UNIVERSITY OF CALIFORNIA
RIVERSIDE

Identification and Characterization of Novel Alleles Required for Antiviral RNA Interference in *Caenorhabditis elegans*

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Genetics, Genomics and Bioinformatics

by

Stephanie Renee Coffman

December 2015

Dissertation Committee:
Dr. Shou-wei Ding, Chairperson
Dr. Morris Maduro
Dr. ALN Rao
Dr. Weifeng Gu

The Dissertation of Stephanie Renee Coffman is approved:

Committee Chairperson

University of California, Riverside

ACKNOWLEDGEMENTS

First and foremost I would like to thank my advisor Dr. Shou-wei Ding for inviting me into his lab and providing guidance throughout the years. Several times, my project has hit roadblocks and Dr. Ding was always optimistic and encouraging. Dr. Ding has taught me how to be an academic scientist and I am forever grateful for his mentorship. I would also like to thank Dr. Morris Maduro for welcoming me into his lab during the first half of my studies. Dr. Maduro has always made himself available with advice and has provided solutions to problematic experiments countless times. I would also like to thank Dr. ALN Rao and Dr. Weifeng Gu for serving on my dissertation committee and for their time and guidance throughout the years.

Over the past five years, I have had several mentors and colleagues. I am particularly grateful to our *C. elegans* team within the Ding Lab and to Gina Briotman-Maduro in Dr. Maduro's lab. Gina taught me how to work with *C.* elegans, without which my work would not have been possible. I am especially grateful to Dr. Zhihuan Gao for his mentorship during my early years as a PhD student. Dr. Gao taught me countless molecular biology techniques and from my first day involved me in scientific discussions. I worked most closely with another graduate student, Yuanyuan Guo, and I would like to thank her for countless hours of discussion and her support on my project. Additionally, I would like to thank Dr. Xunyang Guo and Dr. Jing Zhong for discussions and advice on *C. elegans*. Dr. Yanhong Han has been a constant source of knowledge about small RNA techniques, such as small RNA sequencing and small RNA Northern Blots; I am thankful to her for passing her knowledge on to me and her patience in

teaching me these laborious techniques. Additionally, I am grateful to both Dr. Juan Jovel and Jinfeng Lu, who helped me develop basic programming skills in Perl, without which Chapter 3 would have been impossible. I would also like to thank my bench-mate, Dr. Zhongxin Guo for discussions, especially about forward genetic screens and mapping.

Lastly, I would like to thank my husband, Alex, and my son, Nick, who have made this journey with me. Without their support none of this would be possible.

This project was funded by NIH grant GM94396 and the Video Bioinformatics IGERT Program at UCR DGE 0903667.

ABSTRACT OF THE DISSERTATION

Identification and Characterization of Novel Alleles Required for Antiviral RNA Interference in *Caenorhabditis elegans*

by

Stephanie Renee Coffman

Doctor of Philosophy, Graduate Program in Genetics, Genomics and Bioinformatics
University of California, Riverside, December 2015
Dr. Shou-wei Ding, Chairperson

Antiviral RNAi is a conserved antiviral pathway in plants and invertebrates that mediates clearance of viral RNA. Antiviral RNAi has been extensively studied in plants and insects, but *Caenorhabditis elegans* has only recently emerged as a model to study antiviral immunity. *C. elegans* is an ideal model to study antiviral RNAi, because it shares key commonalities with mammals, including a single Dicer and an expanded class of duplex RNA activated ATPases (DRAs). While several genetic screens have been done to identify the genes required for RNAi induced by exogenous long dsRNA, this unbiased forward genetic screen is the first with the ability to reveal genes required uniquely for antiviral RNAi or novel antiviral pathways. This dissertation develops an EMS screen based on a viral replicon that is defective in RNAi suppression. Viral replication is rescued in animals defective in RNAi and visualized by GFP. Thirteen worm mutants were identified that are defective in antiviral RNAi and four novel alleles of *drh-1*, a DRA homologous to the mammalian gene RIG-I, were revealed.

Both RIG-I and DRH-1 function in antiviral immunity through the activation of

type-I interferon signaling in mammals and antiviral RNAi in *C. elegans*, respectively. This work investigates the function of DRH-1 in the production of virus-derived siRNAs (vsiRNAs). Deep sequencing of wild-type and *drh-1* mutants challenged with Orsay Virus demonstrated that *drh-1* mutants produced primary vsiRNAs that are dependent on DCR-1 and the dsRNA-binding protein RDE-4. However, these DRH-1 independent vsiRNAs are predominantly derived from the 5' terminal regions of both OrV RNA1 and RNA2 and were depleted from the internal regions of viral RNA. In contrast, wild-type and other mutant animals did not exhibit a bias for primary vsiRNA biogenesis from the 5' terminal regions. Mammalian RIG-I acts as an ATP-powered translocase, though the biological function of this activity is unknown. Therefore, DRH-1 may function to promote the production of primary vsiRNAs spreading from the 5' terminal regions to the internal and 3' regions of the viral RNA using translocase activity.

TABLE OF CONTENTS

Chapter 2

Forward Genetic Screen to Identify the Genetic Requirements of Antiviral RNAi

List of Figures

List of Tables

CHAPTER 1

INTRODUCTION

1.1 RNAinterference (RNAi)

RNAinterference (RNAi) broadly defines a diverse group of small RNA mediated pathways that target RNA for silencing in a sequence specific manner. In retrospect, it is clear that RNAi was the driving force behind co-suppression, a phenomenon first observed in ornamental petunias whereby the introduction of a transgene would suppress expression of both the transgene and the endogenous gene of homologous sequence (1, 2). Similarly, incorporation of viral sequences into a plant genome could protect against future infections with the same virus and mediate sequence-specific degradation of viral RNA (3). In 1998, a Nobel-prize winning set of experiments in *Caenorhabditis elegans* embryos conclusively demonstrated that gene silencing mediated by RNAi was triggered by double-stranded RNA (dsRNA) (4) and a few years later, small RNAs were identified in *Arabidopsis thaliana* as the effector of RNAi (5). Subsequently, RNAi was demonstrated to mediate gene silencing in most eukaryotes and to mediate potent antiviral immunity in plants, insects, nematodes and some fungi (6–9). Over the last two decades, our understanding of RNAi has grown at an exponential rate and the use of RNAi as a laboratory tool to knockdown genes has revolutionized molecular biology.

The potential applications of RNAi seem infinite and include the production of disease resistant crops, gene therapy and alternative vaccination or antiviral strategies.

RNAi is conserved in plants, animals and fungi and has evolved a diverse array of RNA silencing functions. Despite this diversity, the RNAi pathways are conserved in mechanism and maintain many of the same key components, most notably Dicer and Argonaute proteins. The endoribonuclease III, Dicer, is conserved throughout eukaryotes, though it has been lost in some single-celled protozoans and some fungi (10, 11). Dicer functions as a molecular ruler and can processively "dice" dsRNA into small RNAs to initiate RNAi (12). Argonaute (AGO) proteins bind to small RNAs and targets RNA for silencing through base pairing of the guide small RNA with a target RNA molecule. Originally thought of only as "slicers", some AGO proteins cleave the target RNA; however, many AGOs mediate gene silencing by recruiting factors for translational repression or chromatin remodeling. While animals have one or two dicer proteins, they generally have a more diverse repertoire of AGO proteins that can function in unique small RNAi pathways. Plants have more redundancy for both proteins, with four dicer-like proteins and 10 AGOs in *Arabidopsis thaliana*. *C. elegans* has the largest number of Argonaute family members with 27 annotated Argonaute genes divided into three subfamilies: (1) Argonautes (AGO), found in all three kingdoms, (2) PIWI, found in metazoans, and (3) an expanded class of worm-specific Argonaute (WAGO). In plants, *C. elegans* and fungi, RNA-dependent RNA Polymerases (RdRP) play an essential role in RNAi and function to amplify the small RNA signal through the production of secondary small RNAs.

1.2 *C. elegans* RNAi

In eukaryotes there are two main RNAi pathways: microRNAs (miRNAs) and small interfering RNAs (siRNAs), which mediate Post-Transcriptional Gene Silencing (PTGS) through mRNA degradation or translational inhibition. siRNAs can also mediate Transcriptional Gene Silencing (TGS) through chromatin remodeling. In animals, there is a third class of small RNAs, piwi-interacting RNAs (piRNAs), which are largely responsible for silencing transposons in the germline, though piRNAs can also mediate silencing of endogenous genes in *C. elegans* (13, 14).

Given the multitude of argonaute proteins in *C. elegans,* it comes as no surprise that *C. elegans* hosts several RNAi pathways. With the exception of microRNAs, RNAi is a two-step process in *C. elegans* that involves initiation through the production of primary siRNAs (or piRNAs) followed by amplification of the silencing signal through the production of secondary siRNAs (commonly referred to as 22G-RNAs). While a diverse array of triggers (e.g. endogenous or exogenous dsRNA, viral RNA, piRNA transcripts) can initiate RNAi via unique biogenesis mechanisms they ultimately feed into a more conserved secondary siRNA pathway.

1.2.1 Exogenous RNAi

Exogenous RNAi (exo-RNAi), also known as experimental RNAi is triggered by synthetic dsRNA that can be administered to worms by feeding, injecting or soaking. The dsRNA is processed by the single *C. elegans* dicer, DCR-1, in complex with RDE-4, RDE-1 and DRH-1 (15, 16), into primary siRNAs. RDE-4 is a dsRNA binding protein

that preferentially binds long dsRNA over short dsRNA, such as miRNAs hairpins, effectively guiding DCR-1 into siRNA pathways (17). Though present in the DCR-1 complex, the function of DRH-1 is only essential during viral infection (18–20). The DCR-1 complex produces primary siRNAs one after the other along the dsRNA, in a process known as phasing. Primary siRNAs derived from long dsRNA are 23-nt duplexed RNA, with 19 complementary base pairs and 2-nt overhangs at the 3'-ends. Primary siRNAs are loaded into the argonaute, RDE-1, with the help of RDE-4 (21).

RDE-1 preferentially binds to small RNAs derived from dsRNA with perfect complementarity. In contrast, miRNA argonautes bind to dsRNA hairpins with bulges caused by imperfect complementarity (21). Interestingly, in the absence of foreign dsRNA, RDE-1 scavenges endogenous small RNAs, such as miRNAs that have few mismatches and mediates silencing of perfectly complementary target mRNA via the exo-RNAi pathway (22). The RNase H, or slicer activity, of RDE-1 is required for removal of the passenger strand from the siRNA duplex, exposing the guide strand for base pairing (23). Once the RDE-1-siRNA complex base pairs with a target mRNA, it recruits, either directly or indirectly, at least three complexes that function in the production of 22G-RNAs.

At the heart of 22G-RNA production is an RdRP complex. The RdRP complex is required for the synthesis of secondary siRNAs and contains the RdRP, RRF-1, in the soma and redundant RdRPs RRF-1 and EGO-1 in the germline, along with DRH-3 and EKL-1 (24, 25). Secondary siRNAs are synthesized antisense to the target RNA and consistent with RdRP synthesis, they have a guanine as a 5' nucleotide and a 5'-

triphosphate (26). These features, along with a length preference of 22-nt, gave rise to the designation 22G-RNAs. The primer-independent synthesis of 22G-RNAs is transitive and displays a strong bias towards spreading in an upstream direction (21, 27, 28). In fact, triggers designed to match the 5' end of an RNA molecule are unable to spread downstream to effectively silence an adjacent gene (21). In some cases transitive RNAi in a 5' to 3' direction have been observed, though to a lesser degree (28, 29). 22G-RNAs are not produced in *drh-3* mutants, though the exact function of DRH-3 remains unknown. The distribution of endogenous siRNAs in *drh-3* mutants suggest that DRH-3 functions in the spread of 22G-RNAs in a 3' to 5' direction (25). Additionally, DRH-3 forms dimers on dsRNA and preferentially binds 22-nt duplex RNAs *in vitro* (30). Taken together these experiments suggest that DRH-3 specifically binds to 22G-RNAs and promotes production of upstream 22G-RNAs.

The RDE-10/RDE-11 complex is a nematode specific complex that functions downstream of RDE-1 but upstream of RRF-1. The RDE-10/RDE-11 complex promotes 22G-RNA production and target mRNA degradation (31, 32). RDE-10 directly interacts with target mRNA through an RDE-1 dependent mechanism and the complex shows a preference for partially degraded mRNA targets (31). Both *rde-10* and *rde-11* mutants are RNAi defective at lower concentrations of dsRNA, but proficient at higher levels (32). *rsd-2* and *rsd-6* display similar dosage-dependent and small RNA phenotypes and physically interact with the RDE-10/RDE-11 complex (32). Taken together these results suggest that the RDE-10/RDE-11 complex binds to target mRNA downstream of RDE-1-siRNA target binding in order to promote the production of 22G-RNAs.

In 2015, the third and newest complex was discovered. The RDE-8 complex contains the riboendonuclease, RDE-8, and the nucleotidyl transferase, RDE-3, and is required for the efficient synthesis of 22G-RNAs by RdRPs, and for the uridylation of the 5'-fragment of target cleavage products (33). The RDE-8 complex is recruited to the target mRNA by RDE-1; however, the order of its recruitment in comparison to the RdRP and RDE-10/RDE-11 complexes is yet to be determined. This complex is likely responsible for cleavage of the mRNA target, either through RDE-8 or another subunit. The discovery of the RDE-8 complex is the beginning of an answer to a long-standing question in the field; namely how are target mRNA molecules cleaved if exo-RNAi argonautes either do not utilize their cleavage activity (e.g. RDE-1) or do not encode cleavage activity (e.g. WAGOs)? Though the discovery of a riboendonuclease is certainly hopeful, we await experimental confirmation that RDE-8 cleaves the target mRNA.

After production, 22G-RNAs are loaded into several different WAGOs. Cytoplasmic WAGOs mediate PTGS of the target mRNA. However, WAGOs that load 22G-RNAs do not have endonuclease activity and must recruit other factors for mRNA destruction, possibly the RDE-8 complex (33). Cytoplasmic WAGOs function redundantly and strains containing up to 12 mutated WAGO genes, called MAGO-12, have been built to better understand their role. Another WAGO, NRDE-3, shuttles 22G-RNAs into the nucleus to initiate TGS through chromatin remodeling and can promote to inheritance of gene silencing (34). While gene silencing is readily inherited in *C. elegans*,

silencing of germline genes can by inherited indefinitely (13, 14, 35, 36), whereas gene silencing of somatic genes wears off in a couple generations.

1.2.2 Endogenous RNAi

C. elegans supports all three of the small RNA pathways known to function in eukaryotes: miRNAs, siRNAs and piRNAs.

1.2.2.1 miRNA Pathway

The first miRNA gene, *lin-4*, was discovered in 1993 using forward genetics and was thought to be a novelty of worm biology (37). Ambros and colleagues discovered that the *lin-4* gene was transcribed from an intron, but was short and unlikely to encode a protein; furthermore, the sequence of *lin-4* was antisense to *lin-14* and its sole function was to negatively regulate *lin-14* (37). In 2000, after the formal discovery of RNAi, a conserved miRNA gene, *let-7*, was discovered in worms (38) and miRNAs were found to be an almost universal mechanism for fine-tuning gene regulation.

Similar to protein coding genes, miRNA genes undergo complex spatial and temporal regulation by specific transcription factors. Primary miRNA transcripts (pri-mRNA) are transcribed by RNA POL II and are capped and polyadenylated in the nucleus. pri-miRNAs are processed into precursor miRNAs (pre-miRNAs) by the RNAase III enzyme, DRSH-1 and a dsRNA binding protein, PASH-1; pre-miRNAs are about 70-nt hairpins and are exported from the nucleus. In the cytoplasm, the single *C. elegans* DCR-1 cleaves the pre-miRNA into a duxplex RNA that is 22-nt in length with a

3' 2-nt overhang. This duplex is comprised of a mature miRNA and it's complement, known as miRNA*. Both the miRNA and the miRNA* can initiate silencing of mRNA targets, though they have different sequence specificity. In general, loading of miRNA into RISC is favored over the miRNA*.

The mature miRNA is loaded into the argonaute proteins, ALG-1 or ALG-2, which will base pair to a similar sequence in the 3'-UTR of a mRNA molecule. The seed region of the miRNA, from the second through the eight nucleotide, is the most important for determining target binding, though even the seed region does not bind with 100% complementarity. Since miRNA binding sites do not require perfect complementarity, one miRNA can have multiple targets and predicting the targets is still an imperfect, but ever improving, process. The miRNA-RISC complex will ultimately repress translation of or deadenylate the target mRNA, completely or partially silencing the gene. The miRNA pathway is the most conserved RNAi pathway and many miRNA genes are conserved.

1.2.2.2 siRNA Pathways

Several endogenous RNAi (endo-RNAi) pathways that produce endogenous siRNAs (endo-siRNAs) have been described in *C. elegans*. The first experiments to elucidate RNAi in *C. elegans* centered on experimentally induced RNAi, the natural function of which was thought to be antiviral. The existence of an endogenous RNAi pathway was not suspected until the discovery of *C. elegans* mutants that exhibited an

Enhanced RNAi (ERI) phenotype, suggesting competition existed for the core RNAi proteins (39).

Primary endo-siRNAs are more commonly referred to as 26G-RNAs, because they are 26-nt in length with a guanine as the 5'-nucleotide (40). 26G-RNAs are synthesized antisense to a target mRNA or linc RNA by the RdRP, RRF-3; interestingly, these RdRP products have a 5'-monophosphate, suggesting the 5' end is processed after synthesis, perhaps by DCR-1 or the RNA phosphatase, PIR-1 (15). 26G-RNA production also requires DCR-1 and the RNA binding protein, RDE-4, as well as the dicer related helicase, DRH-3. While the role of dicer in the production of 26G-RNAs in poorly understood, loss of DCR-1, or the DCR-1 ATPase domain, results in a loss of 26G-RNAs (41). 26G-RNAs are sorted into two classes that differ biochemically and in the AGO protein that they associate with. 26G-RNAs that associate with the AGO ERGO-1 are initiated during oogenesis (42, 43). Interestingly, the ERGO 26G-RNAs do not target germline transcripts, but instead target recently duplicated genes, pseudogenes, lincRNAs and some unannotated regions of the genome (44). The second group of 26G-RNAs associate with two redundant AGO proteins, ALG-3 and ALG-4, in males and function exclusively in spermatogenesis (45). 26G-RNAs found in hermaphrodites are methylated at the 3' end, while those found in males are not (46, 47). 26G-RNAs guide their respective AGO proteins, and other associated factors, to a target mRNA molecule to amplify the small RNA signal by production of 22G-RNAs, discussed above. Similar to exo-RNA, execution of gene silencing can be mediated by cytoplasmic or nuclear WAGOs to induce PTGS or TGS, respectively. The competition observed between

endogenous and exogenous RNAi is likely over secondary WAGOs, as their overexpression increases susceptibility to RNAi (48).

1.2.2.3 piRNA Pathway

The primary function of piRNAs is to protect the integrity of the genome by distinguishing non-self from self nucleic acids (49). For this reason, piRNAs act exclusively in the germline, protecting the viability of following generations. While piRNA pathways tend to vary substantially in terms of biogenesis and amplification between organisms, none of them are known to utilize dsRNA precursors or to require Dicer. In general, piRNAs are 26-30 nucleotides in length, have a uracil as the 5'-nucleotide and are methylated at the 3' end. Because *C. elegans* piRNAs are unusually short in comparison, 21-nt, they are more commonly referred to as 21U-RNAs.

The *C. elegans* genome encodes more than 16,000 21U-RNAs, which are mostly found in two distinct clusters on chromosome IV with each 21U-RNA located downstream of the Ruby motif (40). Transcription of 21U-RNA precursors by RNA POL II begins precisely 2-nt upstream of the 5' uracil and is regulated by the Forkhead family of transcription factors (50). Downstream of transcription, several factor are required for the biogenesis and/or stability of 21U-RNA precursors, including piRNA defective 1 (PRDE-1) and Twenty-one U Fouled Up 3, 4 and 5 (TOFU-3/4/5) (51, 52). TOFU-1, TOFU-2 and PID-1 appear to be required for the processing of 21U-RNA precursors into mature 21U-RNAs, though their function in this process in unknown (52, 53). Presumably, the precursors are exported from the nucleus before processing into mature

21U-RNAs. The PIWI argonaute, PRG-1, interacts with 21U-RNAs and either functions directly in mature 21U-RNA processing or in their stability (54). Mature 21U-RNAs are methylated at the 3'-end by HENN-1 (55, 47).

In contrast to *Drosophila melanogaster* and mice, piRNAs in *C. elegans* are not amplified via the canonical "ping pong" model; instead, 21U-RNAs are amplified in *C. elegans* through the production of 22G-RNAs. PRG-1 is loaded with a 21U-RNA and guided to a target RNA where it recruits the complexes required for 22G-RNA synthesis. Interestingly, 21U-RNAs do not directly map to transposable elements (56). Instead, they can initiate the production of 22G-RNAs by imperfect base pairing between the 21U-RNA and the target RNA (13, 14). This imperfect pairing allows for the production of 22G-RNAs targeting transposons, but also a number of endogenous genes. The 22G-RNAs are loaded into a germline specific argonaute protein, heritable defective 1 (HRDE-1), which shuttles the 22G-RNA into the nucleus where it initiates TGS by H3K9 methylation. The mechanism of TGS mediated by 21U-RNAs in the germline is likely similar to TGS in the soma, as it requires NRDE-1, NRDE-2 and NRDE-4 (36, 57, 58).

1.2.2.4 CSR-1 22G-RNAs

In contrast to the 22G-RNAs dependent on 26G-RNAs or piRNAs, there is a second less abundant class of 22G-RNAs that specifically interacts with one specific WAGO called CSR-1. CSR-1 22G-RNAs are antisense to protein coding genes and not transposons or other genomic loci targeted by the other WAGO class 22G-RNAs. CSR-1 22G-RNAs are synthesized solely by EGO-1, with DRH-3 and EKL-1, and require the

nucleoltidyltransferase, CDE-1, which uridylates the 22G-RNAs and interacts directly with EGO-1 (25, 59, 60). Despite their reliance on nearly identical machinery for their biogenesis, CSR-1 22G-RNAs may not be bona fide secondary siRNAs, because corresponding primary siRNAs are not known. Interestingly, in *csr-1* mutants, gene expression of CSR-1 22G-RNA targets are not up-regulated and may even be down-regulated, suggesting the CSR-1 22G-RNAs may actually promote gene expression (60, 61). CSR-1 binds euchromatin corresponding to CSR-1 22G-RNA targets, in a 22G-RNA and target mRNA dependent manner and may play a direct role in chromatin remodeling (62). A recent GRO-seq study demonstrated that CSR-1 directly interacts with RNA POL II machinery to promote sense transcripts from CSR-1 22G-RNA targets (63). Interestingly, in *csr-1* mutants transcription occurs in a sense and antisense direction (63). Taken together, these results suggest that CSR-1 22G-RNAs actually promote transcription, perhaps as a mechanism to defend against 21U-RNAs. In contrast to other organisms, piRNAs in *C. elegans* can target self-RNA for silencing and CSR-1 22G-RNAs may be a counter-mechanism to protect self-RNA from the piRNA pathway.

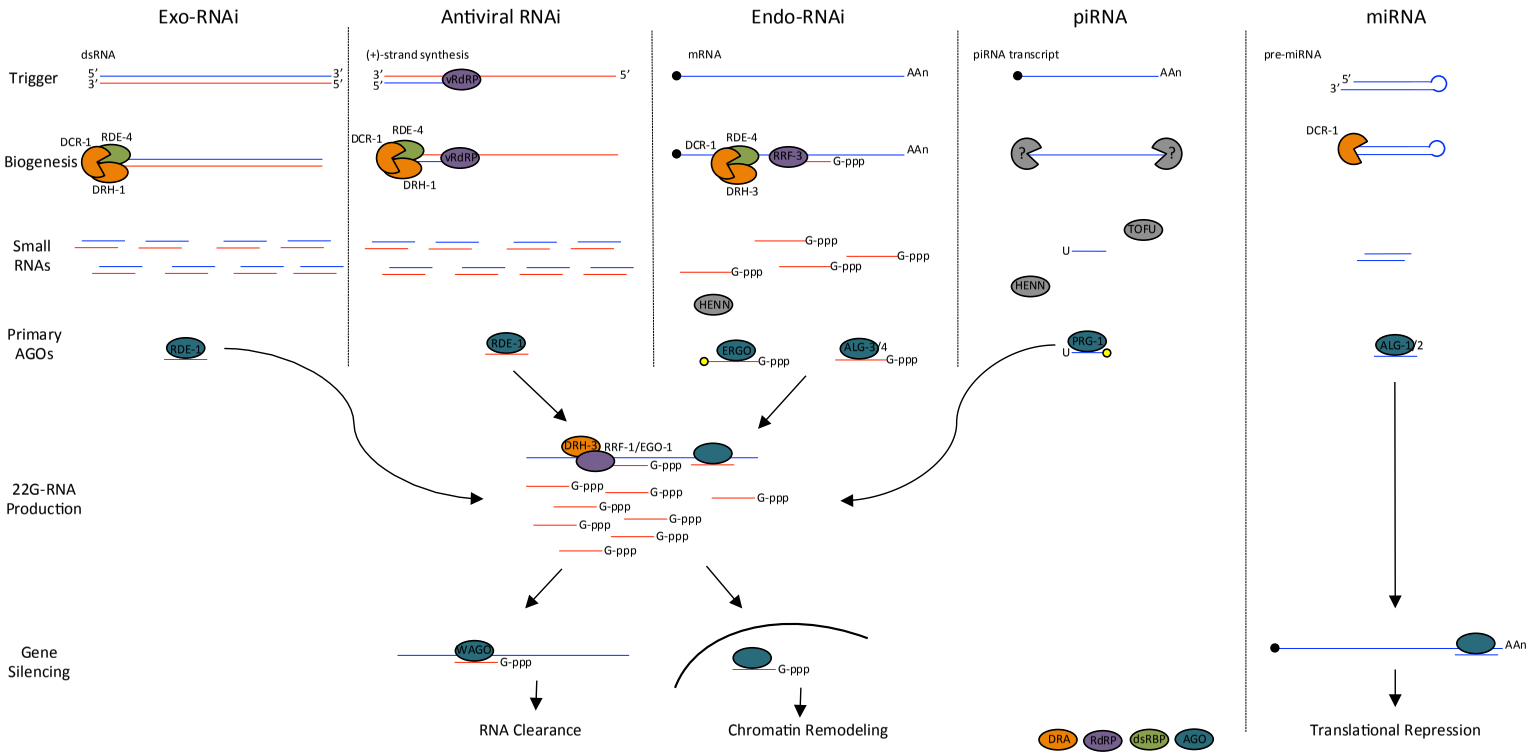

Figure 1-1. Small RNA Pathways in *C. elegans*. Small RNA pathways in *C. elegans* proceed in two steps. In the first step, primary siRNAs are processed by DCR-1 from different dsRNA precursors. piRNAs are DCR-1 independent and processed by unknown factors from a piRNA transcript. Primary siRNAs or piRNAs loaded into their respective argonaute proteins bind to target RNA and recruit the necessary proteins for 22G-RNA production. 22G-RNAs are loaded into worm specific argonautes that can mediate PTGS or TGS. miRNAs are processed by DCR-1 from dsRNA hairpins. miRNAs do not feed into the 22G-RNA pathway, but directly target mRNA for translational repression. Black dots represent 5'-capped RNA and yellow dots represent methyl groups.

1.3 Antiviral RNAi

Antiviral RNAi was first discovered in plants in 1993, when transgenic plants expressing the coat protein from Tobacco Etch Virus (TEV) were found to be resistant to TEV infection and upon recovering from TEV infection no longer expressed the coat protein (3). In 1998, a few viruses were found to encode proteins that could alleviate transgene silencing indicating that gene silencing had an antiviral function (64, 65). Soon after, in 1999, virus derived siRNAs (vsiRNAs) were detected by Northern Blot in virus infected plants (5) and in 2002, vsiRNAs were first detected in animals (6). RNAi was ultimately found to be a wide-spread antiviral mechanism in plants and insects, as well as some fungi and *C. elegans* (7–9, 11). In 2013, two independent studies demonstrated that antiviral RNAi is also active in mice (66, 67).

In mammals, there are two types of immunity: innate and adaptive. Adaptive immunity has been extensively studied for centuries, starting with variolation and vaccination, whereas the paradigm-shifting discovery of the innate immune system was not made until the 1990s. The innate immune system is an immediate response to pathogens that is encoded in the germline and has no specificity and no memory. In contrast, the adaptive immune system takes time to mount a response and arises from genetic rearrangements and mutations in specialized cells, making it highly specific and providing memory (i.e. protection against future infections with the same virus). Most infections are cleared by the innate immune system and the adaptive immune response is required only when a pathogen has evolved to suppress innate immunity. Antiviral RNAi is traditionally considered an innate immune mechanism and its antiviral activity is most

potent in plants and invertebrates, which lack an adaptive immune system. In a typical innate immune response, a Pathogen Associated Molecular Pattern (PAMP) is recognized by a Pattern Recognition Receptor (PRR), which acts as a sensor to initiate an effector mechanism and bring about a host immune response. In response to RNA virus replication, dsRNA is a non-self signature that acts as a PAMP and is detected by the PRR Dicer, which cleaves the long dsRNA into small RNAs. The small RNAs are loaded into an Argonaute protein, and together these two are the Effector that mediates an immune response, namely destruction of viral genomic RNA. Dicer can also be considered an effector, as it directly destroys viral genomic RNA; however, its role as an effector is limited as dicing alone is insufficient to clear viral infection. In contrast to traditional innate immunity, antiviral RNAi is specific, like adaptive immunity and a mounted RNAi response against one virus is powerless to clear a second virus. In contrast, activation of interferon-based innate immunity by one virus can protect against a second virus.

Viral dsRNA can be recognized by the host as a PAMP, either through extensive secondary structures in the viral genomic RNA, or through the dsRNA produced during RNA virus replication. vsiRNAs that are predominantly sense polarity originate from secondary structures, while vsiRNAs that are equally sense and antisense polarity are derived from long dsRNA. During RNA replication, viral dsRNA replication intermediates are formed during the synthesis of the genomic, antigenomic or sub-genomic RNAs. If we consider a (+)-sense RNA virus, the viral dsRNA replication intermediates are the most populous during synthesis of the genomic (+)-sense RNA

from the antigenomic (-)-sense RNA template. Viral dsRNA replication intermediates are also made during synthesis of the (-)-sense RNA from the (+)-sense RNA, though at a lower abundance. In *D. melanogaster,* vsiRNAs produced against some viruses are predominantly derived from the 5' terminal region of the viral RNA, indicating they originate from viral dsRNA replication intermediates from (+)-sense RNA synthesis (68). Dicer likely recognizes different sources of RNA in response to different RNA viruses.

Antiviral RNAi has been extensively studied in two model organisms, *A. thaliana* and *D. melanogaster*. Over 80% of plant viruses have an RNA genome and antiviral RNAi is the central antiviral pathway in plants. In the model plant, *A. thaliana*, antiviral RNAi is initiated by the production of primary vsiRNAs followed by amplification through secondary vsiRNA production. *A. thaliana* encodes four Dicer-like (DCL) proteins that have distinct functions in parallel small RNA pathways. DCL1 is required for the biogenesis of miRNAs, while DCL4, DCL2 and DCL3 produce three groups of siRNAs that are 21-nt, 22-nt and 24-nt in length, respectively (69). The most abundant vsiRNAs are 21-nt in length and are produced by DCL4, though 22-nt vsiRNAs made by DCL2 are sufficient to mount an antiviral response in the absence of DCL4 (70–73). vsiRNAs produced from viral dsRNA are primary vsiRNAs that initiate amplification of the small RNA signal via secondary vsiRNAs. Secondary vsiRNAs are produced from long dsRNA synthesized by endogenous RdRPs, RDR1 or RDR6. Primary vsiRNAs alone do not confer a robust antiviral response, as *rdr1*/*rdr6* double mutants are unable to clear viral infection. Plant secondary siRNAs differ from *C. elegans* 22G-RNAs, as they are dependent on the same biogenesis machinery as primary siRNAs, such as DCL4, and

they are biochemically indistinguishable from primary siRNAs. Antiviral Argonaute proteins load vsiRNAs and can mediate clearance or inhibit translation of viral RNA or regulate host gene expression. AGO1 and AGO2 appear to be the main antiviral argonautes, each acting as the primary antiviral argonaute in response to different viruses (74, 75). AGO4 can also function in antiviral RNAi against both RNA and DNA viruses and AGO5, AGO7 and AGO10 have also been reported to have a minor antiviral role. *A. thaliana* continues to be an important model for antiviral RNAi, as many crops rely on RNAi to defend against both RNA and DNA viruses.

