2014-09-26

Small RNAs and Argonautes Provide a Paternal Epigenetic Memory of Germline Gene Expression to Promote Thermotolerant Male Fertility: A Dissertation

Colin C. Conine
University of Massachusetts Medical School

SMALL RNAS AND ARGONAUTES PROVIDE A PATERNAL EPIGENETIC MEMORY OF GERMLINE GENE EXPRESSION TO PROMOTE THERMOTOLERANT MALE FERTILITY

A Dissertation Presented

By

COLIN C. CONINE

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

September 26, 2014

Molecular Biology and Genetics

SMALL RNAS AND ARGONAUTES PROVIDE A PATERNAL EPIGENETIC MEMORY OF GERMLINE GENE EXPRESSION TO PROMOTE THERMOTOLERANT MALE FERTILITY

A Dissertation Presented By

Colin Christopher Conine

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

_____
Craig Mello, Ph.D., Thesis Advisor

_____
Victor Ambros, Ph.D., Member of Committee

_____
David Page, M.D., Member of Committee

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

_____
Eric Sontheimer, 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

_____
Marian Walhout, 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
September 26, 2014

ACKNOWLEDGEMENTS

This thesis would not be possible without the support of many colleagues, friends, and family. Many thanks to my advisor, Craig Mello, who not only granted me the intellectual freedom to pursue anything that interested me scientifically, but enthusiastically encouraged me to do so. This freedom paired with his support and creative thinking, has been critical to my development as an independent scientist. To that end, I am also indebted to former Mello lab members Julie Claycomb, Elaine Youngman, Pedro Batista, Darryl Conte, Weifeng Gu, and Hsin-Yue Tsai, who have been indispensable mentors in my training as an experimentalist and a thinker. I would also like to thank current lab members Daniel Chaves, Sandra Vergara, Masaki Shirayama, Rita Sharma, and Takao Ishidate who have either aided directly in experimental support or in the ideology of my research through scientific discussion. With regards to the writing of this thesis, I am grateful to Darryl Conte, Pedro Batista, Julie Claycomb and Heidi Hehnly, who have read and commented on various sections of this manuscript.

I want to express my gratitude to the chair of my thesis committee, Marian Walhout, as well as Victor Ambros and Sean Ryder, who have informally served as co-advisors during my graduate research, providing constant support and great ideas that have influenced the course of this thesis. As well, as the other members of my thesis committee, Eric Sontheimer and David Page of M.I.T. for taking the time to read this thesis and participate in its defense.

Finally, my deepest appreciation goes out to my friends and family who have supported me immensely not only during my time in graduate school, but also my entire life. Most importantly, I thank my parents, who have continuously sacrificed for my greater well-being and instilled in me lessons and values that are fundamental to all of my success. I also would like to give special thanks to my Grandfather Ken who inspired my passion for biology at a very young age. Without my family I would not be where I am today.

ABSTRACT

During each life cycle, gametes must preserve and pass on both genetic and epigenetic information, making the germline both immortal and totipotent. In the male germline the dramatic morphological transformation of a germ cell through meiosis, into a sperm competent for fertilization, while retaining this information is an incredible example of cellular differentiation. This process of spermatogenesis is inherently thermosensitive in numerous metazoa ranging from worms to man. Here, I describe the role of two redundant AGO-class paralogs, ALG-3/4, and their small RNA cofactors, in promoting thermotolerant male fertility in *Caenorhabditis elegans*. *alg-3/4* double mutants exhibit temperature dependent sterility resulting from defective spermiogenesis, the postmeiotic differentiation of haploid spermatids into spermatozoa competent for fertilization. The essential Argonaute CSR-1 functions with ALG-3/4 to positively regulate target genes required for spermiogenesis by promoting transcription via a small RNA positive feedback loop. Our findings suggest that ALG-3/4 functions during spermatogenesis to amplify a small-RNA signal loaded into CSR-1 to maintain transcriptionally active chromatin at genes required for spermiogenesis and to provide an epigenetic memory of male-specific gene expression. CSR-1, which is abundant in mature sperm, appears to transmit this memory to offspring. Surprisingly, in addition to small RNAs targeting male-specific genes, we show that males also harbor an extensive repertoire of CSR-1 small RNAs targeting oogenesis-specific mRNAs. The ALG-3/4 small RNA pathway also initiates silencing small RNA signals loaded into WAGO

Argonautes, which function to posttranscripitonally silence their target mRNAs. Silencing WAGO/small RNA-complexes are present in sperm and presumably transmitted to offspring upon fertilization. Together these findings suggest that *C. elegans* sperm transmit not only the genome but also epigenetic activating and silencing signals in the form of Argonaute/small-RNA complexes, constituting a memory of gene expression in preceding generations.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

PREFACE

The chapters of this dissertation have appeared in separate publications:

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. PNAS *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

For the sperm proteomic data presented in chapter III we collaborated with James Moresco from John Yates' lab at The Scripps Research Institute in La Jolla, California. James performed Multidimensional Protein Identification Technology on sperm protein samples to quantitate the proteins present.

CHAPTER I

The *C. elegans* Germline: A Small RNA World

Introduction

The idea that life originated from self-replicating RNA molecules assembled from the 'primordial soup' present on primitive Earth was first described nearly 50 years ago by Francis Crick, Leslie Orgel, and Carl Woese (Crick, 1968; Orgel, 1968; Woese, 1967). After the discovery of catalytic RNAs, or ribozymes, in the 1980s the idea gained undeniable recognition and was termed 'The RNA World' hypothesis (Gilbert, 1986). Since then hundreds of different ribozymes have been isolated with diverse activities, including self-replicating RNAs and RNAs with the ability to catalyze peptide bond formation, ultimately resulting in protein synthesis (Bartel and Unrau, 1999; Cech, 2009).

In the late 1990s the field of RNA biology exploded again with the discovery of RNA interference (RNAi) in the nematode *Caenorhabditis elegans*, wherein double-stranded RNA molecules processed into small interfering RNAs (siRNAs) repress gene expression of mRNAs with complementary nucleotide sequence (Fire et al., 1998; Zamore et al., 2000). In parallel this phenomenon was described in plants (Hamilton and Baulcombe, 1999), and subsequently discovered to be present in fungi, protists, and most metazoans, including humans (Cerutti and Casas-Mollano, 2006; Elbashir et al., 2001). siRNAs can repress gene expression by affecting mRNA stability, mRNA translation, or mRNA synthesis through changes in chromatin structure, depending on the organism and the molecular machinery present (Denli and Hannon, 2003).

Shortly thereafter it was realized that nearly all eukaryotes harness the power and machinery of RNAi to regulate the expression of genomically encoded mRNAs, to regulate a vast array of biological process including but not limited to, developmental timing, meiosis, and cellular differentiation. Cells employ endogenous RNAi (endo-RNAi) by generating small RNAs from genomically encoded double stranded RNAs in the case of endo-siRNAs, genomically encoded hairpin RNAs in the case of microRNAs (miRNAs), nucleolytic processing of longer RNA transcripts in the case of piRNAs, and by RNA-dependent-RNA polymerases that can generate small RNAs *de novo* from an RNA template (Ghildiyal and Zamore, 2009).

Recently, it has been demonstrated that endogenous small RNAs are present in the gametes of plants and numerous metazoans, including worms, flies, mice, and humans (Conine et al., 2010; Garcia-Lopez et al., 2014; Grant-Downton et al., 2009; Gu et al., 2009; Krawetz et al., 2011; Malone et al., 2009; Peng et al., 2012; Stoeckius et al., 2014; Tam et al., 2008; Watanabe et al., 2006). In the gametes, small RNAs can act as carriers of epigenetic information delivered during fertilization, through the regulation of transcripts important for the early embryo, producing phenotypes that persist across development (Brennecke et al., 2008; Burton et al., 2011; Liu et al., 2012). Small RNAs can even be transmitted through the gametes across multiple generations, to elicit regulatory outcomes and phenotypes transgenerationally (Alcazar et al., 2008; Ashe et al., 2012; Buckley et al., 2012; Conine et al., 2013; Fang et al., 2012;

Gapp et al., 2014; Jablonka and Raz, 2009; Shirayama et al., 2012). Epigenetic inheritance of small RNAs could potentially explain some cases of inheritance, where phenotypic traits that cannot be explained by genotypic polymorphisms are clearly transmitted from parent to progeny (genetic inheritance) (Maher, 2008). Whether the 'RNA World' existed or not may be impossible to scientifically prove, however, it is clear that RNA can no longer be viewed merely as a transient carrier of information from DNA to protein, or as a structural entity in tRNAs or rRNAs, but now must be viewed as a master regulator of all aspects of cellular and developmental biology. No place is this more evident than where RNAi was originally discovered: the *C. elegans* germline, where a 'Small RNA World' dictates almost all aspects of its biology. Here, I describe the various roles of small RNA pathways in silencing foreign genetic elements, promoting the expression of essential germline genes, and promoting both male and female fertility. To do so, I will describe the factors and mechanisms involved in the biogenesis of each class of small RNA and the Argonautes they are associated with. Where applicable I will relate the worm small RNA pathways to analogous pathways in other organisms, particularly mammals.

Small RNA's best friend: Argonaute

Small RNAs can be defined by numerous physical properties, as well as by the machinery required for their synthesis. However, the most defining character for any small RNA is the Argonaute protein it associates with.

Argonautes are the direct binding partners of small RNAs and are required to impose regulatory outcomes when the Argonaute/small RNA complex finds a target RNA through complementary base pairing with the small RNA or ‘guide’. Therefore, Argonautes are at the heart of all small RNA pathways and actually define the functional outcome of the small RNA that is bound.

The Argonaute protein family was first identified in plants as being important for development, with the mutants resembling an argonaute squid, hence the name (Bohmert et al., 1998), and in *Drosophila melanogaster* as being required for germline stem-cell division (Lin and Spradling, 1997). Argonautes were first associated with small RNA pathways in *C. elegans* through genetic screens for mutants deficient for RNAi, where the Argonaute gene *rde-1* was found to be absolutely required for RNAi (Tabara et al., 1999). Later it was shown that Argonautes directly interact with the small RNA and use it as a guide to find target RNAs with complementary sequences (Wang et al., 2008a; Wang et al., 2008b).

Argonautes are ~100kDa highly basic proteins that are characterized by having PAZ (Piwi-Argonaute-Zwille), MID (middle) and Piwi domains. Based on structural studies the PAZ domain anchors the 3′ end of the small RNA while the MID domain binds the 5′ end. The Piwi domain, which is structurally similar to RNase H, can act as an endonuclease (slicer) to cleave (slice) target RNAs through a DDX catalytic triad (X can be H or D) of amino acids (Meister, 2013). Using these credentials, Argonautes have been identified (and found associated

with small RNA pathways) across all forms of life including plants, archaea, bacteria, fungi, and metazoa (Cerutti and Casas-Mollano, 2006; Makarova et al., 2009), suggesting that Argonautes and small RNA pathways might have been present in the last universal common ancestor.

Eukaryotic Argonaute proteins can be separated into 3 phylogenetic clades (Tolia and Joshua-Tor, 2007): AGO, Piwi, and worm-specific Argonautes (WAGO). The AGO clade, based on homology to *Arabidopsis thaliana* AGO1, includes all of the plant Argonautes, the human Argonautes AGOs-1-4, the *C. elegans* miRNA effectors ALG-1/2, and primary endogenous small RNA effectors ALG-3/4 (described below). The Piwi clade, based on homology to Drosophila Piwi, interact exclusively with a type of endogenously encoded small RNAs termed piRNAs (for Piwi-interacting RNAs), expressed predominantly in the germline of metazoans, but also present in protists. The *C. elegans* Argonautes PRG-1/2 (described in detail below) are members of the Piwi clade. Finally, the WAGO clade is an expanded family of worm-specific Argonautes (Yigit et al., 2006), many of which lack one or more of the DDX residues in the Piwi domain (Joshua-Tor and Hannon, 2011). Of 27 annotated Argonaute genes in *C. elegans*, 12 encode WAGO Argonautes (Gu et al., 2009; Yigit et al., 2006).

The beginning: RNAi

RNAi can be induced experimentally in the *C. elegans* germline by direct injection of double-stranded RNA (dsRNA) (Fire et al., 1998), soaking worms in dsRNA (Tabara et al., 1998), feeding with bacteria expressing dsRNA (Timmons and Fire, 1998), or by expressing a transgene that produces dsRNA (Tavernarakis et al., 2000). RNAi in worms is systemic: after initial exposure to dsRNA, the silencing signal can spread throughout the body (including the germline) of the animal (Fire et al., 1998). The silencing signal can also be inherited (Vastenhouw et al., 2006). Most worm germline-expressed mRNAs appear to be susceptible to RNAi triggered by any of the four methods, with the exception of mRNAs expressed during spermatogenesis, which are refractory to RNAi for currently unknown reasons (Kamath et al., 2003; Reinke et al., 2004). One hypothesis for the inability of sperm genes to be silenced by RNAi is that some components of the RNAi pathway are also used for endogenous small RNA pathways functioning during spermatogenesis, leading to competition between the pathways ((Conine et al., 2010; Duchaine et al., 2006; Han et al., 2009; Kennedy et al., 2004; Pavelec et al., 2009; Welker et al., 2010; Yigit et al., 2006), see below).

To elicit the silencing response the dsRNA is first recognized by a dsRNA binding protein RDE-4 and the RNase III related enzyme Dicer (DCR-1) that ‘Dices’ (endonucleolytic cleavage) the dsRNA into 21-23nt primary (1°) siRNAs, that have a 5′ monophosphate and a 3′ hydroxyl (Ketting et al., 2001; Tabara et

al., 2002) (Figure 1.1 and Table 1). Dicer is a highly conserved enzyme required for the vast majority of exogenous RNAi pathways and for some endogenous RNAi-related pathways ((Cerutti and Casas-Mollano, 2006), see below). For example, in *C. elegans* DCR-1 is the only Dicer and is essential for both RNAi and miRNA pathways (Grishok et al., 2001; Ketting et al., 2001; Knight and Bass, 2001; Tabara et al., 2002).

After processing by Dicer, 1° siRNAs are loaded onto the Argonaute RDE-1 (Tabara et al., 1999; Yigit et al., 2006), which then uses the sequence of the small RNA as a guide to basepair with RNA targets. Upon finding a target RNA with complementarity to the small RNA loaded in the Argonaute, RDE-1 recruits a multi-protein molecular machine containing an RNA-dependent-RNA polymerase (RdRP) (Figure 1.1). This machine uses the target RNA as a template to synthesize secondary (2°) small RNAs named 22G-RNAs, which are predominantly 22 nt long with a triphosphorylated guanosine at the 5′ end (Pak and Fire, 2007; Sijen et al., 2007; Yigit et al., 2006) (Table 1). In the germline, 22G-RNAs are produced by a combination of two RdRPs, RRF-1 and EGO-1 (Gu et al., 2009). This secondary small RNA production step acts to amplify the siRNA response (Pak and Fire, 2007; Sijen et al., 2007; Yigit et al., 2006) (Figure 1.1). The 2° small RNAs are loaded onto WAGO Argonautes, WAGO-(1-12), as it requires removing all 12 of these WAGO genes to fully disrupt the RNAi response (Gu et al., 2009). The molecular mechanism for how these Argonautes

function to repress the expression of mRNAs is unknown, especially since many lack the catalytic residues for slicer activity.

In *C.elegans* the induction of germline RNAi leads to the inherited silencing of the target RNA for at least 3-4 generations and in some cases upon selection, indefinitely (Alcazar et al., 2008; Claycomb et al., 2009; Grishok et al., 2000; Luteijn et al., 2012; Shirayama et al., 2012; Vastenhouw et al., 2006; Ashe, et al., 2012). Long-term silencing (>4 generations) requires the nuclear RNAi pathway, and the Argonaute WAGO-9/HRDE-1. In this pathway WAGO-9 loaded with 22G-RNAs amplified from the initially dsRNA targeted locus, enters the nucleus and along with the NRDE (nuclear RNAi deficient) machinery recruits chromatin modifying enzymes that mark the gene of the targeted RNA with silencing Histone H3 Lysine 9 methylation (H3K9me) marks (Buckley et al., 2012; Burkhart et al., 2011; Guang et al., 2008; Ashe et al., 2012; Luteijn et al., 2012; Shirayama et al., 2012) (Figure 1.1). This silencing can then be inherited indefinitely (>30 generations) through transmission of the silenced chromatin and small RNAs (Ashe et al., 2012; Claycomb et al., 2009; Gassmann et al., 2012; Luteijn et al., 2012; Shirayama et al., 2012).

The systemic and heritable properties of RNAi are characteristics of RNAi in both nematodes and plants, perhaps consistent with the presence of RdRP in both lineages (Calarco and Martienssen, 2011; Dunoyer et al., 2010; Van Ex et al., 2011). The heritability of experimentally induced RNAi has yet to be explored

in many organisms, but may yet be found in other forms of life, as it's becoming clear that gametes can transmit small RNAs.

The actual beginning: miRNAs

In the early 1990's before the discovery and characterization of RNAi, Victor Ambros and Gary Ruvkun's laboratories determined that a posttranscriptional repressor of *lin-14* (an important regulator of postembryonic development in *C. elegans*), *lin-4* was actually a noncoding small RNA gene (Lee et al., 1993; Wightman et al., 1993). It was shown previously that *lin-4* temporally controlled postembryonic development by negatively regulating LIN-14 protein levels (Ambros, 1989). Interestingly, along with the discovery that *lin-4* encodes a small RNA, they also found that the *lin-14* mRNA 3′ UTR contained 7 stretches of 10 nt complementarity to *lin-4*, and that these sequences were required for the posttranscriptional regulation of *lin-14* by *lin-4* (Lee et al., 1993; Wightman et al., 1993). They went on to postulate that the posttransriptional or translational regulation of *lin-14* by *lin-4* small RNA occurred by an antisense mechanism where *lin-4* binds to its complementary sites in the *lin-14* 3′ UTR.

With the discovery of a second genomically encoded small RNA that regulates *C. elegans* development, *let-7,* which is conserved in Drosophila and humans (Pasquinelli et al., 2000; Reinhart et al., 2000), genomically encoded regulatory small RNAs with biological functions were discovered in organisms ranging from plants to mammals, and were named microRNAs (miRNAs) (Lagos-

Quintana et al., 2001; Lau et al., 2001; Lee and Ambros, 2001; Reinhart et al., 2002). Prior to that it was also demonstrated in *C. elegans* that Dicer, and the Argonautes ALG-1 and ALG-2, are required for the maturation and function of both the *let-7* and *lin-4* miRNAs (Grishok et al., 2001). These findings set the groundwork for defining the miRNA pathway across phyla, which is now intimately connected with most aspects of biology.

miRNAs are genomically encoded 21-23 nt small RNAs that are processed from short dsRNA hairpins occurring in long primary transcripts (pri-mRNAs) or from the introns of pre-mRNAs (Cai et al., 2004; Lee et al., 2002; Lee et al., 2004; Ruby et al., 2007) (Figure 1.1). Maturation of a pri-miRNA into a miRNA requires two processing steps. In the nucleus, a pri-mRNA transcript is processed by the RNAse III-related endonuclease Drosha into a 60-70nt long pre-miRNA (Denli et al., 2004; Lee et al., 2003). A correctly processed pre-miRNA is bound by Exportin-5 and transported to the cytoplasm (Yi et al., 2003)(Figure 1.1). In the cytoplasm, the loop of the pre-miRNA hairpin is removed by Dicer, creating a dsRNA duplex approximately 22nt in length with 5′ monophosphate and 3′ hydroxyl groups (Hutvagner et al., 2001) (Table 1). One strand, the mature miRNA or guide strand, is then loaded into Argonaute (Chendrimada et al., 2005; Grishok et al., 2001; Hutvagner et al., 2001; Martinez et al., 2002; Saito et al., 2005; Schwarz et al., 2002) (Figure 1.1). The other strand, known as miRNA* or passenger strand, is normally released and degraded. In some cases, however, mature miRNAs are generated from both

arms of the pre-miRNA hairpin. The choice of miRNA strand loaded into Argonaute is largely determined by the thermodynamic properties of the miRNA/miRNA* duplex (Khvorova et al., 2003; Schwarz et al., 2003).

Argonaute then uses the miRNA as a guide to find target RNAs with complementary sequences. When a target site with extensive sequence complementarity is bound, animal miRNAs can direct Argonaute catalyzed nucleolytic cleavage of the RNA (Hutvagner and Zamore, 2002; Song et al., 2004; Yekta et al., 2004). Targets like this are rare in animals. More commonly, complementarity between nucleotides 2-7 of the miRNA, known as the 'seed' sequence, and a target mRNA is sufficient to direct mRNA destabilization or translation repression (Bartel, 2009; Lai, 2002) (Figure 1.1). Therefore, a given miRNA can direct the regulation of multiple target RNAs resulting in most RNAs being regulated by miRNAs in some context (Friedman et al., 2009).

In *C. elegans* miRNAs are loaded onto the AGO-clade Argonautes ALG-1 and ALG-2, which are also required for the maturation of pre-miRNAs into mature miRNAs (Figure 1.1). They are at least partially redundant: mutants lacking either gene produce only mild phenotypes, while a mutant lacking both genes is lethal at the embryo or early larval stages of development (Grishok et al., 2001). The *C. elegans* genome encodes 159 miRNAs, nearly all of which are conserved in related nematodes, but only half of which share sequence homology with miRNAs encoded by the fly and human genomes (Ibanez-Ventoso et al., 2008; Lim et al., 2003). In the worm, the only microRNA or miRNA family (miRNAs with

related seed sequences) that is clearly enriched in the germline is *mir-35-42*, which is the most highly expressed microRNA in oocytes (Alvarez-Saavedra and Horvitz, 2010; Gu et al., 2009; Wu et al., 2010). Consistent with *mir-35-42* expression in oocytes, embryonic lethality (Alvarez-Saavedra and Horvitz, 2010), low fecundity (McJunkin and Ambros, 2014), and RNAi hypersensitivity (Massirer et al., 2012) of *mir-35-42* mutants show a partial maternal rescue. Interestingly, maternal and early embryonic expression of *mir-35-42* is required to promote full sperm production and function; thus, microRNA expression in the female germline promotes development of the male germline in the subsequent generation (McJunkin and Ambros, 2014). The *mir-35-42* family sequence does not share homology with any fly or human miRNAs (Ibanez-Ventoso et al., 2008). miR-34c, on the other hand, is conserved in mammals and worms and is expressed in sperm. In mice, miR-34 in the sperm is transmitted to the zygote, where it is important for the first zygotic cell division (Liu et al., 2012). *mir-34* is abundant in *C. elegans* sperm, but its function in sperm is unexplored (Conine et al., 2010).

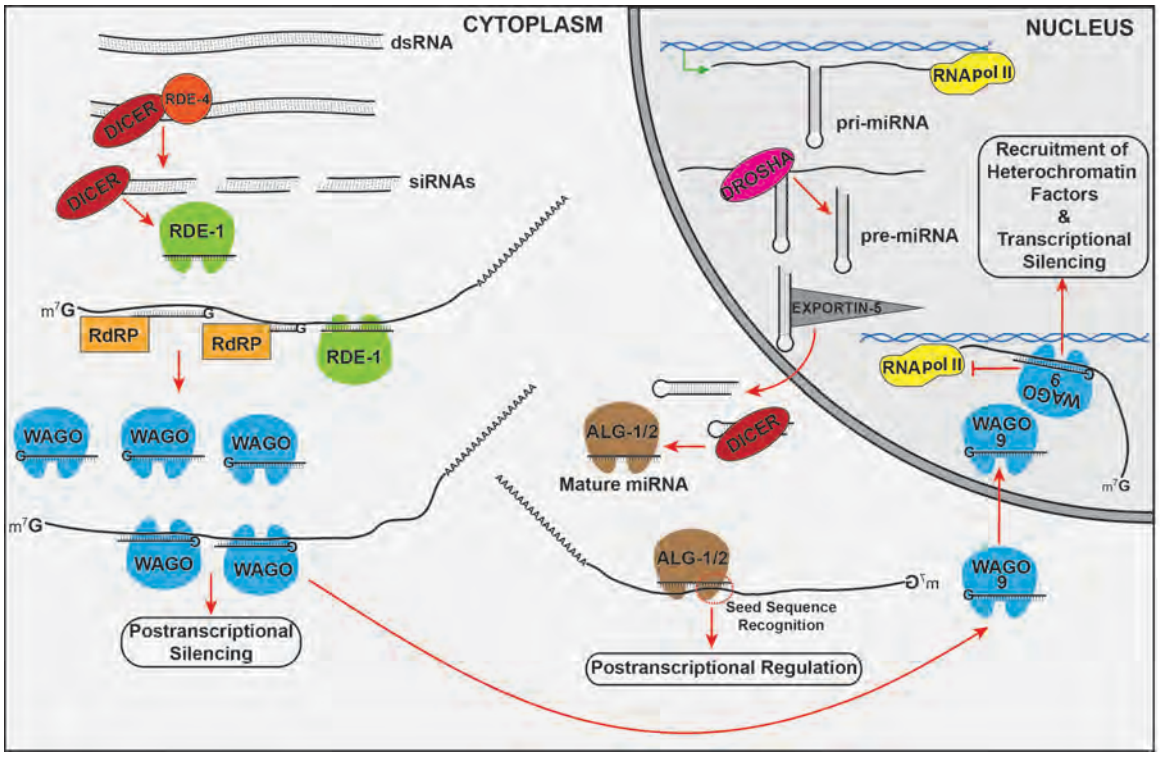

Figure 1.1. The RNAi and miRNA pathways in *C. elegans*
The RNAi pathway to the left, beginning by the recognition of dsRNA in the cytoplasm and the miRNA pathway to the right, beginning with the transcription of a miRNA gene in the nucleus.

piRNAs: Surveyors of germline RNAs

piRNAs (Piwi-interacting RNAs) are endogenously derived small RNAs, expressed predominantly in the metazoan germline. piRNAs were initially identified in Drosophila as rasiRNAs (repeat-associated small interfering RNAs) complementary to repetitive transposable element sequences (Aravin et al., 2003; Chen et al., 2005). rasiRNAs were later named piRNAs after they were found to associate with Piwi-clade Argonautes (Brennecke et al., 2007; Houwing et al., 2007; Saito et al., 2006; Vagin et al., 2006). Shortly thereafter, piRNAs were identified in the testes of mammals, in association with Piwi-clade Argonautes (Aravin et al., 2006; Girard et al., 2006; Grivna et al., 2006a; Lau et al., 2006). With the advent of next generation sequencing, piRNAs have been identified in metazoa ranging from sponges to humans (Lau, 2010).

Unlike miRNAs, piRNA production does not require Dicer processing (Vagin et al., 2006). Rather piRNAs are processed from single-stranded RNA precursors that may originate from transposons or repetitive sequences, non-coding loci, or protein-coding loci, depending on the organism, cell-type, or Piwi protein being examined. Mutations that interfere with the biogenesis of piRNAs result in fertility defects in worms, flies, and mammals.

piRNAs have been well characterized in Drosophila and mouse model systems. In Drosophila, 3 classes of piRNAs exist depending on their origin (Han and Zamore, 2014): the first two derived from distinct genomic regions termed

‘piRNA clusters’ and the third from the 3′ UTRs of mRNAs. The first are piRNAs generated from transposon and repetitive sequences in the germline by ‘ping-pong’ amplification. In the ping-pong model, bidirectional transcription leads to the production of 24-30 nt 1° piRNAs antisense to transposons that bind the Argonautes Aubergine (Aub) and Piwi which direct cleavage of the sense RNA. This cleavage generates sense piRNAs that are loaded into Argonaute 3 (Ago3), in turn directing cleavage of the antisense strand and generation of more antisense piRNAs, thus creating a feed-forward amplification loop of piRNA biogenesis (Brennecke et al., 2007; Gunawardane et al., 2007). This process occurs in the nucleus and the perinuclear ‘nuage’ (germline granules), as Piwi is localized to the nucleus and Aub and Ago3 to nuage. Loss of function of Piwi proteins and piRNAs leads to derepression of transposons and severe defects in gametogenesis and fertility (Klattenhoff and Theurkauf, 2008). Most Drosophila piRNA-pathway mutants reduce male fertility, which is linked to the massive overexpression of the Stellate protein encoded by a repeated gene on the X-chromosome, leading to crystal formation in the testes (Aravin et al., 2001). Overexpression of Stellate in piRNA mutants is due to the loss of piRNAs derived from the Y-chromosome-linked *Suppressor of Stellate* locus, which function to silence *Stellate* in trans in the male germline (Aravin et al., 2001).

The second type of piRNAs in Drosophila are derived from single-stranded RNA precursors and loaded specifically into Piwi. They are transcribed mostly from one piRNA cluster, the *flamenco* locus containing embedded transposon

fragments, in the somatic follicle cells of the drosophila ovary. In the soma, the ping-pong mechanism does not occur and the piRNA signal is not further amplified (Li et al., 2009; Malone et al., 2009). The *flamenco* locus was first identified as a suppressor of transposon expression in somatic follicle cells, supporting the evidence that piRNAs derived from this locus function to silence transposons (Pelisson et al., 1994).

Finally, the third type of piRNAs, also found in ovarian somatic cells, derive from the 3′ UTRs of hundreds of mRNAs and are loaded onto Piwi (Robine et al., 2009). These 1° piRNAs do not lead to ping-pong amplification, and their function is mostly mysterious. Nonetheless, 3′ UTR piRNAs from the *traffic jam* (*tj*) mRNA may negatively regulate *tj* mRNA, which is elevated in *piwi* mutants (Robine et al., 2009).

In mammals three similar classes of piRNA pathways also exist, and function predominantly in the male germline. Like Drosophila, most piRNAs that are non-genic are expressed from clusters that can exceed 100kb, with each cluster producing numerous piRNAs that can have overlapping sequences, generating an extraordinary diversity of piRNA sequences. In prenatal mice, piRNAs are first expressed in the primordial germ cells (PGCs) undergoing mitosis (Aravin et al., 2007). Here, the majority of piRNAs are derived from transposons, with many matching short interspersed elements (SINEs), long interspersed elements (LINEs), and long terminal repeat (LTR) retrotransposons. Although, a significant percentage of prenatal piRNAs also map to genic

sequences, concentrated mostly in 3′ UTRs analogous to those in Drosophlia, however their functions are unknown (Aravin et al., 2007; Robine et al., 2009). Prenatal piRNAs also named the ‘pre-pachytene’ (for their expression prior to the pachytene stage of meiosis) piRNAs are ~26 nts long, in the sense orientation of the derived RNA, and bind the Argonaute MILI, which is expressed beginning in the PGCs, in spermatogonial stem cells, and throughout meiosis up until the formation of haploid round spermatids (Aravin et al., 2006; Aravin et al., 2008; Aravin et al., 2007). Another Argonaute, MIWI2 is expressed specifically in PGCs and spermatogonial stem cells, slightly later than MILI, and binds ~28nt piRNAs antisense to the same transposons as MILI (Aravin et al., 2008). Sequence analysis of the piRNAs bound to these Argonautes indicates that ping-pong amplification also happens in the mouse male germline with pre-pachytene piRNAs, between MILI-MIWI2 and MILI-MILI (Aravin et al., 2008). MILI is localized to perinuclear germ granules analogous to those in Drosophila, while MIWI2 is expressed in the nucleus, perinuclear germ granules, and cytoplasm, indicating that ping-pong amplification occurs in similar cellular compartments in both organisms (Aravin et al., 2008)

Loss-of-function MILI and MIWI2 mutants are sterile due to a complete block in spermatogenesis in the early prophase of meiosis I (Carmell et al., 2007; Kuramochi-Miyagawa et al., 2004). This is potentially due to increased expression of transposable elements, as several classes are up-regulated in the testes of MILI and MIWI2 mutants (Aravin et al., 2008; Aravin et al., 2007;

Carmell et al., 2007; Kuramochi-Miyagawa et al., 2008). In mouse, transposons are repressed transcriptionally via DNA methylation at promoter regions, and it has been shown that nuclear MIWI2 initiates transcriptional silencing downstream of primary targeting by MILI by promoting the deposition of these marks through the DNA methyltransferase Dnmt3L. MILI also degrades transposon RNAs posttranscriptionally by utilizing piRNA guided slicer activity (De Fazio et al., 2011).

During spermatogenesis, another class of mammalian piRNAs, the 'pachytene piRNAs' are highly expressed beginning at the pachytene stage of meiosis and persisting through sperm development. Pachytene piRNAs are ~30 nt long and interact predominantly with a third Piwi clade member MIWI, which is coordinately expressed during spermatogenesis (Aravin et al., 2006; Grivna et al., 2006a; Robine et al., 2009). MIWI is required for the production or stability of pachytene piRNAs, as the testes of *miwi* mutant mice are depleted for these small RNA species (Grivna et al., 2006a). MIWI is localized to the cytoplasm and to the perinuclear chromatoid body, which is a germ granule related to nuage in Drosophila and P granules in *C. elegans* (Beyret and Lin, 2011; Kotaja et al., 2006; Voronina et al., 2011). Pachytene piRNAs are processed in the sense orientation from long-noncoding RNAs transcribed by RNA polymerase II, and unlike pre-pachytene piRNAs their sequences map uniquely to their genomic origin (Li et al., 2013). These piRNA genes are genomically located in clusters and their transcription is driven by the transcription factor A-MYB. A similar

regulatory mechanism exists in roosters, indicating the pathway predates the divergence of birds and mammals (Li et al., 2013). MIWI also associates with piRNAs derived from the 3′ UTRs of mRNAs, many overlapping with MILI (as mentioned above), however, the biogenesis and function of these are also unknown (Robine et al., 2009).

Loss-of-function mutants for *miwi* are sterile as a result of a complete arrest at the round spermatid stage of spermiogenesis (haploid cellular differentiation) (Deng and Lin, 2002). The function of MIWI and pachytene piRNAs during spermatogenesis is a mystery. However, two conflicting hypotheses have been postulated. The first is that MIWI stabilizes spermiogenic mRNAs (mRNAs required post-meiotically) and regulates their translation, as MIWI interacts with mRNAs and has also been shown to interact with the translational machinery (Grivna et al., 2006b; Nishibu et al., 2012; Vourekas et al., 2012). The second is that MIWI and piRNAs through miRNA-like basepairing (seed recognition) are required to eliminate mRNAs during the post-meiotic differentiation of round spermatids to motile spermatozoa (known as spermiogenesis) (Gou et al., 2014). At this point, neither mechanism clearly explains the biological function of Miwi and pachytene piRNAs or why *miwi* and other pachytene piRNA mutants arrest at the round spermatid stage in meiosis.

In *C. elegans* only one class of piRNA exists, termed 21U-RNAs because they are predominantly 21 nt long and begin with a uridine 5′-monophosphate (Table 1). The bias to begin with a uridine 5′-monophosphate is a character also

exhibited in the primary piRNAs of other organisms, including flies and mammals (Aravin et al., 2008; Aravin et al., 2007; Batista et al., 2008; Brennecke et al., 2007; Das et al., 2008; Robine et al., 2009). Their 3′ termini is a hydroxyl group, however, they are posttranscriptionally modified by 2′-O-methylation at the 3′ most nucleotide via the methyltransferase HENN-1 (Billi et al., 2012; Kamminga et al., 2012; Montgomery et al., 2012) (Table 1). This modification is also present at the 3′ termini of piRNAs in flies, fish and mammals, where it is deposited by homologs of the same enzyme (known as Hen1 in other organisms) and is thought to stabilize the small RNA by preventing 3′-to-5′ exonucleolytic processing (Houwing et al., 2007; Kamminga et al., 2010; Kirino and Mourelatos, 2007a, b; Saito et al., 2007; Vagin et al., 2006).

21U-RNAs were originally identified in large-scale sequencing datasets as a unique class of small RNAs apparently expressed from minigenes that share a conserved sequence motif upstream of the 21U-RNA sequence. The loci are present in two broad regions on chromosome IV, with many in dense intergenic clusters, some near protein coding genes, and others within introns (Ruby et al., 2006). 21U-RNAs were identified as piRNAs, because of their association with the Piwi-clade Argonaute PRG-1 (Batista et al., 2008; Das et al., 2008; Wang and Reinke, 2008) (Table 1). The *C. elegans* genome encodes two canonical Piwi Argonautes, PRG-1 and PRG-2, although PRG-2 likely has little or no function (Batista et al., 2008; Das et al., 2008). Like mammalian pachytene piRNAs, 21U-RNAs are diverse in sequence, with greater than 30,000 unique

small RNAs—the overwhelming majority lacking perfectly complementary targets.

PRG-1 and 21U-RNAs are expressed exclusively in the male and female germlines. Intriguingly, PRG-1 is localized to P granules, the perinuclear germ granules of *C. elegans,* analogous to the localization of PIWIs in flies and mammals. Loss-of-function mutants for *prg-1* exhibit a temperature sensitive (TS) fertility defect. At the permissive temperature of 20°C, *prg-1* mutants produce around one-third the number of the progeny as wild type (WT), and at the non-permissive temperature of 25°C *prg-1* mutants are essentially sterile, with some animals producing a handful of progeny compared to the 150-200 progeny produced by WT at 25°C (Batista et al., 2008; Wang and Reinke, 2008). At all temperatures, *prg-1* mutants exhibit a significant reduction in both mitotic and meiotic germ cells (Batista et al., 2008; Das et al., 2008). Providing WT sperm to *prg-1* females by mating, only partially rescues the fertility phenotypes, indicating that the phenotype is a result of male and female germline (sperm and oocyte) defects (Batista et al., 2008; Das et al., 2008; Wang and Reinke, 2008). *prg-1* mutant sperm have defects in haploid post-meiotic differentiation (spermiogenesis), and arrest as round spermatids at 25°C (Wang and Reinke, 2008), very similar to mouse *miwi* (pachytene piRNA) mutants (Deng and Lin, 2002), suggesting that worm and mouse Piwi proteins regulate a conserved aspect of sperm development. 21U-RNAs require PRG-1 for their biogenesis at all temperatures, as 21U-RNAs are significantly depleted in *prg-1* mutants

(Batista et al., 2008; Das et al., 2008; Wang and Reinke, 2008). These findings suggest that, besides maintaining the number of germ cells at all temperatures, PRG-1 and 21U-RNAs function to facilitate a temperature-dependent germline process required for normal fertility (Batista et al., 2008).

Two types of 21U-RNAs exists, type-1 21U-RNAs generated from piRNA loci mostly on chromosome IV and type-2 21U-RNAs derived from short capped small RNAs (csRNAs) that are transcribed bidirectionally at RNA polymerase II transcriptional start sites (TSSs) (Gu et al., 2012) (Figure 1.2). Unlike mammalian piRNAs, type-1 21U-RNAs do not appear to be processed from long noncoding RNA precursors, instead, they derive from ~15,000 individual gene-like loci. Upstream (~40bp) from the majority of type-1 piRNAs is an 8nt motif ‘CTGTTTCA’, referred to as the ‘Ruby motif’ (Ruby et al., 2006; Weick et al., 2014). This motif is thought to act as a promoter for 21U-RNA transcription, as it is required for the biogenesis of 21U-RNAs, and is recognized by forkhead transcription factors (Billi et al., 2013; Cecere et al., 2012). However, forkhead transcription factors are expressed mostly in somatic cells, indicating that other transcriptional activators are also required to transcribe type-1 21U-RNAs (Cecere et al., 2012). An A/T-rich spacer sequence is also present between the Ruby motif and the 21U-RNA sequence, which is thought to create nucleosome-free DNA to promote the transcription of the 21U-RNA precursor (Cecere et al., 2012).

