SURVEILLANCE AND TARGETING OF ABERRANT TRANSCRIPTS IN *CAENORHABDITIS ELEGANS*

Luciana E. Leopold

A dissertation submitted to the faculty at the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Curriculum in Genetics and Molecular Biology in the Graduate School.

Chapel Hill
2014

Approved by:

Shawn Ahmed

William Marzluff

Frank Conlon

Victoria Bautch

David Reiner

©2014
Luciana E. Leopold
ALL RIGHTS RESERVED

ABSTRACT

Luciana E. Leopold: Surveillance and Targeting of Aberrant Transcripts
in *Caenorhabditis elegans*
(Under the direction of Shawn Ahmed)

An organism requires mechanisms with which to protect the fidelity and integrity of the genome, transcriptome, and proteome in order not only to remain healthy and alive, but also to pass on viable genetic material to its offspring. To this end, the cell is constantly surveying the transcripts being expressed. When aberrant transcripts are produced, whether they be from endogenous or exogenous sources, mechanisms are required in order to silence or eradicate such messages before they become toxic to the cell. Two such mechanisms are nonsense-mediated mRNA decay (NMD), which targets transcripts bearing a premature termination codon (PTC), and RNA-induced epigenetic silencing (RNAe), which silences foreign loci in the genome in a heritable, stable manner.

We have identified targets that require unique kinds of silencing that utilize already established NMD or RNAe machinery, but represent deviations from the standard. For example, the *unc-54(r293)* transcript was previously thought to be silenced by standard NMD, however I have shown that *unc-54(r293)* does not contain the canonical target PTC, and instead has a more complex molecular target. I also show that the integral SMG-1 component of the NMD machinery interacts with this target in a manner that is distinct from that of traditional NMD, as illustrated by the novel *smg-1(yp3)* separation-of-function

allele. In the next example, we show that permanent, stable transgene silencing like that generated by RNAe in a single generation, can be attained in certain transgenes only after several generations of successive crossing. These processes are molecularly similar, as they both require the Argonaute protein PRG-1. However, they differ in that single-generation RNAe silencing becomes independent of PRG-1 after initiation, whereas the silencing we observe in multigenerational RNAe requires PRG-1 in order to maintain its silent state.

ACKNOWLEDGEMENTS

To my family & friends – thank you for putting up with me during this process. I know it hasn't been easy, but I am very lucky to have so many people in my life who support me. I am grateful for all of you every single day. To my labmates, who are like family and friends combined – thank you for your constant teamwork, I would never have made it without you. To my mentor – Shawn, you are a brilliant man full of ideas. I hope you never lose your zest for science. Finally, to perseverance.

TABLE OF CONTENTS

ABSTRACT iii

ACKNOWLEDGEMENTS v

LIST OF ABBREVIATIONS x

INTRODUCTION 1

Figures 9

CHAPTER 1: *SMG-1(YP3)* TARGETS UNUSUAL TRANSCRIPTS 13

Background 13

Results 15

Discussion 23

Materials and Methods 27

Figures 30

CHAPTER 2: LACK OF PAIRING DURING MEIOSIS TRIGGERS MULTIGENERATIONAL TRANSGENE SILENCING IN *CAENORHABDITIS ELEGANS* 37

Background 37

Results 40

Discussion 48

Materials and Methods 52

Figures 55

CHAPTER 3: SUMMARY & FUTURE DIRECTIONS.....63

REFERENCES.....69

LIST OF FIGURES

Figure 0.1. Viruses are capable of injecting their own genetic material into the cell 9

Figure 0.2. The mechanism RNA interference for silencing induced by small dsRNA. 10

Figure 0.3. The RNAe silencing mechanism of the piRNA sensor. 11

Figure 0.4. Silencing mechanisms located in the cytoplasm and the nucleus. 12

Figure 1.1. *smg-1(yp3)* displays several unique phenotypes 30

Figure 1.2. Identification of mutations in *smg-1* alleles and corresponding residue conservation 31

Figure 1.3. *unc-54(r293)* serves as an indicator for NMD function in smg-1 mutants 32

Figure 1.4. The *unc-54(r293)* message is bicistronic 33

Figure 1.5. Structure and expression of the pCM1.4 transgene. 34

Figure 1.6. Examination of NMD activity in *smg-1* mutant backgrounds 35

Figure 1.7. Model displaying the possible separation-of-function occurring in *smg-1(yp3)* due to different mRNA targets.. 36

Figure 2.1. mCherry-expressing transgenes are capable of being silenced in a multigenerational manner via crossing. 55

Figure 2.2. GFP-expressing transgenes are capable of being silenced in a transgenerational manner via crossing 56

Figure 2.3. *prg-1* and *rde-2* coordinately promote maintenance of the silencing of the *pot-1::mCherry* transgene 57

Figure 2.4. Model depicting (A) single generation RNAe, as previously described, and (B) multigenerational RNAe. 58

Figure 2.5. Widefield microscopy images of bright field (left column), GFP fluorescence (center column), and a merged (right column). 59

Figure 2.6. Crossing with various markers induces silencing of *cpSi10* and *cpIs10* transgenes 60

Figure 2.7. Schema used to cross the *his-72*::GFP transgene using Chromosome III markers *vab-7 dpy-18*. 61

Figure 2.8. Widefield microscopy depicting bright field (left column) and mCherry fluorescent (right column) images of (top-bottom) in the zim-1 mutant in the pot-1::mCherry transgenic background 62

LIST OF ABBREVIATIONS

bp	base pairs
DMD	Duchenne muscular Dystrophy
dsRNA	double-stranded RNA
EMS	ethyl methanesulfonate
F1/F2/F3	filial generation 1/2/3
GFP	green fluorescent protein
miRNA	micro RNA
MSUD	meiotic silencing by unpaired DNA
mRNA	messenger RNA
N2	wildtype Bristol strain of *C. elegans*
NMD	nonsense-mediated mRNA decay
ORF	open reading frame
PI3K	phosphatidylinositol 3-kinase
piRNA	piwi-interacting RNA
POT1	protection-of-telomeres 1
PTC	premature termination codon
Pvl	protruding vulva phenotype
RdRP	RNA-dependent RNA polymerase
RNAe	RNA-induced epigenetic silencing
RNAi	RNA interference
rtPCR	reverse transcription polymerase chain reaction

siRNA	small interfering RNA
snRNA	small nuclear RNA
snoRNA	small nucleolar RNA
SL1	splice leader 1
SMG1	suppressor with morphological effect on genetalia 1
uORF	upstream open reading frame
UTR	untranslated region

INTRODUCTION

In order for an organism to remain viable and fulfill its ultimate evolutionary purpose of passing along its genetic code, it must be able to protect both itself and the code from harm. While behavioral instincts like the 'fight-or-flight' response, or the ability to camouflage, were being programmed in via evolution across many millennia, molecular protective mechanisms were also emerging. Cells are equipped with numerous pathways that work to ensure the maintenance of genome, transcriptome, and proteome integrity. Each level of information – DNA, RNA, and protein – is vulnerable to attack from various exogenous sources, such as viruses and radiation, as well as endogenous errors that can occur during transcription and translation. Without protective protocols in place, these attacks can be toxic, and have the potential to cause disease and even death. For example, many cancers are caused by mutations in oncogenes resulting from transcriptional errors.

A key subset of surveillance and subsequent protection mechanisms concern the stability and integrity of RNA. RNAs are vital to the cell and serve myriad purposes. The best understood RNAs are perhaps messenger RNAs (mRNAs), which carry information from the DNA and undergo translation in order to produce all of the functional proteins in the cell. However, RNAs serve many other purposes such as endogenous gene regulation by micro RNAs (miRNAs) (1), protection from exogenous double-stranded RNAs (dsRNAs) as facilitated by (small interfering RNAs) siRNAs (2), regulation of retrotransposons via piwi-interacting RNAs (piRNAs) (3), post-transcriptional modifications like splicing which

requires small nuclear RNAs (snRNAs) and RNA nucleotide changes mediated by small nucleolar RNAs (snoRNAs) (4, 5). The ends of eukaryotic chromosomes are maintained by telomerase, which requires an RNA template to replicate the telomeric sequence (6). Many of these pathways have been very well described in both their function and in the identities of the protein players involved. However, it is logical to believe that there may exist variations and nuances of these pathways in order to achieve the level of specificity required to regulate all of the possible targets. Herein I discuss two different projects that concern RNA stability and investigate mechanisms of RNA surveillance and defense.

In Chapter 1, I address a pathway related to nonsense-mediated decay (NMD), which is a method the cell has for eradicating endogenous abberantly processed mRNAs in the cytoplasm. NMD is a pathway that is highly conserved among eukaryotes and canonically identifies RNA transcripts bearing a premature termination codon (PTC) and targets them for degradation, thus alleviating the cell of any toxic stress that may result from a faulty transcript (7). These errors may arise either by a mutation in the DNA source code, or by inaccurate processing. In *Caenorhabditis elegans*, many of the components of NMD have been identified, notably the seven *smg* genes (8). The SMG-2 (Upf1) helicase serves as the first alarm bell of NMD and identifies transcripts containing a PTC. SMG-2 is then phosphorylated by SMG-1 and downstream activators of NMD are triggered, and various decay pathways come into play to degrade the message (9, 10).

In humans, diseases like Duchenne Muscular Dystrophy (DMD) (11) can be caused by nonsense mutations that result in an intermediate stop codon. The subsequent mutant transcripts are degraded by NMD, and no viable protein is made, causing the disease phenotype. However, NMD and/or its machinery has been implicated in a much broader

range of protective mechanisms as well. For instance, NMD components like the SMG/UPF proteins have been shown to be involved in responses to damage induced by UV or γ-radiation (12), double strand breaks (13), as well as the oxidative stress response to TNF-α-induced apoptosis (14).

I specifically investigate an NMD-related pathway that targets the mRNA transcript of *unc-43(r293)*, which in the past has been thought of as a canonical target of NMD. However, upon observation, it is clear that the molecular structure of the *unc-54(r293)* RNA transcript is not that of a standard RNA with a PTC. Rather, *unc-54(r293)* contains an extra-long 3' untranslated region (3'UTR) which reads through to the next gene, *aex-5,* and may possibly result in a bicistronic message. We have isolated a novel allele of the *smg-1* gene, *smg-1(yp3).* While *smg-1(yp3)* functions like other known *smg-1* alleles to stabilize messages carrying a standard PTC, a long 3'UTR, and even a bicistron, it still degrades the *unc-43(r293)* transcript, and thus suppresses its function. Therefore, *smg-1(yp3)* is a separation-of-function allele that defines a pathway that is related to NMD, but is not identical in the way that it processes this specific transcript. I believe that there is possibly a unique molecular trigger contained within the *unc-54(r293)* transcript that requires a variation adjustment in the way that SMG-1 is utilized by the NMD process in order to be target the transcript for degradation.

In Chapter 2, I investigate how the cell handles foreign intruders in its DNA. The cell needs to keep its DNA intact and functional in order to a) keep on living disease-free and b) pass on its genes properly. In order to do this it has many different methods of defense against a variety of exogenous attackers, such as viruses. Viruses inject their own genetic material into the cell, which, in the case of some viruses, is inserted into the host DNA and

hijacks the cell for its own purposes, and uses the native machinery to transcribe its own RNA, which is then translocated to the cytoplasm to be made into viral protein (**Figure 0.1**). In order to protect itself from foreign sequences or other rogue elements in the DNA, the cell must have some kind of immunity, and it is only more recently that we are really starting uncover how this is done. My work makes several significant contributions towards understanding how the cell identifies and silences foreign intruders.

mRNA is transcribed from the DNA in the nucleus, but is then translated into protein in the cytoplasm, so at any given time, foreign RNA species can be residing in either of these compartments (**Figure 0.1**). In order to completely rid itself of the threat of foreign DNA taking over the cell, causing chaos, and maybe even disease, the cell must find ways of degrading unwanted RNAs that is found in both locations.

There is a known mechanism which targets and degrades foreign cytoplasmic RNA, called RNA interference, or RNAi (2) (**Figure 0.2**). Fire and Mello won the Nobel Prize in 2006 for the discovery of RNAi in *C. elegans*. They uncovered how foreign double stranded RNA in the cytoplasm – which could potentially come from something like a virus, or something injected in the lab, as Fire and Mello did – is silenced by cellular machinery. First, Dicer, an endoribonuclease, binds to this dsRNA with the help of RDE-4, a protein known to bind dsRNA, and the RNA is diced into smaller pieces, called siRNAs, which are made single-stranded and loaded onto the RDE-1 Argonaute protein, which is the catalytic component of RISC, a complex of multiple proteins that incorporates this small RNA and is capable of inducing silencing by binding to target mRNA and cleaving it, essentially degrading it and rendering it useless, so the RNA expression of the foreign gene in the cytoplasm is silenced (15-18) (**Figure 0.2**).

This is how the cell takes care of foreign mRNA found in the cytoplasm, but there is another mechanism which targets the mRNA still found in the nucleus. The identification of this process, called RNAe, or RNA-induced epigenetic silencing, is a newer advance in the field, as it was only discovered in 2012 (**Figure 0.3**). A series of papers from three different groups describing this pathway were published in close succession (19-21). In these studies, a process similar to RNAi was outlined, the difference being that this pathway has the potential to epigenetically mark the DNA itself for silencing, preventing transcription of foreign DNA not only in that particular cell, but any cell it passes its genetic code onto. These groups all saw that the silencing was irreversible and occurred in a single generation. Many of the components of RNAe are familiar as there are several similar players in RNAi (**Figure 0.4**). I had observed a similar phenomenon and when these studies came out, I realized this mechanism might be related somehow.

Notably the the specific molecular trigger for this pathway is unknown, but it is known that piRNAs, which are expressed in clusters throughout the genome interact with the specialized Argonaute PRG-1, which participates in the first phase of RNAe, called initiation (20, 22). Over 10000 piRNAs are expressed in the cell and have a broad range of gene silencing functions, the best understood of which is probably the silencing of transposons or transgenes (23, 24). piRNAs are made from specific sites in the genome. The piRNAs are loaded onto the PRG-1 Argonaute, and then binds target RNAs. An RNA-dependent RNA polymerase (RdRP) moves along the sequence nand transcribes small secondary RNA intermediates from the target sequence, in this case the RNA of the piRNA sensor, a transgene containing an engineered piRNA-binding site. These small RNAs are then loaded onto a secondary Argonaute, and are able to silence gene expression with the

help of other factors like MUT-7, a putative exoribonuclease, and its binding partner RDE-2 (**Figure 0.3**).

What makes this pathway particularly unique is that silencing doesn't simply stop at degrading the RNA message, but ultimately results in epigenetic marks, which prevent the transcription of the foreign DNA from happening in the first place. These marks are heritable, so even the progeny of an animal with these marks would essentially "remember" this foreign intruder and silence it. It is important to note that these transgenes have a piRNA binding site, and that this process happens in a single generation, and results in an irreversible state of silencing, which requires PRG-1 for initiation, but it is dispensible for the maintenance phase.

It has been well described in the literature that PRG-1 is a piwi-specific Argonaute that binds to piRNAs, and is capable of silencing transgenes that were used in the three studies mentioned above, via this RNAe mechanism. The idea behind these transgene studies is that they serve as a model for how the cell may cope with foreign genes, like those from viruses, inserted into the genome. However, the transgenes that they used contained a piRNA sensor, which has sites of exact homology to piRNAs being made in the cell (**Figure 0.3**). It is very likely that this specific binding of the piRNA to the piRNA sensor is what is inducing the single-generation silencing witnessed in these previous studies, but not every invader will have a piRNA binding site. The pathway that I investigate here also implicates PRG-1 in this silencing, even though our transgenes do not contain a homologous piRNA sensor. Thus, our findings extend the studies from the other groups, and provide an explanation of how an organism ultimately develops immunity over several generations.

To summarize, the cell needs two forms of defense, one in the cytoplasm, where mRNA is waiting to be translated into protein, and one in the nucleus where RNA is being actively transcribed from the DNA. The pathway that identifies and degrades mRNA targets in the cytoplasm is called RNAi. The pathway that targets transcripts in the nucleus, and is capable of adding silencing marks to prevent mRNA from even being transcribed from the DNA in the first place, in addition to the degradation in the cytoplasm, is called RNAe (**Figure 0.4**). In Chapter 2, my investigation identifies a new RNAe pathway called multigenerational RNAe, which is distinct from the single-generation RNAe described by other groups. Firstly, in single-generation RNAe, silencing occurs in a single generation, whereas in our multigenerational RNAe process, it takes multiple generations of crossing in order to induce silencing, presumably because there are not piRNA sites in the these genes. In addition, their mechanism shows clearly that PRG-1 is involved only in the initiation of silencing, and is not required for maintenance. What I see on the other hand is that PRG-1 is involved in both the initiation and maintenance and that those two steps are not separable, but linked. The silencing that they see is permanent and irreversible. I find that our silencing is often reversible under certain conditions. Lastly, the molecular trigger for RNAe is unknown, but I have evidence that the molecular trigger of our multigenerational mechanism may involve a lack of chromosomal pairing and demonstrated by silencing in the *zim-1* mutant background.

In summary, the identification and eradication of unwanted sequences and transcripts is vital to an organism. There are many complex RNA processing systems that are customized for different kinds of targets, whether they be endogenous errors or foreign interlopers. Herein I describe two unique RNA surveillance pathways. Our findings

represent advances in the understanding of how vital, conserved processes like NMD and RNAi are able to tailor their machinery to the silencing of various diverse targets. The pathway outlined in Chapter 1 protects against endogenous genetic errors and the processing of this RNA, while the pathway illustrated in Chapter 2 silences foreign transgenic sequences and has the potential to integrate this silencing in a stable heritable fashion, possibly via epigenetic changes.

Figures

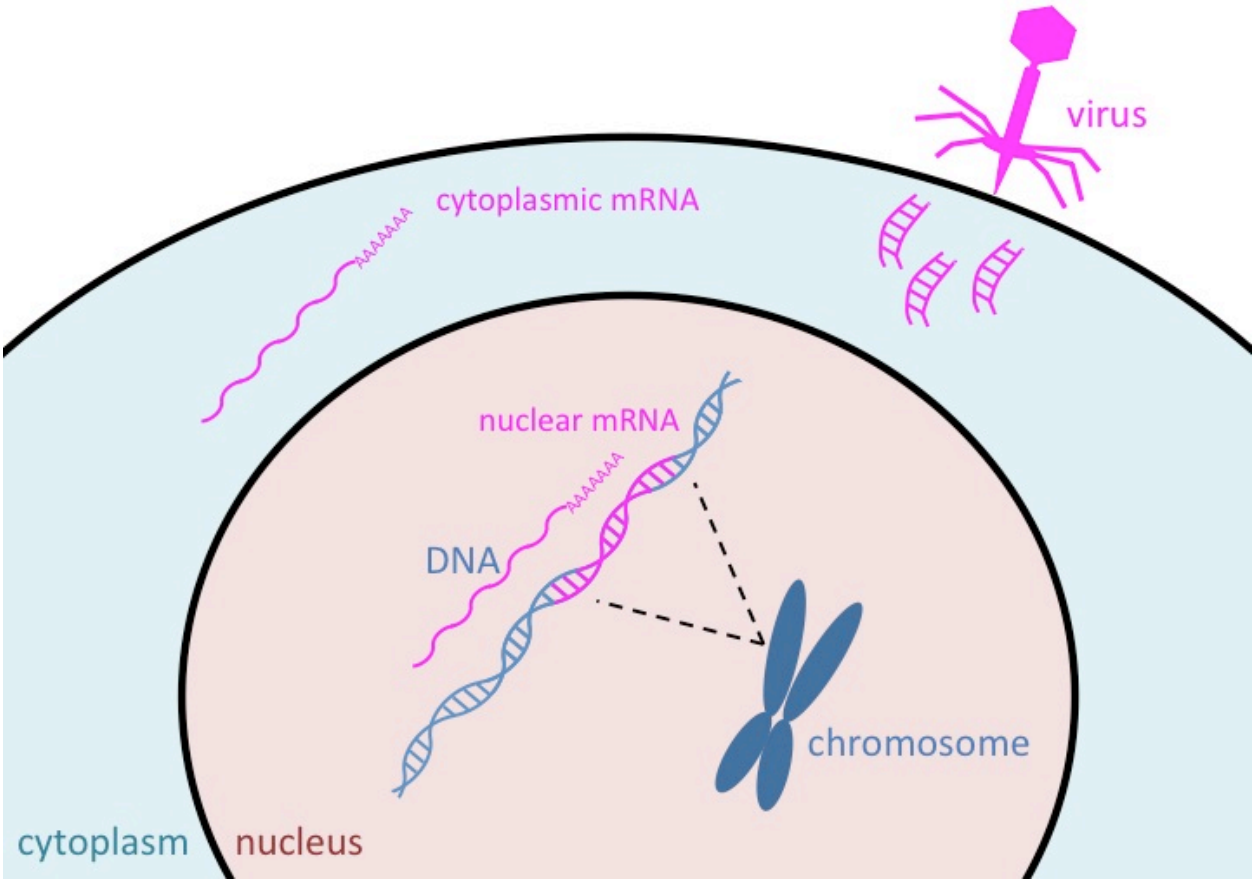


Figure 0.1. Viruses are capable of injecting their own genetic material into the cell, which in some viruses is then incorporated into the DNA and hijacks the native cellular machinery to transcribe its mRNA. This foreign RNA can then be detected in either the cytoplasm as mRNA or the nucleus precursor RNAs.

a

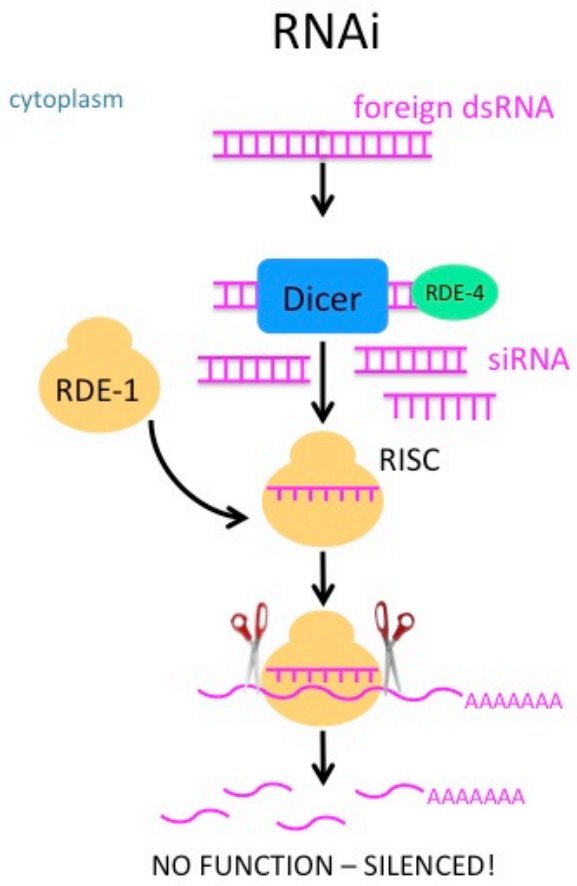


Figure 0.2. The mechanism RNA interference for silencing induced by small dsRNA.