The model organism, *D. melanogaster*, also encodes multiple Dicer proteins, Dicer 1 (DCR-1) and DCR-2 that function in the miRNA and siRNA pathway, respectively (69). In response to viral replication, viral dsRNA replication intermediates are processed by DCR-2 and Loq-PD into vsiRNAs that are predominantly 21-nt in length. vsiRNAs are methylated by HEN1 at the 3' end and loaded into AGO2 by DCR-2 and the dsRNA binding protein, R2D2. The AGO2-vsiRNA complex mediates clearance of viral genomic RNA by cleaving the target RNA. Because DCR-2 is specifically required for the siRNA pathway, *dcr-2* mutant flies are viable and provide a unique opportunity to study the function of dicer in antiviral RNAi in animals. Additionally, antiviral RNAi can be studied without compromising an important endogenous pathway.

The study of antiviral RNAi in the nematode, *C. elegans*, has lagged behind *A. thaliana* and *D. melanogaster*, because there was no known virus that could naturally infect *C. elegans* until 2011. However, *C. elegans* is an excellent system to study antiviral RNAi, because it more closely mirrors the RNAi machinery found in mammals. Both

mammals and *C. elegans* only encode one dicer protein that must function in competing RNAi pathways and both encode an expanded class of duplex RNA activated ATPases (DRAs), which include dicer, the mammalian RIG-I-like receptors and *C. elegans* Dicer Related Helicases. Additionally, RNAi in *C. elegans* requires the production of secondary siRNAs by endogenous RdRPs. However in contrast to plants, in *C. elegans* the populations of primary and secondary siRNAs are biochemically and physically distinct, allowing researchers to uncouple the two steps in a manner that is significantly more difficult in *A. thaliana*.

1.4 *C. elegans* as a model to study antiviral immunity

Invertebrate organisms lack an interferon-based innate immune system and an adaptive immune system making them ideal systems to study antiviral RNAi as an innate immune mechanism. In fact key components of both innate immunity and RNAi have been revealed by genetic screens in invertebrate models (16, 76, 77). *C. elegans* has been a choice model to study a variety of complex cellular pathways for nearly fifty years (78) and several features contribute to its popularity. The worms grow to be just 1mm as adults with a short life cycle of only three days and a brood size of around 300. They are easily maintained in the laboratory on agar plates and can be frozen at -80°C for decades. Additionally, *C. elegans* are mostly hermaphrodites, with less than 0.5% of the population being males, allowing for easy maintenance of strains, but also genetic crosses when necessary. Harnessing such a tractable model to study antiviral RNAi will undoubtedly lead to new discoveries.

1.4.1 Artificial Infection systems

In 2005, three simultaneous studies demonstrated an antiviral role for the RNAi machinery in *C. elegans* using viruses that do not naturally infect *C. elegans*. Two groups utilized an artificial infection system using Vesicular Stomatitis Virus (VSV). VSV is a (-)-sense RNA virus with a single genomic RNA that naturally infects biting flies and mammalian livestock. VSV can also infect cells derived from *C. elegans*, but not the animals themselves. VSV infection is relatively weak in cells derived from wild-type animals, but robust viral infection was found in cells derived from RNAi-defective (RDE) mutants (8, 9). These results indicate that RNAi is required to clear VSV infection from *C. elegans* cells. Additionally, cells derived from mutants with an ERI phenotype, displayed enhanced resistance to VSV (8, 9). These experiments demonstrated that viruses could infect and replicate in worm cells and that *C. elegans* encodes the necessary RNAi machinery to clear viral infection.

1.4.2 Flock House Virus

In order to efficiently study antiviral immunity in *C. elegans*, our lab developed a transgenic system utilizing Flock House Virus (FHV), which naturally infects insects, but is also capable of replicating in mammalian cells, yeast and plants (7, 79). FHV is an insect Nodavirus isolated from the grass grub in Australia and while it has no economic or medical significance, FHV has been extensively utilized as a model to better understand fundamental aspects of virology applicable to many RNA virus families. The FHV genome consists of two (+)-sense RNAs (Fig. 1-2A). RNA1 encodes the viral RdRP

and produces a sub-genomic RNA3 (sgRNA3) during replication. sgRNA3 encodes for the B1 and B2 proteins, which encode a protein of unknown function and a Viral Suppressor of RNAi (VSR), respectively. The B2 protein is capable of suppressing RNAi in *D. melanogaster* and *A. thaliana* (6) and is essential for replication to occur in organisms that use RNAi as an antiviral defense mechanism. Loss of B2 function results in an inability for the virus to efficiently replicate that can be rescued by genetic suppression of the host RNAi machinery. FHV RNA2 encodes the α coat protein that is required for particle assembly and is then self-cleaved into proteins β and γ during the maturation process required for infectivity.

FHV RNA1 and RNA2 are capable of efficient autonomous replication when integrated into the *C. elegans* genome on a heat-inducible transgene (7). Consistent with studies in *D. melanogaster* and *A. thaliana*, when the viral replicon contains a missense mutation in B2 (FR1ΔB2) or B2 is replaced by enhanced GFP (FR1gfp) replication is abolished (18). Replication of the FR1ΔB2 or FR1gfp transgenes can be rescued by introducing a mutation in *rde-1* or *rde-4*, key components of exogenous RNAi, indicating that B2 functions as a suppressor of RNAi in *C. elegans* (7, 18). Additionally, replication of the FHV replicon depends on the viral RdRP and replicates autonomously days after heat induction. When used in *C. elegans*, the FR1gfp replicon has key characteristics that make it ideal for the genetic analysis of the antiviral RNAi pathway. First, the viral replicon is capable of replication only when antiviral RNAi is genetically suppressed in the host. Second, *C. elegans* animals are transparent and viral replication can be observed throughout the living, intact animal via GFP. Third, *C. elegans* is a genetically tractable

model with efficient methods for both forward and genetic screens. The use of FR1gfp in a targeted feeding RNAi screen successfully identified several genes required for antiviral RNAi, including DRH-1 and RSD-2 (18).

1.4.3 Caenorhabditis Viruses: Orsay, Santueil and La Blanc Viruses

The discovery of the Orsay Virus (OrV) was reported in 2011 and is the only known natural viral pathogen of *C. elegans*. OrV was isolated from the JU1580 strain of *C. elegans* in an orchard in Orsay, France, while two other viruses, Santueil and La Blanc, were isolated from *C. briggsae* in eponymous cities (80). All three viruses infect intestinal cells and infection is species specific (80, 81). Serendipitously, these three *Caenorhabditis* viruses are most closely related to Nodaviruses, such as FHV. However, they lack several characteristics common of both fish and insect Nodaviruses and likely comprises a novel virus family. The OrV genome consists of two (+)-sense genomic RNAs (Fig 1-2B). OrV RNA1 encodes only a single protein, the RdRP and, unlike Nodaviruses, there is no evidence of a sgRNA3 or the VSR, B2. OrV RNA2 contains two open reading frames, α and δ, which encode the viral coat protein and a novel protein, respectively. The δ protein is expressed by ribosomal frameshifting as a α-δ fusion protein in OrV, and likely Santeuil and La Blanc as well (82). Neither the α or δ protein individually appears to encode for a VSR (83); however the α-δ fusion protein was not tested as the original study predated its discovery.

While the Orsay Virus efficiently infects the wild isolate of *C. elegans* in which it was found, JU1580, it is only capable of limited replication in the wild-type laboratory

strain, N2 (80). In order observe efficient replication of OrV in the N2 strain it must be genetically deficient in its core RNAi machinery. In fact, the JU1580 strain contains a 159bp deletion in *drh-1* (20). Interestingly, this *drh-1* allele is found at an intermediate frequency wild isolates of *C. elegans*. While it is possible that the deletion is linked to another mutation or that viral infection is unlikely in nature, JU1580 also has an altered endogenous small RNA population that promotes antibacterial gene regulation (84). As bacterivores, protection against bacterial pathogens is essential for *C. elegans* and may explain the prevalence of this allele in nature. The inability of OrV to efficiently infect N2 provides another line of evidence that OrV does not encode a VSR.

The relationship between OrV and *C. elegans* is a natural example that demonstrates an antiviral role for RNAi and provides another invaluable tool for researchers to further study the phenomenon. Additionally, the discovery of OrV has the potential to answer questions unrelated to RNAi and virus replication, such as immune mechanisms that target other steps of the virus life cycle or the study of viral transmission. In addition to investigating host immunity, a recent study has demonstrated that OrV can be reverse engineered directly in *C. elegans* animals, allowing researchers to probe the function of viral sequences and proteins. Amazingly, infectious, mutant virus particles can be obtained from this system and used to infect different *C. elegans* strains (85). This study established *C. elegans* and OrV as a unique and convenient system that allows for manipulation of both the animal host and the virus in parallel.

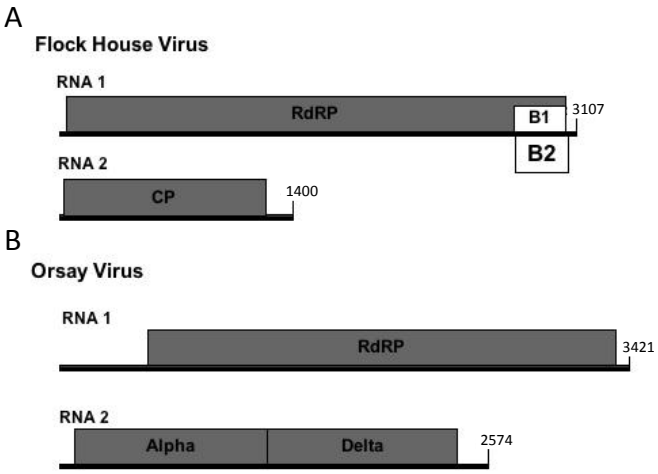

Figure 1-2. Viruses used to study antiviral RNAi in *C. elegans*. Genome oraganization of Flock House Virus (A) and Orsay Virus (B).

1.5 Antiviral RNAi in *C. elegans*

1.5.1 The Duplex RNA Activated ATPase, DRH-1

DRH-1 is required for antiviral RNAi (18–20), but it is not required for exo-RNAi induced by the artificial introduction of dsRNA (18). To date, there has been debate about the placement of DRH-1 in the antiviral RNAi pathway, which this work resolves. The first analysis of vsiRNA production in *drh-1* mutants by small RNA Northern, demonstrated that vsiRNAs targeting FR1gfp were made in the absence of DRH-1, placing DRH-1 downstream of initial vsiRNA biogenesis (18). Analysis of small RNAs sequenced from OrV infected *drh-1* animals, revealed that both the production of primary and secondary vsiRNAs were compromised (20). In particular, small RNA libraries that captured only primary vsiRNAs, showed a depletion in primary vsiRNAs, leading to the hypothesis that DRH-1 functions upstream of initial primary vsiRNA biogenesis, perhaps

in viral sensing homologous to RIG-I (20). This work analyzes the function of DRH-1 in primary vsiRNA production, utilizing the recently published, full-length OrV genome. Our results resolves these conflicting reports and suggests the DRH-1 functions in the production of primary vsiRNAs from the internal regions of viral dsRNA, downstream of the initial production of primary vsiRNAs from the 5' terminal regions that is dependent on DCR-1 and RDE-4. A model for DRH-1 function in primary vsiRNA biogenesis is discussed in Chapter 3.

Mello and colleagues (16) were the first to note the strong sequence similarity among the helicase/ATPase domains encoded by Dicer ribonucleases, *C. elegans* DRH-1 and DRH-3, and the mammalian RIG-I-like receptors (RLRs). Mammalian RLRs include the retinoic acid inducible gene I (RIG-I) and melanoma differentiation-associated gene 5 (MDA5) that act as cytosolic innate immune receptors to detect two distinct sets of RNA viruses (86). Binding of the viral dsRNA by the ATPase and C-terminal regulatory domains (RD) of RIG-I causes a conformational change that releases the N-terminal tandem caspase activation and recruitment domains (CARDs) for the downstream signaling events, leading to the production of type-1 interferons and the expression of numerous interferon-stimulated genes (87). RIG-I is also known to translocate along dsRNA powered by ATP hydrolysis (88), but the biological function of the translocase activity of RIG-I is unclear (87, 89). Interestingly, both the ATPase domain and RD of human RIG-I can functionally replace the corresponding domains of DRH-1 to mediate antiviral RNAi in *C. elegans* (19). Moreover, the conserved KWK motif in the RD that is implicated in dsRNA specificity of RIG-I is also required for the antiviral activity of both

DRH-1 and the chimeric protein containing the N-terminal domain from DRH-1 and the ATPase and RD of RIG-I (19). These findings suggest that DRH-1 and RIG-I may share the same RNA specificity and other biochemical properties such as translocation along dsRNA (19, 89). Recently, Dicer, DRHs and RLRs have been collectively named as double-stranded RNA-activated ATPases (DRAs) because of several shared properties including the presence of a unique α-helical insertion domain (HEL2i) and the lack of dsRNA unwinding activity (89, 90).

1.5.2 Genetic Requirements of Antiviral RNAi

Primary vsiRNAs are produced by DCR-1, RDE-4 and DRH-1 and are identical to primary siRNAs produced from exogenous long dsRNA. In the absence of either DRH-1 or RDE-4, some primary and secondary vsiRNAs are still produced, but are completely depleted in the double mutant (20). Primary vsiRNAs are loaded in the argonaute, RDE-1, which base pairs with a target viral RNA molecule to recruit factors required for 22G-RNA biogenesis. Analysis of OrV infection in *rsd*-2, *rsd-6, rde-10, rde-11* and *rde-8* mutants, suggest that the RDE-10/RDE-11 and RDE-8 complexes function in antiviral RNAi (33, 91, 92). The RdRP complex carrying RRF-1, but not EGO-1, and DRH-3 are essential for the production of viral 22G-RNAs (18–20). Several WAGOs have been implicated in antiviral RNAi and vsiRNAs in the MAGO-12 mutant demonstrate that the WAGOs are required for 22G-RNA production or for their stability (20). Given the instability of short, ssRNA it is likely that immediate WAGO loading is

required to stabilize 22G-RNAs. A robust secondary vsiRNA response is essential to effective clear viral infection mediated by RNAi.

Antiviral RNAi in plants is systemic and travels short distances from cell to cell and long distances from roots to shoots through the phloem. In *C. elegans* exo-RNAi is systemic; however Antiviral RNAi does not require SID-1, a key component of systemic RNAi (91, 92). Furthermore, in a transgenic system the silencing signal induced in a single cell by OrV infection does not spread to neighboring cells (93). Exo-RNAi in *C. elegans* is also inherited. Two studies in *C. elegans* suggest that a silencing signal targeting both FR1gfp and OrV can be inherited (94, 95). Researchers demonstrated that vsiRNAs corresponding to FR1gfp were inherited in subsequent generations that were unable to produce vsiRNAs, such as *rde-4* (94). However, FR1gfp contains a genomic multi-copy array of the FR1gfp replicon, which may be targeted for silencing. Indeed, we have found that vsiRNAs are produced in this strain even in the absence of heat treatment (Chapter 3). The evidence for inheritance of OrV resistance relies on increasing the sensitivity of wild-type animals to OrV and demonstrated decreased levels of OrV infection in subsequent generations (95). However, a more recent study was unable to demonstrate the propagation of inherited vsiRNAs from OrV infection, but was able to detect inherited siRNAs from exo-RNAi targeting a somatic gene (93). Additionally, when a genomic copy of OrV was present, inheritance was observed in F1s, but subsequent generations were not protected (93). Taken together, these data suggest that the limited inheritance of antiviral RNAi observed in FR1gfp is dependent on the

genomic copy and inheritance of OrV silencing is either unlikely or not dependent on small RNAs.

1.6 Conclusions

Today is an exciting time to be studying antiviral RNAi in *C. elegans*. Before the discovery of OrV in 2011, five papers on antiviral RNAi in *C. elegans* were published over the course of seven years. In contrast, twelve papers have been published in the last four years since its discovery. This recent body of work has pieced together an antiviral RNAi pathway similar to the exo-RNAi pathway, with the exception of DRH-1, which is uniquely essential for antiviral RNAi. This dissertation aims to harness the genetic tractability of *C. elegans* to develop an unbiased genetic screen to identify the genetic requirements of antiviral RNAi. While many unbiased genetic screens have been conducted to identify the genetic requirements of exo-RNAi (16, 77, 96, 97), this screen is the first with the ability to identify genes uniquely required for antiviral RNAi, as well as novel antiviral pathways. Additionally, chemical mutagenesis by Ethyl Methanesulfonate causes single base changes that can generate a variety of phenotypically unique alleles in a single gene. A targeted feeding RNAi screen has demonstrated the feasibility of using FR1gfp in *C. elegans* to identify genes required for antiviral RNAi, such as DRH-1 (18). However the precise function of DRH-1 remains unknown. The homology between DRH-1 and RIG-I suggests a role as sensor of viral RNA and/or as a translocase that can move along dsRNA. Using the recently-published full-length sequence of the OrV genome (82), this work utilizes several alleles of *drh-1* to

address the role of DRH-1 in primary vsiRNA production. Furthermore, the pattern of primary vsiRNA biogenesis uncovered in the *drh-1* alleles leads to a discussion of the source of dsRNA PAMPs in *C. elegans*. The work presented in this dissertation establishes a genetic screen that can unveil important factors of innate immunity and delves into the function of an important conserved factor of innate antiviral immunity.

1.7 References

1. Napoli C, Lemieux C, Jorgensen R (1990) Introduction of a Chimeric Chalcone Synthase Gene into Petunia Results in Reversible Co-Suppression of Homologous Genes in trans. *Plant Cell* 2(4):279–289.

2. van der Krol a R, Mur L a, Beld M, Mol JN, Stuitje a R (1990) Flavonoid genes in petunia: addition of a limited number of gene copies may lead to a suppression of gene expression. *Plant Cell* 2(4):291–299.

3. Lindbo J, Silva-Rosales L, Proebsting W, Dougherty W (1993) Induction of a Highly Specific Antiviral State in Transgenic Plants: Implications for Regulation of Gene Expression and Virus Resistance. *Plant Cell* 5(12):1749–1759.

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

5. Hamilton a J, Baulcombe DC (1999) A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science* 286(5441):950–952.

6. Li H, Li WX, Ding SW (2002) Induction and suppression of RNA silencing by an animal virus. *Science* 296(5571):1319–1321.

7. Lu R, et al. (2005) Animal virus replication and RNAi-mediated antiviral silencing in Caenorhabditis elegans. *Nature* 436(7053):1040–1043.

8. Wilkins C, et al. (2005) RNA interference is an antiviral defence mechanism in Caenorhabditis elegans. *Nature* 436(7053):1044–1047.

9. Schott DH, Cureton DK, Whelan SP, Hunter CP (2005) An antiviral role for the RNA interference machinery in Caenorhabditis elegans. *Proc Natl Acad Sci U S A* 102(51):18420–18424.

10. Lye LF, et al. (2010) Retention and Loss of RNA interference pathways in trypanosomatid protozoans. *PLoS Pathog* 6(10):e1001161.

11. Drinnenberg I a, Fink GR, Bartel DP (2011) Compatibility with killer explains the rise of RNAi-deficient fungi. *Science* 333(6049):1592.

12. Bernstein E, Caudy a a, Hammond SM, Hannon GJ (2001) Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* 409(6818):363–366.

13. Bagijn MP, et al. (2012) Function, Targets and Evolution of Caenorhabditis elegans piRNAs. 337(574). doi:10.1126/science.1220952.

14. Lee H-C, et al. (2012) C. elegans piRNAs Mediate the Genome-wide Surveillance of Germline Transcripts. *Cell* 150(1):78–87.

15. Duchaine TF, et al. (2006) Functional proteomics reveals the biochemical niche of C. elegans DCR-1 in multiple small-RNA-mediated pathways. *Cell* 124(2):343–354.

16. Tabara H, Yigit E, Siomi H, Mello CC (2002) The dsRNA binding protein RDE-4 interacts with RDE-1, DCR-1, and a DExH-Box helicase to direct RNAi in C. elegans. *Cell* 109(7):861–871.

17. Parker GS, Eckert DM, Bass BL (2006) RDE-4 preferentially binds long dsRNA and its dimerization is necessary for cleavage of dsRNA to siRNA. *RNA* 12(5):807–818.

18. Lu R, Yigit E, Li WX, Ding SW (2009) An RIG-I-like RNA helicase mediates antiviral RNAi downstream of viral siRNA biogenesis in Caenorhabditis elegans. *PLoS Pathog* 5(2):8–14.

19. Guo X, Zhang R, Wang J, Ding S-W, Lu R (2013) Homologous RIG-I-like helicase proteins direct RNAi-mediated antiviral immunity in C. elegans by distinct mechanisms. *Proc Natl Acad Sci U S A* 110(40):16085–90.

20. Ashe A, et al. (2013) A deletion polymorphism in the Caenorhabditis elegans RIG-I homolog disables viral RNA dicing and antiviral immunity. *Elife* 2013(2):1–21.

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

22. Corrêa RL, Steiner F a, Berezikov E, Ketting RF (2010) MicroRNA-directed siRNA biogenesis in Caenorhabditis elegans. *PLoS Genet* 6(4):e1000903.

23. Ketting F, Steiner FA, Okihara KL, Hoogstrate SW, Sijen T (2009) RDE-1 slicer activity is required only for passenger-strand cleavage during RNAi in Caenorhabditis elegans. 16(2):207–211.

24. Aoki K, Moriguchi H, Yoshioka T, Okawa K, Tabara H (2007) In vitro analyses of the production and activity of secondary small interfering RNAs in C. elegans. *EMBO J* 26(24):5007–19.

25. Gu W, et al. (2009) Distinct Argonaute-Mediated 22G-RNA Pathways Direct Genome Surveillance in the C. elegans Germline. *Mol Cell* 36(2):231–244.

26. Kao CC, Singh P, Ecker DJ (2001) De novo initiation of viral RNA-dependent RNA synthesis. *Virology* 287(2):251–260.

27. Sijen T, et al. (2001) On the role of RNA in gene amplification. *Cell* 107:465–476.

28. Pak J, Maniar JM, Mello CC, Fire A (2012) Protection from Feed-Forward Amplification in an Amplified RNAi Mechanism. *Cell* 151(4):885–899.

29. Sapetschnig A, Sarkies P, Lehrbach NJ, Miska E a. (2015) Tertiary siRNAs Mediate Paramutation in C. elegans. *PLOS Genet* 11(3):e1005078.

30. Fitzgerald ME, Vela A, Pyle AM (2014) Dicer-related helicase 3 forms an obligate dimer for recognizing 22G-RNA. *Nucleic Acids Res* 42(6):3919–3930.

31. Yang H, et al. (2012) The RDE-10/RDE-11 complex triggers RNAi-induced mRNA degradation by association with target mRNA in C. elegans. *Genes Dev* 26(8):846–856.

32. 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.

33. Tsai H-Y, et al. (2015) A Ribonuclease Coordinates siRNA Amplification and mRNA Cleavage during RNAi. *Cell* 160(3):407–419.

34. Guang S, et al. (2008) An Argonaute transports siRNAs from the cytoplasm to the nucleus. *Science* 321(5888):537–541.

35. Shirayama M, et al. (2012) piRNAs Initiate an Epigenetic Memory of Nonself RNA in the C. elegans Germline. *Cell* 150(1):65–77.

36. Ashe A, et al. (2012) piRNAs Can Trigger a Multigenerational Epigenetic Memory in the Germline of C. elegans. *Cell* 150(1):88–99.

37. Lee RC, Feinbaum RL, Ambros V (1993) The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell* 75(5):843–54.

38. Reinhart BJ, et al. (2000) The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. *Nature* 403(6772):901–906.

39. Simmer F, et al. (2002) Loss of the putative RNA-directed RNA polymerase RRF-3 makes C. elegans hypersensitive to RNAi. *Curr Biol* 12(15):1317–9.

40. Ruby JG, et al. (2006) Large-Scale Sequencing Reveals 21U-RNAs and Additional MicroRNAs and Endogenous siRNAs in C. elegans. *Cell* 127(6):1193–1207.

41. Welker NC, et al. (2010) Dicer's helicase domain is required for accumulation of some, but not all, C. elegans endogenous siRNAs. *Rna* 16(5):893–903.

42. Reinke V, Gil IS, Ward S, Kazmer K (2004) Genome-wide germline-enriched and sex-biased expression profiles in Caenorhabditis elegans. *Development* 131(2):311–323.

43. Han T, et al. (2009) 26G endo-siRNAs regulate spermatogenic and zygotic gene expression in Caenorhabditis elegans. *Proc Natl Acad Sci U S A* 106(44):18674–18679.

44. Vasale JJ, et al. (2010) Sequential rounds of RNA-dependent RNA transcription drive endogenous small-RNA biogenesis in the ERGO-1/Argonaute pathway. *Proc Natl Acad Sci* 107(8):3582–3587.

45. Conine CC, et al. (2010) Argonautes ALG-3 and ALG-4 are required for spermatogenesis-specific 26G-RNAs and thermotolerant sperm in Caenorhabditis elegans. *Proc Natl Acad Sci U S A* 107(8):3588–3593.

46. Billi AC, et al. (2012) The Caenorhabditis elegans HEN1 Ortholog, HENN-1, Methylates and Stabilizes Select Subclasses of Germline Small RNAs. *PLoS Genet* 8(4):e1002617.

47. Kamminga LM, Wolfswinkel JC Van, Luteijn MJ, Kaaij LJT (2012) Differential Impact of the HEN1 Homolog HENN-1 on 21U and 26G RNAs in the Germline of Caenorhabditis elegans. 8(7). doi:10.1371/journal.pgen.1002702.

48. 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.

49. Weick E-M, Miska E a (2014) piRNAs: from biogenesis to function. *Development* 141(18):3458–3471.

50. Cecere G, Zheng GXY, Mansisidor AR, Klymko KE, Grishok A (2012) Promoters Recognized by Forkhead Proteins Exist for Individual 21U-RNAs. *Mol Cell* 47(5):734–745.

51. Weick E-MM, et al. (2014) PRDE-1 is a nuclear factor essential for the biogenesis of Ruby motif-dependent piRNAs in C. elegans. *Genes Dev* 28(7):783–96.

52. Goh W-SS, et al. (2014) A genome-wide RNAi screen identifies factors required for distinct stages of C. elegans piRNA biogenesis. *Genes Dev* 28(7):797–807.

53. de Albuquerque BFM, Placentino M, Ketting RF (2015) Maternal piRNAs Are Essential for Germline Development following De Novo Establishment of Endo-siRNAs in Caenorhabditis elegans. *Dev Cell* 34(4):448–456.

54. 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.

55. Montgomery T a., et al. (2012) PIWI Associated siRNAs and piRNAs Specifically Require the Caenorhabditis elegans HEN1 Ortholog henn-1. *PLoS Genet* 8(4):e1002616.

56. 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.

57. Buckley BA, et al. (2012) A nuclear Argonaute promotes multigenerational epigenetic inheritance and germline immortality. *Nature* 489(7416):447–451.

58. Luteijn MJ, et al. (2012) Extremely stable Piwi-induced gene silencing in Caenorhabditis elegans. *EMBO J* 31(16):3422–30.

59. van Wolfswinkel JC, et al. (2009) CDE-1 Affects Chromosome Segregation through Uridylation of CSR-1-Bound siRNAs. *Cell* 139(1):135–148.

60. Claycomb JM, et al. (2009) The Argonaute CSR-1 and Its 22G-RNA Cofactors Are Required for Holocentric Chromosome Segregation. *Cell* 139(1):123–134.

61. Avgousti DC, Palani S, Sherman Y, Grishok A (2012) CSR-1 RNAi pathway positively regulates histone expression in C. elegans. *EMBO J* 31(19):3821–32.

62. Wedeles CJ, Wu MZ, Claycomb JM (2013) A multitasking Argonaute: Exploring the many facets of C. elegans CSR-1. *Chromosom Res* 21(6-7):573–586.

63. Cecere G, Hoersch S, O'Keeffe S, Sachidanandam R, Grishok A (2014) Global effects of the CSR-1 RNA interference pathway on the transcriptional landscape. *Nat Struct Mol Biol* 21(4):358–65.

64. Kasschau KD, Carrington JC (1998) A Counterdefensive Strategy of Plant Viruses: Suppression of Posttranscriptional Gene Silencing. *Cell* 95(4):461–470.

65. Anandalakshmi R, et al. (1998) A viral suppressor of gene silencing in plants. *Proc Natl Acad Sci U S A* 95(October):13079–13084.

66. Maillard P V, et al. (2013) Antiviral RNA interference in mammalian cells. *Science* 342(6155):235–8.

67. Li Y, Lu J, Han Y, Fan X, Ding S-W (2013) RNA interference functions as an antiviral immunity mechanism in mammals. *Science* 342(6155):231–4.

68. Aliyari R, et al. (2008) Mechanism of Induction and Suppression of Antiviral Immunity Directed by Virus-Derived Small RNAs in Drosophila. *Cell Host Microbe* 4(4):387–397.

69. Ding S-W (2010) RNA-based antiviral immunity. *Nat Rev Immunol* 10(9):632–644.

70. Diaz-Pendon J a, Li F, Li W-X, Ding S-W (2007) Suppression of antiviral silencing by cucumber mosaic virus 2b protein in Arabidopsis is associated with drastically reduced accumulation of three classes of viral small interfering RNAs. *Plant Cell* 19(6):2053–2063.

71. Bouché N, Lauressergues D, Gasciolli V, Vaucheret H (2006) An antagonistic function for Arabidopsis DCL2 in development and a new function for DCL4 in generating viral siRNAs. *EMBO J* 25(14):3347–3356.

72. Deleris A, et al. (2006) Hierarchical action and inhibition of plant Dicer-like proteins in antiviral defense. *Science* 313(5783):68–71.

73. Fusaro AF, et al. (2006) RNA interference-inducing hairpin RNAs in plants act through the viral defence pathway. *EMBO Rep* 7(11):1168–1175.

74. Carbonell A, Carrington JC (2015) Antiviral roles of plant ARGONAUTES. *Curr Opin Plant Biol* 27(Dcl):111–117.

75. Wang X-B, et al. (2011) The 21-nucleotide, but not 22-nucleotide, viral secondary small interfering RNAs direct potent antiviral defense by two cooperative argonautes in Arabidopsis thaliana. *Plant Cell* 23(4):1625–1638.

76. Lemaitre B, Nicolas E, Michaut L, Reichhart JM, Hoffmann JA (1996) The dorsoventral regulatory gene cassette spatzle/Toll/Cactus controls the potent antifungal response in Drosophila adults. *Cell* 86(6):973–983.

77. Tabara H, et al. (1999) The rde-1 gene, RNA interference, and transposon silencing in C. elegans. *Cell* 99(2):123–132.

78. Brenner S (1974) The genetics of Caenorhabditis elegans. *Genetics* 77(1):71–94.

79. Venter P, Jovel J, Schneeman A (2010) Chapter 11: Insect Nodaviruses. *Insect Virology*, eds Asgari S, Johnson K (Caister Academic Press), pp 251–282.

80. Félix MA, et al. (2011) Natural and experimental infection of Caenorhabditis nematodes by novel viruses related to nodaviruses. *PLoS Biol* 9(1). doi:10.1371/journal.pbio.1000586.

81. Franz CJ, et al. (2014) Orsay, santeuil and le blanc viruses primarily infect intestinal cells in caenorhabditis nematodes. *Virology* 448:255–264.

82. Jiang H, et al. (2014) Orsay virus utilizes ribosomal frameshifting to express a novel protein that is incorporated into virions. *Virology* 450-451:213–221.

83. Guo X, Lu R (2013) Characterization of virus-encoded RNA interference suppressors in Caenorhabditis elegans. *J Virol* 87(10):5414–23.

84. Sarkies P, Ashe A, Le Pen J, McKie M a., Miska E a. (2013) Competition between virus-derived and endogenous small RNAs regulates gene expression in Caenorhabditis elegans. *Genome Res* 23(8):1258–1270.

85. Jiang H, Franz CJ, Wang D (2014) Engineering Recombinant Orsay Virus Directly in the Metazoan Host C. elegans. *J Virol* (July). doi:10.1128/JVI.01630-14.

86. Reikine S, Nguyen JB, Modis Y (2014) Pattern Recognition and Signaling Mechanisms of RIG-I and MDA5. *Front Immunol* 5:1–7.

87. Yoneyama M, Onomoto K, Jogi M, Akaboshi T, Fujita T (2015) Viral RNA detection by RIG-I-like receptors. *Curr Opin Immunol* 32:48–53.

88. Myong S, et al. (2009) Cytosolic Viral Sensor RIG-I Is a5'-triphosphate-dependent translocase on double-stranded RNA. *Science (80-)* 323:1070–1074.

89. Ahmad S, Hur S (2015) Helicases in Antiviral Immunity : Dual Properties as Sensors and Effectors. *Trends Biochem Sci* 40(10):576–585.

90. Luo D, Kohlway A, Pyle AM (2013) Duplex RNA activated ATPases (DRAs): platforms for RNA sensing, signaling and processing. *RNA Biol* 10(1):111–20.

91. Guo X, Zhang R, Wang J, Lu R (2013) Antiviral RNA silencing initiated in the absence of RDE-4, a double-stranded RNA binding protein, in Caenorhabditis elegans. *J Virol* 87(19):10721–9.

