI pathway competes with the RNAi pathway for the limited pool of WAGOs. So it would be interesting to test if disruption of both *prg-1* and *rde-1* prevents super-silencing.

CONCLUSION

In summary the studies described in this dissertation have revealed two pathways in which Argonaute/small-RNA complexes serve as memories of gene

expression from one generation to the next. One of these is the RNAe pathway, in which single-copy transgenes are permanently silenced. The PIWI Argonaute PRG-1 and its genomically encoded piRNA cofactors initiate RNAe, and maintenance depends on Chromatin factors and the WAGOs. The other pathway, referred as RNA-mediated gene activation RNAa, protects endogenous mRNAs from piRNA induced silencing and involves the Argonaute CSR-1. CSR-1 and the WAGO Argonautes engage antisense small RNAs produced on mRNA templates by RdRP. Our findings support a model in which PRG-1 scans for foreign sequences and two Argonaute pathways serve as epigenetic memories of “self” and “non-self” RNAs. Further we show PRG-1 and CSR-1 AGO pathways are in constant competition on certain transgene targets. Whether a transcript is expressed or silenced depends on degree of targeting by each Argonaute. Our findings are beginning to reveal new components and insights into remarkable multi-Argonaute system that scans whole transcriptome to transmit memories of previous gene expression states.

These findings pose many questions for the future: Do similar mechanisms exist in other organisms? Do other, perhaps entirely different self-recognition, “licensing,” pathways play a role in protecting mRNAs from AGO surveillance? Could these paradigms be important in the stable maintenance of somatic differentiated fates? To what extent are “hard-epigenetic” mechanisms of this kind important for inheritance and evolution in plants and animals?

BIBLIOGRAPHY

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

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

Antonovics, J., Boots, M., Abbate, J., Baker, C., McFrederick, Q., and Panjeti, V. (2011). Biology and evolution of sexual transmission. Ann N Y Acad Sci *1230*, 12-24.

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

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

Aravin, A.A., and Hannon, G.J. (2008). Small RNA silencing pathways in germ and stem cells. Cold Spring Harb Symp Quant Biol *73*, 283-290.

Aravin, A.A., Hannon, G.J., and Brennecke, J. (2007). The Piwi-piRNA pathway provides an adaptive defense in the transposon arms race. Science *318*, 761-764.

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

Axtell, M.J., Jan, C., Rajagopalan, R., and Bartel, D.P. (2006). A two-hit trigger for siRNA biogenesis in plants. Cell *127*, 565-577.

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

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D.A., and Horvath, P. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. Science *315*, 1709-1712.

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

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

Bessler, J.B., Andersen, E.C., and Villeneuve, A.M. (2010). Differential localization and independent acquisition of the H3K9me2 and H3K9me3 chromatin modifications in the Caenorhabditis elegans adult germ line. PLoS Genet *6*, e1000830.

Birchler, J., Pal-Bhadra, M., and Bhadra, U. (2003). Transgene cosuppression in animals. In RNAi: A guide to gene silencing (ed. G. Hannon). Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 23–42.

Bird, A. (2007). Perceptions of epigenetics. Nature *447*, 396-398.

Bosher, J.M., and Labouesse, M. (2000). RNA interference: genetic wand and genetic watchdog. Nat Cell Biol *2*, E31-36.

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

Brennecke, J., Malone, C.D., Aravin, A.A., Sachidanandam, R., Stark, A., and Hannon, G.J. (2008). An epigenetic role for maternally inherited piRNAs in transposon silencing. Science *322*, 1387-1392.

Brenner, S. (1974). The genetics of Caenorhabditis elegans. Genetics *77*, 71-94.

Brodersen, P., and Voinnet, O. (2006). The diversity of RNA silencing pathways in plants. Trends Genet *22*, 268-280.

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

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

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

Carmell, M.A., and Hannon, G.J. (2004). RNase III enzymes and the initiation of gene silencing. Nat Struct Mol Biol *11*, 214-218.

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

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

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

Cerutti, H., and Casas-Mollano, J.A. (2006). On the origin and functions of RNA-mediated silencing: from protists to man. Curr Genet *50*, 81-99.

Chalker, D.L., and Yao, M.C. (2011). DNA elimination in ciliates: transposon domestication and genome surveillance. Annu Rev Genet *45*, 227-246.