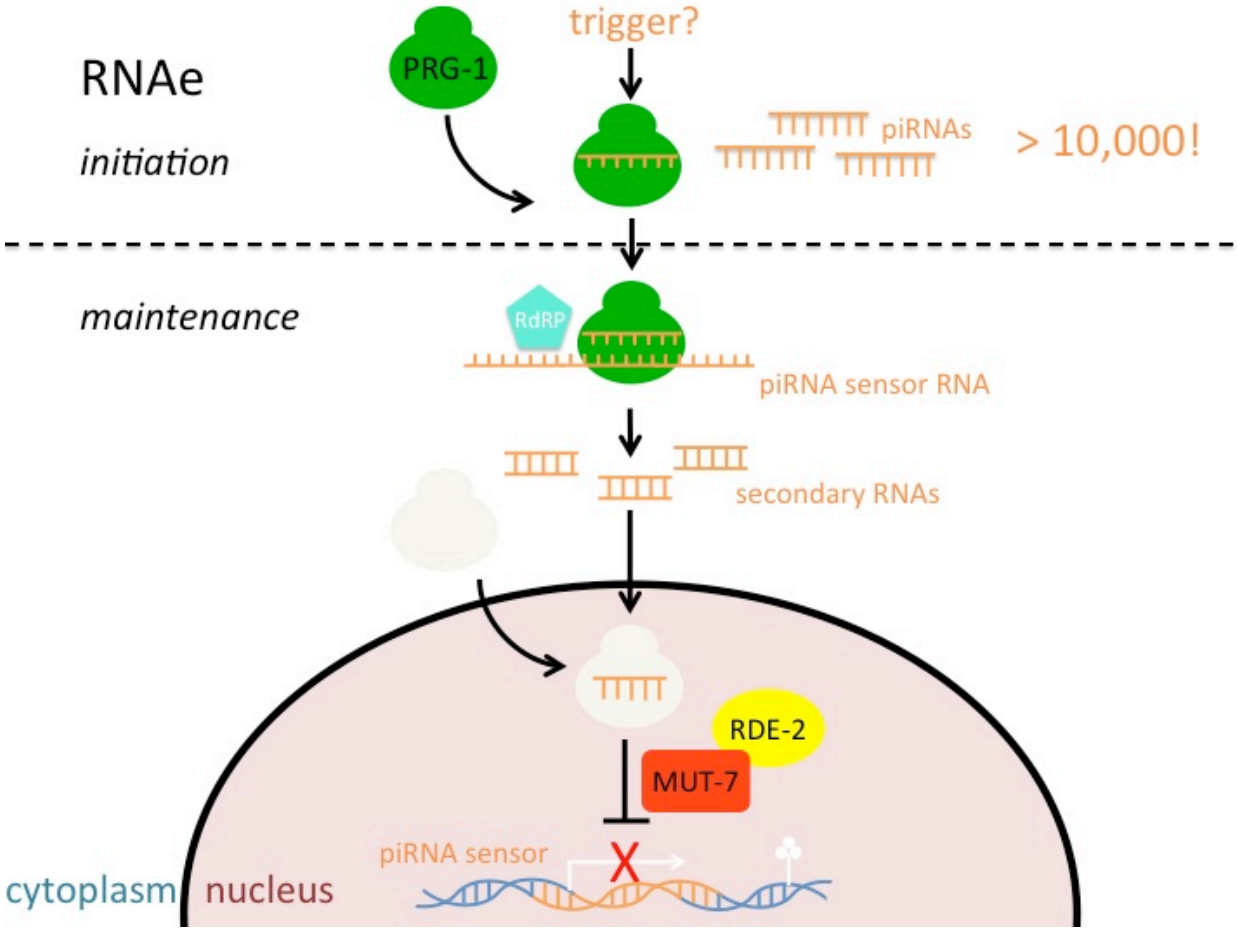


Figure 0.3. The RNAe silencing mechanism of the piRNA sensor.

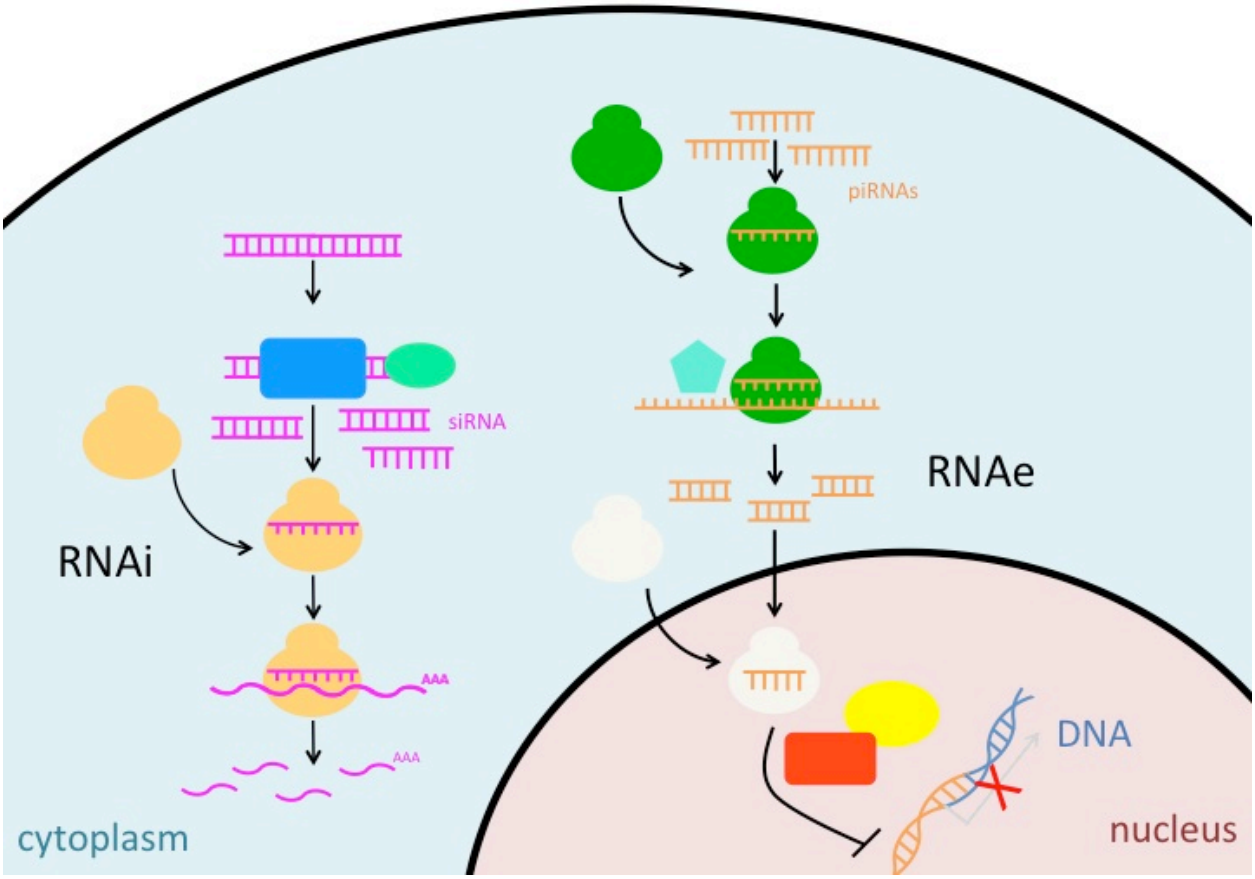


Figure 0.4. Silencing mechanisms located in the cytoplasm and the nucleus.

CHAPTER 1: *SMG-1(YP3)* TARGETS UNUSUAL TRANSCRIPTS

Background

For times of stress as well as regular maintenance, the cell is equipped with several surveillance mechanisms that ensure that RNA is being properly transcribed, and that any aberrant messages are quickly identified as they may have deleterious effects and be toxic to the cell. Nonsense-mediated decay (NMD) is one such mechanism that targets a very specific kind of error – an erroneous premature termination codon (PTC) inserted into the RNA message (7). NMD and its machinery was first identified in yeast, which have three NMD components (25), whereas multicellular eukaryotes have 7 or more, whose identities have largely been determined based on studies in *D. melanogaster* and *C. elegans* (26, 27). While NMD has proven to be a vital surveillance mechanism in eukaryotes, particularly in humans where a failure to identify and degrade PTC-containing mRNAs has been linked to several diseases, it is also know to function on endogenous transcripts (11, 28).

In mammals, research has shown that SMG-1 and its phosphorylation target UPF-1 (SMG-2 in *C. elegans*) are essential, and function in multiple pathways that maintain both genome and transcriptome integrity (11, 29). SMG-1 initiates a cellular stress response when either genome integrity, mRNA translation, or nutrient availability have been compromised, rendering it a multi-tasking safeguard (11).

SMG-1 is a PI3-kinase-like kinase (PIKK) and its function in NMD surveillance has been well described (9). For example, it has been shown that upon DNA damage inflicted by UV or γ-radiation, SMG-1 is activated. Depleting SMG-1 increases the cell's sensitivity to radiation and results in DNA damage (12). SMG-1 is also implicated in p53 phosphorylation upon G1 checkpoint activation in response to double-stranded DNA breaks (13). The oxidative stress response pathway is also mediated by SMG-1 by protecting against TNF-α-induced apoptosis (14). SMG-1-interacting protein UPF-1 (SMG-2 in *C. elegans*) has been shown to be required for the degradation of histone mRNAs in response to replicative stress in response to hydroxyurea-induction of the S-phase checkpoint (30).

In *C. elegans*, mutations in NMD genes were initially discovered by their ability to suppress mutations within genes that function in diverse processes. Seven genes with essential roles in NMD in *C. elegans* have been identified: *smg-1, 2, 3, 4, 5, 6,* and *7*. The *smg* genes were first identified for their action as recessive allele-specific suppressors – i.e. 'suppressor of morphological genitalia' – which additionally results in a protruding vulva (Pvl) phenotype (8). The key regulatory aspect of NMD processing is the identification of PTC on mRNAs, which is centrally regulated by the SMG-2 helicase. SMG-2 is phosphorylated and dephosphorylated by SMG-1, which is necessary for the function of NMD (9, 10). There is a high level of conservation of NMD machinery throughout species, although the mechanisms of PTC identification are thought to differ somewhat (31).

Importantly, genetic suppressors of *unc-54(r293)* are not limited to the *smg* genes, but also include uncharacterized mutations in genes that while disrupting *unc-54(r293)* expression, do not produce a Pvl phenotype when null. This second category of

suppressors suggests that *unc-54(293)* may not be identified in an identical manner to the canonical PTC target of NMD.

We have identified an allele of the SMG-1 kinase that is different from any known *smg* mutation. Exhaustive screens for mutations that affect NMD have been performed using the *unc-54(r293)* mutation, a small deletion in the 3'UTR of *unc-54* which causes a strong locomotive defect and decreased brood size. Unlike these other NMD-interacting mutations, *smg-1(yp3)* failed to suppress *unc-54(r293)* and certain satellite and tandem repeat sequences, but can suppress mutations that result in mRNAs with premature stop codons. These results suggest that the NMD machinery may target unusual RNAs, and that their regulatory effects may reach far beyond that of simple surveillance of a single type of target.

Results

smg-1 yp3 displays a temperature-sensitive defect in NMD

The *yp3* strain was isolated from and ethyl methanesulfonate (EMS) mutagenesis screen to define *mortal germline (mrt)* mutants that become sterile after growth for a number of generations (Liu, Boerkel, & Ahmed, unpublished). *yp3* displays a Mrt phenotype at the non-permissive temperature of 25°C, but not at the permissive temperature of 20°C (**Figure 1.1 A**). Additionally, *yp3* displays a variety of germline phenotypes at 25°C (**Figure 1.1 B**). Phenotypes range from a complete lack of germline to an overproliferation of the mitotic germline cells to form either an asymmetrical or symmetrical tumor (**Figure 1.1 B**). Quantification of the various aberrant germline phenotypes at both 20°C and 25°C over several generations reveals that a) this phenomenon is temperature-induced and b) there

is a progressive increase of aberrant germlines of every type as the animals are propogated over generations (**Figure 1.1 C**).

Identification & sequencing of smg-1 alleles

Initial mapping the *mrt* mutation in *yp3* suggested that the mutation resides on Chromosome *I*, and most likely in the *smg-1* gene. Genomic DNA from the *yp3* strain was isolated and sequenced, and a missense mutation was identified. The mutation results in a G1774E amino acid change in the kinase domain of the protein (**Figure 1.2 A**) – a residue that is highly conserved amongst eukaryotes (**Figure 1.2 B**). Outcrossing of *smg-1(yp3)* revealed that the Mrt and aberrant germline phenotypes discussed above are not attributable to the *yp3* mutation found in *smg-1* and must be due to an additional unlinked EMS-induced mutation.

The identities of other *smg-1* mutant alleles have been unknown, until now. Sequencing of genomic DNA from both *smg-1(cc545)* and *smg-1(cc546)* revealed that like *smg-1(yp3),* both carry missense mutations resulting in T671I and M1957L conversions respectively. The *smg-1(cc546)* mutation is quite close in proximity to that of *smg-1(yp3)*, as they are both harbored in the PI3K domain of SMG-1 (**Figure 1.2 A**). The *smg-1(cc545)* mutation resides much farther toward the N-terminus of the protein, and does not fall in any known domains.

smg-1(cc545) and smg-1(cc546) but not smg-1(yp3) suppress unc-54(r293) transcripts

unc-54 encodes for a myosin heavy chain II (MHC B), one of the two myosin heavy chains expressed in the body wall muscle, and is necessary for proper locomotion (32). Despite it's distinct lack of a PTC, the mRNA of the *unc-54(r293)* allele is traditionally

thought of as a substrate of nonsense-mediated decay (NMD) and has even been used in screens to identify NMD components (8). Under normal circumstances, the *unc-54(r293)* mRNA is a targeted, the message is degraded and the animal displays the *unc-54* paralyzed phenotype. In certain NMD-defective backgrounds, including *smg* mutant alleles, this does not occur - a full-length mutant *unc-54* mRNA is made and remains stable, and as a result the animal has little to no motility defects – this effect is termed 'suppression-of-Unc' (8).

I created double mutants of the *smg-1* alleles - *r904, cc545, cc546, yp3* – with *unc-54(r293)* and subsequently assayed these strains visually for the *unc-54(r293)* phenotype (**Figure 1.3 B**), as well as quantitatively using a locomotion assay(**Figure 1.3 C & D**). I found that the *unc-54(293)* phenotype was smg-suppressible in the presence of the *r904*, *cc545,* and *cc546* mutations. Animals appeared almost wildtype as observed under the microscope, moving around the plate and taking on a typical sinusoidal shape during locomotion.

However, unlike those canonical targets, *yp3* failed to suppress the *unc-54(r293)* phenotype. The animals appeared rigid and failed to move in a wild-type manner, as evidenced by the small amount of tracks made in the bacteria (**Figure 1.3 B**). Quantification of locomotion indicated that movement in the *smg-1(yp3) unc-54(r293)* double mutant is not restored and is within the parameters of the *unc-54(r293)* mutation alone (**Figure 1.3 D**). Thus, the mutation in *unc-54(r293)* is suppressible via the *smg-1* mutations *r904, cc545 and cc546*, but not by *yp3.* This indicates that *smg-1(yp3)* is a separation-of-function allele that is proficient in the degradation of *unc-54(r293)* mRNA.

Molecular nature of unc-54(r293)

The allele *unc-54(r293)* was identified as a small spontaneous deletion located entirely within the 3' UTR, including the 3' cleavage and poly-adenylation sites (33, 34) (**Figure 1.3 A**). However, when assayed via Northern Blot, it has been shown that rather than a short mRNA being produced, an extra long message is present (33). I have determined via reverse transcription polymerase chain reaction (rtPCR) and sequencing that the deletion in *unc-54(293)* actually results in a read-through to the next gene downstream – *aex-5.* The sequence contains the full open reading frames (ORFs), including the start and stop codons, of both *unc-54* and *aex-5* transcripts, which creates an in-frame bicistronic message with three stop codons, one at the 3' ends of unc-54 and two in a small intergenic linker region (**Figure 1.4**). The large transcript also contains a linker region between the two ORFs consisting of a 198 bp long track between the stop codon of *unc-54* and the start codon of *aex-5*. Notably, it contains two small putative ORFs which are 21 bp and 66 bp in length. This linker region is such that *aex-5* coding region is in frame. If the *unc-54(r293)* mRNA is intact, *aex-5* is presumably also translated, as null mutants of this gene should produce no viable progeny, according to RNAi experiments (35).

Thus, at the molecular level, the *unc-54(r293)* does not appear to result in the canonical PTC target of NMD, but rather produces a more complex target. While still regulated by SMG-1, the lack of suppression via *smg-1(yp3)* indicates that a different type of regulation – governed by NMD machinery, but not identical to NMD – is the real watchdog of this transcript.

smg-1(yp3) stabilizes mRNAs with long 3'UTRs

It was previously suspected that a unique property of the *unc-54(r293)* is that it potentially contains an extra-large UTR – or at least a large portion of sequence on the mRNA after the stop codon of the *unc-54* gene. I postulated that the *aex-5* sequence on the end of the *unc-54(r293)* transcript could be behaving simply as an aberrant UTR. D. Reiner contributed the GFP insertion strain pCM1.4 containing a UTR of substantial length that is known to be suppressible in a temperature-sensitive manner by the *smg-1(cc546)* allele (**Figure 1.5**) (correspondence, D. Reiner). pCM1.4 consists of a *lin-26* promoter, a GFP sequence, and *smg*-suppressible 3'UTR, where neither the poly-A site nor the stop codon of the GFP sequence are intact and not disrupted by the UTR (**Figure 1.5 A**). This strain is particularly useful as GFP serves as a visible marker as to whether the transcript is stable or not – if NMD is functioning in a wildtype manner, there will be no GFP expressed, however if the transcript is suppressed, GFP will be evident, particularly in comma stage embryos.

Pairing the pCM1.4 insertion with our novel *smg-1(yp3)* allele would determine whether or not the pattern of suppression was similar to that of the surprising result yielded by the *unc-54(r293)* transcript. Upon making the double mutant with *smg-1* alleles *r904*, *cc545*, *cc546* and *yp3*, I saw equal levels of suppression of degradation of the transcript resulting in the expression of GFP (**Figure 1.5 B**). From these results I can determine that the simple inclusion of an extra long-UTR is not a sufficient target of the alternative *smg-1*-mediated mechanism that degrades the *unc-54(r293)* transcript, and does not further illuminate the nature of the *yp3* separation-of-function.

smg-1(yp3) stabilizes the cha-1 unc-17 gene complex

Few examples of truly bi- or multi-cistronic messages in *C. elegans*, exist, however the *cha-1 unc-17* complex presents a fairly similar scenario, but is a complicated and unique transcript in the *C. elegans* genome. *cha-1* encodes for the structural gene for choline acetyletransferase, and the *unc-17* encodes the synaptic vesicle-associated acetylcholine transporter. Thus, these two genes are related in their function as they are both required for the metabolism acetylcholine. As such, when mutated, both have similar but not identical phenotypes (36).