92. Zhong J (2014) Genetics characterization of antiviral RNA interference in Caenorhabditis elegans. Dissertation (University of California, Riverside).

93. Ashe A, Sarkies P, Le Pen J, Tanguy M, Miska EA (2015) Antiviral RNAi against Orsay virus is neither systemic nor transgenerational in Caenorhabditis elegans. *J Virol*:JVI.03664–14.

94. Rechavi O, Minevich G, Hobert O (2011) Transgenerational inheritance of an acquired small RNA-based antiviral response in C. elegans. *Cell* 147(6):1248–1256.

95. Sterken MG, et al. (2014) A heritable antiviral RNAi response limits orsay virus infection in Caenorhabditis elegans N2. *PLoS One* 9(2). doi:10.1371/journal.pone.0089760.

96. Winston WM, Molodowitch C, Hunter CP (2002) Systemic RNAi in C. elegans requires the putative transmembrane protein SID-1. *Science* 295(5564):2456–2459.

97. Tijsterman M, May RC, Simmer F, Okihara KL, Plasterk RH a (2004) Genes required for systemic RNA interference in Caenorhabditis elegans. *Curr Biol* 14(2):111–116.

CHAPTER 2

FORWARD GENETIC SCREEN TO IDENTIFY THE GENETIC REQUIREMENTS OF ANTIVIRAL RNAI

2.1 Abstract

RNA interference (RNAi) is a potent antiviral pathway in plants and invertebrates that mediates the clearance of viral infection in a sequence specific manner. However, little is known about the specific genetic requirements of antiviral immunity in *C. elegans*. While several forward genetic screens have been conducted in *C. elegans* to identify the genetic components of RNAi, our screen is the first to isolate genes required specifically for antiviral RNAi. We developed a strain carrying a viral replicon, defective in the suppression of RNAi, which is rescued in *C. elegans* animals defective in antiviral RNAi. The 13 isolated mutants were defective in antiviral immunity against both the viral replicon and natural infection with Orsay virus. These 13 mutants included those defective in canonical RNAi, as well as those uniquely required for antiviral RNAi. Using whole genome resequencing, we identify four novel alleles of DRH-1, a homolog of the cytoplasmic viral receptor, RIG-I, as well as a novel allele of RDE-1, the argonaute protein required for canonical RNAi. Importantly, we identify a novel gene, MYS-2, required for antiviral immunity. Our results demonstrate the feasibility of a replicon-based screen to identify new genes required for antiviral immunity and effectively

isolated genes that are uniquely required for antiviral RNAi, as well as components of the core RNAi machinery.

2.2 Introduction

2.2.1 Forward Genetic Screens by Chemical Mutagenesis using EMS

Forward genetic screens continue to be a powerful and irreplaceable tool for piecing together the genetic components of cellular pathways. In particular, single nucleotide substitutions induced by chemical mutagenesis allows for isolation of a gradient of phenotypes and the identification of a variety of alleles within the same gene. Coupling the power of unbiased genetic screens with the genetic tractability of model organisms has led to ground breaking discoveries, such as identifying key components of Toll-like receptor signaling and the RNA interference (RNAi) pathways (1–6).

Exposure to Ethyl-MethaneSulfonate (EMS) induces single base transitions in the genome. These Single Nucleotide Polymorphisms (SNPs) are caused by the conversion of a guanine to O-6-ethylguanine by EMS. During DNA replication, O-6-methylguanine frequently pairs with thymine, resulting in either a guanine to adenine and cytosine to thymine transition, depending on the DNA strand. Mutations in coding regions or important regulatory regions can cause deleterious effects to the protein product and subsequently, to the organism. Only about 13% of EMS-induced mutations are insertions or deletions (7), making it an ideal agent to isolate a variety of subtle alleles.

Caenorhabditis elegans is particularly amenable to genetic analyses at the organismal level, because they have a short life cycle of only three days and a brood size

of around 300. Their hermaphroditic nature further simplifies the mutagenic process by eliminating the need for crosses. Hermaphroditic worms that are treated with EMS at a young adults stage have already produced and stored sperm, but eggs are still being produced in the ovary by meiosis and are the most susceptible to chemical mutagenesis. Regardless of which gamete is mutagenized, the F1 generation will be heterozygous for newly induced mutations and a desired phenotype can be screened in the F2 population (8). The time from EMS treatment to the isolation of potential mutant alleles can be as little as two weeks in *C. elegans*.

2.2.2 C. elegans as a model to study Antiviral RNAi

While *C. elegans* has long been established as a successful model to study a multitude of complex biological processes, it has only emerged recently as a model for virus pathogenesis and antiviral immunity. The first studies utilized transgenic and cell-based systems with viruses that do not naturally infect *C. elegans*, such as Flock house virus (FHV) (9–12). However, the complete infection cycle can now be investigated in *C. elegans* with the exciting recent discovery of the Orsay, Le Blanc and Santeill viruses that naturally infect *Ceanhorhabditis* species (13). Interestingly, these new viruses are most closely related to FHV and other members of the *Nodaviridae* (13).

C. elegans encodes conserved pathways for the biogenesis and function of the three classes of small silencing RNAs identified in eukaryotes: small interfering RNAs (siRNAs), microRNAs (miRNAs) and PIWI-interacting RNAs (piRNAs). Resembling mammals, there is a single Dicer gene (DCR-1) in *C. elegans* responsible for the

biogenesis of both siRNAs and miRNAs. Similar to plants, *C. elegans* has evolved a family of RNA-dependent RNA polymerases (RdRP) required for the amplification of siRNAs. In RNAi induced by exogenous long dsRNA (exo-RNAi), DCR-1 and the dsRNA-binding protein, RDE-4, produce primary siRNAs, which are predominantly 23 nucleotides (nt) in length with 2-nt overhangs at the 3' ends and contain 5'-monophosphates. These primary siRNAs are loaded in the Argonaute protein, RDE-1, to trigger the synthesis of secondary siRNAs by the RdRP RRF-1 using the target mRNA as a template in a process dependent on the Dicer-related helicase 3 (DRH-3) (14–18). Also referred to as 22G RNAs, secondary siRNAs are antisense to the target mRNA with a 5'-triphosphate and are predominantly 22-nt in length with a guanine as the 5' terminal nucleotide (15). In the final step of exo-RNAi, 22G RNAs are loaded into secondary argonaute proteins to guide target RNA clearance.

RNAi directs the principal antiviral mechanism in fungi, plants, insects and *C. elegans* (19) and recent studies indicate that antiviral RNAi is also active in mice (20, 21). Antiviral RNAi is initiated by the immune detection of viral dsRNA, leading to Dicer-dependent biogenesis of virus-derived siRNAs (vsiRNAs) to mediate specific virus resistance by RNAi. Viral secondary siRNAs essential for antiviral RNAi are produced in plants and *C. elegans* that encode RdRPs. As a counter defense, plant and animal viruses encode viral suppressors of RNAi (VSRs) to promote infection at the level of either single cells or whole organisms. Induction of the *C. elegans* antiviral RNAi was first detected in cultured primary cells and whole animals targeting the replication of Vesicular stomatitis virus and FHV, respectively (9–12, 22). Subsequent studies have

shown the same antiviral RNAi pathway is also induced by the natural infection of Orsay virus (OrV) (13, 23–25). Although OrV does not appear to encode a VSR, *C. elegans* antiviral RNAi is efficiently suppressed by the VSR B2 encoded by either FHV or Nodamura virus (9, 26). Antiviral RNAi in *C. elegans* is mechanistically similar to exo-RNAi since it requires DCR-1 and RDE-4 for the production of primary vsiRNAs from viral dsRNA replicative intermediates and RDE-1, RRF-1, and DRH-3 for the biogenesis of secondary vsiRNAs (22–24). In contrast to exogenous RNAi, however, DRH-1 is uniquely required for the biogenesis of primary vsiRNAs (22–24).

2.2.3 FHV-derived viral replicon in C. elegans

We have previously demonstrated that *C. elegans* supports efficient and complete replication of FHV launched from a heat-inducible transgene integrated into the *C. elegans* genome (9). FHV naturally infects insects and encodes an essential viral suppressor of RNAi (VSR), the B2 protein. B2 inhibits antiviral RNAi by binding viral dsRNA replicative intermediates and preventing them from being diced into vsiRNAs (9, 27). Consistent with studies in *Drosophila* (28), when the heat-inducible transgene contains a missense mutation in B2 (FR1ΔB2) virus accumulation is abolished in *C. elegans*. The FR1ΔB2 viral replicon can be rescued by introducing a mutation in the host genes, *rde-1* or *rde-4*, key components of exogenous RNAi (9). We adapted the FHV system to a feeding RNAi screen by replacing the B2 protein with eGFP (FR1gfp) and identified both DRH-1 and RSD-2 as components of antiviral RNAi (22). Similar to Drosophila cells (29), FR1gfp replication induces potent antiviral RNAi in *C. elegans* so

that in the absence of RNAi suppression by B2, eGFP expression is detectable only when the host antiviral RNAi is genetically inactivated.

With recent studies emerging to support an antiviral role for RNAi in mammals (20, 21), research of homologous proteins in such a tractable model system as *C. elegans* has limitless potential. Importantly, genetic studies in *C. elegans* are carried out in intact, living animals and in the case of our screen, viral replication can be visualized via eGFP without disrupting the animal. Additionally, the FR1gfp viral replicon is under the control of a ubiquitous promoter, allowing us to screen every cell type in the context of a living animal. Building upon the success of a targeted feeding RNAi screen, we adapted our system to a forward genetics screen using chemical mutagenesis, which has the potential to isolate subtle alleles and is more amenable to studying the RNAi pathway.

2.2.4 Mapping EMS induced mutations using whole genome resequencing

The advent of massively parallel sequencing platforms has revolutionized the methods available for mapping mutants obtained from forward genetics screens (30). In *C. elegans*, traditional mapping strategies relied on crosses to the well-characterized, Hawaiian isolate and required prior knowledge of SNPs (31). Comparison of a reference sequence to genome resequencing of a mutant genome resulted in too many SNPs to successfully map the allele without prior rough mapping to a 4.0Mb region using the Hawaiin strain (32). In 2010, two strategies were described that did not require a priori knowledge of a mapping strain. In the EMS-based approach, mutants are sequenced after several backcrosses to the parent strain and the density of EMS-induced damage across

the genome is used to map the location of a causal mutation (33). In the SNP-based approach, the mutant is crossed to the Hawaiin isolate, 20-50 mutant F2s are pooled together (bulk segregation) and the causal mutation is mapped using the recombination frequency of SNPs spanning the genome (34). While this study utilized the Hawaiin strain, any non-parental strain with a reference genome will suffice. Additionally, EMS-based and SNP-based approaches have been merged to produce Variant Discovery Mapping, which relies on bulk segregation to any non-parental strain and analyzes both naturally occurring SNPs from the parent strain and EMS-induced SNPs (35).

We found that worm genetic mutants defective in antiviral RNAi were readily obtained by EMS mutagenesis of FR1gfp worms. However, we were not able to clone the causal mutations using the classical map-based cloning approach (31), because of a distorted segregation of the antiviral RNAi phenotype in the progeny crossed with the Hawaiian mapping strain (our unpublished data). Therefore, we adopted a mapping approach recently developed in plants that identifies the causal mutations by whole-genome resequencing of pooled F2s from a backcross with the parental strain (36, 37). This method would utilize an EMS-based approach, combined with the bulk segregation utilized by Variant Discovery Mapping in *C. elegans*.

Since the monumental discovery of RNAi in *C. elegans* (38), several forward genetic screens have been performed to identify the genetic requirements (6, 39, 40). Our approach is unique in that it is the first with the ability to reveal genes that are required exclusively for antiviral RNAi, as well as potentially identify new antiviral pathways.

2.3 Materials and Methods

2.3.1 Worm Maintenance and Genetics

The *C. elegans* isolate, Bristol N2, was used as the wild-type strain in this work (8). Worms were maintained on standard Nematode Growth Medium (NGM) at 20°C, unless otherwise stated. Mutants referred to in the text refer to the alleles in Table 2-1. *C. elegans* media and buffers were prepared as described in Appendix A.

Table 2-1. ***C. elegans*** **strains by Chromosome**

Referenced in text	Genotype
N2	Bristol N2 strain (8)
LGI	
rrf-1	*rrf-1(pk1417)*
ucr7	*mys-2(ucr7); irSi18[$_p$hsp-16.41::FR1gfp::Rz]*
LGII	
FR1gfp	*irSi18[$_p$hsp-16.41::FR1gfp::Rz]*
LGIII	
rde-4	*rde-4(ne301); irSi18[$_p$hsp-16.41::FR1gfp::Rz]*
LGIV	
drh-1	*drh-1(tm1329); irSi18[$_p$hsp-16.41::FR1gfp::Rz]*
ucr2	*drh-1(ucr2); irSi18[$_p$hsp-16.41::FR1gfp::Rz]*
ucr3	*drh-1(ucr3); irSi18[$_p$hsp-16.41::FR1gfp::Rz]*
ucr4	*drh-1(ucr4); irSi18[$_p$hsp-16.41::FR1gfp::Rz]*
ucr6	*drh-1(ucr6); irSi18[$_p$hsp-16.41::FR1gfp::Rz]*
LGV	
rde-1	*rde-1(ne219)*
ucr11	*rde-1(ucr11)*
Unmapped	
*ucr5**	*ucr5; irSi18[$_p$hsp-16.41::FR1gfp::Rz]*
ucr8	*ucr8; irSi18[$_p$hsp-16.41::FR1gfp::Rz]*
ucr9	*ucr9; irSi18[$_p$hsp-16.41::FR1gfp::Rz]*
ucr10	*ucr10; irSi18[$_p$hsp-16.41::FR1gfp::Rz]*
ucr12	*ucr12; irSi18[$_p$hsp-16.41::FR1gfp::Rz]*
ucr13	*ucr13; irSi18[$_p$hsp-16.41::FR1gfp::Rz]*
ucr14	*ucr14; irSi18[$_p$hsp-16.41::FR1gfp::Rz]*

2.3.2 FR1gfp transgene Construct and MosSCI

The $_{p}$*hsp-16.41::FR1gfp::Rz* transgene used in this study is a derivative of the $_{p}$*hsp-16.41*::FR1gfp replicon previously described (9, 22), in which the self-cleaving ribozyme from the tobacco ringspot virus satellite RNA was replaced with the self-cleaving ribozyme from hepatitis delta virus (41), as described previously (42). MosSCI (43) was used to generate a single copy insertion of $_{p}$*hsp-16.41::FR1gfp::Rz*, called *irSi18* in the Maduro Lab at UCR. The MosSCI plasmid was injected into Strain EG4322, which contains the mosI insertion ttTi5605 on LGII and *unc-119(ed9)* III(44). PCR primers on either side of the insertion site and primers in the FR1gfp transgene were used to verify the presence of the transgene on LGII (Table 2-2). *irSi18* was crossed with *drh-1(tm1329)* and *rde-1(ne219)* to verify that viral replication, detected by GFP expression, was induced by heat shock. Northern Blot detection of sgRNA3 48 hours after heat shock confirmed that FR1gfp is capable of autonomous replication when launched from the *irSi18* transgene.

2.3.3 EMS Mutagenesis

N2;*irSi18* young adults were treated for 4 hours in 50mM EMS using standard mutagenesis techniques (45). After thirty minutes of recovery time, ten P_0 animals were singled onto large (100mm) plates. The following day, each P_0 animal was removed from the plates once they had each laid approximately ten eggs. When the F1 animals reached adulthood and had each laid an average of twenty eggs, the F1 adults were removed from the plate. The resultant 2,000 F2s were heat shocked at 34°C for four hours at the L4

stage. 24 hours and 48 hours after induction of FR1gfp, worms expressing any eGFP were singled to their own plate to establish independent lines. Each mutagenized strain was tested two more times for to ensure they failed to suppress FR1gfp.

2.3.4 Preparation of Whole Genome Libraries

100 F2s from a backcross of each mutant to N2;*irSi18* parent were transferred to individual plates and given time to lay eggs. Each F2 was used in single-worm DNA extractions to be genotyped later. FR1gfp replication was induced when the F3s reached the L4 stage and the segregation of the viral GFP expression phenotype was scored to determine if the parent F2 was homozygous wild-type, heterozygous or homozygous mutant. The F3 populations from 20 homozygous mutant F2s and 20 homozygous wild-type F2s were pooled and DNA was extracted using the Gentra Puregene kit (Qiagen). The DNA was sheared using a biorupter (30sec on, 30 sec off, for 15 minutes). To construct libraries for sequencing, 1 µg from each pool was used to generate a library using the PCR-free Truseq DNA kit (Illumina). The samples were multiplexed to have six libraries in each lane for pair-end sequencing of 100-nt reads and sequenced on an Illumina 2500. Mutant and wildtype libraries for *ucr2* were constructed using NEXTflex™ PCR-Free DNA Library Prep Kit for Illumina (Bioo Scientific), which gave lower genome coverage than the other mutants.

2.3.5 Mapping by whole genome re-sequencing and computational analyses.

Results from the whole-genome sequencing were analyzed using a custom pipeline to map the libraries and call the SNPs, coupled with a custom perl script to filter the SNPs based on user defined criteria. Fastq libraries were mapped to the *ce10* version of the *C. elegans* genome using Bowtie2.0, then SNPs were called using Samtools0.1.19. After mapping both the mutant and wild-type pools, the two resultant VCF files were compared to identify SNPs unique to the mutant pool. The candidate SNPs were then narrowed down to EMS induced mutations (G to A or C to T). Then, the EMS-induced mutations were filtered based on a user defined allele frequency for the mutant and wild-type pools, as well as a variance for both. Additional user defined parameters include depth of coverage and quality score.

2.3.6 Single worm DNA extraction and genotyping

Genotyping primers of EMS-induced SNPs were designed to utilize natural occurring restriction enzyme sites, such as *ucr7(mys-2),* when possible. Otherwise, mutagenic primers were designed using dCAP Finder 2.0 (46) to induce a restriction site in either the wild type or mutant PCR product. Single worm PCR was used to isolate single F2 worms for genotyping after they laid eggs. Single worms were places in PCR-tubes containing 10µL DNA extraction buffer with Proteinase K (Appendix A) and incubated at 65°C for one hour, followed by 95°C for ten minutes. Samples were diluted 2X and 2µL was used in a PCR reaction. 10µL PCR product was used in a restriction enzyme digest and incubated at 37°C for 1 hour or overnight. Digest products were

analyzed on 4% agarose gels for mutagenic primers and 1% argarose gels for regular primers. The *drh-1(ucr3)* SNP is linked to a EMS-induced 400bp deletion immediately upstream of the *drh-1* coding region.

Table 2-2. **Genotyping Primers**

	Primers	**Type**
irSi18	*irSi18*_ext_F: CTGAATATCCATGGCACATGCTTCGTGC *irSi18* ext_R: AGAAGACCGTTACGAAACAGACTGAT *irSi18*_int_F: CCGAAGTGCGGTGATGATGGTCTT *irSi18*_int_R: TCAGGGTGGTCACGAGGGTGGGCC	Normal
ucr2 DRH-1	*ucr2*_AseI_F: GTAAATGATCACACA**T**TAA *ucr2*_R: TTCGCTCACTCACGGGGAAAAGGCAA	Mutagenic AseI, Mt
ucr3 DRH-1	*ucr3*_F: ATCTGCAAATTTAGATTTTTGAGAAAAACTAAGAGAAA *ucr3*_R: CCAATTTAGTATTTGATAATTTCTACGAGGGTC	Regular
ucr4 DRH-1	*ucr4*_F: CCTTGTCTGTGATCCGGGATTCTG *ucr4*_NcoI_R: AAAGAAAACCCACAAAGAGCTGATAG**C**CATG	Mutagenic NcoI, Wt
ucr5 SUMV-2	*ucr5*_F: CAACAGAATTCGAGATTTCCGAAG *ucr5*_RsaI_R: ACTAGTGACAGTATGCCACTGGCATCC**G**T	Mutagenic RsaI, Mt
ucr6 DRH-1	*ucr6*_F: GTGGAAGAGCTCTCAACTCAGAAT *ucr6*_XhoI_R: TGCGCTTTATCGGTATCTTCTCTCAGGA**C**TC	Mutagenic XhoI, Wt
ucr7 MYS-2	*ucr7*_F: AGCATTTTTCTCGAAAAATCCCATG *ucr7*_R: GCTAAAAATTTGGCAATTTTTGAG	Regular HpaII,Wt
ucr10 MYS-2	*ucr10*_F: GTTACGAGCTATCCCGCCGCGAAG *ucr10*_BamHI_R: TTTCAGATTTTTTCGGCGAAAAACCGCG**G**A	Mutagenic BamHI, Mt

2.3.7 Preparation of Small RNA Libraries

Small RNAs were extracted using 4M LiCl followed by precipitation in 2 volumes ethanol. Up to 5µg of purified small RNAs was used in a RNA 5' Polyphosphatase (RPP) reaction to remove the 5'-triphosphate groups from the secondary small RNAs. The reactions were incubated for 30 minutes at 37°C, per the manufacture's method (Epicentre). RPP was removed from the sample by Trizol extraction and the small RNAs were re-suspended in 7µL of DEPC-H_2O. 6µL of the recovered, RPP-treated, purified small RNAs were used in the generation of small RNA libraries using

the Truseq Small RNA sample preparation kit (Illumina). Samples were multiplexed and sequenced on an Illumina 2500 for 50-nt reads.

2.3.8 Small RNA Analysis

Trimming and removal of structural small RNAs. Small RNA analysis was performed using Bowtie (0.12.7), custom Perl (5.16.3) scripts and Wormbase (WS240). Adapter sequences ('tggaattctcgggtgccaaggc') were removed from the 3' end of each read and filtered by size to obtain inserts 18-30nt long, using Skewer (47). Reads from each library that aligned with zero mismatches to the sense strand of structural small RNAs were removed (17); sequences for 140 snoRNAs, 631 tRNAs, 1 scRNA, 22 rRNAs and 114 snRNAs were extracted from WormBase and used in this analysis. The resultant nonstructural siRNA reads were used in the following analyses. The nonstructural small RNAs were aligned to *C. elegans* miRNAs obtained from WormBase (WS240) and used to normalize vsiRNA reads.

Analysis of vsiRNA profiles. Small RNA reads were aligned to the Orsay Virus genome (Genbank IDs: HM030970.2 and HM030971.2) using Bowtie, allowing zero mismatches. Aligned reads were then parsed by size, polarity, and 5' nucleotide identity using a custom perl script.

2.3.9 Feeding RNAi

Feeding RNAi clones for *dpy-7*, *unc-22*, *pop-1* and *pos-1* were obtained from the *C. elegans* ORF-RNAi library (48) (Source BioScience). As a negative control, an empty

L4440 vector was transformed into the HT115 strain of *E. coli*. NGM plates with IPTG and carbenicillin (Appendix A) were seeded with the feeding RNAi clones. For *dpy-7* and *unc-22* feeding RNAi, three gravid adults were placed on feeding RNAi plates to lay eggs and the next generation was grown at 20°C until they reached the L4 stage, about two days. At the L4 stage, each animal was scored for either a dumpy or unc phenotype. For *pop-1* and *pos-1,* three L4 worms were placed on feeding RNAi plates and grown at 20°C to lay eggs. The number of viable L1s were scored and compared with the number of viable L1s obtained from three L4s plated on vector only control plates over the same time period.

2.3.10 Viral RNA detection by Northern Blot

Total RNA was extracted from homogenized worms using Trizol (Sigma-Aldrich) by following the manufacturer's protocol, with the addition of an extra chloroform separation with a 1:1 ratio before precipitation. 4µg of high molecular weight RNA was ran on a 1.2% agarose gel with Formaldehyde. [α-^{32}P]dCTP labeled cDNA corresponding to OrV RNA 1 was used to detect OrV infection, and a PCR fragment of eGFP was used to detect both RNA1 and sgRNA3 of FR1gfp. Both probes were labeled with [α-^{32}P]dCTP using the Ready to Go DNA Labelling kit (-dCTP) (Amersham Biosciences). Hybridization of both probes was done at 65°C for three hours. Buffers and media used for Nothern Blotting can be found in Appendix B.

2.3.11 FR1gfp Induction and OrV infection

For detection of FR1gfp, animals were synchronized and grown at 20°C for two days, until they reached the L4 stage. L4 staged animals were heat shocked at 34°C for 4 hours and then returned to 20°C. GFP was analyzed under a dissecting microscope 48 hours later and collected for RNA extractions. For OrV infections, 10-30 young adult worms were transferred to new 60mm NGM plates seeded with OP50 *E. coli*. 10 young adults were sufficient for the N2 strain; however more worms were used for difference mutants, due to a smaller brood size. 50µL of OrV filtrate was added to each plate and worms were incubated at 20°C. Five days post infection, the next generation of worms were collected for RNA extractions. OrV was originally provided by the labs of David Wang and Marie-Ann Felix. OrV filtrates were prepared by growing infected *drh-1(tm1329)* animals on ten 100mm plates. The mixed stage worms, mostly adults, were washed off in 3mL of M9 buffer into two 1.5mL ependorf tubes. The infected worms were homogenized and centrifuged for 1 minute at 13,000rpm. The supernatant was then filtered through a 20µm filter. The filtrates from each plate were mixed together and aliquoted into 1.5mL ependorf tubes to be used in experiments.

2.3.12 Transgenic Rescue of DRH-1 and MYS-2

ucr2, ucr3, ucr4 and *ucr6* were injected with a *psur-5::DRH-1* plasmid previously described (24), along with *myo-2::mCherry* using a standard microinjection protocol. *mys-2* cDNA was amplified by RT-PCR and cloned into the pd51 plasmid, between the *sur-5* promoter and *unc-54* 3'-UTR. *ucr7* animals were injected with *psur-5::MYS-2* and

myo-2::mCherry. F1s expressing mCherry in the pharynx were selected and F2s expressing mCherry were isolated to establish individual lines carrying the rescue construct. For analysis by Northern Blot, ten animals expressing mCherry in the pharynx were selected and their progeny were heat shocked at the L4 stage to induce replication of FR1gfp. 48 hours later, worms were imaged using a fluorescent microscope and then collected for RNA extractions.

The DNA injection mix for DRH-1 rescue experiments had a concentration of 200ng/µL, composed of 20ng/µL *psur-5::DRH-1* and 180ng/µL *myo-2::mCherry*. However, for the MYS-2 rescue initial attempts with the same concentrations (20ng/µL *psur-5::MYS-2* and 180µg/uL *myo-2::mCherry*) were unsuccessful and F1s expressing mCherry in the pharynx were not viable. Adjusting the DNA injection mix to a concentration of 5ng/µL *psur-5::MYS-2*, 2.5ng/µL *myo-2::mCherry* and 192.5ng/µL *pd51* resulted in viable F1s expressing mCherry in the pharynx and gave rise to several F2 lines carrying the extrachromosomal array.

2.3.13 Genetic Complementation

For genetic complementation, about twenty *rde-1(ne219);irSi18* males and five late L4 stage hermaphrodites (*ucr* mutant) were placed on a single 35mm plates. After two days of mating at 15°C, each hermaphrodite was singled to its own 35mm plate and left to lay eggs overnight. The hermaphrodite was removed from the plate when there were 50-100 F1s (mix of eggs and L1s). If approximately 50% of the F1s were males, then the mating was considered successful and those F1s were subjected to heat treatment

to induce FR1gfp replication when they reached the L4 stage. eGFP expression from FR1gfp was scored 48 hours later in the F1s and in wild-type and mutant worms.

2.4 Results

2.4.1 EMS mutagenesis of a C. elegans strain exhibiting antiviral RNAi

To identify mutants defective in RNAi-mediated antiviral immunity, we adapted the FR1gfp replicon previously described (22), to a forward genetic screen. Our original heat-inducible FR1gfp viral transgene had a ribozyme derived from a plant satellite virus at the 3'-end of the replicon and was integrated into the *C. elegans* genome as a multicopy array (22). We found that suppression of FR1gfp in wild-type animals was incomplete, complicating the selection of genuine mutant alleles. We replaced the plant ribozyme with the antigenomic ribozyme of Hepatitis delta virus (41) as described previously (42) and constructed a single copy insertion on chromosome II, called *irSi18[$_p$hsp-16.41::FR1gfp::Rz]* (Fig 2-1A). None of the *irSi18* animals in wild type N2 background exhibited any visible eGFP expression after heat treatment and deep sequencing identified an abundant population of the canonical vsiRNAs in response to FR1gfp replication in the N2;*irSi18* animals (Fig 3-6A). In contrast, we observed GFP expression in *C. elegans* mutants defective in antiviral RNAi after heat induction of the transgene, either in the whole body (e.g. *drh-1(tm1329)*) or in the pharynx (e.g. *rde-1(ne219)*). Northern blot analysis further revealed that while both RNA1 and RNA3 of FR1gfp accumulated to higher levels in *rde-1* and *drh-1* animals, while neither viral RNA was detectable in N2 animals (Fig 2-1C). These results indicate that eGFP expression

from the heat-inducible *irSi18* transgene was potently inhibited by antiviral RNAi, establishing the N2;*irSi18* strain as a model for the genetic characterization of the antiviral RNAi pathway.

We subjected N2;*irSi18* young adults to standard EMS-induced mutagenesis (45) and screened approximately 28,400 haploid genomes (Appendix C). When the mutagenized F2 generation reached the L4 stage, we induced replication of FR1gfp by heat treatment and 48 hours later screened for eGFP expression. Single worms expressing eGFP were transferred to individual plates to establish independent F2 mutant lines.

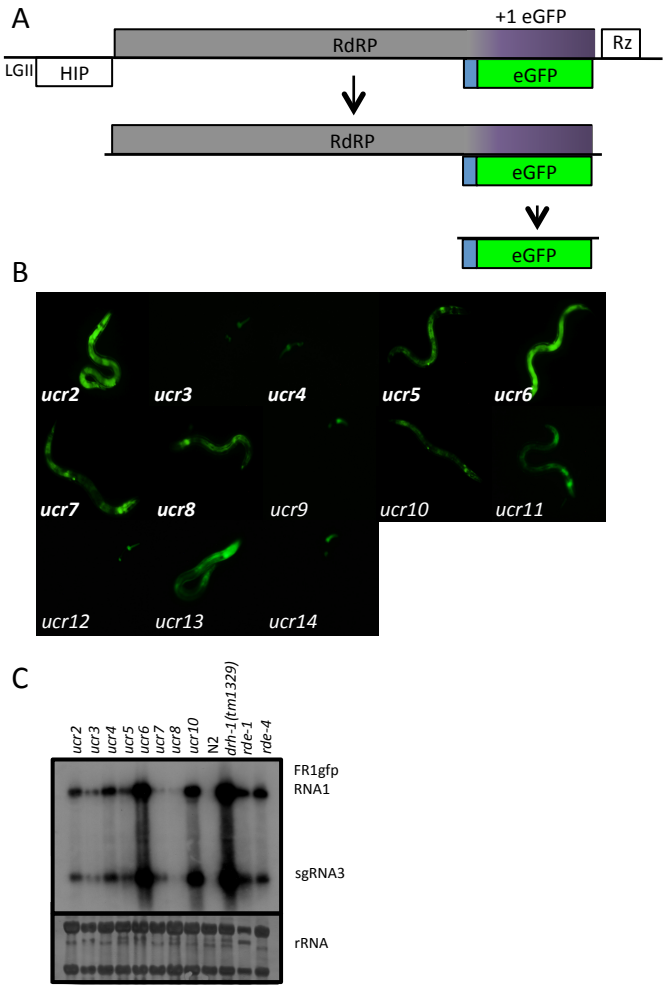

Figure 2-1. Isolation of thirteen mutants defective in antiviral immunity. (A) Schematic of the FR1gfp viral replicon launched from the *irSi18* transgene. The C-terminal region of the viral RNA-dependent RNA polymerase (RdRP) is effectively truncated and replaced by the -1 reading frame of the eGFP coding sequence. Rz: HDV ribozyme, HIP: Heat inducible promoter. (B) Synchronized L1 wildtype (N2) and mutant animals were plated and heat treated for 4 hours at 34°C at the L4 stage. GFP was visualized by fluorescent microscopy 48 hours after heat induction. (C) Northern blot detection of the accumulation of FR1gfp RNA1 and RNA3 in mutants proficient in RNAi.

Although most of the identified mutant animals were not viable, we isolated 13 viable mutants that reproducibly failed to suppress GFP expression from FR1gfp in subsequent generations. These mutant animals, designated *ucr2* to *ucr14*, displayed distinct patterns of GFP expression. For example, *ucr3* and *ucr4* expressed eGFP mostly in the pharynx, while most *ucr6* animals exhibited eGFP in the whole body (Fig 2-1B). Furthermore, we found that the phenotypic penetrance of eGFP expression was less than 100% in all of the isolated mutants.