Chen, C., Ridzon, D.A., Broomer, A.J., Zhou, Z., Lee, D.H., Nguyen, J.T., Barbisin, M., Xu, N.L., Mahuvakar, V.R., Andersen, M.R.*, et al.* (2005a). Real-time quantification of microRNAs by stem-loop RT-PCR. Nucleic Acids Res *33*, e179.

Chen, C.C., Simard, M.J., Tabara, H., Brownell, D.R., McCollough, J.A., and Mello, C.C. (2005b). A member of the polymerase beta nucleotidyltransferase superfamily is required for RNA interference in C. elegans. Curr Biol *15*, 378-383.

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

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

Conine, C.C., Moresco, J.J., Gu, W., Shirayama, M., Conte, D., Jr., Yates, J.R., 3rd, and Mello, C.C. (2013). Argonautes promote male fertility and provide a paternal memory of germline gene expression in C. elegans. Cell *155*, 1532-1544.

Couteau, F., Guerry, F., Muller, F., and Palladino, F. (2002). A heterochromatin protein 1 homologue in Caenorhabditis elegans acts in germline and vulval development. EMBO Rep *3*, 235-241.

Cox, D.N., Chao, A., Baker, J., Chang, L., Qiao, D., and Lin, H. (1998). A novel class of evolutionarily conserved genes defined by piwi are essential for stem cell self-renewal. Genes Dev *12*, 3715-3727.

Czech, B., and Hannon, G.J. (2016). One Loop to Rule Them All: The Ping-Pong Cycle and piRNA-Guided Silencing. Trends Biochem Sci *41*, 324-337.

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

Daxinger, L., and Whitelaw, E. (2012). Understanding transgenerational epigenetic inheritance via the gametes in mammals. Nat Rev Genet *13*, 153-162.

de Carvalho, F., Gheysen, G., Kushnir, S., Van Montagu, M., Inze, D., and Castresana, C. (1992). Suppression of beta-1,3-glucanase transgene expression in homozygous plants. Embo J *11*, 2595-2602.

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

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

E.A. Russo, R.A.M., Arthur D. Riggs. (1996). Epigenetic mechanisms of gene regulation. xii, 692 p.

Egger, G., Liang, G., Aparicio, A., and Jones, P.A. (2004). Epigenetics in human disease and prospects for epigenetic therapy. Nature *429*, 457-463.

Erhard, K.F., Jr., and Hollick, J.B. (2011). Paramutation: a process for acquiring trans-generational regulatory states. Curr Opin Plant Biol *14*, 210-216.

Fang, W., Wang, X., Bracht, J.R., Nowacki, M., and Landweber, L.F. (2012). Piwi-interacting RNAs protect DNA against loss during Oxytricha genome rearrangement. Cell *151*, 1243-1255.

Fire, A. (1986). Integrative transformation of Caenorhabditis elegans. Embo J *5*, 2673-2680.

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

Flajnik, M.F., and Kasahara, M. (2010). Origin and evolution of the adaptive immune system: genetic events and selective pressures. Nat Rev Genet *11*, 47-59.

Frokjaer-Jensen, C., Davis, M.W., Hopkins, C.E., Newman, B.J., Thummel, J.M., Olesen, S.P., Grunnet, M., and Jorgensen, E.M. (2008). Single-copy insertion of transgenes in Caenorhabditis elegans. Nat Genet *40*, 1375-1383.

Frokjaer-Jensen, C., Jain, N., Hansen, L., Davis, M.W., Li, Y., Zhao, D., Rebora, K., Millet, J.R., Liu, X., Kim, S.K.*, et al.* (2016). An Abundant Class of Non-coding DNA Can Prevent Stochastic Gene Silencing in the C. elegans Germline. Cell *166*, 343-357.

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

Ghildiyal, M., and Zamore, P.D. (2009). Small silencing RNAs: an expanding universe. Nat Rev Genet *10*, 94-108.

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

Goh, W.S., Falciatori, I., Tam, O.H., Burgess, R., Meikar, O., Kotaja, N., Hammell, M., and Hannon, G.J. (2015). piRNA-directed cleavage of meiotic transcripts regulates spermatogenesis. Genes Dev *29*, 1032-1044.

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

Goldberg, A.D., Allis, C.D., and Bernstein, E. (2007). Epigenetics: a landscape takes shape. Cell *128*, 635-638.