The molecular structure of the *cha-1 unc-17* complex has been deduced and actually consists of an alternatively spliced transcript that is processed from a pool of one large transcript (36). The *cha-1 unc-17* gene complex is not a true bicistronic message like that of *unc-54(r293)* in that there are not distinct start and stop codons that follow each other sequentially, but rather the two different transcripts arise from alternative splicing – both utilize the splice leader 1 (SL1) at the 5' end of the complex, and both include the first short exon – it is only then that the transcripts are alternatively spliced into sequential portions of the transcript, as described in previously (36) and additionally shown in **Figure 1.6 A**.

There exist overlapping deletions at this locus which lie between the cha-1 and unc-17 genes – *unc-17(md1447)* and *cha-1(p1156).* EMS mutagenesis defined suppressors of both deletions in *smg* genes, and this was confirmed by testing a variety of different *smg* mutations genetically. The *p1156* mutation resulted in a bicistronic transcript, as *cha-1(p1156)* was confirmed by sequencing to be a 747 bp deletion that removes the entire *unc-17* 3'UTR, the poly-adenylated site, and 285 bp of non-coding genomic sequence. It has been determined that this deletion should also remove splicing sequences vital for the

downstream expression *cha-1* (37). The *unc-17(md1447)* allele is a deletion of 465 bp – which removes a portion of the 3'UTR, the poly-adenylated site, and 245 bp of non-coding genomic sequence (36) (**Figure 1.6 A**). Both the *p1156* and *md1447* mutations result in phenotypes with limited locomotion and resistance to cholinesterase inhibitors, which *cha-1(p1156)* carrying the more severe locomotive phenotype (38).

Testing our novel *smg-1(yp3)* allele against such a target would certainly yield information regarding the type of target necessary to induce the unique pattern of suppression seen in the *smg*-suppressibility of *unc-54(r293).* The cause of the dysfunction of the *p1156* and *md1447* transcripts was previously unknown, however it is now clear that NMD machinery is involved, as the alleles of *smg-1 – r904, cc545, cc546* and *yp3* are all capable of suppressing this degradation (**Figure 1.6 B**). and as a result, the *cha-1 unc-17* hybrid transcript is produced. The phenotypic consequence of this suppression is an animal which appears more wildtype (**Figure 1.6 B**). Upon sequencing, the *cha-1(p1156)* mutation was confirmed to be a deletion between the two transcripts. However if NMD-mediated degradation is suppressed in the *smg-1(cc546) cha-1(p1156)* double mutant, I see that there is read-through to the downstream *cha-1* gene. This deletion does cause the insertion of a putative stop codon, but it is unknown whether this is the direct target of degradation. Thus, the *cha-1 unc-17* gene complex is stabilized in an NMD, SMG-1-dependent manner.

Although both the *cha-1 unc*-17 gene complex and *unc-54(r293)* appear to result in long transcripts with potentially two individually expressed mRNAs, they are not targeted in the same manner. Simply producing a read-through which contains two transcripts is

not sufficient to illicit the form of degradation targeting the *unc-54(r293)* transcript which results in the *yp3* separation-of-function phenotype.

smg-1(yp3) suppresses a transcript with a canonical PTC

In order to confirm that *smg-1(yp3)* is indeed a functional mutation of *smg*-1 and can suppress the NMD-mediated degradation of mRNA messages bearing a PTC, I assayed its NMD function against a target bearing a canonical PTC – *unc-30(e191). unc-30* encodes for a transcription factor that controls the terminal differentiation of the GABA-ergic motor neurons (39), and the mutation *unc-30(e191)* bears a C to T substitution in the first exon, resulting in a Q to amber stop conversion (**Figure 1.6 C**). This mutation results in a moderate locomotive defect, which is easily discernable by eye, making it an advantageous choice for analysis.

Assaying the effect of *smg-1(yp3)* on these strains was done by creating double mutants of our various *smg*-1 alleles with *unc-30(e191).* If the alleles of *smg*-1 tested were abrogating NMD function, we would see a rescue of the mutant phenotype, as the mutant mRNA message will still be made, and translated into what would most likely be viable, functional protein. Although the purpose of NMD is to rid the cell of mutant transcripts that may present a dominant negative or otherwise harmful adversary to the cell, in many cases a single amino acid change will still result in a functional protein. All alleles of *smg-1* tested (*r904, cc545, cc546,* and *yp3)* were capable of restoring some function to the mutant *unc-30* strain at 25°C (**Figure 1.6 D**). At 20°C, only the null mutant, *smg-1(r904)* was capable of the suppression, indicating that the *cc545, cc546* and *yp3* alleles are all temperature-sensitive in this regard. These results are in keeping with past research demonstrating that the

cc545 and *cc546* alleles are temperature-sensitive, as well as with our expectations regarding the nature of the *yp3* mutation, give it's proximity to *cc545* (**Figure 1.2 A**). Therefore, *smg-1(yp3)* disrupts canonical NMD suppression of an mRNA message containing PTCs in a temperature-dependent manner.

Tandem and satellite repeats affected by smg-1(yp3)

Microarray analysis was performed to identify possible endogenous targets of *smg-1* that could be stabilized by *r904, cc545,* and *cc546* alleles, but were suppressed by *yp3.* Our analysis yielded results and I found several satellite and tandem repeats that were suppressed by *smg-1(yp3)*. There appear to be two classes of satellite regulation. One, in which satellite repeats are up-regulated in *r904, cc545,* and *cc546*, and down-regulated in *yp3* as compared to wild-type, and another where repeats are up-regulated by all *smg-1* alleles tested, but to a consistently lesser extent in *yp3.* Further analysis is needed.

Discussion

Re-thinking NMD – is it really just about nonsense?

It has become clear in recent years that the mechanism traditionally known as NMD is not merely a simple quality-control mechanism, but is also an important factor in the regulation of many physiological mRNA targets as a part of their normal function. Microarray experiments have strongly suggested that approximately 10% of transcripts in *S. cerevisiae* are affected when NMD is inactivated (40-42). In yeast, this has lead to the understanding that there are several distinct classes of structural targets available to the

NMD mechanism – (1) the canonical PTC inserted into an mRNA regions (43, 44), (2) mRNAs with abnormally long 3'UTRs (45), (3) certain mRNAs that contain small upstream open reading frames (uORFs) (46, 47), (4) pre-mRNAs that contain introns and have mistakenly entered the cytoplasm (48), and (5) mRNAS where an out-of-frame AUG start codon is used for translation initiation (41).

Although higher eukaryotic studies do not go quite as far in attributing these kinds of surveillance to NMD, mammalian experiments do suggest that one-third of splice events may produce a possible NMD target (49, 50), indicating that NMD is functioning not only as a surveillance mechanism, but as a regulatory system that responds to changes in transcription environments. For example, it is known that T-cell receptor genes can form rearrangements that result in their mRNA messages containing PTCs (51). It is highly likely that various components of the NMD machinery are functioning in many different types of regulation and surveillance as in yeast – the system is likely more complex, making it more difficult to elucidate exactly which components and mechanisms are acting on particular breeds of substrate.

A specific molecular target within unc-54(r293)

After testing several types of *smg*-suppressible targets against our novel *smg-1(yp3)* allele, I can discern that there is something especially unique about the *unc-54(r293)* transcript that makes it a non-canonical *smg*-suppressible target whose degradation utilizes NMD machinery (**Figure 1.7**). Given what we now know about the molecular structure and sequence of the *unc-54(r293)* transcript - as well as other NMD targets that

are not suppressible via the *smg-1(yp3)* allele - one can contemplate several options for the specific molecular target that lies within *unc-54(r293)* transcript.

A possible molecular target of *unc-54(r293)* is in fact the small 21 bp and 66 bp putative ORFs found in the linker region between *unc-54* and *aex-5* in the *r293* transcript (**Figure 1.4**). Initially it was thought that translation machinery would initiate upon locating the most proximal AUG codon (52), but in recent years it has come to light that there exist transcripts with long 5' regions upstream of the "true" start that contain functional AUG codons that presumably have a regulatory function (53). In fact, bioinformatic surveys have identified uORFs in as many as 35-49% of human and rodent transcripts (54-56). Few of these uORFs have been studied in detail (57, 58) and their mechanism of action is largely unknown. Already found to initiate NMD in yeast (46, 47), uORFs have recently been implicated in functions requiring NMD machinery in human cells. Microarray analysis indicates that when NMD components UPF-1 (SMG-2 in *C. elegans*) and SMG-6 are knocked down, more than 35.5% of all upregulated genes contain at least one uORF (59).

Positive identification of the uORFs at the beginning of our *unc-54(r293)* transcript may prove difficult as uORFs are very diverse and vary greatly in length, number per transcript, position relative to the "true" start, sequence and secondary structure (60). If the putative ORFs upstream of *aex-5* are indeed functional ORFs, this would give us great insight into now only how the *unc-54(r293)* transcript is being regulated but also strongly implicate NMD machinery in the mechanism by which uORFs regulate translation in a variety of transcripts across the genome.

It is also possible that the linker region in the *unc-54(r293)* transcript (**Figure 1.4**) is being targeted and degraded via miRNAs. Using miRBase, it was found that four separate miRNAs are predicted to bind to the linker region – cel-miR-85, cel-miR-65, cel-miR-1829a and cel-miR392. As of yet none of these miRNAs have known targets or functions. However, it is plausible that miRNAs could bind to the small linker region and target the entire *unc-54(r293)* transcript for degradation in a *smg*-mediated manner. This would implicate a new mechanism for miRNA action and could help to elucidate the function and purpose of many miRNAs that are currently known.

Disease implications

Understanding canonical NMD in and of itself has large repercussions in terms of understanding disease biology. It is estimated that approximately 1800, or 25% of all human genetic diseases have alleles that are due to an aberrant mRNA transcript that contains a PTC and subsequent degradation of the RNA message. Diseases like Duchenne muscular dystrophy (DMD), cystic fibrosis and many others are of this type. Currently treatments that bypass the NMD mechanism are under investigation. Depending on the location of the mutation that causes the PTC, if read-through is allowed, a functional protein could be made (11). For example, in rare cases of DMD, the truncating mutation occurs near the 3' end of the dystrophin gene. In theory, if one could over-ride PTC identification by NMD, a functional protein could be made, rescuing the DMD phenotype (61). A new class of drugs is emerging to attempt to rescue the protein production of diseases like DMD via NMD suppression – primary among them is Ataluren, which is currently in Phase III trials (41, 62).

The isolation of the separation-of-function allele *smg-1(yp3)* may make it possible to further advance therapies that would specifically inactivate NMD, thus rescuing disease phenotypes, leaving other regulation pathways governed by some NMD machinery – like that which degrades the *unc-54(r293)* transcript - intact and functional.

There could also be more that *smg-1(yp3)* could teach us about disease. The novel form of degradation that seems to target the *unc-54(r293)* transcript as well as certain satellite and tandem repeats, could help to explain several kinds of human cancer that have been linked to the de-repression of these heterochromatic regions. Demethylation at Sat2 repeats was found to be a hallmark in (63), and a genome-wide profile of lung carcinomas indicated that satellite repeat regions were also hypomethylated (64). It has also been shown that loss of Brca1 in mice results in a de-repression of tandem satellite repeat DNA and that similar de-repression was also observed in human BRCA-1-deficient breast cancer cells (65).

Delving further into the source of the degradation of the unique transcript *unc-54(r293)* - and possibly certain satellite and tandem repeats - could be the first small step to uncovering a slew of knowledge pertaining not only to the understanding of basic biological and genetic regulatory mechanisms, but could have a direct impact on disease knowledge that could improve current technology and drug development.

Materials and Methods

Strains

Unless noted otherwise, all strains were cultured at 20°C on Nematode Growth Medium (NGM) plates seeded with *E. coli* OP50. Strains used include Bristol N2 wild type, *smg-*

1(r904) I, PD8119 *smg-1(cc545) I*, PD8120 *smg-1(cc546) I, smg-1(yp3) I, unc-54(r293) I, dpy-5(e907) I*, CB61 *dpy-5(e61) I*, MT7929 *unc-13(e51) I*, CB450 *unc-13(e450) I, unc-17(md1447) IV*, CB845 *unc-30(e191) IV, cha-1(p1156) IV, pCM1.4*

smg-1(yp3) was outcrossed versus an outcrossed stock of *dpy-5 e61, unc-13 e450*, and freshly isolated homozygous F2 lines were established and outcrossed a further five times by the same method.

Double mutants were typically constructed by marking each mutant with a physical marker, selecting one mutant-marker strain and crossing with N2 to yield males, and crossing said males with the other mutant-marker strain. F1 progeny of successful crosses were then isolated, and each marker selected against in subsequent F2 and F3 generations.

Locomotion Assay

As previously described in Reiner et al, 1999 (66). To assay for locomotion, individual animals were placed in the center of a 60 mm plate containing a lawn of *Escherichia coli* OP50 with the origin marked. The plates were incubated for 15 minutes at room temperature (approximately 23°C), and then briefly incubated at 4°C to stop movement. The radial distance from the origin to the farthest distance away the animal moved was measured to the nearest 0.5mm.

GFP Microscopy

Comma stage embryos were mounted onto glass slides using M9 and glass beads. Images were taken under 100X magnification using a Nikon Eclipse E800 microscope.

RNA Preparation and cDNA Synthesis

RNA for rtPCR was prepared using mixed staged animals collected with M9 buffer. Total RNA was purified by the Guanidinium thiocyanate-phenol-chloroform (TRIzol) extraction method. First-strand cDNA was synthesized using Super Script III Reverse Transcriptase (Invitrogen).

Figures

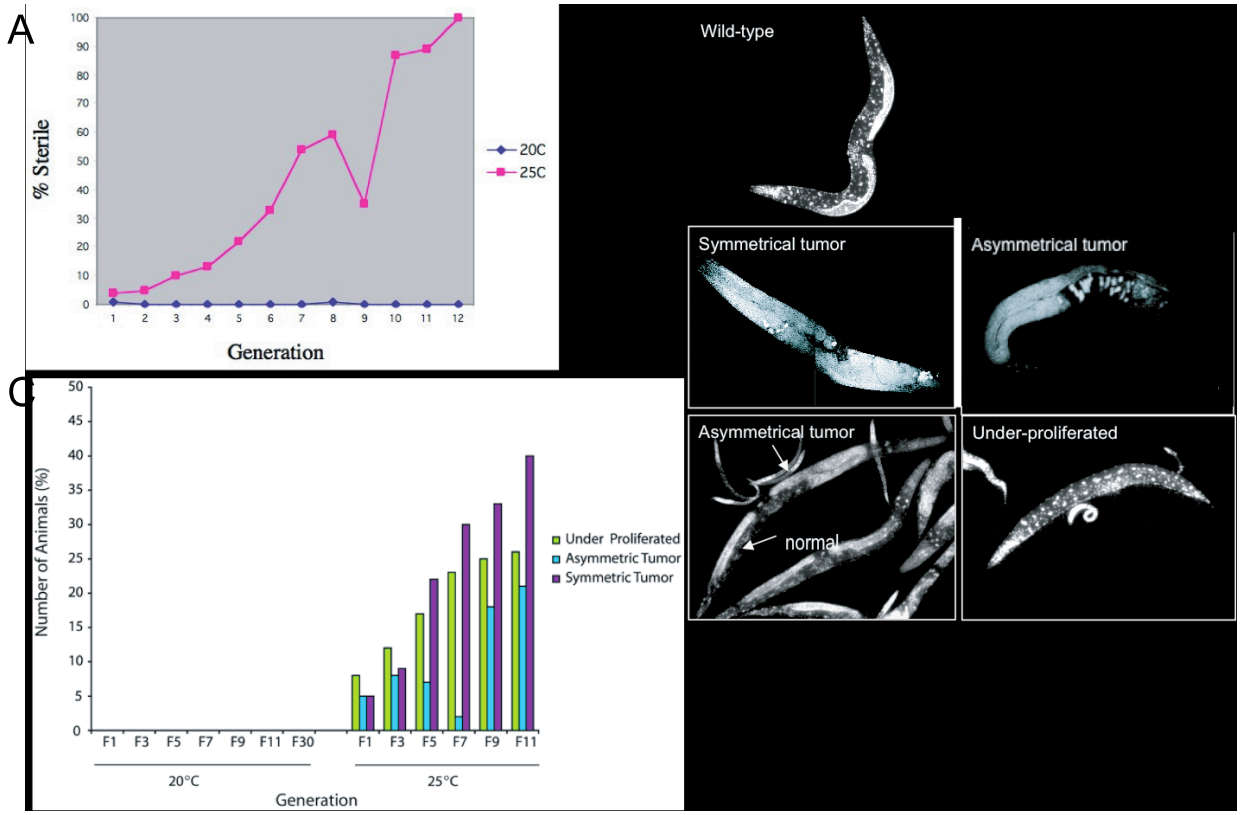


Figure 1.1. *smg-1(yp3)* displays several unique phenotypes. (A) *smg-1(yp3)* has a temperature-sensitive sterility phenotype, where progressive sterility over 10+ generations is observed at 25°C but not at 20°C. (B) *smg-1(yp3)* animals also display various tumorous phenotypes at 25°C, which are quantified in (C).

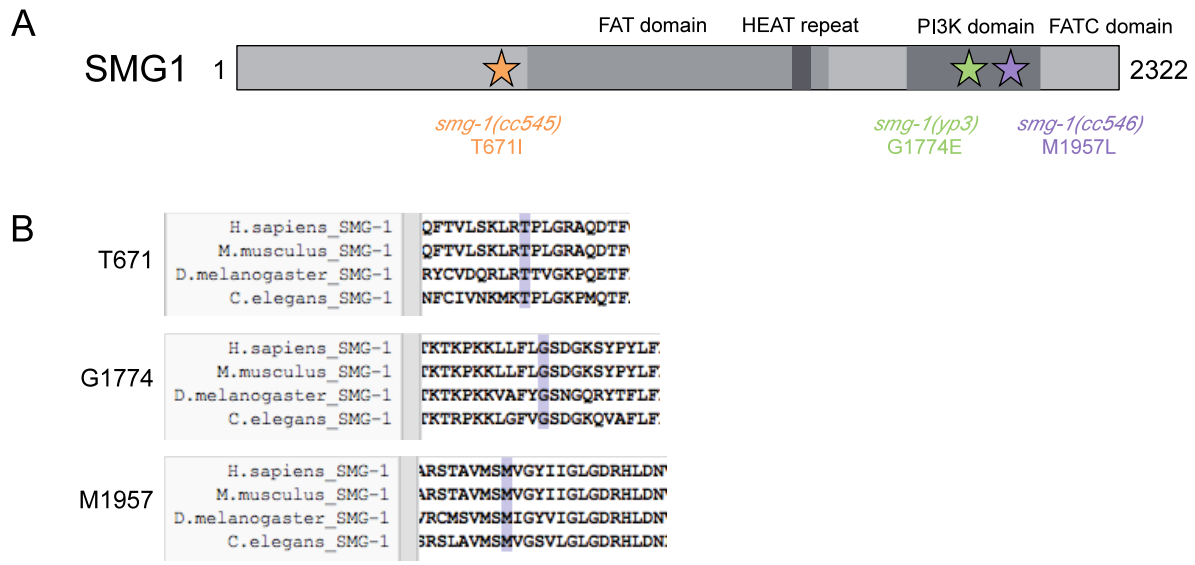


Figure 1.2. Identification of mutations in *smg-1* alleles and corresponding residue conservation. The mutations in several *smg-1* alleles were (A) sequenced and identified, and are (B) shown to be highly conserved across species.