2.4.2 Characterization of exogenous and antiviral RNAi in isolated mutants

To characterize the identified mutants, we first determined if they exhibited an altered susceptibility to exo-RNAi targeting either somatic genes (*dpy-7* and *unc-22*) or germline genes (*pop-1* and *pos-1*). Most of these mutants remained proficient to support RNAi against the somatic and germline genes by feeding the worms with *E. coli* expressing gene-specific dsRNA (Fig 2-2A). We found that five of the 13 mutants were resistant to exo-RNAi against both of the germline genes (Fig 2-2A). However, only four of these five mutants, *ucr9*, *ucr11*, *ucr13* and *ucr14*, were also as resistant as *rde-1* animals to the exo-RNAi targeting both of the somatic genes (Fig 2-2A). *ucr12* was proficient in somatic RNAi but defective in germline RNAi, similar to genes required for dosage dependent RNAi or systemic RNAi (39, 40, 49, 50). These results divided the genetic mutants recovered from our screen into distinct groups according to their susceptibility to exo-RNAi.

We next determined the susceptibility of the 13 mutants to infection by OrV. Wildtype N2 animals are resistant to OrV infection (13) so that the accumulation of OrV RNAs is not readily detectable by Northern blotting (Fig 2-2B). All of the five mutants defective in exo-RNAi exhibited an enhanced susceptibility to OrV (Fig 2-2B), providing independent genetic evidence for a shared pathway between exo- and antiviral RNAi in *C. elegans*. Notably, we found that all of the eight mutants proficient in exo-RNAi were also more susceptible to OrV and reproducibly supported high level replication of OrV

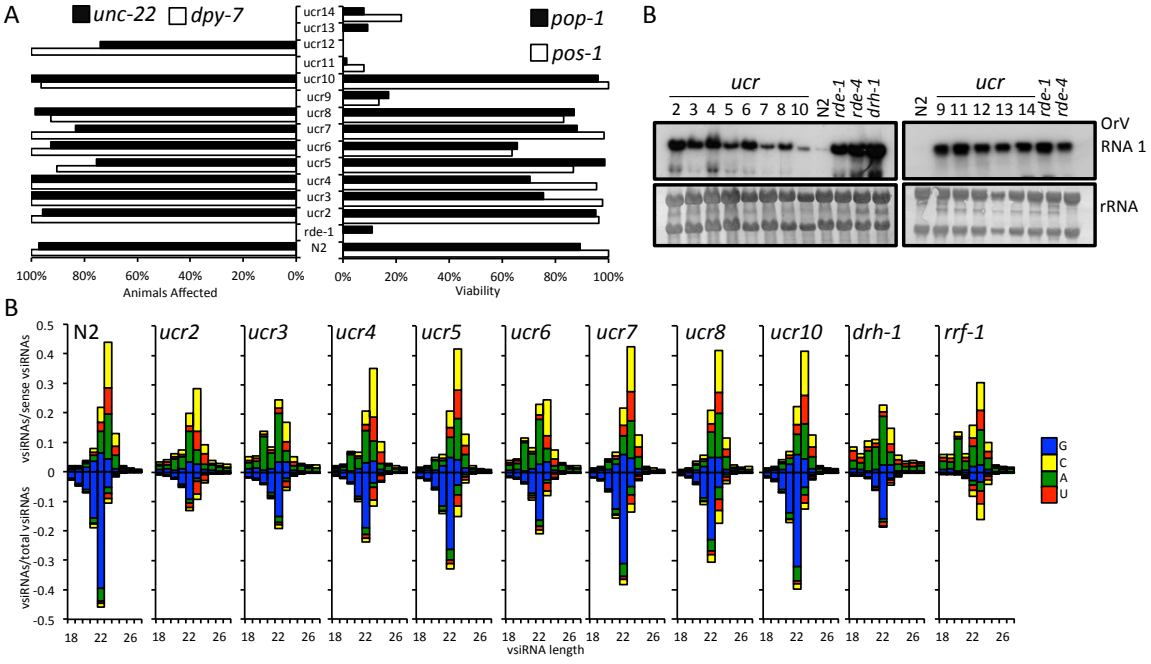

Figure 2-2. Characterization of exo- and antiviral RNAi in mutants. (*A*) Susceptibility of the isolated mutants to exogenous RNAi to target the somatic gene *dpy-7* (left) or the germline gene *pop-1* (right). Shown are the percentage of dumpy or viable animals after RNAi of *dpy-7* and *pop-1*, respectively. (*B*) Northern blot detection of OrV RNA1 accumulation in N2 and the isolated mutants (designated *ucr*) grouped according to their susceptibility (left) or resistance (right) to exogenous RNAi, using known mutants as controls. (*C*) Size distribution, polarity, and the 5'-terminal nucleotide of virus-derived small RNAs sequenced from N2 and mutant worms 5 day after infection with Orsay virus. The prevalence of the antisense vsiRNAs (bottom) is shown as the proportion of total vsiRNAs whereas that of sense vsiRNAs (top) is presented as the proportion of total sense vsiRNAs.

(Fig 2-2B). These findings demonstrate the feasibility to discover genes involved in host antiviral defense by the replicon-based genetic screen.

Among the worm genes identified in the antiviral RNAi pathway, only *drh-1* is dispensable for exo-RNAi (22). Thus, we sequenced the total small RNAs of the eight mutants proficient in exo-RNAi after OrV infection to determine if they included mutants defective in antiviral RNAi. The small RNA libraries were constructed by a protocol that clones both the primary and secondary vsiRNAs containing 5' mono- and tri-phosphates, respectively (23). As controls, we also sequenced the small RNAs from N2, *rrf-1* and *drh-1(tm1329)* worms infected with OrV. Although the vsiRNAs in the size range of 21- to 24-nt were detectable in all of the mutants, the prevalence of both the primary vsiRNAs (23-nt) and secondary vsiRNAs (antisense 22G-RNAs) was markedly reduced in *ucr2*, *ucr3*, *ucr4* and *ucr6* mutants compared to the remaining mutants or N2 worms (Fig 2-2C). The vsiRNA profiles of *ucr2*, *ucr3*, *ucr4* and *ucr6* mutants were similar to that found in *drh-1* mutant animals, indicating that these mutants exhibited specific defects in antiviral RNAi.

These results demonstrate the feasibility to discover the genes involved in host antiviral defense by a replicon-based genetic screen. Notably, the genetic mutants recovered from our screen included both those required for core RNAi and uniquely for antiviral RNAi.

2.4.3 Identification of four novel alleles of drh-1 by whole genome resequencing

We chose to map the causal mutations in *ucr2*, *ucr3*, *ucr4* and *ucr6* mutants defective in antiviral RNAi but proficient in exo-RNAi. Selection of the four RNAi-proficient mutants for mapping would avoid these mutants characterized extensively in previous RNAi screens (5, 6, 39). However, several attempts to map the *ucr* alleles by the classical map-based cloning approach were unsuccessful (31). Unlike the backcross with the parental N2 strain, outcrossing with Hawaiian isolate for genetic mapping gave rise to a distorted segregation of the viral eGFP expression phenotype. Therefore, we adopted a mapping strategy developed recently in plants that identifies the causal mutations by whole-genome resequencing of pooled F2 progenies from backcrosses with the parental strain (37).

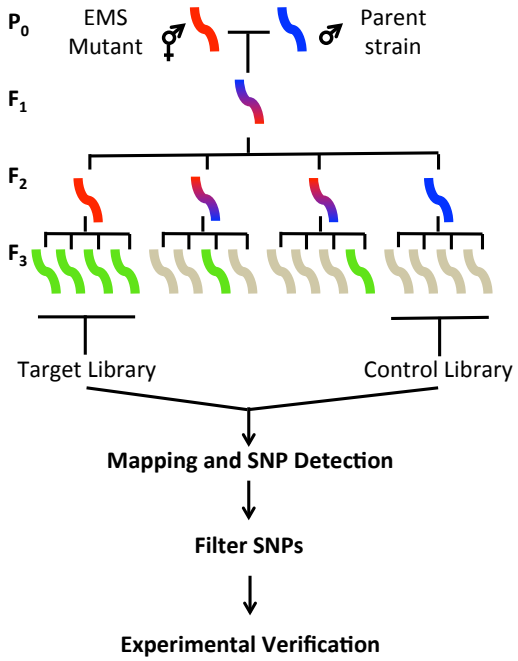

Figure 2-3. Schematic of the mapping strategy used to identify the causal, EMS-induced mutations.

Briefly, we backcrossed each mutant to the parental strain, N2;*irSi18,* and identified 20 F2 animals homozygous for the wildtype or mutant allele by scoring the segregation of the eGFP phenotype in the F3 generation (Fig 2-3). The DNA from each group of animals were pooled for genome resequencing with at least 20x coverage. The causal mutation was identified by interrogating the whole-genome resequencing data from each pair of the wildtype and mutant pools using a custom pipeline developed essentially as described (37). We narrowed down the candidate SNPs unique to the mutant pool by first identifying only those mutations that could be EMS-induced (G→A or C→T) before filtering the remaining SNPs with a user defined allele frequency.

To our surprise, the final candidate SNPs for each of these four worm mutants included one in the *drh-1* gene on chromosome IV (Appendix E). *ucr3* and *ucr4* both contain missense mutations, while *ucr6* contains a nonsense mutation and *ucr2* contains a splice site mutation. Together, these alleles represent an array of consequences to the DRH-1 protein, which will be discussed in Chapter 3, along with further analysis of the function of DRH-1 in antiviral RNAi.

2.4.4 Rescue of the mutant alleles by ectopic expression of wild type DRH-1

To determine if the mutant phenotypes in these four mutants were indeed caused by the SNPs identified in *drh-1*, we injected the gonads of the mutant worms with a *DRH-1* transgene along with a *mCherry* marker transgene as previously described (Fig 2-4A) (24). FR1gfp replication and the marker transgene expression were examined by fluorescence microscopy in the progeny of the injected mutant animals, some of which

did not inherit the extrachromosomal *mCherry/DRH-1* transgene array. For all four worm-mutants, we detected eGFP expression only in animals that lacked red fluorescence from the transgene array (Fig 2-4C). In contrast, all of the mutant animals that carried the *mCherry/DRH-1* transgene array and expressed red fluorescence in the pharynx tissue did not exhibit green fluorescence from FRgfp (Fig 2-4C). Therefore, the viral eGFP expression was detectable only in those mutant animals that did not carry the *DRH-1* transgene whereas transgenic expression of DRH-1 restored antiviral RNAi, resulting in the suppression of the viral eGFP expression in all of the four worm mutants.

We further isolated 10 animals that carried the transgene array from each mutant and determined the accumulation of FR1gfp in the next generation following heat treatment to induce FR1gfp replication. In all four mutants, FR1gfp replicated to drastically reduced levels following the ectopic expression of the *DRH-1* transgene (Fig

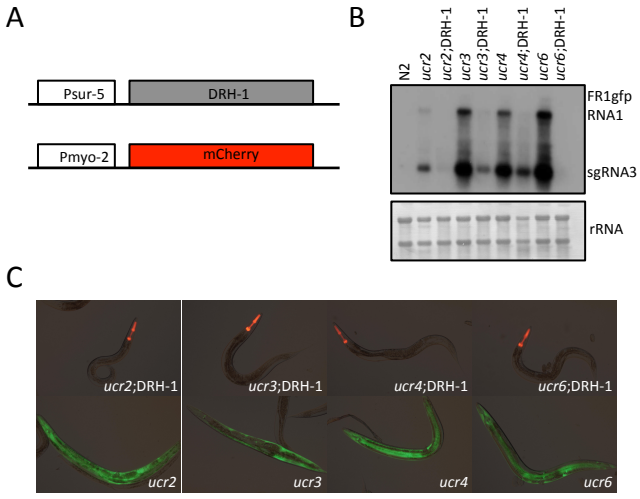

Figure 2-4. Transgenic rescue of *ucr2*, *ucr3*, *ucr4* and *ucr6*. *(A)* Schematic of the rescue transgene carrying wildtype DRH-1 and the co-injected marker transgene. *(B)* Select images overlaid with a red and green exposure from fluorescence microscopy of the rescued animals carrying DRH-1 transgene marked by the red fluorescence from the marker mCherry (top) and the mutant animals before the injection with the rescue constructs (bottom). Note that strong green fluorescence from FR1gfp replication in the mutant worms (bottom) disappeared following the introduction of the DRH-1 transgene and exhibited only red fluorescence (top).

2-4B). Therefore, we conclude that the increased virus susceptibility in *ucr2*, *ucr3*, *ucr4* and *ucr6* is caused by distinct genetic defects in DRH-1. Identification of multiple alleles

in the *drh-1* gene from our unbiased genetic screen further supports an indispensable role of DRH-1 in antiviral RNAi in *C. elegans* revealed by previous studies (22–24).

The computational identification of the candidate SNPs in our approach used a relaxed allele frequency for the wildtype (0-40%) and mutant (60-100%) pools. In contrast, the mutant allele frequency was predicted as 0 for the wild-type pool and 100% for the mutant pool because the two phenotypically distinct pools of F3 families were selected for genome resequencing. Since the causal mutation was identified in each of the four worm mutants, it was possible to compare the frequency of the specific alleles obtained from sequencing the F3 families pooled by their phenotypes with those calculated by genotyping. To this end, we individually genotyped the original F2s that made up the pools for *ucr4* and *ucr6*. The frequency of the specific *drh-1* mutant allele was 0 in the sequenced wildtype pool of both *ucr4* and *ucr6* mutants. The twenty F2s in the wildtype pool of both mutants also were all genotyped as homozygotes for the wildtype allele. However, the frequency of the specific *drh-1* mutant allele obtained by sequencing the mutant pool of *ucr4* and *ucr6* mutants was 83.7% and 75%, respectively, both of which were lower than the predicted value. Consistently, genotyping revealed that the twenty F3 families in the mutant pool of either *ucr4* or *ucr6* included 17 homozygotes and 3 heterozygotes for the respective mutant allele, yielding a mutant allele frequency of 92.5% for both mutants. These results indicate that our use of a relaxed allele frequency will facilitate the identification of candidate SNPs in genetic screens in which the mutant phenotype cannot be identified with 100% accuracy.

Table 2-3. **Comparison of Allele Frequencies**

	Whole Genome Sequencing		Individual Genotyping	
	WT (C)	**MT (T)**	**WT (C)**	**MT (T)**
drh-1(ucr4) MT	16.3%	83.7%	7.5%	92.5%
drh-1(ucr4) WT	100%	0	100%	0
drh-1(ucr6) MT	25%	75%	7.5%	92.5%
drh-1(ucr6) WT	100%	0	100%	0

2.4.5 Identification of novel RDE-1 allele by genetic complementation

The function of the argonaute protein, RDE-1, is well documented in antiviral RNAi and is completely resistant to exo-RNAi (6, 22). Additionally, *rde-1* is a fairly large gene, 3.2kb of coding sequence that when mutated gives rise to healthy animals. For these two reasons, *rde-1* alleles are commonly recovered in EMS screens. In order to determine if any of the RNAi defective alleles were caused by a defect in RDE-1, we crossed *rde-1(ne219)* males with hermaphrodites from *ucr9*, *ucr11*, *ucr13* and *ucr14* and heat shocked the F1 generation to induce FR1gfp replication. *ucr9, ucr13* and *ucr14* all produced F1 progeny that were successfully complemented (Fig 3-5B). In contrast, *ucr11,* failed to compliment and the F1s failed to suppress FR1gfp replication (Fig 3-5B). *ucr11* had a weaker eGFP expression phenotype than *rde-1(ne219)*, and the recovered

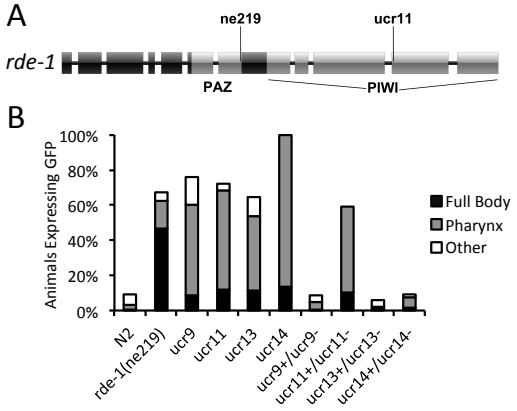

Figure 2-5. Identification of a novel *rde-1* allele by genetic complementation. (A) Schematic of the *rde-1* gene structure with the identified *ucr11* allele indicated. (B) F1 animals that were heterozygous for *rde-1(ne219)* and four *ucr* alleles were heat treated to induce FR1gfp replication. F1 animals, as well as homozygous animals were scored for GFP expression in the full body (black), pharynx (gray) and other (white). Other GFP expression includes specs of GFP in the pharynx as well as specs located in other areas of the body.

F1s exhibited a phenotype comparable to *ucr11*. Based on these results, we sequenced the *rde-1 mRNA* from *ucr11* and discovered an EMS-induced nonsense mutation that changes Arg^{800} to a stop codon, effectively truncating 220 amino acids from the C-terminal (Fig 3-5A). This truncation would delete a third of the PIWI domain, essential for anchoring the 5' end of the small RNA and the catalytic slicing activity. The catalytic activity of RDE-1 is required for the removal of the passenger strand (51).

2.4.6 Identification of a novel gene required for antiviral immunity

Wild type and mutant pools were also sequenced for *ucr5*, *ucr7* and *ucr10* using the method described. Consistent with their small RNA profiles (Fig 2-2C), none of these alleles mapped to *drh-1*, despite being susceptible to exogenous RNAi. *ucr7* was strongly linked to LG I and returned only one candidate, a missense mutation that changes $Glycine^{246}$ to a Glutamic Acid in MYS-2, a histone acteyltransferase (Appendix E). Interestingly, *ucr10* was also strongly linked to LG I and contained three candidates, two missense mutations and a splice site mutation at the 5' splice site in MYS-2 (Appendix I). Analysis of the *ucr5* libraries also resulted in three candidate SNPs, one missense and two nonsense mutations (Appendix I). One of the nonsense mutations was in SUMV-2, a subunit of the MYS-2 histone actyltransferase. Interestingly, knockdown of both MYS-2 and SUMV-2 cause loss of heritable RNAi and transgene silencing, respectively (52, 53). Additionally, two mutants isolated by Dr. Xunyang Guo mapped to distinct nonsense alleles in SUMV-2 (Appendix I). Based on the strong correlation between these five mutants, we decided to investigate MYS-2 as a potential candidate.

Previously we mapped four alleles of *drh-1*; however *drh-1* was obvious to choose from a list of candidates. Therefore, we sought to identify a novel gene required for antiviral RNAi.

In order to determine if the causal mutation in *ucr7* was caused by the SNP in *mys-2*, we constructed a plasmid expressing wild-type MYS-2 under the control of the ubiquitous *sur-5* promoter. Several attempts to amplify the *mys-2* cDNA from independent RNA samples revealed a 51-nt deletion in the *mys-2 mRNA* sequence compared to the reference sequence (WormBase WS249). *mys-2 mRNA* without these 51-nts would cause an in frame deletion of 17 amino acids and thus not severely affect the predicted protein product. The experimentally verified sequence of *mys-2 mRNA* is provided in Appendix G.

We injected the gonads of *ucr7* worms with *psur5::MYS-2* and *pmyo-2::mCherry* to establish transgenic lines ectopically expressing wild-type MYS-2, marked by

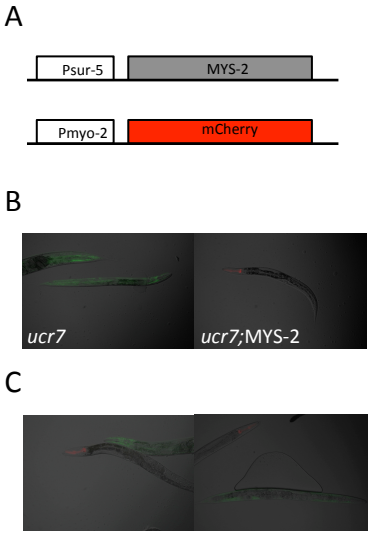

Figure 2-6. Identification of a novel gene required for antiviral Immunity. (A) Schematic of the rescue transgene carrying wildtype MYS-2 and the co-injected marker transgene. (B) Select images overlaid with a red and green exposure from fluorescence microscopy of the rescued animals carrying MYS-2 transgene marked by the red fluorescence from the marker mCherry (right) and the mutant animals that aren't carrying the rescue construct (left). (C) Select images of worms carrying the MYS-2 rescue construct (red) and worms supporting FR1gfp replication (green).

mCherry expression in the pharynx (Fig 2-6A). Several lines were established and heat shocked at the L4 stage to induce FR1gfp replication. The extrachromosomal array carrying *MYS-2/mCherry* was inherited by approximately 60% of the progeny. We used fluorescent microscopy to examine the expression of GFP from FR1gfp and mCherry from the marker transgene. In worms that expressed GFP from FR1gfp, we did not observe any instances of *MYS-2/mCherry* expression (Fig 2-6B-C). Additionally, in worms carrying the *MYS-2/mCherry* transgene, we did not observe any eGFP from FR1gfp replication. These results indicate that the susceptibility to viral replication and infection in *ucr7* is dependent on the genetic lesion in *mys-2*.

2.4.7 Mapping of mutant alleles without a wild-type population

We next investigated whether it is prudent to have both a mutant and wild type library for each allele. We analyzed the *ucr7* mutant library again using the *ucr5* mutant library as a control and obtained nearly identical results (Appendix H). The causal mutation of *ucr7* was experimentally verified to be located on LGI and both methods returned seven EMS-induced SNPs located on LGI, including the causal mutation in *mys-2*. Using the *ucr5* library did return five additional candidates randomly distributed across the other chromosomes; however these are easily identified as unlinked. One caveat to this method would be if two identical alleles are independently isolated, but comparing the mutant library to more than one control library easily rectifies this dilemma. Therefore, we conclude that a wild-type library is not necessary to successfully

map an EMS-induced mutation, similar to existing methods that rely on out-crosses to a non-parental strain (32, 34, 35).

2.5 Discussion

In this study, we developed a small animal model for a forward genetic screen by chemical mutagenesis to identify genes required for antiviral RNAi in *C. elegans*. We isolated 13 viable mutants that markedly enhanced the *in vivo* accumulation the FR1gfp viral replicon. Interestingly, all of the isolated mutants also exhibited enhanced susceptibility to natural infection by OrV, demonstrating the potential of the FHV replicon approach to studying antiviral immunity in *C. elegans.* Complications with the traditional mapping strain as well as an incomplete phenotypic penetrance complicated early attempts to establish a forward genetic screen using FR1gfp in *C. elegans*.

We show that EMS-induced mutations can be mapped and cloned by genome resequencing of pooled F2 progeny from backcrosses with the parental strain. EMS-based mapping has been effective in identifying causal EMS induced mutations after a series of backcrosses (33). In contrast, we removed the initial series of backcrosses, which saves the researcher time and continuing backcrosses can be preformed while the samples are being sequenced or can wait until the researcher has identified desirable alleles. Importantly, our approach was successful even when the phenotype under observation is not 100% penetrant. However, sequencing a wild-type pool is costly. In our case, we were able to obtain nearly identical results comparing two different mutant pools. Here we demonstrate an effective method to add to the EMS mapping tool box, which is

especially desirable in cases where crosses to a parental strain is warranted, phenotypic penetrance is variable and time is limiting.

In total, we sequenced mutant and wild-type pools from 7 mutants. *ucr2*, *ucr3*, *ucr4* and *ucr6* were all found to be capable of exo-RNAi and displayed a defect in antiviral RNAi by small RNA sequencing. EMS-induced mutations were identified and verified in *drh-1* for all four mutants. *ucr2* causes a 30 amino acid deletion in the Hel2i, while *ucr4* and *ucr6* affect the RD domain. Both the Hel2i and RD domain are required for binding dsRNA. The *ucr2* allele should also be defective in ATPase activity. These four alleles are discussed more in Chapter 3. Since DRH-1 was previously identified be a feeding RNAi screen as a component of antiviral RNAi (22), we validated out mapping method by mapping a new gene required for antiviral immunity. To this end, a fifth allele, *ucr7*, was mapped to a SNP in *mys-2*. *ucr5* and *ucr10* are yet to be mapped and the candidates obtained from our mapping method can be found in Appendix I. The remaining *ucr* alleles were backcrossed to the parental strain, but only the mutant library was sequenced. These candidates, as well as candidates to other *ucr* alleles obtained by Dr. Xunyang Guo can be found in Appendix I.

The histone acetyltransferase, MYS-2, is homologous to the Lysine(K) Acetyl Transferase 8 (KAT8), and is required for the H4K16ac associated with active transcription. KAT8 can also acetylate other proteins, though to a lesser extent. MYS-2 is the most extensively studied as a suppressor of the multi-vulva phenotype, which is regulated by a host of chromatin remodeling genes. Interestingly, MYS-2 was also identified in a feeding RNAi screen to identified genes required for inheritance of RNAi

(52); however, the small RNA profiles from *mys-2(ucr7)* do not suggest an obvious defect in vsiRNAs. Additionally, *mys-2(ucr7)* does not host as robust an OrV infection as animals defective in antiviral RNAi and may represent an alternative antiviral pathway. Because MYS-2 activates transcription, it may not directly function in antiviral immunity and the antiviral phenotype of *mys-2(ucr7)* result of indirect inactivation of innate immunity genes. Nevertheless, study of the genes regulated by MYS-2 will reveal genes that are required for antiviral immunity.

2.6 References

1. Anderson K V, Jürgens G, Nüsslein-Volhard C (1985) Establishment of dorsal-ventral polarity in the Drosophila embryo: genetic studies on the role of the Toll gene product. *Cell* 42(3):779–789.

2. Lemaitre B, Nicolas E, Michaut L, Reichhart JM, Hoffmann JA (1996) The dorsoventral regulatory gene cassette spatzle/Toll/Cactus controls the potent antifungal response in Drosophila adults. *Cell* 86(6):973–983.

3. Poltorak A, et al. (1998) Defective LPS signaling in C3H/HeJ and C57BL/10ScCr mice: mutations in Tlr4 gene. *Science* 282(5396):2085–2088.

4. Cogoni C, Macino G (1999) Gene silencing in Neurospora crassa requires a protein homologous to RNA-dependent RNA polymerase. *Nature* 399(6732):166–169.

5. Tabara H, Yigit E, Siomi H, Mello CC (2002) The dsRNA binding protein RDE-4 interacts with RDE-1, DCR-1, and a DExH-Box helicase to direct RNAi in C. elegans. *Cell* 109(7):861–871.

6. Tabara H, et al. (1999) The rde-1 gene, RNA interference, and transposon silencing in C. elegans. *Cell* 99(2):123–132.

7. Anderson P (1995) Mutagenesis. *Methods Cell Biol* 48:31–58.

8. Brenner S (1974) The genetics of Caenorhabditis elegans. *Genetics* 77(1):71–94.

9. Lu R, et al. (2005) Animal virus replication and RNAi-mediated antiviral silencing in Caenorhabditis elegans. *Nature* 436(7053):1040–1043.

10. Schott DH, Cureton DK, Whelan SP, Hunter CP (2005) An antiviral role for the RNA interference machinery in Caenorhabditis elegans. *Proc Natl Acad Sci U S A* 102(51):18420–18424.

11. Wilkins C, et al. (2005) RNA interference is an antiviral defence mechanism in Caenorhabditis elegans. *Nature* 436(7053):1044–1047.

12. Liu W-H, et al. (2006) Restriction of vaccinia virus replication by a ced-3 and ced-4-dependent pathway in Caenorhabditis elegans. *Proc Natl Acad Sci U S A* 103(11):4174–4179.

13. Félix MA, et al. (2011) Natural and experimental infection of Caenorhabditis nematodes by novel viruses related to nodaviruses. *PLoS Biol* 9(1). doi:10.1371/journal.pbio.1000586.

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

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

16. Nakamura M, et al. (2007) Dicer-related drh-3 gene functions in germ-line development by maintenance of chromosomal integrity in Caenorhabditis elegans. *Genes to Cells* 12(9):997–1010.

17. Gu W, et al. (2009) Distinct Argonaute-Mediated 22G-RNA Pathways Direct Genome Surveillance in the C. elegans Germline. *Mol Cell* 36(2):231–244.

18. 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.

19. Ding S-W (2010) RNA-based antiviral immunity. *Nat Rev Immunol* 10(9):632–644.

20. Li Y, Lu J, Han Y, Fan X, Ding S-W (2013) RNA interference functions as an antiviral immunity mechanism in mammals. *Science* 342(6155):231–4.

21. Maillard P V, et al. (2013) Antiviral RNA interference in mammalian cells. *Science* 342(6155):235–8.

22. Lu R, Yigit E, Li WX, Ding SW (2009) An RIG-I-like RNA helicase mediates antiviral RNAi downstream of viral siRNA biogenesis in Caenorhabditis elegans. *PLoS Pathog* 5(2):8–14.

23. Ashe A, et al. (2013) A deletion polymorphism in the Caenorhabditis elegans RIG-I homolog disables viral RNA dicing and antiviral immunity. *Elife* 2013(2):1–21.

24. Guo X, Zhang R, Wang J, Ding S-W, Lu R (2013) Homologous RIG-I-like helicase proteins direct RNAi-mediated antiviral immunity in C. elegans by distinct mechanisms. *Proc Natl Acad Sci U S A* 110(40):16085–90.

25. Guo X, Zhang R, Wang J, Lu R (2013) Antiviral RNA silencing initiated in the absence of RDE-4, a double-stranded RNA binding protein, in Caenorhabditis elegans. *J Virol* 87(19):10721–9.

26. Guo X, Lu R (2013) Characterization of virus-encoded RNA interference suppressors in Caenorhabditis elegans. *J Virol* 87(10):5414–23.

27. Aliyari R, et al. (2008) Mechanism of Induction and Suppression of Antiviral Immunity Directed by Virus-Derived Small RNAs in Drosophila. *Cell Host Microbe* 4(4):387–397.

28. Li H, Li WX, Ding SW (2002) Induction and suppression of RNA silencing by an animal virus. *Science* 296(5571):1319–1321.

29. Li W-X, et al. (2004) Interferon antagonist proteins of influenza and vaccinia viruses are suppressors of RNA silencing. *Proc Natl Acad Sci U S A* 101(5):1350–1355.

30. Moresco EMY, Li X, Beutler B (2013) Going forward with genetics: Recent technological advances and forward genetics in mice. *Am J Pathol* 182(5):1462–1473.

31. Davis MW, et al. (2005) Rapid single nucleotide polymorphism mapping in C. elegans. *BMC Genomics* 6:118.

32. Sarin S, et al. (2010) Analysis of multiple ethyl methanesulfonate-mutagenized Caenorhabditis elegans strains by whole-genome sequencing. *Genetics* 185(2):417–430.

33. Zuryn S, Le Gras S, Jamet K, Jarriault S (2010) A strategy for direct mapping and identification of mutations by whole-genome sequencing. *Genetics* 186(1):427–430.

34. Doitsidou M, Poole RJ, Sarin S, Bigelow H, Hobert O (2010) C. elegans mutant identification with a one-step whole-genome-sequencing and SNP mapping strategy. *PLoS One* 5(11). doi:10.1371/journal.pone.0015435.

35. Minevich G, Park DS, Blankenberg D, Poole RJ, Hobert O (2012) CloudMap: A cloud-based pipeline for analysis of mutant genome sequences. *Genetics* 192(4):1249–1269.

36. Abe A, et al. (2012) Genome sequencing reveals agronomically important loci in rice using MutMap. *Nat Biotechnol* 30(2):174–178.

37. Zhu Y, et al. (2012) Gene discovery using mutagen-induced polymorphisms and deep sequencing: application to plant disease resistance. *Genetics* 192(1):139–46.

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

39. Winston WM, Molodowitch C, Hunter CP (2002) Systemic RNAi in C. elegans requires the putative transmembrane protein SID-1. *Science* 295(5564):2456–2459.

40. Tijsterman M, May RC, Simmer F, Okihara KL, Plasterk RH a (2004) Genes required for systemic RNA interference in Caenorhabditis elegans. *Curr Biol* 14(2):111–116.

41. Perrotta AT, Been MD (1991) A pseudoknot-like structure required for efficient self-cleavage of hepatitis delta virus RNA. *Nature* 350(6317):434–436.

42. Johnson KL, Ball LA (1999) Induction and maintenance of autonomous flock house virus RNA1 replication. *J Virol* 73(10):7933–7942.

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

44. Maduro M, Pilgrim D (1995) Identification and cloning of unc-119, a gene expressed in the Caenorhabditis elegans nervous system. *Genetics* 141(3):977–988.

45. Jorgensen EM, Mango SE (2002) The art and design of genetic screens: caenorhabditis elegans. *Nat Rev Genet* 3(5):356–369.

46. Neff MM, Turk E, Kalishman M (2002) Web-based primer design for single nucleotide polymorphism analysis. *Trends Genet* 18(12):613–615.

47. Jiang H, Lei R, Ding S-W, Zhu S (2014) Skewer: a fast and accurate adapter trimmer for next-generation sequencing paired-end reads. *BMC Bioinformatics* 15(1):182.

48. Rual J-F, et al. (2004) Toward Improving Caenorhabditis elegans Phenome Mapping with an ORFeome-Based RNAi Library. *Genome Res* 14:2162–2168.