21U-RNA precursors are transcribed by RNA Pol II ~40nt downstream of the Ruby motif at YRNT (Y=pyrimidine, N=any nucleotide, R=purine) sequences, where the thymidine will be the first nucleotide of the 21U-RNA. YR resembles the initiator element required for RNA Pol II transcription initiation in mammals, plants, and flies (Gu et al., 2012). This produces a capped small RNA with two additional 5′ nucleotides, that is 26-29 nucleotides long (Cecere et al., 2012; Gu et al., 2012; Weick et al., 2014). In order to generate the mature 21U-RNA loaded on to PRG-1, the cap and 2nt must be removed from the 5′ termini, and 3′-to-5′ trimming must occur, however, the mechanism and factors involved in these steps are unknown. One factor that is required for the maturation of precursor to mature 21U-RNA has been identified, PRDE-1, as 26-29nt capped precursors accumulate in *prde-1* mutants. However, the function of PRDE-1 in the biogenesis of these 26-29nt precursors into 21U-RNAs is currently unknown (Weick et al., 2014). PRDE-1 is expressed in the nucleus suggesting that the pre-21U-RNA maturation occurs in the nucleus (Weick et al., 2014) (Figure 1.2). The presence of a uracil at position 3 of the precursor 21-RNA, leading to a mature 5′ uridine monophosphate is required for the processing of all 21U-RNAs, generating the 5′ U bias (Billi et al., 2013; Gu et al., 2012)

The majority of type-1 piRNAs are differentially enriched in either the male or female germline, with 56% of type-1 piRNAs preferentially expressed in the male germline and 15.8% in the female germline (Billi et al., 2013). This correlates with a CTGTTTCA Ruby motif for male enriched piRNAs and an

ATGTTTCA Ruby motif for female enriched piRNAs. More definitively, swapping of the motifs to a piRNA of the opposite sex disrupts the sex-specific expression of 21U-RNAs (Billi et al., 2013). This indicates there could be sex-specific Piwi/piRNA functions in *C. elegans.*

Type-2 piRNAs are associated with the promoters of protein coding genes, as their sequences are oriented divergently, just upstream of transcriptional start sites (TSSs), with the sense RNAs often corresponding to the major TSS (Gu et al., 2012). They were first discovered as capped small RNAs (csRNAs) that are ~26 nt (Gu et al., 2012), reminiscent of csRNAs transcribed by paused RNA Pol II or through early termination that have been described in a variety of organisms (Haussecker et al., 2008; Nechaev et al., 2010; Seila et al., 2008). These csRNAs are defined as piRNA precursors because they are processed into 21U-RNAs and loaded onto PRG-1. They exist as greater than 15,000 unique small RNA species, bringing the total number of unique piRNAs in *C. elegans* to greater than 30,000 (Gu et al., 2012). It is not known if type-2 piRNAs exert any regulatory function on the genes from which they are derived.

As mentioned above, very few 21U-RNAs have perfectly complementary target RNAs expressed in the *C. elegans* germline, making identifying potential targets complicated. However, it has been demonstrated that PRG-1/piRNAs can target RNAs with imperfect complementarity of up to four mismatched base pairs (Bagijn et al., 2012; Lee et al., 2012). Much like RDE-1/siRNAs and the exogenous RNAi pathway this initial targeting recruits RdRPs (EGO-1 and RRF-

1) to generate secondary (2°) 22G-RNAs with 5′ guanosine triphosphates that are antisense and of perfectly complementarity to the target (Bagijn et al., 2012; Lee et al., 2012) (Figure 1.2). These 2° 22G-RNAs are then thought to be loaded into WAGO Argonautes, which go on to silence their targets (Figure 1.2 and Table 1). This type of silencing is required to regulate transposable elements (Bagijn et al., 2012; Batista et al., 2008; Das et al., 2008), analogous to Drosophila piRNAs and mammalian pre-pachytene piRNAs. However, 21U-RNAs and PRG-1 also negatively regulate germline mRNAs through this mechanism, the function for this regulation is unknown (Bagijn et al., 2012; Lee et al., 2012).

The WAGO pathway was first identified as a genomic surveillance pathway in the *C. elegans* germline required to silence transposons, pseudogenes, cryptic loci (aberrant transcription), and greater than 1,000 coding genes (Gu et al., 2009). The machinery required to make these silencing 22G-RNAs and the association of these RNAs with the WAGO Argonaute WAGO-1 defined the targets of the pathway. WAGO-1 is localized to P granules, and the majority of these small RNAs are maternally inherited, as they are expressed in *C. elegans* oocytes (Gu et al., 2009). How the biogenesis of these small RNAs is triggered is a mystery; however, it was discovered that PRG-1 and piRNAs target some WAGO target mRNAs recruiting 2° RDRPs to produce a subset of WAGO-1 22G-RNAs (as described above) (Lee et al., 2012), as *prg-1* is required for the production of these small RNAs. Interestingly, the targets of this pathway are

upregulated in *prg-1* and 2° small RNA production/function mutants (Gu et al., 2009; Lee et al., 2012). Not all WAGO pathway 22G-RNAs are dependent on *prg-1* and their targets are not downregulated by piRNAs, indicating that there are both PRG-1-dependent and PRG-1-independent WAGO genomic surveillance pathways (Lee et al., 2012).

One established function of *C. elegans* piRNAs and PRG-1 is to initiate an epigenetic memory of 'nonself' RNAs (foreign sequences) not normally present in the germline. When transgenes containing nonself sequences such as green fluorescent protein (GFP) are introduced into the genome, even in single copy, they are often silenced in the *C. elegans* germline. Intriguingly, a germline silenced GFP transgene confers dominant epigenetic silencing in trans in the F1 generation when crossed to a strain that expresses a germline GFP transgene (Shirayama et al., 2012). Silencing persists in the F2 generation, even after crossing out the original silenced germline GFP transgene, and is preserved for greater than 30 generations. This phenomenon has been coined RNAe for RNA-induced epigenetic silencing (Shirayama et al., 2012), and has been demonstrated in multiple related transgene-silencing phenomena in the worm germline (Ashe et al., 2012; Buckley et al., 2012; Luteijn et al., 2012). This phenomenon is reminiscent of a form of epigenetic allelic interaction, also involving small RNAs, known as paramutation that has been described in organisms ranging from corn to mice (Erhard and Hollick, 2011).

Initiation of RNAe requires PRG-1 and piRNAs, as transgenes containing foreign sequences are readily expressed when introduced into *prg-1* mutant germlines (Shirayama et al., 2012). The machinery to produce 2° 22G-RNAs and at least 2 WAGO Argonautes (WAGO-1 and WAGO-9) are required to maintain silencing indefinitely, as crossing in one of these mutants into a germline silenced GFP transgenes, induces the expression of the transgene, however, piRNAs mutants do not, suggesting that they are only involved in initiation of RNAe (Bagijn et al., 2012; Luteijn et al., 2012; Shirayama et al., 2012). Downstream targeting of the WAGO pathway on silenced transgenes leads to H3K9me silencing marks to be deposited on the chromatin associated with the transgene, reinforcing the maintenance of silencing. Accordingly, chromatin modifying enzymes required for H3K9me and heterochromatin formation are also required for the maintenance of silencing (Bagijn et al., 2012; Luteijn et al., 2012; Shirayama et al., 2012).

The endogenous function of these pathways remains an open question in the *C. elegans* small RNA field. Why do PRG-1/piRNAs and the downstream WAGO pathways silence endogenous mRNAs, many of which could have germline function? Fascinatingly, WAGO Argonautes epigenetically transmit these silencing signals transgenerationally through both the sperm and oocyte (Conine et al., 2010; Gu et al., 2009). But if these targets are just silenced, why do they remain in the genome? It could be a matter of spatial or temporal regulation: piRNA targets might be expressed in distinct cells or times in the

germline, and the piRNA pathway might silence them where or when they are not required. It is a wonder why the pathway does not target more endogenous germline mRNAs. If PRG-1 and its greater than 30,000 unique piRNA cofactors target mRNAs with up to 4 mismatches, then tens of thousands of mRNAs would be targeted, including many essential to germline function. However, the *C. elegans* germline has a solution for this in another endogenous germline small RNA pathway: The CSR-1 pathway.

CSR-1: Master regulator of the *C. elegans* germline

CSR-1 (*C*hromosome *S*egregation, *R*NAi defective) is the only essential Argonaute in *C. elegans* (Yigit et al., 2006). *csr-1* mutant hermaphrodites are completely sterile, sometimes producing a few embryos that display fully penetrant embryonic lethality. This sterility is a result of chromosome segregation defects, as *csr-1* mutant embryos exhibit mitotic chromosomes that fail to properly condense and assemble on the metaphase plate, kinetochores that fail to orient with the chromosomes and mitotic spindle, and aberrant mitotic divisions resulting in anaphase bridging and aneuploidy (Claycomb et al., 2009). In the hermaphrodite germline, loss of CSR leads to defects in mitotic proliferation, meiotic progression, and chromosome condensation (She et al., 2009). Surprisingly, *csr-1* males are not fully sterile, but exhibit temperature sensitive sterility: at 20°C (permissive temperature) *csr-1* males produce half as many progeny as a WT male, but at 25°C *csr-1* males are completely sterile. This

sterility is a result of a spermiogenic defect, similar to *prg-1* mutant sperm, and *miwi* mutants in mice (Conine et al., 2013). CSR-1 localizes to P granules and is also required to maintain the structure of these germ granules (Claycomb et al., 2009; Updike and Strome, 2010).

CSR-1 interacts with 22G-RNAs complementary to ~6,000 genes expressed in both the hermaphrodite and male germline, accounting for the vast majority of mRNAs expressed in the germline (Claycomb et al., 2009; Conine et al., 2013). With a few exceptions, CSR-1 does not overlap with WAGO Argonaute targets, indicating that the CSR-1 and piRNA/WAGO endogenous small RNAs represent distinct pathways (Claycomb et al., 2009; Gu et al., 2009). The production of these 22G-RNAs requires the RdRP EGO-1, the Dicer-related helicase DRH-3, and the Tudor domain protein EKL-1 (Claycomb et al., 2009) (Table 1). The triggers or upstream factors that signal the production of these small RNAs are unknown, except in the case of the male germline (see below). Also, it remains a mystery how the CSR-1 and WAGO 22G-RNAs are correctly sorted, even though the CSR-1 and WAGO pathways share the machinery required for 22G-RNA biogenesis, and define very different endogenous small RNA pathways present in the worm germline.

Interestingly, unlike canonical small RNA pathways, CSR-1 does not silence (downregulate) its targets, as target mRNAs are not overexpressed in *csr-1* mutants. Conversely, it has been demonstrated in the hermaphrodite germline that CSR-1 and 22G-RNAs promote the transcription of their targets,

while also preventing antisense transcription and the ectopic transcription of silent chromatin domains (Cecere et al., 2014). CSR-1 also promotes the transcription of spermiogenic mRNAs (mRNAs required for haploid post-meiotic differentiation) in the male germline ((Conine et al., 2013), see below). It is unclear how CSR-1 promotes transcription; however, CSR-1 interacts with the chromatin of its targets, and also directly interacts with RNA Pol II in an RNA-dependent manner, indicating that it associates with nascent transcripts (Cecere et al., 2014; Conine et al., 2013; Wedeles et al., 2013b).

It is unclear why *csr-1* mutants exhibit chromosome segregation defects and if it is related to the misregulation of germline transcription. CSR-1 target genes are found in domains dispersed along the chromosomes, comprising 10-15% of the genome. CSR-1 genomic domains are associated with euchromatic histone marks, such as H3K4 methylation (mono-di-tri), H3K36 methylation, H3K9 acetylation, and histone H4K8 and K16 acetylation, as well as the histone variant H2AZ (HTZ-1), which is enriched at the majority of germline-expressed genes (Wedeles et al., 2013a). These domains are inversely correlated with the histone variant CENP-A (HCP-3), a histone variant that serves as an epigenetic mark for the formation of centromeres, conserved in eukaryotes (Claycomb et al., 2009; Gassmann et al., 2012). As *C. elegans* chromosomes are holocentric (centromeres are organized across the entire length of the chromosome), with each CENP-A domain a site of kinetochore attachment. Intriguingly, in *csr-1* mutants CENP-A localization on the chromosomes is disorganized and the

kinetochores fail to properly orient towards the spindles, indicating that the CSR-1 pathway participates in organizing CENP-A domains used as centromeres during mitosis (Claycomb et al., 2009). This could be by promoting the correct transcription of the germline as it has been demonstrated that patterns of transcription dictate where CENP-A is incorporated, CENP-A being excluded from regions where germline transcription occurs (Gassmann et al., 2012).

Another function of CSR-1 is to protect germline-expressed mRNAs from silencing by the piRNA surveillance pathway. Tethering of CSR-1 to the RNA of an actively expressed germline GFP transgene prevents piRNA-mediated RNAe (Wedeles et al., 2013b). This establishes that CSR-1 can protect RNAs from PRG-1 and piRNA initiated silencing. Interestingly, tethering of CSR-1 for several generations to a silenced germline GFP transgene or RNAe allele, activates its expression, indicating that CSR-1 targeting is not only protective but can also activate gene expression over several generations (Wedeles et al., 2013b).

In a related phenomenon, some actively expressed germline GFP transgenes dominantly activate the expression of RNAe silenced germline GFP transgenes in trans. This phenomenon requires CSR-1 and was named ‘RNAa’, for RNA-induced gene activation (Seth et al., 2013). Strains containing GFP transgenes that confer RNAa acquire 22G-RNAs complementary to GFP that accumulate in CSR-1, rather than WAGOs, identifying GFP as a ‘self’ RNA sequence, thereby protecting it from targeting by PRG-1/21U-RNAs and becoming silenced (Seth et al., 2013)(Figure 1.2). These findings suggest a

model where CSR-1 interacts with ‘self’ RNAs required for germline function, and PRG-1/21U-RNAs scan for ‘nonself’ or foreign sequences. When an RNA is targeted by PRG-1/21U-RNA and CSR-1 is not also present, it then recruits RdRP to generate silencing 22G-RNAs loaded into WAGOs that posttranscriptionally degrade the RNA (Figure 1.2). Nuclear WAGOs such as WAGO-9 can also use these silencing 22G-RNAs as guides to target the nascent transcripts of nonself RNAs to recruit chromatin modifiers that promote the formation of silenced heterochromatin, thereby preventing transcription (Seth et al., 2013; Shirayama et al., 2012; Wedeles et al., 2013b) (Figure 1.2). Fascinatingly, these Argonautes and small RNAs are present in the gametes of both sexes and are transmitted epigenetically from one generation to the next. Therefore, CSR-1 transmits a memory of self and WAGOs a memory of nonself, providing a highly adaptable, innate, sequence-specific genome-defense mechanism in the germline (Seth et al., 2013). However, this model is supported only by data from an unnatural transgene system, without answering the question of what these pathways’ natural functions are in coordinating the positive and negative regulation of nearly all mRNAs expressed in the *C. elegans* germline. The answer, at least in the male germline, comes from another endogenous small RNA pathway.

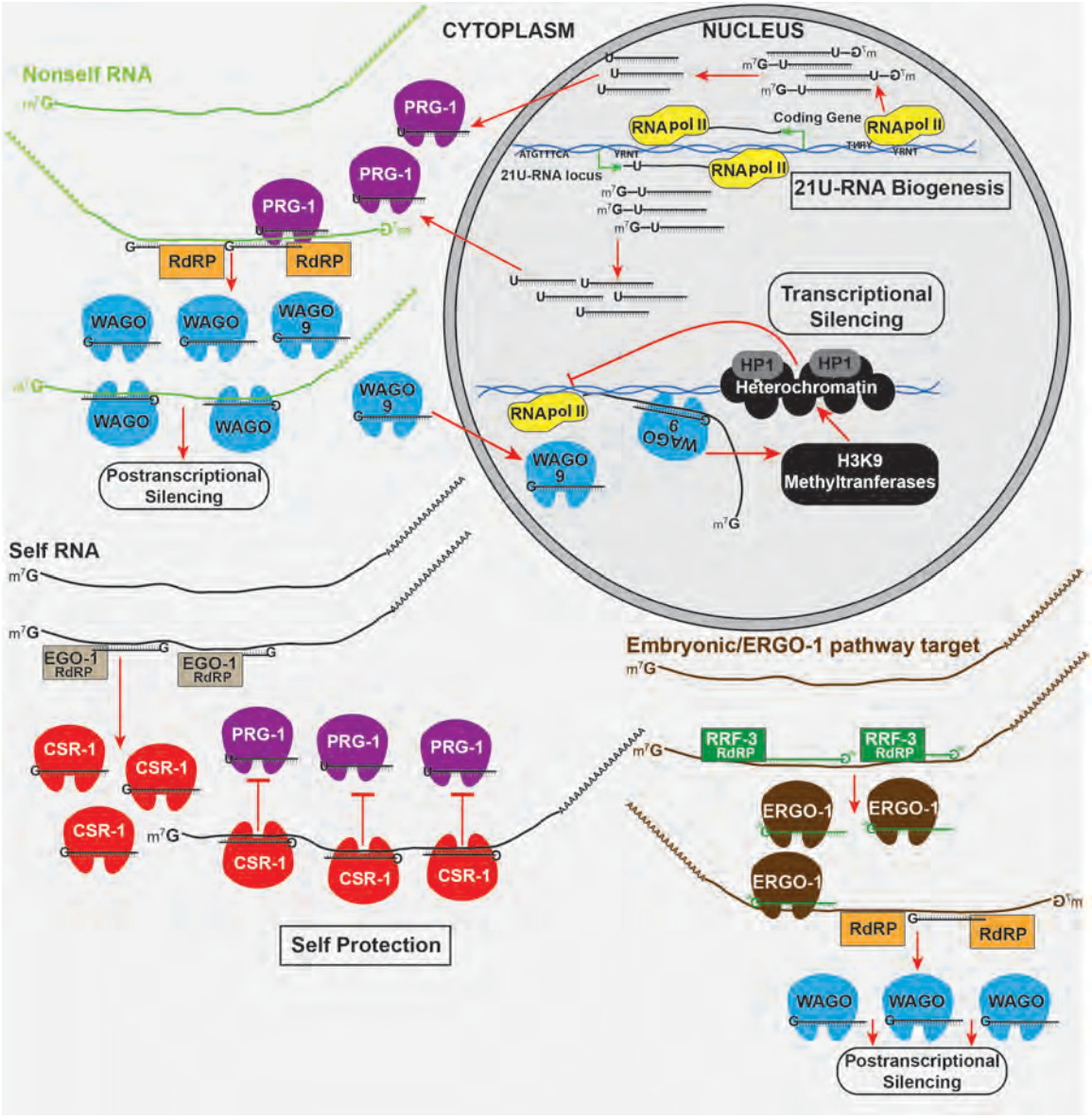

Figure 1.2. Endogenous small RNA pathways in the *C. elegans* hermaphrodite germline/oocyte.
Top, 21U-RNAs (piRNAs) are transcribed by RNA Pol II in the nucleus, Type I from 21U-RNA genes with upstream promoter elements, and Type II 21U-RNAs from bidirectional transcription at TSSs. In the cytoplasm 21U-RNAs bound to PRG-1 scan for nonself RNAs, recruiting RdRP to generate 2° 22G-RNAs that are loaded into WAGO Argonautes, which can posttranscriptionally or transcriptionally silence their targets. *Bottom Ieft*, CSR-1 and 22G-RNAs target self RNAs protecting from PRG-1/21U-RNA initiated silencing. *Bottom right*, the ERGO-1 embryonic 26G-RNA pathway. Not shown: CSR-1 promoting transcription of germline targets in the nucleus.

The ERI Endogenous Small RNA Pathway: Sperm and Egg

The final major small RNA pathway functioning in the *C. elegans* oocyte/embryo and male germline is the ERI endogenous small RNA pathway. Named ERI for *E*nhanced *R*NA*i*, as proteins required for the biogenesis and function of this pathway were first identified in genetic screens for mutants that exhibited enhanced sensitivity to exogenously triggered RNAi (Kennedy et al., 2004; Simmer et al., 2002). The ERI-pathway actually consists of two pathways, distinguished by developmental expression and the 1° Argonaute that functions in the pathway: the ERGO-1 (or embryonic) 26G-RNA pathway and the ALG-3/4 (or sperm) 26G-RNA pathway (Conine et al., 2010; Han et al., 2009; Pavelec et al., 2009; Vasale et al., 2010; Yigit et al., 2006). Many of the proteins required for the biogenesis and function of these endogenous pathways are also required for exogenous RNAi, most notably Dicer and the WAGOs (Conine et al., 2010; Duchaine et al., 2006; Gent et al., 2010; Pavelec et al., 2009; Vasale et al., 2010; Welker et al., 2010). The enhancement of RNAi in *eri* mutants is thought to be a result of competition for use of the shared components, in the mutants the endogenous pathways are lost, freeing up the machinery for exogenous RNAi (Duchaine et al., 2006; Yigit et al., 2006). Interestingly, a majority of *eri* mutants also exhibit a temperature-sensitive (TS), male-sterile phenotype that results from a nonfunctional sperm 26G-RNA pathway (see below).

Both pathways are defined by 1° endogenous small RNAs, termed 26G-RNAs, because they have a bias to begin with a 5′ guanosine monophosphate, and are predominantly 26 nt long (Conine et al., 2010; Han et al., 2009; Ruby et al., 2006; Vasale et al., 2010) (Table 1). The embryonic 26G-RNAs begin expression in the oocyte, and persists throughout the development of the embryo (Han et al., 2009; Vasale et al., 2010). Expression continues in the germline, as *glp-4(bn2)* mutants which lack a germline, exhibit significant depletion of embryonic 26G-RNAs (Han et al., 2009). These small RNAs interact with the Piwi-clade 1° Argonaute ERGO-1 (Vasale et al., 2010). ERGO-1 is also required for the biogenesis of embryonic 26G-RNAs (Han et al., 2009; Vasale et al., 2010) (Table 1). Interestingly, *ergo-1* mutants exhibit only an *eri* phenotype, and not TS male sterility like other *eri* mutants, suggesting that ERGO-1 functions exclusively in the embryonic pathway (Han et al., 2009; Pavelec et al., 2009; Vasale et al., 2010). The embryonic/ERGO-1 26G-RNA pathway also functions in the soma, as many of the phenotypes induced by enhanced RNAi in *eri* mutants, are somatic defects, particularly in the nervous system, and a subset of these small RNAs persists in mutants that lack germline tissue (Gent et al., 2010; Kennedy et al., 2004; Simmer et al., 2002). Enhanced RNAi is the only obvious phenotype displayed by *ergo-1* mutants, and therefore, assigning a biological function to the pathway has been a challenge (Han et al., 2009; Vasale et al., 2010). This is similar to endo-siRNAs in the fly soma and mouse oocyte, which are dispensable

for viability and reproduction (Ghildiyal et al., 2008; Kawamura et al., 2008; Okamura et al., 2008; Tam et al., 2008; Watanabe et al., 2008).

In the oocyte/embryo the RdRP RRF-3 generates antisense 26G-RNAs targeting ~50 genes and ~100 unannotated genomic regions that produce likely noncoding RNAs (Gent et al., 2010; Han et al., 2009; Vasale et al., 2010) (Figure 1.2). The 26G-RNAs produced in the embryo are loaded onto ERGO-1 and modified by 2′O-methylation at the 3′ most nucleotide via the methyltransferase HENN-1, providing an interesting correlation with piRNAs, as ERGO-1 is a Piwi-clade Argonaute (Billi et al., 2012; Kamminga et al., 2012; Montgomery et al., 2012) (Table 1). Interestingly, the majority of the genes targeted by the ERGO-1 26G-RNA pathway have no known function, however, the genomic loci producing these target RNAs exhibit a nonrandom distribution in the genome and appear to include many genes resulting from genomic duplications, suggesting that this pathway may control overexpression from gene expansion (Vasale et al., 2010). ERGO-1 using 26G-RNAs as a guide then engage the target RNAs that were initially used as a template to produce the 26G-RNA. This targeting then recruits a secondary 2° RdRP (RRF-1 or EGO-1) to generate 2° 22G-RNAs that are then loaded into WAGO Argonautes (Figure 1.2). These 22G-RNAs have 5′ guanosine triphosphates analogous to other classes of 22G-RNAs (Gent et al., 2010; Vasale et al., 2010) (Table 1). It should be noted that ERGO-1 contains an active catalytic triad required for endonucleolytic cleavage in its Piwi domain, allowing it to slice target RNAs (Yigit et al., 2006). Whether ERGO-1 possesses

Slicer activity and if it is required for 2° small RNA production is unknown. Targeting by the ERGO-1 26G-RNA pathway leads to potent silencing of its targets, but the function of silencing is unknown (Gent et al., 2010; Han et al., 2009; Vasale et al., 2010) (Figure 1.2).

In the meiotic male germline during spermatogenesis RRF-3 generates 26G-RNAs targeting ~1400 mRNAs, including the majority of mRNAs expressed specifically in the male germline (sperm genes) (Conine et al., 2010; Conine et al., 2013; Gent et al., 2009; Han et al., 2009; Pavelec et al., 2009) (Figure 1.3 and Table 1). Interestingly, these sperm-specific 26G-RNAs do not have any 3′ modification, providing a physical distinction between them and the ERGO-1 embryonic 26G-RNAs (C.C. Unpublished Data, Billi et al., 2012; Kamminga et al., 2012; Montgomery et al., 2012). These 26G-RNAs are thought to be loaded into the redundant AGO-clade 1° Argonautes ALG-3 and ALG-3/4 (referred to collectively as ALG-3/4), as *alg-3/4* are required for the biogenesis of sperm-specific 26G-RNAs (Conine et al., 2010; Han et al., 2009) (Figure 1.3). Accordingly, *alg-3/4* mutants exhibit a TS male sterility phenotype, and do not exhibit an Eri phenotype, as they are not required for the production of ERGO-1/embryonic 26G-RNAs (Conine et al., 2010; Han et al., 2009).

Analogous to the ERGO-1/embryonic 26G-RNA pathway, ALG-3/4 using 26G-RNAs as guides target sperm genes from which the 26G-RNAs were derived, to recruit 2° RdRPs (RRF-1 and EGO-1) to generate 2° 22G-RNAs, again with a 5′ guanosine triphosphate (Conine et al., 2010) (Figure 1.3 and

Table 1). Like ERGO-1, ALG-3/4 both contain an active catalytic triad required for endonucleolytic cleavage in its Piwi domain, allowing it to potential slice target RNAs, as in the ERGO-1 pathway, whether slicing occurs and if it's required for 2° small RNA production is unknown (Yigit et al., 2006). A subset of ALG-3/4 26G-RNA target mRNAs are posttranscriptionally silenced downstream of 1° targeting by 22G-RNAs and WAGOs. While targeting by ALG-3/4 26G-RNAs to another subset of mRNAs generates 22G-RNAs that are loaded into CSR-1, which in turn promotes the transcription of the targeted mRNA during spermatogenesis, creating a positive feedback to drive expression of the mRNA and complementary small RNAs (Conine et al., 2013) (Figure 1.3 and Table 1).

The proceeding chapters will describe the expression of the Argonautes ALG-3/4 and CSR-1 and their small RNA cofactors in the male germline, the biogenesis of the small RNAs, and their regulatory affects on their target mRNAs. These findings support a model for the ALG-3/4 26G-RNA pathway functioning to promote thermotolerant male fertility and to provide a paternally transmitted memory of germline gene expression. These findings underlie the capability of sperm to transmit an extraordinary amount of epigenetic information via small RNAs representing binary silencing and activating signals to successive generations. This dissertation provides the first comprehensive overview of small RNA pathways in the *C. elegans* male germline presenting a framework for their future studies.

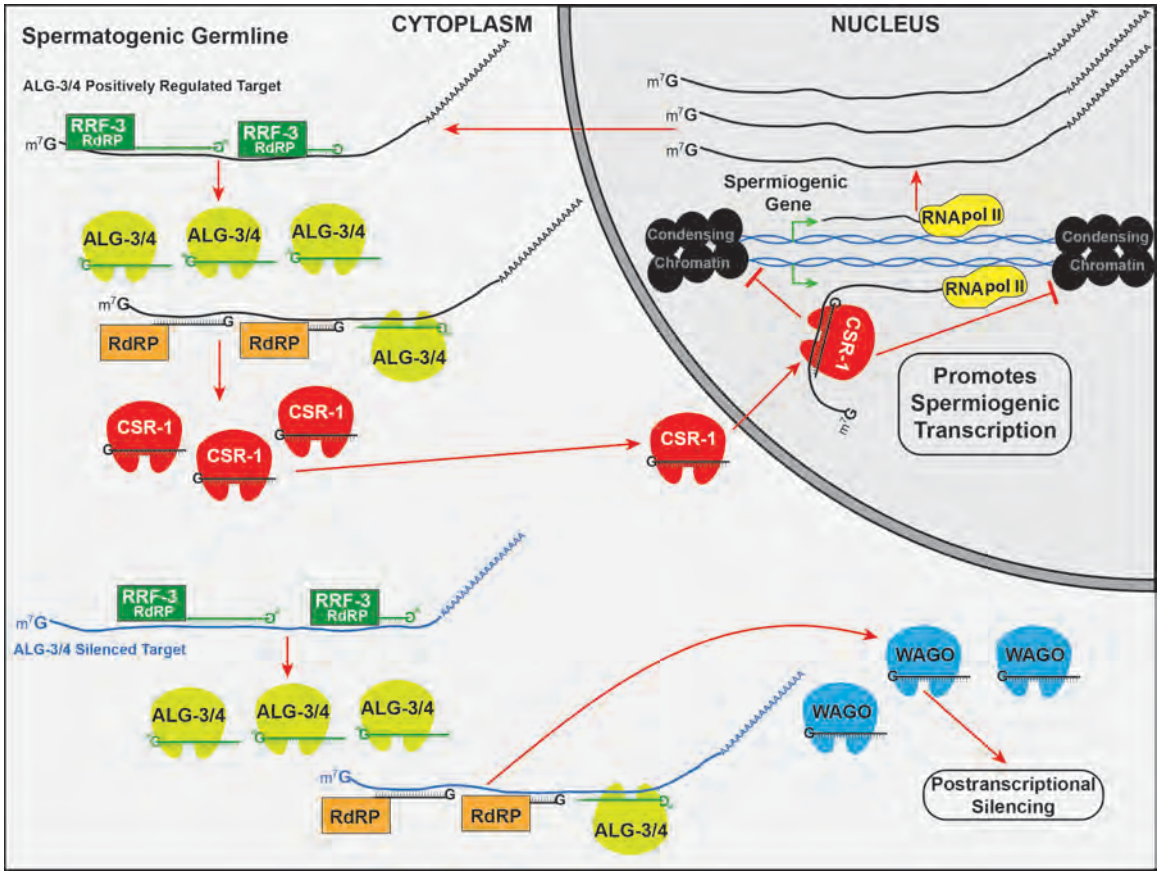

Figure 1.3. The ALG-3/4 sperm 26G-RNA pathway.
The RdRP RRF-3 uses mRNAs expressed during spermatogenesis as templates to create 26G-RNA which are thought to be loaded into the 1° Argonautes ALG-3/4. ALG-3/4 and 26G-RNAs then target the original template mRNA to recruit a 2° RdRP to generate 2° 22G-RNAs loaded into CSR-1. CSR-1 in the nucleus promotes the transcription of these targets. Another subset of targets also produces 22G-RNAs downstream of ALG-3/4 and 26G-RNAs that are loaded into WAGOs leading to the posttranscriptional silencing of the targeted mRNA.

Small RNA Pathway	1° Small RNA	Length (nt)	5' Nucleotide bias	3' modification	1° Argonaute	2° Small RNA	RdRP	2° Argonaute	Germline Expression	Function
Exogenous RNAi	siRNA	21-22	none	-OH	RDE-1	22G-RNA	RRF-1 & EGO-1	WAGOs	Male and Female	postranscriptional & transcripitonal silencing
miRNA pathway	miRNA	21-23	U	-OH	ALG-1/2	NA	NA	NA	Male and Female	postranscriptional silencing & translational repression
piRNA pathway	21U-RNA	21	U	2'-OCH_3	PRG-1	22G-RNA	EGO-1	WAGOs	Male and Female	postranscriptional & transcriptonal silencing
CSR-1 pathway (female)	- -	- -	- -	- -	- -	22G-RNA	EGO-1	CSR-1	Female	promoting transcription
ALG-3/4 Sperm pathway	26G-RNA	26	G	-OH	ALG-3/4	22G-RNA	EGO-1	CSR-1 & WAGOs	Male	promoting transcription & postranscriptional silencing
ERGO-1 Embryonic pathway	26G-RNA	26	G	2'-OCH_3	ERGO-1	22G-RNA	RRF-1 & EGO-1	WAGOs	Female	postranscriptional silencing

Table 1. Small RNA pathways in the *C. elegans* germline.
The small RNA pathways are characterized by the physical characteristics of the small RNAs, the Argonautes they interact with, expression, function, and whether they elicit a secondary small RNA response.

CHAPTER II

THE ARGONAUTES ALG-3 AND ALG-4 ARE REQUIRED FOR SPERMATOGENESIS-SPECIFIC 26G-RNAS AND THERMOTOLERANT SPERM IN C. ELEGANS

Summary

Gametogenesis is a thermosensitive process in numerous metazoans, ranging from worms to man. In *C. elegans,* a variety of RNA-binding proteins that associate with germ-line nuage (P granules), including the Piwi-clade Argonaute PRG-1, have been implicated in maintaining fertility at elevated temperature. Here we describe the role of two AGO-class paralogs, *alg-3* (T22B3.2) and *alg-4* (ZK757.3), in promoting thermotolerant male fertility. A rescuing GFP::*alg-3* transgene is localized to P granules beginning at the late pachytene stage of male gametogenesis. *alg-3/4* double mutants lack a subgroup of small RNAs, the 26G-RNAs which target and appear to downregulate numerous spermatogenesis-expressed mRNAs. These findings add to a growing number of AGO pathways required for thermotolerant fertility in *C. elegans* and support a model in which AGOs and their small RNA cofactors function to promote robustness in gene-expression networks.

INTRODUCTION

Argonaute (AGO) proteins have been implicated in gene silencing in fungi, plants, protozoans, and metazoans including humans. They are ~100-kD highly basic proteins that are characterized by the presence of PAZ and PIWI domains and by their association with small RNA species (reviewed in (Hutvagner and Simard, 2008)). Argonautes can be classified into three clades (Tolia and Joshua-Tor, 2007): (i) the AGO clade, which includes the human AGOs1-4, the *C. elegans* miRNA effectors ALG-1/2, and all of the *Arabidopsis thaliana* AGOs; (ii) the Piwi clade, which more closely resembles *Drosophila* Piwi and includes the *C. elegans* piRNA AGO PRG-1; and (iii) an expanded family of worm-specific AGOs (WAGOs) that lack residues in the PIWI domain thought to be necessary for slicer endonuclease activity.

In *C. elegans*, null mutations are available for the entire family of over 24 AGO genes (Yigit et al., 2006), and at least 5 different combinations of these mutants result in lethal or sterile phenotypes. *alg-1/2* mutants display heterochronic defects and lethality that arise from loss of miRNA-mediated regulation of developmentally-important transcripts (Grishok et al., 2001). *csr-1* mutants have severe chromosome segregations defects that result in embryonic lethality (Claycomb et al., 2009; Yigit et al., 2006), and *prg-1* mutants have severe defects in the development of the germline (Batista et al., 2008; Das et al., 2008; Wang and Reinke, 2008). In addition, the simultaneous deletion of twelve WAGOs leads to a temperature-sensitive (ts) sterile phenotype, where

mutants are viable and fertile at 20°C but sterile at 25°C (Gu et al., 2009).

Here we describe two paralogous AGO-clade family members, *alg-3* and *alg-4* that are together critical for sperm development and function. While single mutants have only a minor reduction in brood size, double *alg-3/4* mutants exhibit drastically reduced brood sizes at elevated temperatures. We show that a rescuing GFP::ALG-3 protein is expressed in germ-cells undergoing spermatogenesis, and that, while both male and hermaphrodite *alg-3*/*4* mutants produce near wild-type numbers of spermatids, these spermatids exhibit severe defects in the activation process called spermiogenesis that converts spermatids into motile ameboid sperm.

A class of 26nt small RNAs called the 26G-RNAs were first identified in deep sequencing data sets in *C.elegans* as part of the previously described Dicer-ribonuclease-dependent endogenous small RNA pathway called the ERI pathway, and a subset of these were noted to be enriched for spermatogenesis-expressed mRNA targets (Duchaine et al., 2006; Gent et al., 2009; Pavelec et al., 2009; Ruby et al., 2006). Here we show that, like their targets, a subset of 26G-RNAs are specifically expressed in the male germline and that their accumulation depends on the ALG-3/4 AGOs. Our findings suggest that ALG-3/4 function directly or indirectly in concert with 26G-RNAs and with other components of the ERI pathway to negatively regulate the steady-state levels of their target transcripts. Together with findings described by Vasale et al. (Vasale et al., 2010) our study supports a two-step AGO model in which an initial round of

AGO/26G-RNA-mediated targeting triggers the production of secondary small RNAs (called 22G-RNAs) that engage a distinct argonaute(s) to amplify the silencing signal.

Our findings add to a growing number of Argonaute-mediated pathways that promote robust-thermotolerant fertility in *C. elegans*. Argonautes acquire specificity through their RNA cofactors, and thus in principle have virtually unlimited capacity for sequence-specific gene regulation. We propose that AGO systems may utilize their versatile and highly adaptable nature to promote robustness in gene expression networks.

Results

Mutations in the *alg-3/4* AGOs result in male-associated temperature-sensitive sterility

The AGO-clade argonautes *alg-3* and *alg*-4 exhibit 96% sequence identity at the amino acid level (Figure 2.1), as well as nucleotide homology that extends into the 5´ and 3´ noncoding regions. Consistent with the idea that *alg-3* and *alg-4* are recently duplicated genes, our genetic tests suggested that they retain partially overlapping functions. When compared to wild type, the *alg-3(tm1155)* single mutant exhibited a 2-fold decrease in brood size at 20°C, and a 2.5- fold decrease at 25°C (Figure 2.2A). Similarly, *alg-3(ok1041)* and *alg-4 (tm1184)* displayed a 2-fold decrease in brood size at 20°C, and 3-fold decrease at 25°C

(Figure 2.2A). However, when *alg-3(tm1155)* was combined with *alg-4(ok1041)* , the resulting double mutants exhibited a 3-fold decrease in brood size at 20°C, and complete sterility at 25°C (Figure 2.2A).