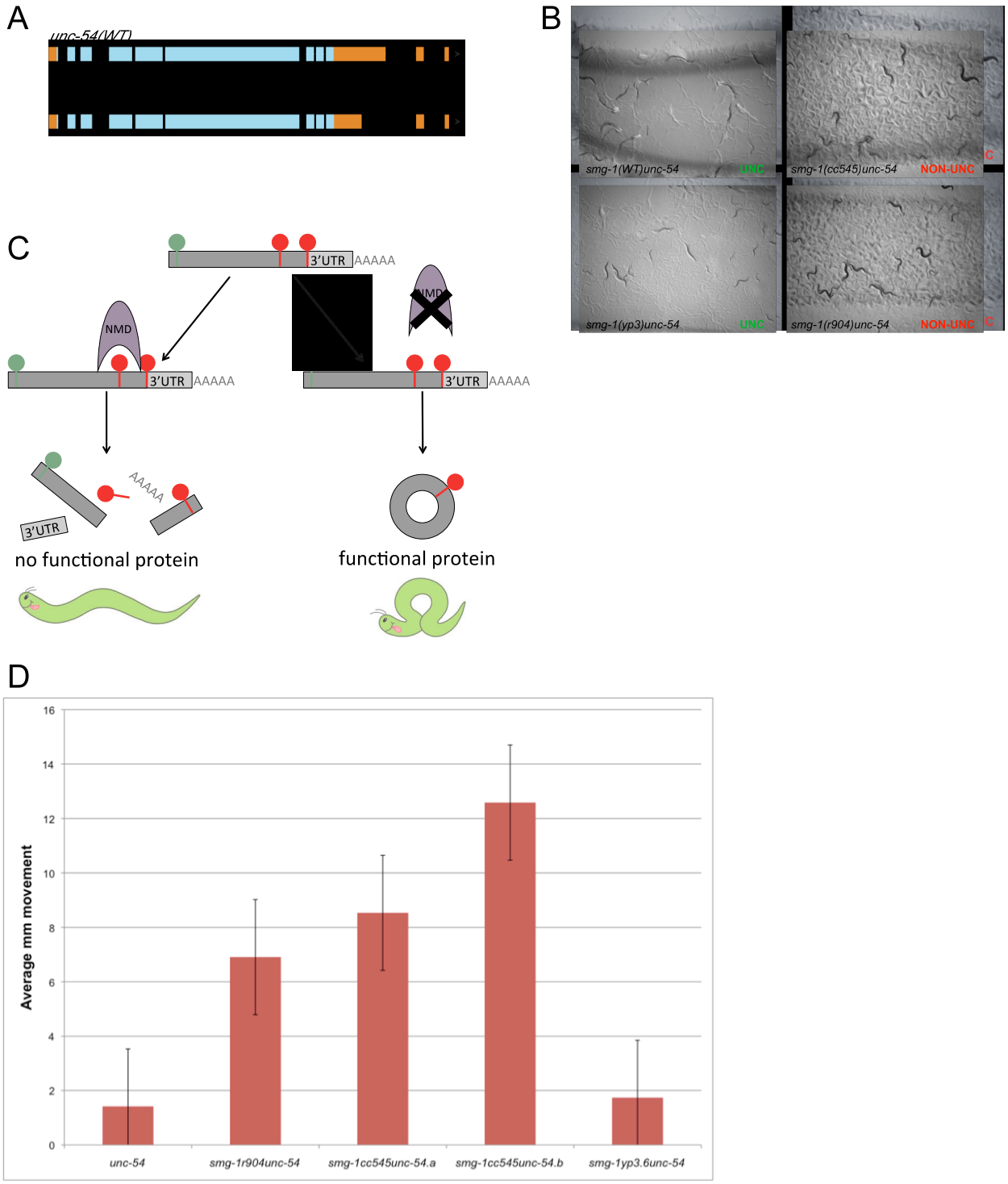


Figure 1.3. *unc-54(r293)* serves as an indicator for NMD function in smg-1 mutants. (A) Molecular structure of *unc-54(r293)* versus the wild-type gene. (B) Images of *unc-54* animals in various *smg-1* allele backgrounds. (C) Locomotion assay to assess whether NMD is occurring in *unc-54(r293)* animals. (D) Quantification of results of locomotion assay in various *smg-1* allele backgrounds.

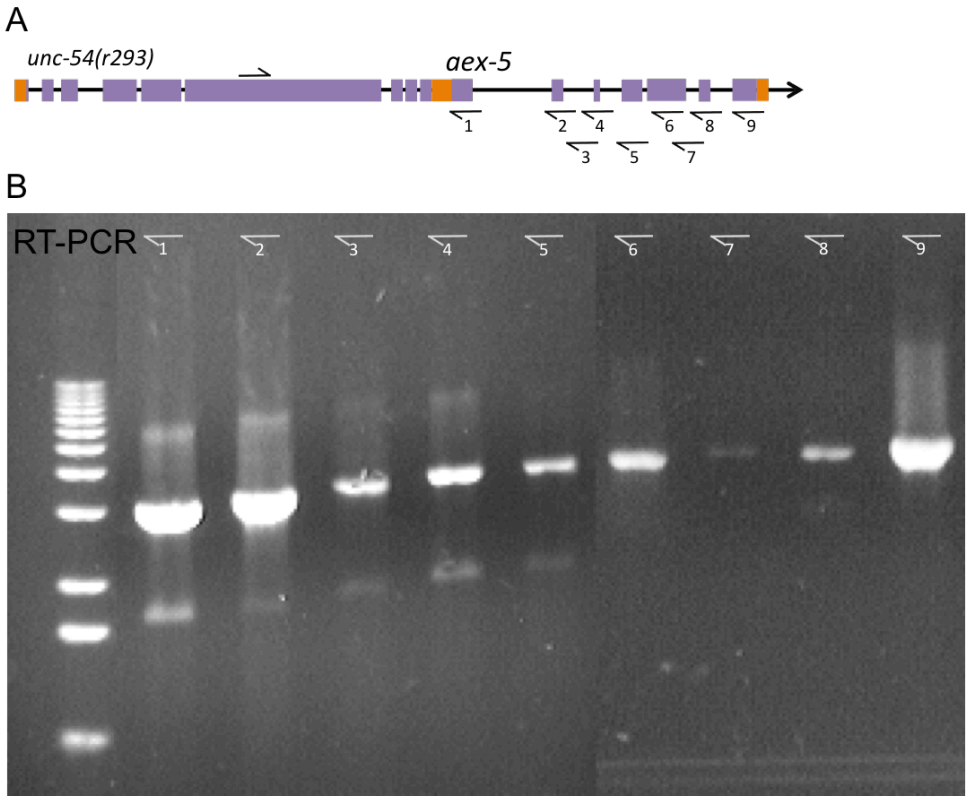


Figure 1.4. The *unc-54(r293)* message is bicistronic. (A) The structure of the *unc-54-aex-5* bicistron, in addition to the location of primers used to identify this read-through. (B) rtPCR of the *unc-54(r293)* message.

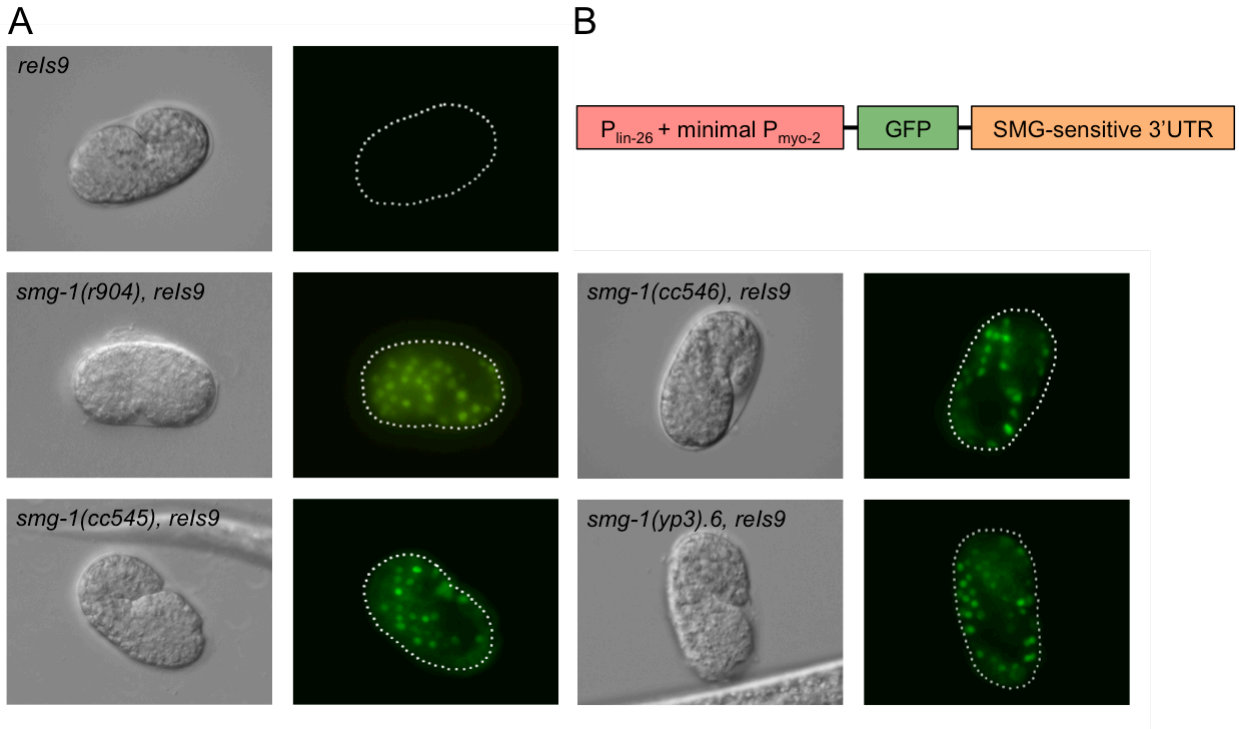
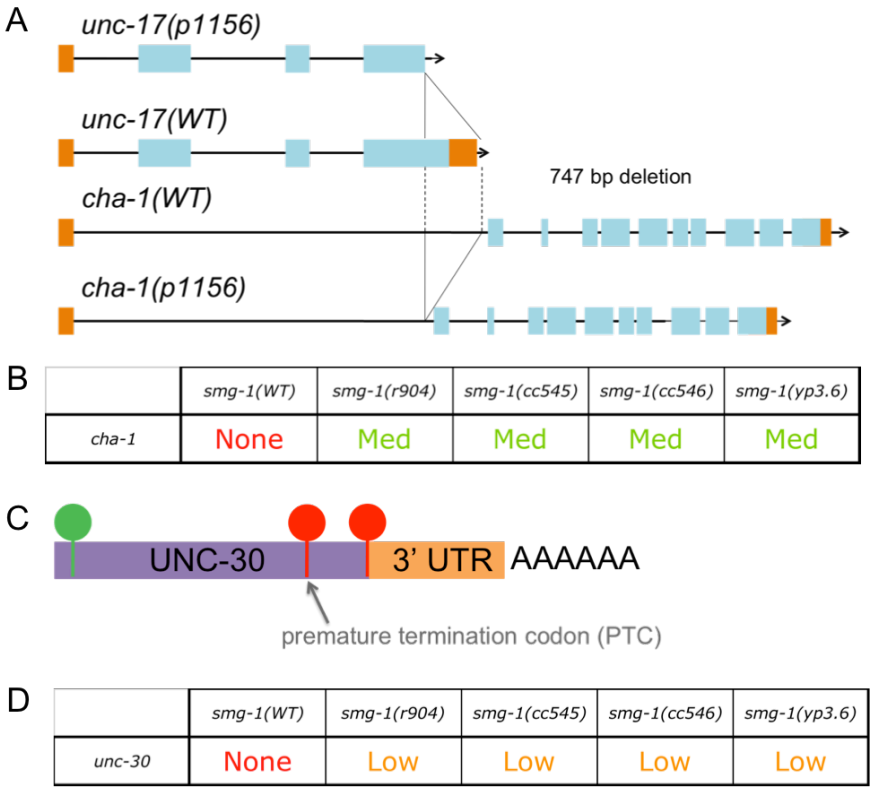


Figure 1.5. Structure and expression of the pCM1.4 transgene. (A) Microscopy images of representative comma-stage embryos containing pCM1.4 (here denoted as reIs9) in various *smg-1* mutant backgrounds. (B) The structure of the pCM1.4 transgene.



	smg-1(WT)	smg-1(r904)	smg-1(cc545)	smg-1(cc546)	smg-1(yp3.6)
cha-1	None	Med	Med	Med	Med

	smg-1(WT)	smg-1(r904)	smg-1(cc545)	smg-1(cc546)	smg-1(yp3.6)
unc-30	None	Low	Low	Low	Low

Figure 1.6. Examination of NMD activity in *smg-1* mutant backgrounds. (A) Structure of the *unc-17 cha-1* complex. (B) Results of examining various *smg-1* alleles in the cha-1 background for the suppression-of-Unc phenotype. (C) Structure of *unc-30(e191)* and (D) suppression-of-Unc phenotypes in various *smg-1* mutant backgrounds.

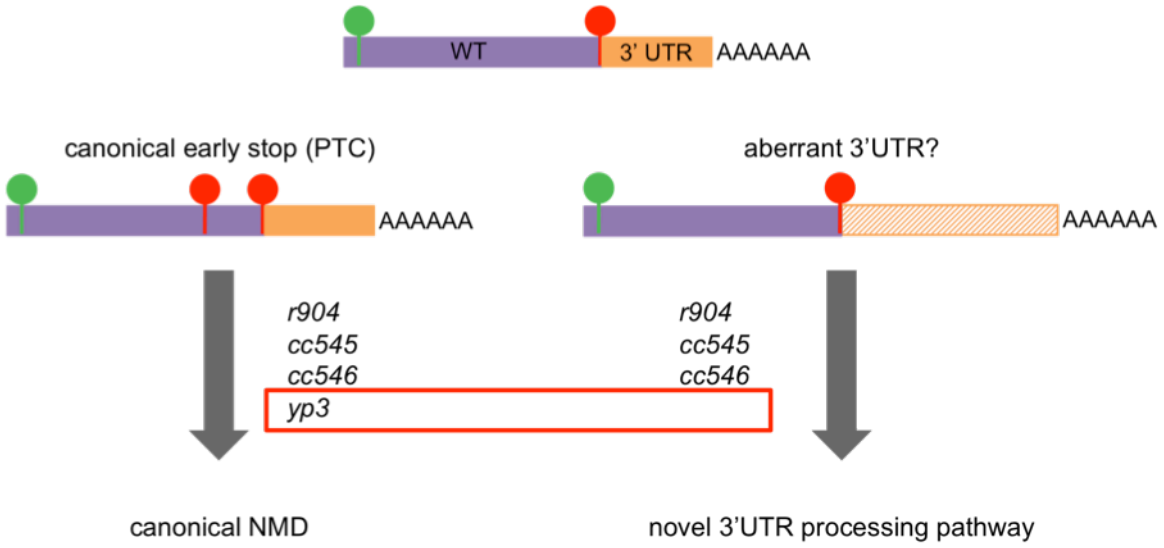


Figure 1.7. Model displaying the possible separation-of-function occurring in *smg-1(yp3)* due to different mRNA targets..

CHAPTER 2: LACK OF PAIRING DURING MEIOSIS TRIGGERS MULTIGENERATIONAL TRANSGENE SILENCING IN *CAENORHABDITIS ELEGANS*

Background

Small RNAs can repress expression of endogenous genes as well as parasites such as transposons or viruses. One form of small RNA-mediated repression is epigenetic silencing of genomic loci that can result in a permanent, heritable state of expression in germ cells.

RNA interference (RNAi) is a conserved biological process in which small non-coding RNA molecules promote gene silencing (2). RNAi was originally identified in *Caenorhabditis elegans*, but has been observed in a large number of eukaryotes ranging from yeast to plants to humans (1, 2, 67-70). Small interfering RNAs (siRNAs) can be produced from a variety of sources (71, 72) and can be divided into two classes – endogenous siRNAs, which are produced by genes, transposons, or aberrant transcripts, and exogenous siRNAs, which are derived from foreign double-stranded RNA (dsRNA). In *C. elegans*, siRNAs interact with Argonuate proteins that potentiate their functions (16). When exogenous double-stranded RNA is introduced into *C. elegans*, it is processed by the Dicer nuclease into primary 5' monophosphorylated 22G siRNA duplexes that interact with the Argonaute RDE-1 whose slicer activity promotes degradation of one strand of the duplex (15). RDE-1 and associated primary siRNAs then interact with target mRNAs to recruit RNA-dependent RNA polymerases (RdRPs) that synthesize secondary 5' triphosphorylated

siRNAs (17, 73-76). Secondary Argonaute proteins (SAGOs) then bind with the secondary siRNAs, and it is this complex that directly degrades the target mRNA message (18).

A second class of primary siRNAs in *C. elegans* is the Piwi-interacting RNAs (piRNAs) that are highly abundant in the germline and interact with the *C. elegans* Piwi Argonaute protein PRG-1 (22, 23). *C. elegans* piRNAs are termed 21U-RNAs as they are 21 nucleotides long and possess a 5' uracil. PRG-1 and associated piRNAs target transposons and some genes based on imperfect homology, which recruits RdRPs to promote biogenesis of 22G secondary siRNAs that bear perfect homology to their targets. These secondary 22G-RNAs interact with WAGO-class Argonaute proteins to promote silencing of germline loci. The vast repertoire of *C. elegans* piRNAs and their ability to target nucleic acids with mismatches may allow them to target both endogenous loci as well as foreign nucleic acids such as transposons or viruses (24, 77). Although *C. elegans* piRNAs can target many endogenous transposons, *prg-1* mutants displayed transposition for only one out of three DNA transposons tested, even though all three transposons become active if secondary siRNA populations are disrupted with *Mutator* gene mutations (78). These results imply that small RNA-mediated epigenetic silencing of many transposons may initially depend on PRG-1 and associated piRNAs, but then a downstream secondary siRNA system is capable of maintaining silencing of most transposons in the absence of piRNAs.

A process that relies on a secondary silencing system is 'cosuppression', where microinjection of plasmids into the germline of *C. elegans* often leads to creation of repetitive extrachromosomal transgenic arrays that are expressed in somatic cells but become rapidly and permanently silenced in the germline (79). In cosuppression, germline silencing can act in trans on the endogenous locus, leading to its silencing as well (80, 81).

The diffusible transposon silencing factor depends on *Mutator* class proteins that promote secondary siRNA production, but not on proteins that initiate the response to exogenous dsRNA such as RDE-1 or RDE-4 (80). Cosuppression was originally observed in plants in response to high copy number transgenes and also acts via siRNAs (82-84). Methylation-mediated silencing of multicopy transgenes has also been observed in zebrafish, where transgenes inserted as high copy concatemeric arrays become completely silenced in two to three generations (85, 86). Similar results are seen in mice, where high copy number is inversely correlated to expression (87, 88). Therefore, epigenetic silencing of high copy number transgenes in the *C. elegans* germline is consistent with a natural defense response to foreign nucleic acids such as transposons that is seen in diverse organisms.

In order to stably express transgenes in *C. elegans,* an elegant method to create single-copy transgene insertions was developed with the aid of a unique copy of the *Drosophila melanogaster* transposon MosI to induce a chromosomal double-strand break that promotes site-specific recombination with a plasmid-derived template (19, 21, 89, 90). Although single-copy transgenes can be stably expressed in the germline using this method, a GFP transgene with a piRNA target site in its 3'UTR (piRNA sensor) as well as other transgenes possessing GFP, can be subjected to rapid and permanent epigenetic silencing in the germline (19-21). Initiation of silencing depends on PRG-1, as single-copy transgenes are expressed if created in a *prg-1* mutant background, but become silenced within a single generation when wildtype PRG-1 is introduced. Transgene silencing then becomes independent of *prg-1* and is subsequently maintained for many generations by Mutator class secondary siRNA proteins as well as nuclear RNAi and chromatin proteins (19-21). This implies that initial targeting of transgenes by PRG-1 and associated piRNAs elicits

production of a secondary small RNA population whose maintenance, in conjunction with siRNA-directed with histone marks, is sufficient to enforce a stable, heritable silent state. This rapid process is termed 'RNAe' (small RNA-induced epigenetic silencing), though we suggest that the term 'single-generation RNAe' might be an appropriate descriptor.

Here I describe a process that we term 'multigenerational RNAe'. We identified single-copy transgenes that become silent if crossed in the hemizygous state for multiple generations. We found that multigenerational transgene silencing can become permanent and that small RNA silencing proteins previously shown to initiate or maintain silencing of transgene epialleles play unexpected roles in this multigenerational process. Our study provides new insights into the genesis and maintenance of epialleles, relevant not only to transgenes but also potentially to dynamic regulation of germline gene expression across generations in the vast majority of metazoans.

Results

Discovery of multigenerational transgene silencing

The Protection of Telomeres-1 (POT-1) single-stranded telomere binding protein inhibits telomerase and forms discrete foci at *C. elegans* telomeres *in vivo* (91). We previously created three independent single-copy transgene insertions that express POT-1::mCherry, *ypSi1*, *ypSi2* and *ypSi3* (91), each inserted via a transposon-induced double-strand break at the *MosI* locus *ttTi5605*, located near the center of Chromosome *II* (**Figure 2.1 A**) (89, 90). POT-1::mCherry fluorescence can be detected throughout the germline, and is particularly evident in meiotic pachytene germ cells as fluorescent punctae (**Figure 2.1 C**

& E) (91). Although several single-copy transgenes inserted in the *ttTi5605 MosI* locus are silenced in the F1 cross-progeny of a single outcross (20), we previously found that *pot-1::mCherry* transgenes remained robustly expressed following two successive crosses: one to combine *pot-1::mCherry* with a marker mutation, and a second to cross in a mutation with relevance to telomere biology (91). However, when we crossed *pot-1::mCherry* in trans to the recessive marker mutation *rol-6(e189)* for nine successive crosses - by crossing *pot-1::mCherry/rol-6* heterozygous males with *rol-6 -/-* homozygous hermaphrodites and then selecting for *pot-1::mCherry* homozygotes in the F3 - we found that POT-1::mCherry fluorescence was either weak or abolished (n=2 crosses each for three independent *pot-1::mCherry* insertions). Further propagation of these strains for a number of generations revealed that POT-1::mCherry expression was robustly restored for *ypSi1.9* and *ypSi3.9*, but that the *ypSi2.9* transgene had become permanently silenced. Neither nuclear nor telomeric POT-1::mCherry fluorescence was observed for the *ypSi2.9* strain during three years of culture in our laboratory, whereas the original non-crossed *ypSi2* strain displayed robust expression during this time. The presence of the *pot-1::mCherry* transgene in *ypSi2.9* was confirmed by PCR from genomic DNA, and rtPCR indicated that *pot-1::mCherry* mRNA is present in the original *ypSi2* transgene, hereafter referred to as *pot-1::mCherry,* but not for the crossed *ypSi2.9* transgene, hereafter referred to as *pot-1::mCherry.9* (**Figure 2.1 B**). Together, these results led us to hypothesize that a partially penetrant multigenerational silencing process can occur when single-copy transgenes are repeatedly crossed in *C. elegans*. This can ultimately result in a state of permanent transgene silencing that is stable for many generations.