49. Yang H, et al. (2012) The RDE-10/RDE-11 complex triggers RNAi-induced mRNA degradation by association with target mRNA in C. elegans. *Genes Dev* 26(8):846–856.

50. 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.

51. Ketting F, Steiner FA, Okihara KL, Hoogstrate SW, Sijen T (2009) RDE-1 slicer activity is required only for passenger-strand cleavage during RNAi in Caenorhabditis elegans. 16(2):207–211.

52. Vastenhouw NL, et al. (2006) Long-term gene silencing by RNAi. *Nature* 442:881.

53. Cui M, Kim EB, Han M (2006) Diverse Chromatin Remodeling Genes Antagonize the Rb-Involved SynMuv Pathways in C. elegans. *PLoS Genet* 2(5):e74.

CHAPTER 3

NEW FUNCTION OF THE RIG-I HOMOLOG DRH-1 IN RNAI-MEDIATED ANTIVIRAL IMMUNITY

3.1 Abstract

The *Caenorhabditis elegans* protein DRH-1 exhibits the highest sequence similarity in the ATPase and C-terminal domains of the mammalian virus immune receptor retinoic acid-inducible gene I (RIG-I). Both RIG-I and DRH-1 bind to viral dsRNA in the cytoplasm to activate type-1 interferon signaling in mammals and antiviral RNA interference (RNAi) in *C. elegans*. RIG-I can also translocate on long dsRNA powered by ATP hydrolysis, but the biological function of this translocase activity is unclear. Here we identified *C. elegans* mutants defective in antiviral RNAi and uncovered four loss-of-function alleles of *drh-1*, three of which were located in the ATPase and C-terminal domains. We show that *drh-1* mutants produce primary virus-derived small interfering RNAs (vsiRNAs) dependent on Dicer-1 and dsRNA-binding protein RDE-4, which are known to interact with each other and associate with DRH-1. However, vsiRNAs found in *drh-1* mutants were overwhelmingly mapped to the 5' terminal regions of the bipartite positive-strand RNA genome of Orsay virus. In contrast, such an abrupt decrease in the density of vsiRNAs to target the internal regions of the

viral RNA genome was not detected in any of the wildtype and mutant *C. elegans* strains that encode a functional DRH-1, indicating a new function of DRH-1 in the biogenesis of vsiRNAs. We propose that DRH-1 enhances the processing of the internal regions of viral dsRNA replicative intermediates by a mechanism involving dsRNA translocation of Dicer-1 in complex with DRH-1.

3.2 Introduction

3.2.1 Duplex RNA-activated ATPases (DRAs)

Conserved dicer endonucleases, mammalian RIG-I-like Receptors (RLRs), and *C. elegans* Dicer-related Helicases (DRH) proteins belong to a class of conserved Duplex RNA-activated ATPases (DRAs). Originally classified as a group of RNA helicases, DRAs do not have bona fide helicase activity. Helicases are nucleic acid-dependent ATPases that unwind double-stranded nucleic acids and were originally grouped into multiple helicase families based on sequence similarity. Modern classification recognizes that these proteins are nucleic acid-dependent ATPases that function as translocases with diverse activities, helicase activity being only one of them (1). Translocases are molecular motors that can move along single or double stranded nucleic acids in either a 3'-5' (α) or 5'-3' (β) direction (1, 2). In light of this recent scientific discussion, we will refer to the conserved "Helicase domain" as the "ATPase core domain" (Fig 3.1). These terms are currently being used interchangeably in the literature, but the individual sub-domains have retained helicase nomenclature (e.g. Hel 1, Hel 2).

DRA proteins exhibit ATPase activity that is dependent on dsRNA binding and coupled to functions such as a conformational change in the DRA protein, changes in the bound nucleic acid, remodeling of the protein-nucleic acid complex or translocation of the DRA along the bound nucleic acid. The ATPase domain of DRAs is similar in sequence and structure to the DEAD BOX family of helicases, containing 2 RecA-like folds, called Hel 1 and Hel 2. In contrast to DEAD Box Helicases that bind ssRNA proximal to dsRNA, DRAs preferentially bind dsRNA. Binding of duplexed RNA is dependent on three unique features of DRAs. The first is a α-helical insertion domain, called Hel 2i that is located within the second RecA-like fold of the ATPase core domain. The second and third are two semi-conserved motifs called IIa and Vc that are conserved

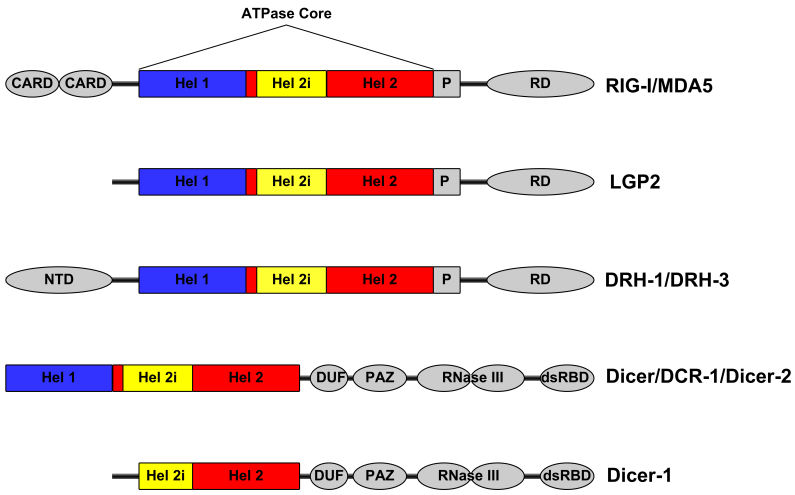

Figure 3-1. Duplex RNA Activated ATPases. The ATPase core domain is comprised of the Hel 1 (Blue), Hel 2 (Red) and the Hel 2i (Yellow) domain inserted within the Hel 2 domain. Mammalian RIG-I and MDA5 share domain architecture, while LGP2 differs only in the lack of CARD domains. The *C. elegans* DRH-1 and DRH-3 also share key domains with mammalian RLRs, though they contain a nematode specific N-terminal domain. Dicer proteins contain the conserved ATPase core domain, but contain domains specific to its function as an RNase. Interestingly, Dicer-1 from *D. melanogaster* does not process long dsRNA and encodes a truncated ATPase core domain.

in structure more than sequence. These motifs make contacts with the 5'-3' strand, the second strand, allowing for duplexed RNA binding (2). Many other translocases contain domains that only contact the 3'-5' strand, or tracking strand. Finally, DRAs also lack a crucial β-hairpin motif required for strand separation by DEAD Box helicases. It is therefore not surprising that experiments to detect helicase activity in DRAs have been unsuccessful.

3.2.2 Function of DRAs in Innate Immunity

In the innate immune system, Pattern Recognition Receptors (PRRs) detect molecular structures that are not shared by the host and are relatively constant features of groups of pathogens, called Pathogen Associated Molecular Patterns (PAMP). Examples of Bacterial PAMPs include the lipopolysaccharides of gram-negative bacteria, peptidoglycan of gram-positive bacteria or the flagellin protein of flagellated bacteria. On the other hand, viruses lack widely conserved proteins and other macromolecules, so viral PAMPs are limited to foreign nucleic acids, such as unmethylated DNA, dsRNA or 5'-triphosphate RNA. In mammals, RLRs are PRRs that function as cytosolic immune receptors to detect viral RNA and activate signaling pathways. RLRs are conserved in vertebrates and RLR-like proteins are also found in some invertebrates, including nematodes, sea anenomes and sponges (3). RLR proteins contain two tandem caspase recruitment domains (CARD) at the N-terminus and a Regulatory Domain (RD) at the C-terminus, which are required for signaling and viral RNA detection and binding, respectively. Additionally, the ATPase core domain of RLRs is tethered to the RD by a

Pincer domain comprised of two α-helices required to couple RNA binding to ATPase activity (4). Mammals encode three RLRs: Retinoic Acid-Inducible Gene I (RIG-I), Melanoma Differentiation-Associated protein 5 (MDA5) and Laboratory of Genetics and Physiology 2 (LGP2). Both RIG-I and MDA5 share domain architecture, while LGP2 lacks the tandem CARD domains (Fig. 3-1).

RIG-I was initially discovered as retinoic acid induced gene during the differentiation of acute promyelocytic leukemia in 1997 (5) and soon after in pigs as RHIV-1, a gene induced by infection of Porcine Reproductive and Respiratory Syndrome Virus (6). In 2002, the RLR-like protein, DRH-1, was isolated in complex with Dicer in *C. elegans* and the conservation of their ATPase domains was discovered (7). In 2004, RIG-I was discovered to be essential for activating the interferon pathway in response to viral dsRNA (8) and in the decade since RIG-I has been extensively studied. RIG-I recognizes viruses with 5'-triphosphate as a non-self signature, such as Influenza A, Newcastle disease virus, sendai virus, and vesicular stomatitis virus, among others (9, 10). A recent study, demonstrated that RIG-I can also detect reovirus dsRNA with 5'-diphosphates (11). In the absence of an appropriate PAMP, the Hel 2i domain sequesters the signaling CARD domains of RIG-I. When the RD domain recognizes and tightly binds to the 5'-triphosphate end of the RNA, the Hel 2i domain makes contacts with the RNA backbone, releasing the tandem CARD domains. This conformational change also allows ATP to bind RIG-I, a step required for its activation. Once exposed, the CARD domains are polyubiquitnated at lysine 63, which promotes formation of tetrameric filaments along the viral RNA. In its role as a sensor, ligand-bound RIG-I is the

beginning of a signaling cascade that propagates to MAVS/IPS-1 at the mitochondria via the exposed CARD domains, and then to the cytoplasmic kinases IKK and TBK1 that activate the transcription factors, NF-κB and IRF3, respectively. Phosphorylated NF-κB and IRF3 move to the nucleus to activate transcription of Type I interferons and other inflammatory genes, establishing a local immune response as well as an intracellular antiviral state to protect uninfected cells. Recent studies support a model in which RIG-I has an additional function as an effector of antiviral immunity by displacing key viral proteins from dsRNA and either directly interfering with viral RNA replication or indirectly promoting detection by other antiviral proteins, such as PKR (12–14).

The RLR MDA5 recognizes internal regions of dsRNA as a non-self signature and detects viral infection by picornoviruses, such as Polio, which do not contain a 5'-triphosphate. Some viruses, such as West Nile and Dengue, can be detected by either MDA5 or RIG-I. The RD domain of MDA5 differs slightly from RIG-I, lacking the 5'-end capping loop, so that MDA5 is indifferent to the dsRNA ends. Unlike RIG-I, the CARD domains of MDA5 are not sequestered in the absence of a ligand (15). Instead, the CARD domains allow MDA5 to cooperatively form long filaments along the dsRNA (15). Ligand-bound MDA5 relays the signal downstream to MAVS-IPS-1 at the mitochondria to activate RLR-mediated immunity. MDA-5 can also be polyubiquinated; however, the role of this ubiquitination remains unknown.

LGP2 is the third and least studied RLR; consequently, little is known about its function. LGP2 is unique in that it lacks the N-terminal CARD domains. Early studies suggested that LGP2 negatively regulated RIG-I mediated by a high affinity for dsRNA

(16). More recent studies suggest that LGP2 may act cooperatively with MDA-5 by sensing viral dsRNA from picornovirses and by aiding in MDA5 filament formation (17, 18). Studies with LGP2 knockouts also report conflicting results (19–21) and more research is needed to elucidate the role of LGP2 in RLR signaling.

3.2.3 Function of DRAs in RNAi

3.2.3.1 Dicer

The RNaseIII, Dicer, is the central component of conserved RNAi pathways. In plants, insects, nematodes and some fungi, RNAi is a potent antiviral pathway, with Dicer playing dual roles as a sensor and a more limited role as an effector. Dicer encodes an N-terminal ATPase core domain homologous to RLRs, in addition to a domain of unknown function (DUF283), a PAZ domain, two tandem RNase III and dsRNA binding domain (Fig. 3-1). The ATPase domain is highly conserved amongst dicer proteins, yet it is phylogeneticaly distinct from the ATPase domains found in mammalian and nematode RLRs (2). As an effector of antiviral immunity, Dicer recognizes viral dsRNA replication intermediates, usually with the assistance of a dsRNA binding protein and cleaves the viral RNA into viral-derived siRNAs (vsiRNAs). However, the initial cleavage of the viral RNA is not sufficient to clear viral infection and the vsiRNA is loaded into an argonaute protein to form an RNA-induced Silencing Complex (RISC). RISC is guided by the vsiRNA to genomic viral RNA and the argonaute protein cleaves the viral genomic RNA or recruit additional factors to mediate viral clearance. It is in this capacity to initiate RNAi that dicer is considered a sensor, analogous to RLRs.

Many organisms encode multiple dicer proteins to mediate different small RNA pathways. The fruit fly *D. melanogaster* encodes two dicer proteins that function in unique RNAi pathways. Dicer-1 processes miRNA precursors, while Dicer-2 processes long dsRNA, including viral RNA. Interestingly, Dicer-2, but not Dicer-1, encodes a full ATPase core domain (Fig 3-1). Considering that miRNA processing requires only one cleavage event, the ATPase core domain would be dispensable for this function. In the single *C. elegans* dicer (DCR-1) mutations in the helicase domain also affects endo-siRNA biogenesis, but not miRNAs (22). Additionally, *in vitro* assays suggest that the helicase domain is required to process small RNAs from the internal region of dsRNA (23). Interestingly, antiviral RNAi in *C. elegans* requires two additional DRAs to effectively clear viral infection.

3.2.3.2 RIG-I-like Proteins

Mello and colleagues (7) were the first to note the strong sequence similarity among the ATPase domains of *C. elegans* DRH proteins with Dicer proteins and the mammalian RLRs. Interestingly, the ATPase domain of DRH proteins, particularly DRH-1, is more closely related to RIG-I than Dicer. In addition to the ATPase core domain, both DRH-1 and DRH-3 encode a pincer domain and a RD domain homologous to RLRs, but lack N-terminal CARD domains. Interestingly, either the ATPase or RD of human RIG-I can functionally replace the homologous domain of DRH-1 to mediate antiviral RNAi in *C. elegans* (24). Moreover, the conserved KWK motif in the RD that is implicated in dsRNA specificity of RIG-I is also required for the antiviral activity of both

DRH-1 and the chimeric protein containing the N-terminal domain from DRH-1 and the ATPase and RD of RIG-I. These findings suggest that DRH-1 and RIG-I share activities such as RNA specificity and translocation along dsRNA and raise an inttriguing possibility for a role of mammalian RLRs in antiviral RNAi (25, 26). The N-terminal domain (NTD) of the DRH proteins is nematode specific and is essential for their function (24, 27). In DRH-3, the NTD is required for size specificity, binding preferentially to 22-nt duplexes (27), indicating the NTD may specify different small RNA binding specificities of DRH-1 and DRH-3.

DRH-1 and DRH-3 function in sequential steps of antiviral RNAi. DRH-1 is found in a complex with DCR-1 and RDE-4 (7, 28) and functions in the production of primary vsiRNAs from viral dsRNA (24, 29, 30). Interestingly, the DCR-1/RDE-4 complex does not require the function of DRH-1 in the production of primary vsiRNAs from synthetic or endogenous dsRNA (29). DRH-3 is found in complex with the RdRP-complex required for the production of 22G-RNAs targeting viral RNA, as well as endogenous mRNA targeted for silencing (24, 30, 31). In the absence of DRH-3, 22G-RNAs are depleted, but high levels of primary small RNAs remain intact, indicating that DRH-3 may function directly in the production of or in the stability of 22G-RNAs. Analysis of 22G-RNA distribution along target genes revealed that 22G-RNAs are skewed towards the 3' end of the transcript (31), indicating the DRH-3 may function in promoting the production of 22G-RNAs from the internal regions and the 5' end.

The *C. elegans* RNAi pathway has a unique requirement for three DRAs to effectively mediate antiviral RNAi. Both DCR-1 and DRH-1 are required for effective

production of primary vsiRNAs, while DRH-3 is required for secondary vsiRNA production. In contrast to other organisms that require secondary vsiRNAs to efficiently clear viral infection, the production of 22G-RNAs is DCR-1 independent. It is intriguing that an alternative DRA is required in its place, perhaps to compensate for specific Dicer functions, such as translocation. In this chapter we investigate the specific roles of DCR-1 and DRH-1 in the production of primary vsiRNAs.

3.3 Materials and Methods

3.3.1 Worm Maintenance and Genetics

Worms were maintained on standard Nematode Growth Medium (NGM) at 20°C, unless otherwise stated. Mutants referred to in the text refer to the alleles in Table 3-1. *C. elegans* media and buffers were prepared as described in Appendix A.

Table 3-1. ***C. elegans*** **strains by Chromosome**

Referenced in text	**Genotype**
N2	Bristol N2 strain (32)
LGI	
drh-3	*drh-3(ne4253)*
rrf-1	*rrf-1(pk1417)*
LGII	
FR1gfp	*irSi18[$_p$hsp-16.41::FR1gfp::Rz]*
LGIII	
rde-4	*rde-4(ne301); irSi18[$_p$hsp-16.41::FR1gfp::Rz]*
LGIV	
drh-1(tm1329)	*drh-1(tm1329); irSi18[$_p$hsp-16.41::FR1gfp::Rz]*
drh-1	*drh-1(ok3495)*
drh-1(ucr2)	*drh-1(ucr2); irSi18[$_p$hsp-16.41::FR1gfp::Rz]*
drh-1(ucr3)	*drh-1(ucr3); irSi18[$_p$hsp-16.41::FR1gfp::Rz]*
drh-1(ucr4)	*drh-1(ucr4); irSi18[$_p$hsp-16.41::FR1gfp::Rz]*
drh-1(ucr6)	*drh-1(ucr6); irSi18[$_p$hsp-16.41::FR1gfp::Rz]*
LGV	
rde-1	*rde-1(ne219)*

3.3.2 OrV infections

For OrV infections, 10-30 young adult worms were transferred to new 60mm NGM plates seeded with OP50 *E. coli*. 10 young adults were sufficient for the N2 strain; however more worms were used for difference mutants, due to a smaller brood size. 50μL of OrV filtrate was added to each plate and worms were incubated at 20°C. Five days post infection, the next generation of worms were collected for RNA extractions. OrV filtrate was prepared by growing infected *drh-1(tm1329)* animals on ten 100mm plates. The mixed stage worms, mostly adults, were washed off in 3mL of M9 buffer into two 1.5mL ependorf tubes. The infected worms were homogenized and centrifuged for 1 minute at 13,000rpm. The supernatant was then filtered through a 20μm filter. The filtrates from each plate were mixed together and aliquoted into 1.5mL ependorf tubes to be used in experiments.

3.3.3 Preparation of Small RNA Libraries

Small RNAs were extracted using 4M LiCl followed by precipitation in 2 volumes ethanol. Up to 5μg of purified small RNAs was used in a RNA 5' Polyphosphatase (RPP) reaction to remove the 5'-triphosphate groups from the secondary small RNAs. The reactions were incubated for 30 minutes at 37°C, per the manufacture's method (Epicentre). RPP was removed from the sample by Trizol extraction (see below) and the small RNAs were re-suspended in 7μL of DEPC-H_2O. 6μL of the recovered, RPP-treated, purified small RNAs were used in the generation of small RNA libraries

using the Truseq Small RNA sample preparation kit (Illumina). Samples were multiplexed and sequenced on an Illumina 2500.

3.3.4 Small RNA Analysis

Trimming and removal of structural small RNAs. Small RNA analysis was performed using Bowtie (0.12.7), custom Perl (5.16.3) scripts and Wormbase (WS240). Adapter sequences ('tggaattctcgggtgccaaggc') were removed from the 3' end of each read and filtered by size to obtain inserts 18-30nt long, using Skewer (33). Reads from each library that that aligned with zero mismatches to the sense strand of structural small RNAs were removed (31); sequences for 140 snoRNAs, 631 tRNAs, 1 scRNA, 22 rRNAs and 114 snRNAs were extracted from WormBase and used in this analysis. The resultant nonstructural siRNA reads were used in the following analyses. The nonstructural small RNAs were aligned to *C. elegans* miRNAs obtained from WormBase (WS240) and used to normalize vsiRNA reads.

Analysis of vsiRNA profiles. Small RNA reads were aligned to the Orsay Virus genome (Genbank IDs: HM030970.2 and HM030971.2) using Bowtie, allowing zero mismatches. Aligned reads were then parsed by size, polarity, and 5' nucleotide identity using a custom perl script.

Analysis of vsiRNA distribution. Small RNA reads that aligned to the OrV genome were binned into groups of 20-nt, based on the position of their 5’ nucleotide. Reads per million miRNAs (x1000) were graphed.

Table 3-2. **Properties of Small RNA Libraries**

Samples	**Total Reads**	**vsiRNAs**	**Reference**
5'-independent libraries			
N2	12,267,058	160,012 (1.3%)	This study
rrf-1	20,840,478	869,173 (4.2%)	This study
drh-1(tm1329)	9,363,236	404,969 (4.3%)	This study
drh-1(ucr2)	2,219,398	105,584 (4.8%)	This study
drh-1(ucr3)	2,827,190	62,102 (2.2%)	This study
drh-1(ucr4)	1,830,021	102,451 (5.6%)	This study
drh-1(ucr6)	7,107,381	200,351 (2.8%)	This study
N2	2,775,912	14,969 (0.5%)	Ashe et al. 2013
drh-1(ok3495)	3,251,771	30,690 (0.9%)	Ashe et al. 2013
rde-1	2,776,556	13,798 (0.5%)	Ashe et al. 2013
rde-4	4,075,692	17,067 (0.4%)	Ashe et al. 2013
drh-3	2,131,804	352,876 (16.6%)	Ashe et al. 2013
rde-1; drh-1(ok3495)	3,628,014	14,825 (0.4%)	Ashe et al. 2013
drh-3; drh-1(ok3495)	2,550,313	139,831 (5.5%)	Ashe et al. 2013
rde-4; drh-1(ok3495)	3,398,365	143,757 (4.2%)	Ashe et al. 2013
5'-dependent libraries			
N2	2,116,311	18,571 (0.9%)	Ashe et al. 2013
drh-1(ok3495)	3,858,387	95,984 (2.5%)	Ashe et al. 2013
drh-3	1,375,217	166,949 (12.1%)	Ashe et al. 2013
rde-1	2,776,556	13,798 (0.5%)	Ashe et al. 2013

Total reads represent the total nonstructural reads between 18 and 30-nt in length. Total nonstructural reads were calculated by filtering out sense reads that align to structural small RNAs (see methods). Reads that aligned to the OrV genome and were 21 to 24-nt in length were considered genuine vsiRNAs. The percentage of total nonstructural reads is shown.

3.3.5 Viral RNA detection by Northern Blot.

Total RNA was extracted from homogenized worms using Trizol (Sigma-Aldrich) by following the manufacturer's protocol, with the addition of an extra chloroform separation with a 1:1 ratio before precipitation. 4µg of high molecular weight RNA was ran on a 1.2% agarose gel with Formaldehyde. [α-^{32}P]dCTP labeled cDNA corresponding to OrV RNA 1 was used to detect OrV infection. OrV probes were labeled with [α-

^{32}P]dCTP using the Ready to Go DNA Labelling kit (-dCTP) (Amersham Biosciences). Hybridization of both probes was done at 65°C for three hours. Buffers and media used for Nothern Blotting can be found in Appendix B.

3.4 Results

3.4.1 Isolation of four novel drh-1 alleles

Both *ucr3* and *ucr4* contained a missense mutation, leading to single amino acid substitution in the N- and C-terminal domains of DRH-1, respectively (Fig. 3.2A). The N-terminal domain of DRH-1 shares no detectable similarity to the N-terminal CARDs of mammalian RIG-I/MDA5. By contrast, the proline residue at position 966 (Pro^{966}) mutated in *ucr4* is conserved in the RD of all of the known RLRs encoded by *C. elegans* and mammals (Fig. 3-2A). The nonsense mutation in *ucr6* induced a premature stop codon at residue position 865 of DRH-1, resulting in a putative protein that lacks the entire RD (Fig. 3-2A).

The G→A transition identified in *ucr2* was located at the 5'-terminal nucleotide position of intron 13 immediately after exon 14 of *drh-1*, predicted to disrupt the 5' consensus splice site, 5'-AG|GURAGUU-3'. Indeed, our characterization of the mature mRNA of *drh-1* in *ucr2* mutant worms revealed the use of an alternative, upstream splice site, 5'-GG|GUUUGUA, leading to a 90-nt deletion in the coding sequence (Fig. 3-2C). The resultant mRNA encodes a protein with a predicted in-frame deletion of 30 amino acids from position 596 to 625. This deletion would disrupt the Hel2i domain, which is

conserved in all of the known RLRs encoded by *C. elegans* and mammals (Fig. 3.2B) and a defining characteristic of DRAs that specialize in duplex RNA recognition (2).

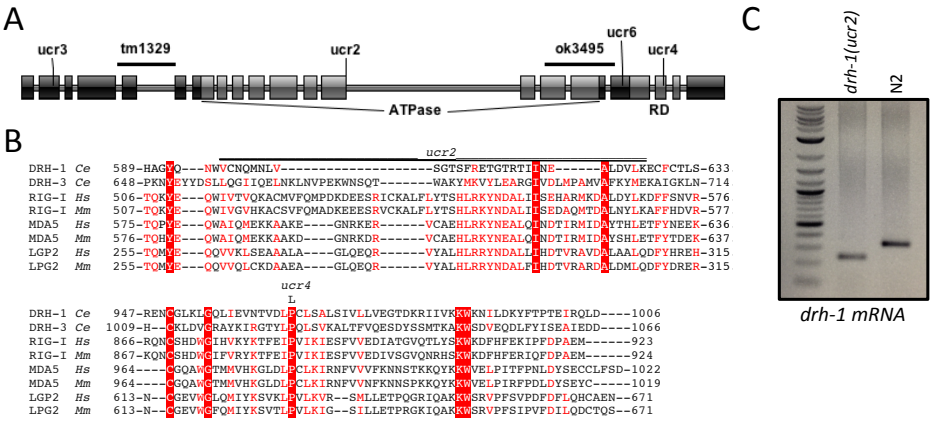

Figure 3-2. Identification of four loss-of-function alleles of *drh-1*. *(A)* Schematic of the *drh-1* gene structure with the identified *ucr* alleles indicated. The ATPase domain (divided into Hel1, Hel2i and Hel2 motifs) and the C-terminal regulatory domain (RD) are conserved in *C. elegans* DRH-1/DRH-3 and mammalian RIG-I/MDA5. Also indicated are the regions deleted in two known *drh-1* mutants. (B) Alignments showing the region deleted by the splice site mutation in *ucr2* and the amino acid substitution at a conserved Proline in *ucr4*. *(C)* Reverse transcription and PCR to detect a 90-nt deletion in the *drh-1* mRNA in *drh-1(ucr2)* mutant compared to that in N2 animals.

3.4.2 Characterization of vsiRNAs produced during Orsay infection

Using a potocol to isolate both the mono- and tri-phosphate small RNAs, we sequenced the small RNAs from wild-type, N2, *rrf*-1, *drh-1(tm1329)*, *drh-1(ucr2)*, *drh-1(ucr3)*, *drh-1(ucr4)* and *drh-1(ucr6)* animals infected with OrV, discussed in the previous chapter (Fig. 2-2B-C). Wild-type animals are able to effectively clear the viral infection and over 90% of the vsiRNAs detected in N2 animals are antisense, indicative of secondary vsiRNAs and a robust antiviral RNAi response to OrV infection. Nearly half of the vsiRNAs are 22G-RNAs, and another fourth are 21G-RNAs during a

functional antiviral RNAi response (Fig. 3-2A). In contrast, in *rrf-1* and all of the *drh-1* mutants, between 50-60% of the vsiRNAs were antisense, indicating a defect in the production of vsiRNAs. As expected, in *rrf-1*mutants only primary vsiRNAs are produced (Fig. 2-2C; Fig. 3-2C). While some 22G-RNAs are still produced in all the *drh-1* mutants, their prevalence in the vsiRNA population is markedly reduced. Surprisingly, we also detected a 23-nt sense peak in all of the *drh-1* mutants (Fig. 2-2C; Fig. 3-2C). Additionally, in wild-type animals most vsiRNAs are derived from RNA1 of OrV. In contrast, most of the vsiRNAs in *rrf-1* and all of the *drh-1* mutants are derived from OrV RNA2, despite RNA2 being 800-nt shorter than RNA1 (Fig. 3-3B). The preference for RNA2 may reflect an increase in abundance of RNA 2 in animals undergoing robust viral infection or the loss of 22G-RNA production, which is biased towards RNA1 in N2 animals.

Next we sorted the vsiRNAs into primary vsiRNAs (23-nt vsiRNAs) and secondary vsiRNAs (antisense 22G-RNAs). In N2, the abundance of 22G-RNAs is much greater than 23-nt vsiRNAs, with a ratio of 2.8 22G-RNAs/23-nt vsiRNAs (Fig. 3-3C). In contrast, *rrf-1* is essential for the biogenesis of 22G-RNAs, but dispensable for primary vsiRNA production and has a ratio of .06. All of the *drh-1* mutants fall in between and have either more primary vsiRNAs than 22G-RNAs or comparable amounts (Fig. 3-3C). These results suggest that primary vsiRNAs are produced in the absence of *drh-1;* however, these DRH-1 independent vsiRNAs are unable to effectively produce 22G-RNAs.

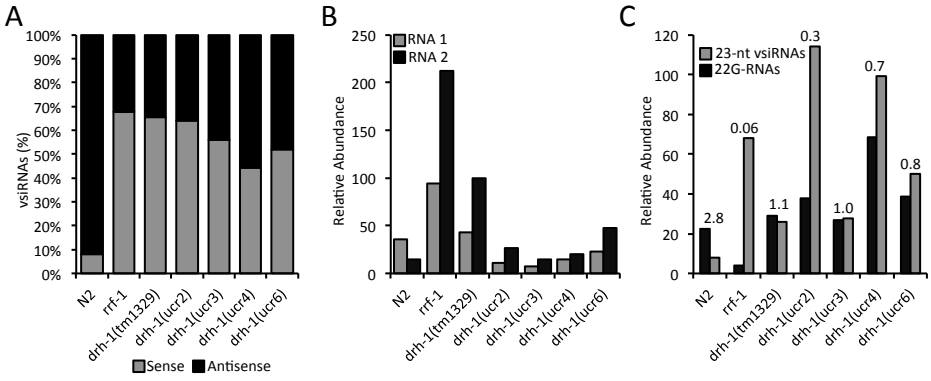

Figure 3-3. Characterization of OrV-derived vsiRNAs. (A) The percentage of vsiRNA (21-24-nt) that map to the genomic (+)-sense OrV RNA (black) and the antigenomic (-)-sense OrV RNA. (B) The relative abundance of 21-24-nt vsiRNAs that map to OrV RNA1 (gray) and OrV RNA2 (black). (C) The relative abundance of sense and antisense 23-nt vsiRNAs (gray) and antisense 22G-RNAs (black). The ratio of 22G-RNAs/23-nt vsiRNAs is shown. The relative abundance is calculated as reads per million miRNA (x1000).

3.4.3 Production of abundant 23-nt vsiRNAs in absence of DRH-1 targeting the terminal regions of viral genomic RNAs

Identification of multiple alleles in the *drh-1* gene from our unbiased genetic screen further supports an indispensable role of DRH-1 in antiviral RNAi in *C. elegans* revealed by previous studies (24, 29, 30). Curiously, we noted a peak for both the sense and antisense 23-nt vsiRNAs in the four new *drh-1* mutants as well as the control *drh-1(tm1329)* mutant (Fig. 2-2C), which was not detected previously in *drh-1(ok3495)* mutant infected with OrV (30). Absence of a distinct 23-nt peak among other sizes of OrV-derived small RNAs led to the idea that *drh-1* mutants are deficient in primary vsiRNA production (30), in disagreement with a previous study (29). We noted that the complete 5' and 3' terminal sequences of both OrV RNAs 1 and 2 (34) were not available for the initial mapping of vsiRNAs (30). Thus, we hypothesized that the population of the

23-nt vsiRNAs detected in *drh-1* mutants was derived from these terminal regions of the OrV genomic RNAs. To test this idea, we examined the distribution patterns of 23-nt vsiRNAs along the full-length genomic RNAs 1 and 2 of OrV sequenced from infected N2, *drh-1(ucr2), drh-1(ucr3), drh-1(ucr4)* and *drh-1(ucr6)* animals. Indeed, we found that in all four *drh-1* mutants, abundant 23-nt vsiRNAs in both positive and negative polarities were mapped to the 5' terminal region of both RNA1 and RNA2, and to the 3' terminal region of RNA2 although the density of the 3'-terminal vsiRNAs was low in *drh-1(ucr2)* and *drh-1(ucr6)* (Fig. 3-4A).