Gou, L.T., Dai, P., Yang, J.H., Xue, Y., Hu, Y.P., Zhou, Y., Kang, J.Y., Wang, X., Li, H., Hua, M.M.*, et al.* (2014). Pachytene piRNAs instruct massive mRNA elimination during late spermiogenesis. Cell Res *24*, 680-700.

Grentzinger, T., Armenise, C., Brun, C., Mugat, B., Serrano, V., Pelisson, A., and Chambeyron, S. (2012). piRNA-mediated transgenerational inheritance of an acquired trait. Genome Res *22*, 1877-1888.

Grewal, S.I., and Elgin, S.C. (2007). Transcription and RNA interference in the formation of heterochromatin. Nature *447*, 399-406.

Grishok, A., Sinskey, J.L., and Sharp, P.A. (2005). Transcriptional silencing of a transgene by RNAi in the soma of C. elegans. Genes Dev *19*, 683-696.

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

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

Gu, W., Lee, H.C., Chaves, D., Youngman, E.M., Pazour, G.J., Conte, D., Jr., and Mello, C.C. (2012b). CapSeq and CIP-TAP identify Pol II start sites and reveal capped small RNAs as C. elegans piRNA precursors. Cell *151*, 1488-1500.

Gu, W., Shirayama, M., Conte, D., Jr., Vasale, J., Batista, P.J., Claycomb, J.M., Moresco, J.J., Youngman, E.M., Keys, J., Stoltz, M.J.*, et al.* (2009). Distinct argonaute-mediated 22G-RNA pathways direct genome surveillance in the C. elegans germline. Mol Cell *36*, 231-244.

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

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

Gunawardane, L.S., Saito, K., Nishida, K.M., Miyoshi, K., Kawamura, Y., Nagami, T., Siomi, H., and Siomi, M.C. (2007). A slicer-mediated mechanism for repeat-associated siRNA 5' end formation in Drosophila. Science *315*, 1587-1590.

Halfmann, R., and Lindquist, S. (2010). Epigenetics in the extreme: prions and the inheritance of environmentally acquired traits. Science *330*, 629-632.

Hammond, S.M., Boettcher, S., Caudy, A.A., Kobayashi, R., and Hannon, G.J. (2001). Argonaute2, a link between genetic and biochemical analyses of RNAi. Science *293*, 1146-1150.

Han, B.W., Wang, W., Li, C., Weng, Z., and Zamore, P.D. (2015). Noncoding RNA. piRNA-guided transposon cleavage initiates Zucchini-dependent, phased piRNA production. Science *348*, 817-821.

Harold, F.M. (2005). Molecules into cells: specifying spatial architecture. Microbiol Mol Biol Rev *69*, 544-564.

Henderson, I.R., and Jacobsen, S.E. (2007). Epigenetic inheritance in plants. Nature *447*, 418-424.

Hinas, A., Wright, A.J., and Hunter, C.P. (2012). SID-5 is an endosome-associated protein required for efficient systemic RNAi in C. elegans. Curr Biol *22*, 1938-1943.

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

Hornung, V., and Latz, E. (2010). Intracellular DNA recognition. Nat Rev Immunol *10*, 123-130.

Huang, V., Zheng, J., Qi, Z., Wang, J., Place, R.F., Yu, J., Li, H., and Li, L.C. (2013). Ago1 Interacts with RNA polymerase II and binds to the promoters of actively transcribed genes in human cancer cells. PLoS Genet *9*, e1003821.

Iwasaki, Y.W., Siomi, M.C., and Siomi, H. (2015). PIWI-Interacting RNA: Its Biogenesis and Functions. Annu Rev Biochem *84*, 405-433.

Jackstadt, P., Wilm, T.P., Zahner, H., and Hobom, G. (1999). Transformation of nematodes via ballistic DNA transfer. Mol Biochem Parasitol *103*, 261-266.

Janowski, B.A., Younger, S.T., Hardy, D.B., Ram, R., Huffman, K.E., and Corey, D.R. (2007). Activating gene expression in mammalian cells with promoter-targeted duplex RNAs. Nat Chem Biol *3*, 166-173.

Johnson, C.L., and Spence, A.M. (2011). Epigenetic licensing of germline gene expression by maternal RNA in C. elegans. Science *333*, 1311-1314.