To determine whether the sterility of *alg-3; alg-4* (*alg-3/4)* double mutant hermaphrodites reflects a general deficit in gametogenesis, or alternatively, is specific to either spermatogenesis or oogenesis, we created obligate male/female populations. To do this we utilized a *fog-2* mutant that results in the feminization of hermaphrodites (Schedl and Kimble, 1988). In the following crosses, all the individuals are *fog-2* mutants, and for simplicity we refer to animals with wild-type *alg-3* and *alg-4* activities as "*fog-2* males" or "*fog-2* females" and to *alg-3*; *alg-4* double mutants as "*alg-3*/*4* males" or "*alg-3/4* females". When mated to either *fog-2* females or *alg-3/4* females, *alg-3*/*4* males sired half as many progeny as *fog-2* males at 20°C, and were almost completely sterile at 25°C (Figure 2.2B). In contrast, *alg-3*/*4* females were only slightly less fertile than *fog-2* females when mated to *fog-2* males. These data indicate that the sterility of *alg-3*/*4* mutants arises primarily from defects specific to the male germline.

A

```
ALG-3   MSRRNATNFVDNNTLTSSGISGSGSLSPPITSRPASGQASPLSSNGSLSPP--VDDQGSV 58
ALG-4   MSRRNATSFVDNNTLTSSGISGSGSMSPPITSRPASGQASPLTSNGSLSPPQYADDQGSV 60

ALG-3   SYNSDSPRDLSPLLLSELACLNMREVVARPGLGTIGRKIPVKSNFFAVDLKNPKMVVVQY 118
ALG-4   SYNLDSPRDLSPLLLSELACLNMREVVARPGLGTIGRQIPVKSNFFAMDLKNPKMVVIQY 120

ALG-3   HVEVHHPGCRKLDKDEMRIIFWKAVSDHPNIFHNKFALAYDGAHQLYTVARLEFPDDQGS 178
ALG-4   HVEIHHPGCRKLDKDEMRIIFWKAVSDHPNIFHNKFALAYDGAHQLYTVARLEFPDDQGS 180

ALG-3   VRLDCEATLPKDNRDRTRCAISIQNVGPVLLEMQRTRTNNLDERVLTPIQILDIICRQSL 238
ALG-4   VRLDCEASLPKDNRDRTRCAISIQNVGPVLLEMQRTRTNNLDERVLTPIQILDIICRQSL 240

ALG-3   TCPLLKNSANFYTWKSSCYRIPTAAGQALDLEGGKEMWTGFFSSAHIASNYRPLLNIDVA 298
ALG-4   TCPLLKNSANFYTWKSSCYRIPTAAGQALDLEGGKEMWTGFFSSAHIASNYRPLLNIDVA 300

ALG-3   HTAFYKTRITVLQFMCDVLNERTSKPNRNNPRGPGAPGG---YRGGRGGARGGSYQNFGN 355
ALG-4   HTAFYKTRITVLQFMCDVLNERTSKPNRNNPRGPGGPGGPGGYRGGRGGGRGGSYGNFGN 360

ALG-3   RGPPGANVRDDFGGNGLTFTMDTLSRDTQLSSFETRIFGDSIRGMKIRATHRPNAIRVYK 415
ALG-4   RGPPGANVRDDFGGNGLTFTMDTLSRDTQLSSFETRIFGDAIRGMKIRAAHRPNAIRVYK 420

ALG-3   VNSLQLPADKLMFQGIDEEGRQVVCSVADYFSEKYGPLKYPKLPCLHVGPPTRNIFLPME 475
ALG-4   VNSLQLPADKLMFQGIDEEGRQVVCSVADYFSEKYGPLKYPKLPCLHVGPPTRNIFLPME 480

ALG-3   HCLIDSPQKYNKKMTEKQTSAIIKAAAVDATQREDRIKQLAAQASFGTDPFLKEFGVAVS 535
ALG-4   HCLIDSPQKYNKKMSEKQTSAIIKAAAVDATQREDRIKQLAAQASFGTDPFLKEFGVAVS 540

ALG-3   SQMIETSARVIQPPPIMFGGNNRSINPVVFPKDGSWSMDHQTLYMPATCRSYSMIALVDP 595
ALG-4   SQMIQTTARVIQPPPIMFGGNNRSVNPVVFPKDGSWTMDNQTLYMPATCRSYSMIALVDP 600

ALG-3   RDQTSLQTFCQSLTMKATAMGMNFPRWPDLVKYGRSKEDVCTLFTEIADEYRVTNTVCDC 655
ALG-4   RDQTSLQTFCQSLTMKATAMGMNFPRWPDLVKYGRSKEDVCTLFTEIADEYRVTNTVCDC 660

ALG-3   IIVVLQSKNSDIYMTVKEQSDIVHGIMSQCVLMKNVSRPTPATCANIILKLNMKMGGINS 715
ALG-4   IIVVLQSKNSDIYMTVKEQSDIVHGIMSQCVLMKNVSRPTPATCANIVLKLNMKMGGINS 720

ALG-3   RIVADQITNKYLVDQPTMVVGIDVTHPTQAEMRMNMPSVAAIVANVDLLPQSYGANVKVQ 775
ALG-4   RIVADKITNKYLVDQPTMVVGIDVTHPTQAEMRMNMPSVAAIVANVDLLPQSYGANVKVQ 780

ALG-3   KKCRESVVYLLDAIRERIITFYRHTKQKPARIIVYRDGVSEGQFSEVLREEIQSIRTACL 835
ALG-4   KKCRESVVYLLDAIRERIITFYRHTKQKPARIIVYRDGVSEGQFSEVLREEIQSIRTACL 840

ALG-3   AIAEDFRPPITYIVVQKRHHARIFCKFPNDMVGKAKNVPPGTTVDTGIVSPEGFDFYLCS 895
ALG-4   AIAEDFRPPITYIVVQKRHHARIFCKYQNDMVGKAKNVPPGTTVDTGIVSPEGFDFYLCS 900

ALG-3   HYGVQGTSRPARYHVLLDECKFTADEIQNITYGMCHTYGRCTRSVSIPTPVYYADLVATR 955
ALG-4   HYGVQGTSRPARYHVLLDECKFTADEIQSITYGMCHTYGRCTRSVSIPTPVYYADLVATR 960

ALG-3   ARCHIKRKLGLADNNDCDTNSLSSSLASLLNVRTGSGKGKKSHAPSVDDESYSLPDAASD 1015
ALG-4   ARCHVKRKLGLADNNDCDTNSRSSTLASLLNVRTGSGKGKKSYAPSVDDESYSLSDATSD 1020

ALG-3   QILQDCVSVAADFKSRMYFI                                          1035
ALG-4   QILQDCVSVATDFKSRMYFI                                          1040
```

B

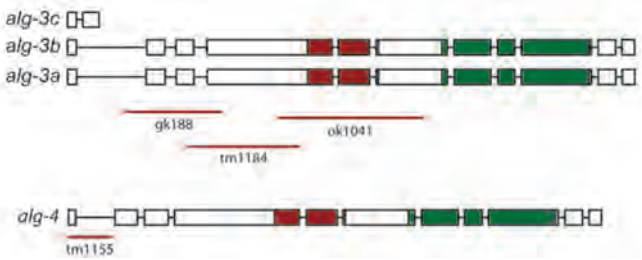

Figure 2.1. Amino acid alignment and gene models for *alg-3 and alg-4.*
(A) Alignment of amino acid sequences for *alg-3* and *alg-4* using clustalW2. **(B)** Gene models for *alg-3* and *alg-4*. Three predicted isoforms of *alg-3* are shown.

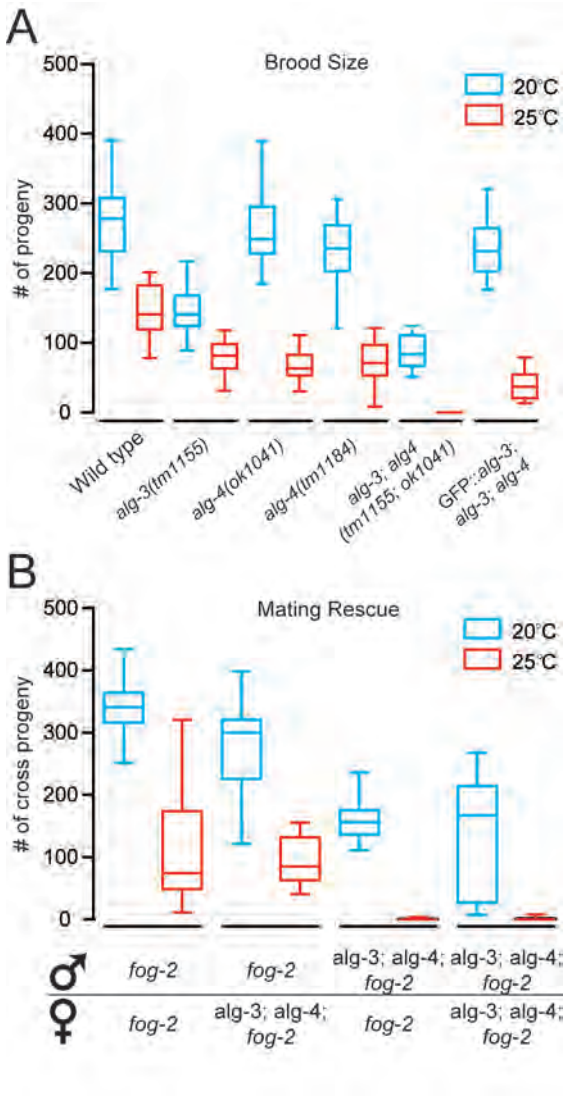

Figure 2.2. *alg-3/4* mutants exhibit temperature-sensitive sterility associated with the male germline.
(A-B) Box-and-whiskers plots of brood size in wild-type and mutant strains as indicated for (n>20) animals cultured at 20°C (blue) and 25°C (red). In these and all subsequent box-and-whisker plots, the top and bottom ends of each box represent the 75th and 25th percentile; respectively, the line in the middle represents the median value, and the extended lines illustrate the range (highest and lowest value). The temperature of cultivation is indicated by color. **(A)** Self-crosses; **(B)** crosses between *fog-2* (wild-type) and *alg-3/4*;*fog-2* (mutant) worms.

To determine the temperature-sensitive period (tsp) for *alg-3*/*4*-associated sterility, hermaphrodites reared at either 20°C or 25°C were shifted to the converse temperature (either up to 25°C or down to 20°C) as L2, L3, or L4 larva, or as gravid adults (Ward and Miwa, 1978). When shifted up from 20°C to 25°C at or prior to the L4 stage, *alg-3*/*4* hermaphrodites were completely sterile (Figure 2.3) Conversely, fertility was partially restored when *alg-3*/*4* animals were shifted down from 25°C to 20°C anytime prior to the L4 stage (Figure 2.3). Fertility could not, however, be restored by shifting adults reared at 25°C down to 20°C, demonstrating that the temperature-sensitive sterility is at that point irreversible. Taken together, these data place the male-fertility defect at the L4 stage, coincident with the timing of spermatogenesis.

	Shift	L2/L3	L3/L4	L4/YA	GA
embryo hatched at 25°C	to 20°C	++	++	+	-
embryo hatched at 20°C	to 25°C	-	-	+	+

++ average progeny similar to worms maintained at 20˚C;
+ average progeny <50% of worms maintained at 20˚C;
- no progeny

Figure 2.3. The temperature sensitive period for *alg-3/4* double mutants. Worms were hatched at either 20°C or 25° and then shifted to the opposite temperature at the indicated larval stage and the brood size was determined.

ALG-3 is expressed in the region of the germline undergoing spermatogenesis

To examine the localization of ALG-3, we generated a full-length, N-terminally tagged GFP::*alg-3* transgenic line under the control of the *alg-3* promoter and 3´UTR. This transgenic line rescues *alg-3*; *alg-4* brood size to a level near the single *alg-4 (ok1041)* mutant (Figure 2.2A). In GFP::*alg-3* hermaphrodites and males, GFP expression was observed in the proximal germline beginning at the L4 stage (Figure 2.4A & Figure 2.5A), coincident with the onset of spermatogenesis. In adult hermaphrodites, after the switch to oogenesis, GFP::ALG-3 expression was restricted to the spermatheca. Within the spermatheca, in hermaphrodites, and testis in males, GFP::ALG-3 was localized within residual bodies, which are enucleate cytoplasts produced by the budding off of mature spermatids at the end of meiosis II. Little if any GFP::ALG-3 could be detected in mature spermatids (Figure 2.4C & Figure 2.5A-C). As expected, the proximal expression of GFP::ALG-3 continued through adulthood in males (Figure 2.5B). Within the developing spermatocytes GFP::ALG-3 was first apparent in the post-pachytene germline prior to the point where chromosomes appear to aggregate into a single mass (Figure 2.4A) (Shakes et al., 2009). GFP signal was localized throughout the cytoplasm and enriched at perinuclear foci coincident with P granules based on colocalization with RFP::PGL-1 (Figure 2.4B) (Kawasaki et al., 1998). The developmental expression profile of *alg-3* mRNA mirrored that of GFP::ALG-3 protein expression

and was enriched in male-enriched populations and depleted in female populations (Figure 2.5D).

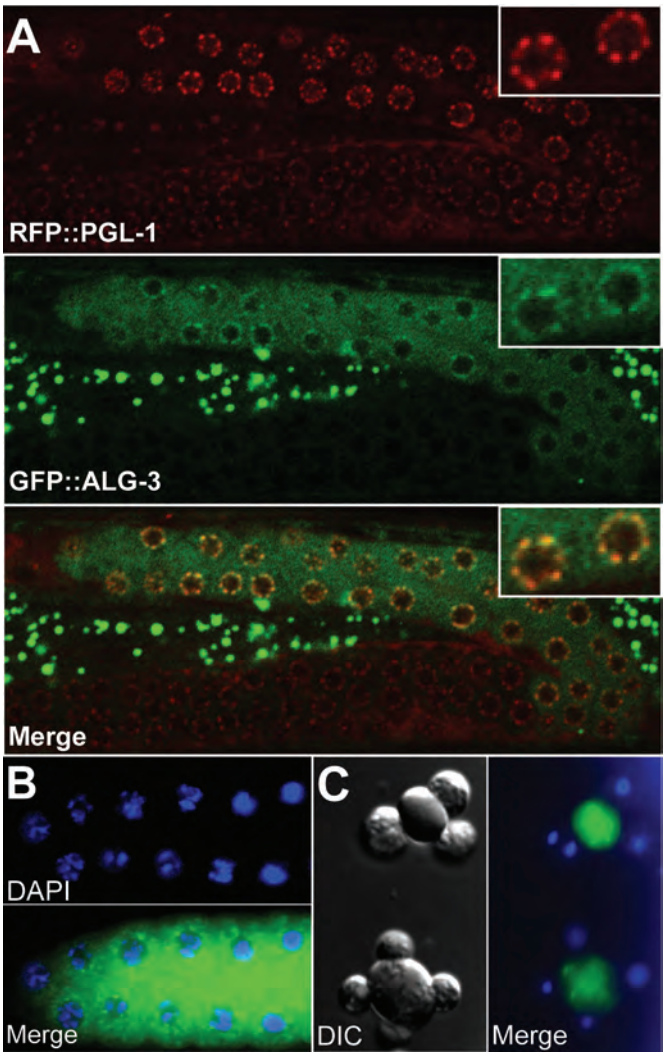

Fig. 2.4. GFP::ALG-3 is expressed during spermatogenesis.
(A-C) Micrographs of ALG-3::GFP (green), nuclear staining/DAPI (blue) and PGL-1::RFP (red). **(A)** Confocal images in a young adult hermaphrodite. **(B)** Onset of expression in the proximal region of a male gonad. **(C)** Nomarski and fluorescence images of spermatids attached to residual bodies.

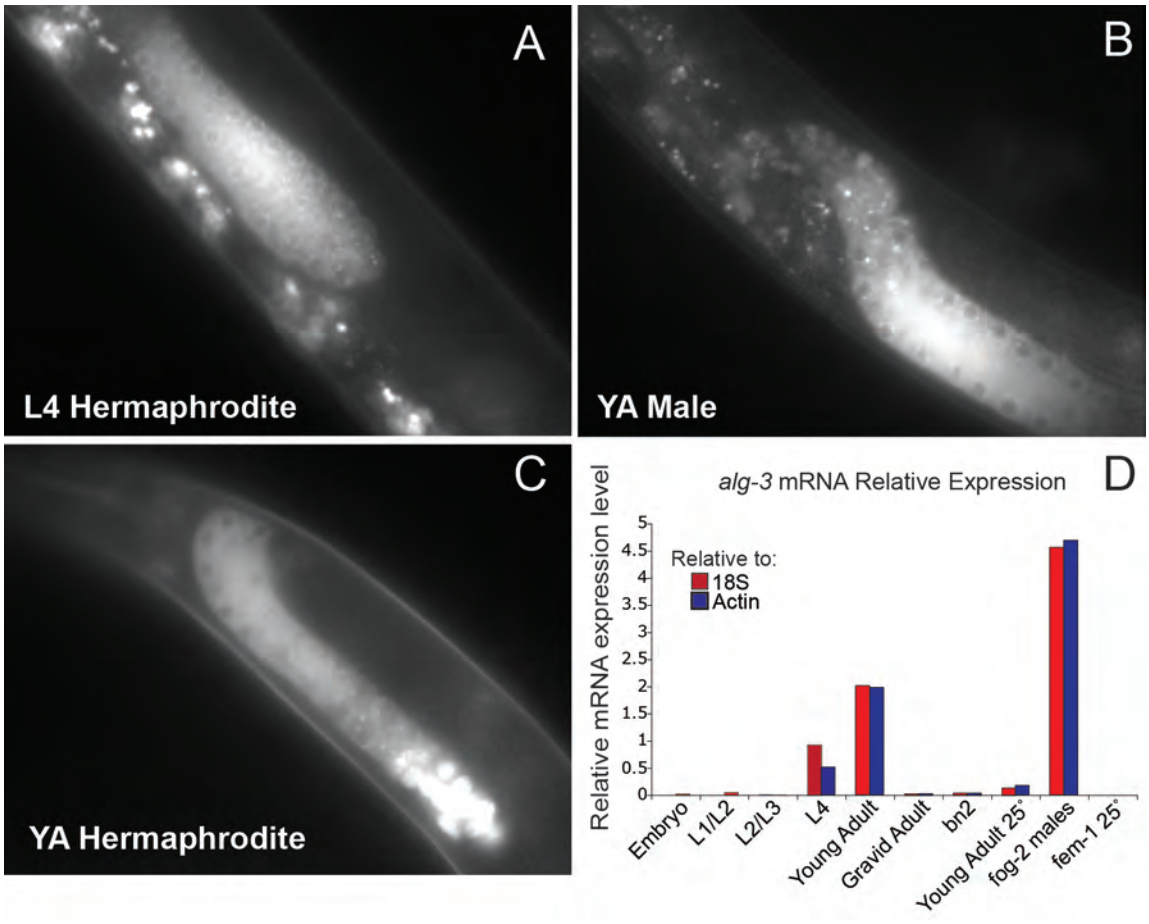

Figure 2.5. ALG-3 is expressed during spermatogenesis.
(A-C) Fluorescence microscopy of GFP::ALG-3 in a Wild-type L4 hermaphrodite and in Young Adult (YA Male and Hermaphrodite) animals as indicated. **(D)** Quantitative real time PCR analysis of *alg-3* mRNA expression using *actin* mRNA (Blue bars) or 18s rRNA (Red bars) as a normalization standard.

alg-3/4 double mutants exhibit defects in spermatogenesis and spermiogenesis

To investigate the underlying cause of the *alg-3/4* male-germline-associated sterile phenotype, we examined the development and numbers of *alg-3/4* spermatids in more detail. Relative to *fog-2* males, *alg-3/4* males produced approximately wild-type numbers of spermatids at 20°C, and at 25°C produced 29% fewer than wild type (Figure 2.6A). This reduction in mature spermatids at the nonpermissive temperature was correlated with the persistence of abnormal secondary (2°) spermatocytes into adulthood (Figure 2.7). These 2° spermatocytes appeared to be arrested as large dinucleate masses or as spermatids that fail to bud from the residual body. Chromosome bridging was observed in 20% of abnormal secondary spermatocytes in both males and hermaphrodites, indicating possible segregation defects during meiosis (Figure 2.7). Hence, a reduction in the number of spermatids and defects in meiotic cell division could contribute to the fully sterile phenotype of *alg-3/4* double mutants at the non-permissive temperature. However, many apparently normal-spermatids are produced in *alg-3/4* mutants at both the permissive and non-permissive temperatures. We therefore wondered whether a defect in spermiogenesis also existed. During spermiogenesis, activated spermatids form spikes that protrude from the cell and subsequently rearrange into a single pseudopod (Shakes and Ward, 1989). *In vivo*, male spermatids are activated upon ejaculation, while hermaphrodite spermatids are activated when an oocyte

enters the spermetheca (L'Hernault, 2006). *In vitro,* spermiogenesis can be induced by treating isolated spermatids with pronase (Shakes and Ward, 1989). Relative to *fog-2*, 20% fewer *alg-3*/*4* spermatids isolated from males were activated by pronase, as indicated by the formation of spike-like structures at permissive temperature (Figure 2.6B). Strikingly, only 10% of activated *alg-3/4* spermatids progressed to form pseudopods, compared to 57% of activated wild type spermatids. Instead, the *alg-3*/*4* spermatids accumulated many abnormal, long spike-like structures and failed to become motile (Figure 2.6C). At non-permissive temperature this defect increased, as less than 2% of spermatids examined formed pseudopods compared to 54% for WT.

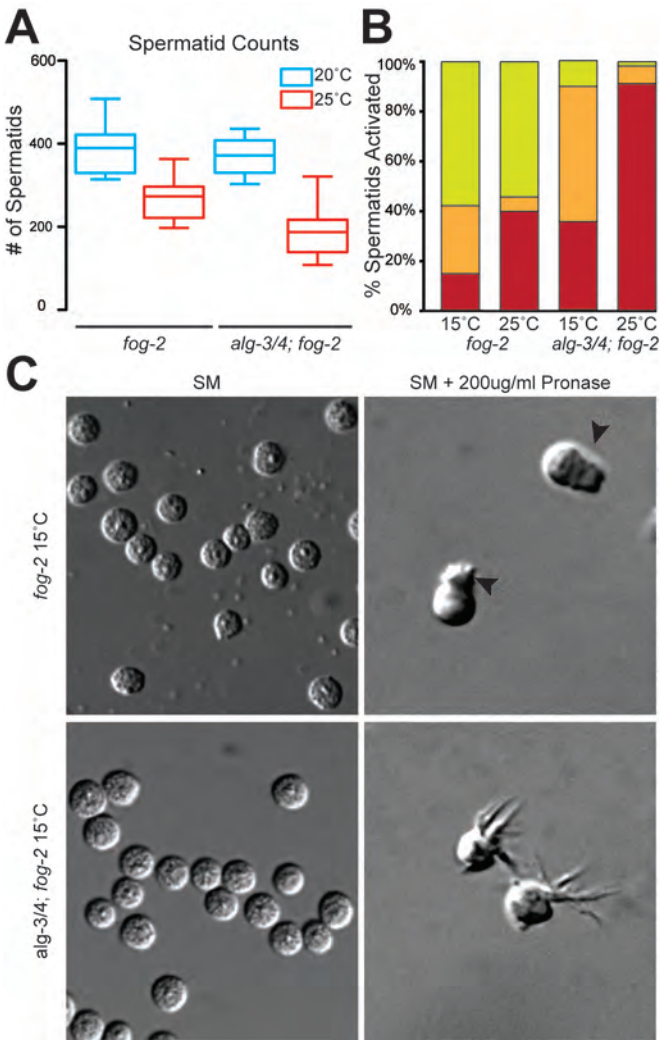

Figure 2.6. *alg-3/4* mutants exhibit defects in sperm activation.
(A) Box-and-whisker plots of spermatid counts performed on (n>20) wild-type (*fog-2*) and *alg-3/4*;*fog-2* animals (as indicated) cultured at 20°C (blue) or 25°C (red). **(B)** Graphic depiction of spermatid activation, illustrating the percent of spermatids with pseudopods (green), spikes (orange), or unactivated (red) (n>200). **(C)** DIC images of wild-type (*fog-2)* and *alg-3/4*;*fog-2* spermatids before activation (left panels) and after (right panels). Black arrowheads denote pseudopods.

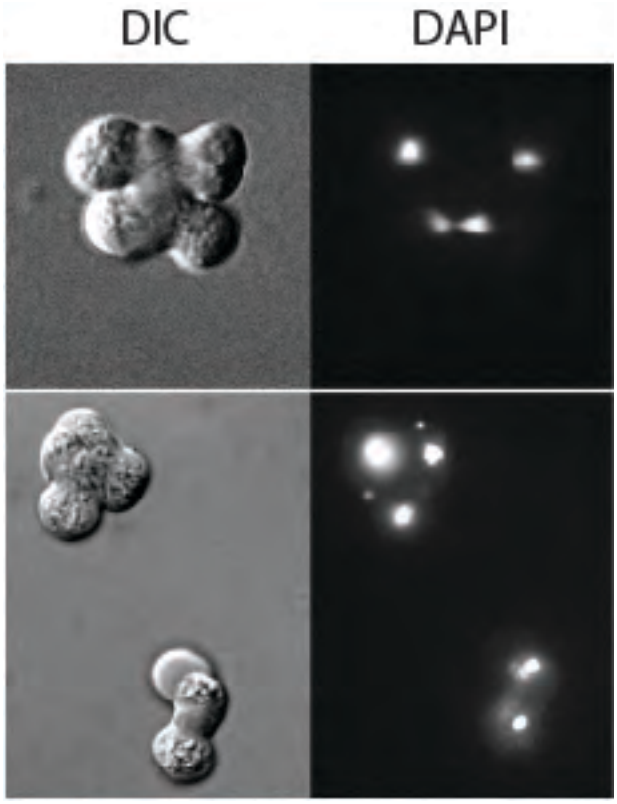

Figure 2.7. *alg-3/4* mutant defects in spermatogenesis.
DIC and DAPI images of *alg-3/4* developing spermatocytes with defects in DNA intergrity and budding from the residual body.

A class of sperm-specific 26G-RNAs requires *alg-3* and *alg-4*

Small RNAs that engage specific AGOs are often depleted in the corresponding AGO mutant backgrounds (Batista et al., 2008; Grishok et al., 2001; Gu et al., 2009). Therefore, we used Illumina deep sequencing to identify small RNA species that were depleted in the *alg-3/4* mutant background. Small RNAs were isolated from male-enriched *fog-2* and *alg-3*/4 populations, as well as mature sperm isolated from *fem-3* hermaphrodites, which overproduce sperm. To ensure that both 5′ monophosphorylated and 5′ tri-phosphorylated small RNAs would be amenable to ligation, samples were pretreated with tobacco acid pyrophosphatase (TAP) prior to cloning (Gu et al., 2009). In both the wild-type and *alg-3/4* datasets, the majority of small RNAs were found within a large peak in the 21-23nt size range corresponding to microRNAs, 21U-RNAs, and 22G-RNAs (Figure 2.8A) (Batista et al., 2008; Gu et al., 2009; Ruby et al., 2006). Strikingly, a second, much smaller peak at 26nt was observed in the wild-type dataset but was completely absent in the *alg-3/4* dataset (Figure 2.8A). A strong bias for guanine at the 5´ residue of the 26 nt RNA species (86% of the 67,206 reads in wild type) allowed us to identify these small RNAs as 26G-RNAs (Ruby et al., 2006). Interestingly, *alg-3*/*4*-dependent 26G-RNAs exhibited a strikingly lower abundance in mature sperm when compared to the whole male samples. Overall, 26G-RNAs were 7-fold less abundant in mature sperm, a finding consistent with the expression pattern of GFP::ALG-3 protein, which was detected in developing spermatocytes but not in spermatids.

Most male-enriched 26G-RNAs were antisense to the expressed regions of endogenous *C. elegans* genes (Figure 2.8B and (Ruby et al., 2006)). Altogether, 3191 genes were targeted by 26G-RNAs found in our wild-type samples, including 733 genes whose mRNAs were previously identified as enriched during spermatogenesis (Reinke et al., 2004). In contrast, only 1093 genes (279 spermatogenesis-enriched) were targeted by 26G-RNAs in the *alg-3/4* mutant data sets. To create a more stringently defined set of 26G-targeted genes, we applied a cutoff that included only genes targeted by at least 10 26G-RNA reads per million (RPM) total reads. This higher-confidence set included 400 genes targeted by 26G-RNAs in wild-type, while only 4 genes were targeted in *alg-3*/*4* mutants. Among the 400 stringently-defined 26G-RNA targets, 397 genes (>98%) were at least 2-fold depleted of 26G-RNAs in *alg-3*/*4* mutants (Figure 2.8E). The 26G-RNAs that were not dependent on *alg-3*/*4* activity were dependent on *ergo-1*, which encodes an AGO required for a subset of 26G-RNAs that are abundantly expressed in embryos (Vasale et al., 2010).

We used microarray analysis of young adult RNA preparations to ask if the mRNAs targeted by 26G-RNAs were misregulated in *alg-3*/*4* mutants (Figure 2.8C). Strikingly, we found that 109 of the 397 most stringently defined targets were upregulated by 2 fold or more in the *alg-3*/*4* mutants (Fig. 2.8C, Fig. 2.8E). Real-time quantitative RT-PCR (RT-qPCR) analysis of seven *alg*-3/4 26G-RNA targets confirmed the microarray analysis: Five targets that were upregulated in the microarrays were also upregulated by RT-qPCR, while two targets that were

unchanged in the microarray analysis were also unchanged in the RT-qPCR assay (Figure 2.8D). As expected, an ERGO-1-dependent 26G-RNA target was not upregulated in *alg-3*/*4* mutants.

ALG-3/4-dependent 26G-RNAs were not distributed randomly along their targets. Instead we noted a marked bias for accumulation at the termini of the transcripts (Figure 2.9). Approximately half of the targets exhibited a bias for 5´ accumulation of 26G-RNAs, while 34% exhibited a 3´ bias (Figure 2.8E). Strikingly, the mRNA levels of genes targeted by 26G-RNAs with a bias for 5´ accumulation showed higher levels of upregulation: 85% of the 109 targets upregulated in the *alg-3*/*4* mutants based on microarray analysis exhibited a 5´ bias for 26G-RNA accumulation. In contrast, only 5.5% of upregulated targets exhibited a 3´ bias. The remaining 9.5% exhibited no bias (Fig. 2.8E).

Finally, we used Northern-blot analysis to examine the expression of representative 26G-RNA species. We found that 26G-RNAs targeting *f36h12.4* and *ssp-16* were not detectable in *alg-3*/*4* mutants when compared to wild type, but were unaltered in *ergo-1* mutants (Figure 2.8F). In *alg-3(tm1155) and alg-4(ok1041)* single mutants these 26G-RNAs were only partially reduced (Figure 2.8F). Conversely, a probe designed to detect 26G-RNAs from an *ergo-1*-dependent 26G-RNA target (Vasale et al., 2010) *c40a11.10,* revealed strong depletion in the *ergo-1* mutants, but showed no change in abundance in *alg-3*/*4* mutants. Spermatogenesis-expressed 26G-RNAs were also missing in *rrf-3 and eri-1,* supporting the placement of these small RNAs in the ERI pathway.

Interestingly, the presence of these 26G-RNAs were only partially reduced in, a multiple mutant lacking all 12 WAGO genes (MAGO-12). These findings are consistent with the idea that spermatogenesis-specific 26G-RNAs depend on *alg-3* and *alg-4* while 26G-RNAs expressed at other stages depend on *ergo-1* (Figure 2.8F, and (Vasale et al., 2010)).

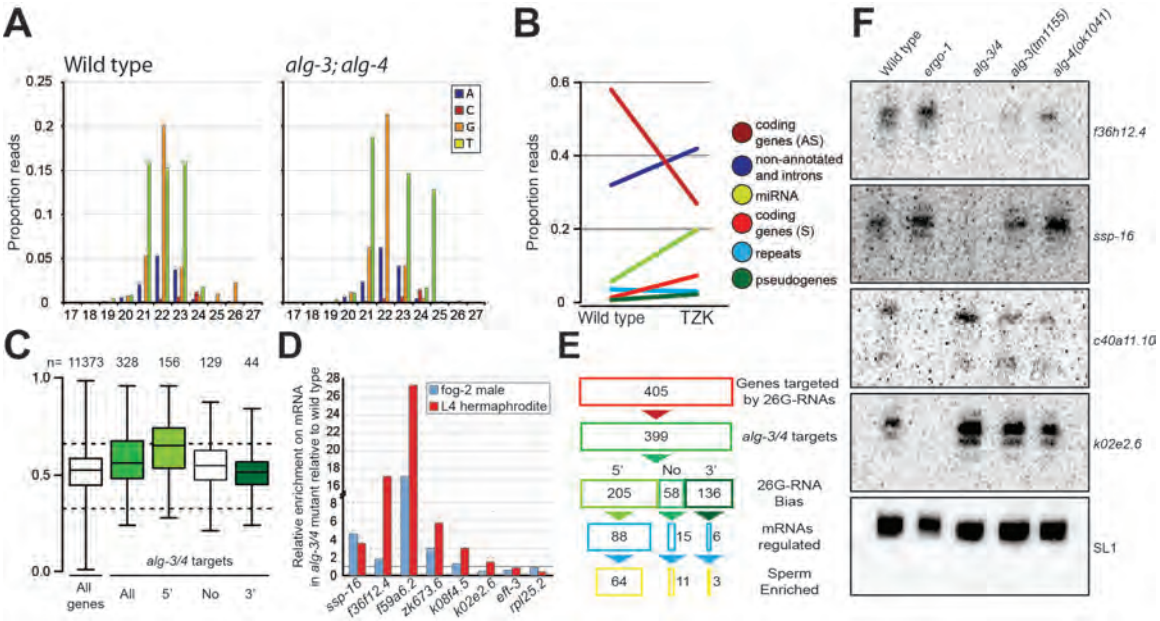

Figure 2.8. Analysis of 26G-RNA expression and targeting.
(A) Length and first-nucleotide distribution of deep-sequencing reads from wild-type (*fog-2) (left),* and *alg-3/4; fog-2 (right).* (B) Two-point plots comparing the relative proportions of various small RNA classes (as indicated by color) in wild-type (*fog-2)* [left point] and *alg-3/4*; *fog-2* mutants [right point]. (C) Box-and-whisker plots depicting relative mRNA levels in microarray assays on *alg-3/4* (mutant) and N2 (wild-type) populations. Here and in figure 2.10B, the Y axis represents the relative proportion of reads (measured as[*alg-3/4* mutant value divided by (wild-type plus *alg-3/4* mutant values)] for any given locus). Dotted lines indicate the values corresponding to 2-fold enrichment (a value of 0.66) or depletion (a value of 0.33). (D) RT-qPCR measurement of target mRNAs upregulated (bold type) or not regulated (regular type) based on microarray analysis. *k02e2.6* is an *ergo-1* dependent 26G-RNA target. (F) Schematic representation of 26G-RNA targets defined using a cutoff of 10 reads per million. (G) Northern blot analysis of small RNAs in wild-type and various mutant backgrounds as indicated. SL1 and *mir-66* are loading controls.

Figure 2.9. 26G-RNA distribution on targets.
f59a6.2 exhibits *5′* bias, *eft-3* exhibits 3′ bias, and *ssp-16* exhibits no bias. Cyan arrows represent 1-3 reads, violet 3-10 reads, and red greater than 10 reads. To emphasize this bias the number of reads in some cases are labeled.

wago-1 and *alg-3/4*-dependent 22G-RNAs are expressed in mature spermatids

Two recent studies identified distinct 22G-RNA populations that interact with the Argonautes WAGO-1 and CSR-1 (Claycomb et al., 2009; Gu et al., 2009), but did not explore 22G-RNA expression in sperm. We found that 22G-RNAs represented ~70% of all small RNAs cloned from sperm (Figure 2.10A), a proportion comparable to that observed in small RNAs recovered from gravid adult samples (Gu et al., 2009). Among the 397 targets that exhibited depletion of 26G-RNAs in *alg-3/4* mutants, 185 also exhibited a 2-fold or greater depletion in 22G-RNAs (Figure 2.10B). In contrast to the expression pattern of 26G-RNAs, which were less abundant in mature sperm samples than in male-enriched samples, the overall level of 22G-RNAs that target *alg-3/4* 26G-RNA targets was not changed in the sperm sample relative to the male-enriched sample.

Because WAGO proteins interact with 22G-RNAs (Gu et al., 2009), we asked whether WAGOs were present in and required for spermatid function. Consistent with this possibility MAGO-12 mutants exhibit a temperature-sensitive sterile phenotype (Gu et al., 2009). We found that the sterility of MAGO-12 hermaphrodites could be partially rescued by mating with wild-type males: Crosses with wild-type males at 25°C produced an average brood of 29.8 progeny, compared to an average brood of 5.7 for self-mated hermaphrodites. Further, we found that a GFP::WAGO-1 translational fusion was expressed throughout the germline. And unlike GFP::ALG-3 was also present in mature

spermatids (Figure 2.10C).

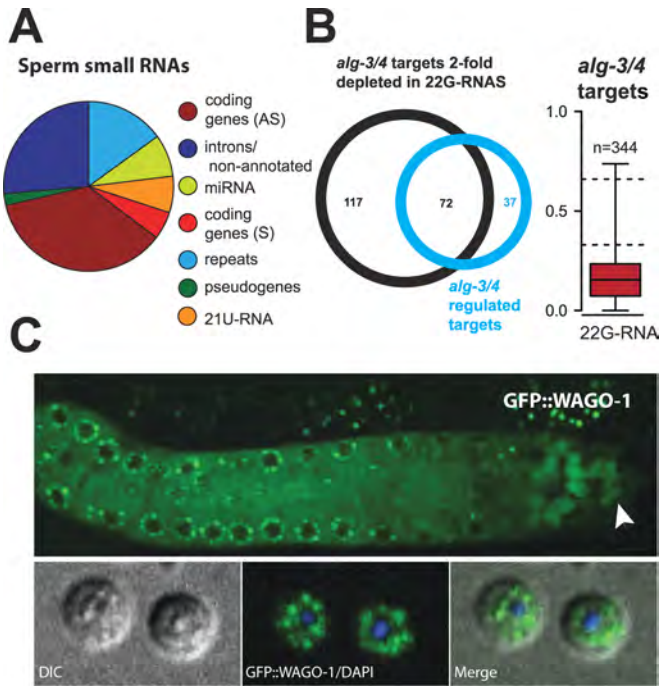

Figure 2.10. Analysis of Small RNA pathways in mature sperm.
(A) Pie chart representing the distribution of different classes of small RNA present in isolated spermatids. The 26nt RNAs represent less than 2% of the total small RNA reads. **(B)** *alg-3/4* targets are also 22G-RNA targets. The Venn diagram shows intersection between regulated *alg*-3/4 targets (based on microarray) and targets that are depleted 2-fold or greater of 22G-RNAs. The box-and-whisker diagram shows depletion of 22G-RNAs relative to wild-type on *alg-3/4* targets. **(C)** GFP::WAGO-1 expression in an adult male (*upper panel*). Spermatids are marked with a white arrowhead. GFP::WAGO-1 expression in individual spermatids, also stained with DAPI (blue) (*lower panel*).