We further analyzed the total vsiRNAs mapped to the complete genome of OrV that were cloned similarly using a protocol to capture both primary siRNAs (with 5'-monophosphate) and secondary siRNAs (with 5'-triphosphate) by Miska and colleagues (30). Examining the size distribution of the total small RNAs derived from OrV revealed a peak for both the sense and antisense 23-nt vsiRNAs in *drh-1(ok3495)* animals (Fig. 3-4C). Similarly, the 5' terminal 23-nt vsiRNAs of both polarities were abundant in *drh-1(ok3495)* animals while the enrichment for the 3' terminal 23-nt vsiRNAs of RNA2 was modest as found for *drh-1(ucr2)* and *drh-1(ucr6)* animals (Fig. 3-4A).

By contrast, 23-nt vsiRNAs were not enriched in any of these three terminal regions in infected N2 animals (Fig. 3-4A). Instead, hot spots of 23-nt vsiRNAs were distributed randomly across the terminal and internal regions of the two genomic RNAs in the presence of DRH-1. We further examined the vsiRNAs from *rde-1*, *rde-4*, *rrf-1* and *drh-3* mutant animals, which are all defective in antiviral RNAi and support high levels of OrV replication (24, 29, 30, 35–39). All of these mutant animals produced

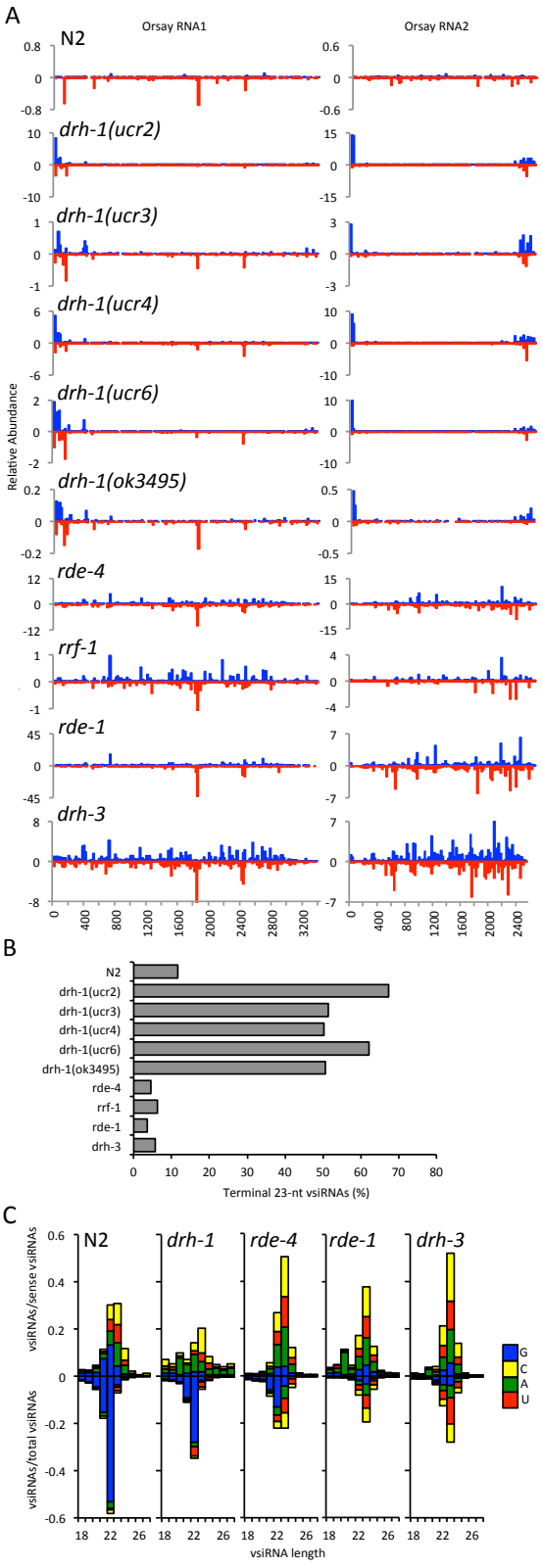

Figure. 3-4. Abundant 23-nt vsiRNAs target the terminal regions in *drh-1*. (A) Mapping of 23-nt sense (blue) and antisense (red) vsiRNAs to the full-length genomic RNAs of OrV obtained from N2 and mutant animals by a protocol to clone both primary and secondary vsiRNAs. The relative abundance was calculated as the 23-nt vsiRNA reads per million miRNAs (x1000). (B) % of 23-nt vsiRNAs mapped to the terminal regions (including 3'-terminal 200-nt region of RNA2 and 5'-terminal 200-nt regions from both RNAs 1 and 2) in N2 and mutant animals. (C) Size distribution, polarity, and the 5'-terminal nucleotide of total 18- to 28-nt vsiRNAs from N2 and mutant animals that were mapped to the full-length OrV genome RNAs. The libraries from *drh-1(ok3495)*, *rde-4*, *rrf-1*, *rde-1* and *drh-3* were published previously (Ashe et al. 2013).

abundant 23-nt vsiRNAs (Fig. 4C) as reported previously (24, 30). However, we detected no enrichment of the terminal 23-nt vsiRNAs in any of these mutant animals as found in N2 animals (Fig. 3-4A). These findings indicate that the terminal regions were not preferentially targeted for the production of the 23-nt vsiRNAs in the presence of wildtype DRH-1, regardless of whether or not there was robust viral RNA replication.

We next divided the 23-nt vsiRNAs from each library into terminal and internal vsiRNAs. Terminal vsiRNAs included the positive and negative strand 23-nt vsiRNAs mapped to the 5'-terminal 200-nt region of both RNA1 and RNA2 as well as those from the 3'-terminal 200-nt region of RNA2 whereas the remaining 23-nt vsiRNAs were all designated as internal vsiRNAs. 11.6% of the total 23-nt vsiRNAs produced in N2 animals were derived from the three terminal regions (or 10%) of the OrV genome and less than 10% of the total 23-nt vsiRNAs from *rde-1*, *rde-4*, *rrf-1* or *drh-3* mutant animals were terminal vsiRNAs (Fig. 3-4B). In contrast, more than 50% the total 23-nt vsiRNAs from the five distinct *drh-1* mutants were terminal vsiRNAs (Fig. 3-4B). These findings demonstrate production of highly abundant terminal vsiRNAs in the infected *drh-1* mutant animals, suggesting that DRH-1 functions to enhance the production of the vsiRNAs to target the internal regions of the viral RNA genome.

3.4.4 Terminal vsiRNAs are primary vsiRNAs produced by DCR-1 and RDE-4

The terminal vsiRNAs detected in *drh-1* mutant animals were 23-nt long and displayed no preference for either polarity or any 5'-terminal nucleotide, suggesting that they were primary vsiRNAs processed from dsRNA replicative intermediates by DCR-1.

Small RNAs made by Dicer also contain 5'-terminal monophosphate. Thus, we analyzed the previously published libraries of vsiRNAs cloned from infected N2, *rde-1*, *drh-1(ok3495)* and *drh-3* animals by a protocol to capture only small RNAs with 5'-monophosphate (30). The total sense and antisense vsiRNAs in the Dicer product size range of 21 to 24 nt in each library were mapped to the full length genomic RNAs 1 and 2 of OrV. We found that the internal vsiRNAs with 5'-monophosphate were as abundant as the terminal vsiRNAs in N2, *rde-1*, or *drh-3* animals (Fig. 3-5A), similar that found for the 23-nt vsiRNAs cloned from these animals by the protocol independent of the 5'-monophosphate (Fig. 3-4A). In *drh-1* mutant animals, however, abundant 5'-monophosphate vsiRNAs were mapped only to the 5'-terminal region of both OrV RNAs 1 and 2, and there was an abrupt decrease in the density of the 5'-monophosphate vsiRNAs to target the internal regions so that they were hardly visible even with an enlarged scale (Fig. 3-5A). Consistently, less than 10% of the total 5'-monophosphate vsiRNAs were derived from the 5'-terminal regions in N2 and *drh-3*, whereas more than 50% of the total 5'-monophosphate vsiRNAs from *drh-1* animals were mapped to the 5'-terminal regions (Fig. 3-5C). Moreover, the size distribution of the total 5'-monophosphate virus-specific small RNAs mapped to the 5'-terminal regions of OrV genomic RNAs from *drh-1* animals revealed peaks in the size range (22- and 23-nt) of Dicer products without strong preference for either polarity or any 5'-terminal nucleotide (Fig. 3-5B). These findings together indicate that the abundant 5'-terminal vsiRNAs detected in *drh-1* animals are the primary vsiRNAs processed viral dsRNA replicative intermediates by DCR-1 and that *drh-1* animals are defective in the production of Dicer-

dependent primary vsiRNAs to target the internal regions of the viral RNA genome. Intriguingly, the 3'-terminal peak of 23-nt vsiRNAs to target RNA2 detected in some *drh-1* mutants (Fig. 3-4A) disappeared when only 5'-monophosphate vsiRNAs from *drh-1* animals were examined (Fig. 3-5A), suggesting that the 3'-terminal vsiRNAs may not follow the same biogenesis pathway as the 5'-terminal vsiRNAs.

Miska and colleagues have published the total primary and secondary vsiRNAs sequenced from *drh-1(ok3495);drh-3*, *drh-1(ok3495);rde-1* and *drh-1(ok3495);rde-4*

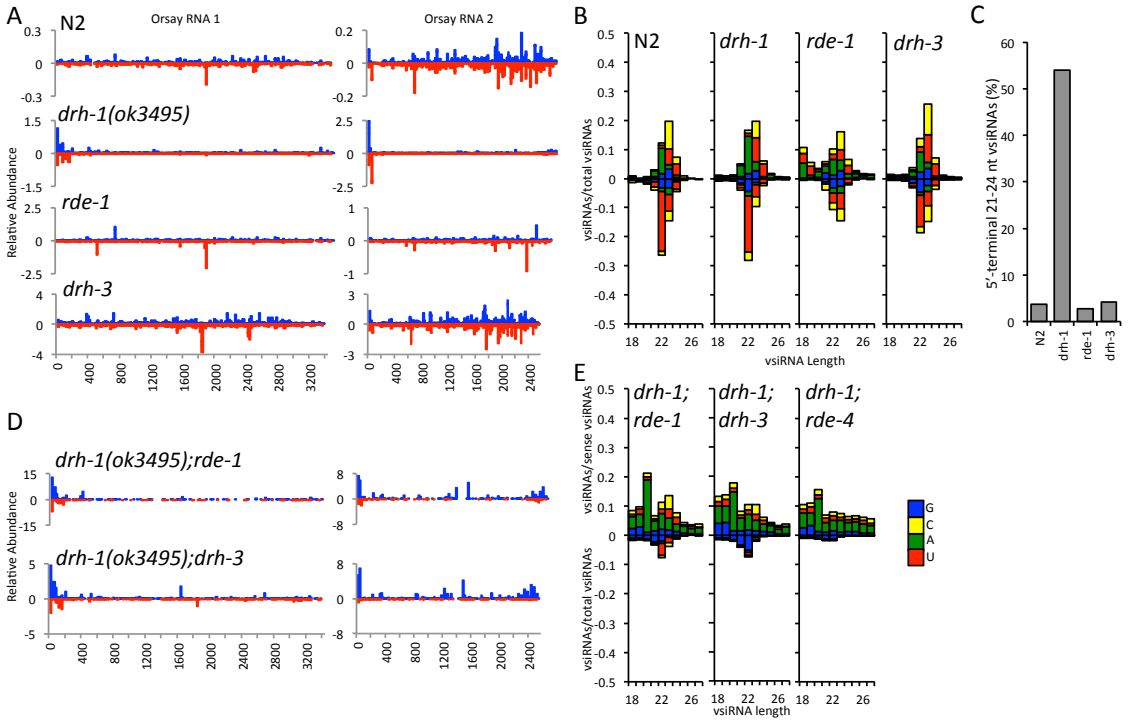

Fig. 3-5. Primary vsiRNAs from 5'-termini are dependent on DCR-1 and RDE-4. (A) Mapping of 21- to 24-nt sense (blue) and antisense (red) vsiRNAs to the genomic RNAs of OrV obtained from N2 and mutant animals by a protocol to clone only primary vsiRNAs. The relative abundance was calculated as the reads per million miRNAs (x1000). (B,C) Size distribution, polarity, and the 5'-terminal nucleotide of total vsiRNAs (*B*) or % of total 21- to 24-nt vsiRNAs (C) mapped to the 5'-terminal 200-nt regions of OrV RNAs 1 and 2 from the same libraries shown in (A). (D) Mapping of 23-nt sense (blue) and antisense (red) vsiRNAs to the genomic RNAs of OrV obtained from double mutants by a protocol to clone both primary and secondary vsiRNAs. (E) Size distribution, polarity, and the 5'-terminal nucleotide of total 18- to 28-nt vsiRNAs mapped to OrV RNAs 1 and 2 from the same libraries shown in (D). Small RNA libraries examined were published previously (Ashe et al. 2013).

double mutants infected with OrV (30). Re-examination of the total vsiRNAs mapped to the complete genome of OrV revealed the presence of 22- and 23-nt populations of the positive- and negative-strand vsiRNAs in both *drh-1;drh-3* and *drh-1;rde-1* mutants (Fig. 3-5E). Mapping of the 23-nt vsiRNAs to the genome of OrV further revealed the production of the abundant 5' terminal vsiRNAs in *drh-1;drh-3* and *drh-1;rde-1* mutants (Fig. 3-5D). These results indicate that neither RDE-1 nor DRH-3 is essential for the biogenesis of the terminal primary vsiRNAs. In contrast, virus-specific small RNAs cloned from infected *drh-1;rde-4* animals did not exhibit strong preference in the size range of Dicer products and were predominantly positive strands (Fig. 5E), suggesting that RDE-4 is necessary for the DCR-1-mediated production of primary vsiRNAs, including the terminal vsiRNAs.

3.4.5 Analysis of FR1gfp-derived vsiRNAs

In order to further determine if the production of 5'-terminal vsiRNAs in *drh-1* mutants was unique to OrV, we sequenced primary and secondary vsiRNAs from N2, *drh-1(tm1329)*, and *rde-1* animals carrying *irSi18,* forty eight hours after heat treatment. Similar to OrV-infected animals, vsiRNAs were more abundant in *drh-1(tm1329)* and *rde-1* animals than N2 animals (Table 3-3). In N2 animals, 98.4% of the vsiRNAs were antisense to FR1gfp, with a strong bias for 22G-RNAs (Fig 3-6A). Interestingly, we detected a weak signal of 22G-RNAs with a sense polarity, which may represent secondary vsiRNAs synthesized from antigenomic RNA (Fig 3-6A). In contrast, 41.3% and 49.9% of vsiRNAs were antisense in *drh-1(tm1329)* and *rde-1* with a strong primary

vsiRNA signature (Fig 3-6A). These results differ from OrV in the absence of any 22G-RNAs in *drh-1(tm1329)* mutant. We mapped the 23-nt vsiRNAs to FR1gfp and found that in N2 and *rde-1*, 23-nt vsiRNAs are evenly distributed across the genome. Similar to OrV, 23-nt vsiRNAs in *drh-1(tm1329)* were more abundant from the 5'-terminal region of FR1gfp RNA1. Additionally, abundant 23-nt vsiRNAs were detected along sgRNA3 encoding GFP, likely due to the higher abundance of sgRNA3 compared to RNA 2 (Fig. 3-6C). 23-nt vsiRNAs from the 200-nt terminal regions, or 15% of FR1gfp genomic RNA, comprise 14.5% and 7.4% of the 23-nt vsiRNAs in N2 and *rde-1*, respectively. In contrast, 37.9% of the 23-nt vsiRNAs are derived from terminal regions in *drh-1(tm1329)* (Fig. 3-6D). These results support the hypothesis that DRH-1 functions to promote primary vsiRNA production from internal regions of viral RNA.

Interestingly, we detected some differences between the profile in *drh-1* animals between FR1gfp and OrV. Notably, that no 22G-RNAs were present in animals undergoing FR1gfp replication (Fig. 3-6A). There are a few key differences between these system, which may contribute. The first is timing. OrV-infected animals are collected 5 days post infection, while animals undergoing FR1gfp replication are collected two days post induction. Additionally, they differ in viral RNA abundance, with FR1gfp viral RNAs being more abundant and they differ in tissue tropism, with OrV being restricted to the intestine, while FR1gfp is ubiquitous.

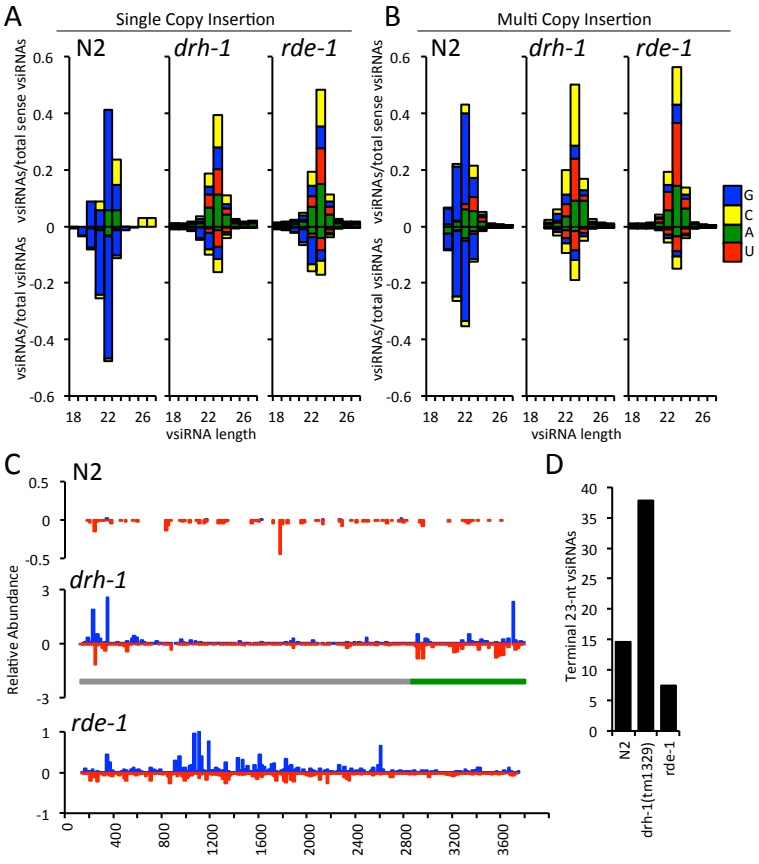

Figure 3-6. Characterization of vsiRNAs derived from FR1gfp. (A-B) Size distribution, polarity, and the 5'-terminal nucleotide of FR1gfp-derived primary and secondary vsiRNAs from N2 and mutant worms 48 hours after induction of FR1gfp lauched from *irSi18* (A) or *irIs91* (B). The prevalence of the antisense vsiRNAs (bottom) is shown as the proportion of total vsiRNAs whereas that of sense vsiRNAs (top) is presented as the proportion of total sense vsiRNAs. (C) Mapping of 23-nt sense (blue) and antisense (red) vsiRNAs to FR1gfp obtained from N2 and mutants undergoing FR1gfp replication by a protocol to clone both primary and secondary vsiRNAs. (D) Percentage of 23-nt vsiRNAs mapped to the terminal regions (including the 3'-terminal 200-nt region of FR1fgp RNA1 and 5'-terminal 200-nt regions from both FR1gfp RNA1 and sgRNA3) in N2 and mutant animals.

Previously, we developed a multi-copy insertion of FR1gfp, called *irIs91*, on Chromosome V. When we sequenced the primary and secondary small RNAs from N2, *drh-1(tm1329)* and *rde-1* animals, forty-eight hours after heat treatment, we found that the vsiRNAs profiles were nearly identical to the single copy insertion (Fig. 3-6A-B), though the vsiRNAs were markedly more abundant (Table 3-3). Differing from both OrV infected animals and animals undergoing FR1gfp replication launched from *irSi18*, we detected more vsiRNAs in N2 animals than in either mutant (Table 3-3). Additionally, in animals carrying *irIs9* that were not heat treated, we still detected vsiRNAs; in contrast, animals carrying *irSi18* that were not heat treated did not produce vsiRNAs (Table 3-3).

These vsiRNAs may pre-immunize the animals and explain why we observed enhanced FR1gfp replication in animals launched from the *irSi18* transgene compared to *irIs91*. Taken together, these results indicate that *irSi18* is a better model to study antiviral RNAi, because it shares key features of vsiRNA profiles with OrV infected animals, which are lacking in the *irIs91* model.

Table 3-3. Properties of FR1gfp-derived vsiRNAs

	Total Reads	**vsiRNAs**	**Abundance**
Single Copy Insertion			
N2	994,708	1,814 (0.2%)	19,837.5
drh-1(tm1329)	8,095,541	52,365 (0.6%)	33,645.1
rde-1(ne219)	7,797,697	41,904 (0.5%)	35,781.1
No heat shock	2,219,398	2 (0.00009%)	7.8
Multi-copy Insertion			
N2	5,114,508	268,026 (5.2%)	174,679.8
drh-1(tm1329)	829,759	21,686 (2.6%)	90,461.2
rde-1(ne219)	1,458,152	37,197 (2.6%)	99,428.8
No heat shock	9,874,041	942 (.01%)	4,278.9

Total reads represent the total nonstructural reads between 18 and 30-nt in length. Total nonstructural reads were calculated by filtering out sense reads that align to structural small RNAs (see methods). Reads that aligned to the FR1gfp and were 21 to 24-nt in length were considered genuine vsiRNAs. The percentage of total nonstructural reads is shown. Abundance was calculated as the vsiRNA reads per million miRNAs.

3.5 Discussion

We found that four *C. elegans* mutants selected for mapping because of their contrasting antiviral and exogenous RNAi phenotypes were all caused by independent loss-of-function mutations in DRH-1. Whereas the *drh-1(ucr-3)* missense allele (A72T) is located in the worm-specific N-terminal domain, the *drh-1(ucr-4)* missense allele alters a residue conserved within the RD domain of all *C. elegans* DRHs and mammalian RIG-I/MDA5s. The *drh-1(ucr6)* nonsense mutation results in a premature translational termination of DRH-1 to remove the entire RD domain. Similarly, use of an alternative

intron splice site induced by the *drh-1(ucr2)* lesion leads to an in-frame deletion of 30 amino acids in the Hel2i domain conserved in DRAs. Isolation of the four distinct alleles of *drh-1* from our genetic screen further supports a central role of the host protein in antiviral RNAi indicated by previous studies (24, 29, 30).

We discovered an abundant population of terminal vsiRNAs in distinct *drh-1* mutants that was undetectable by previous analysis of small RNA libraries with the incomplete genome sequence of OrV (12). These terminal vsiRNAs contain 5'-monophosphates, exhibit size preference in the range of Dicer products and display no strand bias, suggesting that they correspond to primary vsiRNAs processed by DCR-1 from dsRNA. Examination of the published libraries of small RNAs from *C. elegans* double mutants infected with OrV further reveals that the terminal vsiRNAs are RDE-4 dependent, but independent of RDE-1 and DRH-3, indicating that they indeed follow the biogenesis pathway of the primary vsiRNAs. Our finding resolves the discrepancy between the results of published studies (9,12) and illustrates that DRH-1 is not indispensable for the biogenesis of all primary vsiRNAs. Enrichment of the 5' terminal primary vsiRNAs was reproducibly detected for both of the genomic RNAs and in five distinct *drh-1* mutants as well as in *drh-1;rde-1* and *drh-1;drh-3* double mutants. By comparison, the 3'-terminal vsiRNAs were markedly enriched only for OrV RNA2 in some of the *drh-1* mutants, but not in *drh-1* animals when only Dicer products were examined. These findings are consistent with a previously proposed model in which the viral dsRNA replicative intermediates formed between the 5′-terminal nascent progeny (+)RNA and the (–)RNA template are the main precursor of vsiRNAs (40, 41).

In contrast to *drh-1* mutants, we found that primary vsiRNAs targeting the internal regions of the viral genomic RNAs were as abundant as the 5' terminal vsiRNAs in all of the examined *C. elegans* strains that encode a functional DRH-1. These strains include wildtype N2 animals that produce low levels of primary vsiRNAs possibly because the synthesis of the viral dsRNA precursors is potently inhibited by antiviral RNAi. Also *C. elegans* mutants (e.g., *rde-1* and *drh-3*) that are defective in secondary vsiRNA biogenesis, support efficient virus RNA replication, and accumulate highly abundant primary vsiRNAs. Notably, the abrupt decrease in the density of primary vsiRNAs to target the internal regions of the viral genomic RNAs found in *drh-1* animals was also not detected in *rde-4* mutant animals even though RDE-4 plays a role in the biogenesis of primary siRNAs from viral and injected dsRNA precursors (7). These findings indicate a new function of DRH-1 in the Dicer-mediated production of primary vsiRNAs to target the internal regions of the viral dsRNA precursors. This new function of DRH-1 may explain why DRH-1 exists in a complex *in vivo* with DCR-1 and RDE-4 (7).

We show that the new function of DRH-1 was genetically inactivated by independent, EMS-induced mutations in the ATPase and RD domains highly conserved in DRAs, including RIG-I (2, 42). Given that RIG-I can translocate on long dsRNA powered by ATP hydrolysis (43), we propose that a similar translocase activity of DRH-1 is required to facilitate the translocation of the DCR-1 complex from the 5'-termini to the internal and 3' regions of the viral dsRNA precursors (Fig. 3-7). In the case of exo-RNAi, 3'-terminal dsRNA would be readily available even in the absence of DRH-1and the

translocase function of DRH-1 may not be essential (Fig. 3-7). Given a bias for 22G-RNA production upstream of the trigger (44–47), it is possible that production of primary vsiRNAs to target the 3'-terminal and/or internal regions is necessary to initiate potent RNAi in *C. elegans*.

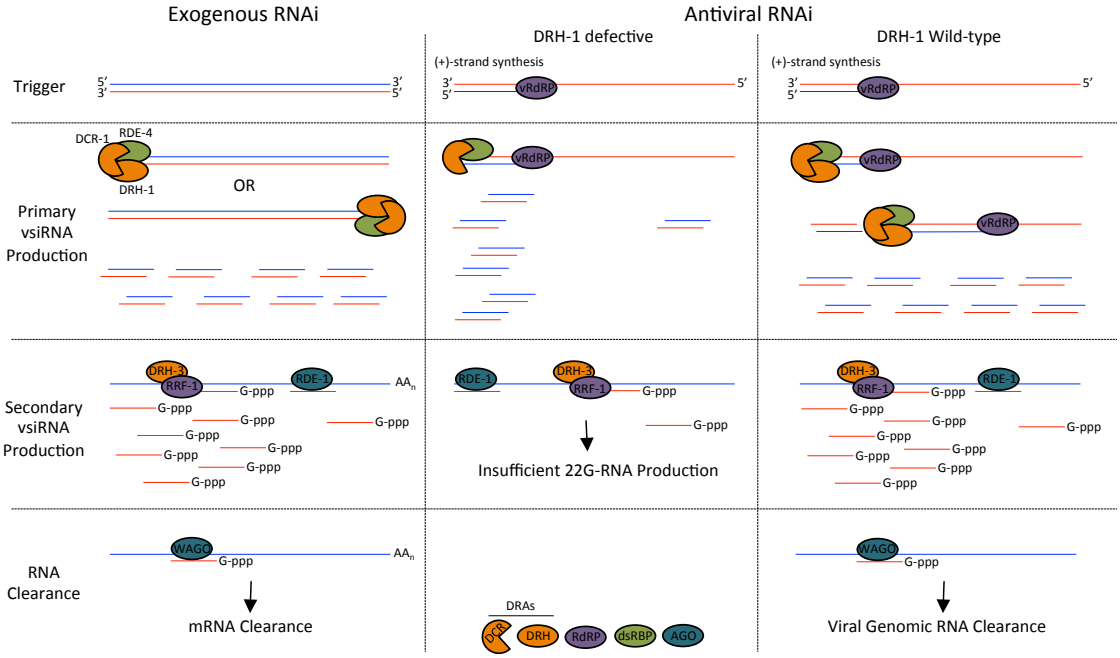

Figure 3-7. Model of Antiviral RNAi in *C. elegans*. Exogenous RNAi in *C. elegans* is triggered by long dsRNA, where as antiviral RNAi is triggered by dsRNA-replication intermediates. The production of primary siRNAs does not require the function of DRH-1, while DRH-1 function in essential for primary vsiRNA production. In DRH-1(-) animals, primary vsiRNAs are not produced from the internal or 3'-terminal regions of viral RNA. Primary vsiRNAs from the 5' terminal regions are unable to mount an efficient antiviral response mediated by 22G-RNAs. In DRH-1(+) animals, primary vsiRNAs are made along the entire viral RNA. We propose a model by which DRH-1 is required for the production of primary vsiRNAs from the internal and 3' terminal regions of viral dsRNA. In exogenous RNAi, this function of DRH-1 would be dispensable because the DCR-1 complex has equal access to the 5' and 3' ends. Primary siRNAs from the 3' end would be sufficient to mount efficient 22G-RNA production.

3.5 References

1. Singleton MR, Dillingham MS, Wigley DB (2007) Structure and mechanism of helicases and nucleic acid translocases. *Annu Rev Biochem* 76:23–50.

2. Luo D, Kohlway A, Pyle AM (2013) Duplex RNA activated ATPases (DRAs): platforms for RNA sensing, signaling and processing. *RNA Biol* 10(1):111–20.

3. Zou J, Chang M, Nie P, Secombes CJ (2009) Origin and evolution of the RIG-I like RNA helicase gene family. *BMC Evol Biol* 9:85.

4. Rawling DC, Kohlway AS, Luo D, Ding SC, Pyle AM (2014) The RIG-I ATPase core has evolved a functional requirement for allosteric stabilization by the Pincer domain. *Nucleic Acids Res* 42(18):11601–11.

5. Sun Y (1997) RIG-I, a human homolog gene of RNA helicase, is induced by retinoic acid during the differentiation of acute promyelocytic leukemia cell. Dissertation (Shanghai Second Medical Univeristy).

6. Zhang X, Wang C, Schook LB, Hawken RJ, Rutherford MS (2000) An RNA helicase, RHIV-1, induced by porcine reproductive and respiratory syndrome virus (PRRSV) is mapped on porcine chromosome 10q13. *Microb Pathog* 28(5):267–278.

7. Tabara H, Yigit E, Siomi H, Mello CC (2002) The dsRNA binding protein RDE-4 interacts with RDE-1, DCR-1, and a DExH-Box helicase to direct RNAi in C. elegans. *Cell* 109(7):861–871.

8. Yoneyama M, et al. (2004) The RNA helicase RIG-I has an essential function in double-stranded RNA-induced innate antiviral responses. *Nat Immunol* 5(7):730–737.

9. Reikine S, Nguyen JB, Modis Y (2014) Pattern Recognition and Signaling Mechanisms of RIG-I and MDA5. *Front Immunol* 5(July):1–7.

10. Yoneyama M, Onomoto K, Jogi M, Akaboshi T, Fujita T (2015) Viral RNA detection by RIG-I-like receptors. *Curr Opin Immunol* 32:48–53.

11. Goubau D, et al. (2014) Antiviral immunity via RIG-I-mediated recognition of RNA bearing 5′-diphosphates. *Nature* 514(7522):372–375.

12. Weber M, et al. (2015) Influenza virus adaptation PB2-627K modulates nucleocapsid inhibition by the pathogen sensor RIG-I. *Cell Host Microbe* 17(3):309–19.

13. Sato S, et al. (2014) The RNA Sensor RIG-I Dually Functions as an Innate Sensor and Direct Antiviral Factor for Hepatitis B Virus. *Immunity* 42(1):123–132.

14. Yao H, et al. (2015) ATP-Dependent Effector-like Functions of RIG-I-like Receptors. *Mol Cell* 58:1–8.

15. Berke IC, Modis Y (2012) MDA5 cooperatively forms dimers and ATP-sensitive filaments upon binding double-stranded RNA. *EMBO J* 31(7):1714–1726.

16. Yoneyama M, et al. (2005) Shared and unique functions of the DExD/H-box helicases RIG-I, MDA5, and LGP2 in antiviral innate immunity. *J Immunol* 175(5):2851–2858.

17. Deddouche S, et al. (2014) Identification of an LGP2-associated MDA5 agonist in picornavirus-infected cells. *Elife* 2014(3):1–20.

18. Bruns AM, Leser GP, Lamb RA, Horvath CM (2014) The Innate Immune Sensor LGP2 Activates Antiviral Signaling by Regulating MDA5-RNA Interaction and Filament Assembly. *Mol Cell* 55(5):771–781.

19. Venkataraman T, et al. (2007) Loss of DExD/H box RNA helicase LGP2 manifests disparate antiviral responses. *J Immunol* 178(10):6444–6455.

20. Satoh T, et al. (2010) LGP2 is a positive regulator of RIG-I- and MDA5-mediated antiviral responses. *Proc Natl Acad Sci U S A* 107(4):1512–1517.

21. Suthar MS, et al. (2012) The RIG-I-like receptor LGP2 controls CD8(+) T cell survival and fitness. *Immunity* 37(2):235–48.

22. Welker NC, et al. (2010) Dicer's helicase domain is required for accumulation of some, but not all, C. elegans endogenous siRNAs. *Rna* 16(5):893–903.