I repeated the above cross by placing the *pot-1:mCherry* transgene in trans to the *rol-6* marker mutation and found that a fraction of *pot-1:mCherry / rol-6* animals were mCherry-negative starting at cross 3, although complete silencing of the *pot-1::mCherry* transgene was not achieved by cross 7 (**Figure 2.1 H**). Thus, a partially penetrant multigenerational silencing process occurs when the *pot-1::mCherry* transgene is crossed in trans to the *rol-6* marker mutation.

I next placed *pot-1::mCherry* in trans to the marker mutations *dpy-10 unc-4*, which flank the *pot-1::mCherry* transgene, and repeatedly crossed *pot-1::mCherry / dpy-10 unc-4* crossed heterozygous males with *dpy-10 unc-4* hermaphrodites (**Figure 2.1 D**). A proportion of silent *pot-1::mCherry* heterozygotes were observed for crosses 2 and 3, followed by uniform silencing for crosses 4 through 6 (**Figure 2.1E & I**). When F2 animals were revived from starved plates of cross 6 and homozygosed by selecting against the *dpy-10 unc-4* balancer mutations, the vast majority of *pot-1::mCherry* homozygotes showed POT-1::mCherry expression (28/30). However, 2/30 *pot-1::mCherry* homozygous lines were completely silent and remained silent for 8 additional generations (**Figure 2.1 I columns 6S**). I therefore conclude that a hemizygous single-copy transgene can become strongly silenced following multiple crosses, and that although this silencing reverts for many F3 homozygotes, a fraction of *pot-1::mCherry* homozygotes become permanently epigenetically silenced.

Independent transgenes succumb to multigenerational silencing

I next tested multigenerational silencing of an independent single-copy transgene, *oxSi487*, which expresses both mCherry- and GFP-tagged proteins from an operon

integrated into the *ttTi5605 MosI* locus, *Pmex-5::mCherry::H2B::tbb-2: 3'UTR::gpd-2 operon::GFP::H2B::cye-1 3'UTR* (**Figure 2.1 A & E**). The mCherry expression in this transgene was highly robust, whereas the GFP expression was substantially weaker, therefore, I used the mCherry fluorescence to score the animals. Crosses of *oxSi487* in trans to *dpy-10 unc-4* mutations resulted in silencing for some progeny in cross 2 and complete silencing by cross 3. F2 animals from starved plates of cross 3 were singled, and I examined the F3 progeny of four *oxSi487* homozygotes, all of which displayed uniform silencing of mCherry::H2B (**Figure 2.1 E & G columns 3S**). These strains displayed fully penetrant silencing for at least 8 generations without reversion, suggesting permanent transgene silencing. I also crossed *oxSi487* in trans to *rol-6* and found that a fraction of animals for four independent crosses were mCherry-negative from crosses 4 to 11, indicating weaker multigenerational silencing than for the *dpy-10 unc-4* marker mutations. *oxSi487* homozygotes isolated from starved *oxSi487/rol-6* cross 11 plates, where significant levels of silencing were observed, gave rise to F3 progeny in which 2/6 *oxSi487* homoyzygotes uniformly expressed mCherry, indicating reversion of silencing (**Figure 2.1F columns 11S**). However, 4/6 *oxSi487* homozygotes derived from cross 11 plates gave rise to uniformly mCherry-negative F3 progeny, which remained silent upon propagation for 8 additional generations, suggestive of permanent transgene silencing (**Figure 2.1E & F columns 11S**). Together, the above *pot-1::mCherry* and *oxSi487* experiments suggest that multigenerational transgene silencing occurs more quickly in trans to the *dpy-10 unc-4* marker mutations than in trans to *rol-6*, and that permanent transgene silencing can occur in either circumstance.

We next asked if multigenerational transgene silencing also occurs for transgenes that only contain GFP-epitope tags, and if transgene insertion via a double-strand break created by the Drosophila *MosI* transposon *ttTi5605* has any impact on transgenerational silencing. We therefore tested identical transgenes that express *Pmex-5::GFP::tbb-2 3'UTR* that were inserted either via *MosI*-mediated gene conversion at *ttTi5605*, *cpSi10[Pmex-5::GFP::tbb-2 3'UTR + unc-119(+)],* or via CRISPR/Cas9-mediated gene conversion at the identical position on Chromosome *II*, *cpIs10[Pmex-5::GFP::tbb-2 3'UTR + unc-119(+)]* (**Figure 2.2 A**) (92). We found that *cpSi10* and *cpIs10* underwent multigenerational silencing with very similar kinetics (**Figure 2.2 B & C**), indicating that the presence of the Drosophila *MosI* transposon at the *ttTi5605* locus prior to transgene insertion does not affect transgene silencing. When *cpSi10* and *cpIs10* were crossed in trans to our *dpy-10 unc-4* marker strain, c*pSi10* was silenced by cross 3, and *cpIs10* at cross 6 (**Figure 2.2 B, C & E**). Chunking of starved cross plates yielded F3 progeny homozygous for the transgene, where transgene expression was restored (**Figure 2.2 B column 3S, 2.2 C column 3S**). To determine if starvation was necessary for restoration of transgene expression in transgene homozygotes, we singled unstarved F1 progeny from cross 6 of *cpSi10/ dpy-10 unc-4* and *cpIs10/ dpy-10 unc-4* heterozygotes and found that the majority of F2 *cpSi10* or *cpIs10* homozygotes gave rise to F3 progeny that expressed GFP in the absence of starvation (**Figure 2.2 B, Figure 2.2 C columns H**). Several of the homozygous F3 progeny that did not express GFP were isolated and observed (**Figure 2.2 B column 9H, Figure 2.2 C columns 8H, 9H**). These lines remained GFP-negative, and fully penetrant silencing was transmitted for at least 5 generations. We conclude that a strong yet reversible

multigenerational silencing process occurs in transgene heterozygotes, and that this silencing can become permanent for a fraction of transgene homozygotes.

We next asked if a GFP insertion at a distinct genomic location, the *his-72* locus on Chromosome *III*, could also be subjected to multigenerational silencing (92). *his-72::GFP* was placed in trans to the marker mutations *vab-7 dpy-18* and crossed versus these mutations (**Figure 2.7**). Similar to our other transgenes tested, a significant fraction of *his-72::GFP / vab-7 dpy-18* heterozygotes displayed transgene silencing for crosses 3 to 22 (**Figure 2.2 D & E**). Thus, multigenerational transgene silencing is not specific to single-copy transgenes at the *ttTi5605* locus on Chromosome *II*, and can regulate the expression of GFP cassettes inserted at endogenous genomic loci.

Outcrossing cpSi10 and cpIs10 with other markers

We decided to test a variety of balancer mutations related to the *dpy-10 unc-4* mutations used, which had been outcrossed numerous times to our laboratory stock of N2 (wildtype). We tested an unoutcrossed *dpy-10 unc-4* obtained from the CGC, *dpy-2 unc-4*, and a strain containing an inversion between *dpy-10* and *unc-4* (*mIn1; unc-4 dpy-10*). When *cpSi10* and *cpIs10* were crossed in trans to *dpy-2 unc-4* we found a fully penetrant multigenerational transgene silencing phenotype similar to our *dpy-10 unc-4* result (**Figure 2.5, Figure 2.6**). When *cpSi10* and *cpIs10* were crossed in trans to either the unoutcrossed *dpy-10 unc-4* or the inversion *unc-4 dpy-10* marker strains, we found a partially penetrant multigenerational transgene silencing phenotype, similar to that observed with *rol-6* for the *mCherry* transgenes above (**Figure 2.5, Figure 2.6**). Since several marker strains

resulted in multigenerational transgene silencing, it is likely that unidentified secondary mutations are not responsible for this effect.

Coupling of initiation and maintenance of multigenerational transgene silencing via small RNA proteins

RDE-2 is a component of the Mutator-class of RNA interference proteins that promote biogenesis of secondary siRNA populations, which were previously shown to be required for silencing of single-copy transgenes that become silenced in a single cross (19-21), as well as for cosuppression-mediated silencing of repetitive transgenes (23, 80, 93-95). I found that POT-1::mCherry fluorescence was partially restored when the silenced *pot-1::mCherry.9* transgene was placed in an *rde-2*–deficient background (**Figure 2.3 A & B**). Thus, an siRNA-mediated process promotes multigenerational transgene silencing.

Secondary siRNA proteins such as RDE-2 have been shown to promote maintenance of single-generation transgene silencing (78, 80). In contrast, the piRNA-interacting Argonaute protein PRG-1 is required for initiation of single-generation transgene silencing but is dispensable for maintenance of transgene silencing (20, 96). I found that independent alleles of *prg-1*, *tm872* and *n4357*, elicited partial desilencing of *pot-1::mCherry*.9 (**Figure 2.3 A & B**). As *prg-1* and *rde-2* mutations both resulted in partial desilencing of *pot-1::mCherry.9*, I constructed *rde-2 prg-1(tm872); pot-1::mCherry.9* and *rde-2 prg-1(n4357); pot-1::mCherry.9* strains and found that POT-1::mCherry fluorescence was fully restored in both cases (**Figure 2.3 A & B**). Thus, proteins that play distinct roles in initiation or maintenance of single-generation single-copy transgene silencing coordinately promote maintenance of multigenerational transgene silencing.

As a control, I tested the RSD-6 Tudor domain protein, which has been previously described to promote spreading of RNAi from the soma to the germline and generation of robust secondary siRNA populations in response to exogenous dsRNA triggers, but is dispensable for single-generation silencing of the piRNA sensor transgene (21, 97, 98). *pot-1::mCherry* silencing was maintained when *rsd-6* was mutant (**Figure 2.3 A & B**). I also found that the NRDE-3 Argonaute protein, which promotes transcriptional silencing in response to exogenous dsRNAs (99), was dispensable for *pot-1::mCherry* silencing (**Figure 2.3 A & B**).

Lack of chromosomal pairing induced rapid silencing

ZIM-1 is a member of a family of four related C2H2 zinc-finger domains proteins that promote pairing of specific chromosome homologs during meiosis (100). ZIM-1 itself promotes pairing of Chromosomes *II* and *III,* as meiotic pairing of these chromosomes is completely abrogated in *zim-1* mutants (101, 102). When *pot-1::mCherry; zim-1* males were crossed with *dpy-10 unc-4; zim-1* hermaphrodites, uniform mCherry silencing occurred in the progeny of cross 1 (n=10, F1 hermaphrodites scored) (**Figure 2.8**). The same result was observed for progeny of *oxSi487; zim-1* males crossed with *dpy-10 unc-4; zim-1* hermaphrodites. Therefore, complete lack of chromosome pairing leads to rapid single-generation transgene silencing for transgenes that normally succumb to multigenerational transgene silencing.

Discussion

Lack of a pairing partner in meiosis is known to alert an RNA directed-silencing mechanism termed Meiotic Silencing of Unpaired DNA (MSUD), which was first identified in *Neurospora crassa* (103). MSUD occurs when a diploid organism suppresses expression of genes not paired with their homologs during the first prophase of meiosis (104-106). MSUD requires small RNA factors that include an Argonaute protein and an RNA-dependent RNA polymerase that likely act on unpaired meiotic DNA to produce dsRNA that feeds into a siRNA-mediated silencing system (103, 107-109).

We found that transgene silencing occurred when *single-copy transgene / marker mutation* heterozygotes were crossed with *marker mutation* homozygotes for multiple generations. While potent, this multigenerational transgene silencing process was reversible, as transgene activation was commonly observed when transgene homozygotes were isolated from silent *single-copy transgene / marker mutation* heterozygotes. These results suggest that the trigger of multigenerational transgene silencing corresponds to unpaired transgene DNA during meiosis, which occurs if one homolog possesses a transgene and the other does not. Thus, the multigenerational transgene silencing that we report here possesses a physical hallmark of MSUD.

Initiation and maintenance of transgene silencing are clearly separable for single-copy GFP transgenes that are rapidly silenced in a single outcross via a process that is initiated by the germline Argonaute protein PRG-1 but then maintained by secondary siRNA biogenesis proteins and small RNA-mediated nuclear silencing factors (20, 96). This rapid silencing process, which we term 'single-generation RNAe', is related to the

multigenerational silencing process that we define here. We found that although multigenerational silencing of hemizygous transgenes was often reversed when a transgene becomes homozygous, although permanent transgene silencing was observed for some transgene homozygotes. We term this reversible silencing process that is potent enough to become permanent 'multigenerational RNAe'.

Maintenance of silencing for transgenes silenced by multigenerational RNAe can require both PRG-1, which is dispensable for maintenance of single-generation RNAe, as well as the secondary siRNA protein RDE-2. Deficiency for either of these proteins led to partial transgene desilencing for the *pot-1::mCherry.9* transgene, and full desilencing was observed when both PRG-1 and RDE-2 were removed. These results imply that multigenerational RNAe represents a state of epigenetic silencing that is distinct from single-generation RNAe, where initiation and maintenance phases of transgene silencing can remain mechanistically coupled once the silent state has been established..

Recent work has identified an anti-silencing Argonaute protein CSR-1 that promotes germline transgene expression in *C. elegans*, and could generally protect endogenous germline genes from silencing (96, 110-112). Transgenes that are active in the germline and licensed by CSR-1 have been shown elicit a multigenerational anti-silencing effect, in which distinct silent transgenes with homologous sequences can be activated (96, 110). The direct competition between CSR-1 and PRG-1 to mark a gene for either licensing and protection, or silencing and repression, may explain why several transgenes examined neither reached a state of complete silencing, nor demonstrated reversibility of the silencing upon isolation of the F3 homozygotes. I speculate that transgenes resistant to

permanent and stable silencing may be regulated by an abundance of CSR-1, which would prevent PRG-1 from initiating multigenerational RNAe.

Our results are consistent with a model derived from studies of epigenetic silencing in plants, where initiation and maintenance loops can promote silencing mediated by small RNA-directed DNA methylation, for example in response to viral DNA (83, 113). Once silencing is established, the initiation loop can become dispensable for silencing, as is the case for PRG-1-initiated rapid silencing of single-copy GFP transgene. However, the initiation loop can contribute to maintenance of a silent state (83), as I show for PRG-1 in multigenerational silencing of the *pot-1::mCherry.9* transgene. These results are consistent with a known role for PRG-1 in maintenance of silencing for a subset of transposons: the Tc3 DNA transposon becomes desilenced and active when *prg-1* is mutant, but not Tc1 or Tc4 transposons, all of which become active when secondary siRNA biogenesis is disrupted (24). Overall, our results are consistent with genome silencing activities that can remain at least partially dependent on PRG-1 function, in contrast to the discrete role of PRG-1 in initiation, but not maintenance, of single-generation RNAe. Further, the reversibility of multigenerational hemizygous transgene silencing strongly contrasts with permanent and immediate transgene silencing of single-generation RNAe, indicating the multigenerational RNAe is a distinct, though clearly related, form of genome silencing.

I postulate that unpaired meiotic transgene DNA that persists for multiple generations in the germline is identified by PRG-1 as an unwanted non-self element, which leads to secondary siRNA production via *Mutator*-class proteins such as RDE-2, and that heritable silencing of multigenerational RNAe across generations remains dependent on PRG-1/piRNA-driven production of secondary siRNAs that target the transgene locus (Fig.

4B). Since rapid transgene silencing, like that which acts on the 'piRNA sensor,' is promoted by PRG-1/Piwi and opposed by the CSR-1 Argonaute protein in *C. elegans*, the multigenerational silencing process defined in our study likely represents a competition between pro- and anti-silencing small RNA systems, where silencing is promoted by maintenance of hemizygosity for multiple generations (**Figure 2.4**).

The formation of epialleles in reaction to foreign DNA has been well documented in *A. thaliana*. Although plants possess a system of DNA methylation and *C. elegans* lacks such mechanisms, other epigenetic alterations that affect chromatin states do occur in the nematode (114-116). Epigenetic modification in plants is accomplished via the RNA-directed DNA methylation pathway, where RNA-dependent RNA polymerases bind single-stranded non-coding RNAs and create small dsRNAs that are then cleaved by Dicer and loaded into an Argonaute complex (117-121), which target both CG and non-CG sites for methylation (122). Once established, the cytosine methyl marks are associated with an epigenetic state that is transmitted to future generations. Maintenance is mediated by specific DNA methyltransferases which copy epigenetic information to the daughter strand of DNA during meiosis resulting in the establishment of an epiallele that can be passed on from one generation to the next (117, 123).

Genome-wide association studies have revealed that DNA sequence polymorphisms explain only a minor fraction of variability for a number of common human traits such as Type II Diabetes. Some of this missing heritability may be due to epigenetic modification of the genome (124-126), as reflected by creation of numerous de novo epialleles in F2 lines derived from crosses between different *Arabidopsis thaliana* ecotypes (117, 127). Although there are a number of explanations for creation of such epialleles, the genomes of diploid

species typically display numerous copy number changes between homologous chromosomes. Gain or loss of copies of a specific segment of the genome has been estimated to account for up to 5% of the variability between human individuals (128, 129). This implies that significant levels of unpaired DNA occur during most meioses. I propose that the multigenerational epigenetic silencing process described in our study, likely triggered by lack of pairing during meiosis, can elicit both permanent and reversible forms of silencing that may be relevant to heritable epigenomic variation in many species, with pertinent yet complex implications for understanding and predicting heritability in the context of human disease.

Materials and Methods

Strains

All strains were cultured and maintained at 20°C on nematode growth medium plates seeded with Escherichia coli OP50. Strains used include: Bristol N2 ancestral, WM161 *prg-1(tm872) I*, SX922 *prg-1(n4357) I*, WM29 *rde-2(ne221) I*, *rsd-6(yp11) I*, *dpy-2(e8) unc-4(e120) II*, DR103 *dpy-10(e128) unc-4(e120) II*, *mIn1[unc-4(e120) dpy-10(e128)] II* derived from DR2054, YA1198 *ypIn3 [Pdaz-1:pot-1::mCherry::tbb-2utr]*, EG6787 *oxSi487 [Pmex-5::mCherry::H2B::tbb-2: 3'UTR::gpd-2 operon::GFP::H2B::cye-1 3'UTR + unc-119(+)] II*, CB187 *rol-6 (e187) II*, LP135 *cpSi10 [Pmex-5::GFP::tbb-2 3'UTR + unc-119(+)] II; unc-119(ed3) III*, LP136 *cpIs10 [Pmex-5::GFP::tbb-2 3'UTR + unc-119(+)] II; unc-119(ed3) III*, vab-*7(e1562) dpy-18(e364) III*, LP148 *unc-119(ed3) his-72(cp10[his-72::gfp + LoxP unc-119 (+) LoxP])* III, CA324 *zim-1(tm1813) IV;* YY158 *nrde-3 (gg66) X*.

Microscopy

Live 1-day old young adult worms were mounted on 2% agarose pads in 7μL of 2mg/mL levamisole. Strains were examined in widefield using a Nikon Eclipse E800 microscope both under differential interference contrast (DIC) and a 595-nm excitation wavelength at x20, x60 or x100 magnifications. Confocal microscopy was performed using a Zeiss LSM710 laser-scanning microscope with a C-APO 40x/1.2 NA objective at 2.0x optical zoom. Images were obtained using a 561nm diode laser for excitation, and emission was collected at 563-701nm. Bright field images were collected simultaneously.

Fluorescence quantification

Animals were mounted as described above, and Z-stacks were taken using confocal microscopy as described above, within 2 hours of mounting. Zeiss ZEN 2009 software was used to collect and compile the stacks. Analysis of fluorescence in the nuclear volume (arbitrary units) was performed using ImageJ v1.45S and the "Measure Stack" volume measurement plugin (developed by RF Dougherty and available at http://www.optinav.com/imagej.html). For each genotype, 30 animals were selected and 5 nuclei were quantified from each animal.

rtPCR

Animals were collected in Trizol®. Control and experimental animals were grown alongside each other under identical conditions. RNA extraction was performed using standard protocols. cDNA was synthesized using SuperScript III reverse transcriptase (Invitrogen) with random hexamers. Primer pair sequences used are as follows: actin control (F-GATATGGAGAAGATCTGGCATCA and R-GGGCAAGAGCGGTGATT); mCherry (F-TGGTCCAATTTCGTGGTTTATATCCTC and R-CTTTGCTCTTCGCCATTGTTTCC).