Discussion

Argonautes and their associated small RNAs function in gametogenesis in all metazoans. The best studied examples are Piwi-clade AGOs, which function in the suppression of transposon activity in mammals, insects (reviewed in (Aravin et al., 2007)), and nematodes (Batista et al., 2008; Das et al., 2008). Here, we have shown that two AGO-class paralogs, *alg-3* and *alg-4,* function during male gametogenesis to promote fertility at elevated temperatures (thermotolerance). Interestingly, these AGOs are required for a species of 26nt RNA called (26G-RNAs) that are anti-sense to hundreds of spermatogensis-enriched mRNAs (rather than to transposons or repetitive elements). Moreover, many of the genes targeted appear to be downregulated in response.

A subset of 26G-RNAs are expressed during embryogenesis and depend on a distinct Argonaute, ERGO-1 (Han et al., 2009; Vasale et al., 2010). Vasale et al, propose a model for ERGO-1-dependent 26G-RNAs that involves two rounds of RNA templated small RNA production and mRNA targeting. In this two-step model, ERGO-1 interacts directly with 26G-RNAs produced by the ERI-complex-associated RNA dependent RNA polymerase (RdRP), RRF-3. The resulting ERGO-1/26G-RNAs then recognize mRNA targets. However, rather than directly down regulating these targets, the initial targeting recruits a second RdRP (RRF-1 and/or EGO-1), which then produces 22G-RNAs that are loaded onto WAGO-class Argonautes to mediate silencing (Vasale et al., 2010). Finally, whereas ERGO-1 and its associated 26G-RNAs are abundant during

embryogenesis, the corresponding 22G-RNAs on the ERGO-1 targets are abundant throughout later larval development and into adulthood.

Several of our findings are consistent with a similar two-step model for ALG-3/4-dependent 26G-RNAs (Figure 2.11A). First of all, ALG-3 expression is required for and coincides with expression of the spermatogenesis-expressed 26G-RNAs. Second, as was the case for ERGO-1, ALG-3/4 are also required for the expression of 22G-RNAs on their targets. Finally, whereas both ALG-3 and 26G-RNAs are depleted in mature sperm, WAGO-1 and the ALG-3/4-dependent 22G-RNAs remain abundant in mature sperm. Taken together these findings support a model in which ALG-3/4 are loaded with 26G-RNAs produced by the ERI-Dicer-complex during spermatogenesis. These in turn, may induce the recruitment of a second RdRP to produce 22G-RNAs that persist as WAGO-1 complexes in mature sperm. This two step AGO system could function to control the levels of mRNAs important for sperm function, perhaps by downregulating in mature sperm a set of transcripts expressed during spermatogenesis (Figure 2.11B). Persistence of WAGO-1 22G-RNA complexes in mature sperm may also provide a mechanism for the inheritance of epigenetic silencing signals important for fertility.

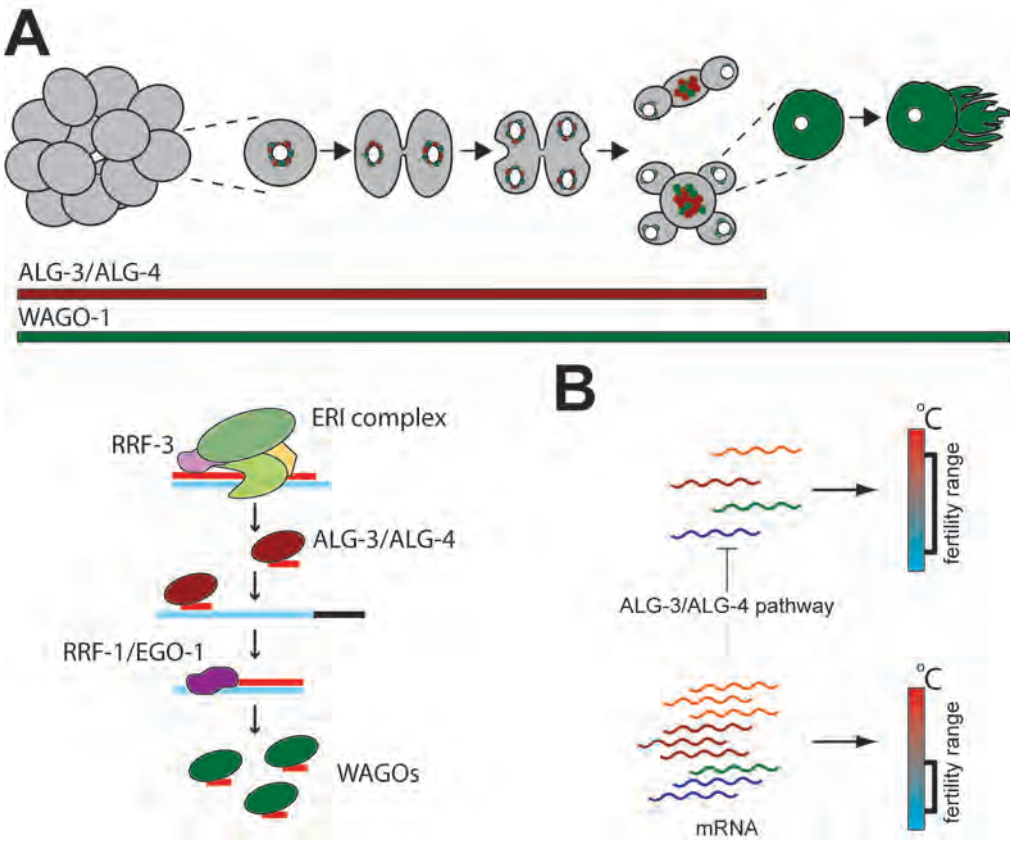

Figure 6. Model of ALG-3/4 and WAGO-1 expression during sperm development
(A) Depiction of spermatogenesis and spermiogenesis, with ALG-3/4 and WAGO-1 expression indicated by the red and green bars (respectively), as well as a model for the biogenesis of 26G and 22G RNAs. The schematic in **(B)** illustrates a potential role for ALG-3/4 in lowering target-mRNA levels to increase robustness to temperature.

AGOs, temperature sensitive sterility, and P granules

Wild-type animals exhibit sterility when cultured just 2°C above the optimum growth range of 20-25°C, suggesting that one or more aspect of gametogenesis involves an inherently temperature-sensitive process. In addition to the *alg-3*; *alg-4* mutants and the 12-fold WAGO mutants discussed above, mutations in the Piwi-related AGO, *prg-1,* cause temperature-dependent sterility. Mutations in a fourth AGO, *csr-1,* cause nonconditional sterility and chromosome segregation defects.

Interestingly, ALG-3, WAGO-1, PRG-1 and CSR-1 all localize to germline nuage structures called P granules, and *csr-1* mutants exhibit dramatic mislocalization of the P granules away from the periphery of germ-cell nuclei (Batista et al., 2008; Claycomb et al., 2009; Gu et al., 2009). Furthermore, mutations in other P granule components that cause defects in the localization of the P granules, including *pgl-1* and *glh-1* (Kawasaki et al., 1998; Kuznicki et al., 2000), also result in temperature-dependent sterility. P granules appear to dock with nuclear pores and are thought to be processing centers for germline transcripts (Anderson and Kedersha, 2009). Thus it seems likely that P granules function broadly to organize post-transcriptional regulation (and perhaps aspects of transcriptional regulation) in the germline, and that at least some of these regulatory mechanisms are inherently sensitive to temperature.

Adaptation to temperature and small-RNA pathways

AGO systems appear to have independently evolved genome-scale regulatory capacities in diverse organisms. For example, the AGO-mediated micro-RNA (miRNA) systems of plants and animals appear to have independent evolutionary origins, and yet in both systems miRNAs have acquired hundreds of targets, many with subtle affects on gene expression (Bartel, 2004; Bushati and Cohen, 2007). The ALG-3/4 system provides another example of an independently evolved AGO system that has acquired hundreds of targets. The modularity of AGO/small-RNA-mediated targeting permits a single class of AGO protein to interact with hundreds or thousands of different small RNA cofactors, each of which can in turn regulate the expression of multiple targets. Consequently, the expression and sequence of the AGO itself, as well as the expression and sequence of each specificity-determining small-RNA, can evolve independently. We hypothesize that these features provide AGO systems with the capacity to evolve rapidly, giving them the ability to superimpose new regulation on existing gene-expression networks.

The rules that govern the recognition of a transcript and trigger the biogenesis of 26G-RNAs are not yet known. The structure or expression of the target gene or transcript may promote recognition. Alternatively, all transcripts might be sampled stochastically during nuclear export and processing in the P granule as discussed above. Whatever the mechanism for initial targeting, the presence of 22G-RNAs in mature sperm could drive a feed-forward mechanism

that reinforces the recognition and silencing of ALG-3/4 targets in the next generation. Thus the ALG-3/4 system may function both broadly and heritably to promote robustness to temperature by functioning broadly in the regulation of a multitude of targets whose silencing improves the robustness of sperm to temperature.

Gametogenesis appears to be an inherently thermosensitive process both in *C.elegans* and in other metazoans. In most mammals core body temperatures are lethal to sperm, and external male gametogenesis appears likely to represent an adaptation that was basal to the evolution of endothermy in the vertebrate lineage (Werdelin and Nilsonne, 1999). Understanding how the *alg*-3/4 pathway promotes thermotolerant sperm development is likely to uncover general principals of gene regulation important for fertility, development and germline maintenance in diverse organisms.

EXPERIMENTAL PROCEDURES

Worm strains and genetics

C. elegans culture and genetics were as described in (Brenner, 1974). The Bristol N2 strain was used as the standard wild-type strain. Alleles used, listed by chromosome, include: Unmapped: *neIs23[unc-119(+) GFP::ALG-3]*; LGII: *neSi1[cb-unc-119(+) GFP::WAGO-1]*; LGIII: *alg-4(ok1041)*, *alg-4(1184)*, *unc-119(ed3)*; LGIV: *alg-3(tm1155)*, *fem-(q20)*; LGV: *fog-2(q71)*, *ergo-1(tm1860)*.

Strains: WM200: *(alg-3(tm1155); alg-4(ok1041)),* **WM201**: *(alg-3(tm1155); alg-4(ok1041); fog-2(q71)),* **WM202**: *(unc-119(ed3); neIs23[unc-119(+) GFP::ALG-3]; alg-3(tm1155); alg-4(ok1041)),* **WM203**: *(unc-119(ed3); neIs23[unc-119(+)* GFP::ALG-3]; *alg-3(tm1155)*; *alg-4(ok1041)*; *fog-2(q71)),* **WM204**: (unc-119(ed3); ttTi5605; neSi1[cbunc-119(+) GFP::WAGO-1] II), RFP::PGL-1; **WM191:** MAGO-12 (Gu et al., 2009).

Transgenic constructs: GFP::ALG-3 as follows: A Not1 site was engineered between the ALG-3 promoter and ORF-3′UTR using PCR, followed by Not1 digestion and ligation of the promoter and ORF-3′UTR, followed by cloning into the TOPO TA vector (invitrogen). GFP with Not1 sites on both termini was then ligated in between the promoter and ORF after digestion. The Gateaway cloning system (invitrogen) was then used to transfer the GFP fusion to the pCT045 destination vector. The resulting plasmids were introduced into *unc-119(ed3)* strain using biolistic transformation according to (Praitis et al., 2001). Transgenic strains were identified and integrated lines were crossed into

alg-3/4(tm1155;ok1041) background.

Phenotypic characterization of *alg-3/4* mutants

Brood size analysis was performed as described (Batista et al., 2008). Males were enriched from *fog-2* or *alg-3/4;fog-2* cultures as described (Miller, 2006). Spermatid activation was performed as described (Shakes and Ward, 1989). Spermatid numbers were determined by imaging male worms stained with DAPI. Microscopy was performed as described (Claycomb et al., 2009).

Spermatid numbers were determined by imaging male worms stained with DAPI. Worms were grown synchronously for 65-70 hrs at 20◦C or for 46-51 hrs at 25◦C. Worms were washed off the plate with M9, and then incubated in 70% ethanol for 5 minutes, followed by 3 washes with PBS. Samples were then incubated with DAPI (5 ug/ml) for 5 minutes, and again washed 3 times with PBS. Finally worms were put onto indentation slides, and DAPI-stained spermatid nuclei were counted using fluorescence microscopy.

Gonads and sperm were excised from worms in 1x sperm salts containing 2mM levamisole and DAPI (5ug/ml) on poly-L-lysine coated slides. Images in were acquired using Solamere Technology Group CSU10B Spinning Disk Confocal System scan head mounted on a Nikon TE-2000E2 inverted microscope with a 40x Plan/APO Oil lens and a Roper Coolsnap HQ2 camera. Metamorph, ImageJ, and Adobe Photoshop software were used to analyze the images. Z sections ranging from 0.2 to 0.3µm were collected from live worms

immobilized by 2mM levamisole (Sigma).

Molecular Biology

Enrichment of small RNA and Northern blot analysis were performed as described (Gu et al., 2009). Probe sequences (northern blot):

f36h12.4, tcatgttgccaatgattgcaattttcgtgctacttggatc

ssp-16, ttattagcattgtattcatacctatcatagaaaacc

c40a11.10, cggaatctcaaacttttccatcttgc

SL1, ctcaaacttgggtaattaaacc

Small RNAs extracted from isolated sperm, *fog-2* males and *alg-3/4;fog-2* males were pretreated with Tobacco Acid Pyrophosphatase (Epicenter Biotechnologies) and cloned using a 5′ ligation-dependent protocol (Gu et al., 2009). cDNA libraries were sequenced by the UMass Deep Sequencing Core using an Illumina GAII. RT-qPCR was performed as described in (Batista et al., 2008). Analyses of deep-sequencing and tiling array data were performed as described (Gu et al., 2009).

CHAPTER III

ARGONAUTES PROMOTE TRANSCRIPTION IN THE MALE GERMLINE AND PROVIDE A PATERNAL MEMORY OF GERMLINE GENE EXPRESSION IN C. ELEGANS

SUMMARY

During each life cycle germ cells preserve and pass on both genetic and epigenetic information. In *C. elegans,* the ALG-3/4 Argonaute proteins are expressed during male gametogenesis and promote male fertility. Here we show that the CSR-1 Argonaute functions with ALG-3/4 to positively regulate target genes required for spermiogenesis. Our findings suggest that ALG-3/4 functions during spermatogenesis to amplify a small-RNA signal that represents an epigenetic memory of male-specific gene expression. CSR-1, which is abundant in mature sperm, appears to transmit this memory to offspring. Surprisingly, in addition to small RNAs targeting male-specific genes, we show that males also harbor an extensive repertoire of CSR-1 small RNAs targeting oogenesis-specific mRNAs. Together these findings suggest that *C. elegans* sperm transmit not only the genome but also epigenetic binary signals in the form of Argonaute/small-RNA complexes that constitute a memory of gene expression in preceding generations.

INTRODUCTION

The transmission of information independently of the DNA sequence of the genome is termed epigenetic inheritance. During sexual reproduction both genetic and epigenetic information is passed to the zygote via specialized germ cells known as gametes. Gametogenesis involves dynamic molecular and morphological changes, culminating in the creation of highly specialized sperm and egg cells that package a haploid genome and all of the cellular machinery and epigenetic information necessary to launch zygotic development upon fertilization. Although many of the pathways required for gametogenesis are phylogenetically conserved (Eddy, 2002), especially those that mediate the partitioning of genetic information, very little is known about how gametes package and transmit epigenetic inheritance.

Male gametogenesis is an amazing example of cellular differentiation, in which undifferentiated male germ cells proceed through meiosis and develop into motile spermatozoa. In mammals the process of spermiogenesis, when spermatids differentiate into highly polarized motile spermatozoa, is initiated by a massive wave of gene expression essential for post-meiotic differentiation (Sassone-Corsi, 2002). Shortly thereafter, transcription ceases, and compaction of the haploid male genome ensues. Genome compaction within differentiating spermatids is facilitated by the replacement in chromatin of histones with small basic proteins called protoamines (Wykes and Krawetz, 2003).

The culmination of male gametogenesis is a motile gamete capable of initiating fertilization and delivering a paternal genome complement to an egg. However, genetic material is not the only information packaged in the sperm. Epigenetic information is also transmitted in the form of chromatin (DNA and/or histone) modifications, and RNA. In humans and mice paternal epigenetic factors have been shown to influence metabolism, stress response, and reproduction (Rando, 2012).

In *C. elegans* a major source of epigenetic inheritance involves Argonaute/small RNA pathways. Argonautes are structurally related to ribonuclease H and gain sequence specificity via small guide RNAs. Upon binding, Argonautes can direct endonucleolytic cleavage of target mRNAs, or can recruit cofactors that mediate post-transcriptional or transcriptional silencing (Ghildiyal and Zamore, 2009). In *C. elegans* mutations that perturb Argonaute pathways often result in infertility (Batista et al., 2008; Buckley et al., 2012; Claycomb et al., 2009; Conine et al., 2010; Gu et al., 2009; Han et al., 2009; Pavelec et al., 2009). For example, the Piwi Argonaute *prg-1* is required for both male and hermaphrodite fertility, and has been linked to transposon and transgene silencing (Batista et al., 2008; Ruby et al., 2006). PRG-1 engages over 30,000 distinct species of genomically-encoded small RNAs, termed Piwi-interacting (pi) RNAs (Batista et al., 2008; Gu et al., 2012). PRG-1/piRNA complexes are thought to utilize imperfect base pairing to scan germline-expressed mRNAs (Bagijn et al., 2012; Lee et al., 2012). When PRG-1/piRNA

complexes bind to foreign RNA sequences, such as those produced by a transgene, they initiate the production, via RNA-dependent RNA polymerase (RdRP), of amplified small RNAs called 22G-RNAs (Gu et al., 2009). These 22G-RNAs are in turn loaded onto members of an expanded group of Worm-specific Argonaute (WAGO) proteins (Yigit et al., 2006), which silence gene expression transcriptionally and post-transcriptionally (Buckley et al., 2012; Gu et al., 2009). This form of RNA-induced epigenetic silencing (referred to as RNAe) is then stably transmitted via both the sperm and the egg, apparently indefinitely through subsequent generations (Shirayama et al., 2012).

A major question related to the mechanism by which PRG-1 surveys germline gene expression is how certain mRNAs are recognized as self and protected from silencing. The CSR-1 Argonaute is a candidate factor for mediating self-recognition. CSR-1 is related to WAGOs but engages RdRP-derived small RNAs antisense to most if not all germline-expressed mRNAs. Therefore, it is possible that targeting by CSR-1 prevents PRG-1 recognition of self mRNA. If this model is correct, then a mechanism must exist during gametogenesis to package a cache of CSR-1 22G-RNAs reflecting the state of gene expression during each phase of the germline life cycle.

In this study, we investigate the role of Argonaute small RNA pathways during spermatogenesis in *C. elegans*. Previous work identified ALG-3 and ALG-4 (ALG-3/4) as redundant AGO-clade Argonautes that promote male fertility (Conine et al., 2010; Han et al., 2009). ALG-3/4 engage a class of Dicer and

RdRP-dependent small RNAs termed 26G-RNAs that are antisense to mRNAs expressed during spermatogenesis (Conine et al., 2010; Han et al., 2009; Pavelec et al., 2009). Here we show that sperm transcripts that are targeted by ALG-3/4 26G-RNAs are also targeted by CSR-1 22G-RNAs in the male germline. We show that *alg-3/4* and *csr-1* mutant males exhibit identical temperature-sensitive sterile phenotypes that result from failed spermiogenesis. Both ALG-3/4 and CSR-1 are required for robust transcription of spermiogenic mRNAs, suggesting that ALG-3/4 and CSR-1 function in the same pathway. Consistent with this, CSR-1 associates with the chromatin of spermiogenic transcripts and localizes to chromatin at the periphery of sperm nuclei in an ALG-3/4–dependent manner.

ALG-3/4 and 26G-RNAs are absent or greatly reduced in mature sperm (Conine et al., 2010). However, we show that CSR-1 and associated 22G-RNAs antisense to ALG-3/4 targets are abundant in mature sperm. Surprisingly, sperm also contain CSR-1 small RNAs antisense to female-specific germline mRNAs. We show that heterozygous offspring of homozygous *alg-3/4* or *csr-1* males exhibit reduced fertility. Moreover, repeatedly backcrossing heterozygous hermaphrodites to homozygous mutant males results in a progressive loss of fertility (germline-mortal phenotype) that can be rescued by wild-type sperm. Taken together, these findings are consistent with a model in which ALG-3/4 and CSR-1 and their associated small RNAs provide an epigenetic memory of

paternal gene expression, thereby ensuring the trans-generational continuity of a robust spermiogenic program.

RESULTS

Ultrastructural characterization of the *alg-3/4* mutant spermiogenic defect

Mutations in the Argonautes ALG-3/4 and in other enhanced RNAi (ERI) pathway components exhibit a temperature-sensitive sperm-defective phenotype (Conine et al., 2010; Han et al., 2009; Pavelec et al., 2009). At 20°C, *alg-3/4* mutant males produce many functional spermatozoa and are able to produce a brood size equal to approximately 50% (~150) of wild-type levels. However, at 25°C they produce haploid spermatids that fail to undergo spermiogenesis, the post-meiotic differentiation process where a pseudopod is formed thus creating spermatozoa competent for fertilization. At 25°C, *alg-3/4* spermatids occasionally make aberrant, multibranched ‘spiked’ pseudopods, but normally no activation is seen (Conine et al., 2010) which results in arrest as round-spermatids and complete male-specific sterility at 25°C.

To better characterize the spermiogenesis defect of *alg-3/4* mutants, we used transmission electron microscopy (TEM) to examine *alg-3/4* spermatozoa at an ultrastructural level (Figure 3.1). Wild-type *C. elegans* spermatids and spermatozoa exhibit several striking features that are visible using TEM. These include an electron-dense perinuclear halo containing RNA that encases the

nucleus and centrosome of the spermatids and persists through spermiogenesis in spermatozoa (Ward et al., 1981), as well as the Fibrous Body-Membraneous Organelle complex (FB-MO)(Roberts et al., 1986) (Figure 3.1). The morphogenesis of FB-MOs begins in 1° spermatocytes, where large organelles derived from the Golgi become associated with fibrous bodies comprised of polymerized Major Sperm Protein (MSP), the major structural component of the pseudopod (Roberts et al., 1986; Ward and Klass, 1982). In haploid spermatids, MOs localize to the plasma membrane while FBs dissolve, releasing depolymerized MSP into the cytoplasm. During spermiogenesis or activation of the spermatid the MOs fuse with the plasma membrane on one side of the spermatid, releasing their contents from the cell for oocyte signalling, while MSP protein polymerizes to form the pseudopod on the opposite pole.

At 20°C many *alg-3/4* spermatozoa form WT-looking pseudopods. However, MOs fail to become polarized and often fuse with the plasma membrane adjacent to the pseudopod rather than with the outer membrane on the opposite side of the spermatid (Figure 3.1). At 25°C *alg-3/4* spermatids never form a WT pseudopod and the RNA halo is frequently absent or appears abnormal. In addition, large tubule-like structures are present in the cytoplasm (Figure 3.1). These tubules are too large in diameter to be comprised of known cytoskeletal components, but instead are thought to be made up of polymerized perinuclear halo components (Ward et al., 1981). We also found that FBs fail to

disassemble in *alg-3/4* mutants at 25°C (Figure 3.1), a phenotype characteristic of *fer-6* spermatozoa, which also fail to make pseudopods (Ward et al., 1981).

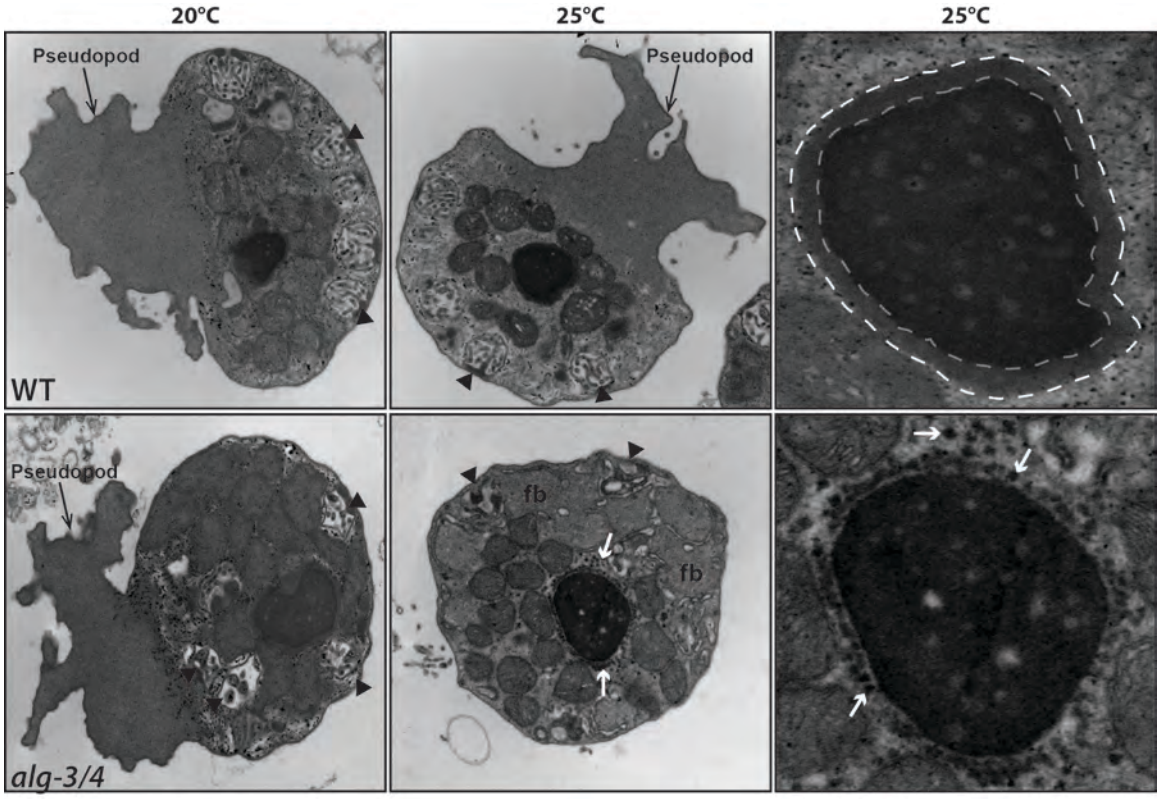

Figure 3.1. Ultrastructural analysis of *alg-3/4* spermatozoa.
Transmission electron micrographs of spermatozoa from wild-type (WT, top row) or *alg-3/4* males (bottom row) cultured at 20°C (left) or 25°C (center and right). Examples of membranous organelles (black arrowheads) and Fibrous bodies(fb) are indicated. Representative nuclei of WT and *alg-3/4* spermatozoa are shown at right. Dotted lines emphasize the RNA halo surrounding the wild-type nucleus. White arrows label large tubule-like structures (visible in cross-section (Ward et al., 1981)) adjacent to *alg-3/4* nuclei.

The previously identified *fer* genes function in the ALG-3/4 pathway

The fertility phenotypes of *alg-3/4* mutant sperm are identical to those of temperature-sensitive fertilization-defective (*fer*) mutants isolated in genetic screens more than 30 years ago (Argon and Ward, 1980; Hirsh and Vanderslice, 1976; Ward et al., 1981). To date, most of the *fer* genes remain molecularly uncharacterized except for *fer-1*, which encodes a member of the Ferlin family of membrane proteins required for the fusion of MOs during spermiogenesis (Washington and Ward, 2006).

The similarity of *alg-3/4* and *fer* mutant phenotypes extends to the ultra-structural level as assayed by electron microscopy (Figure 3.1). For example, *alg-3/4* mutants, like the previously characterized *fer* mutants arrest as round non-polarized spermatids in which sperm-specific organelles fail to undergo a series of stereotypic fusion events and structural reorganizations that drive pseudopod formation. Mutant spermatids also exhibit a missing or abnormal peri-nuclear RNA halo (Figure 3.1, Ward et al., 1981).

Given the striking similarities between the *fer* and *alg-3/4* mutant phenotypes, we asked whether the wild-type activities of the *fer* genes are required for the production of ALG-3/4 pathway 26G-RNAs. Northern blot analysis of small RNAs isolated from L4/Young adult staged hermaphrodites revealed that *fer-2, -4, -6*, and *-15* are required for the production of two abundant ALG-3/4 pathway 26G-RNAs targeting the mRNA products of *f36h12.4* and *ssp-16* (Figure 3.2A). In the *fer-3* mutant, 26G-RNAs targeting *f36h12.4* and

ssp-6 were significantly reduced or absent, suggesting that *fer-3(hc3)* may be a hypomorphic allele of a gene required for 26G-RNA biogenesis. Neither *fer-1* nor *fer-7* were required for the expression of either *f36h12.4* or *ssp-6* 26G-RNAs (Figure 3.2A). Consistent with this finding, *fer-1* and *fer-7* appear to act later in sperm development than *alg-3/4* and the remaining *fer* mutants tested here (Argon and Ward, 1980; Conine et al., 2010). Furthermore, *fer-1* mutant spermatozoa form short-stubby pseudopods with normal looking projections at 25°C, whereas *alg-3/4* and *fer-2, -4, -6* and *-15* mutant spermatozoa do not develop pseudopods at 25°C (Ward et al., 1981).

We found that *fer-3(hc3)* and *fer-15(b26)* exhibit an Eri phenotype (Figure 3.2B) and, like other Eri mutants, are required for ERGO-1 pathway 26G-RNAs targeting the gene *c44b11.6* (Figure 3.2A). Indeed, *fer-3* and *fer-15* mutations map nearby the ERI-pathway genes *eri-3* and *rrf-3*, respectively (Roberts and Ward, 1982; Ward et al., 1981). *fer-3(hc3)* failed to complement the Fer and Eri phenotypes of *eri-3(tm1361)* (Figure 3.2B). Consistent with this finding, and the partial loss of male-specific 26G-RNAs, we identified a nucleotide substitution in exon 3 of *eri-3* that is predicted to result in a serine-to-proline missense mutation at amino acid 69 (Figure 3.2C). *fer-15(b26)* failed to complement the Fer and Eri phenotypes of *rrf-3(pk1426)*, and we identified a 223 bp deletion that removes exon 6 of *rrf-3* and is predicted to shift the reading frame and introduce a premature termination codon (Figure 3.2C). *fer-2*, *fer-4* and *fer-6* mutants were not Eri, and their molecular identities remain to be determined. Taken together,

our findings indicate that several previously isolated Fer mutants define genes that function in the ALG-3/4-26G-RNA pathway.

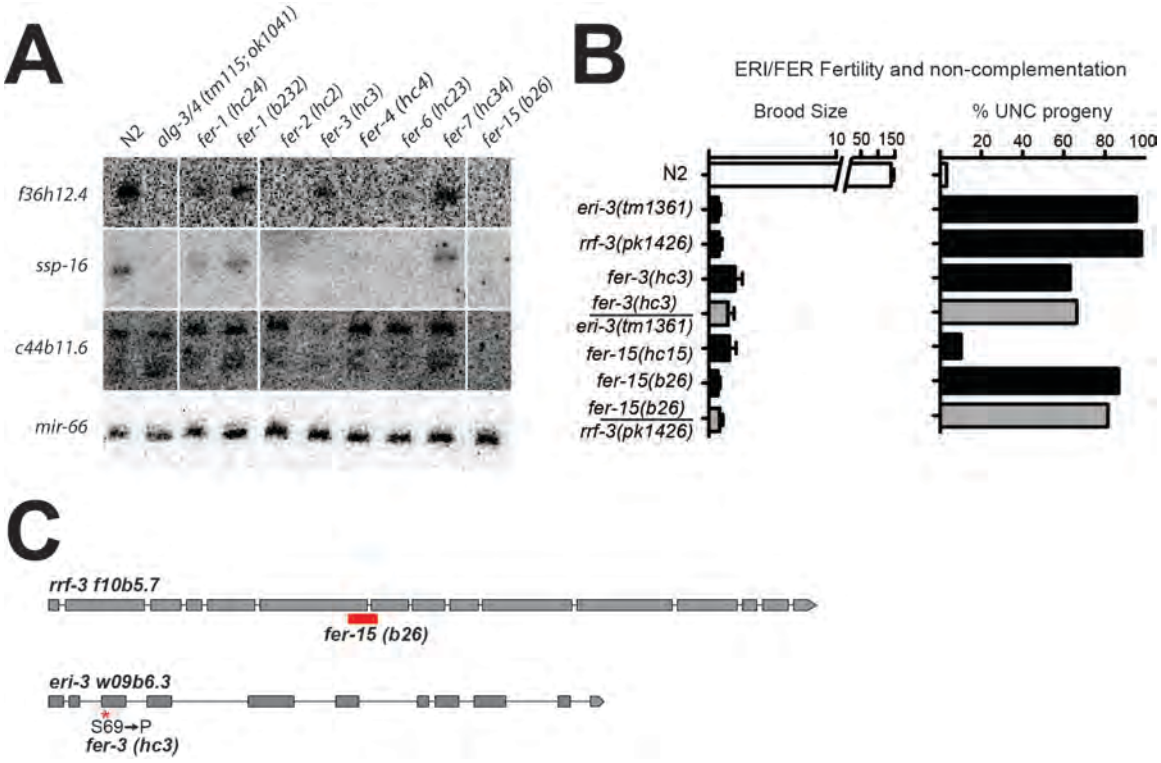

Figure 3.2. Identification of *fer* genes as ALG-3/4 pathway components.
(A) Northern blot analysis of small RNAs from whole male RNA extracts. Radiolabeled probes were used for two ALG-3/4 targets, *f36h12.4* and *ssp-16*, and for an ERGO-1–dependent 26G-RNA target *c44b11.6*. 22G-RNAs are also visible targeting *c44b11.6*. *mir-66* was probed as a loading control. **(B)** Genetic complementation tests for the fertility (Fer) (left panel) and Enhanced RNAi (Eri) phenotypes (right panel) between *fer-3* and *eri-3* and *fer-15* and *rrf-3*. Fertility was assayed at 25°C. Enhanced RNAi was assayed by *unc-73(RNAi)* feeding. **(C)** The exon-intron structures of *rrf-3* and *eri-3* are indicated by the boxes and lines, respectively. The *fer-15(b26)* allele corresponds to a 223 nt deletion (red box) that spans intron 6, results in a frame shift and introduces a premature termination codon in the *rrf-3* coding region The *fer-3(hc3)* corresponds to a missense mutation in *eri-3* that changes Serine 69 to Proline.

Target regulation by ALG-3/4 and 26G-RNAs

Our previous deep-sequencing studies identified 397 genes as high-confidence targets of ALG-3/4–dependent 26G-RNAs using a 10 reads-per-million (RPM) cut-off (Conine et al., 2010). To identify the full repertoire of ALG-3/4 targets, we cloned and deep sequenced small RNAs isolated from *alg-3/4* males and WT males cultured at both 20° and 25°C. In total, we identified 1497 genes targeted by 26G-RNAs in wild-type males at a density of at least 5 RPM. We found that 94% (1408) of these genes exhibited a >2-fold reduction in 26G-RNAs in *alg-3/4* mutants and were thus designated as ALG-3/4 target genes. The remaining 6% of genes targeted by 26G-RNAs that were unaffected in *alg-3/4* mutants are mostly ERGO-1–dependent 26G-RNA targets (Vasale et al., 2010). The expanded list of 1408 ALG-3/4 26G-RNA targets accounts for 63% (617/970) of genes with sperm-specific expression (Reinke et al., 2004).

Many ALG-3/4 targets are required for spermiogenesis and are involved in pseudopod formation and sperm motility. For example, these targets include genes that encode MSPs and MSP-related proteins (Burke and Ward, 1983), as well as factors required for sperm motility (Buttery et al., 2003). Because the activities of ALG-3/4 and of many of their target genes are required for spermiogenesis, it therefore seemed unlikely that ALG-3/4 directs the silencing of these targets.

To determine how ALG-3/4 and 26G-RNAs regulate target mRNAs, we performed mRNA deep sequencing (mRNA-seq) of poly-A purified mRNA

isolated from WT and *alg-3/4* mutant males grown at 20° and 25°C. Comparing the relative levels of ALG-3/4 targeted mRNAs between *alg-3/4* mutant males and wild-type males cultured at 25°C, we found that 214 target mRNAs were positively regulated by 2-fold or more (decreased by 2-fold or more in the mutant), 204 were negatively regulated by 2-fold or more, and 991 target mRNAs did not change (Figure 3.3A). Strikingly, we found that positively-regulated mRNAs were much more abundant in WT males than were non-regulated (4-fold lower) or negatively-regulated (6-fold lower) mRNAs, when measuring reads per million per thousand basepairs (RPMK) from the 25°C mRNA-Seq datasets (Figure 3.3B). This enrichment of positively-regulated mRNAs was completely dependent on ALG-3/4 activity (was abolished in *alg-3/4* mutants, Figure 3.3B).

To determine whether the changes in male mRNA expression correlate with changes in the sperm proteome, we compared the proteomes of *alg-3/4* mutant spermatids and wild-type spermatids purified from males grown at 20°C or 25°C. To do this, we used stable isotope labeling by amino acids (Ong et al., 2002) to label wild-type proteins with the heavy isotope of nitrogen, ^{15}N. Unlabeled protein extracted from *alg-3/4* or wild-type sperm was combined with a known quantity of labeled wild-type sperm protein, and the mixture was analyzed by mass spectrometry (Moresco et al., 2010). For each mixture, we obtained a ratio of unlabeled peptides (from *alg-3/4* or wild-type at 20°C or 25°C) to labeled peptides (from wild-type sperm at 20°C) for each protein, which allowed us to

measure relative protein levels between *alg-3/4* mutant and wild-type sperm, or between temperatures.

Our proteomic analysis revealed 122 proteins that decreased by at least 1.5 fold in *alg-3/4* mutants relative to WT, and 43 that increased (Figure 3.4A). For each protein identified in our proteomic analysis, we compared the change in protein level in *alg-3/4* males at 25°C to the change in mRNA (Figure 3.4E). For most ALG-3/4-dependent positively-regulated proteins (81/122, 66%), the decrease in protein level in the *alg-3/4* mutant could be explained by a reduction in the corresponding mRNA. By contrast, for 33% (14/43) of ALG-3/4 negatively-regulated proteins, we found that the increase in protein could be attributed to changes in corresponding mRNAs (3.6B & 3.6E). These data indicate that the ALG-3/4 pathway promotes the expression of many of its targets and does so by increasing their mRNA levels.