23. Welker NC, et al. (2011) Dicer's Helicase Domain Discriminates dsRNA Termini to Promote an Altered Reaction Mode. *Mol Cell* 41(5):589–599.

24. Guo X, Zhang R, Wang J, Ding S-W, Lu R (2013) Homologous RIG-I-like helicase proteins direct RNAi-mediated antiviral immunity in C. elegans by distinct mechanisms. *Proc Natl Acad Sci U S A* 110(40):16085–90.

25. Li Y, Lu J, Han Y, Fan X, Ding S-W (2013) RNA interference functions as an antiviral immunity mechanism in mammals. *Science* 342(6155):231–4.

26. Maillard P V, et al. (2013) Antiviral RNA interference in mammalian cells. *Science* 342(6155):235–8.

27. Fitzgerald ME, Vela A, Pyle AM (2014) Dicer-related helicase 3 forms an obligate dimer for recognizing 22G-RNA. *Nucleic Acids Res* 42(6):3919–3930.

28. Duchaine TF, et al. (2006) Functional proteomics reveals the biochemical niche of C. elegans DCR-1 in multiple small-RNA-mediated pathways. *Cell* 124(2):343–354.

29. Lu R, Yigit E, Li WX, Ding SW (2009) An RIG-I-like RNA helicase mediates antiviral RNAi downstream of viral siRNA biogenesis in Caenorhabditis elegans. *PLoS Pathog* 5(2):8–14.

30. Ashe A, et al. (2013) A deletion polymorphism in the Caenorhabditis elegans RIG-I homolog disables viral RNA dicing and antiviral immunity. *Elife* 2013(2):1–21.

31. Gu W, et al. (2009) Distinct Argonaute-Mediated 22G-RNA Pathways Direct Genome Surveillance in the C. elegans Germline. *Mol Cell* 36(2):231–244.

32. Brenner S (1974) The genetics of Caenorhabditis elegans. *Genetics* 77(1):71–94.

33. Jiang H, Lei R, Ding S-W, Zhu S (2014) Skewer: a fast and accurate adapter trimmer for next-generation sequencing paired-end reads. *BMC Bioinformatics* 15(1):182.

34. Jiang H, et al. (2014) Orsay virus utilizes ribosomal frameshifting to express a novel protein that is incorporated into virions. *Virology* 450-451:213–221.

35. Guo X, Zhang R, Wang J, Lu R (2013) Antiviral RNA silencing initiated in the absence of RDE-4, a double-stranded RNA binding protein, in Caenorhabditis elegans. *J Virol* 87(19):10721–9.

36. Félix MA, et al. (2011) Natural and experimental infection of Caenorhabditis nematodes by novel viruses related to nodaviruses. *PLoS Biol* 9(1). doi:10.1371/journal.pbio.1000586.

37. Ashe A, Sarkies P, Le Pen J, Tanguy M, Miska EA (2015) Antiviral RNAi against Orsay virus is neither systemic nor transgenerational in Caenorhabditis elegans. *J Virol*:JVI.03664–14.

38. Sarkies P, Miska E a (2013) RNAi pathways in the recognition of foreign RNA: antiviral responses and host-parasite interactions in nematodes. *Biochem Soc Trans* 41(4):876–880.

39. Ding SW, Lu R (2011) Virus-derived siRNAs and piRNAs in immunity and pathogenesis. *Curr Opin Virol* 1(6):533–544.

40. Aliyari R, et al. (2008) Mechanism of Induction and Suppression of Antiviral Immunity Directed by Virus-Derived Small RNAs in Drosophila. *Cell Host Microbe* 4(4):387–397.

41. Ding S-W (2010) RNA-based antiviral immunity. *Nat Rev Immunol* 10(9):632–644.

42. Ahmad S, Hur S (2015) Helicases in Antiviral Immunity : Dual Properties as Sensors and Effectors. *Trends Biochem Sci* 40(10):576–585.

43. Myong S, et al. (2009) Cytosolic Viral Sensor RIG-I Is a5'-triphosphate-dependent translocase on double-stranded RNA. *Science (80-)* 323:1070–1074.

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

45. Sijen T, et al. (2001) On the role of RNA in gene amplification. *Cell* 107:465–476.

46. Pak J, Maniar JM, Mello CC, Fire A (2012) Protection from Feed-Forward Amplification in an Amplified RNAi Mechanism. *Cell* 151(4):885–899.

47. Sapetschnig A, Sarkies P, Lehrbach NJ, Miska E a. (2015) Tertiary siRNAs Mediate Paramutation in C. elegans. *PLOS Genet* 11(3):e1005078.

CHAPTER 4

CONCLUSIONS AND FUTURE DIRECTIONS

4.1 Conclusions

The FR1gfp viral replicon launched from a single copy transgene in *C. elegans* is readily visualized as GFP only when the host immune system has been compromised, making this system ideal for dissecting the genetic requirements of antiviral RNAi. Through the establishment of an organism-level unbiased genetic screen, I isolated thirteen mutants that failed to suppress FR1gfp replication. Notably, all thirteen mutants failed to suppress OrV infection, establishing the viral replicon-based screen as an appropriate method for the isolation of mutants that are defective in antiviral immunity. The overlap between the genetic requirements of antiviral RNAi with exo-RNAi led me to sort the mutants based their responses to exogenous dsRNA. Eight of the mutants were capable of RNAi against both somatic and germlines genes. These eight mutants represent genes that are uniquely required for antiviral RNAi, such as DRH-1, or may lead to the discovery of novel antiviral pathways in *C. elegans*. Four mutants were resistant to both somatic and germline RNAi, while the final mutant, *ucr12*, was resistant to germline RNAi, but not RNAi against somatic genes. While I was able to demonstrate that *ucr11* corresponds to a nonsense mutation in *rde-1*, the remaining RDE alleles may

identify novel genes required for exo- and antiviral RNAi

Previous EMS screens using FR1gfp in *C. elegans* were stalled at the mapping stage. The traditional mapping strain Hawaiian was found to be susceptible to FR1gfp replication and the causal mutation could not be reliably followed after outcrossing. Because wild isolates varies significantly in the potency of their RNAi responses, as well as their susceptibility to viral infection (1), I opted to map the alleles using a method that relies on backcrosses to the parental strain that was pioneered in plants (2, 3). Using this method, I was able to map four novel and unique alleles of *drh-1*, recapitulating the important role that DRH-1 plays in antiviral RNAi. Additionally, I mapped an allele of *mys-2* using the same approach, demonstrating that this mapping method can isolate novel genes as well.

Next, I analyzed the primary and secondary vsiRNAs in the four *drh-1* alleles and obtained results that varied from published vsiRNA profiles (4), most notably that the *drh-1* mutants had more 23-nt vsiRNAs in my analysis. At the time of the initial analysis the OrV reference genome was missing 492-nt and 248-nt from the 5' and 3' end of RNA 1 and 51-nt and 161-nt from the 5' and 3' ends of RNA 2, suggesting that the 23-nt vsiRNAs in my mutants came from these terminal regions. Analysis of 23-nt vsiRNA distribution along the OrV genome supported this hypothesis. I re-analyzed the vsiRNA profiles that were previously published (4) and found that 23-nt vsiRNAs in *drh-1* mutants were derived from the 5' terminal region of OrV RNA1 and RNA2 and sometimes the 3' end of OrV RNA2. 23-nt vsiRNAs form the internal and 3' terminal regions were depleted in the absence of DRH-1. To more precisely analyze the DCR-1

dependent primary vsiRNAs, I analyzed libraries that relied on a 5'-dependent cloning protocol to only sequence mono-phosphate small RNAs (4). Analysis of the mono-phosphate vsiRNAs in *drh-1(ok3495)* demonstrated that the 5'-terminal vsiRNAs produced in the absence of DRH-1 are bona fide primary vsiRNAs dependent on DCR-1. Additionally, analysis of *drh-1(ok3495);rde-1* and *drh-1(ok3495);drh-3* supports a role for DRH-1 upstream of RDE-1 and DRH-3. In contrast, analysis of *drh-1(ok3495);rde-4* revealed that the 5'-terminal primary vsiRNAs are also dependent on RDE-4. Taken together these results suggest that in the absence of DRH-1, primary vsiRNAs are made from viral dsRNA replication intermediates by DCR-1 and RDE-4. DRH-1 belongs to a family of translocases and is homologous to mammalian RIG-I. Therefore, I propose a model for DRH-1 function in which DRH-1 promotes production of internal primary vsiRNAs using translocase activity. Due to a bias of 22-G RNA production upstream of the trigger, 5'-terminal primary vsiRNAs would be insufficient to promote an effective secondary vsiRNA response. This function also explains why DRH-1 is only required for antiviral RNAi and is dispensable for exo-RNAi, as exogenous dsRNA is equally abundant at the 3' end and primary siRNAs from the 3' end would be sufficient to produce an effective secondary vsiRNA response.

PAMPS derived from RNA viruses are commonly dsRNA derived from either secondary structures in the genomic viral RNA or dsRNA produced during RNA replication. Three lines of evidence support the biogenesis of primary vsiRNAs from viral dsRNA replication intermediates rather than secondary RNA structures in *C. elegans*. The first is the 50/50 sense and antisense ratio in 5'-dependent wild-type

libraries. Dicer processing of dsRNA produces vsiRNAs that are equally sense and antisense, whereas Dicer processing of secondary structures results in vsiRNAs that are of the sense polarity. The second is the production of abundant antisense 22G-RNAs using the (+)-sense genomic RNA as a template. Sense vsiRNAs derived from secondary structures would be unable to target the (+)-sense genomic RNA strand for 22G-RNA production. The third is the production of 5'-terminal primary vsiRNAs in the absence of DRH-1. During viral replication 5'-terminal dsRNA results from (+)-strand synthesis, which is much more abundant that 3'-terminal dsRNA that results from (-)-strand synthesis. Taken together, these results suggest that vsiRNAs in *C. elegans* are derived from dsRNA replication intermediates in response to at least RNA viruses, OrV and FHV.

4.2 Continued Screening to isolate mutants defective in antiviral immunity

The screen developed in this dissertation demonstrates that the FR1gfp viral replicon launched from a single copy insertion is suitable for identifying alleles required for antiviral RNAi during natural viral infection by Orsay. After completing 10 rounds of screening, I adapted the screen for continued screening by undergraduate students. During my time in the lab, I taught three undergraduates, two seniors and one sophomore, to conduct the EMS screen and characterize the acquired mutants. A forward genetic screen using *C. elegans* is an ideal project for undergraduate students. Students learn about classical genetics principles and techniques, as well as more modern principles, such as GFP and RNAi. One round of screening takes 8 days, followed by verification

and characterization of OrV and exo-RNAi susceptibility. A promising undergraduate can isolate and characterize at least one mutant during a 10-week quarter. Additionally, most days of screening require less than an hour of work, making it ideal for fitting around class schedules. The initial screen utilized Northern Blots to detect OrV infection, which requires radioactive training and is inappropriate for undergraduates. Instead, OrV infection is easily detectable by a forgiving and rewarding small molecule fluorescent *in situ* hybridization protocol (smFISH) perfect for undergraduates that was developed by Gina Broitman-Maduro in the Maduro Lab at UCR. OrV infection can also be detected using qRT-PCR. A detailed protocol intended for undergraduates can be found in Appendix J. This protocol could easily be adapted into a senior-level laboratory exercise.

During my initial screen, most of the F2s that were isolated were not viable (Appendix C). These mutants could be isolated and identified using a clonal F1 screen (5). In this type of screen, F1s are cloned to individual plates instead of sharing a large plate. The F2s would be heat treated and screened for GFP expression. F2 populations that demonstrate 25% lethal F2s that allowed replication of FR1gfp would indicate the presence of a lethal allele required for FR1gfp suppression. The lethal mutation can then be retrieved from the heterozygous siblings and maintained by examining the progeny every generation. Once a lethal mutation known to reproducibly fail to suppress FR1gfp is isolated, it can be maintained using a balancer chromosome (6, 7). Some components of the core RNAi machinery, such as DCR-1, are not viable and would provide valuable information of the overlap between the genetic requirements of endogenous RNAi and

antiviral RNAi. Essential developmental factors can also be repurposed to have roles in antiviral immunity later in the life cycle and could be lethal (8).

The results from mapping the alleles of *drh-1* and *mys-2* suggest that similar results can be obtained from whole genome sequencing of only the mutant pools. This method would eliminate the need to score the F3 population and FR1gfp replication in the F2s would suffice. Additionally, it would cut the cost of sequencing in half. Several libraries were prepared using this method, mostly by Dr. Xunyang Guo who continued to the screen and isolated several RNAi proficient alleles. The candidates obtained for these alleles can be found in Appendix I and await verification and characterization.

4.3 Characterize *mys-2(ucr7)*

In addition to identifying four novel alleles of *drh-1,* this work also verified an allele of *mys-2*, a MYST histone acytl-transferase. Though they await validation, four other alleles isolated by myself and Dr. Xunyang Guo contain candidates in *mys-2* and a subunit of the MYS-2 complex, SUMV-2. MYS-2 is the homolog of the mammalian KAT8 histone acetyltransferase that is the primarily responsible for H4K16Ac, a marker of active transcription (9). Interestingly, MYS-2 and SUMV-2 have been implicated in transgene silencing via RNAi (10, 11); however the role of histone marks that activate transcription in gene silencing have not been explored. Additionally, available evidence suggests that antiviral RNAi is not inherited (12). It is yet to be determined if MYS-2 functions in antiviral RNAi or a novel innate immunity pathway, but small RNA sequencing does not show an obvious defect in small RNAs. However, the OrV-derived

vsiRNAs in *mys-2* are much more abundant than in other mutants such as *drh-1*, even though they do not support as high a replication levels of OrV genomic RNA. This defect may suggest that the vsiRNAs made in *mys-2* are only partially functional.

In mammals, MYS-2 forms two complexes MOF-MSL and MOF-MSLv1 that acetylate histone and other proteins, respectively. The MOF-MSL complex consist of MSL1, MSL2 and MSL3, while MOF-MSLv1 consists of NSL1, NSL2, NSL3, MCSR1, PHF20, OGTI, WDR5 and HCF1 (13). These subunits are conserved in *C. elegans* and the first question I would ask about the function of MYS-2 in antiviral immunity is whether it requires acetylation of histones or other proteins. Using feeding RNAi or genetic mutants, the subunits would be tested for their ability to suppress FR1gfp replication and OrV infection. Identifying the genetic requirements of MYS-2 mediated antiviral immunity will suggest future experiments. For example, if the results suggest that histone acetylation is required, then the transcriptome and H4K16Ac profile could be examine using next generation sequencing under conditions of viral infection in *mys-2(ucr7)* and wild-type animals. Even if the antiviral phenotype observed in *mys-(ucr7)* is indirectly caused by suppression of antiviral genes, transcriptome analysis can identify these genes.

4.4 Investigate the translocation ability of *C. elegans* DRAs

All three *C. elegans* DRAs are essential to mount an effective antiviral RNAi response. It would be interesting to learn if all three *C. elegans* DRAs have bona fide translocase activity that is required for antiviral RNAi. Translocation activity of purified

proteins can be monitored *in vitro* by Protein-induced Fluorescent Enhancement (PIFE) (14, 15). PIFE monitors the intensity of a single fluorophore attached to a substrate, such as dsRNA. The intensity of the fluorophore increases upon binding of a protein in its proximity. PIFE does not require labeling the protein, so several purified proteins, such as DCR-1, DRH-1 and DRH-3 can be assayed using the same labeled substrate. During translocation, the intensity of fluorophore will oscillate, peaking repeatedly as proteins bind (14). In the absence of ATP, RIG-I binds to the dsRNA substrate only once, whereas in the presence of ATP, RIG-I repeatedly bind the molecule indicating translocation (14). Mutant DRH-1 proteins purified from the *drh-1* mutants acquired in this study, such as *ucr4* that alters a single amino acid in the RD domain, could also be tested to learn about which amino acids or protein domains are required for translocation. Because *C. elegans* requires three DRAs to mediate an effective antiviral response, all three may not contain efficient translocase activity. If *C. elegans* DCR-1 has inefficient translocase activity then this would explain the requirement of DRH-1 to target internal regions of viral RNA for primary vsiRNA production.

4.5 Further examine the role of DRH-1 in antiviral RNAi

In the absence of DRH-1, primary vsiRNAs corresponding to the internal and 3' regions of the viral RNA are depleted, while primary vsiRNAs from the 5' terminal region are unaffected, indicating that the primary vsiRNAs from the 5' end are defective in initiating an effective antiviral response. Here we consider two possibilities to explain why primary vsiRNAs produced from 5' terminal regions are insufficient to initiate

RNAi. While my model presumes that the 5'-terminal primary vsiRNAs are insufficient to promote effective production of viral 22G-RNAs because of a bias for upstream production of 22G-RNAs, we cannot rule out the possibility that DRH-1 functions in loading the primary vsiRNAs into the argonaute protein, RDE-1. In order to answer this question, I would clone and sequence RDE-1 associated small RNAs from *drh-1* mutants. RDE-1 associated small RNAs have been immunoprecipitated from other mutant backgrounds (16, 17). Briefly, a HA or GFP tagged RDE-1 protein could be expressed in *drh-1;rde-1* mutant backgrounds and in the *rde-1* background, both infected with OrV. RNA extractions from the immunoprecipitated RDE-1 protein would then be sequenced using a 5'-dependent cloning protocol to sequence only the primary, mono-phosphate, small RNAs. The primary vsiRNAs could be analyzed for length, 5'-nucleotide identity and polarity and mapped to the OrV genome. If DRH-1 is not required for RDE-1 loading, then we would expect to find primary vsiRNAs produced form the 5' terminal regions in the RDE-1 co-IP in *drh-1* mutants. In the control, we would not expect to find enrichment at the 5' ends. If DRH-1 is required for RDE-1 loading, then we would not find primary vsiRNAs in the RDE-1 co-IP in *drh-1* mutants. I predict that results from a RDE-1 co-IP would support my model and that DRH-1 is not required for RDE-1 loading, because this function would be essential for exo-RNAi, as well as antiviral RNAi.

In order to test if the primary vsiRNAs in *drh-1* mutants are defective because they are derived from the 5' terminal region, I proposed compensating the lack of available internal and 3' end dsRNA by providing synthetic dsRNA targeting the internal

and 3' terminal regions of viral genomic RNA. Synchronized wild-type and *drh-1* mutants would be fed HT115 *E. coli* expressing dsRNA against (1) internal region of FR1gfp and (2) 3'-terminal region of FR1gfp. Animals would be heat shocked at the L4 stage and analyzed forty eight hours later for GFP expression. If the dsRNA against the internal or 3' regions of FR1gfp is able to suppress replication of FR1gfp in *drh-1* mutants, then the failure of *drh-1* to suppress FR1gfp is dependent on the lack of primary vsiRNAs from the internal and/or 3' regions, supporting my model. As a control, I would repeat the same experiment in *rde-4* and *rde-1* animals to ensure that the dsRNA is not acting as an immunization against FR1gfp and that the rescue is dependent on compensating for DRH-1 function. A similar experiment could be attempted with OrV as well; however in this case, the viral replicon allows for delayed induction of viral replication and offers more robust replication. Wild-type and *drh-1* mutants challenged with OrV could be fed HT115 *E. coli* expressing dsRNA against the 3' end of either OrV RNA1 or RNA2 or both. Similar to the FR1gfp experiment, the ability of dsRNA from the internal or 3' terminal regions of OrV to rescue the OrV susceptibility of *drh-1* mutants would support my model.

The proline that is changed to a leucine in *drh-1(ucr4)* is located in the Regulatory Domain and is conserved in all of the RIG-I and RIG-I like proteins. Crystal structure of the RIG-I RD Domain indicates that this amino acid, P^{885}, is in the β8 sheet, which along with β5, β6 and β7, make up one of the three leaves that structurally comprise the flat RD domain (18). However, the function of this amino acid is unknown. The RD domain of RIG-I can functionally replace the RD domain of DRH-1 to mediate viral clearance by

RNAi (19). In order to determine if P^{885} is required for the function of the mammalian RIG-I RD domain, I modified this chimeric construct to change P^{885} to a Leucine, called D1RIG-1 P885L. As a control, I modified the DRH-1 rescue plasmid described in Chapter 2 to change P^{966} of DRH-1 to a Leucine, called DRH-1 P966L. If ectopic expression of D1RIG-1 P885L fails to rescue suppression of FR1gfp, then the proline identified in *C. elegans* may play a fundamental role in the function of mammalian RIG-I.

4.6 Examine the role of DRH-1 in exo-RNAi

DRH-1 is found in the DCR-1/RDE-4 complex in the absence of viral infection, but every known *drh-1* allele is dispensable for exo-RNA (20). In response to viral infection, I propose a model by which DRH-1 is required for the production of primary vsiRNAs from the internal region of viral dsRNA replication intermediates. In the absence of DRH-1, primary vsiRNAs are made from the 5'-terminal regions of dsRNA replication intermediates, where dsRNA is the most abundant and easily accessible. However, because of a bias for 22G-RNAs to be made upstream of primary siRNAs, 5'-terminal vsiRNAs would be insufficient to mount an antiviral response. If this model is correct, then DRH-1 could also be required for the production of internal siRNAs from dsRNA to trigger exo-RNAi. However, the production of internal primary siRNAs from long dsRNA would not be essential for effective exo-RNAi, because the DCR-1/RDE-4 complex would have equal access to the 5' and 3' ends for production of primary siRNAs. Primary vsiRNAs processed form the 3'-end of the dsRNA would be suffice to

target the mRNA molecules for 22G-RNA production and subsequent clearance. To test this hypothesis, I would use a 5'-dependent small RNA sequencing method to sequence the monophosphate small RNAs from *drh-1* mutant animals undergoing exo-RNAi against a somatic gene and a germline gene, such as *unc-22* and *pop-1*. Alignment of the dsRNA-derived small RNAs would determine if DRH-1 is required for the production of primary small RNAs from the internal regions of long dsRNA. I expect to find that in the absence of DRH-1, siRNAs targeting the internal regions, but not the 5' or 3' terminal regions, will be depleted.

4.7 References

1. Félix MA, et al. (2011) Natural and experimental infection of Caenorhabditis nematodes by novel viruses related to nodaviruses. *PLoS Biol* 9(1). doi:10.1371/journal.pbio.1000586.

2. Zhu Y, et al. (2012) Gene discovery using mutagen-induced polymorphisms and deep sequencing: application to plant disease resistance. *Genetics* 192(1):139–46.

3. Abe A, et al. (2012) Genome sequencing reveals agronomically important loci in rice using MutMap. *Nat Biotechnol* 30(2):174–178.

4. Ashe A, et al. (2013) A deletion polymorphism in the Caenorhabditis elegans RIG-I homolog disables viral RNA dicing and antiviral immunity. *Elife* 2013(2):1–21.

5. Jorgensen EM, Mango SE (2002) The art and design of genetic screens: caenorhabditis elegans. *Nat Rev Genet* 3(5):356–369.

6. Herman RK (1978) Crossover suppressors and balanced recessive lethals in Caenorhabditis elegans. *Genetics* 88:49–65.

7. Rosenbluth RE, Baillie DL (1981) The genetic analysis of a reciprocal translocation, eT1(III; V), in Caenorhabditis elegans. *Genetics* 99(3-4):415–28.

8. Lemaitre B, Nicolas E, Michaut L, Reichhart JM, Hoffmann JA (1996) The dorsoventral regulatory gene cassette spatzle/Toll/Cactus controls the potent antifungal response in Drosophila adults. *Cell* 86(6):973–983.

9. Yang Y, Han X, Guan J, Li X (2014) Regulation and function of histone acetyltransferase MOF. *Front Med* 8(1):79–83.

10. Cui M, Kim EB, Han M (2006) Diverse Chromatin Remodeling Genes Antagonize the Rb-Involved SynMuv Pathways in C. elegans. *PLoS Genet* 2(5):e74.

11. Vastenhouw NL, et al. (2006) Long-term gene silencing by RNAi. *Nature* 442:881.

12. Ashe A, Sarkies P, Le Pen J, Tanguy M, Miska EA (2015) Antiviral RNAi against Orsay virus is neither systemic nor transgenerational in Caenorhabditis elegans. *J Virol*:JVI.03664–14.

13. Cai Y, et al. (2010) Subunit composition and substrate specificity of a MOF-containing histone acetyltransferase distinct from the Male-specific Lethal (MSL) complex. *J Biol Chem* 285(7):4268–4272.

14. Myong S, et al. (2009) Cytosolic Viral Sensor RIG-I Is a5'-triphosphate-dependent translocase on double-stranded RNA. *Science (80-)* 323:1070–1074.

15. Hwang H, Myong S (2014) Protein induced fluorescence enhancement (PIFE) for probing protein–nucleic acid interactions. *Chem Soc Rev* 21(43):1221–1229.

16. Corrêa RL, Steiner F a, Berezikov E, Ketting RF (2010) MicroRNA-directed siRNA biogenesis in Caenorhabditis elegans. *PLoS Genet* 6(4):e1000903.

17. Shirayama M, Stanney W, Gu W, Seth M, Mello CC (2014) The Vasa homolog RDE-12 engages target mRNA and multiple argonaute proteins to promote RNAi in C. elegans. *Curr Biol* 24(8):845–851.

18. Cui S, et al. (2008) The C-Terminal Regulatory Domain Is the RNA 5′-Triphosphate Sensor of RIG-I. *Mol Cell* 29(2):169–179.

19. Guo X, Zhang R, Wang J, Ding S-W, Lu R (2013) Homologous RIG-I-like helicase proteins direct RNAi-mediated antiviral immunity in C. elegans by distinct mechanisms. *Proc Natl Acad Sci U S A* 110(40):16085–90.

20. Lu R, Yigit E, Li WX, Ding SW (2009) An RIG-I-like RNA helicase mediates antiviral RNAi downstream of viral siRNA biogenesis in Caenorhabditis elegans. *PLoS Pathog* 5(2):8–14.

APPENDICES

Appendix A. Worm Medias and Buffers

Nematode Growth Media (NGM)

3 Liters

	Thin Lawn	Thick Lawn
Agar	51g	60g
NaCl	9g	9g
Peptone	7.5g	7.5g
2g/L Uracil	3mL	30mL
$CaCl_2$	.44g	.44g
10g/L Cholesterol (in EtOH)	1.5mL	1.5mL
Dextrose	N/A	6.1g

Autoclave, let cool and then add:

Phosphate Buffer*	75mL
1M $MgSO_4$	3mL

*1M KH_2PO_4 pH 6.0. Combine 5 parts Buffer A with 2 parts Buffer B (Buffer A = 1M KH_2PO_4, Buffer B = 1M K_2HPO_4)

For Feeding RNAi make thin lawn plates and add:

1M IPTG	3mL
100mg/mL Carbenicillin	3mL

M9 Buffer

Two Liters

	Amount
KH_2PO_4	6g
$Na_2HPO_4.7H_2O$	22.6g (11.4g anhydrous)
NaCl	7.5g
1M $MgSO_4$	2mL*

*add after everything else is in solution

DNA extraction buffer for single worm PCR

	Amount
10x PCR Buffer (100 mM Tris, 500 mM KCl, 20 mM MgCl2 pH 8.)	50µL
ddH_2O	445µL
20mg/µL Proteinase K	5µL

Appendix B. Buffers for Northern Blot Analysis

Northern Blot Buffers	
RNA Gel	1.2% (w/v) agarose, 1.8% (v/v) Formaldehyde in 1X MOPS buffer
10X MOPS Buffer	0.5M MOPS, 0.01M EDTA, pH 7.0
6X Loading Buffer	50% (v/v) glycerol, 0.25% (w/v) bromophenol blue, 0.25% (w/v) xylene cyanol FF
2X Formamide Loading Buffer	2 parts Formamide, 1 part 6X loading Buffer
20X SCC	3M NaCL, 0.3M Trisodium citrate
Methylene Blue	0.04% (w/v) Methylene Blue, .5M Sodium Acetate, pH 5.2
Hybridization Buffer	PerfectHyb Plus Hybridization Buffer (Sigma Aldrich)
Wash Buffer	0.01% (w/v) SDS, .1X SSC

Appendix C. Results from EMS mutagenesis of *irSi18*

Round	P0s	F1s	F2s	(+)GFP	Viable	Reproducible (+)-GFP	Susceptible to OrV	*unc-22 (RNAi)*	*pop-1 (RNAi)*
1	100	1,000	20,000	24 (4)	3 (1)	0	N/A	N/A	N/A
2	100	1,000	20,000	55 (5)	4 (3)	0	N/A	N/A	N/A
3	200	2,000	40,000	95 (10)	12 (7)	3	3	3	3
4	140	1,400	28,000	23 (6)	4(3)	2	2	2	2
5	150	1,500	30,000	63(7)	21(6)	4	4	2	2
6	150	1,500	30,000	101(10)	8(5)	1	1	1	1
7	150	1,500	30,000	114(10)	17(10)	0	N/A	N/A	N/A
8	150	1,500	30,000	63(8)	5(4)	0	N/A	N/A	N/A
9	150	1,500	30,000	118(10)	17(7)	2	2	1	0
10	150	1,500	30,000	168(15)	28(11)	1	1	0	0
Total	**1,440**	**14,400**	**288,000**			**13**	**13**	**9**	**8**

Appendix D. Whole Genome Resequencing Libraries

Library	Reads (Paired End)	Reads used in Analysis	Coverage (Expected)*	Coverage (Actual)[§]
ucr2 mutant	21,984,304	21,984,304	22.0	7.1
ucr2 wild-type	141,353,864	141,353,864	141.4	7.8
ucr3 mutant	71,356,804	71,356,804	71.4	71.6
ucr3 wild-type	61,895,606	61,895,606	61.9	62.1
ucr4 mutant	105,113,956	50,000,000	50.0	50.1
ucr4 wild-type	64,690,326	50,000,000	50.0	49.6
ucr5 mutant	71,176,126	71,176,126	71.2	68.4
ucr5 wild-type	43,593,798	43,593,798	43.6	42.1
ucr6 mutant	48,169,648	48,169,648	48.2	47.1
ucr6 wild-type	59,001,946	59,001,946	59.0	58.9
ucr7 mutant	54,277,136	54,277,136	54.3	50.0
ucr7 wild-type	71,593,122	50,000,000	50.0	46.9
ucr10 mutant	42,597,394	42,597,394	42.6	37.6
ucr10 wild-type	42,418,692	42,418,692	42.4	40.4

Expected coverage was calculated as (Reads X 100bp Read Length)/(100Mb Genome Size). Average coverage at each base in the genome, calculated with Bedtools and custom Perl script**.** The large discrepancy between the expected and actual coverage in the *ucr2* libraries was caused by abundant adapter-dimers that were undetectable by bioanalyzer or gel electrophoresis.