Figures

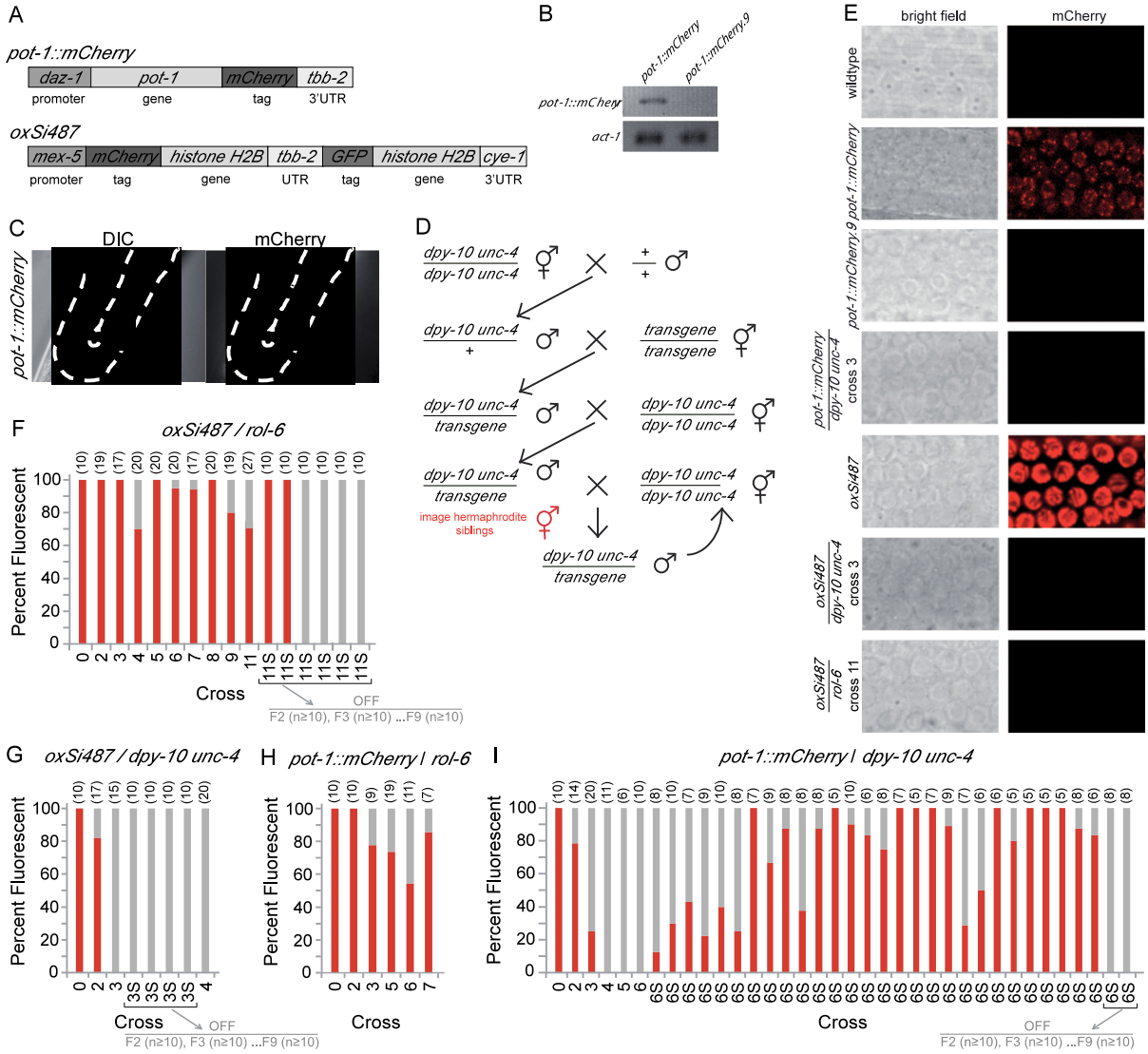


Figure 2.1. mCherry-expressing transgenes are capable of being silenced in a multigenerational manner via crossing. (A) Structure of *pot-1::mCherry* and *oxSi487* transgenes. (B) rtPCR of strains containing the *pot-1::mCherry* transgene. Primer sequences for *pot-1::mCherry* and *act-1* control in Methods. (C) DIC (left) and mCherry fluorescent image (right) of *pot-1::mCherry*. Dotted white line outlines the mitotic and meiotic germline. (D) Crossing schema used to cross Chromosome *II* transgenes. (E) Confocal images of bright field (left column) and mCherry fluorescence (right column) in the mitotic germline nuclei in various strains. Wildtype is the N2 Bristol ancestral strain. (F-I) Percent of population fluorescent at each round of crossing for *pot-1::mCherry* and *oxSi487* using either *rol-6* or *dpy-10 unc-4* marker strain. Numbers at top in parenthesis indicate number of worms scored. Bar 0 indicates original transgenic strain prior to crossing. S denotes homozygous F3 worms derived from starved heterozygous *transgene / marker*.

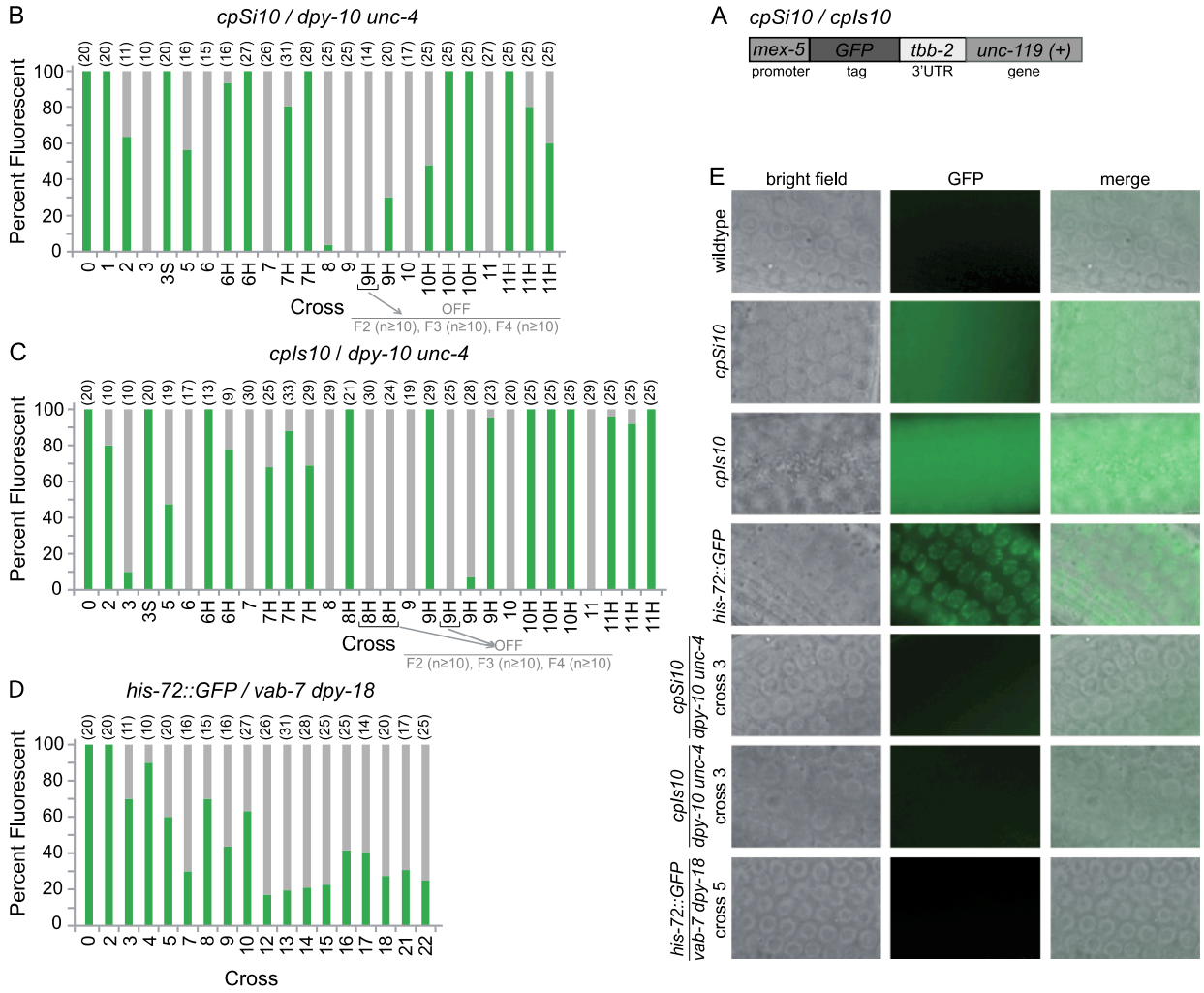


Figure 2.2. GFP-expressing transgenes are capable of being silenced in a transgenerational manner via crossing. (A) Structure of *cpSi10* and *cpIs10* transgene. *cpSi10* was inserted in an inverse orientation to that of *cpIs10.* (B-D) Percent of population fluorescent at each round of crossing for *cpSi10, cpIs10,* and *his-72::GFP* using *dpy-10 unc-4* or *vab-7 dpy-18* marker strain. Numbers at top in parenthesis indicate number of worms scored. Bar 0 indicates original transgenic strain prior to crossing. S denotes homozygous F3 worms derived from starved heterozygous *transgene / marker*. H denotes homozygous F3 worms derived from heterozygous *transgene / marker.* (E) Widefield microscopy images of the mitotic germline nuclei in bright field (left column), GFP fluorescence (center column), and a merged (right column). Wildtype is the N2 Bristol ancestral strain.

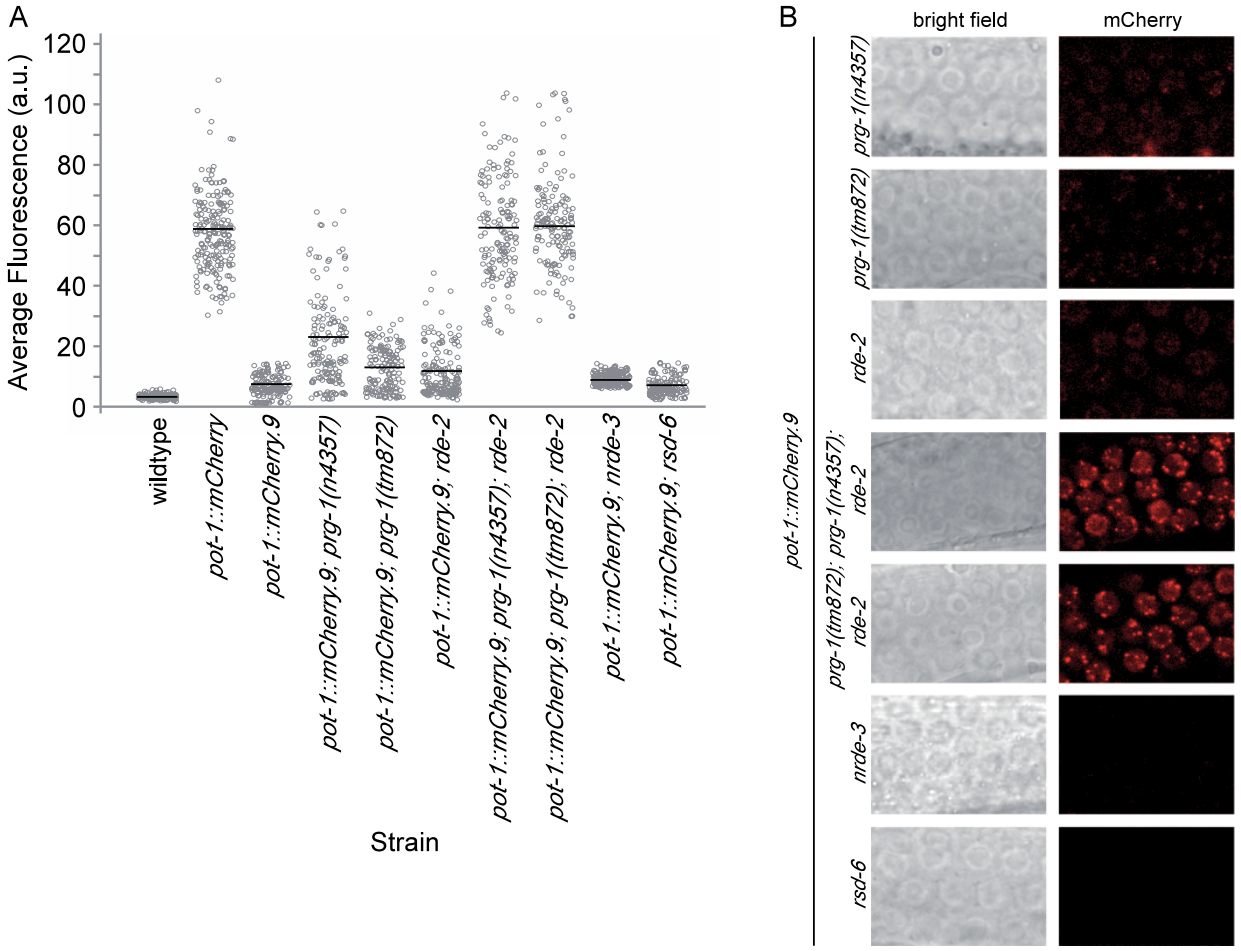


Figure 2.3. *prg-1* and *rde-2* coordinately promote maintenance of the silencing of the *pot-1::mCherry* transgene. (A) Average fluorescence quantified for individual nuclei (n=150 for each genotype, 5 nuclei counted per animal). Solid horizontal bars indicate the mean intensity for each population. Wildtype is the N2 Bristol ancestral strain. (B) Confocal images of both bright field (left column) and mCherry fluorescence (right column) in the mitotic germline nuclei in various strains in the *pot-1::mCherry.9* transgenic background.

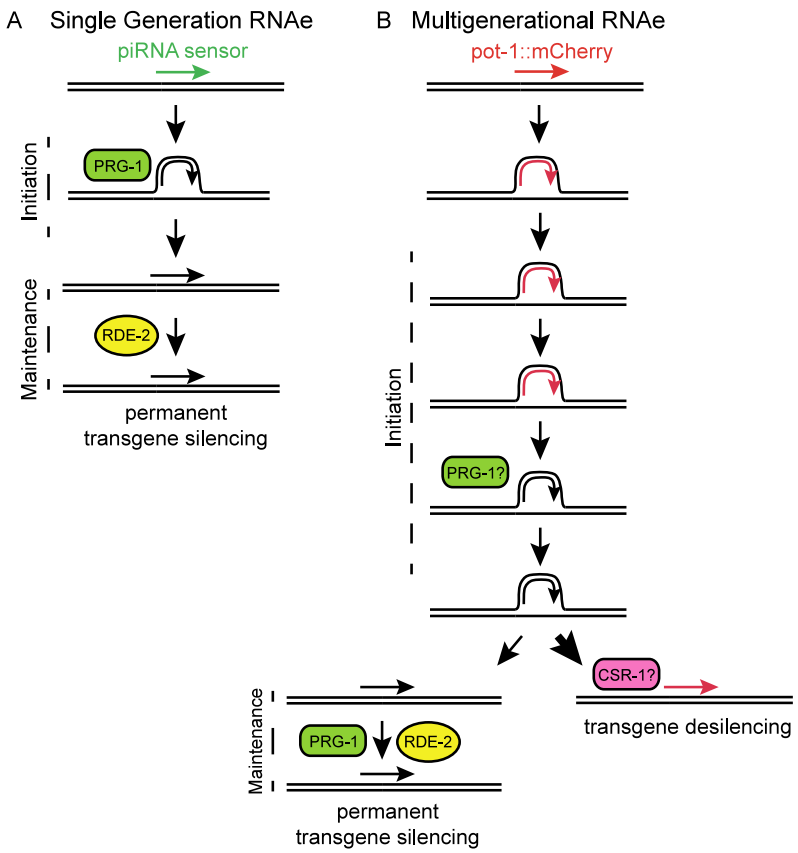


Figure 2.4. Model depicting (A) single generation RNAe, as previously described, and (B) multigenerational RNAe.

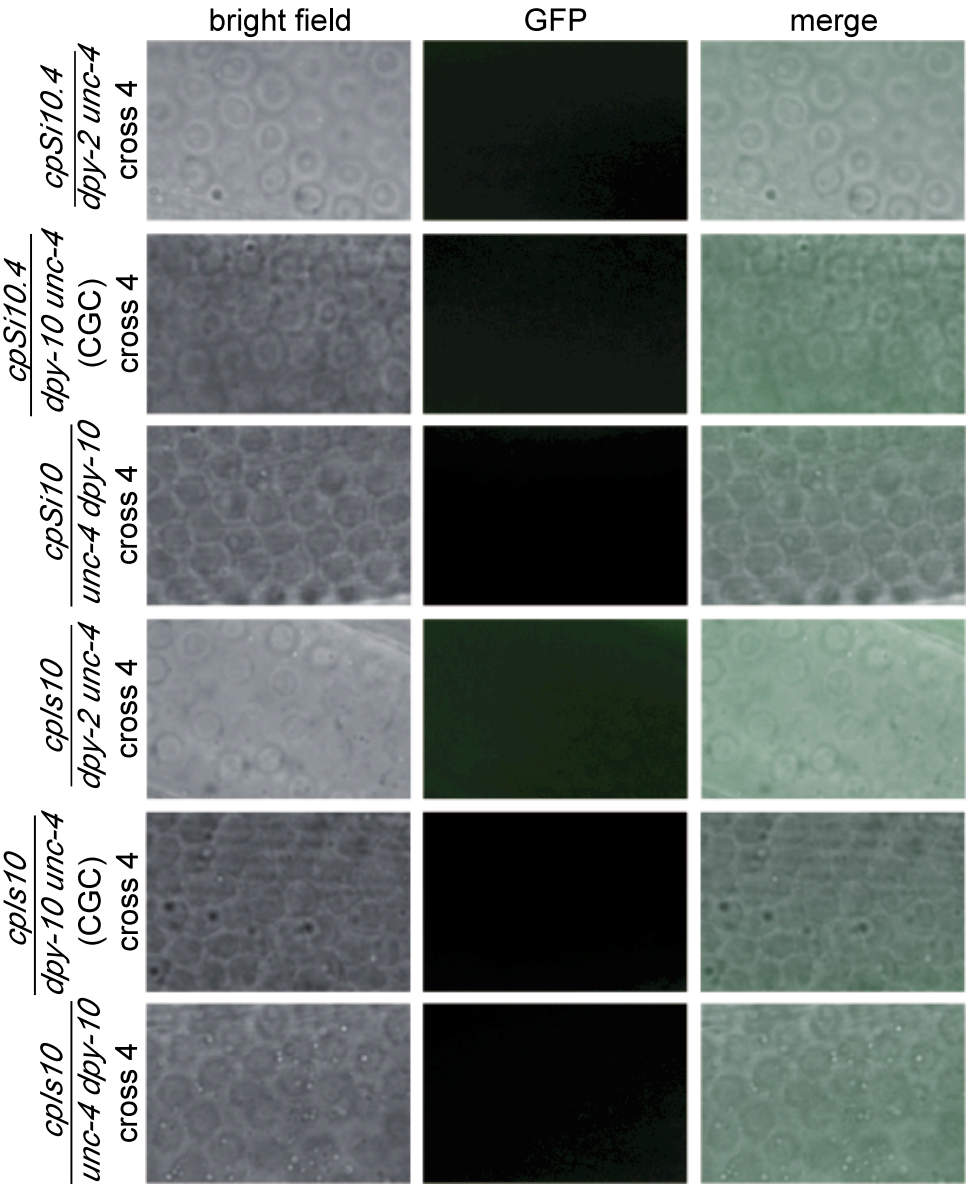


Figure 2.5. Widefield microscopy images of bright field (left column), GFP fluorescence (center column), and a merged (right column).