Juliano, C., Wang, J., and Lin, H. (2011). Uniting germline and stem cells: the function of Piwi proteins and the piRNA pathway in diverse organisms. Annu Rev Genet *45*, 447-469.

Kalmykova, A.I., Klenov, M.S., and Gvozdev, V.A. (2005). Argonaute protein PIWI controls mobilization of retrotransposons in the Drosophila male germline. Nucleic Acids Res *33*, 2052-2059.

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

Kawaoka, S., Izumi, N., Katsuma, S., and Tomari, Y. (2011). 3' end formation of PIWI-interacting RNAs in vitro. Mol Cell *43*, 1015-1022.

Kelly, W.G., and Fire, A. (1998). Chromatin silencing and the maintenance of a functional germline in Caenorhabditis elegans. Development *125*, 2451-2456.

Kelly, W.G., Schaner, C.E., Dernburg, A.F., Lee, M.H., Kim, S.K., Villeneuve, A.M., and Reinke, V. (2002). X-chromosome silencing in the germline of C. elegans. Development *129*, 479-492.

Kelly, W.G., Xu, S., Montgomery, M.K., and Fire, A. (1997). Distinct requirements for somatic and germline expression of a generally expressed Caernorhabditis elegans gene. Genetics *146*, 227-238.

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

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

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

Khurana, J.S., and Theurkauf, W. (2010). piRNAs, transposon silencing, and Drosophila germline development. J Cell Biol *191*, 905-913.

Kim, H., Ishidate, T., Ghanta, K.S., Seth, M., Conte, D., Jr., Shirayama, M., and Mello, C.C. (2014). A co-CRISPR strategy for efficient genome editing in Caenorhabditis elegans. Genetics *197*, 1069-1080.

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

Kim, V.N., Han, J., and Siomi, M.C. (2009). Biogenesis of small RNAs in animals. Nat Rev Mol Cell Biol *10*, 126-139.

Kimble, J., Hodgkin, J., Smith, T., and Smith, J. (1982). Suppression of an amber mutation by microinjection of suppressor tRNA in C. elegans. Nature *299*, 456-458.

Knight, S.W., and Bass, B.L. (2002). The role of RNA editing by ADARs in RNAi. Mol Cell *10*, 809-817.

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

Kuhn, C.D., and Joshua-Tor, L. (2013). Eukaryotic Argonautes come into focus. Trends Biochem Sci *38*, 263-271.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol *10*, R25.

Lau, N.C., Seto, A.G., Kim, J., Kuramochi-Miyagawa, S., Nakano, T., Bartel, D.P., and Kingston, R.E. (2006). Characterization of the piRNA complex from rat testes. Science *313*, 363-367.

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

Lessing, D., and Lee, J.T. (2013). X chromosome inactivation and epigenetic responses to cellular reprogramming. Annu Rev Genomics Hum Genet *14*, 85-110.

Li, L.C., Okino, S.T., Zhao, H., Pookot, D., Place, R.F., Urakami, S., Enokida, H., and Dahiya, R. (2006). Small dsRNAs induce transcriptional activation in human cells. Proc Natl Acad Sci U S A *103*, 17337-17342.

Li, X.Z., Roy, C.K., Dong, X., Bolcun-Filas, E., Wang, J., Han, B.W., Xu, J., Moore, M.J., Schimenti, J.C., Weng, Z.*, et al.* (2013). An ancient transcription factor initiates the burst of piRNA production during early meiosis in mouse testes. Mol Cell *50*, 67-81.

Lim, J.P., and Brunet, A. (2013). Bridging the transgenerational gap with epigenetic memory. Trends Genet *29*, 176-186.

Lin, H. (2007). piRNAs in the germ line. Science *316*, 397.

Lin, H., and Spradling, A.C. (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.

Lippman, Z., and Martienssen, R. (2004). The role of RNA interference in heterochromatic silencing. Nature *431*, 364-370.

Liu, J., Carmell, M.A., Rivas, F.V., Marsden, C.G., Thomson, J.M., Song, J.J., Hammond, S.M., Joshua-Tor, L., and Hannon, G.J. (2004a). Argonaute2 is the catalytic engine of mammalian RNAi. Science *305*, 1437-1441.