In *alg-3/4* mutants, sperm defects are much more severe at high temperature (Conine et al., 2010). We therefore wished to know how ALG-3/4 target mRNA levels changed with temperature in both the wild-type and *alg-3/4* mutant strains. We found that, in wild-type males positively-regulated mRNAs increased at 25°C relative to 20°C, whereas negatively-regulated mRNAs tended to decrease at 25°C (Figure 3.3C). The converse was true for *alg-3/4* mutant males: positively-regulated mRNAs decreased dramatically at 25°C compared to 20°C, whereas negatively-regulated targets increased, albeit less dramatically (Figure 3.3D). Similar analyses of our proteomic data also corroborate these

findings (Figure 3.4C-D). The described changes in transcript levels by mRNA-seq were confirmed using RT-qPCR (Figure 3.5). These data support the existence of at least two distinct classes of ALG-3/4 target genes: one that is dramatically up-regulated by ALG-3/4 at elevated temperature, and one that is moderately down-regulated by ALG-3/4 at elevated temperature.

ALG-3/4-positively-regulated targets are clearly genes required for spermiogenesis and motility, including many of the MSP and MSP-related genes involved in pseudopod formation, as well as the PP1 phosphatases encoded by *gsp-3/4*, which are required for multiple aspects of *C. elegans* sperm development and when mutated exhibit temperature-sensitive (TS) sperm defects similar to *alg-3/4* (Wu et al., 2012). ALG-3/4 negatively-regulated targets include genes associated with mitochondrial function (electron transport chain and cellular respiration), but a correlation between reproductive genes and negatively-regulated ALG-3/4 targets was not observed.

To confirm that ALG-3/4 promotes the expression of target genes in the germline, we used immunofluorescence (IF) assays to detect MSP and GSP-3 proteins and fluorescence in situ hybridization (FISH) to detect *msp* transcripts. Interestingly, IF of MSP and GSP-3, two positively-regulated targets, suggested that these proteins were only slightly reduced in *alg-3/4* germlines during spermatogenesis and in haploid spermatids at 20°C compared to WT (Figure 3.6A). At 25°C, however, MSP and GSP-3 were severely reduced throughout the spermatogenic germline of *alg-3/4* males compared to WT and failed to

accumulate in spermatids (Figure 3.6A). Although GSP-3 protein was depleted overall, we found that it also accumulated abnormally in the nuclei of *alg-3/4* mutants, particularly at 25°C (Figure 3.6A). In addition, we observed more nuclei undergoing chromatin condensation in *alg-3/4* males compared to wild-type males (see below). Examining spermatids released from the testes of *alg-3/4* mutants raised at 25°C and activated to undergo spermiogenesis in vitro revealed that 12% of *alg-3/4* spermatozoa accumulate near WT levels of MSP and GSP-3. Unlike WT spermatozoa, however, these *alg-3/4* spermatozoa never formed pseudopods of normal morphology or polarity, with MSP and GSP-3 localized to the leading edge of the pseudopod (Figure 3.6B). Instead, we observed that: 1) MSP and GSP-3 remained in large punctae that resemble sperm organelles called Fibrous Bodies (FBs) (5%); 2) MSP localized to the cortex and GSP-3 to the cytoplasm of the spermatid but lacked polarity (3%); or 3) GSP-3 remained in the cytoplasm, and spermatids formed a multibranched, spikey pseudopod containing MSP (4%). Consistent with our mRNA-seq and RT-qPCR data, FISH revealed that *msp* mRNA was indeed reduced in the spermatogenic germline of *alg-3/4* males compared to wild-type males (Figure 3.6C). This reduction was more pronounced at 25°C, with very little signal detected in *alg-3/4* male germlines, consistent with the temperature-dependent regulation of expression (Figure 3.6C).

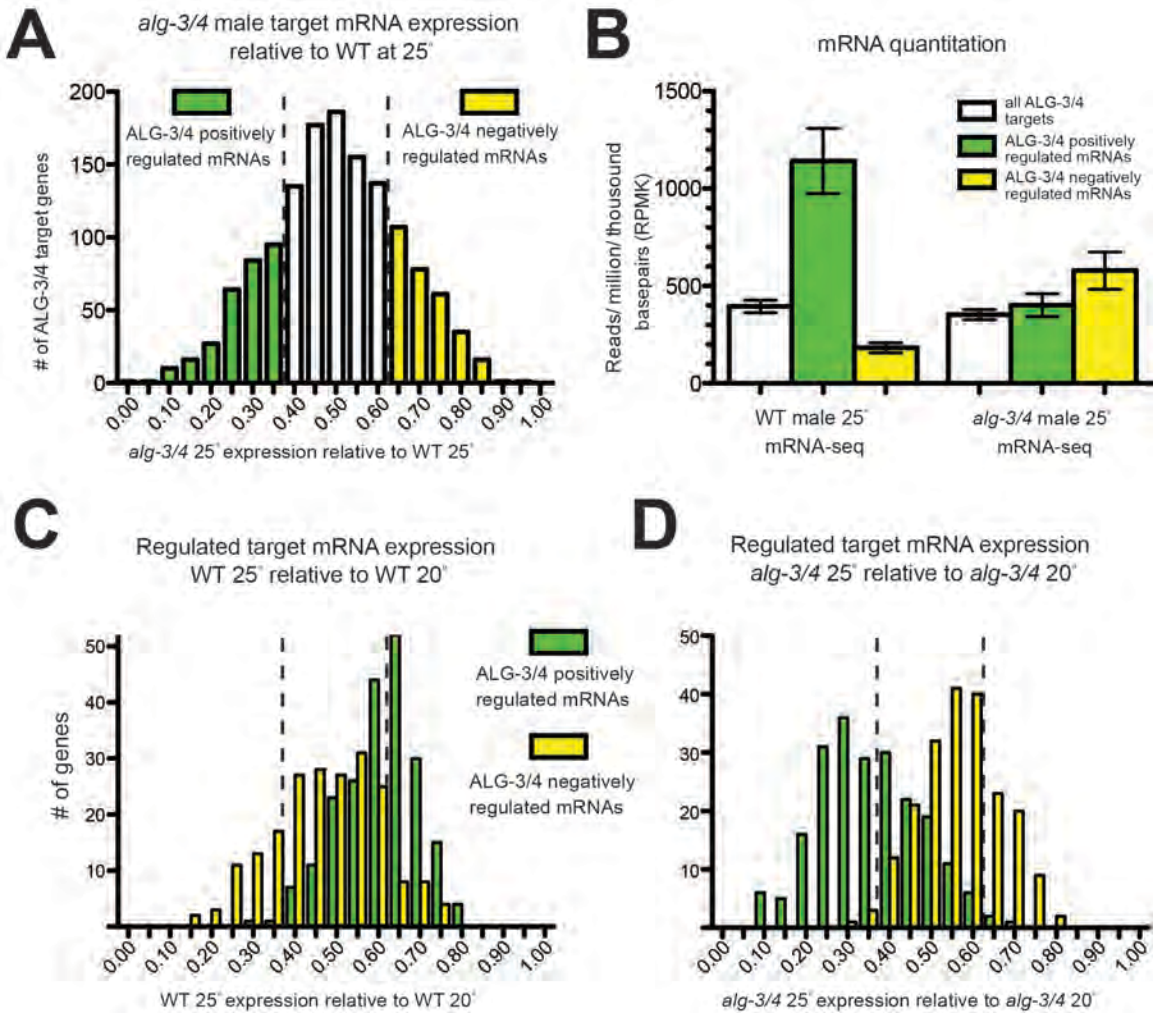

Figure 3.3. ALG-3/4 positively and negatively regulates hundreds of target mRNAs. **(A)** Histogram illustrating the enrichment or depletion of ALG-3/4 target mRNAs in *alg-3/4* mutants relative to WT at 25°C. Dotted lines indicate a two-fold change, values approaching 1 indicate enrichment (ALG-3/4 negatively-regulated mRNAs), and values approaching 0 indicate depletion (ALG-3/4 positively-regulated mRNAs). In A-D, colored bars indicate ALG-3/4 targets positively-regulated (green) or negatively-regulated (yellow) more than two-fold. Enrichment was calculated as the Reads Per Million (RPM) ratio of *alg-3/4* / (*alg-3/4* + WT). **(B)** Bar graphs showing the average RPM per kilobase (RPMK) for ALG-3/4 regulated targets in WT or *alg-3/4* mutant at 25°C. **(C and D)** Histogram illustrating the enrichment of positively- or negatively-regulated ALG-3/4 target mRNAs in WT (**C**) or *alg-3/4* (**D**) males at 25°C relative to 20°C. Enrichment was calculated as the RPM ratio of 25°C / (25°C + 20°C).

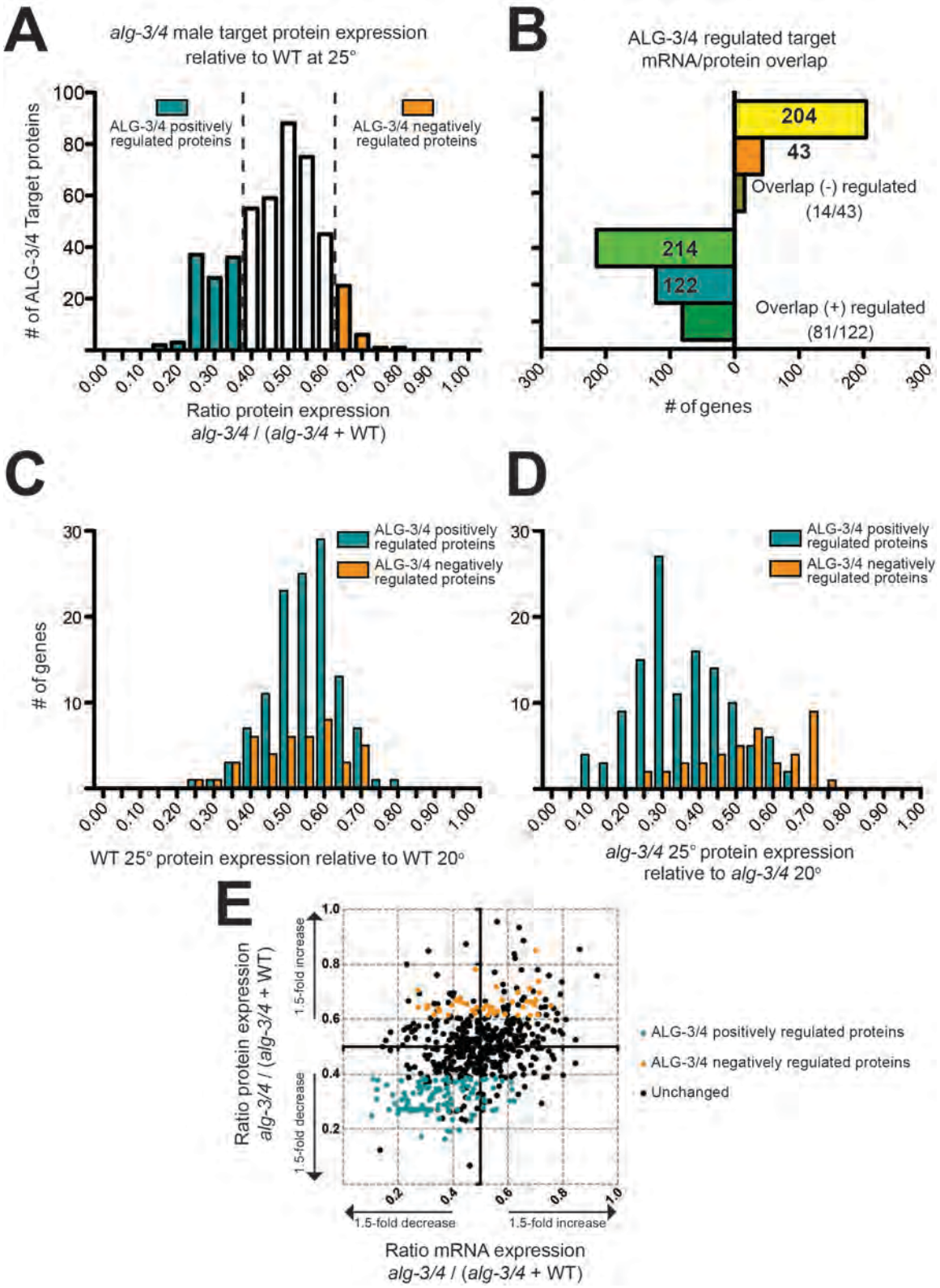
A
alg-3/4 male target protein expression relative to WT at 25°
# of ALG-3/4 Target proteins
ALG-3/4 positively regulated proteins
ALG-3/4 negatively regulated proteins
Ratio protein expression
alg-3/4 / (alg-3/4 + WT)
B
ALG-3/4 regulated target mRNA/protein overlap
204
43
Overlap (-) regulated (14/43)
214
122
Overlap (+) regulated (81/122)
# of genes
C
# of genes
ALG-3/4 positively regulated proteins
ALG-3/4 negatively regulated proteins
WT 25° protein expression relative to WT 20°
D
alg-3/4 25° protein expression relative to alg-3/4 20°
E
Ratio protein expression alg-3/4 / (alg-3/4 + WT)
1.5-fold increase
1.5-fold decrease
Ratio mRNA expression
alg-3/4 / (alg-3/4 + WT)
ALG-3/4 positively regulated proteins
ALG-3/4 negatively regulated proteins
Unchanged

Figure 3.4. Quantitative proteomic analysis of ALG-3/4 target in isolated sperm.
(A) Histogram illustrating the distribution of ALG-3/4 target protein expression in *alg-3/4* mutant sperm relative to WT sperm at 25°C, calculated by peptide ratio in *alg-3/4* divided by the peptide ratio in *alg-3/4* plus the peptide ratio in WT. Dotted lines represent either 2-fold enrichment (right, orange) or depletion (left, turquois) in *alg-3/4*. **(B)** Overlap between the ALG-3/4 positively-regulated proteins with the positively-regulated mRNAs. As well as the ALG-3/4 negatively-regulated proteins with the negatively-regulated mRNAs. **(C)** Histogram illustrating the distribution of ALG-3/4 target protein expression from WT sperm grown at 25°C relative to WT sperm at 20°C, calculated as the peptide ratio in WT 25°C divided by peptide ratio in WT 25°C plus the peptide ratio in WT 20°C. **(D)** Histogram illustrating the distribution of ALG-3/4 target protein expression in *alg-3/4* sperm grown at 25°C relative to *alg-3/4* sperm at 20°C. **(E)** Scatter plot comparing the expression of proteins (proteomic analysis) in *alg-3/4* sperm relative to WT at 25°C (y-axis) to the expression of mRNAs (mRNA-seq) in *alg-3/4* males relative to WT at 25°C (x-axis). Turquois dots represent ALG-3/4 positively-regulated proteins and orange dots represent ALG-3/4 negatively-regulated proteins.

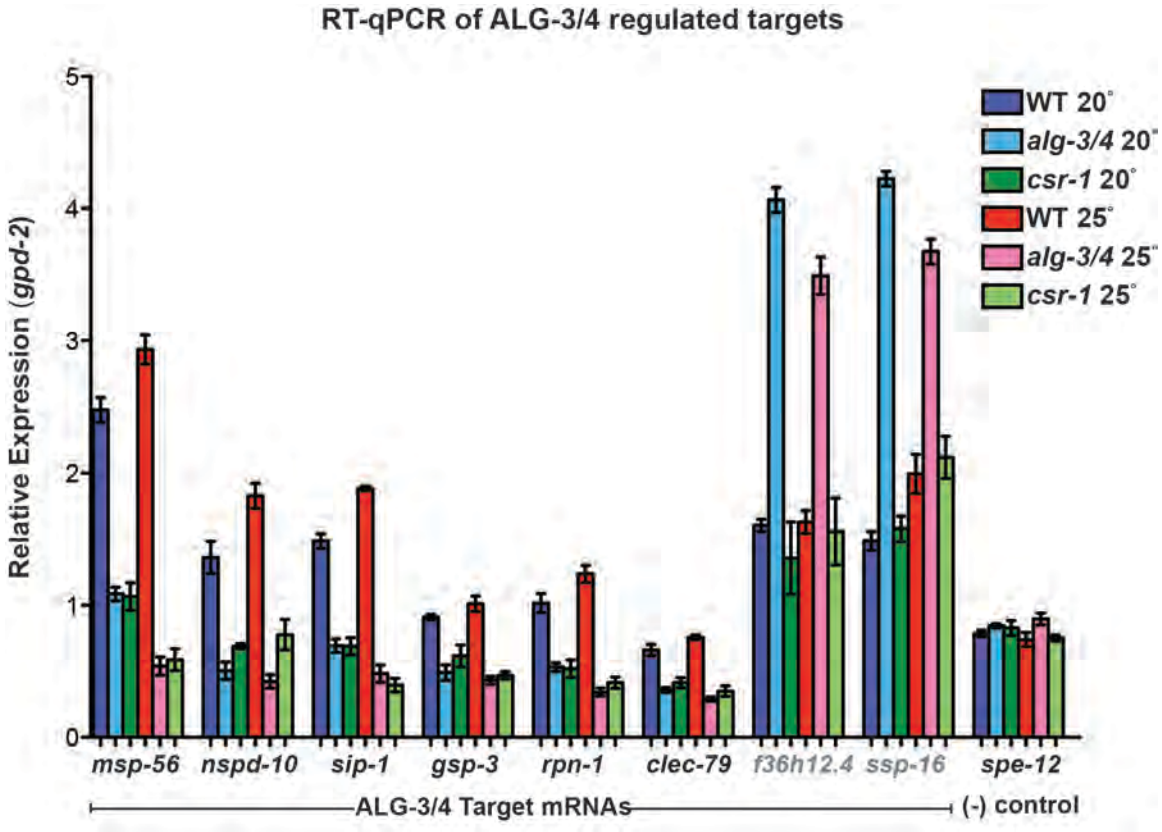

Figure 3.5. Target mRNA regulation by ALG-3/4 and CSR-1.
RT-qPCR analysis of ALG-3/4 targets normalized to *gpd-2*. Data are represented as mean +/- SEM.

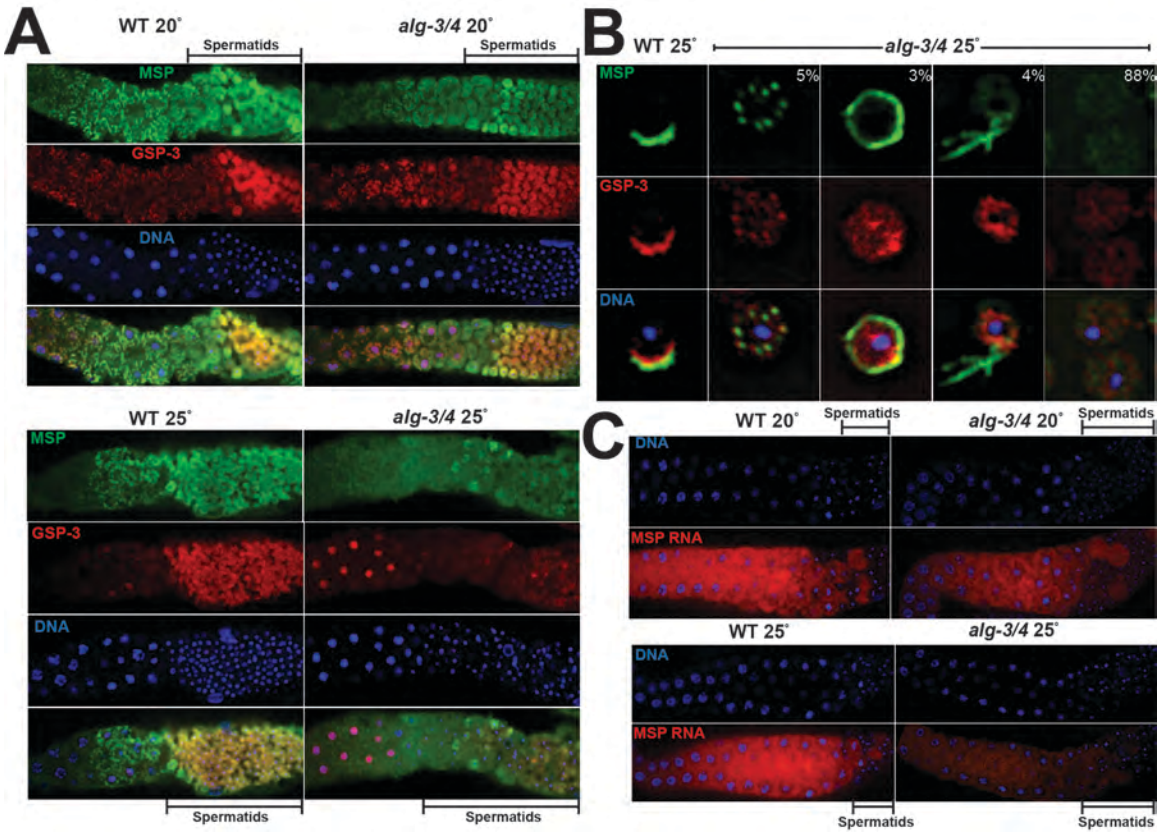

Figure 3.6. ALG-3/4 positively-regulates genes required for spermiogenesis.
(A-C) Confocal images of WT and *alg-3/4* mutant gonads **(A, C)** and activated spermatocytes **(B)**. Each column of photographs represent a single specimen imaged with different fluorescence channels (as indicated). A merged image is shown at the bottom. The genotype and temperature at which spermatogenesis occurred are indicated. DNA was visualized by DAPI staining (blue). In **(A and B)** the sperm proteins MSP (green) and GSP-3 (red) are visualized by Immunofluorescence. In **(B)** the percentage of WT and *alg-3/4* spermatozoa that exhibit the staining pattern shown is indicated. In **(C)** *msp* mRNA is visualized by fluorescence in situ hybridization (FISH).

ALG-3/4 promotes the transcription of target genes

Our findings suggest that ALG-3/4 positively regulates targets at the level of the mRNA. We therefore sought to determine whether ALG-3/4 promotes the transcription of positively-regulated targets. Using RNA Polymerase II (Pol II) chromatin immunoprecipitation (ChIP) followed by qPCR, we found that Pol II occupancy of positively-regulated ALG-3/4 targets decreased 1.5- to 3-fold in *alg-3/4* males compared to wild-type males cultured at 20°C. Notably, Pol II occupancy was reduced 3- to 7-fold in *alg-3/4* males grown at 25°C (Figure 3.7A). Consistent with the reduced Pol II occupancy in *alg-3/4* males, the pre-mRNAs of positively-regulated targets were also decreased in *alg-3/4* and *fer-15/rrf-3* males by 2- to 4-fold at 20°C and by 4- to 8-fold at 25°C. These findings suggest that both the Argonaute (ALG-3/4) and RdRP (FER-15/RRF-3) promote the transcription of positively-regulated targets (Figure 3.7B). Pol II occupancy of *spe-12* and *col-122*, two control genes expressed in male germline and somatic tissues (respectively) but not targeted by ALG-3/4 26G-RNAs, was not dependent on *alg-3/4* or temperature (Figure 3.7A).

Curiously, Pol II occupancy of the negatively-regulated gene *ssp-16* was reduced by 2- and 3-fold in *alg-3/4* males at 20° and 25°C, respectively, relative to wild-type levels (Figure 3.7A). Pol II occupancy of a second negatively-regulated target *f36h12.4* was unchanged at 20°C and reduced nearly 2-fold at 25°C compared to wild-type levels (Figure 3.7A). The expression of the *f36h12.4*

pre-mRNA mirrored these findings. Each of the five negatively-regulated targets that we examined showed a similar reduction in pre-mRNA at 25°C (Figure 3.7B). These results, therefore, suggest that negatively-regulated ALG-3/4 targets are silenced at a post-transcriptional level.

Because the WAGO 22G-RNA pathway is required for silencing (Gu et al., 2009), we tested whether the WAGO pathway is required to silence negatively-regulated ALG-3/4 targets at a post-transcriptional level. Indeed, we found that the mRNA levels of several negatively-regulated targets were increased in males lacking *rde-3* or deleted for all 12 WAGOs (MAGO-12; Figure 3.8). These increased levels were similar to those observed in *alg-3/4* males (Conine et al., 2010). Strikingly, the pre-mRNA levels of these negatively-regulated targets remained unchanged in *rde-3* or MAGO-12 males. Thus as previously hypothesized (Conine et al., 2010), ALG-3/4 negatively regulates a subset of its targets at a post-transcriptional level via the WAGO pathway.

To examine the temporal and spatial regulation of transcription in the male germline by ALG-3/4, we stained male germlines with an antibody that recognizes elongating Pol II (Ahn et al., 2004). In wild-type germlines at 25°C, elongating Pol II was detected at the periphery of condensing spermatocyte nuclei and in small discrete foci in spermatids (Figure 3.7C). By contrast, in *alg-3/4* germlines at 25°C, elongating Pol II disappeared from spermatocyte nuclei at the onset of the condensation process (Figure 3.7C), but was detected later in spermatogenesis throughout the nucleus in spermatocytes, not just at the

nuclear periphery (Figure 3.7C). We next examined a histone modification associated with actively transcribed chromatin (Histone H3 dimethylated on Lysine 4, H3K4me2, (Kelly et al., 2002). This analysis revealed H3K4me2 localization throughout the nucleus in wild-type spermatocytes at 25°C. However, H3K4me2 was dramatically reduced in *alg-3/4* spermatocytes and was present only in small patches of chromatin (Figure 3.7D).

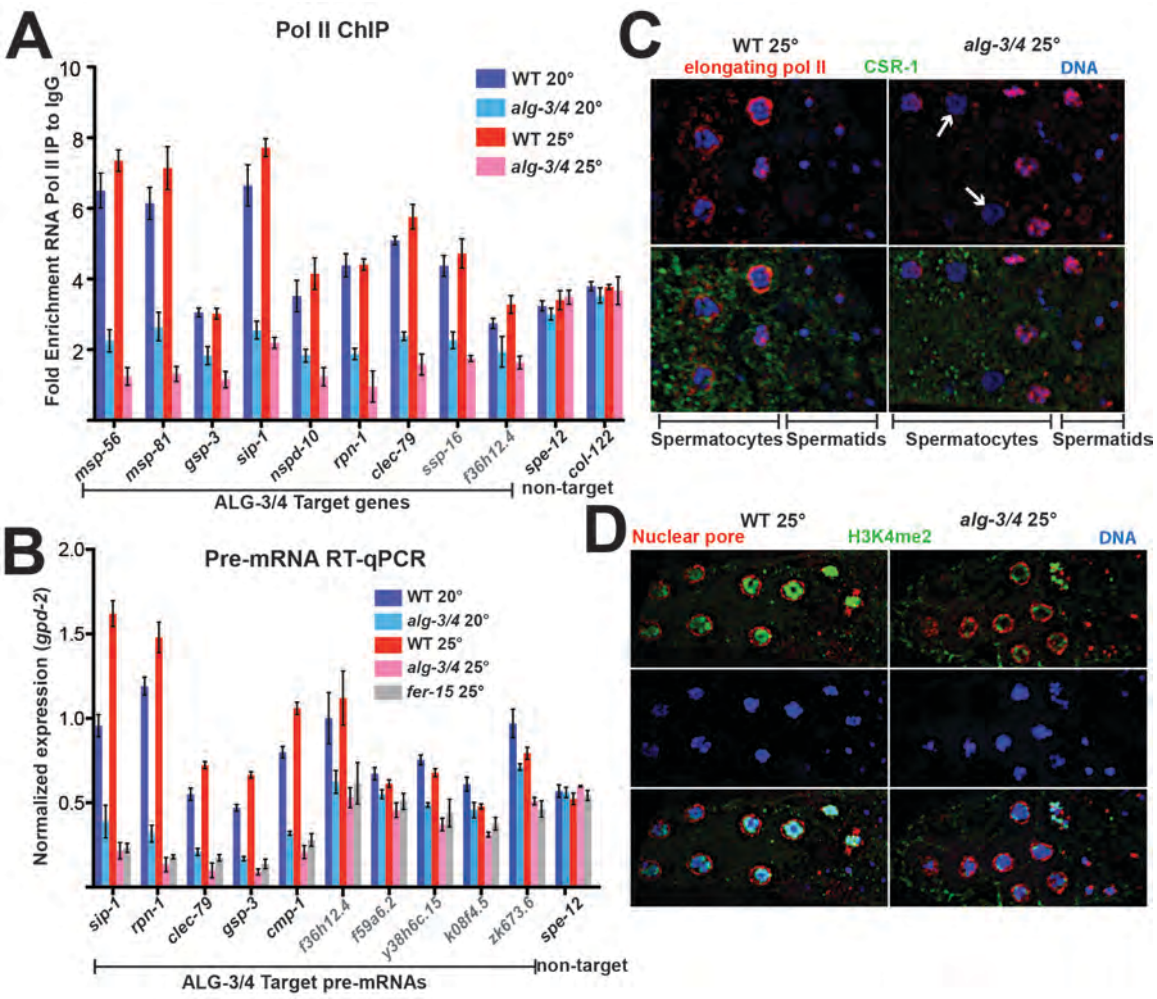

Figure 3.7. ALG-3/4 promote transcription and CSR-1 nuclear localization in condensing meiotic nuclei.

(A) qPCR analysis of RNA Pol II ChIP at ALG-3/4 target genes in wild-type (WT) and *alg-3/4* mutants at 20°C and 25°C, normalized to an intergenic region not occupied by pol II. Data are represented as mean +/- SEM. ALG-3/4 negatively-regulated targets are labeled in grey. **(B)** RT-qPCR analysis of ALG-3/4 target pre-mRNAs. Data are normalized to the non-target *gpd-2* pre-mRNA and represented as mean +/- SEM. **(C and D)** Confocal IF images of dissected WT (left) or *alg-3/4* (right) spermatogenic germlines at 25°C stained with antibodies against **(C)** elongating RNA Pol II (red) and CSR-1 (green), or **(D)** antibodies against H3K4me2 (green) and Nuclear pore protein (red). White arrows denote spermatocyte nuclei lacking pol II staining. DNA was stained with DAPI (blue).

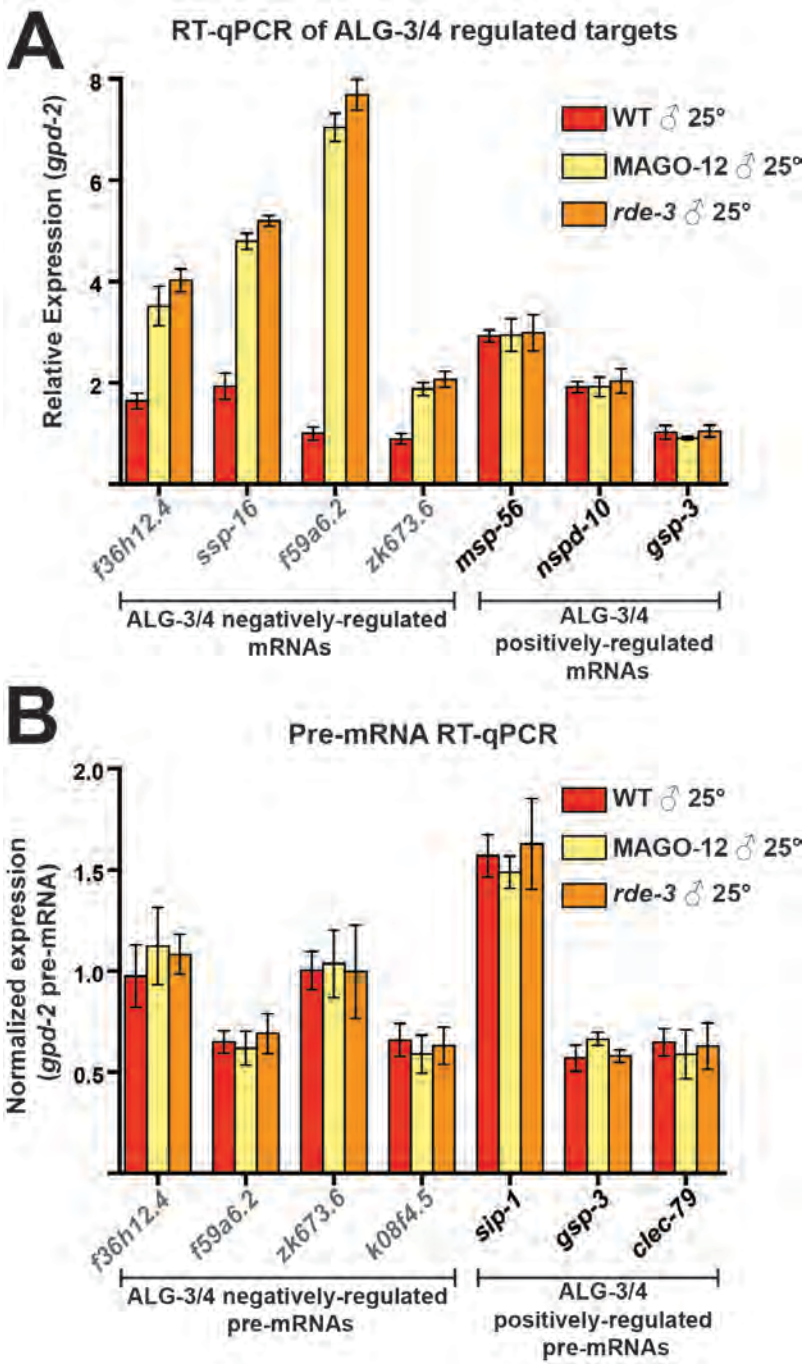

Figure 3.8. ALG-3/4 negatively-regulated targets are silenced by the WAGO small RNA pathway.

(A) RT-qPCR analysis of ALG-3/4 targets normalized to *gpd-2* in WAGO pathway mutants. **(B)** RT-qPCR analysis of ALG-3/4 target pre-mRNAs in WAGO pathway mutants. Values are normalized to the non-target *gpd-2* pre-mRNA**.** Data are represented as mean +/- SEM.

CSR-1 acts in the ALG-3/4 pathway to promote sperm development

Previous work suggested that the Argonaute CSR-1 and its 22G-RNA cofactors target but do not silence germline-expressed genes (Claycomb et al., 2009), and recent work suggests that CSR-1 promotes germline transgene expression (Seth et al., 2013). Homozygous *csr-1(tm892)* mutant hermaphrodites are essentially sterile, but produce a few embryos that die due to chromosome segregation defects (Yigit et al., 2006, Claycomb et al., 2009). Although *csr-1* males were sterile at 25°C, we were surprised to find that they were ~50% fertile relative to wild-type males at 20°C (Figure 3.9A). This temperature-dependent sterility was similar to that observed for *alg-3/4*. Indeed, like *alg-3/4* mutant spermatids, *csr-1* spermatids failed to complete spermiogenesis at 25°C and arrested as either round spermatids or spermatids with non-motile or spiky pseudopods (Figure 3.9A).

Consistent with the possibility that CSR-1 functions in the same pathway as ALG-3/4, we found that triple *csr-1 alg-3; alg-4* mutant males were completely sterile at 25°C with defects in spermiogenesis almost identical to those observed in the *csr-1* single, and *alg-3/4* double mutants respectively (Figure 3.9A; Conine et al., 2010). At 20°C *csr-1 alg-3; alg-4* males exhibited fertility defects similar to, but slightly more severe, than *csr-1* or *alg-3/4* males (Figure 3.9A). Furthermore, *csr-1* males grown at 25°C exhibited defects in elongating Pol II and H3K4me2 localization in the spermatogenic germline that were indistinguishable from phenotypes observed in *alg-3/4* males at 25°C (Figure 3.10A). Taken together

these findings suggest that CSR-1 functions along with ALG-3/4 to promote a chromosomal environment compatible with transcription. Perhaps consistent with this idea, we observed an increase in the number of spermatocytes undergoing nuclear condensation in both *alg-3/4* and *csr-1* males. The increase was moderate at 20°C (1.2- to 1.4-fold), but enhanced at 25°C, with both *alg-3/4* and *csr-1* mutants showing a nearly 2-fold increase in the number of condensing nuclei relative to wild-type (Figure 3.10B). These observations suggest that spermatocyte nuclei begin to condense prematurely, or fail to complete condensation appropriately, in *alg-3/4* and *csr-1* mutants.

In the hermaphrodite germline, CSR-1 localizes to perinuclear P-granules and to condensed chromatin in oocytes (Claycomb et al., 2009). In addition CSR-1 was identified in a proteomic study as a protein associated with sperm chromatin (Chu et al., 2006). To examine the expression pattern and localization of CSR-1 during male gametogenesis, we performed immunofluorescence using an antibody that recognizes endogenous CSR-1 (Claycomb et al., 2009). CSR-1 was associated with P-granules throughout the syncytial male germline and into differentiating spermatocytes, where P-granules disperse and disappear. In developing gametes, we found that CSR-1 localized to large cytoplasmic foci and in discrete chromatin domains of spermatocytes undergoing nuclear condensation as well as in haploid spermatids (Figure 3.9B). In *alg-3/4* mutants cultured at 25°C, CSR-1 was present in large cytoplasmic foci in spermatocytes undergoing nuclear condensation but was largely absent from chromatin.

Notably, in haploid spermatids from *alg-3/4* males we observed fewer CSR-1 cytoplasmic and nuclear foci than in haploid spermatids from wild-type males (Figure 3.9B). Taken together, these findings suggest that CSR-1 functions with ALG-3/4 to promote gene expression in developing spermatocytes.

CSR-1 22G-RNAs target genes also targeted by ALG-3/4 26G-RNAs

The similarities between the cellular and developmental phenotypes of *alg-3/4* and *csr-1* males are consistent with the possibility that ALG-3/4 and CSR-1 function in the same small RNA pathway. To determine whether CSR-1 is required for the expression of small RNAs antisense to ALG-3/4 targets, we cloned and deep sequenced small RNAs from *csr-1* males grown at 25°C. Remarkably, we found that ALG-3/4-dependent 26G-RNAs – regardless of class, i.e., antisense to positively- or negatively-regulated targets – were strongly depleted in *csr-1* males. Indeed, this depletion was to a level resembling that of *alg-3/4* mutant males (Figure 3.9C). On the other hand, 22G-RNAs were only slightly depleted in *alg-3/4* and *csr-1* males (Figure 3.9C). A similar minor reduction of 22G-RNAs was observed previously when small RNAs were cloned from *csr-1* hermaphrodites (Claycomb et al., 2009). While these findings are consistent with the idea that ALG-3/4 and CSR-1 function in the same pathway, they also suggest that the pathway may not be linear (see Discussion).