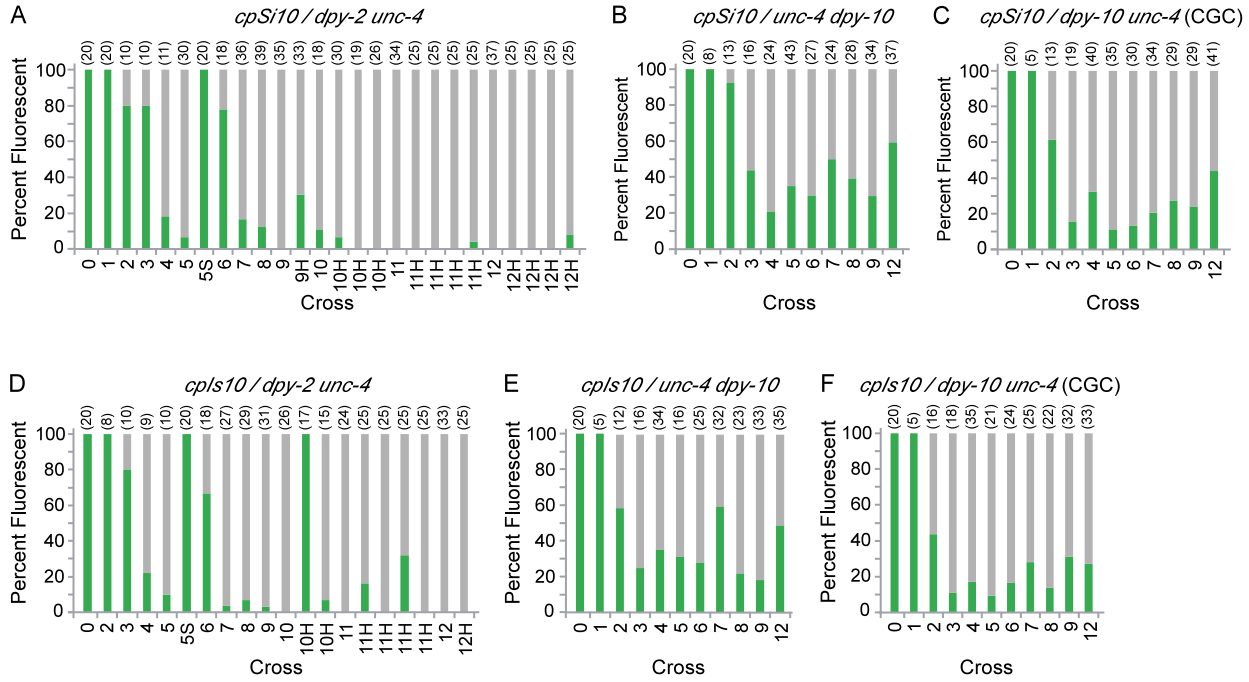


Figure 2.6. Crossing with various markers induces silencing of *cpSi10* and *cpIs10* transgenes. (A-F) Percent of population fluorescent at each round of crossing for cpSi10 and *cpIs10* using either dpy-2 unc-4, unc-4 dpy-10, or dpy-10 unc-4 (obtained from Caenorhabditis Genetics Center) marker strain. Numbers at top in parenthesis indicate number of worms scored. Bar 0 indicates original transgenic strain prior to crossing. S denotes homozygous F3 worms derived from starved heterozygous transgene / marker. H denotes homozygous F3 worms derived from heterozygous transgene / marker.

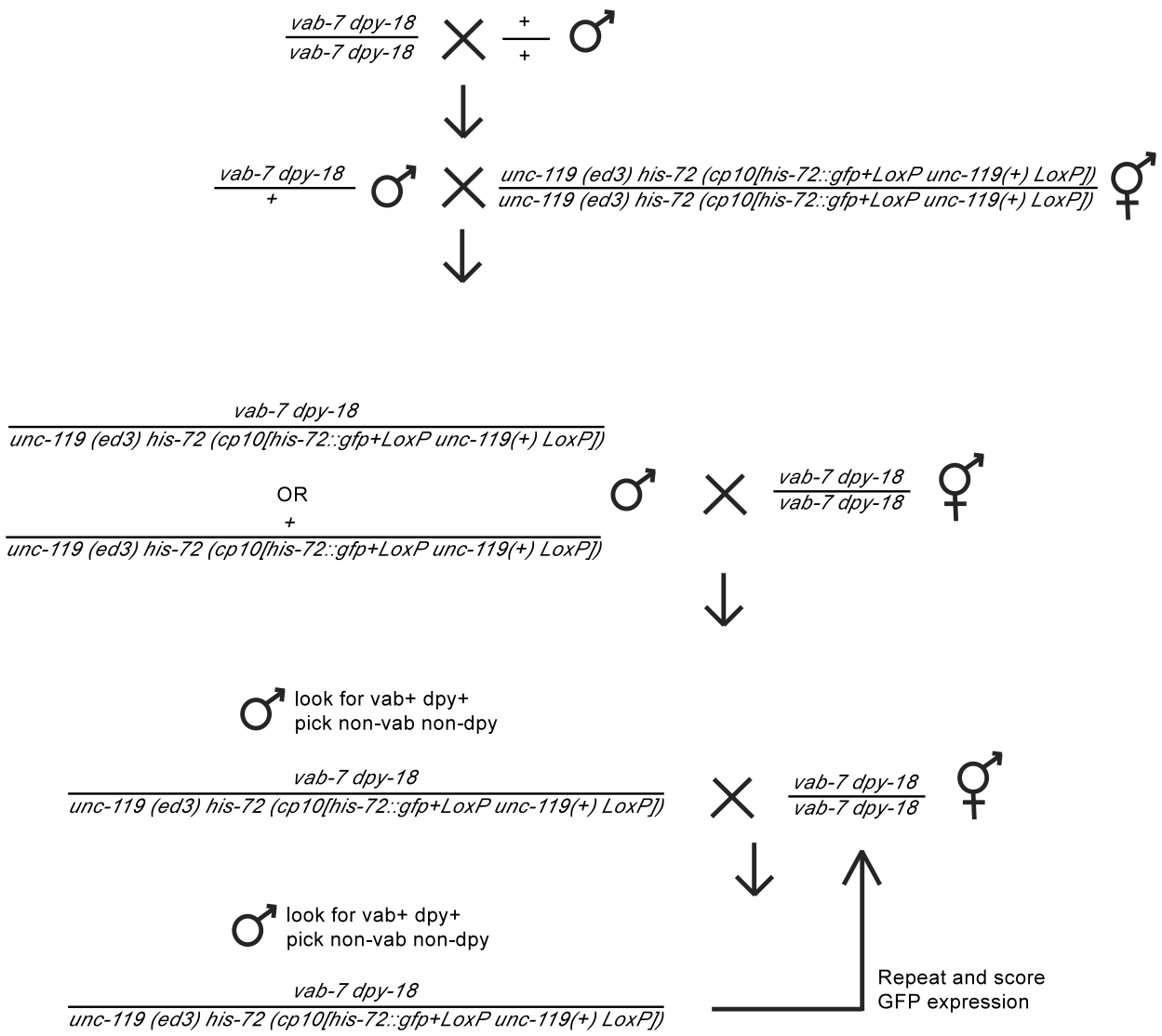


Figure 2.7. Schema used to cross the *his-72*::GFP transgene using Chromosome III markers *vab-7 dpy-18*.

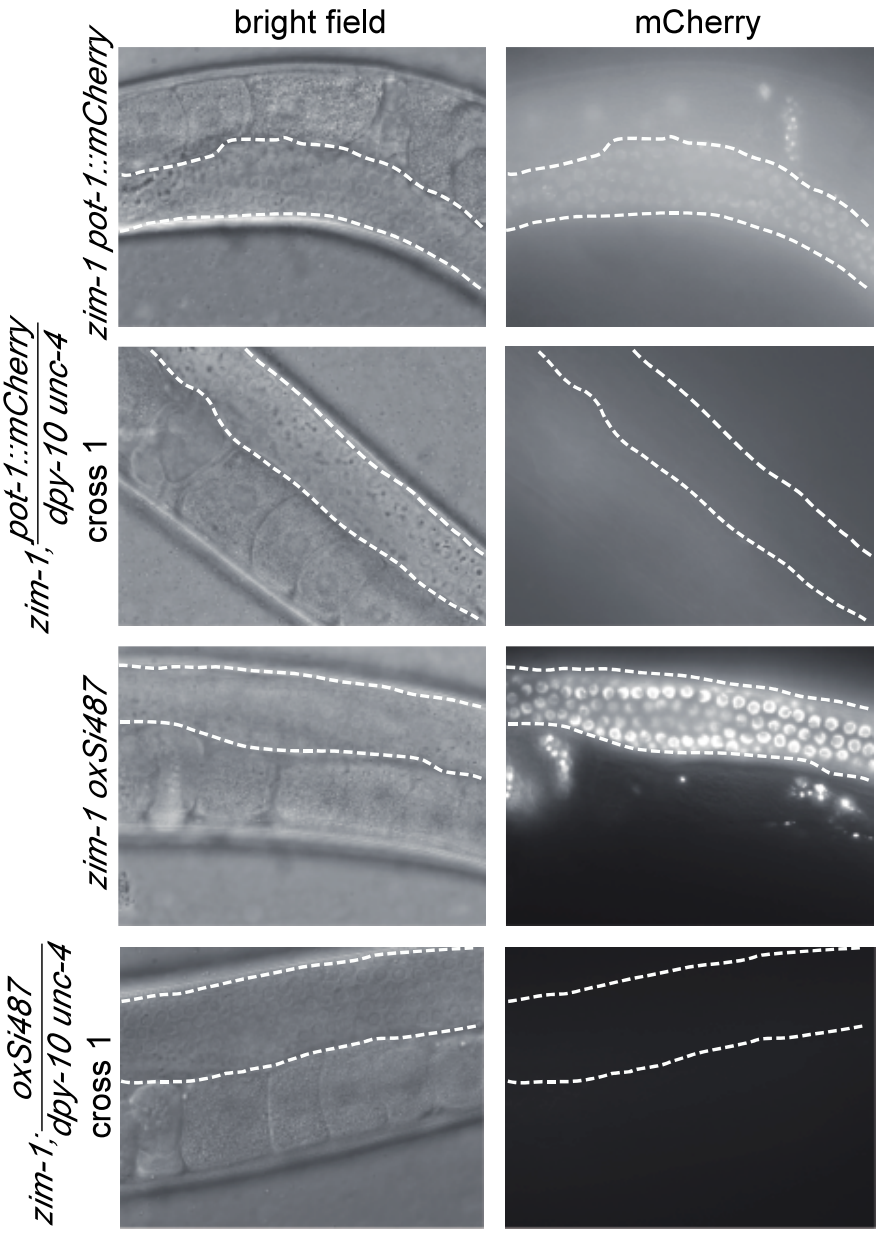


Figure 2.8. Widefield microscopy depicting bright field (left column) and mCherry fluorescent (right column) images of (top-bottom); the zim-1 mutant in the pot-1::mCherry transgenic background alone (n=24), and crossed once with dpy-10 unc-4 (n=11); and the oxSi487 transgene in the pot-1::mCherry transgenic background alone (n=11), and crossed once with dpy-10 unc-4 (n=10). Dotted white line outlines the mitotic germline.

CHAPTER 3: SUMMARY & FUTURE DIRECTIONS

CHAPTER 1

We isolated a novel allele of *smg-1*, *smg-1(yp3),* from a large EMS mutagenesis screen. SMG-1 is a PI3K-like kinase known to be involved in NMD (nonsense-mediated mRNA decay) (9). Upon sequencing, I found that our novel allele contained a missense mutation in the PI3K domain – a G1774E change (**Figure 1.2**). Like other *smg-1* alleles *r904, cc545* and *cc546*, the *yp3* mutation was found to stabilize transcripts normally degraded by NMD, for example the *unc-30(e191)* transcript which contains the canonical PTC (premature termination codon) target of NMD (**Figure 1.6**), as well as the pCM1.4 transgene which contains an extra long 3'UTR (**Figure 1.5**), and the more complex *unc-17 cha-1* bicistronic message (**Figure 1.6**).

However *smg-1(yp3)* is unique in that it displays a separation-of-function phenotype in that it does not stabilize the *unc-54(r293)* transcript, whereas the other alleles of *smg-1* do (**Figure 1.3**). This indicates that the mechanism of regulation of *unc-54(r293)* is a process related to NMD in that SMG-1 (and presumably other NMD machinery) is used, but has different requirements than canonical NMD (**Figure 1.7**). I found that although the mutation in *unc-54(r293)* resides in the 3'UTR, deleting both the 3' cleavage and poly-adenylation sites, the mRNA produced is extra-long rather than too short (**Figure 1.4**). This is because instead of transcription halting at the end of the *unc-54* gene, the code is read through the subsequent gene, *aex-5*.

In order to further elucidate the exact nature of this novel mRNA processing pathway, I would like to examine other alleles with canonical PTCs and aberrant UTRs whose phenotypes been shown to be *smg*-suppressible. If others demonstrate a separation-of-function for *smg-1(yp3),* we can build a more complete picture of the abnormalities targetted by this novel pathway. Candidates include *unc-97(su110)* and *tra-2(e1209).*

Determining the exact molecular trigger for this pathway would go a long way towards helping our understanding. We know the sequence of the linker region in between the *unc-54* and *aex-5* sequences in the *unc-54(r293)* transcript, and that there are putative binding sites for miRNAs contained within it. An *in vitro* reporter-gene assay would give us a better idea as to whether any of the miRNAs are actually binding to the linker region and perhaps targetting it. Another method for elucidating whether or not the linker region contains the molecular target would be to make transgenes that contain the linker region – or versions thereof – and inject them into animals to see whether they would be targeted in the same manner as the native *unc-54(r293)* transcript.

Another interesting avenue to explore would be the verification of satellite and tandem repeats that preliminary work shows might be regulated in a similar manner to *unc-54(r293)*. The first step would be to create a method to systematically go through the microarray data already accumulated to find locations that show the same pattern of regulation – stabilization via *smg-1(r904), smg-1(cc545)* and *smg-1(cc546)*, and suppression via *smg-1(yp3).* These putative targets could then be confirmed via PCR. These results would presumably give us more information about what kinds of targets are utilized by our novel pathway, and therefore give us insight into how it functions to regulate the expression of certain mRNAs.

In this study, we identify a novel mechanism capable of silencing single-copy transgenes that is triggered by a lack of chromosomal pairing during meiosis. In contrast to a previously described 'RNAe' mechanism, which silences transgenes in a single generation, we define a distinct mechanism that requires several generations of successive crossing to elicit complete silencing. The PRG-1 Argonaute was found to promote maintenance of silencing, which expands current knowledge of its role in initiation to include maintenance of single-generation transgene silencing. Our analysis of 'multigenerational RNAe' offers new insights into how organisms facilitate permanent, heritable epigenetic changes to the genome.

Upon crossing the single-copy transgene insertion expressing POT-1::mCherry, *ypSi2*, with the *rol-6* marker 9 times successively, we observed that its expression had become silenced – and remained so during 3 years of culture in our laboratory. I determined via rtPCR that this resulting transgene – now called *pot-1::mCherry.9* – does not express the corresponding mRNA, and therefore the mechanism governing this regulation is operating at the RNA level. I was able to repeat this result using the same crossing system with the *dpy-10 unc-4* marker, with uniform silencing appearing at generations 4 through 6. However, expression of POT-1::mCherry was rescuable in the majority of the cases upon homozygosing of the strains (**Figure 2.1**). Several other transgenes – *oxSi487* (**Figure 2.1**), as well as GFP-expressing *cpSi10* and *cpIs10* (**Figure 2.2**) – underwent multigenerational silencing with very similar kinetics. These transgenes not only expressed different fluorescent tags, but were created using various insertion techniques – either the MosI transposon insertion system or the CRISPR/Cas9-mediated gene conversion – to insert the

transgene into a specific Chromosome *II* locus. This suggests that the transgenerational silencing effect that we observe is not merely a result of the insertion technique. Even a *his-72::GFP* transgene inserted into a different location on Chromosome *III* displays a level of silencing (**Figure 2.2**).

In order to elucidate the mechanism which promotes this multigenerational transgene silencing, I adopted a genetic strategy. Double mutants were created using various candidates in the *pot-1::mCherry.9* background. I found that transgene fluorescence was partially restored in both *rde-2 and prg-1* deficient backgrounds, and complete restored where both were simultaneously mutant in the *pot-1::mCherry.9* background (**Figure 2.3**). RDE-2 promotes the biogenesis of secondary siRNA populations, and has been previously shown to promote the maintenance of traditional single-generation transgene silencing (78, 80). However, PRG-1 is a piRNA Argonaute protein which has been shown to be required for the initiation of transgene silencing, but is dispensable for the maintenance (20, 96). Thus, we have identified a novel situation in which proteins that act in initiation or maintenance act in tandem to promote maintenance of transgene silencing.

A potential molecular target for this silencing mechanism is the lack of paired DNA at the transgene locus during meiosis. To test this I repeated the successive crossing of the *pot-1::mCherry* and *oxSi487* transgenes with the *dpy-10 unc-4* marker, however in this case both the male and the hermaphrodite being crossed were deficient for *zim-1.* In the *zim-1* mutant background, Chromosomes *II* and *III* completely fail to pair in meiosis. In both cases, the onset of silencing was much more rapid and was observed in the progeny of the first cross (**Figure 2.8**).

To further understand the nature of this multigenerational transgene silencing, it would be prudent to determine if this effect was being transmitted through the male or female germline, or whether a contribution from both is required. One possibility is to repeat the *zim-1* experiment and exclude the *zim-1* mutation from the *dpy-10 unc-4* marker background to determine whether the male contribution is necessary and sufficient. Very preliminary experiments indicate that this is most likely the case. We also have preliminary data demonstrating that when transgenic hermaphrodites are crossed once into the mutant background, and then propagated without crossing for 5-8 generations, silencing fails to occur. Continuing these experiments and replicating with other transgenes and markers would solidify the conclusion that the trigger for silencing is transmitted through the male germline.

Additionally, testing our transgenes for rapid silencing in the other *zim* mutant backgrounds would indicate whether or not the accelerated silencing effect I see is due to lack of pairing of the specific chromosome the transgene is located on, as *zim-1, 2,* and *3* govern the pairing of different chromosomes.

The genetic results from **Figure 2.3** are a powerful indicator of what mechanisms might be governing the silencing we observe. In addition to the results described above, I have preliminary evidence that *rde-1* is also governing the silencing of *pot-1::mCherry.9.* This is surprising as RDE-1 is the Argonaute known to function in response to the presence of foreign dsRNA, and a dual-Argonaute system has yet to be observed for one target. I would like to test whether RDE-4, which forms a complex with RDE-1 to bind dsRNA, is involved in the silencing of *pot-1::mCherry.9* by creating a double mutant with *rde-4* in the transgenic background. If *rde-1* and *rde*-4 are both shown to regulate the silencing of *pot-1::mCherry.*9,

this would suggest that a dsRNA intermediate is being derived from the homozygous transgene sequence. I could then assay directly for the presence of the transgenic dsRNA using a biochemical assay currently under development in our laboratory. If such a dsRNA species were to be identified, this would strongly implicate it as the source of the silencing trigger. This dsRNA would most likely be upstream of piRNAs or another kind of secondary siRNA in this particular system.

I would also like to test the *pot-1::mCherry.9* transgene with other candidates, for instance other players involved secondary siRNA production and targeting like *hrde-1, nrde-1, nrde-2, nrde-3* and *nrde-4.* It would also be interesting to investigate chromatin modifying genes that may be acting downstream from the RNA processing pathway in order to create a permanently encoded form of heritable silencing, like *rbr-2, spr-2,* and *spr-5.*

To ensure that the same mechanism is governing all of the transgenes we are studying, it would be prudent to test the other transgenes – *oxSi487, cpsI10, cpIs10,* and perhaps even *his-72::GFP* – in the various candidate backgrounds. Discovery of additional transgenes that undergo similar silencing upon successive crossing would also aid in the understanding of this mechanism.

An intriguing idea is the possibility that there is a mutagenic effect that causes gene loss in the events where we witness permanent, stable silencing. Comparing the sequence of active transgenes versus silent transgenes, we could determine quite simply whether or not this is the case.

REFERENCES

1. Ambros V (2004) The functions of animal microRNAs. *Nature* 431(7006):350-355.

2. Fire A*, et al.* (1998) Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature* 391(6669):806-811.

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

4. Matera AG, Terns RM, & Terns MP (2007) Non-coding RNAs: lessons from the small nuclear and small nucleolar RNAs. *Nat Rev Mol Cell Biol* 8(3):209-220.

5. Mattick JS & Makunin IV (2006) Non-coding RNA. *Hum Mol Genet* 15 Spec No 1:R17-29.

6. Greider CW & Blackburn EH (1985) Identification of a specific telomere terminal transferase activity in Tetrahymena extracts. *Cell* 43(2 Pt 1):405-413.

7. Fasken MB & Corbett AH (2005) Process or perish: quality control in mRNA biogenesis. *Nat Struct Mol Biol* 12(6):482-488.

8. Hodgkin J, Papp A, Pulak R, Ambros V, & Anderson P (1989) A new kind of informational suppression in the nematode Caenorhabditis elegans. *Genetics* 123(2):301-313.

9. Grimson A, O'Connor S, Newman CL, & Anderson P (2004) SMG-1 is a phosphatidylinositol kinase-related protein kinase required for nonsense-mediated mRNA Decay in Caenorhabditis elegans. *Mol Cell Biol* 24(17):7483-7490.

10. Johns L, Grimson A, Kuchma SL, Newman CL, & Anderson P (2007) Caenorhabditis elegans SMG-2 selectively marks mRNAs containing premature translation termination codons. *Mol Cell Biol* 27(16):5630-5638.

11. Nicholson P & Muhlemann O (2010) Cutting the nonsense: the degradation of PTC-containing mRNAs. *Biochem Soc Trans* 38(6):1615-1620.