Liu, Y., Mochizuki, K., and Gorovsky, M.A. (2004b). Histone H3 lysine 9 methylation is required for DNA elimination in developing macronuclei in Tetrahymena. Proc Natl Acad Sci U S A *101*, 1679-1684.

Lopez-Larrea, C. (2012). Self and Nonself, 1 edn (Springer-Verlag New York).

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

Malone, C.D., and Hannon, G.J. (2009). Small RNAs as guardians of the genome. Cell *136*, 656-668.

Marraffini, L.A., and Sontheimer, E.J. (2008). CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. Science *322*, 1843-1845.

Marraffini, L.A., and Sontheimer, E.J. (2010). CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. Nat Rev Genet *11*, 181-190.

Matzke, M.A., Matzke, A.J., Pruss, G.J., and Vance, V.B. (2001). RNA-based silencing strategies in plants. Curr Opin Genet Dev *11*, 221-227.

McClintock, B. (1951). Chromosome organization and genic expression. Cold Spring Harb Symp Quant Biol *16*, 13-47.

Meister, G., Landthaler, M., Patkaniowska, A., Dorsett, Y., Teng, G., and Tuschl, T. (2004). Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs. Mol Cell *15*, 185-197.

Mello, C., and Fire, A. (1995). DNA transformation. Methods Cell Biol *48*, 451-482.

Mello, C.C., Kramer, J.M., Stinchcomb, D., and Ambros, V. (1991). Efficient gene transfer in C.elegans: extrachromosomal maintenance and integration of transforming sequences. Embo J *10*, 3959-3970.

Mercer, T.R., and Mattick, J.S. (2013). Structure and function of long noncoding RNAs in epigenetic regulation. Nat Struct Mol Biol *20*, 300-307.

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

Mochizuki, K., Fine, N.A., Fujisawa, T., and Gorovsky, M.A. (2002). Analysis of a piwi-related gene implicates small RNAs in genome rearrangement in tetrahymena. Cell *110*, 689-699.

Murray, N.E. (2002). 2001 Fred Griffith review lecture. Immigration control of DNA in bacteria: self versus non-self. Microbiology *148*, 3-20.

Nakamura, K., Kim, S., Ishidate, T., Bei, Y., Pang, K., Shirayama, M., Trzepacz, C., Brownell, D.R., and Mello, C.C. (2005). Wnt signaling drives WRM-1/beta-catenin asymmetries in early C. elegans embryos. Genes Dev *19*, 1749-1754.

Napoli, C., Lemieux, C., and Jorgensen, R. (1990). Introduction of a Chimeric Chalcone Synthase Gene into Petunia Results in Reversible Co-Suppression of Homologous Genes in trans. Plant Cell *2*, 279-289.

Nicoll, M., Akerib, C.C., and Meyer, B.J. (1997). X-chromosome-counting mechanisms that determine nematode sex. Nature *388*, 200-204.

O'Donnell, K.A., and Boeke, J.D. (2007). Mighty Piwis defend the germline against genome intruders. Cell *129*, 37-44.

Ohta, H., Fujiwara, M., Ohshima, Y., and Ishihara, T. (2008). ADBP-1 regulates an ADAR RNA-editing enzyme to antagonize RNA-interference-mediated gene silencing in Caenorhabditis elegans. Genetics *180*, 785-796.

Ortiz, M.A., Noble, D., Sorokin, E.P., and Kimble, J. (2014). A new dataset of spermatogenic vs. oogenic transcriptomes in the nematode Caenorhabditis elegans. G3 (Bethesda) *4*, 1765-1772.

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

Pancer, Z., Amemiya, C.T., Ehrhardt, G.R., Ceitlin, J., Gartland, G.L., and Cooper, M.D. (2004). Somatic diversification of variable lymphocyte receptors in the agnathan sea lamprey. Nature *430*, 174-180.

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

Place, R.F., Li, L.C., Pookot, D., Noonan, E.J., and Dahiya, R. (2008). MicroRNA-373 induces expression of genes with complementary promoter sequences. Proc Natl Acad Sci U S A *105*, 1608-1613.

Praitis, V. (2006). Creation of transgenic lines using microparticle bombardment methods. Methods Mol Biol *351*, 93-107.

Rhind, N.R., Miller, L.M., Kopczynski, J.B., and Meyer, B.J. (1995). xol-1 acts as an early switch in the C. elegans male/hermaphrodite decision. Cell *80*, 71-82.