To further explore the idea that CSR-1 and ALG-3/4 share targets, we cloned and deep sequenced CSR-1–associated 22G-RNAs. To do this we

performed CSR-1 Immunoprecipitation (IP) followed by small-RNA deep-sequencing from male-enriched (>95%) populations expressing a rescuing FLAG::CSR-1 transgene and grown at 25°C. We found that CSR-1 IP enriched small RNAs targeting 5575 genes, including 90% (3775/4190) of the genes previously identified as CSR-1 targets in the hermaphrodite (Claycomb et al., 2009). As expected, CSR-1 IP did not enrich miRNAs or 21U-RNAs, nor did it enrich 22G-RNAs previously identified as WAGO-1-associated in hermaphrodites (Claycomb et al., 2009; Gu et al., 2009). We also recovered small RNAs that target ~1800 genes not previously identified as CSR-1 targets, including most annotated male-germline-expressed genes. Notably, using a 2-fold cutoff, FLAG::CSR-1 IP enriched for small RNAs antisense to 82% (1156/1408) of all ALG-3/4 targets defined in this study (Figure 3.9D), including 88% (188/214) of ALG-3/4 positively-regulated genes, 83% (821/991) of non-regulated genes and 72% (147/204) of ALG-3/4 negatively-regulated genes (Figure 3.9D). Thus, just as CSR-1 is required for all ALG-3/4 26G-RNAs, these findings indicate that CSR-1 associates with 22G-RNAs targeting most and perhaps all ALG-3/4 targets.

Surprisingly, while analyzing our CSR-1 IP data, we found that the FLAG::CSR-1 IP from males enriched small RNAs that target female-specific genes (Reinke et al., 2004). Remarkably, these oogenesis-specific small RNAs were present at levels similar to those of small RNAs targeting male-specific genes (Figure 3.9D). This finding is striking given that the corresponding female-

specific mRNAs were at least 13-fold less abundant than male-specific mRNAs in our mRNA-seq data (Figure 3.9D). These results suggest that males might inherit maternally-derived small RNAs that target female-specific genes and actively transmit them to their offspring (see Discussion).

CSR-1 promotes the expression of ALG-3/4 targets

CSR-1 was previously shown to associate with the chromatin of its 22G-RNA target genes in hermaphrodites (Claycomb et al., 2009). To ask whether CSR-1 associates with the chromatin of genes targeted by ALG-3/4 in males, we performed CSR-1 ChIP on WT and *alg-3/4* male populations grown at 25°C. We found that in WT males CSR-1 ChIP enriched all of the ALG-3/4 targets assayed, including both positively and negatively regulated targets, by 1.5- to 3-fold relative to a no-antibody control (Figure 3.9E). This enrichment was not detected in *alg-3/4* mutant animals (Figure 3.9E). CSR-1 ChIP from males also enriched several genes that were previously identified as CSR-1 targets in hermaphrodites, including *daf-21, cgh-1* and *oma-1* (Claycomb et al., 2009). The latter two, *cgh-1* and *oma-1,* were only weakly expressed in males (Data Not Shown) and were not ALG-3/4 targets. As expected, we found that the association of CSR-1 with these two loci was independent of ALG-3/4 (Figure 3.9E). CSR-1 ChIP did not enrich *spe-12* or *f46f5.5*, genes that are not targeted by ALG-3/4– or CSR-1–associated small RNAs (Figure 3.9E). Using RT-qPCR we found that the mRNA and pre-mRNA levels of positively-regulated ALG-3/4

targets were reduced by similar amounts in *csr-1* males (see Figure 3.11C below). By contrast, we found that mRNA levels of ALG-3/4 negatively-regulated targets were unchanged in *csr-1* mutant males at 25°C (Figure 3.5), while pre-mRNA levels decreased (see Figure 3.11C below). This latter finding – that pre-mRNA levels decrease for ALG-3/4 negatively-regulated targets – was also observed, paradoxically, in *alg-3/4* mutant males (Figure 3.7B). Taken together these findings indicate that targets positively regulated by ALG-3/4 are also positively regulated by CSR-1, and suggest that ALG-3/4 and CSR-1 promote the expression of their targets at a transcriptional level, including a subset of ALG-3/4 targets whose net expression is negatively regulated due to post-transcriptional silencing.

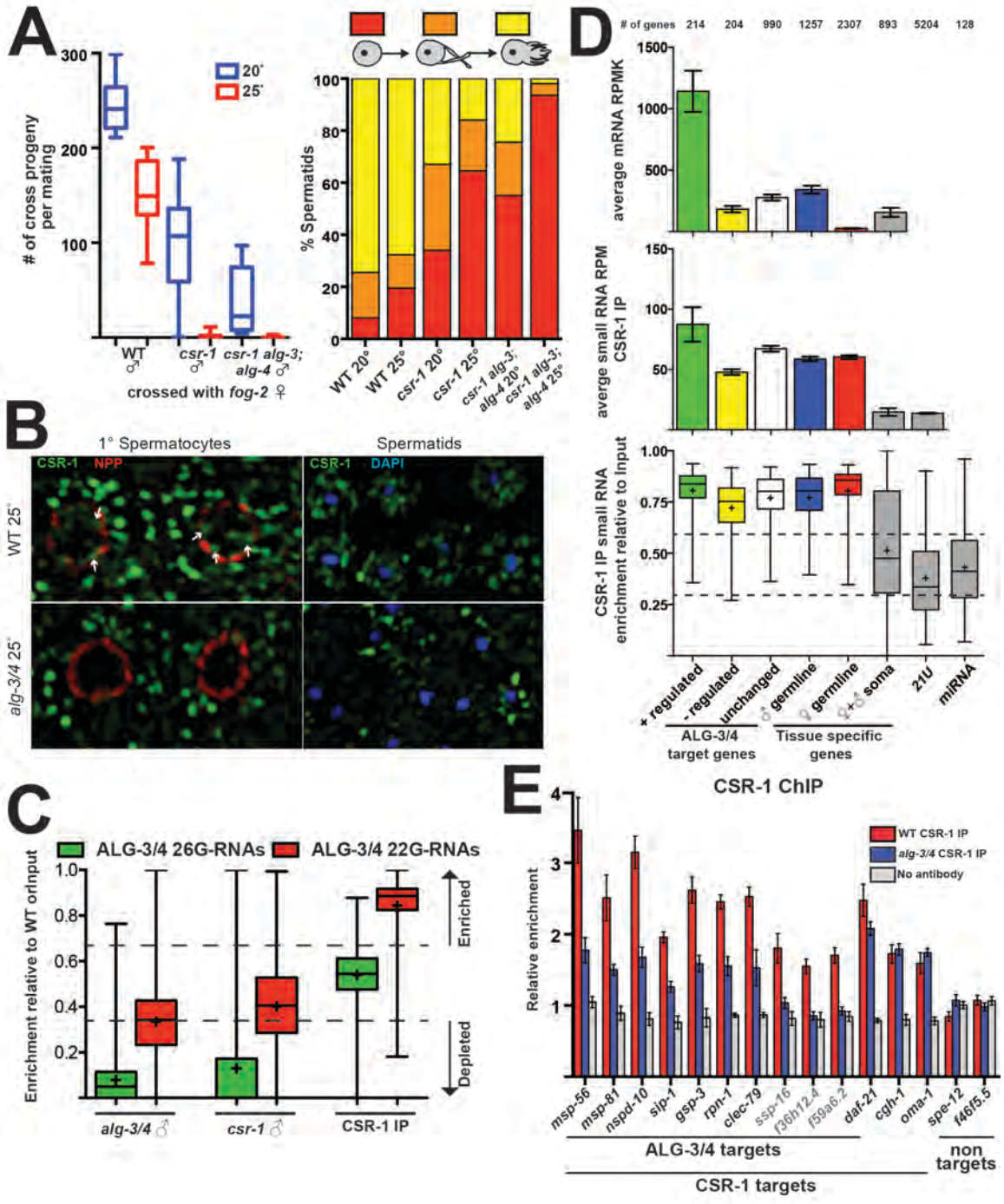
A
# of cross progeny per mating
20°
25°
WT ♂
csr-1 ♂
csr-1 alg-3; alg-4 ♂
crossed with fog-2 ♀
% Spermatids
WT 20°
WT 25°
csr-1 20°
csr-1 25°
csr-1 alg-3; alg-4 20°
csr-1 alg-3; alg-4 25°
B
1° Spermatocytes
Spermatids
CSR-1 NPP
CSR-1 DAPI
WT 25°
alg-3/4 25°
C
ALG-3/4 26G-RNAs
ALG-3/4 22G-RNAs
Enrichment relative to WT orInput
Enriched
Depleted
alg-3/4 ♂
csr-1 ♂
CSR-1 IP
D
# of genes 214 204 990 1257 2307 893 5204 128
average mRNA RPMK
averge small RNA RPM CSR-1 IP
CSR-1 IP small RNA enrichment relative to Input
+ regulated
- regulated
unchanged
♂ germline
♀ germline
♀+♂ soma
21U
miRNA
ALG-3/4 target genes
Tissue specific genes
CSR-1 ChIP
E
Relative enrichment
WT CSR-1 IP
alg-3/4 CSR-1 IP
No antibody
msp-56
msp-81
nspd-10
sip-1
gsp-3
rpn-1
clec-79
ssp-16
f36h12.4
f59a6.2
daf-21
cgh-1
oma-1
spe-12
f46f5.5
ALG-3/4 targets
non targets
CSR-1 targets

Figure 3.9. CSR-1 associates with both male- and female-specific small RNAs in males and positively regulates spermiogenic gene expression.
(A) (Left) Fertility of WT, *csr-1*, or *csr-1 alg-3; alg-4* males at 20°C or 25°C measured in crosses with *fog-2* females. Box and whisker plots represent a range of cross progeny from at least 15 independent crosses for each genotype. (Right) *In vitro* activation of spermatids isolated from *csr-1*, *csr-1 alg-3; alg-4,* or WT males grown at 20°C or 25°C. Bar graphs show the percent of spermatids that fail to activate (red), partially activate and form spikey projections (orange), or fully activate and form a pseudopod (yellow) (n>300). **(B)** Confocal IF images of primary (1°) spermatocytes (left) or spermatids (right) in dissected *alg-3/4* or WT germlines stained with antibodies against CSR-1 (green) and a nuclear pore protein (red). DNA was stained with DAPI (blue). White arrows denote chromatin domains where CSR-1 localizes. **(C)** Box and whisker plots indicate depletion of 26G-RNAs and 22G-RNAs antisense to ALG-3/4 targets in *alg-3/4* and *csr-1* males at 25°C. Box and whisker plots to the right indicate enrichment of 22G-RNA and lack of enrichment of 26G-RNAs antisense to ALG-3/4 targets in the CSR-1 IP. **(D)** Analysis of mRNA expression and CSR-1 small RNA levels in wild-type males. Categories of gene expression are indicated below the bottom graph. ALG-3/4 targets as defined in Figure 3.3: positively-regulated (green), negatively-regulated (yellow), and unchanged (White). Male-germline specific genes (blue), includes many ALG-3/4 targets. Female-germline specific genes (red) – not targeted by ALG-3/4. Soma-specific mRNAs and 21U-RNAs are control sequences not enriched by CSR-1 IP. The number of genes in each category is indicated above the graphs. Top graph: mRNA expression as monitored by RNA-Seq in reads per million per kilobase (RPMK). Middle graph: small RNA levels in reads per million (RPM) per gene. Bottom graph: box and whisker plot indicating the enrichment or depletion in CSR-1 IP of small RNAs antisense to genes. Enrichment was calculated as the RPM ratio of FLAG::CSR-1 IP / (FLAG::CSR-1 IP + Input). Dotted lines indicate 2-fold enrichment (upper) or depletion (lower). '+' indicates the average enrichment value for all genes in the category. **(E)** qPCR analysis of FLAG::CSR-1 ChIP at genes targeted by ALG-3/4, CSR-1, or neither. ALG-3/4 negatively-regulated targets are indicated in grey text. Data were normalized to *y47h10a.3*, which is not targeted by ALG-3/4 or CSR-1. Data are represented as mean +/- SEM.

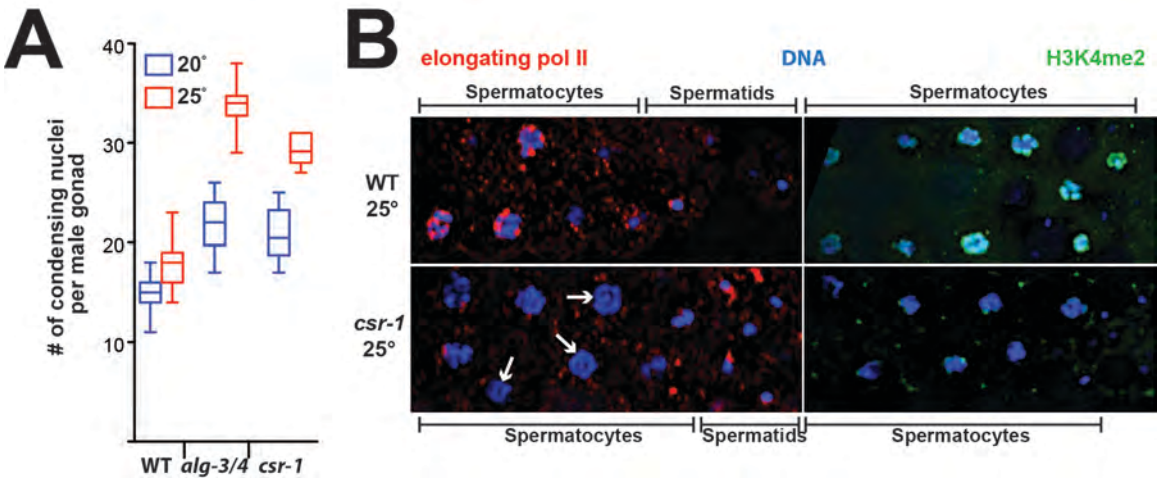

Figure 3.10. ALG-3/4/CSR-1 pathway mutants exhibit defects in spermatogenic chromatin condensation and transcription.
(A) *alg-3/4* and *csr*-1 mutants exhibit similar defects in spermatogenic nuclear condensation. Comparison of the number of condensing nuclei present in the meiotic germline of WT, *alg-3/4* and *csr-1* males at 20°C or 25°C. Box and whisker plots indicate the range of data from at least twelve DAPI-stained confocal image stacks per strain and condition. **(B)** Confocal IF images of dissected WT (top) or *csr-1* (bottom) spermatogenic germlines at 25°C stained with antibodies against (left) elongating RNA Pol II (red), spermatocyte nuclei completely missing Pol II staining are denoted with a white arrow, or (right) antibodies against H3K4me2 (green). DNA was stained with DAPI (blue).

CSR-1 and ALG-3/4 provide a paternal memory of past gene expression

The ALG-3/4 proteins are present during spermatogenesis but are eliminated from maturing spermatids (Conine et al., 2010). The CSR-1 protein, however, is abundant in mature sperm (Figure 3.9B). We therefore wondered if ALG-3/4 and CSR-1 might function together to pass a memory of male-specific gene expression from one generation to the next via CSR-1. To test this idea, we first analyzed the fertility of *alg-3/4* and *csr-1* heterozygous hermaphrodites (F_1) cultured at 25°C. We then mated heterozygous hermaphrodites to homozygous *alg-3/4* or *csr-1* males, respectively, at the permissive temperature to obtain F_2 heterozygous hermaphrodites. In this and subsequent generations, the resulting heterozygous offspring were cultured at 25°C to assay their fertility before being mated to homozygous males at 20°C (Figure 3.11A). In parallel, as a control, we mated heterozygous hermaphrodites to heterozygous males to obtain heterozygous offspring (Figure 3.11A). Thus, the maternal genotype was always heterozygous throughout this analysis, whereas the paternal genotype was either homozygous (experimental series) or heterozygous (control series). Remarkably, we found that heterozygous offspring (hermaphrodites or males) derived from homozygous *alg-3/4* or *csr-1* males became progressively less fertile with each generation at 25°C (Figure 3.11B). By the sixth generation (F_6), the heterozygous offspring of *alg-3/4* or *csr-1* homozygous males were completely sterile. The sterility of F_6 heterozygous hermaphrodites could be rescued by mating to wild-type males, indicating that the sterility was due to a sperm defect. In the control

series, on the other hand, the heterozygous offspring of heterozygous males maintained a consistent level of fertility throughout the course of the experiment. Importantly, the impaired fertility of heterozygous *alg-3/4* and *csr-1* offspring correlated with declining pre-mRNA expression levels of ALG-3/4 targets. By the F_5 generation, relative to either control heterozygous males or wild-type males, the heterozygous male offspring of homozygous males exhibited a reduction in pre-mRNA similar to *alg-3/4* or *csr-1* males at 25°C (Figure 3.11C). Thus, ALG-3/4 and CSR-1 promote a paternal epigenetic memory of ALG-3/4 target gene expression.

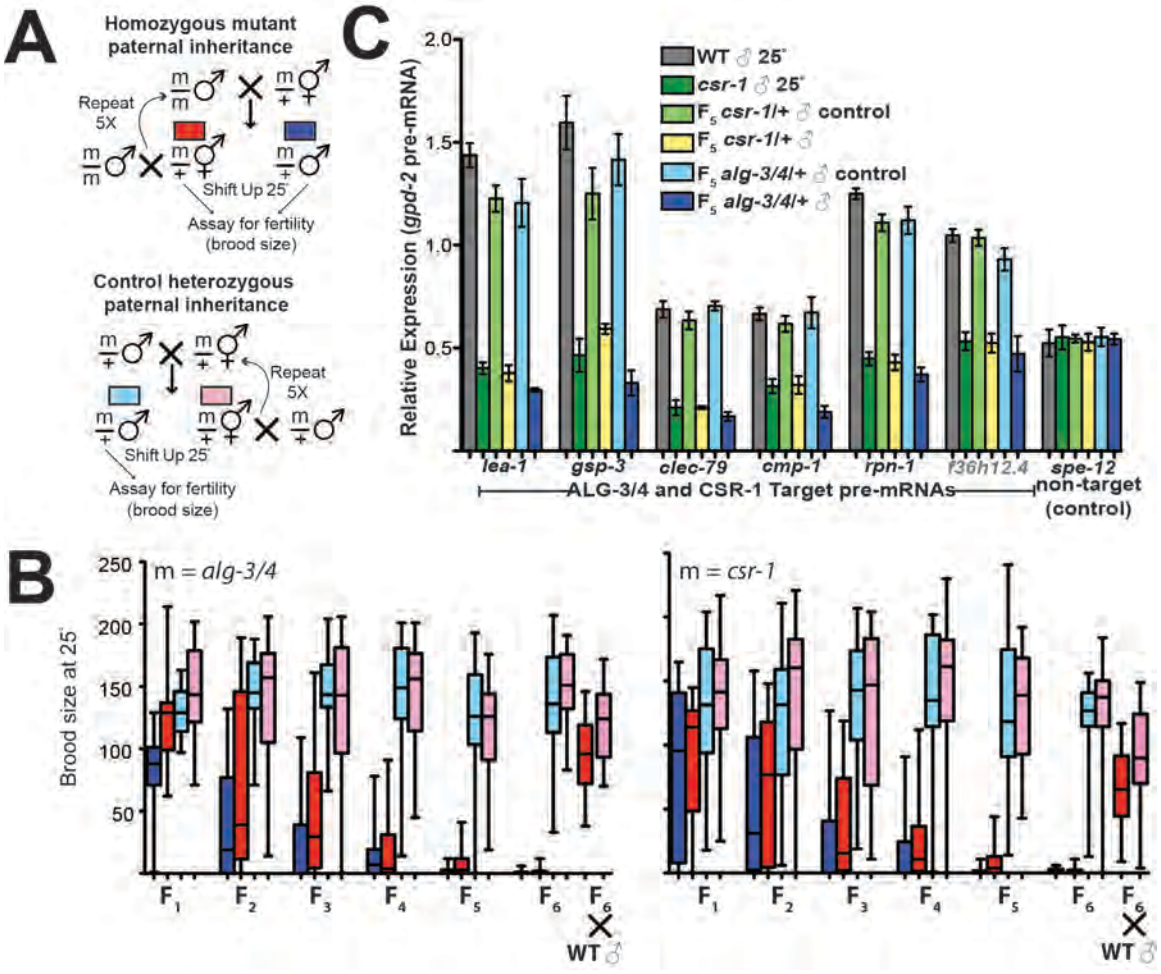

Figure 3.11. ALG-3/4 and CSR-1 provide a paternal memory of germline gene expression.

(A) Schematic of crosses to assay paternal inheritance of gene expression and sperm function. 'm' indicates either *csr-1* or *alg-3/4* allele. **(B)** Repeated mating with *alg-3/4* and *csr-1* males induces a progressive dominant germline-mortal phenotype. Box and whisker plots indicate the brood sizes of heterozygous hermaphrodite (red/pink) and male (blue/aqua) cross progeny of homozygous mutant fathers (red and blue) or control heterozygous hermaphrodites and males (pink and aqua, respectively) determined in successive generations of mating to homozygous (blue) or heterozygous (aqua) *alg-3/4* (left panel) or *csr-1* (right panel) mutant males. Fertility of male cross progeny was assayed by mating to *fog-2* females. **(C)** RT-qPCR analysis of pre-mRNA levels in male cross progeny (as indicated by color). Data are normalized to *gpd-2* pre-mRNA and represented as mean +/- SEM. CSR-1 and ALG-3/4 target mRNAs, and a non-target mRNA were assayed (as indicated). An ALG-3/4 negatively-regulated target is indicated in grey text.

DISCUSSION

A small RNA feed-forward loop transmits a paternal epigenetic memory of past gene expression

During male gametogenesis germ cells proceed through meiosis and undergo dramatic changes in cellular morphology to produce haploid spermatids containing highly compacted, transcriptionally inert chromatin (Ward et al., 1981). The completion of this process and the subsequent transformation of spermatids into polarized motile spermatozoa capable of fertilization depends on the proper execution of an extensive gene-expression program involving thousands of genes. Here, we have shown that the Argonautes ALG-3/4 and CSR-1 are required during this process to promote robust spermatogenic gene expression. Although ALG-3/4 is absent from mature sperm (Conine et al., 2010), we have shown that CSR-1 is abundant in mature sperm. The propagation of strains lacking CSR-1 or ALG-3/4 activities in the paternal lineage caused a progressive loss of fertility (a germline-mortal phenotype), in which even heterozygous descendants, with a wild-type copy of the respective locus, exhibited complete sperm-specific sterility when assayed at 25°C. The observed infertility involved an arrest as round spermatids with decreased transcription of ALG-3/4 targets, a phenotype identical to that observed in *alg-3/4* and *csr-1* homozygous mutants. These findings suggest that ALG-3/4 and CSR-1 are not only required to promote spermatogenic gene expression but also act together to transmit an epigenetic memory of paternal gene expression via the sperm.

How might this work? Paternal CSR-1 22G-RNAs delivered via the sperm could enter the zygotic germline (See Model, Figure 3.12). Later, when spermatogenesis initiates in hermaphrodites and males, CSR-1 targeting could recruit the RdRP-containing ERI complex (Duchaine et al., 2006; Pavelec et al., 2009) to initiate the production of 26G-RNAs that are loaded onto ALG-3/4. In a feed-forward mode, ALG-3/4 could then target cognate transcripts to recruit the EGO-1 RdRP complex to re-amplify CSR-1 22G-RNAs (Claycomb et al., 2009; Conine et al., 2010; Gu et al., 2009). While the initial biogenesis of ALG-3/4 and CSR-1 small RNAs likely requires some template mRNA destruction, the net result of this amplification cycle appears to be increased mRNA levels, perhaps due to feedback on transcription. During spermatogenesis, CSR-1 and its 22G-RNA cofactors might promote gene expression by engaging nascent transcripts on the chromatin of its target genes. For example, CSR-1 could recruit factors that help maintain a transcriptionally active state during spermatogenesis and nuclear condensation, ensuring that spermatids obtain the appropriate level of gene products required for spermiogenesis. Finally, CSR-1 small-RNA complexes could once again become incorporated into mature sperm, thus poised to reinitiate the cycle in the next generation (Figure 3.12).

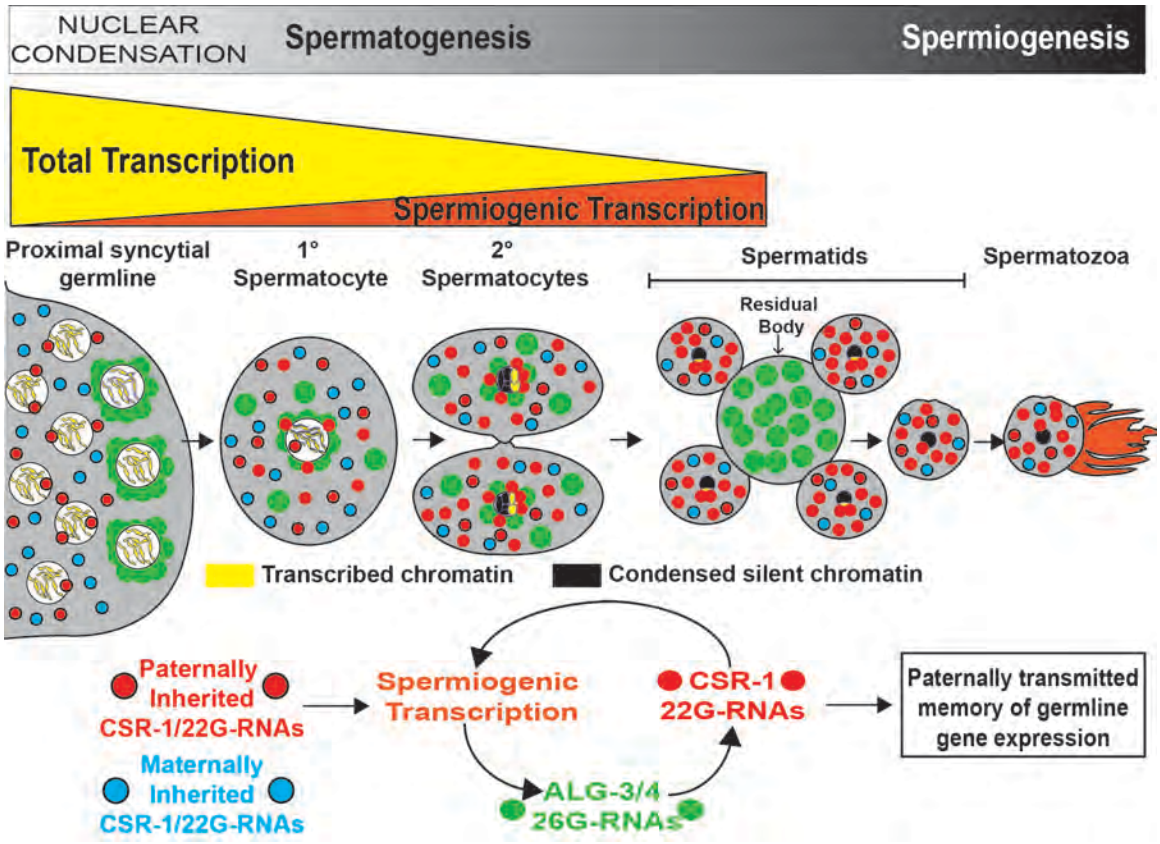

Figure 3.12. Model for ALG-3/4 and CSR-1 function in the male germline.
ALG-3/4 and CSR-1 and their small RNA cofactors promote the transcription of spermiogenic genes and provide a memory of past germline gene expression. Black dots in spermatids represent areas of condensed silent chromatin; yellow areas indicate transcriptionally-active chromatin.

A protective role for piRNAs across phyla

In many animals, sperm development is inherently temperature sensitive (Rockett et al., 2001). The effects of elevated temperature during sperm development can also be epigenetically transmitted, causing developmental defects in mammalian embryos, including decreased embryonic mass and increased mortality (Jannes et al., 1998; Setchell et al., 1988). However, the molecular mechanism behind the temperature sensitivity of sperm development remains unclear.

Our findings indicate that the gene expression programs required for spermatogenesis in *C. elegans* are sensitive to temperature. In wild-type males, the expression of many spermiogenesis genes was higher at 25°C than at 20°C. We have shown that ALG-3/4 and CSR-1 are required to initiate and maintain the activation of spermiogenesis genes at elevated temperatures. Failure to maintain spermiogenic gene expression in *alg-3/4* and *csr-1* mutants correlates with dramatic defects in spermatid activation and infertility in both mutant strains at elevated temperature. Thus, the ALG-3/4 and CSR-1 pathways appear to act as enhancers of gene expression that buffer the effects of temperature on sperm development.

In wild-type animals, the process of spermatogenesis accelerates by greater than 25% at 25°C relative to 20°C, which in turn demands a similar increase in the expression of gene products that support spermiogenesis. Thus, increased gene expression is required in a setting where the rapid onset of

meiosis and subsequent chromatin condensation would be expected to shut down transcription. Our findings suggest that the ALG-3/4 and CSR-1 pathways act together to selectively maintain transcriptionally-active chromatin at spermiogenesis genes during meiotic nuclear condensation, while packaging other regions of the genome not essential for spermiogenesis into transcriptionally-inactive chromatin. This would create a burst of transcription of spermiogenesis genes near the end of spermatogenesis, as seen in mammals (Sassone-Corsi, 2002), when chromatin begins to condense in 1° spermatocytes (Shakes et al., 2009).

Interestingly, mice deficient for the Piwi homolog MIWI display spermatogenic arrest at the round spermatid stage (Deng and Lin, 2002), a phenotype similar to that of *alg-3/4* and *csr-1* mutant sperm. MIWI associates with pachytene piRNAs, ~29-31 nt small RNAs derived from large piRNA genes within non-repetitive genomic regions (Li et al., 2013). Much like ALG-3/4 26G-RNAs, mouse pachytene piRNAs are expressed specifically in developing spermatocytes upon entering the pachytene stage of meiotic prophase I (Girard et al., 2006). Intriguingly, MIWI was shown to associate with the translational machinery as well as with polysomes during early spermiogenesis, leading to speculation that it promotes translation (Lau, 2010). Recent work also suggests that MIWI promotes the stabilization of spermiogenic mRNAs (Nishibu et al., 2012; Vourekas et al., 2012). Interestingly, a protective role for small RNAs and Piwi Argonaute proteins was recently identified in the ciliated protozoan

Oxytricha, where piRNAs prevent DNA elimination during genome rearrangement (Fang et al., 2012). These findings raise a thought-provoking possibility that pachytene piRNAs and ALG-3/4/CSR-1 small RNAs provide analogous functions, protecting the genome and buffering gene expression in protists, nematodes, and mammals.

Whole-genome surveillance by Argonaute/small-RNA pathways

While ALG-3/4 and CSR-1 function together to promote the expression of many genes required for spermiogenesis, ALG-3/4 also functions independently of CSR-1 to negatively regulate a subset of its targets via the WAGO genomic surveillance silencing pathway. Although CSR-1 22G-RNAs target these same mRNAs, CSR-1 is not required for their silencing (Figure 3.5). Indeed, perhaps paradoxically, both CSR-1 and ALG-3/4 appear to moderately promote the transcription of these ALG-3/4 negatively-regulated targets. These observations suggest that in males ALG-3/4 and CSR-1 promote the transcription of some targets that are silenced post-transcriptionally via the WAGO/22G-RNA pathway. This complexity might reflect distinct positive and negative regulation of the same target genes by CSR-1 and WAGO at different times during spermatogenesis. However, it is also possible that transcriptional up-regulation of these targets is required to ensure that sufficient template RNA is produced to re-amplify WAGO-22G-RNAs important for post-transcriptional silencing in the next generation. In either case, these findings hint at additional complexity.

Surprisingly, males also contain abundant CSR-1 22G-RNAs targeting female-specific transcripts. We do not know whether these 22G-RNAs are maternal CSR-1/22G-RNA complexes that persist in the male, or alternatively, if they are generated de novo in males, perhaps using transcripts maternally inherited or produced at low levels. Regardless of their origin, this result could indicate that sperm transmit a memory of both paternal and grand-maternal gene expression. CSR-1 is also abundant in oocytes and it seems likely that it could provide similar functions there, delivering epigenetic signals from the mother and possibly the grandfather. Given the relatively much smaller volume of sperm, it is possible that the ALG-3/4 system exists, in part, to amplify the sperm-specific CSR-1 signal, ensuring both the robust expression of sperm genes and the transmission of a memory of sperm-specific gene expression to offspring.

Like CSR-1-dependent small-RNA signals, WAGO-dependent signals are also transmitted via both the egg and the sperm (Ashe et al., 2012; Buckley et al., 2012; Gu et al., 2009; Shirayama et al., 2012). Our analysis of CSR-1 and a previous study of HRDE-1/WAGO-9 (Buckley et al., 2012) suggest that both Argonaute pathways promote germline immortality. Interestingly, when either small RNA pathway is lost, fertility begins to decline gradually over a few generations. These findings suggest that Argonautes act to reinforce and maintain parallel transgenerational epigenetic signals that might, for example, be chromatin mediated. In the respective Argonaute mutants, loss of the small RNA signals may cause a gradual loss, over a period of several generations, of

chromatin marks associated with active or silent gene-expression states, resulting in the observed gradual onset of germline mortality.

Recent work on transgenes expressed in the *C. elegans* germline (Seth et al., 2013) has identified a CSR-1–dependent transgenerational activating signal (RNAa). This activating signal, which can be transmitted via the sperm, can act within a single generation to reverse a persistent mode of epigenetic silencing referred to as RNA-induced epigenetic gene silencing (RNAe) (Shirayama et al., 2012). Interestingly, while de-silencing was observed immediately after exposure to RNAa, continuous exposure, over many generations, was necessary to render a formerly silent gene capable of durable independent expression (Seth et al., 2013). The maintenance of RNAe requires members of the WAGO clade of Argonautes, as well as repressive heterochromatin factors and histone modifications (Ashe et al., 2012; Buckley et al., 2012; Shirayama et al., 2012). Perhaps germline mortality occurs gradually, after loss of CSR-1 and ALG-3/4 activity, as silencing marks spread into and gradually silence genes required for spermatogenesis. The findings reported here support the idea that CSR-1 transmits a protective small-RNA-induced trans-activating signal, and provide physiological evidence linking CSR-1 and RNAa more globally to the transmission and maintenance of paternal, and possibly maternal epigenetic memory.

EXPERIMENTAL PROCEDURES

Worm strains and genetics

C. elegans culture and genetics were performed as described (Brenner, 1974). Unless otherwise noted, the “wild-type” (WT) strain in this study is the Bristol N2 strain carrying the *fog-2(q71)* allele. Alleles listed by chromosome: LGI: *fer-1(hc24), fer-1(b232), fer-6(hc23), fer-7(hc34);* LGII: *neSi1[cb-unc-119(+) 3xflag::csr-1], eri-3(tm1361), fer-3(hc3), rrf-3(pk1426), fer-15(b26)*; LGIII: *alg-4(ok1041), unc-119(ed3), fer-2(hc2)*; LGIV: *alg-3(tm1155)*, *csr-1(tm892), DnT1[unc(n754dm) let](IV;V);* LGV: *fog-2(q71), fer-4(hc4)*. The *3xflag::csr-1* transgenic strain was generated by Mos-mediated single-copy insertion (Frokjaer-Jensen et al., 2008) and details are provided in Extended Experimental Procedures. Analysis of Eri phenotypes, male fertility and spermatid activation.

Eri, fertility and spermatid activation assays.

Eri phenotypes were assayed by feeding worms with bacteria expressing *unc-73* dsRNA and scoring the F1 progeny for the Unc-73 phenotype (Duchaine et al., 2006; Kennedy et al., 2004). Male fertility was assayed by mating males to virgin *fog-2* females and determining the number of viable cross progeny as described (Batista et al., 2008). Spermatid activation was performed by isolated spermatids from males and incubating with 200 µg/ml pronase as described (Shakes and Ward, 1989).

Molecular Biology

The *flag::csr-1* transgene was generated by Mos-mediated single copy insertion (Frokjaer-Jensen et al., 2008). A 7.0 kb SpeI – BstZ171 fragment from cosmid F20D12 containing the entire *csr-1* gene was inserted into a modified version of pCFJ151 (B1496) for insertion into LGII (Frokjaer-Jensen et al., 2008). The *3xflag* sequence was inserted into a SmaI site created by site-direct mutagenesis immediately after the initiation codon in the second exon of the *csr-1* gene. The *3xflag::csr-1* vector was present at 10 ng/ml in the MosSCI injection mixture. Expression of FLAG::CSR-1 was confirmed by western blot. The single copy *3xflag::csr-1* transgene inserted in LGII fully rescued the *csr-1(tm892)* null mutant.

Males were enriched to >95% homogeneity by filtering through a 35 micron mesh filter (Miller, 2006). Worms were homogenized in a stainless steel dounce in the presence of TRI Reagent (MRC Inc) for RNA isolation or IP buffer (Gu et al., 2009) for IP. RNA was extracted in TRI Reagent (MRC Inc) according to the manufacturer's specifications. The FLAG::CSR-1 IP was performed as described (Gu et al., 2009; Shirayama et al., 2012) using M2 FLAG antibody (Invitrogen). Small RNA enrichment and northern blots were performed as described (Conine et al., 2010). For cloning and deep sequencing, small RNAs were pretreated with Tobacco Acid Pyrophosphatase (Epicenter Biotechnologies) and ligated to an adenylated 3′ linker and then to a 5′ linker containing a 4 nt barcode as described (Gu et al., 2009). Illumina adapters were added by PCR.

For mRNA sequencing, polyA RNA was purified from total RNA using the PolyATtract mRNA Isolation System (Promega); each sample was purified twice. The mRNA was fragmented by base hydrolysis at 95°C for 9.5 min in Sodium carbonate buffer, and 2′,3′-cyclic phosphates were resolved with T4 Polynucleotide Kinase (NEB). mRNA fragments with an average length of ~200 nt were purified from an 8% polyacrylamide/8M urea gel, ligated to a 3′ linker, and reverse transcribed (Superscript III, Invitrogen) using a primer targeting the 3′ linker and containing reverse complement 5′ linker sequence with barcodes for multiplex sequencing. The cDNA was circularized using Circligase (epicentre).

Illumina adapters were added by PCR and cDNA libraries were sequenced by the UMass Deep Sequencing Core using an Illumina GAII or HiSeq. Deep-sequencing data analyses were performed as described (Gu et al., 2012). RPMK analysis was performed on the mRNA-seq libraries using a custom Perl script

Quantitative PCR and Chromatin IP

Reverse transcription followed by qPCR (RT-qPCR) was performed as described (Batista et al., 2008). To measure pre-mRNA, one of the RT-qPCR primers was placed within intron sequence. Primer sequences are available upon request.

For Chromatin Immunoprecipitation (ChIP), male worms were enriched to >95%, crosslinked in 2% paraformaldehyde for 30 min at 20°C, washed and

worm pellets were flash frozen. Worm pellets were resuspended in 4 volumes of FA buffer (Cold Spring Harbor Protocols) containing protease inhibitors (Roche) and sonicated at medium intensity in a Biorupter (Diogenode) for 50 min (30 seconds on, 30 seconds off). Insoluble debris was pelleted by centrifugation at 13,000 x g for 15 min. For Pol II ChIP, 6µg of mouse monoclonal anti-Pol II RNA pol II (8WG16 from Covance) was added to 5 mg of protein lysate and incubated for 2 hr at 4°C. For CSR-1 ChIP, 10 µg of anti-CSR-1 antibody (Claycomb et al., 2009) was added to 10-15 mg of protein lysate and incubated for 2 hr at 4°C. Immune complexes were captured with protein A/G agarose beads (Santacruz) at 4°C for 2 hr and washed, and proteins were eluted in TE containing 1% SDS and 250mM NaCl at 65°C for 15 min in a thermomixer (800-1000 RPM). As an IgG negative control, 6 ug of normal mouse IgG (SantaCruz) was also added to 5 mg of protein lysate, and processed the same as the experimental samples. Crosslinks were reversed by incubating overnight in TE/1% SDS at 65°C, and proteins were digested with proteinase K. DNA was extracted with phenol-chloroform and precipitated with ethanol. Enrichment of target loci was assayed by qPCR relative to the negative control IgG ChIP and a control intergenic region.