12. Brumbaugh KM*, et al.* (2004) The mRNA surveillance protein hSMG-1 functions in genotoxic stress response pathways in mammalian cells. *Mol Cell* 14(5):585-598.

13. Gehen SC, Staversky RJ, Bambara RA, Keng PC, & O'Reilly MA (2008) hSMG-1 and ATM sequentially and independently regulate the G1 checkpoint during oxidative stress. *Oncogene* 27(29):4065-4074.

14. Oliveira V, *et al.* (2008) A protective role for the human SMG-1 kinase against tumor necrosis factor-alpha-induced apoptosis. *J Biol Chem* 283(19):13174-13184.

15. Steiner FA, Okihara KL, Hoogstrate SW, Sijen T, & Ketting RF (2009) RDE-1 slicer activity is required only for passenger-strand cleavage during RNAi in Caenorhabditis elegans. *Nat Struct Mol Biol* 16(2):207-211.

16. Hutvagner G & Simard MJ (2008) Argonaute proteins: key players in RNA silencing. *Nat Rev Mol Cell Biol* 9(1):22-32.

17. Pak J & Fire A (2007) Distinct populations of primary and secondary effectors during RNAi in C. elegans. *Science* 315(5809):241-244.

18. Yigit E, *et al.* (2006) Analysis of the C. elegans Argonaute family reveals that distinct Argonautes act sequentially during RNAi. *Cell* 127(4):747-757.

19. Shirayama M, *et al.* (2012) piRNAs initiate an epigenetic memory of nonself RNA in the C. elegans germline. *Cell* 150(1):65-77.

20. Luteijn MJ, *et al.* (2012) Extremely stable Piwi-induced gene silencing in Caenorhabditis elegans. *EMBO J* 31(16):3422-3430.

21. Ashe A, *et al.* (2012) piRNAs can trigger a multigenerational epigenetic memory in the germline of C. elegans. *Cell* 150(1):88-99.

22. Bagijn MP, *et al.* (2012) Function, targets, and evolution of Caenorhabditis elegans piRNAs. *Science* 337(6094):574-578.

23. Lee RC, Hammell CM, & Ambros V (2006) Interacting endogenous and exogenous RNAi pathways in Caenorhabditis elegans. *RNA* 12(4):589-597.

24. Das PP, *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(1):79-90.

25. Culbertson MR, Underbrink KM, & Fink GR (1980) Frameshift suppression Saccharomyces cerevisiae. II. Genetic properties of group II suppressors. *Genetics* 95(4):833-853.

26. Chang YF, Imam JS, & Wilkinson MF (2007) The nonsense-mediated decay RNA surveillance pathway. *Annu Rev Biochem* 76:51-74.

27. Cali BM & Anderson P (1998) mRNA surveillance mitigates genetic dominance in Caenorhabditis elegans. *Mol Gen Genet* 260(2-3):176-184.

28. Hall GW & Thein S (1994) Nonsense codon mutations in the terminal exon of the beta-globin gene are not associated with a reduction in beta-mRNA accumulation: a mechanism for the phenotype of dominant beta-thalassemia. *Blood* 83(8):2031-2037.

29. Azzalin CM & Lingner J (2006) The double life of UPF1 in RNA and DNA stability pathways. *Cell Cycle* 5(14):1496-1498.

30. Kaygun H & Marzluff WF (2005) Regulated degradation of replication-dependent histone mRNAs requires both ATR and Upf1. *Nat Struct Mol Biol* 12(9):794-800.

31. Conti E & Izaurralde E (2005) Nonsense-mediated mRNA decay: molecular insights and mechanistic variations across species. *Curr Opin Cell Biol* 17(3):316-325.

32. MacLeod AR, Waterston RH, & Brenner S (1977) An internal deletion mutant of a myosin heavy chain in Caenorhabditis elegans. *Proc Natl Acad Sci U S A* 74(12):5336-5340.

33. Pulak RA & Anderson P (1988) Structures of spontaneous deletions in Caenorhabditis elegans. *Mol Cell Biol* 8(9):3748-3754.

34. Karn J, Brenner S, & Barnett L (1983) Protein structural domains in the Caenorhabditis elegans unc-54 myosin heavy chain gene are not separated by introns. *Proc Natl Acad Sci U S A* 80(14):4253-4257.

35. Rual JF*, et al.* (2004) Toward improving Caenorhabditis elegans phenome mapping with an ORFeome-based RNAi library. *Genome Res* 14(10B):2162-2168.

36. Alfonso A, Grundahl K, McManus JR, Asbury JM, & Rand JB (1994) Alternative splicing leads to two cholinergic proteins in Caenorhabditis elegans. *J Mol Biol* 241(4):627-630.

37. Alfonso A, Grundahl K, McManus JR, & Rand JB (1994) Cloning and characterization of the choline acetyltransferase structural gene (cha-1) from C. elegans. *J Neurosci* 14(4):2290-2300.

38. Rand JB & Russell RL (1984) Choline acetyltransferase-deficient mutants of the nematode Caenorhabditis elegans. *Genetics* 106(2):227-248.

39. Jin Y, Hoskins R, & Horvitz HR (1994) Control of type-D GABAergic neuron differentiation by C. elegans UNC-30 homeodomain protein. *Nature* 372(6508):780-783.

40. He F*, et al.* (2003) Genome-wide analysis of mRNAs regulated by the nonsense-mediated and 5' to 3' mRNA decay pathways in yeast. *Mol Cell* 12(6):1439-1452.

41. Welch EM & Jacobson A (1999) An internal open reading frame triggers nonsense-mediated decay of the yeast SPT10 mRNA. *EMBO J* 18(21):6134-6145.

42. Lelivelt MJ & Culbertson MR (1999) Yeast Upf proteins required for RNA surveillance affect global expression of the yeast transcriptome. *Mol Cell Biol* 19(10):6710-6719.

43. Losson R & Lacroute F (1979) Interference of nonsense mutations with eukaryotic messenger RNA stability. *Proc Natl Acad Sci U S A* 76(10):5134-5137.

44. Peltz SW, Brown AH, & Jacobson A (1993) mRNA destabilization triggered by premature translational termination depends on at least three cis-acting sequence elements and one trans-acting factor. *Genes Dev* 7(9):1737-1754.

45. Muhlrad D & Parker R (1999) Aberrant mRNAs with extended 3' UTRs are substrates for rapid degradation by mRNA surveillance. *RNA* 5(10):1299-1307.

46. Oliveira CC & McCarthy JE (1995) The relationship between eukaryotic translation and mRNA stability. A short upstream open reading frame strongly inhibits translational initiation and greatly accelerates mRNA degradation in the yeast Saccharomyces cerevisiae. *J Biol Chem* 270(15):8936-8943.

47. Ruiz-Echevarria MJ & Peltz SW (2000) The RNA binding protein Pub1 modulates the stability of transcripts containing upstream open reading frames. *Cell* 101(7):741-751.

48. He F, Peltz SW, Donahue JL, Rosbash M, & Jacobson A (1993) Stabilization and ribosome association of unspliced pre-mRNAs in a yeast upf1- mutant. *Proc Natl Acad Sci U S A* 90(15):7034-7038.

49. Lewis BP, Green RE, & Brenner SE (2003) Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans. *Proc Natl Acad Sci U S A* 100(1):189-192.

50. Green RE*, et al.* (2003) Widespread predicted nonsense-mediated mRNA decay of alternatively-spliced transcripts of human normal and disease genes. *Bioinformatics* 19 Suppl 1:i118-121.

51. Carter MS*, et al.* (1995) A regulatory mechanism that detects premature nonsense codons in T-cell receptor transcripts in vivo is reversed by protein synthesis inhibitors in vitro. *J Biol Chem* 270(48):28995-29003.

52. Kozak M (1978) How do eucaryotic ribosomes select initiation regions in messenger RNA? *Cell* 15(4):1109-1123.

53. Kozak M (1987) An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Res* 15(20):8125-8148.

54. Mignone F, Gissi C, Liuni S, & Pesole G (2002) Untranslated regions of mRNAs. *Genome Biol* 3(3):REVIEWS0004.

55. Iacono M, Mignone F, & Pesole G (2005) uAUG and uORFs in human and rodent 5'untranslated mRNAs. *Gene* 349:97-105.

56. Matsui M, Yachie N, Okada Y, Saito R, & Tomita M (2007) Bioinformatic analysis of post-transcriptional regulation by uORF in human and mouse. *FEBS Lett* 581(22):4184-4188.

57. Hinnebusch AG (2005) Translational regulation of GCN4 and the general amino acid control of yeast. *Annu Rev Microbiol* 59:407-450.

58. Mueller PP & Hinnebusch AG (1986) Multiple upstream AUG codons mediate translational control of GCN4. *Cell* 45(2):201-207.

59. Yepiskoposyan H, Aeschimann F, Nilsson D, Okoniewski M, & Muhlemann O (2011) Autoregulation of the nonsense-mediated mRNA decay pathway in human cells. *RNA* 17(12):2108-2118.

60. Wethmar K, Smink JJ, & Leutz A (2010) Upstream open reading frames: molecular switches in (patho)physiology. *Bioessays* 32(10):885-893.

61. Kerr TP, Sewry CA, Robb SA, & Roberts RG (2001) Long mutant dystrophins and variable phenotypes: evasion of nonsense-mediated decay? *Hum Genet* 109(4):402-407.

62. Kerem E*, et al.* (2008) Effectiveness of PTC124 treatment of cystic fibrosis caused by nonsense mutations: a prospective phase II trial. *Lancet* 372(9640):719-727.

63. Tilman G*, et al.* (2012) Cancer-linked satellite 2 DNA hypomethylation does not regulate Sat2 non-coding RNA expression and is initiated by heat shock pathway activation. *Epigenetics* 7(8):903-913.

64. Carvalho RH*, et al.* (2012) Genome-wide DNA methylation profiling of non-small cell lung carcinomas. *Epigenetics Chromatin* 5(1):9.

65. Zhu Q*, et al.* (2011) BRCA1 tumour suppression occurs via heterochromatin-mediated silencing. *Nature* 477(7363):179-184.

66. Reiner DJ, Newton EM, Tian H, & Thomas JH (1999) Diverse behavioural defects caused by mutations in Caenorhabditis elegans unc-43 CaM kinase II. *Nature* 402(6758):199-203.

67. Blevins T, *et al.* (2006) Four plant Dicers mediate viral small RNA biogenesis and DNA virus induced silencing. *Nucleic Acids Res* 34(21):6233-6246.

68. Brodersen P & Voinnet O (2006) The diversity of RNA silencing pathways in plants. *Trends Genet* 22(5):268-280.

69. Cogoni C & Macino G (1997) Isolation of quelling-defective (qde) mutants impaired in posttranscriptional transgene-induced gene silencing in Neurospora crassa. *Proc Natl Acad Sci U S A* 94(19):10233-10238.

70. Hamilton AJ & Baulcombe DC (1999) A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science* 286(5441):950-952.

71. Bernstein E, Denli AM, & Hannon GJ (2001) The rest is silence. *RNA* 7(11):1509-1521.

72. Hutvagner G, *et al.* (2001) A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. *Science* 293(5531):834-838.

73. Gent JI, *et al.* (2010) Distinct phases of siRNA synthesis in an endogenous RNAi pathway in C. elegans soma. *Mol Cell* 37(5):679-689.

74. Gu C, *et al.* (2009) Structure-function analysis of mutant RNA-dependent RNA polymerase complexes with VPg. *Biochemistry (Mosc)* 74(10):1132-1141.

75. Sijen T, *et al.* (2001) On the role of RNA amplification in dsRNA-triggered gene silencing. *Cell* 107(4):465-476.

76. Sijen T, Steiner FA, Thijssen KL, & Plasterk RH (2007) Secondary siRNAs result from unprimed RNA synthesis and form a distinct class. *Science* 315(5809):244-247.

77. Batista PJ, *et al.* (2008) PRG-1 and 21U-RNAs interact to form the piRNA complex required for fertility in C. elegans. *Mol Cell* 31(1):67-78.

78. Simon M, *et al.* (2014) Reduced Insulin/IGF-1 Signaling Restores Germ Cell Immortality to Caenorhabditis elegans Piwi Mutants. *Cell Rep*.

79. Adamo A, *et al.* (2012) Transgene-mediated cosuppression and RNA interference enhance germ-line apoptosis in Caenorhabditis elegans. *Proc Natl Acad Sci U S A* 109(9):3440-3445.

80. Dernburg AF, Zalevsky J, Colaiacovo MP, & Villeneuve AM (2000) Transgene-mediated cosuppression in the C. elegans germ line. *Genes Dev* 14(13):1578-1583.

81. Ketting RF & Plasterk RH (2000) A genetic link between co-suppression and RNA interference in C. elegans. *Nature* 404(6775):296-298.

82. Dalmay T, Hamilton A, Rudd S, Angell S, & Baulcombe DC (2000) An RNA-dependent RNA polymerase gene in Arabidopsis is required for posttranscriptional gene silencing mediated by a transgene but not by a virus. *Cell* 101(5):543-553.

83. Baulcombe D (2004) RNA silencing in plants. *Nature* 431(7006):356-363.

84. Jorgensen RA (1995) Cosuppression, flower color patterns, and metastable gene expression States. *Science* 268(5211):686-691.

85. Akitake CM, Macurak M, Halpern ME, & Goll MG (2011) Transgenerational analysis of transcriptional silencing in zebrafish. *Dev Biol* 352(2):191-201.

86. Goll MG, Anderson R, Stainier DY, Spradling AC, & Halpern ME (2009) Transcriptional silencing and reactivation in transgenic zebrafish. *Genetics* 182(3):747-755.

87. Grosveld F, van Assendelft GB, Greaves DR, & Kollias G (1987) Position-independent, high-level expression of the human beta-globin gene in transgenic mice. *Cell* 51(6):975-985.

88. Sharpe JA*, et al.* (1993) Analysis of the human alpha-globin gene cluster in transgenic mice. *Proc Natl Acad Sci U S A* 90(23):11262-11266.

89. Frokjaer-Jensen C, Davis MW, Ailion M, & Jorgensen EM (2012) Improved Mos1-mediated transgenesis in C. elegans. *Nat Methods* 9(2):117-118.

90. Frokjaer-Jensen C*, et al.* (2008) Single-copy insertion of transgenes in Caenorhabditis elegans. *Nat Genet* 40(11):1375-1383.

91. Shtessel L*, et al.* (2013) Caenorhabditis elegans POT-1 and POT-2 repress telomere maintenance pathways. *G3 (Bethesda)* 3(2):305-313.

92. Dickinson DJ, Ward JD, Reiner DJ, & Goldstein B (2013) Engineering the Caenorhabditis elegans genome using Cas9-triggered homologous recombination. *Nat Methods* 10(10):1028-1034.

93. Grishok A, Tabara H, & Mello CC (2000) Genetic requirements for inheritance of RNAi in C. elegans. *Science* 287(5462):2494-2497.

94. Sijen T & Plasterk RH (2003) Transposon silencing in the Caenorhabditis elegans germ line by natural RNAi. *Nature* 426(6964):310-314.

95. Tops BB*, et al.* (2005) RDE-2 interacts with MUT-7 to mediate RNA interference in Caenorhabditis elegans. *Nucleic Acids Res* 33(1):347-355.

96. Seth M, *et al.* (2013) The C. elegans CSR-1 argonaute pathway counteracts epigenetic silencing to promote germline gene expression. *Dev Cell* 27(6):656-663.

97. Han W, Sundaram P, Kenjale H, Grantham J, & Timmons L (2008) The Caenorhabditis elegans rsd-2 and rsd-6 genes are required for chromosome functions during exposure to unfavorable environments. *Genetics* 178(4):1875-1893.

98. Zhang C, *et al.* (2012) The Caenorhabditis elegans RDE-10/RDE-11 complex regulates RNAi by promoting secondary siRNA amplification. *Curr Biol* 22(10):881-890.

99. Zhuang JJ, Banse SA, & Hunter CP (2013) The nuclear argonaute NRDE-3 contributes to transitive RNAi in Caenorhabditis elegans. *Genetics* 194(1):117-131.

100. Phillips CM & Dernburg AF (2006) A family of zinc-finger proteins is required for chromosome-specific pairing and synapsis during meiosis in C. elegans. *Dev Cell* 11(6):817-829.

101. Hawley RS & Gilliland WD (2009) Homologue pairing: getting it right. *Nat Cell Biol* 11(8):917-918.

102. Phillips CM, *et al.* (2009) Identification of chromosome sequence motifs that mediate meiotic pairing and synapsis in C. elegans. *Nat Cell Biol* 11(8):934-942.

103. Maine EM, *et al.* (2005) EGO-1, a putative RNA-dependent RNA polymerase, is required for heterochromatin assembly on unpaired dna during C. elegans meiosis. *Curr Biol* 15(21):1972-1978.

104. Aramayo R & Metzenberg RL (1996) Meiotic transvection in fungi. *Cell* 86(1):103-113.

105. Shiu PK & Metzenberg RL (2002) Meiotic silencing by unpaired DNA: properties, regulation and suppression. *Genetics* 161(4):1483-1495.

106. Shiu PK, Raju NB, Zickler D, & Metzenberg RL (2001) Meiotic silencing by unpaired DNA. *Cell* 107(7):905-916.

107. Hammond TM, *et al.* (2013) Identification of small RNAs associated with meiotic silencing by unpaired DNA. *Genetics* 194(1):279-284.

108. Hammond TM, *et al.* (2011) SAD-3, a Putative Helicase Required for Meiotic Silencing by Unpaired DNA, Interacts with Other Components of the Silencing Machinery. *G3 (Bethesda)* 1(5):369-376.

109. Hynes MJ & Todd RB (2003) Detection of unpaired DNA at meiosis results in RNA-mediated silencing. *Bioessays* 25(2):99-103.

110. Avgousti DC, Palani S, Sherman Y, & Grishok A (2012) CSR-1 RNAi pathway positively regulates histone expression in C. elegans. *EMBO J* 31(19):3821-3832.

111. Conine CC*, et al.* (2013) Argonautes promote male fertility and provide a paternal memory of germline gene expression in C. elegans. *Cell* 155(7):1532-1544.

112. Wedeles CJ, Wu MZ, & Claycomb JM (2013) Protection of germline gene expression by the C. elegans Argonaute CSR-1. *Dev Cell* 27(6):664-671.

113. Waterhouse PM, Graham MW, & Wang MB (1998) Virus resistance and gene silencing in plants can be induced by simultaneous expression of sense and antisense RNA. *Proc Natl Acad Sci U S A* 95(23):13959-13964.

114. Cui M & Han M (2007) Roles of chromatin factors in C. elegans development. *WormBook*:1-16.

115. Gonzalez-Aguilera C, Palladino F, & Askjaer P (2013) C. elegans epigenetic regulation in development and aging. *Brief Funct Genomics*.

116. Wenzel D, Palladino F, & Jedrusik-Bode M (2011) Epigenetics in C. elegans: facts and challenges. *Genesis* 49(8):647-661.

117. Bond DM & Baulcombe DC (2014) Small RNAs and heritable epigenetic variation in plants. *Trends Cell Biol* 24(2):100-107.

118. Haag JR*, et al.* (2012) In vitro transcription activities of Pol IV, Pol V, and RDR2 reveal coupling of Pol IV and RDR2 for dsRNA synthesis in plant RNA silencing. *Mol Cell* 48(5):811-818.

119. Havecker ER*, et al.* (2010) The Arabidopsis RNA-directed DNA methylation argonautes functionally diverge based on their expression and interaction with target loci. *Plant Cell* 22(2):321-334.

120. Law JA, Vashisht AA, Wohlschlegel JA, & Jacobsen SE (2011) SHH1, a homeodomain protein required for DNA methylation, as well as RDR2, RDM4, and chromatin remodeling factors, associate with RNA polymerase IV. *PLoS Genet* 7(7):e1002195.

121. Xie Z*, et al.* (2004) Genetic and functional diversification of small RNA pathways in plants. *PLoS Biol* 2(5):E104.

122. Chan SW*, et al.* (2004) RNA silencing genes control de novo DNA methylation. *Science* 303(5662):1336.

123. Law JA & Jacobsen SE (2010) Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat Rev Genet* 11(3):204-220.

124. Jablonka E & Lamb MJ (2006) The evolution of information in the major transitions. *J Theor Biol* 239(2):236-246.

125. Park S & Lehner B (2013) Epigenetic epistatic interactions constrain the evolution of gene expression. *Mol Syst Biol* 9:645.

126. Weigel D & Colot V (2012) Epialleles in plant evolution. *Genome Biol* 13(10):249.

127. Schmitz RJ*, et al.* (2011) Transgenerational epigenetic instability is a source of novel methylation variants. *Science* 334(6054):369-373.

128. Bailey JA*, et al.* (2002) Recent segmental duplications in the human genome. *Science* 297(5583):1003-1007.

129. Zhang F, Gu W, Hurles ME, & Lupski JR (2009) Copy number variation in human health, disease, and evolution. *Annu Rev Genomics Hum Genet* 10:451-481.