Rieckher, M., Kourtis, N., Pasparaki, A., and Tavernarakis, N. (2009). Transgenesis in Caenorhabditis elegans. Methods Mol Biol *561*, 21-39.

Robert, V.J., Vastenhouw, N.L., and Plasterk, R.H. (2004). RNA interference, transposon silencing, and cosuppression in the Caenorhabditis elegans germ line: similarities and differences. Cold Spring Harb Symp Quant Biol *69*, 397-402.

Romano, N., and Macino, G. (1992). Quelling: transient inactivation of gene expression in Neurospora crassa by transformation with homologous sequences. Mol Microbiol *6*, 3343-3353.

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

Rulicke, T., and Hubscher, U. (2000). Germ line transformation of mammals by pronuclear microinjection. Exp Physiol *85*, 589-601.

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

Shibuya, K., Fukushima, S., and Takatsuji, H. (2009a). RNA-directed DNA methylation induces transcriptional activation in plants. Proc Natl Acad Sci U S A *106*, 1660-1665.

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

Shirayama, M., Stanney, W., 3rd, Gu, W., Seth, M., and Mello, C.C. (2014). The Vasa Homolog RDE-12 engages target mRNA and multiple argonaute proteins to promote RNAi in C. elegans. Curr Biol *24*, 845-851.

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

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

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

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

Song, J.J., Smith, S.K., Hannon, G.J., and Joshua-Tor, L. (2004). Crystal structure of Argonaute and its implications for RISC slicer activity. Science *305*, 1434-1437.

Stinchcomb, D.T., Shaw, J.E., Carr, S.H., and Hirsh, D. (1985). Extrachromosomal DNA transformation of Caenorhabditis elegans. Mol Cell Biol *5*, 3484-3496.

Strome, S., and Lehmann, R. (2007). Germ versus soma decisions: lessons from flies and worms. Science *316*, 392-393.

Strome, S., Powers, J., Dunn, M., Reese, K., Malone, C.J., White, J., Seydoux, G., and Saxton, W. (2001). Spindle dynamics and the role of gamma-tubulin in early Caenorhabditis elegans embryos. Mol Biol Cell *12*, 1751-1764.

Tabara, H., Grishok, A., and Mello, C.C. (1998). RNAi in C. elegans: soaking in the genome sequence. Science *282*, 430-431.

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

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

Tavernarakis, N., Wang, S.L., Dorovkov, M., Ryazanov, A., and Driscoll, M. (2000). Heritable and inducible genetic interference by double-stranded RNA encoded by transgenes. Nat Genet *24*, 180-183.

Timmons, L., Tabara, H., Mello, C.C., and Fire, A.Z. (2003). Inducible systemic RNA silencing in Caenorhabditis elegans. Mol Biol Cell *14*, 2972-2983.

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

Vastenhouw, N.L., Brunschwig, K., Okihara, K.L., Muller, F., Tijsterman, M., and Plasterk, R.H. (2006). Gene expression: long-term gene silencing by RNAi. Nature *442*, 882.

Waddington, C.H. (2012). The epigenotype. 1942. Int J Epidemiol *41*, 10-13.

Wang, W., Han, B.W., Tipping, C., Ge, D.T., Zhang, Z., Weng, Z., and Zamore, P.D. (2015). Slicing and Binding by Ago3 or Aub Trigger Piwi-Bound piRNA Production by Distinct Mechanisms. Mol Cell *59*, 819-830.

Wang, X., Zhao, Y., Wong, K., Ehlers, P., Kohara, Y., Jones, S.J., Marra, M.A., Holt, R.A., Moerman, D.G., and Hansen, D. (2009). Identification of genes expressed in the hermaphrodite germ line of C. elegans using SAGE. BMC Genomics *10*, 213.

Waterhouse, P.M., Wang, M.B., and Lough, T. (2001). Gene silencing as an adaptive defence against viruses. Nature *411*, 834-842.

Wedeles, C.J., Wu, M.Z., and Claycomb, J.M. (2013a). A multitasking Argonaute: exploring the many facets of C. elegans CSR-1. Chromosome Res *21*, 573-586.

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

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

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

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

Youngman, E.M., and Claycomb, J.M. (2014). From early lessons to new frontiers: the worm as a treasure trove of small RNA biology. Front Genet *5*, 416.

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