Fluorescence microscopy

Male germlines were dissected on poly-lysine slides in PBS, fixed in methanol and formaldehyde and washed with PBT (PBS, 0.1% BSA, and 0.1%

Triton-X 100). For immunofluorescence, slides were incubated with primary antibodies in PBT overnight at 4°C, washed with PBT, and incubated with secondary antibody in PBT for 1-2 hrs at room temperature. Slides were washed and mounted with Vectashield (Vector) containing 1 µg/ml DAPI. The following Primaries antibodies were used in this study: anti-MSP (4A5 from DSHB), anti-GSP-3 (kind gift from Dr. Diana Chu (Wu et al., 2012)), anti-elongating RNA pol II (H5 from Covance), anti-CSR-1(Claycomb et al., 2009), anti-nuclear pore (MAB414 from Covance), and anti-H3K4me2 (07-030 Millipore). Secondary antibodies were obtained from Jackson Immunoresearch or Molecular Probes. For mRNA FISH, slides were hybridized with Stellaris RNA FISH probes (Biosearch Technologies) and processed for imaging according to the manufacturer's recommendations for *C. elegans* samples.

The number of condensing nuclei in *alg-3/4* and *csr-1* male germlines was determined using DAPI images from immunofluorescence and RNA FISH experiments by counting all of the nuclei from pachytene to metaphase of the first meiotic division in which chromosomes could not be individually resolved.

Images were acquired with a Solamere Technology Group CSU10B Spinning Disk Confocal System scan head mounted on a Nikon TE-200E2 inverted microscope with a 60X or 100X Plan/APO Oil lens and a Roper Coolsnap HQ2 camera. Images were processed using Metamorph (7.7.4), ImageJ (1.47n5), and Adobe Photoshop (CS4) software. Z sections with slices ranging from 0.2 to 0.4 µm were collected.

Proteomics

Sperm were isolated from *fog-2* males fed with ^{15}N-labeled (or heavy) HB101 bacteria for three generations at 20°C and from *fog-2* or *alg-3/4; fog-2* males fed with unlabeled HB101 bacteria at 20°C or 25°C. Each sperm sample was homogenized and lysates were centrifuged twice at 10,000 x *g* for 20 minutes. Proteins were precipitated in trichloroacetic acid (TCA) and washed in acetone. Proteins were denatured, reduced and alkylated prior to trypsin digestion. The ^{15}N-labeled (or heavy) protein sample served as an internal quantification standard by combining with unlabeled (ie. ^{14}N or light) experimental sperm protein samples. MuDPIT (Washburn et al., 2001) analyses were performed using an Eksigent nano-LC pump and a Thermo LTQ-Orbitrap connected to a homemade electrospray stage. Protein identification and quantification were performed with Integrated Proteomics Pipeline - IP2 (Integrated Proteomics Applications, Inc., San Diego, CA. http://www.integratedproteomics.com/). Tandem mass spectra were extracted from raw files using RawExtract 1.9.9 (McDonald et al., 2004), searched against the Wormbase protein database (release WP180) and reversed sequences using ProLuCID (Peng et al., 2003; Tao Xu, 2006), and peptide candidates were filtered using DTASelect with the parameters –p 1 –y 1 –trypstat –DM 10 –in (McDonald et al., 2004; Tabb et al., 2002). Quantification was performed using Census (Park et al., 2008).

Reagents and Chemicals. Deionized water (18.2 MW, Barnstead, Dubuque, IA) was used for all preparations. Buffer A consists of 5% acetonitrile 0.1% formic acid, buffer B consists of 80% acetonitrile 0.1% formic acid, and buffer C consists of 500 mM ammonium acetate and 5% acetonitrile.

Sample Preparation. 50 ug each of light and heavy proteins were mixed and brought to 200 ul with water. Proteins were precipitated with 60 ul TCA (Sigma-Aldrich, St. Louis, MO, Product number T-0699) at 4°C O/N. After 30 min centrifugation at 18000 x g, protein pellets were washed 2 times with 500 ul ice-cold acetone. Air-dried pellets were dissolved in 8 M urea/ 100 mM Tris pH 8.5. Proteins were reduced with 5 mM Tris(2-carboxyethyl)phosphine hydrochloride (Sigma-Aldrich, St. Louis, MO, product C4706) and alkylated with 10 mM iodoacetamide (Sigma-Aldrich, St. Louis, MO, product I11490). Proteins were digested for 18 hr at 37°C in 2 M urea 100 mM Tris pH 8.5, 1 mM $CaCl_2$ with 2 ug trypsin (Promega, Madison, WI, product V5111). Digest was stopped with formic acid, 5% final concentration. Debris was removed by centrifugation, 30 min 18000 x g.

MudPIT Microcolumn. A MudPIT microcolumn (Wolters et al., 2001) was prepared by first creating a Kasil frit at one end of an undeactivated 250 mm ID/360 mm OD capillary (Agilent Technologies, Inc., Santa Clara, CA). The Kasil frit was prepared by briefly dipping a 20 - 30 cm capillary in well-mixed 300 mL Kasil 1624 (PQ Corporation, Malvern, PA) and 100 mL formamide, curing at 100°C for 4 hrs, and cutting the frit to ~2 mm in length. Strong cation exchange particles (SCX Luna, 5 mm dia., 125 Å pores, Phenomenex, Torrance, CA) were

packed in-house from particle slurries in methanol 2.5 cm. An additional 2.5 cm reversed phase particles (C18 Aqua, 3 mμm dia., 125 Å pores, Phenomenex) were then similarly packed into the capillary using the same method as SCX loading, to create a biphasic column. An analytical RPLC column was generated by pulling a 100 mm ID/360 mm OD capillary (Polymicro Technologies, Inc, Phoenix, AZ) to 5 mm ID tip. Reversed phase particles (Aqua C18, 3 mm dia., 125 Å pores, Phenomenex, Torrance, CA) were packed directly into the pulled column at 800 psi until 12 cm long. The MudPIT microcolumn was connected to an analytical column using a zero-dead volume union (Upchurch Scientific (IDEX Health & Science), P-720-01, Oak Harbor, WA).

LC-MS/MS analysis was performed using an Eksigent nano-LC pump and a Thermo LTQ-Orbitrap using an in-house built electrospray stage. MudPIT experiments were performed where each step corresponds to 0, 10, 20, 30, 40, 50, 60, 70, and 100% buffer C being run for 4 min at the beginning of each gradient of buffer B. The 100% buffer C step was repeated. Electrospray was performed directly from the analytical column by applying the ESI voltage at a tee (150 mm ID, Upchurch Scientific). Electrospray directly from the LC column was done at 2.5 kV with an inlet capillary temperature of 250°C. Data-dependent acquisition of MS/MS spectra with the LTQ -Orbitrap were performed with the following settings: MS/MS on the 6 most intense ions per precursor scan, 1 microscan, charge state 1 reject; dynamic exclusion repeat count, 1, repeat duration, 30 second; exclusion list size 500; and exclusion duration, 180 second.

Data Analysis. Protein and peptide identification and protein quantitation were done with Integrated Proteomics Pipeline - IP2 (Integrated Proteomics

Applications, Inc., San Diego, CA.) Tandem mass spectra were extracted from raw files using RawExtract 1.9.9 (McDonald et al., 2004) and were searched against Wormbase database (WP180) with reversed sequences using ProLuCID (Peng et al., 2003; Tao Xu, 2006). The search space included all fully-tryptic peptide candidates. Carbamidomethylation (+57.02146) of cysteine was considered as a static modification. Peptide candidates were filtered using DTASelect, with these parameters –p 1 –y 1 –trypstat –DM 10 –in (McDonald et al., 2004; Tabb et al., 2002). Quantitation was performed using Census (Park et al., 2008).

CHAPTER IV

Sperm Development, Temperature, and Small RNAs

Overview of Research

In *C. elegans* the redundant AGO-clade paralogs, T22B3.2 (ALG-3) and ZK757.3 (ALG-4) are required for male fertility at elevated temperatures. *alg-3/4* mutants exhibit reduced fertility at 20°C and complete sterility at 25°C. At all temperatures *alg-3/4* mutants lack a class of endogenous small RNAs, named 26G-RNAs that target greater than 1400 spermatogenesis-expressed mRNAs. *alg-3/4* sterility can be rescued by mating with wild type males, suggesting that the infertility results from defects specific to the male germline. *alg-3/4* sperm exhibit a temperature dependent defect in spermiogenesis, which in *C. elegans* is the haploid post-meiotic differentiation of a spermatid (non-motile) into spermatozoa capable of motility, through the formation of a pseudopod (Ward et al., 1981). A rescuing GFP::ALG-3 transgene is localized in P-granules beginning at the late pachytene stage of male gametogenesis, persisting until the haploid stage of meiosis. Two other Argonautes, CSR-1 and WAGO-1 are expressed throughout the male germline and also abundant in mature sperm.

26G-RNAs are synthesized using the targeted mRNAs as template, by the RNA-dependent RNA polymerase (RdRP) RRF-3 (see Figure 1.4). ALG-3/4 and 26G-RNAs can then target the mRNA from which the small RNA was generated, to recruit a 2° RdRP. This leads to the production of secondary (2°) 22G-RNAs. A subset of targets generates 22G-RNAs that are loaded into the essential Argonaute CSR-1 (see Figure 1.4). *csr-1* mutant males are identical phenotypically to *alg-3/4* males and like *alg-3/4* mutants, exhibit temperature

sensitive infertility resulting from defective spemiogenesis. CSR-1 and 22G-RNAs associate with the chromatin of target genes, where they promote the transcription of this subset of ALG-3/4 26G-RNA target mRNAs, creating a positive-feedback loop to drive expression of target mRNAs and additional small RNAs.

ALG-3/4 and 26G-RNAs are also required for the posttranscriptional silencing of a distinct subset of targets in the male germline, that do not appear to be required for spermiogenesis. The 22G-RNAs derived from these targets are thought to be loaded into silencing 2° Argonautes, as factors required for the production of silencing WAGO 2° 22G-RNAs, and the WAGOs themselves are required for the silencing of these targets in the male germline (see Figure 1.4).

alg-3/4 and *csr-1* heterozygous males are normally fertile at all temperatures. However, heterozygous male progeny produced after 3 to 5 generations of paternal homozygosity exhibit the same thermo-intolerant sterile phenotype observed for homozygous males. CSR-1 is localized to the chromatin and cytoplasm of mature sperm, permitting it to transmit small RNAs to the zygote during fertilization. These findings are consistent with a role for the ALG-3/4, CSR-1 pathway in maintaining an epigenetic program for thermotolerant spermiogenesis.

CSR-1 also engages small RNAs antisense to female-specific germline mRNAs that are not expressed in males, suggesting that they could be maternally inherited. Taken together our findings suggest that the ALG-3/4 CSR-

1 pathway promotes thermotolerance by preserving robust spermiogenic gene expression during meiosis, and that CSR-1 functions epigenetically to transmit a memory of both male- and female-specific gene expression to successive generations.

Connecting the dots between chapters II and III

We initially identified 397 ALG-3/4 26G-RNA targets, based on a small RNA sequencing data set from males cultured at 20°C, using a stringent >10 read per million (RPM) cutoff, as described in chapter II. With more extensive sequencing of small RNAs from males cultured at both 20°C and 25°C and a less stringent >5 RPM cutoff, we identified 1408 ALG-3/4 26G-RNA target RNAs in chapter III. The former targets are represented in the latter list, and the latter list of targets includes those regulated at 25°C, where the pathway's function is most critical for sperm development and the targets are regulated most intensely.

The remaining difference between chapters II and III, is that in chapter II I describe the pathway as merely a silencing pathway functioning to downregulate target mRNAs. Mainly, due to the fact that of the 397 ALG-3/4 26G-RNA targets identified in chapter two, 109 were upregulated in *alg-3/4* mutants and only 11 were downregulated, determined by microarray analysis of WT and *alg-3/4* young adult hermaphrodite samples cultured at 20°C. This actually correlates well with the mRNA sequencing data of chapter III, where of the 1408 ALG-3/4 26G-RNA targets, 242 target mRNAs are upregulated and 20 are downregulated

in *alg-3/4* mutants at 20°C, as well as with the sperm proteomic data at 20°C where 74 target proteins are upregulated and 60 are downregulated. The remaining differences are probably a result of differing RNA quantitation techniques, and in chapter II the microarray analysis was performed on young adult hermaphrodites that have ceased sperm production, compared to mRNA-Seq of males in chapter III. Thus, it appears that at permissive temperature (20°C) the ALG-3/4 26G-RNA pathway predominantly functions to silence ALG-3/4 negatively-regulated targets. At 25°C the pathway predominantly promotes the transcription of ALG-3/4 positively-regulated targets, while also silencing the negatively-regulated targets. Interestingly, CSR-1 is required to promote the transcription of positively-regulated targets, but not to silence negatively-regulated targets, and *csr-1* males are temperature sensitive (TS) male sterile, suggesting that the loss of transcription of ALG-3/4 positively-regulated targets underlies the spermiogenic defect at 25°C. *rde-3* and *mago-12* mutants are also sterile at 25°C and required for the silencing of ALG-3/4 negatively-regulated targets, but the sterility in these mutants is not completely due to defects in the male germline (chapter III, Masaki Shirayama personnel communication).

Of the 397 ALG-3/4 targets identified in chapter II, 205 exhibit a 5′ bias for where in the transcript the 26G-RNAs map, and 136 exhibit a 3′ bias. 5′ targeting appears to be predictive of silencing at 20°C as 88/205 genes with 5′ targeting bias are upregulated in their steady state mRNA levels in *alg-3/4* mutants (chapter II). This trend for 26G-RNAs targeting the 5′ and 3′ ends of transcripts

also exists at 25°C when analyzing the chapter III data, as 653 targets exhibit a 5´ bias and 503 exhibit a 3´ bias. However, it is not predictive of silencing, as 48% and 31% negatively-regulated targets, and 46% and 34% of positively-regulated targets, exhibit 5´ and 3´ small RNA targeting bias, respectively (Figure 4.1A). Both 5´ and 3´ biased targets are also targets of CSR-1-assoicated 22G-RNAs in the male germline (Figure 4.1B). It is unclear why ALG-3/4 targets are preferentially targeted at the 5´ and 3´ end of their transcripts; although, it could be potentially correlated with UTRs. Thus, ribosome occupancy and translation may play a role in where RRF-3 (RdRP) can access the transcript to generate 26G-RNAs.

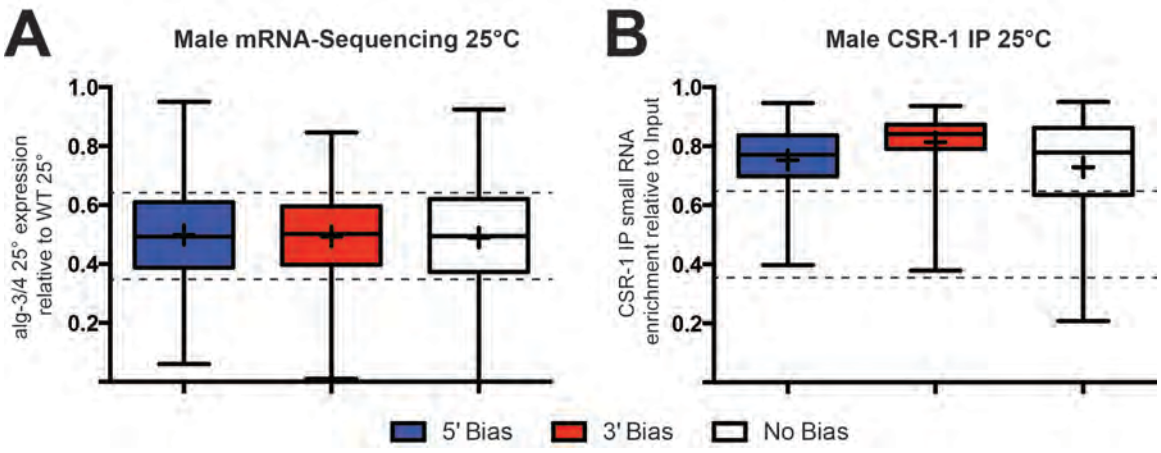

Figure 4.1. 5′ or 3′ biased targeting by ALG-3/4 26G-RNAs is not predictive of the regulatory outcome on that target mRNA.
(A) Box and whisker plot indicating the relative abundance of target mRNA by mRNA-seq at 25°C. Calculated as the RPM ratio of WT RPM / (*alg-3/4* RPM + WT RPM). Dotted lines indicate 2-fold increase (upper) or decrease (lower). '+' indicates the average enrichment value for all genes in the category. **(B)** The relative abundance of small RNAs antisense to 5′ or 3′ targeting biased genes in the CSR-1 IP. Enrichment was calculated as the RPM ratio of FLAG::CSR-1 IP / (FLAG::CSR-1 IP + Input). Dotted lines indicate 2-fold enrichment (upper) or depletion (lower). '+' indicates the average enrichment value for all genes in the category.

To promote transcriptionally or to silence postranscriptionally

The ALG-3/4 26G-RNA pathway targets can be separated into two distinct functional categories: 1) Positively-regulated targets that CSR-1 and 22G-RNAs promote the transcription of during spermatogenesis. 2) Negatively-regulated targets that require the WAGO machinery downstream of ALG-3/4 26G-RNAs for posttranscriptional silencing. Interestingly, the majority of ALG-3/4-positively regulated targets appear to be required for spermiogenesis, sperm motility, and fertilization. These 'spermiogenic' targets include many of the MSP (Major Sperm Protein) genes that are the main structural proteins composing the pseudopod. As well as, all four of the related MSD (Major Sperm Domain) genes required in the nematode *Ascaris suum* for depolymerization of MSP fibers as well as genes encoding cytosolic motility proteins (CMP) which in the nematode *Ascaris suum* complex with MSP to promote polymerization and motility (Buttery et al., 2003). Also included in the ALG-3/4 positively regulated targets are regulators of microtubule dynamics, and several SSP genes that encode for proteins similar to MSP. All appear to be intimately involved in pseudopod formation. The positively-regulated targets also include the PP1 phosphatases *gsp-3/4*, which in *C. elegans* have TS sterility phenotypes similar to *alg-3/4* and are also thought to regulate MSP dynamics (Wu et al., 2012). Genes that function during fertilization are also represented in ALG-3/4 positively regulated targets, including many of the components of the proteasome, which has been shown to be required for fertilization in mammals (Sakai et al., 2004), and lectins which function in sperm-

oocyte binding events (Clark, 2013). A similar correlation with reproductive genes and the ALG-3/4 negatively regulated targets is not present. However, these genes do include many factors associated with the electron transport chain and ribosomal protein genes.

For the positively-regulated targets, ALG-3/4 uses 26G-RNAs as guides to target sperm genes, resulting in the recruitment of 2° RdRPs (RRF-1 and EGO-1) to generate 2° 22G-RNAs, with a 5´ guanosine triphosphate (Conine et al., 2010). These 2° 22G-RNAs are loaded into CSR-1, as CSR-1 associates with the small RNAs derived from ALG-3/4 targets in the male germline (Conine et al., 2013). Downstream of ALG-3/4 and 26G-RNA targeting, CSR-1 uses 2° 22G-RNAs to enter the nucleus and associate with the chromatin of ALG-3/4 target genes, probably via the nascent transcript produced by RNA Pol II (Cecere et al., 2014; Conine et al., 2013; Gu et al., 2009; Wedeles et al., 2013b). A major question remaining in not only the male germline but also the hermaphrodite germline is; how does CSR-1 promote transcription? During spermatogenesis in most organisms, including *C. elegans*, the chromatin of developing sperm undergoes increased condensation, culminating in a transcriptionally inert haploid spermatid (Shakes et al., 2009; Ward et al., 1981; Wykes and Krawetz, 2003). In mammals and other organisms, the process of spermatogeneic chromatin condensation is facilitated by the replacement of histones with small basic proteins called protoamines (Wykes and Krawetz, 2003). It is unclear if a similar replacement occurs in *C.* elegans; however, similar small basic proteins

have been identified in the sperm chromatin proteome of worms (Chu et al., 2006).

One hypothesis for the function of ALG-3/4 and CSR-1 promoting transcription is that it prevents the condensation of chromatin around genes that are required for spermiogenesis, as these genes are required postmeiotically (as condensation is progressing) to function upon the conclusion of meiosis, whereas other loci undergo condensation and silencing (Conine et al., 2013). CSR-1 associated with the chromatin of ALG-3/4 targets may define these domains, by recruiting chromatin modifiers, or histone variants that prevent the formation of transcriptionally inert chromatin. Supporting this model, sperm genes and ALG-3/4 26G-RNA targets are nonrandomly distributed in the *C. elegans* genome, located in several large clusters on two of the autosomes (C.C. unpublished data, (Miller et al., 2004)).

CSR-1 promoting the transcription of ALG-3/4 positively-regulated targets creates a positive-feedback loop, as increased transcription generates more template for RRF-3 to make 26G-RNAs loaded in ALG-3/4, and for ALG-3/4 and 26G-RNAs to recruit 2° RdRPs to generate 2° 22G-RNAs loaded into CSR-1, increasing transcription (Conine et al., 2013). Initiated after the onset of spermatogenesis, this positive-feedback would sufficiently drive the transcription, in a short period of time, of large amounts of spermiogenic mRNAs required for successful postmeiotic differentiation to produce a spermatozoa competent for fertilization.

ALG-3/4 and 26G-RNAs are also required for the posttranscriptional silencing of a separate subset of targets in the male germline, which do not appear to be required for spermiogenesis (Conine et al., 2013). Factors required for the production of silencing WAGO 2° 22G-RNAs, and the WAGOs themselves are also required for the silencing of these targets in the male germline, suggesting that ALG-3/4 and 26G-RNA 1° targeting also produce 2° 22G-RNAs that are loaded into WAGOs (Conine et al., 2013).

Recently, I performed WAGO-1 and WAGO-9 Immunoprecipitation (IP) experiments from males grown at 25°C, using FLAG-tagged (*wago-1* or *wago-9*) *fog-2* (obligate male/female) strains, and sequenced the associated small RNAs. The canonical member of the WAGO pathway, WAGO-1, is localized to the P granules and cytoplasm throughout the male germline, and is thought to posttranscriptionally silence target mRNAs (Conine et al., 2010; Gu et al., 2009). WAGO-1 was previously shown to interact with 22G-RNAs that target greater than a thousand genes in the adult hermaphrodite germline (Gu et al., 2009). WAGO-9 is localized throughout the hermaphrodite germline in the nucleus, where in association with 22G-RNAs it directs epigenetically inherited transcriptional silencing of exogenous RNAi or RNAe (Ashe et al., 2012; Buckley et al., 2012; Luteijn et al., 2012; Shirayama et al., 2012) WAGO-9's localization in the male germline is unknown, but it is known to be associated with the chromatin of sperm (Chu et al., 2006). Prior to this experiment, the endogenous targets of WAGO-9 in either the hermaphrodite or male germline were unknown.

Surprisingly, 22G-RNAs targeting ALG-3/4 negatively-regulated targets are not enriched in either the WAGO-1 or WAGO-9 IP (Figure 4.2A). Only 3 and 24 ALG-3/4 negatively regulated targets are greater than 2-fold enriched in the WAGO-1 and WAGO-9 IP, respectively. These data indicate that other WAGO Argonautes are required for the posttranscriptional silencing of these targets. Using male mRNA-seq from chapter III, and enrichment in the male CSR-1 IP (an indicator of germline expression) to identify WAGOs expressed at the appropriate time and place, we identified potential ALG-3/4 26G-RNA pathway downstream silencing WAGOs in the male germline, including: *wago-3 (ppw-2), wago-4, wago-6 (sago-2), wago-7 (ppw-1),* and *wago-10* (Figure 4.2B).

It is unclear whether ALG-3/4 negatively-regulated targets are silenced spatially or temporally by ALG-3/4 in the male germline, or if these targets are ubiquitously silenced, raising the question of their function/presence in the *C. elegans* genome. Like the endogenous PRG-1/21U-RNA and WAGO targets, ALG-3/4 negatively-regulated targets could be spatially or temporally regulated. At some location or developmental point, they may be required for germline function, while at others they may not be required or detrimental, hence their silencing. To test this, I have attempted to localize two negatively-regulated target transcripts, *ssp-16* and *f36h12.4*, using RNA fluorescence in situ hybridization (RNA-FISH) in the male germline, as in chapter III. I was unable to detect a significant signal, probably because these RNAs are expressed at a relatively low level.

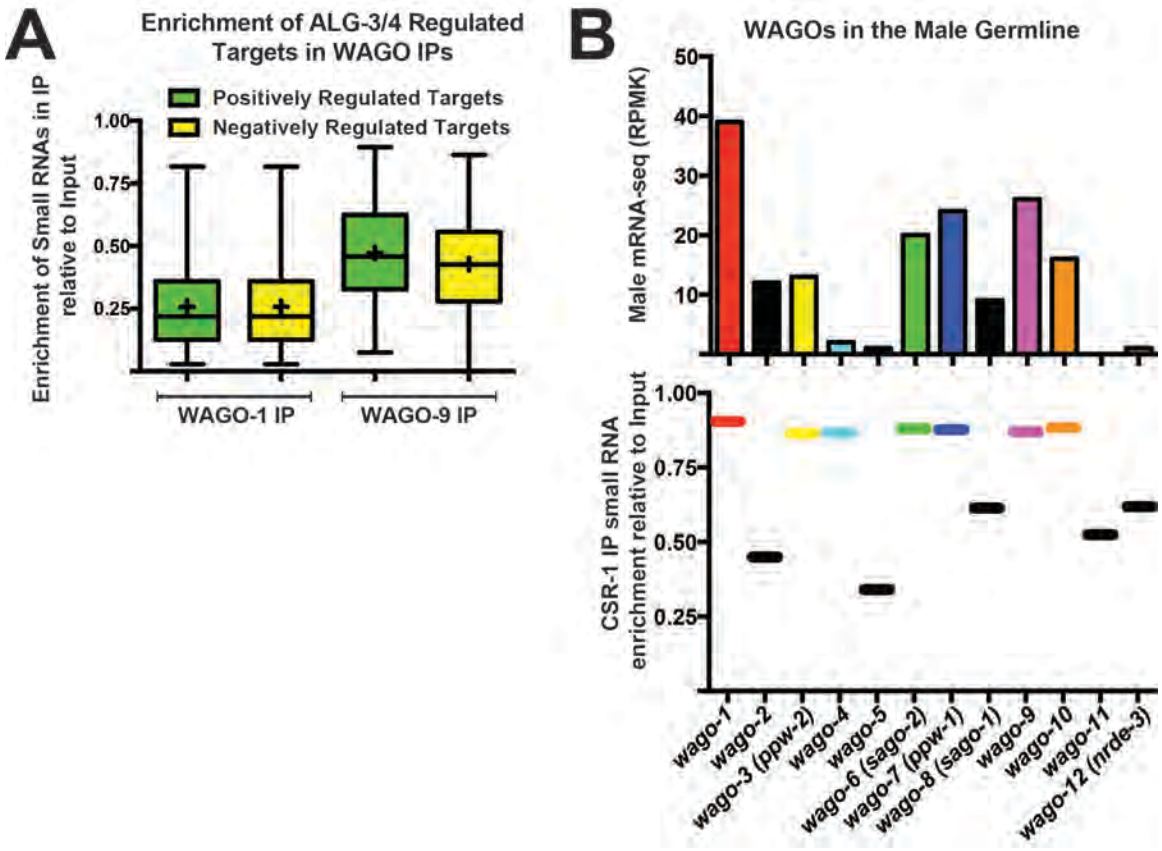

Figure 4.2. WAGO Argonautes in the male germline.
(A) Box and whisker plot indicating the relative abundance of 22G-RNAs antisense to ALG-3/4 targets in the WAGO-1 or WAGO-9 male IP. Enrichment was calculated as the RPM ratio of FLAG::WAGO/ (FLAG::WAGO + Input). Dotted lines indicate 2-fold enrichment (upper) or depletion (lower). '+' indicates the average enrichment value for all genes in the category. **(B)** *Top,* mRNA expression of WAGOs by RNA-Seq of 25°C males samples in reads per million per kilobase (RPMK). *Bottom*, enrichment or depletion of small RNAs in the CSR-1 male IP targeting WAGOs, relative to the input. Enrichment in the CSR-1 IP is a readout for male germline gene expression.

Other small RNAs in the male germline

CSR-1 and its associated 22G-RNAs were originally shown to target 4190 genes in adult hermaphrodites (Claycomb et al., 2009). By pulling down CSR-1 and sequencing the associated small RNAs in the male germline I determined that CSR-1 22G-RNAs target 5574 genes, including 90% (3775/4190) of the hermaphrodite targets. This experiment revealed an additional 1799 CSR-1 targets that were previously unknown, many expressed specifically in the male germline or sperm. The majority of these genes are downstream of the 1408 ALG-3/4 26G-RNA targets, as 82% (1156/1408) are also CSR-1 targets (Figure 4.3).

The WAGO genomic surveillance pathway was originally shown to target 2743 genes. Like CSR-1, this initial analysis was also performed on data obtained only from adult hermaphrodite samples (Gu et al., 2009), thus raising the question of how many additional WAGO targets are in the male germline? From the WAGO-1 (cytoplasmic) and WAGO-9 (nuclear) IP/small RNA-seq (above) from males grown at 25°C, WAGOs/22G-RNAs target 3284 genes, including 1099 of the hermaphrodite WAGO target genes, revealing an additional 2185 targets in the male. WAGO-1 and WAGO-9 target 2134 and 1587 genes respectively, with only 737 shared (Figure 4.3). This suggests that WAGO-1 and WAGO-9 may have distinct regulatory functions, beyond their redundancies, as they have many unique targets.

To determine PRG-1/21U-RNA dependent and independent targets I also performed IP/small RNA-seq of WAGO-1 and WAGO-9 from *prg-1* mutant males grown at 25°C, using FLAG-tagged *prg-1; fog-2* strains, and sequenced the small RNAs associated. Of the 3284 WAGO targets, 2128 are dependent on *prg-1* for their accumulation, with the remaining 1156 being independent of the piRNA pathway. Nearly half (554/1132) of the PRG-1-independent WAGO targets are also CSR-1 targets in the male germline. Conversely, only 23% (495/2128) of the PRG-1/21U-RNA-dependent WAGO-1 male targets overlap with CSR-1 male targets (Figure 4.3). These data support the notion that the male target genes are in functionally distinct pathways, and provide a connection between CSR-1 and PRG-1 independent WAGO pathways.

Together these findings provide a comprehensive list of genes targeted by the CSR-1 22G-RNA, ALG-3/4 26G-RNA pathway, and PRG1/21U-RNA dependent and independent 22G-RNA pathways, in the male germline. These pathways account for between 94-98% of all genes targeted by all small RNAs with greater than 5 RPM in the male germline, depending on the WT small data set analyzed, *fog-2* males 20°C, *fog-2* males 25°C, or N2 males 25°C. Of the 7515 genes that produce greater than 10 RPMK (reads per million per thousand nucleotides) from the 25°C mRNA-seq dataset described in chapter III, 62% are targeted by one of the pathways defined here. This percentage would certainly increase by subtracting genes expressed only in the soma; however, a comprehensive list produced from mRNA-seq data of male somatic mRNAs is

not currently available. Regardless, it is clear that Argonautes and small RNAs target the majority of male mRNAs to function as master regulators of germline gene expression in *C. elegans*.

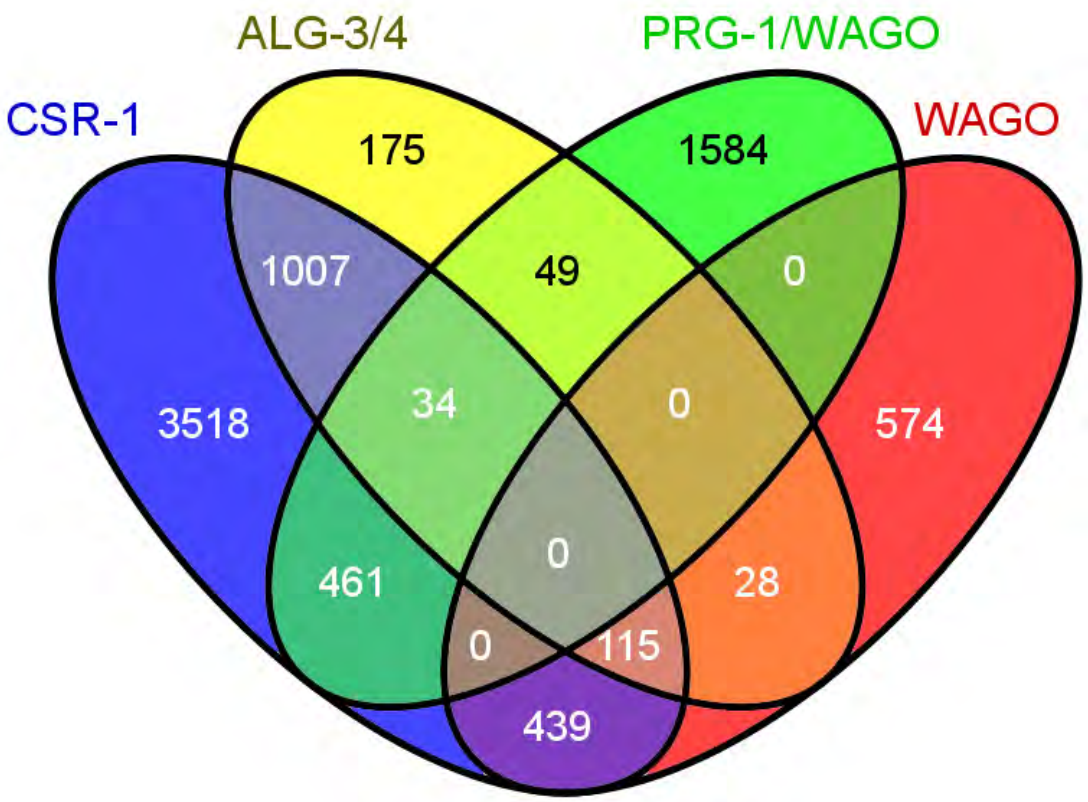

Figure 4.3. Genes targeted by endogenous small RNA pathways in the male germline.
(A) Venn diagram illustrating the overlap between the CSR-1, ALG-3/4, PRG-1 dependent WAGO, and PRG-1 independent WAGO endogenous small RNA pathways in the male germline. Created using http://bioinfogp.cnb.csic.es/tools/venny/index.html.

The FER/ERI mutants and their male germline defects

At 20°C (permissive temperature) *alg-3/4* males produce ~50% of the progeny a WT male produces, while at 25°C they are completely sterile (Conine et al., 2010; Han et al., 2009). At 25°C *alg-3/4* male germlines exhibit some meiotic defects, and produce 29% less sperm than a WT male, however, this does not explain the fully penetrant sterility at increased temperatures. Complete male sterility at 25°C can be explained by a defect in spermiogenesis. At 20°C *alg-3/4* mutants produce a fraction of normal looking spermatozoa, while the majority have deformed pseudopods resembling spike-like structures that are the intermediates of pseudopod development (Ward et al., 1981). At 25°C, greater than 95% of *alg-3/4* spermatids arrest and never undergo any cellular morphogenesis indicative of spermiogenesis.

Several of the *fer* mutants (*fer-2,3,6,&15)*, for *fertilization* defective, that were isolated over 30 years ago in genetic screens for TS sperm-defective mutants (Argon and Ward, 1980; Hirsh and Vanderslice, 1976), are also defective in the ALG-3/4 26G-RNAs, as the small RNAs are depleted in these mutants. Interestingly, these mutants exhibit almost identical phenotypes to *alg-3/4* mutants, as identified by ultrastructural examination of mutant sperm by transmission electron microscopy (Conine et al., 2013; Ward et al., 1981). Both the *fer* and *alg-3/4* phenotype include the disappearance of the RNA halo, which appears to encase the nucleus of *C. elegans* sperm. The function of this structure is currently mysterious. Previously, the genetic lesions associated with

these mutants were unknown, however, after determining that *fer-3* and *fer-15* were required for sperm 26G-RNA biogenesis, they were successfully mapped to genes required for the biogenesis of 26G-RNAs, *eri-3* and the RdRP *rrf-3*, respectively.

Other *eri* mutants, *eri-1, eri-3,* and *eri-4* (an allele of Dicer), also demonstrate similar spermiogenic phenotypes, as well as more subtle meiotic defects, and X-chromosome segregation defects (leading to an increased incidence of male progeny) (Pavelec et al., 2009). These *eri* mutants and *rrf*-3 are required for the biogenesis and/or accumulation of ALG-3/4 sperm-specific 26G-RNAs (Conine et al., 2010; Conine et al., 2013; Gent et al., 2009; Han et al., 2009; Welker et al., 2010). It has also been reported that some allelles of *rrf-3* display more severe defects in meiotic cell divisions and X-chromosome segregation, as well as a paternally transmitted embryonic-lethal phenotype to progeny, which depends on the mutant sperm being produced at elevated temperatures (Gent et al., 2009). This suggests that the RdRP RRF-3 may have pleiotropic functions in the male germline outside of the ALG-3/4 pathway.

Why is Sperm Development Sensitive to Temperature?

In many animals, sperm development is inherently temperature sensitive (Rockett et al., 2001). The effects of elevated temperature during sperm development can also be epigenetically transmitted, causing phenotypes in developing embryos in mammals including decreased embryonic mass and

increased mortality rates (Jannes et al., 1998; Mieusset et al., 1991; Setchell et al., 1988). However, the molecular mechanism behind the temperature sensitivity of sperm development remains unknown.

Our findings indicate that the gene expression programs required for spermatogenesis in *C. elegans* are sensitive to temperature. In wild-type males, the expression of many spermiogenesis genes was higher at 25°C than at 20°C. Intriguingly, WT sperm development is inherently sensitive to temperature, with males exhibiting a greater than 25% reduction in brood size at elevated temperatures (Hirsh and Vanderslice, 1976). The majority of the *eri* mutants (and *alg-3/4*) that exhibit TS male sterility are null alleles and the sperm-specific 26G-RNAs are absent at all temperatures, thus, the TS phenotype cannot be due to conditional inactivation of the pathway. Rather it is more likely that the pathway functions to regulate an inherently temperature sensitive process (Conine et al., 2010; Han et al., 2009; Pavelec et al., 2009).

Here we have shown that ALG-3/4 and CSR-1 are required to initiate and maintain the activation of spermiogenesis genes at elevated temperatures. Failure to maintain spermiogenic gene expression in *alg-3/4* and *csr-1* mutants correlates with dramatic defects in spermatid activation and infertility in both mutant strains at elevated temperature. Thus, the ALG-3/4 and CSR-1 pathways appear to act as enhancers of gene expression that buffer the effects of temperature on sperm development.

Increasing temperature could increase the rate of meiosis and spermatogenic nuclear condensation, resulting in a shortened window to express spermiogenic transcripts required at the completion of meiosis. Thus, an active mechanism to identify and transcribe spermiogenic genes at high production rates, in the face of nuclear condensation would be essential at elevated temperatures, while partly dispensable at permissive temperatures. In this model, CSR-1 downstream of ALG-3/4 and 26G-RNAs would ensure a burst of spermiogenic transcription buffered against increasing temperature, loading the sperm with the factors necessary to undergo spermiogenesis. Loss of the pathway would then lead to a TS spermiogenic defect, due to an incomplete complement of the factors required for spermiogenesis (Conine et al., 2013).

The increased condensation of chromatin during spermatogenesis is a highly conserved aspect of sperm development (Rathke et al., 2014). In order to develop a sperm competent for fertilization in any organism, the meiotic divisions of a 4C 1° spermatocyte into haploid spermatid must precede the postmeiotic differentiation of that sperm, from a round spermatid into a motile spermatozoa (spermiogenesis). Consequently, two separate, coordinately executed gene expression programs must exist. Spermiogenesis occurs during and/or after spermatogenic nuclear condensation, thus, the spermiogenic transcriptional program could occur prior to condensation and meiosis, with the mRNAs required then stored for translation after the completion of meiosis. Alternatively, the transcription of spermiogenic genes could occur during spermatogenic

nuclear condensation creating a stockpile of RNAs that are immediately translated to initiate spermiogenesis. The fact that haploid sperm in *C. elegans* and other organisms are devoid of ribosomes supports the latter hypothesis.

If spermiogenic genes are transcribed specifically while the nucleus is condensing, then transcription would be dependent on the condensation state of chromatin associated with these genes. I hypothesize that this could be the reason for the conserved affect of temperature on sperm development. Therefore, an active process that allowed spermiogenic genes to sustain open chromatin during spermatogenic nuclear condensation would buffer against conditions that alter the dynamics of chromatin, such as increased temperature. In *C. elegans* the ALG-3/4 26G-RNA and CSR-1/22G-RNA pathway fill this niche, while in other organisms, different small RNAs or other mechanisms may serve an analogous function. Identifying these pathways may be of particular relevance to human fertility, as many cases of infertility are associated with unexplained male-specific defects.

Small RNAs as Carriers of Epigenetic Information

Small RNAs play central roles in most of the best-understood transgenerational epigenetic inheritance paradigms, most extensively in *C. elegans*, where phenotypes triggered artificially by small interfering RNAs (siRNAs) can be inherited for greater than 5 generations, and small RNAs provide an ancestral ‘memory’ of germline gene expression (Alcazar et al., 2008;

Ashe et al., 2012; Buckley et al., 2012; Grishok et al., 2000; Luteijn et al., 2012; Rathke et al., 2014; Seth et al., 2013; Shirayama et al., 2012; Vastenhouw et al., 2006; Wedeles et al., 2013a). Transgenerational inheritance mediated by small RNAs is also widespread in plants, playing key roles in phenomena such as paramutation in maize (Arteaga-Vazquez and Chandler, 2010). In mammals, recent studies in mouse demonstrate that early life trauma in males can lead to changes in the RNA payload of sperm, which can epigenetically transmit behavioral and metabolic phenotypes to offspring. Intriguingly, injection of sperm small RNA from males subjected to early life trauma into control embryos can induce a subset of these phenotypes in the offspring, suggesting that the sperm small RNAs are epigenetically transmitting the phenotypes (Gapp et al., 2014).

Here, I show that CSR-1 and 22G-RNAs functioning downstream of ALG-3/4 also transmit epigenetic information about ancestral paternal germline gene expression via the small RNAs to succeeding generations, as CSR-1 and 22G-RNAs are abundant in sperm (Conine et al., 2010; Conine et al., 2013). Remarkably, repeatedly backcrossing heterozygous hermaphrodites to homozygous *alg-3/4* or *csr-1* mutant males results in a progressive loss of fertility (germline-mortal phenotype) that could be rescued by wild-type sperm. This sterility is analogous to that of the *alg-3/4* and *csr-1* male TS fertility defect, resulting in temperature-dependent defective spermiogenesis (Conine et al., 2013). Also, spermiogenic ALG-3/4 positively regulated targets exhibit reduced transcription in the heterozygotes derived from mutant paternal lines, suggesting

that males transmit a CSR-1 small RNA signal downstream of ALG-3/4 and 26G-RNAs that provides a paternal memory of germline gene expression.

As WAGOs associated with silencing 22G-RNAs downstream of PRG-1/22G-RNAs, are also present in the sperm (Conine et al., 2010), and in accord with the idea of CSR-1 representing ‘self’, and WAGOs as ‘non-self’, it suggests that sperm and eggs transmit memories of previous ancestral germline gene expression to succeeding generations, to regulate intrinsic germline factors and to prevent invasion by foreign nucleic acids produced by transposons and viruses. In support of this, it was recently demonstrated in *C. elegans* through mRNA and small RNA sequencing of sperm, oocytes, and 1-cell embyros, that the majority of sperm small RNAs are transmitted to the embryo during fertilization as well as ~200 mRNAs (Stoeckius et al., 2014). These findings reveal a surprising, even breathtaking, scope of epigenetic programming in *C. elegans* sperm in which Argonautes and RNAs transmit not only silencing signals but also positive epigenetic signals that function transgenerationally to promote the expression of many spermiogenesis genes.

Small RNAs are also abundant in the sperm of humans and other mammals (Krawetz et al., 2011; Peng et al., 2012). Exhibiting a complex repertoire of sequences that are delivered upon fertilization. The function of paternally transmitted small RNAs in mammals is unclear, creating exciting future avenues of experimentation and study. It is clear, however, that in organisms ranging from worm to man, that small RNAs could potentially carry vast amounts

of epigenetic information delivered to the zygote by sperm. This information has the potential to provide both inherent and acquired information about the environment to ensure adapted fitness during development.

Conclusion

Since the discovery of RNAi and miRNAs in *C. elegans*, a plethora of endogenous germline small RNAs have been discovered. In total, hundreds of thousands of unique sequences are represented, and associated with over 20 Argonautes. A major challenge in the field is to distinguish the function of each Argonaute and their small RNA cofactors. This problem manifests itself most with the WAGOs, as 12 different Argonautes are required for silencing pathways downstream of RNAi, 21U-RNA (piRNA) targeting, and the ERI endogenous small RNA pathways. All 12 have to be knocked-out to completely block RNAi, however, for the other pathways it is unclear whether particular WAGOs are required. Related to this, a central questions is: How do 22G-RNAs with identical physical properties, 5´ guanosine triphosphate, 22nt long, and 3´ hydroxyl, downstream of at least 4 different pathways, get into the correct downstream Argonaute to elicit the correct biological response? It is unclear how CSR-1 and WAGOs each target thousands of distinct RNAs with little overlap, it is obviously biologically functional, as targeting CSR-1 and WAGOs produce very different outcomes.

Remarkably, small RNAs in the *C. elegans* germline can confer both positive and negative regulation of RNAs, regulating RNAs both transcriptionally and posttranscriptionally. It would not be surprising to discover that germline smalls RNAs also regulate translation or other aspects of gene expression. Some of these functions appear to be unique to the nematode lineage, as RdRPs are rare among metazoans genomes, and the majority of these small RNA pathways require their activity. Still, there are definite correlations between many aspects of *C. elegans* small RNA pathways and those of other well studied model organisms, such as flies and mice. The basics of the miRNA pathway are conserved in each, as well as the exogenous RNAi pathway, besides the potency due to amplification of the initial dsRNA trigger by RdRPs in *C. elegans*. A conserved function of the piRNAs in all three systems is to silence transposons; again it appears that worms may use a unique mechanism via RdRP activity. Intriguingly, Piwi Argonautes localize to germ granules in each of these organisms, leading to the hypothesis that these granules could be factories for piRNA/small RNA production and downstream silencing. Many of the *C. elegans* Argonautes localize to P granules, suggesting that they are indeed important for small RNA biogenesis and/or function.

The function of pachytene piRNAs in mammals is unknown, and mutations abrogating the pathway are sterile due to the complete arrest of haploid sperm during spermatogenesis. Interestingly, a very similar phenotype is exhibited by PRG-1/21U-RNA (piRNA) and ALG-3/4 sperm 26G-RNA mutants in *C. elegans*,

however, in worms the defect is temperature dependent. The ALG-3/4 pathway is required to promote the transcription of spermiogenic genes, and loss of this activity is thought to contribute to the phenotype. A similar activity has been proposed for MIWI and mouse pachytene piRNAs although this has yet to be formally proven. It is possible that the ALG-3/4 26G-RNA pathway *C. elegans* and MIWI/ pachytene piRNAs in mice could be providing analogous functions.

ALG-3/4 are AGO-clade Argonautes; therefore, sperm 26G-RNAs are not considered piRNAs, as they are also Dicer-dependent. Nonetheless, they could be functionally analogous to mammalian piRNAs, raising the question: What is a piRNA? Is binding a Piwi Argonaute an appropriate definition, or should expression and function also be considered? To that end, ERGO-1 is a Piwi-clade Argonaute, embryonic 26G-RNAs RNAs are modified at their 3´ termini with 2′ O-methylation, analogous to piRNAs in flies and mouse; however, they are not considered piRNAs in the literature. ERGO-1 26G-RNAs are Dicer-dependent, perhaps lending to this distinction, other than that it is unclear why they are not considered piRNAs.

The nematode *C. elegans* as a model organism of study has been an essential tool to the understanding of RNAi and small RNAs. Since their discovery small RNAs have been identified as intrinsic regulators of gene expression in all forms of life, suggesting that the ancestor of living things had some form of a small RNA pathway. With their tremendous breadth of conservation, small RNA pathways have evolved to fill a plethora of cellular

niches, achieving many unique functions. No place in biology is this more evident than the *C. elegans* germline, where a ‘Small RNA World’ functions prominently in gene expression, reproduction, genomic immunity and adaptability.

BIBLIOGRAPHY

Ahn, S.H., Kim, M., and Buratowski, S. (2004). Phosphorylation of serine 2 within the RNA polymerase II C-terminal domain couples transcription and 3' end processing. Mol Cell *13*, 67-76.

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.

Alvarez-Saavedra, E., and Horvitz, H.R. (2010). Many families of C. elegans microRNAs are not essential for development or viability. Curr Biol *20*, 367-373.

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

Anderson, P., and Kedersha, N. (2009). RNA granules: post-transcriptional and epigenetic modulators of gene expression. Nat Rev Mol Cell Biol *10*, 430-436.

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

Aravin, A.A., Lagos-Quintana, M., Yalcin, A., Zavolan, M., Marks, D., Snyder, B., Gaasterland, T., Meyer, J., and Tuschl, T. (2003). The small RNA profile during Drosophila melanogaster development. Dev Cell *5*, 337-350.

Aravin, A.A., Naumova, N.M., Tulin, A.V., Vagin, V.V., Rozovsky, Y.M., and Gvozdev, V.A. (2001). Double-stranded RNA-mediated silencing of genomic tandem repeats and transposable elements in the D. melanogaster germline. Curr Biol *11*, 1017-1027.

Aravin, A.A., Sachidanandam, R., Bourc'his, D., Schaefer, C., Pezic, D., Toth, K.F., Bestor, T., and Hannon, G.J. (2008). A piRNA pathway primed by individual transposons is linked to de novo DNA methylation in mice. Mol Cell *31*, 785-799.

Aravin, A.A., Sachidanandam, R., Girard, A., Fejes-Toth, K., and Hannon, G.J. (2007). Developmentally regulated piRNA clusters implicate MILI in transposon control. Science *316*, 744-747.

Argon, Y., and Ward, S. (1980). Caenorhabditis elegans fertilization-defective mutants with abnormal sperm. Genetics *96*, 413-433.

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

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.

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.

Bartel, D.P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. Cell *116*, 281-297.

Bartel, D.P. (2009). MicroRNAs: target recognition and regulatory functions. Cell *136*, 215-233.

Bartel, D.P., and Unrau, P.J. (1999). Constructing an RNA world. Trends Cell Biol *9*, M9-M13.

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.

Beyret, E., and Lin, H. (2011). Pinpointing the expression of piRNAs and function of the PIWI protein subfamily during spermatogenesis in the mouse. Developmental biology *355*, 215-226.

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

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

Bohmert, K., Camus, I., Bellini, C., Bouchez, D., Caboche, M., and Benning, C. (1998). AGO1 defines a novel locus of Arabidopsis controlling leaf development. Embo J *17*, 170-180.

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.

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

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.

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.

Burke, D.J., and Ward, S. (1983). Identification of a large multigene family encoding the major sperm protein of Caenorhabditis elegans. J Mol Biol *171*, 1-29

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

Bushati, N., and Cohen, S.M. (2007). microRNA functions. Annu Rev Cell Dev Biol *23*, 175-205.

Buttery, S.M., Ekman, G.C., Seavy, M., Stewart, M., and Roberts, T.M. (2003). Dissection of the Ascaris sperm motility machinery identifies key proteins involved in major sperm protein-based amoeboid locomotion. Mol Biol Cell *14*, 5082-5088.

Cai, X., Hagedorn, C.H., and Cullen, B.R. (2004). Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. Rna *10*, 1957-1966.

Calarco, J.P., and Martienssen, R.A. (2011). Genome reprogramming and small interfering RNA in the Arabidopsis germline. Curr Opin Genet Dev *21*, 134-139.

Carmell, M.A., Girard, A., van de Kant, H.J., Bourc'his, D., Bestor, T.H., de Rooij, D.G., and Hannon, G.J. (2007). MIWI2 is essential for spermatogenesis and repression of transposons in the mouse male germline. Dev Cell *12*, 503-514.

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.

Cech, T.R. (2009). Crawling out of the RNA world. Cell *136*, 599-602.

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.

Chen, P.Y., Manninga, H., Slanchev, K., Chien, M., Russo, J.J., Ju, J., Sheridan, R., John, B., Marks, D.S., Gaidatzis, D., *et al.* (2005). The developmental miRNA profiles of zebrafish as determined by small RNA cloning. Genes Dev *19*, 1288-1293.

Chendrimada, T.P., Gregory, R.I., Kumaraswamy, E., Norman, J., Cooch, N., Nishikura, K., and Shiekhattar, R. (2005). TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. Nature *436*, 740-744.

Chu, D.S., Liu, H., Nix, P., Wu, T.F., Ralston, E.J., Yates, J.R., 3rd, and Meyer, B.J. (2006). Sperm chromatin proteomics identifies evolutionarily conserved fertility factors. Nature *443*, 101-105

Clark, G.F. (2013). The role of carbohydrate recognition during human sperm-egg binding. Hum Reprod *28*, 566-577.

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.

Crick, F.H. (1968). The origin of the genetic code. J Mol Biol *38*, 367-379.

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

De Fazio, S., Bartonicek, N., Di Giacomo, M., Abreu-Goodger, C., Sankar, A., Funaya, C., Antony, C., Moreira, P.N., Enright, A.J., and O'Carroll, D. (2011). The endonuclease activity of Mili fuels piRNA amplification that silences LINE1 elements. Nature *480*, 259-263.

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

Denli, A.M., and Hannon, G.J. (2003). RNAi: an ever-growing puzzle. Trends Biochem Sci *28*, 196-201.

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

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.

Dunoyer, P., Brosnan, C.A., Schott, G., Wang, Y., Jay, F., Alioua, A., Himber, C., and Voinnet, O. (2010). An endogenous, systemic RNAi pathway in plants. Embo J *29*, 1699-1712.

Eddy, E.M. (2002). Male germ cell gene expression. Recent Prog Horm Res *57*, 103-128.

Elbashir, S.M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K., and Tuschl, T. (2001). Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. Nature *411*, 494-498.

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., 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.

Friedman, R.C., Farh, K.K., Burge, C.B., and Bartel, D.P. (2009). Most mammalian mRNAs are conserved targets of microRNAs. Genome Res *19*, 92-105.

Gapp, K., Jawaid, A., Sarkies, P., Bohacek, J., Pelczar, P., Prados, J., Farinelli, L., Miska, E., and Mansuy, I.M. (2014). Implication of sperm RNAs in transgenerational inheritance of the effects of early trauma in mice. Nat Neurosci *17*, 667-669.

Garcia-Lopez, J., Hourcade Jde, D., Alonso, L., Cardenas, D.B., and del Mazo, J. (2014). Global characterization and target identification of piRNAs and endo-siRNAs in mouse gametes and zygotes. Biochim Biophys Acta *1839*, 463-475.

Gassmann, R., Rechtsteiner, A., Yuen, K.W., Muroyama, A., Egelhofer, T., Gaydos, L., Barron, F., Maddox, P., Essex, A., Monen, J., *et al.* (2012). An inverse relationship to germline transcription defines centromeric chromatin in C. elegans. Nature *484*, 534-537.

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

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

Ghildiyal, M., Seitz, H., Horwich, M.D., Li, C., Du, T., Lee, S., Xu, J., Kittler, E.L., Zapp, M.L., Weng, Z., *et al.* (2008). Endogenous siRNAs derived from transposons and mRNAs in Drosophila somatic cells. Science *320*, 1077-1081.

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

Gilbert, W. (1986). Origin of Life - The RNA World. Nature *319*, 618-618.

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.

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.

Grant-Downton, R., Le Trionnaire, G., Schmid, R., Rodriguez-Enriquez, J., Hafidh, S., Mehdi, S., Twell, D., and Dickinson, H. (2009). MicroRNA and tasiRNA diversity in mature pollen of Arabidopsis thaliana. BMC Genomics *10*, 643.

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

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

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

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

Gu, W., Lee, H.C., Chaves, D., Youngman, E.M., Pazour, G.J., Conte, D., Jr., and Mello, C.C. (2012). 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., 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.

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

Han, B.W., and Zamore, P.D. (2014). piRNAs. Curr Biol *24*, R730-733.

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

Haussecker, D., Cao, D., Huang, Y., Parameswaran, P., Fire, A.Z., and Kay, M.A. (2008). Capped small RNAs and MOV10 in human hepatitis delta virus replication. Nat Struct Mol Biol *15*, 714-721.

Hirsh, D., and Vanderslice, R. (1976). Temperature-sensitive developmental mutants of Caenorhabditis elegans. Dev Biol *49*, 220-235.

Houwing, S., Kamminga, L.M., Berezikov, E., Cronembold, D., Girard, A., van den Elst, H., Filippov, D.V., Blaser, H., Raz, E., Moens, C.B., *et al.* (2007). A role for Piwi and piRNAs in germ cell maintenance and transposon silencing in Zebrafish. Cell *129*, 69-82.

Hutvagner, G., McLachlan, J., Pasquinelli, A.E., Balint, E., Tuschl, T., and Zamore, P.D. (2001). A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. Science *293*, 834-838.

Hutvagner, G., and Simard, M.J. (2008). Argonaute proteins: key players in RNA silencing. Nat Rev Mol Cell Biol *9*, 22-32.

Hutvagner, G., and Zamore, P.D. (2002). A microRNA in a multiple-turnover RNAi enzyme complex. Science *297*, 2056-2060.

Ibanez-Ventoso, C., Vora, M., and Driscoll, M. (2008). Sequence relationships among C. elegans, D. melanogaster and human microRNAs highlight the extensive conservation of microRNAs in biology. PLoS One *3*, e2818.

Jablonka, E., and Raz, G. (2009). Transgenerational epigenetic inheritance: prevalence, mechanisms, and implications for the study of heredity and evolution. Q Rev Biol *84*, 131-176.

Jannes, P., Spiessens, C., Van der Auwera, I., D'Hooghe, T., Verhoeven, G., and Vanderschueren, D. (1998). Male subfertility induced by acute scrotal heating affects embryo quality in normal female mice. Hum Reprod *13*, 372-375.

Joshua-Tor, L., and Hannon, G.J. (2011). Ancestral roles of small RNAs: an Ago-centric perspective. Cold Spring Harb Perspect Biol *3*, a003772.

Kamath, R.S., Fraser, A.G., Dong, Y., Poulin, G., Durbin, R., Gotta, M., Kanapin, A., Le Bot, N., Moreno, S., Sohrmann, M., *et al.* (2003). Systematic functional analysis of the Caenorhabditis elegans genome using RNAi. Nature *421*, 231-237.

Kamminga, L.M., Luteijn, M.J., den Broeder, M.J., Redl, S., Kaaij, L.J., Roovers, E.F., Ladurner, P., Berezikov, E., and Ketting, R.F. (2010). Hen1 is required for oocyte development and piRNA stability in zebrafish. Embo J *29*, 3688-3700.

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

Kawamura, Y., Saito, K., Kin, T., Ono, Y., Asai, K., Sunohara, T., Okada, T.N., Siomi, M.C., and Siomi, H. (2008). Drosophila endogenous small RNAs bind to Argonaute 2 in somatic cells. Nature *453*, 793-797.

Kawasaki, I., Shim, Y.H., Kirchner, J., Kaminker, J., Wood, W.B., and Strome, S. (1998). PGL-1, a predicted RNA-binding component of germ granules, is essential for fertility in C. elegans. Cell *94*, 635-645.

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.

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., Fischer, S.E., Bernstein, E., Sijen, T., Hannon, G.J., and Plasterk, R.H. (2001). Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in C. elegans. Genes Dev *15*, 2654-2659.

Khvorova, A., Reynolds, A., and Jayasena, S.D. (2003). Functional siRNAs and miRNAs exhibit strand bias. Cell *115*, 209-216.

Kirino, Y., and Mourelatos, Z. (2007a). The mouse homolog of HEN1 is a potential methylase for Piwi-interacting RNAs. RNA *13*, 1397-1401.

Kirino, Y., and Mourelatos, Z. (2007b). Mouse Piwi-interacting RNAs are 2'-O-methylated at their 3' termini. Nat Struct Mol Biol *14*, 347-348.

Klattenhoff, C., and Theurkauf, W. (2008). Biogenesis and germline functions of piRNAs. Development *135*, 3-9.

Knight, S.W., and Bass, B.L. (2001). A role for the RNase III enzyme DCR-1 in RNA interference and germ line development in Caenorhabditis elegans. Science *293*, 2269-2271.

Kotaja, N., Bhattacharyya, S.N., Jaskiewicz, L., Kimmins, S., Parvinen, M., Filipowicz, W., and Sassone-Corsi, P. (2006). The chromatoid body of male germ cells: similarity with processing bodies and presence of Dicer and microRNA pathway components. Proc Natl Acad Sci U S A *103*, 2647-2652.

Krawetz, S.A., Kruger, A., Lalancette, C., Tagett, R., Anton, E., Draghici, S., and Diamond, M.P. (2011). A survey of small RNAs in human sperm. Hum Reprod *26*, 3401-3412.

Kuramochi-Miyagawa, S., Kimura, T., Ijiri, T.W., Isobe, T., Asada, N., Fujita, Y., Ikawa, M., Iwai, N., Okabe, M., Deng, W., *et al.* (2004). Mili, a mammalian member of piwi family gene, is essential for spermatogenesis. Development *131*, 839-849.

Kuramochi-Miyagawa, S., Watanabe, T., Gotoh, K., Totoki, Y., Toyoda, A., Ikawa, M., Asada, N., Kojima, K., Yamaguchi, Y., Ijiri, T.W., *et al.* (2008). DNA methylation of retrotransposon genes is regulated by Piwi family members MILI and MIWI2 in murine fetal testes. Genes Dev *22*, 908-917.

Kuznicki, K.A., Smith, P.A., Leung-Chiu, W.M., Estevez, A.O., Scott, H.C., and Bennett, K.L. (2000). Combinatorial RNA interference indicates GLH-4 can compensate for

GLH-1; these two P granule components are critical for fertility in C. elegans. Development *127*, 2907-2916.

Lagos-Quintana, M., Rauhut, R., Lendeckel, W., and Tuschl, T. (2001). Identification of novel genes coding for small expressed RNAs. Science *294*, 853-858.

Lai, E.C. (2002). Micro RNAs are complementary to 3' UTR sequence motifs that mediate negative post-transcriptional regulation. Nat Genet *30*, 363-364.

Lau, N.C. (2010). Small RNAs in the animal gonad: guarding genomes and guiding development. Int J Biochem Cell Biol *42*, 1334-1347.

Lau, N.C., Lim, L.P., Weinstein, E.G., and Bartel, D.P. (2001). An abundant class of tiny RNAs with probable regulatory roles in Caenorhabditis elegans. Science *294*, 858-862.

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.

Lee, R.C., and Ambros, V. (2001). An extensive class of small RNAs in Caenorhabditis elegans. Science *294*, 862-864.

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

Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., Lee, J., Provost, P., Radmark, O., Kim, S., *et al.* (2003). The nuclear RNase III Drosha initiates microRNA processing. Nature *425*, 415-419.

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

Lee, Y., Kim, M., Han, J., Yeom, K.H., Lee, S., Baek, S.H., and Kim, V.N. (2004). MicroRNA genes are transcribed by RNA polymerase II. EMBO J *23*, 4051-4060.

L'Hernault, S.W. (2006). Spermatogenesis. WormBook, 1-14.

Li, C., Vagin, V.V., Lee, S., Xu, J., Ma, S., Xi, H., Seitz, H., Horwich, M.D., Syrzycka, M., Honda, B.M., *et al.* (2009). Collapse of germline piRNAs in the absence of Argonaute3 reveals somatic piRNAs in flies. Cell *137*, 509-521.

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, L.P., Lau, N.C., Weinstein, E.G., Abdelhakim, A., Yekta, S., Rhoades, M.W., Burge, C.B., and Bartel, D.P. (2003). The microRNAs of Caenorhabditis elegans. Genes Dev *17*, 991-1008.

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.

Liu, W.M., Pang, R.T., Chiu, P.C., Wong, B.P., Lao, K., Lee, K.F., and Yeung, W.S. (2012). Sperm-borne microRNA-34c is required for the first cleavage division in mouse. Proc Natl Acad Sci U S A *109*, 490-494.

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.

Maher, B. (2008). Personal genomes: The case of the missing heritability. Nature *456*, 18-21.

Makarova, K.S., Wolf, Y.I., van der Oost, J., and Koonin, E.V. (2009). Prokaryotic homologs of Argonaute proteins are predicted to function as key components of a novel system of defense against mobile genetic elements. Biol Direct *4*, 29.

Malone, C.D., Brennecke, J., Dus, M., Stark, A., McCombie, W.R., Sachidanandam, R., and Hannon, G.J. (2009). Specialized piRNA pathways act in germline and somatic tissues of the Drosophila ovary. Cell *137*, 522-535.

Martinez, J., Patkaniowska, A., Urlaub, H., Luhrmann, R., and Tuschl, T. (2002). Single-stranded antisense siRNAs guide target RNA cleavage in RNAi. Cell *110*, 563-574.

Massirer, K.B., Perez, S.G., Mondol, V., and Pasquinelli, A.E. (2012). The miR-35-41 family of microRNAs regulates RNAi sensitivity in Caenorhabditis elegans. PLoS Genet *8*, e1002536.

McJunkin, K., and Ambros, V. (2014). The Embryonic mir-35 Family of MicroRNAs Promotes Multiple Aspects of Fecundity in C. elegans. G3 (Bethesda).

Meister, G. (2013). Argonaute proteins: functional insights and emerging roles. Nat Rev Genet *14*, 447-459.

Mieusset, R., Quintana Casares, P.I., Sanchez-Partida, L.G., Sowerbutts, S.F., Zupp, J.L., and Setchell, B.P. (1991). The effects of moderate heating of the testes and epididymides of rams by scrotal insulation on body temperature, respiratory rate, spermatozoa output and motility, and on fertility and embryonic survival in ewes inseminated with frozen semen. Ann N Y Acad Sci *637*, 445-458.

Miller, M.A. (2006). Sperm and oocyte isolation methods for biochemical and proteomic analysis. Methods Mol Biol *351*, 193-201.

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

Nechaev, S., Fargo, D.C., dos Santos, G., Liu, L., Gao, Y., and Adelman, K. (2010). Global analysis of short RNAs reveals widespread promoter-proximal stalling and arrest of Pol II in Drosophila. Science *327*, 335-338.

Nishibu, T., Hayashida, Y., Tani, S., Kurono, S., Kojima-Kita, K., Ukekawa, R., Kurokawa, T., Kuramochi-Miyagawa, S., Nakano, T., Inoue, K., *et al.* (2012). Identification of MIWI-associated Poly(A) RNAs by immunoprecipitation with an anti-MIWI monoclonal antibody. Biosci Trends *6*, 248-261.

Okamura, K., Chung, W.J., Ruby, J.G., Guo, H., Bartel, D.P., and Lai, E.C. (2008). The Drosophila hairpin RNA pathway generates endogenous short interfering RNAs. Nature *453*, 803-806.

Orgel, L.E. (1968). Evolution of the genetic apparatus. J Mol Biol *38*, 381-393.

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

Pasquinelli, A.E., Reinhart, B.J., Slack, F., Martindale, M.Q., Kuroda, M.I., Maller, B., Hayward, D.C., Ball, E.E., Degnan, B., Muller, P., *et al.* (2000). Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. Nature *408*, 86-89.

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

Pelisson, A., Song, S.U., Prud'homme, N., Smith, P.A., Bucheton, A., and Corces, V.G. (1994). Gypsy transposition correlates with the production of a retroviral envelope-like protein under the tissue-specific control of the Drosophila flamenco gene. EMBO J *13*, 4401-4411.

Peng, H., Shi, J., Zhang, Y., Zhang, H., Liao, S., Li, W., Lei, L., Han, C., Ning, L., Cao, Y., *et al.* (2012). A novel class of tRNA-derived small RNAs extremely enriched in mature mouse sperm. Cell Res *22*, 1609-1612.

Praitis, V., Casey, E., Collar, D., and Austin, J. (2001). Creation of low-copy integrated transgenic lines in Caenorhabditis elegans. Genetics *157*, 1217-1226.

Rando, O.J. (2012). Daddy issues: paternal effects on phenotype. Cell *151*, 702-708.
Reinke, V., Gil, I.S., Ward, S., and Kazmer, K. (2004). Genome-wide germline-enriched and sex-biased expression profiles in Caenorhabditis elegans. Development *131*, 311-323.

Rathke, C., Baarends, W.M., Awe, S., and Renkawitz-Pohl, R. (2014). Chromatin dynamics during spermiogenesis. Biochim Biophys Acta *1839*, 155-168.

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

Reinhart, B.J., Weinstein, E.G., Rhoades, M.W., Bartel, B., and Bartel, D.P. (2002). MicroRNAs in plants. Genes Dev *16*, 1616-1626.

Reinke, V., Gil, I.S., Ward, S., and Kazmer, K. (2004). Genome-wide germline-enriched and sex-biased expression profiles in Caenorhabditis elegans. Development *131*, 311-323.

Robine, N., Lau, N.C., Balla, S., Jin, Z., Okamura, K., Kuramochi-Miyagawa, S., Blower, M.D., and Lai, E.C. (2009). A broadly conserved pathway generates 3'UTR-directed primary piRNAs. Curr Biol *19*, 2066-2076.

Rockett, J.C., Mapp, F.L., Garges, J.B., Luft, J.C., Mori, C., and Dix, D.J. (2001). Effects of hyperthermia on spermatogenesis, apoptosis, gene expression, and fertility in adult male mice. Biol Reprod *65*, 229-239.

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.

Ruby, J.G., Jan, C.H., and Bartel, D.P. (2007). Intronic microRNA precursors that bypass Drosha processing. Nature *448*, 83-86.

Saito, K., Ishizuka, A., Siomi, H., and Siomi, M.C. (2005). Processing of pre-microRNAs by the Dicer-1-Loquacious complex in Drosophila cells. PLoS Biol *3*, e235.

Saito, K., Nishida, K.M., Mori, T., Kawamura, Y., Miyoshi, K., Nagami, T., Siomi, H., and Siomi, M.C. (2006). Specific association of Piwi with rasiRNAs derived from retrotransposon and heterochromatic regions in the Drosophila genome. Genes Dev *20*, 2214-2222.

Saito, K., Sakaguchi, Y., Suzuki, T., Siomi, H., and Siomi, M.C. (2007). Pimet, the Drosophila homolog of HEN1, mediates 2'-O-methylation of Piwi- interacting RNAs at their 3' ends. Genes Dev *21*, 1603-1608.

Sakai, N., Sawada, M.T., and Sawada, H. (2004). Non-traditional roles of ubiquitin-proteasome system in fertilization and gametogenesis. Int J Biochem Cell Biol *36*, 776-784.

Schedl, T., and Kimble, J. (1988). fog-2, a germ-line-specific sex determination gene required for hermaphrodite spermatogenesis in Caenorhabditis elegans. Genetics *119*, 43-61.

Schwarz, D.S., Hutvagner, G., Du, T., Xu, Z., Aronin, N., and Zamore, P.D. (2003). Asymmetry in the assembly of the RNAi enzyme complex. Cell *115*, 199-208.

Schwarz, D.S., Hutvagner, G., Haley, B., and Zamore, P.D. (2002). Evidence that siRNAs function as guides, not primers, in the Drosophila and human RNAi pathways. Mol Cell *10*, 537-548.

Seila, A.C., Calabrese, J.M., Levine, S.S., Yeo, G.W., Rahl, P.B., Flynn, R.A., Young, R.A., and Sharp, P.A. (2008). Divergent transcription from active promoters. Science *322*, 1849-1851.

Setchell, B.P., D'Occhio, M.J., Hall, M.J., Laurie, M.S., Tucker, M.J., and Zupp, J.L. (1988). Is embryonic mortality increased in normal female rats mated to subfertile males? J Reprod Fertil *82*, 567-574.

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.

Shakes, D.C., and Ward, S. (1989). Initiation of spermiogenesis in C. elegans: a pharmacological and genetic analysis. Developmental biology *134*, 189-200.

Shakes, D.C., Wu, J.C., Sadler, P.L., Laprade, K., Moore, L.L., Noritake, A., and Chu, D.S. (2009). Spermatogenesis-specific features of the meiotic program in Caenorhabditis elegans. PLoS Genet *5*, e1000611.

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

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.

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.

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.

Stoeckius, M., Grun, D., and Rajewsky, N. (2014). Paternal RNA contributions in the Caenorhabditis elegans zygote. EMBO J *33*, 1740-1750.

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.

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

Tam, O.H., Aravin, A.A., Stein, P., Girard, A., Murchison, E.P., Cheloufi, S., Hodges, E., Anger, M., Sachidanandam, R., Schultz, R.M., *et al.* (2008). Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes. Nature *453*, 534-538.

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., and Fire, A. (1998). Specific interference by ingested dsRNA. Nature *395*, 854.

Tolia, N.H., and Joshua-Tor, L. (2007). Slicer and the argonautes. Nat Chem Biol *3*, 36-43.

Updike, D., and Strome, S. (2010). P granule assembly and function in Caenorhabditis elegans germ cells. J Androl *31*, 53-60.

Vagin, V.V., Sigova, A., Li, C., Seitz, H., Gvozdev, V., and Zamore, P.D. (2006). A distinct small RNA pathway silences selfish genetic elements in the germline. Science *313*, 320-324.

Van Ex, F., Jacob, Y., and Martienssen, R.A. (2011). Multiple roles for small RNAs during plant reproduction. Curr Opin Plant Biol *14*, 588-593.

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.

Voronina, E., Seydoux, G., Sassone-Corsi, P., and Nagamori, I. (2011). RNA granules in germ cells. Cold Spring Harb Perspect Biol *3*.

Vourekas, A., Zheng, Q., Alexiou, P., Maragkakis, M., Kirino, Y., Gregory, B.D., and Mourelatos, Z. (2012). Mili and Miwi target RNA repertoire reveals piRNA biogenesis and function of Miwi in spermiogenesis. Nat Struct Mol Biol *19*, 773-781.

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

Wang, Y., Juranek, S., Li, H., Sheng, G., Tuschl, T., and Patel, D.J. (2008a). Structure of an argonaute silencing complex with a seed-containing guide DNA and target RNA duplex. Nature *456*, 921-926.

Wang, Y., Sheng, G., Juranek, S., Tuschl, T., and Patel, D.J. (2008b). Structure of the guide-strand-containing argonaute silencing complex. Nature *456*, 209-213.

Ward, S., and Miwa, J. (1978). Characterization of temperature-sensitive, fertilization-defective mutants of the nematode caenorhabditis elegans. Genetics *88*, 285-303.

Ward, S., Argon, Y., and Nelson, G.A. (1981). Sperm morphogenesis in wild-type and fertilization-defective mutants of Caenorhabditis elegans. J Cell Biol *91*, 26-44.

Washburn, M.P., Wolters, D., and Yates, J.R., 3rd (2001). Large-scale analysis of the yeast proteome by multidimensional protein identification technology. Nat Biotechnol *19*, 242-247.

Washington, N.L., and Ward, S. (2006). FER-1 regulates Ca2+ -mediated membrane fusion during C. elegans spermatogenesis. J Cell Sci *119*, 2552-2562.

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

Watanabe, T., Totoki, Y., Toyoda, A., Kaneda, M., Kuramochi-Miyagawa, S., Obata, Y., Chiba, H., Kohara, Y., Kono, T., Nakano, T., *et al.* (2008). Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. Nature *453*, 539-543.

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., 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.

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

Werdelin, L., and Nilsonne, A. (1999). The evolution of the scrotum and testicular descent in mammals: a phylogenetic view. J Theor Biol *196*, 61-72.

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

Woese, C.R. (1967). The genetic code; the molecular basis for genetic expression (New York,, Harper & Row).

Wolters, D.A., Washburn, M.P., and Yates, J.R., 3rd (2001). An automated multidimensional protein identification technology for shotgun proteomics. Anal Chem *73*, 5683-5690.

Wu, J.C., Go, A.C., Samson, M., Cintra, T., Mirsoian, S., Wu, T.F., Jow, M.M., Routman, E.J., and Chu, D.S. (2012). Sperm development and motility are regulated by PP1 phosphatases in Caenorhabditis elegans. Genetics *190*, 143-157

Wu, E., Thivierge, C., Flamand, M., Mathonnet, G., Vashisht, A.A., Wohlschlegel, J., Fabian, M.R., Sonenberg, N., and Duchaine, T.F. (2010). Pervasive and cooperative deadenylation of 3'UTRs by embryonic microRNA families. Mol Cell *40*, 558-570.

Wykes, S.M., and Krawetz, S.A. (2003). The structural organization of sperm chromatin. J Biol Chem *278*, 29471-29477.

Yekta, S., Shih, I.H., and Bartel, D.P. (2004). MicroRNA-directed cleavage of HOXB8 mRNA. Science *304*, 594-596.

Yi, R., Qin, Y., Macara, I.G., and Cullen, B.R. (2003). Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. Genes Dev *17*, 3011-3016.

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.

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.