Appendix E. Candidates from mapped alleles

ucr3 Mapping Results

	Location	Ref	Mut	Mutant AF	WT AF	Change
1	chrII 12260354	C	T	100		
2	chrIV 4271228	C	T	89	20	F36A4.2 5' UTR
3	chrIV 4449792	C	T	90	0	R08C7.2 3' UTR
4	chrIV 6613075	C	T	100	14	drh-1 Ala -> Thr (aa72)
5	chrIV 7945518	C	T	100	17	czw-1 Val -> Iso (aa317)
6	chrIV 8115409	C	T	97	20	Intragenic
7	chrIV 8318774	C	T	100	0	Intragenic
8	chrIV 8388998	C	T	96	0	Intragenic
9	chrIV 8693703	C	T	100	17	Intragenic
10	chrIV 8873984	C	T	100	22	Intragenic
11	chrIV 10681467	C	T	99	20	Intragenic
12	chrIV 11130931	C	T	91	21	psa-4 3' UTR
13	chrIV 12221484	C	T	92	27	F35G2.5 Intron
14	chrIV 13013870	C	T	97	18	mbk-2 intron
15	chrIV 13093375	C	T	98	24	C39E9.11 Ala -> Ala (aa100)
16	chrIV 13489465	C	T	97	38	tbc-9 Arg -> Arg (aa426)
17	chrIV 14605340	C	T	87	38	Intragenic
18	chrIV 14765535	C	T	88	25	Intragenic
19	chrIV 15050056	C	T	88	40	srt-47 intron
20	chrIV 16271168	C	T	88	30	F38C2.4 intron
21	chrIV 16325147	C	T	82	32	srz-105 Ser -> Phe (aa3)
22	chrIV 16791963	C	T	86	37	Intragenic
23	chrIV 16916049	C	T	79	20	Intragenic
24	chrV1764545	C	T	63	37	

ucr4 Mapping Results

	Location	Ref	Mut	Mutant AF	WT AF	Change
1	chrII 14956568	C	T	82	38	
2	chrIV 3709551	G	A	75	3	Y37E11AL.1 Arg-> Arg(aa14)
3	chrIV 3935434	G	A	92	2	clec-70 Val-> Iso (aa269)
4	chrIV 4271228	C	T	97	7	Y51A2D.28 3 'UTR
5	chrIV 4863257	G	A	88	2	str-173 intron
6	chrIV 5633478	G	A	97	0	Intragenic
7	chrIV 6177776	G	A	94	0	Intragenic
8	chrIV 6342374	G	A	97	0	Intragenic
9	chrIV 6352746	G	A	91	0	Intragenic
10	chrIV 6607879	G	A	83	0	drh-1 Pro -> Leu (aa 996)
11	chrIV 7453991	G	A	97	0	Intragenic
12	chrIV 7942361	G	A	90	0	Intragenic
13	chrIV 7945518	C	T	95	0	czw-1 Val -> Phe (aa317)
14	chrIV 8195003	G	A	91	0	Intragenic
15	chrIV 8318774	C	T	90	0	Intragenic
16	chrIV 8325530	G	A	87	0	Intragenic
17	chrIV 12111533	G	A	75	2	Intragenic
18	chrIV 12577268	C	T	65	2	Intragenic
19	chrIV 17165417	C	T	64	3	Intragenic
20	chrV 254949	C	T	63	39	
21	chrV 1174671	C	T	68	36	
22	chrX 10969351	G	A	74	37	
23	chrX 12730610	G	A	66	40	
24	chrX 14212310	G	A	64	38	

ucr6 Mapping Results

	Location	Ref	Mut	Mutant AF	WT AF	Change
1	ChrIV 3049553	G	A	78	5	eel-1 (aa123)
2	ChrIV 3091320	G	A	82	10	ncRNA
3	ChrIV 4271228	C	T	81	3	F36A4.2 5' UTR
4	ChrIV 4620568	G	A	71	0	Intragenic
5	ChrIV 5296212	G	A	85	2	ZK354.6 Pro->Ser (aa302)
6	ChrIV 5653088	G	A	85	0	skn-1 5' UTR
7	ChrIV 5999473	G	A	83	0	unc-44 Thr->Thr (aa5604)
8	ChrIV 6608233	G	A	75	0	drh-1 Arg->Stop (aa865)
9	ChrIV 6907612	G	A	72	0	rpn-1 Pro->Leu (aa700)
10	ChrIV 7431748	G	A	76	0	Intragenic
11	ChrIV 8304495	G	A	78	0	vha-5 Leu->Leu (aa75)
12	ChrIV 8619739	G	A	72	0	Intragenic
13	ChrIV 9013318	G	A	78	2	Intragenic
14	ChrIV 9212763	G	A	87	0	Intragenic
15	ChrIV 9916702	G	A	70	0	bed-3 Leu->Leu(aa59)
16	ChrIV 10038640	G	A	85	0	Intragenic
17	ChrIV 10132695	G	A	78	0	glt-3 Intron
18	ChrIV 10841181	G	A	77	0	T11G6.5 Tyr->Tyr (aa292)
19	ChrIV 10887238	G	A	70	0	F13E9.16 Pro->Ser (aa43)
20	ChrIV 10897427	G	A	76	0	Intragenic
21	ChrIV 11111193	G	A	77	0	Intragenic
22	ChrIV 11791932	G	A	87	0	ncRNA
23	ChrIV 12086168	G	A	72	0	F11A10.7 Glu->Glu(aa256)
24	ChrIV 12404339	G	A	77	4	Intragenic
25	ChrIV 12785653	C	T	71	10	unc-31 Thr->Ile (aa
26	ChrIV 13167312	C	T	75	12	Intragenic

ucr7 Mapping Results

	Location	Ref	Mut	Mutant AF	WT AF	Change
1	ChrI 12977981	G	A	95	0	gale-1 intron
2	ChrI 13011534	G	A	97	0	Intragenic
3	ChrI 13468565	G	A	92	4	ncRNA
4	ChrI 13660827	G	A	94	0	ncRNA
5	ChrI 14122022	G	A	97	2	mys-2 Gly->Glu (aa246)
6	ChrI 14183654	G	A	95	0	ncRNA
7	ChrI 14719502	G	A	84	4	pbs-5 intron

Appendix F. Causal genetic lesions in mapped *ucr* alleles

SNPs with 20-nt flanking sequences	
drh-1(ucr2)	AGGCATTGGATGTGTTAAAA**a**ttagtgtgtgatcatttaca
drh-1(ucr3)	AAAACAGTATATTGTCAAAC**A**CTGATGATGAGCGGCTTTAC
drh-1(ucr4)	AAGTGAATACAGTGGATCTT**T**CATGTCTATCAGCTCTTTgt
drh-1(ucr6)	AAAGCAATAAAATATGGCCT**T**GAATCCTGAGAGAAGATACC
rde-1(ucr11)	CAGCGCATATTGTAGTCTAT**T**GAGACGGAGTTAGCGATTCG
mys-2(ucr7)	CAAAGTGCGTCAACCTCCCG**A**AAATGAAATCTACCGAAAAG

Appendix G. Experimentally verified *mys-2 mRNA* sequence

>mys-2_K03D10.3_cDNA

ATGAGCAAACGAGAGCCGAGGAAACGAGCACACGACTCAACAAATGAGGCAGCTGCTCC

CGGCGGAAATGACAGTGAACCCCCGCAACAACAAATAATGCCAATAATTAACAAACGAT

ATATGGTCGAGAAATTCATTATTAATACTGTGAAAAAGGTGGTCGCCGTAATTATACAC

ATTGATTTTCGAAAACCGCCGCAAGAAGTCCGCACGAATAACTATAAATCCATTGAGAA

TCACGAGCAAGGAGCGAAAGAGCTCATGTATTATGTGCATTATGAGGAGTTGGATCGGA

GGAATGATGAATGGATTACGCTGGACAAAATTCTCGTAAACGAAGTGGTAGTCGAGCCG

GTGATCAAAGTGCCCACAAAAGCGCCTGTCGAAGATCTACACAAGGAAATTGTCGTCGG

ACCGCCCGTCATAAAGCCATCTGGAGCCCTAACACGAAGTCAGCGGAGGACATTGGAAG

AATATTCACATTTAAAAACGGATCTTAATGATTTGGACGCCACGACCGCCCGGCTAGAA

CGGGAACATGAAGAACGTACAAAAGTGAAGAACATACCACAAATAACGATAGGCGCTCA

CGAGATTCTCGCCTGGTACTATTCACCGTTTCCGCCTGACTGTGAAAATTTGGACATTT

ATATGTGTGAATATTGCCTACTCTACACACCTCATCACGAACGATTCAAGCAGCACATT

GATACGTGCAAAGTGCGTCAACCTCCCGGAAATGAAATCTACCGAAAAGATCATCTATC

AGTGTACGAAGTGGACGGAAGTGGACAGAAGCTCTACTGTCAGTGCTTATGCCTACTTT

CCAAACTTTTCATGGATCACAAGACGCTCTACTTTGACGTCGATGATTTCATGTTTTAT

GTGCTTTGCGAGACAGACGAGCACGGAGCCCATATAGTTGGCTACTTTTCACGAGAAGT

CGAATCGGCAAATAATTTGGCGTGTATAATGGTTTTTCCGCCGTTTCAGAAGAAGGGAT

ACGGAAAATTGTTGATTCAGTTTAGTTACGAGCTATCCCGCCGCGAAGGATACATTGGA

ATGCCCGAAAAACCGTTGTCAGACCTCGGAAAAGTGTCATATCGCAGCTATTGGTGGTG

GCGACTGATGAAATTGTTTCATATACACCAGGGACATACAGTAACCGCCACGTTTTTGT

CGAATGAGTCGGGCATCGCCGTCGATGATATTGTTTCGACGTTGATTACGATGCGCATG

TGCCGACAGTATAAGGAACCCGAATTCATTCCCGGCGAGTGGTACGTACGAATTCATCG

AAAAATAGTGGATCATTGTGTAATGTGTGGCTATGGTAAACCGCCAAGGCTCCTACTCG

ACCGGACCAAAGTCCGCTGGGCTCCGGCCCAAACTCGGCCCGAATTCGAACGACAACAG

CAACGGACGGCCGCAGCGATGACGTCACGCAGGGCCTCGAAATCGCAAAACGTCACACC

CCTGGTCACGCCCCTAGCCACGCCCCCAATTGAGCAGAAGCCTGAAGACACCTACACCC

CGTCGCCGCTGACAGACCACCATGTAGCTACGGATGCTCCCGACCATATCTCATTGCCA

TTGAACTGA

Extra 51-nt present in reference sequence:

AGTGGACAGAAGCTTTTTCAGCCATTTTTTGCCCCTTTCCCCCCAAAATTGCCCATTTT

TCAGCTCTACTG

Appendix H. Raw Data Comparing Mapping Methods

The *ucr7* mutant library was analyzed using the *ucr7* wild-type pool and *ucr5* mutant pool in order to determine if the wild-type libraries are required to map a novel mutation. Below are the raw data obtained using the two different controls.

Target: *ucr7* mutant		Control: *ucr7* wild-type		
chrI 12977981	G	A	222	DP=41;AF1=0.951219512195122;DP4=1,1,12,27
chrI 13011534	G	A	222	DP=39;AF1=0.971428571428571;DP4=0,1,23,11
chrI 13468565	G	A	222	DP=46;AF1=0.928571428571429;DP4=1,2,14,25
chrI 13660827	G	A	222	DP=39;AF1=0.947368421052632;DP4=0,2,18,18
chrI 14122022	G	A	222	DP=50;AF1=0.977777777777778;DP4=1,0,22,22
chrI 14183654	G	A	222	DP=22;AF1=0.95;DP4=1,0,13,6
chrI 14719502	G	A	225	DP=54;AF1=0.847826086956522;DP4=5,2,12,27
chrV 8330573	G	A	48.1	DP=17;AF1=0.857142857142857;DP4=2,0,9,3

Target: *ucr7* mutant		Control: *ucr5* mutant		
chrI 12977981	G	A	222	DP=41;AF1=0.951219512195122;DP4=1,1,12,27
chrI 13011534	G	A	222	DP=39;AF1=0.971428571428571;DP4=0,1,23,11
chrI 13468565	G	A	222	DP=46;AF1=0.928571428571429;DP4=1,2,14,25
chrI 13660827	G	A	222	DP=39;AF1=0.947368421052632;DP4=0,2,18,18
chrI 14122022	G	A	222	DP=50;AF1=0.977777777777778;DP4=1,0,22,22
chrI 14183654	G	A	222	DP=22;AF1=0.95;DP4=1,0,13,6
chrI 14719502	G	A	225	DP=54;AF1=0.847826086956522;DP4=5,2,12,27
chrII 777129	G	A	222	DP=64;AF1=1;DP4=0,0,24,36;DP'=47;AF1'=0.32;DP4'=12,20,8,7
chrIII 12920052	G	A	225	DP=37;AF1=0.628571428571429;DP4=8,5,14,8
chrV 1255545	G	A	74.4	DP=12;AF1=0.8;DP4=0,2,0,8;DP'=21;AF1'=0.38;DP4'=0,13,0,8
chrV 3216742	C	T	225	DP=34;AF1=0.76;DP4=3,5,8,17;DP'=24;AF1'=0.25;DP4'=7,11,4,2
chrV 8330573	G	A	48.1	DP=17;AF1=0.857142857142857;DP4=2,0,9,3
chrM 5287	G	A	222	DP=1595;AF1=0.998683344305464;DP4=1,1,795,722

Appendix I. Catalog of unmapped *ucr* alleles

ucr5 Mapping

	Location	Ref	Mut	Mutant AF	WT AF	Change
1	ChrV 544564	C	T	61	7	C14C6.3 Cys->Cys(aa186)
2	ChrV 1169279	C	T	62	5	ncRNA
3	ChrV 2110567	C	T	61	0	ncRNA
4	ChrV 2145478	C	T	70	0	str-52 intron
5	ChrV 4622824	C	T	67	3	F54D11.2 Gln->Stop (aa640)
6	ChrV 4765854	C	T	73	2	C18G1.8 Ala->Thr (aa27)
7	ChrV 6042333	C	T	62	13	ncRNA
8	ChrV 8773637	C	T	64	10	srx-115 Gln->Stop (aa273)

ucr10 Mapping

	Location	Ref	Mut	Mutant AF	WT AF	Change
1	ChrI 12243282	C	T	62	0	Intragenic
2	ChrI 13558395	C	T	94	0	vars-2 Arg -> Lys (aa99)
3	ChrI 14125134	G	A	100	0	mys-2 5' splice site, I8
4	ChrI 14157604	A	T	100	0	ncRNA
5	ChrI 14542324	G	A	100	0	zoo-1 Glu -> Lys (aa730)

The following alleles were mapped with only mutant libraries comprised of about 50 F2s.

L4-4 Mapping

	Location	Ref	Mut	Mutant AF	WT AF	Change
1	chrI 333538	G	A	81	26	
2	chrI 3263052	G	A	100	33	
3	chrI 4503135	G	A	100	31	
4	chrI 4503136	G	A	100	16	
5	chrII 14384	C	T	100	16	
6	chrIV 2993669	C	T	100	25	
7	chrIV 2993671	C	T	100	31	
8	chrV 3374592	C	T	83	0	srbc-2 Ala -> Thr (aa282)
9	chrV 3388953	C	T	89	0	Intragenic
10	chrV 3535843	C	T	94	0	Intragenic
11	chrV 3839448	C	T	80	0	ugt- Gly -> Ser (aa284)
12	chrV 3962959	C	T	80	0	cyp-34A7 Gln -> Gln (aa194)
13	chrV 4010549	C	T	100	0	srh-4 Ser -> Leu (aa32)

14	chrV 4047301	C	T	81	0	C49G7.3 3' UTR
15	chrV 4180858	C	T	80	0	Y45G5AM.2 Val->Val (aa987)
16	chrV 4202873	C	T	81	0	Intron
17	chrV 4362594	C	T	92	0	Intragenic
18	chrV 4504168	C	T	91	0	unc-62 Pro -> Ser (aa348)
19	chrV 4624184	C	T	86	0	F54D11.2 Gln -> Stop (aa768)
20	chrV 4636373	C	T	88	0	F54D11.2 Asp -> Asn (aa94)
21	chrV 5224839	C	T	87	0	Intragenic
22	chrV 5600288	C	T	86	0	CD4.8 Gln -> Gln (aa281)
23	chrV 6383447	C	T	88	0	Intragenic
24	chrV 6497802	C	T	82	0	W01A11.2 Cys -> Cys (aa205)
25	chrV 6560070	C	T	86	0	spp-19 intron
26	chrV 6647605	C	T	92	0	Intragenic
27	chrV 7161431	C	T	88	0	ugt-41 intron
28	chrV 7169507	C	T	81	0	Intragenic
29	chrV 7352500	C	T	92	0	Intragenic
30	chrV 7683496	G	A	88	0	sre-11 intron
31	chrV 8480370	C	T	100	0	Intragenic
32	chrV 8797080	C	T	83	0	Intragenic
33	chrV 8949388	C	T	83	0	gck-1 intron
34	chrV 9529668	C	T	85	0	Intragenic
35	chrV11470116	C	T	81	0	aat-6 intron
36	chrV11480773	C	T	88	0	T11F9.14 Gly -> Asp (aa15)
37	chrV12316951	C	T	88	0	R04F11.3 Gly -> Arg (aa55)
38	chrV12477884	C	T	83	0	Intragenic
39	chrV13077393	C	T	83	0	Intragenic
40	chrV13301133	C	T	80	0	Intragenic
41	chrV14512705	C	T	80	0	R02D5.6 Gly -> Glu (aa242)
42	chrV14824007	C	T	80	0	Intragenic
43	chrV15401332	C	T	82	0	sri-7 Iso -> Iso (aa220)
44	chrV15580320	C	T	91	0	F28F8.7 Glu -> Lys (aa154)
45	chrV15642385	C	T	92	0	Intragenic
46	chrV16228919	C	T	90	0	F21H7.2 Glu ->Lys (aa94)
47	chrV16244755	C	T	85	0	Intragenic
48	chrX 9438988	C	T	95	0	
49	chrX17561917	C	T	100	0	

L4-2 Mapping

	Location	Ref	Mut	Mutant AF	WT AF	Change
1	chrII 14384	C	T	100	0	
2	chrIII 530089	C	T	92	0	intragenic
3	chrIII 5253731	C	T	100	0	C34E10.8 splice site, 3' (1st nt)
4	chrIII 5925335	C	T	100	0	dnc-2 Ala->Th (aa84)
5	chrIII 6033755	C	T	100	0	ZK328.7 3' UTR (1st nt)
6	chrIII 7159233	C	T	100	0	rnp-7 Gly->Asn (aa286)
7	chrIII 9021143	C	T	100	0	intragenic
8	chrIII 9035944	C	T	100	0	intragenic
9	chrIII 9228523	C	T	100	0	rnr-1 intron
10	chrIII 9433854	C	T	100	0	intragenic
11	chrIII 9900542	C	T	100	0	cyld-1 Gly-> Asn (aa695)
12	chrIII 10306465	G	A	100	16	intragenic
13	chrIII 10341925	G	A	100	8	intragenic
14	chrIII 10907845	C	T	95	33	unc-119 Arg -> stop (aa113)*
15	chrIII 10988992	C	T	94	0	W05B2.7 intron
16	chrIII 11063605	C	T	100	0	rsa-2 intron
17	chrIII 11458120	C	T	90	0	Y47D3B.11 intron
18	chrIII 11655044	C	T	92	0	Y41C4A.2 intron
19	chrV 4353137	C	T	100	0	
20	chrV 12536959	G	A	100	34	
21	chrV 12761680	G	A	100	27	
22	chrV 14988410	G	A	100	33	
23	chrV 15756224	G	A	100	14	

**unc-119(ed9)* allele used in the MosSCI protocol

22-4 Mapping. This library had low coverage, so the causal mutation may be missing.

	Location	Ref	Mut	Mutant AF	WT AF	Change
1	chrI 934758	G	A	95	0	Intragenic
2	chrII 6508136	C	T	100	0	Intragenic
3	chrIII 7425383	C	T	100	0	Intragenic
4	chrIV 10949716	G	A	92	0	Intragenic
5	chrIV 11073998	C	T	92	0	Intragenic
6	chrIV 11660707	C	T	67	0	ZK792.1 intron
7	chrIV 14775292	G	A	69	0	Intragenic
8	chrV 17636498	C	T	90	0	Intragenic
9	chrV 17711555	C	T	85	0	twk-33 Val-> Phe (aa354)
10	chrX 5049263	C	T	87	0	Intragenic

22-5 Mapping

	Location	Ref	Mut	Mutant AF	WT AF	Change
1	chrI 201781	G	A	68	0	
2	chrI 993783	G	A	65	0	
3	chrI 1422149	G	A	61	0	
4	chrI 1940762	G	A	62	0	
5	chrI 5961504	C	T	75	0	
6	chrI 7004926	C	T	70	0	
7	chrI 10849089	G	A	100	0	
8	chrII 14384	C	T	100	0	
9	chrII 10757434	C	T	63	0	
10	chrII 11173482	C	T	62	0	
11	chrII 11441199	C	T	100	0	
12	chrIII 2496833	C	T	62	0	
13	chrIII 9575009	G	A	84	0	
14	chrIII 10306465	G	A	100	0	
15	chrIII 10341925	G	A	100	0	
16	chrIII 10907845	C	T	95	0	
17	chrIV 649826	C	T	71	0	
18	chrIV 1153417	C	T	61	0	
19	chrIV 4094666	C	T	73	0	
20	chrIV 4284633	G	A	64	0	
21	chrIV 10949716	G	A	63	0	
22	chrIV 17445459	C	T	90	0	
23	chrV 72081	C	T	79	0	
24	chrV 968258	C	T	67	0	
25	chrV 1278972	C	T	67	0	
26	chrV 2257914	C	T	67	0	
27	chrV 2948758	C	T	70	0	
28	chrV 3466525	C	T	65	0	
29	chrV 4353137	C	T	100	0	
30	chrV 5162970	G	A	100	14	
31	chrV 5287297	C	T	100	0	
32	chrV 6943841	G	A	69	0	
33	chrV 7715090	C	T	71	0	
34	chrV 8330573	G	A	72	0	
35	chrV 9799754	G	A	70	0	
36	chrV 9799780	C	T	68	0	

37	chrV 10237098	C	T	73	0	
38	chrV 11357335	G	A	68	0	
39	chrV 12217840	G	A	72	0	
40	chrV 12536959	G	A	100	0	
41	chrV 12761680	G	A	100	0	
42	chrV 13205945	G	A	83	0	
43	chrV 13244164	G	A	75	0	
44	chrV 14551461	G	A	65	0	
45	chrV 14988410	G	A	100	0	
46	chrV 15548017	G	A	67	0	
47	chrV 15733107	G	A	65	0	
48	chrV 15756224	G	A	100	0	
49	chrV 17221975	G	A	66	0	
50	chrV 17685310	G	A	65	0	
51	chrV 17867548	C	T	71	0	
52	chrX 165379	C	T	61	0	T08D2.2 intron
53	chrX 190426	C	T	62	0	T08D2.2 Ala -> Val (aa426)
54	chrX 645207	C	T	74	0	ifc-2 Leu -> Leu (aa698)
55	chrX 1035963	C	T	100	0	F55A4.1 Asn -> Asn (aa99)
56	chrX 1442976	C	T	73	0	zfp-3 intron
57	chrX 1490127	C	T	68	0	F09E10.5 splice site mutation
58	chrX 2275969	C	T	68	0	Intragenic
59	chrX 2436214	C	T	78	0	T07D1.2 intron
60	chrX 2442506	C	T	77	0	Intragenic
61	chrX 2455775	C	T	83	0	Intragenic
62	chrX 2785832	C	T	73	0	cgef-1 intron
63	chrX 3455311	C	T	79	0	sax-3 Gly -> Glu (aa455)
64	chrX 3637498	G	A	80	0	Intragenic
65	chrX 3641916	G	A	77	0	R11G1.6 Thr -> Iso (aa348)
66	chrX 3643240	G	A	92	0	R11G1.6 Pro -> Leu (aa162)
67	chrX 4790565	G	A	77	0	Intragenic
68	chrX 4794520	G	A	81	0	F43C9.1 intron
69	chrX 5166285	G	A	71	0	intron
70	chrX 5318007	C	T	69	0	Intragenic
71	chrX 5953144	G	A	69	0	Intragenic
72	chrX 5977929	G	A	80	0	Intragenic
73	chrX 6561238	G	A	70	0	T14E8.1 Gln -> Stop (aa148)
74	chrX 8018590	G	A	73	0	c34D10.2 intron
75	chrX 8566818	G	A	72	0	Intragenic

76	chrX 8839382	G	A	64	0	Intragenic
77	chrX 8839724	C	T	68	0	Intragenic
78	chrX 8947476	G	A	74	0	trk-1 intron
79	chrX 9542150	G	A	64	0	nuc-1 Trp-> Stop (aa362)
80	chrX 10197310	G	A	80	0	Intragenic
81	chrX 11396007	G	A	76	0	slo-2 Leu -> Leu (aa789)
82	chrX 12971725	G	A	70	0	ppk-3 Cys -> Cys (aa912)
83	chrX 13357361	G	A	68	0	Intragenic
84	chrX 14045376	C	T	65	0	Intragenic
85	chrX 14154163	G	A	68	0	Intragenic
86	chrX 15490028	G	A	77	0	H13NO6.2 intron
87	chrX 16596503	G	A	81	0	cdh-9 intron
88	chrX 16827092	G	A	82	0	Intragenic
89	chrX 17000916	G	A	82	0	Intragenic
90	chrX 17017545	G	A	96	0	F35SB3.1 intron
91	chrX 17022561	G	A	86	0	pqn-34 Leu -> Phe (aa1040)
92	chrX 17087682	G	A	100	0	Intragenic
93	chrM 5097	C	T	100	0	

L13-13 Mapping

	Location	**Ref**	**Mut**	**Mutant AF**	**WT AF**	**Change**
1	chrII 14384	C	T	100	0	
2	chrII 11441199	C	T	100	0	
3	chrIII 7564246	G	A	95	0	
4	chrIII 10306465	G	A	100	16	
5	chrIII 10341925	G	A	100	8	
6	chrIII 10907845	C	T	100	33	
7	chrV 557656	C	T	94	0	Intragenic
8	chrV 1105810	C	T	96	0	Intragenic
9	chrV 1486446	C	T	90	0	Intragenic
10	chrV 2816486	C	T	93	0	Intragenic
11	chrV 3332558	C	T	100	0	Intragenic
12	chrV 4134087	C	T	100	0	nhr-57 Asp->Asp (aa313)
13	chrV 4190420	G	A	95	0	Y45G5AM.9 intron
14	chrV 4353137	C	T	100	0	F32D1.3 Ala->Thr (aa382)
15	chrV 4387470	G	A	100	0	F13D2.9 Pro->Ser (aa7)
16	chrV 4625598	C	T	94	0	F54D11.2 Arg->Stop (aa906)
17	chrV 4842761	C	T	100	0	Intragenic

18	chrV 4968602	C	T	100	0	Intragenic
19	chrV 5161006	G	A	95	0	abu-6 Ala->Ala (aa275)
20	chrV 5162970	G	A	100	0	Intragenic
21	chrV 5731017	C	T	100	0	Intragenic
22	chrV 5887121	C	T	100	0	Intragenic
23	chrV 6047539	C	T	100	0	R01B10.4 intron
24	chrV 6408483	C	T	100	0	F57F4.4 Gly->Arg (aa78)
25	chrV 6821849	C	T	100	0	T15B7.8 Leu->Val (aa45)
26	chrV 6870280	C	T	94	0	T19F4.1 5' UTR
27	chrV 7556359	C	T	100	0	Intragenic
28	chrV 7667413	C	T	100	0	Intragenic
29	chrV 7786822	C	T	100	0	C53C11.4 Glu->Stop (aa82)
30	chrV 7859249	C	T	100	0	C53C11.4 Ser->Phe (aa320)
31	chrV 7874133	C	T	100	0	C53C11.4 intron
32	chrV 9138055	C	T	100	0	Intragenic
33	chrV 9504509	C	T	100	0	Intragenic
34	chrV 9876351	C	T	100	0	Intragenic
35	chrV 10389357	C	T	94	0	AC3.5 Pro->Ser (aa54)
36	chrV 10948529	C	T	92	0	F17C11.2 intron
37	chrV 12536959	G	A	100	0	Intragenic
38	chrV 12665695	C	T	95	0	Intragenic
39	chrV 12679527	C	T	94	0	Intragenic
40	chrV 12761680	G	A	100	27	abcf-1 intron
41	chrV 12887922	C	T	100	0	Intragenic
42	chrV 12999330	C	T	97	0	far-4 Ser->Lys (aa102)
43	chrV 13349470	C	T	90	0	twk-43 intron
44	chrV 14988410	G	A	100	33	t01c3.1 intron
45	chrV 15267352	C	T	94	0	Intragenic
46	chrV 15722214	C	T	93	0	srh-140 Leu->Phe (aa16)
47	chrV 15756224	G	A	100	0	str-8 Ser->Asp (aa300)
48	chrV 15881864	C	T	90	0	Intragenic
49	chrV 16098180	C	T	100	0	str-7 Ala->Gly (aa297)
50	chrV 19669647	C	T	92	0	Intragenic

Appendix J. Protocol for EMS mutagenesis of FR1gfp in *C. elegans*

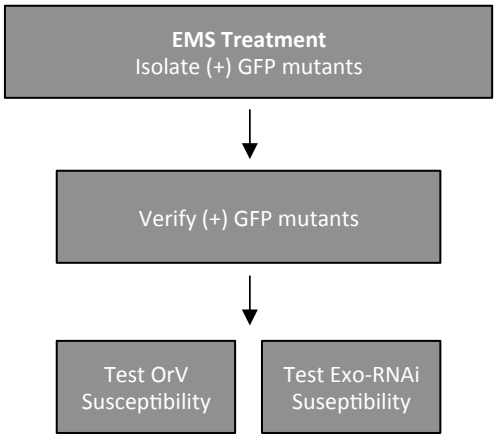

Figure J-1. Schematic of *irSi18* EMS screen. After EMS treatments, F2s expressing GFP are isolated. Mutants that reproducibility fail to suppress FR1gfp are characterized by OrV and exo-RNAi susceptibility.

1. Isolation of mutants that fail to suppress FR1gfp

Day 1- 50mM EMS treatment

1. Work with EMS under the hood! All materials that come in contact with EMS, pipette tips, tubes, etc. should be disposed of in the container labeled "EMS waste". Liquid EMS waste should be placed in a 50mL tube, labeled "EMS Waster". When finished, parafilm the EMS waste tube and dispose of it in the EMS waste.
2. Add late L4/young adult worms to 2mL of M9 in a 15mL tube. In a separate tube add 2mL M9 and 20µL EMS. Close the tube with EMS and Parafilm the lid. Mix the EMS by inverting a couple of time. Combine the contents of the two tubes and parafilm the lid. Incubate for 4 hours at 20°C, on rotisserie.

3. Spin down worms into a 1.5mL tube using the bench top mini-centrifuge. Remove EMS supernatant into the EMS waste. Repeat until all 4mL of EMS-treated worms have been spun down.
4. Add 500 µL of M9 and spin down using the mini-centrifuge. Remove M9 supernantant into the EMS waste. Repeat two more times.
5. Transfer worms to a clean plate and let them recover for at least 30 minutes at room temperature.
6. Transfer ten healthy, late L4s to a large plate. Each plate of 10 P0s is considered one group and should be given a letter designation (A, B, C, D…).

Day 2- Remove P0s

1. Manually remove the 10 P0s from each plate once each parent has laid ~10 eggs. There should be ~100 eggs on every large plate. Make sure you get all 10 P0s.

Day 4- Remove F1s

1. Remove F1s after they have laid ~20 eggs each. Gently wash F1s off of the plates with M9 (eggs will remain stuck in food) and then manually remove any remaining F1s from each of the plates. There should be ~2,000 F2s on each plate.

Day 6- Induce FR1gfp Replication

1. When F2s are at the L4 stage, heat shock the plates by incubating them at 34°C for 4 hours in water bath. If the EMS treatment was successful, then at this point you should be able to observe various phenotypes in the F2s populations, such as unc and dpy.

Day 8-

1. Observe plates under the GFP filter on the dissecting scope.
2. Analyze the *drh-1* controls first to ensure that the heat treatment was successful. Record your observations in your lab notebook.
3. Analyze the N2 controls to ensure that FR1gfp is suppressed in wild-type animals. Record your observations in your lab notebook.
4. Using Figure J-2 as a guide, single F2s expressing GFP to individual small plates and label with designation from the corresponding large plate.
5. Record in your lab notebook how many (+)GFP mutants you isolated.

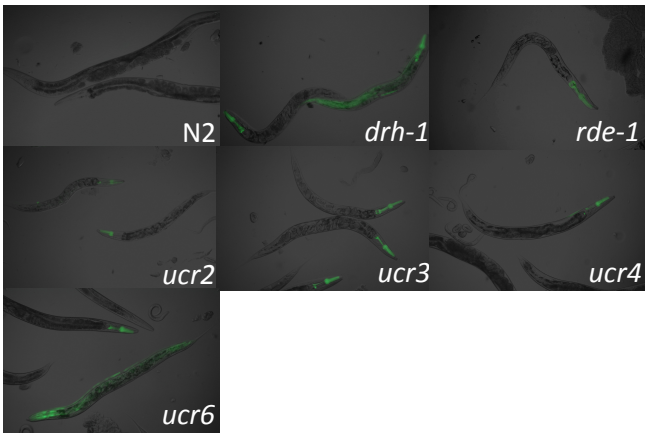

Figure J-2. GFP expression phenotype in mutants defective in antiviral RNAi.

Day 10/11-

1. Check every F2 under the dissecting scope to determine if they were viable. If an F2 was viable, then you will see the next generation. If the F2 was not viable, then there will be no animals on the plate.
2. Chunk viable F2s onto new medium plates and store original plates at 15°C.

3. Record in your lab notebook how many (+)GFP mutants were viable.

2. Re-Test mutants for FR1gfp

Every isolated mutant should undergo the following experiment three times. Three independent replicates will demonstrate that the mutant carries an EMS-induced mutation that reproducibly fails to suppress FR1gfp.

Day 1- Plate worms

1. For each mutant, N2 and *drh-1*, transfer synchronized L1 worms to a large plate.
4. Incubate at 20°C for two days.

Day 3- Induce replication of FR1gfp

1. When F2s are at the L4 stage, heat shock the plates by incubating them at 34°C for 4 hours in water bath.
2. Return to the 20°C incubator.

Day 5-Score GFP

1. Observe the *drh-1* and wild-type worms under the GFP filter on the fluorescent dissecting scope. Write down your observations in your lab notebook.
2. Observe the mutants and write down your observations.

3. smFISH for Orsay RNA1 in *C. elegans*

Day 1- Set up Orsay Infections

1. Pippette 50µL of OrV filtrate onto medium plates. For each mutant, transfer about 200 synchronized L1 worms onto each plate. Set up two plates for each mutant, N2 and *drh-1*.

2. Incubate at 20°C for two days.

Day 3-Fix Worms

1. Check the worms under the dissecting microscope. They should be at the L4 stage or early adults. If the worms are full of eggs then they are too old and the gonad will block the view of the intestinal cells.
2. Wash OrV-infected, late L4 or early adult worms off of plate into a 1.5mL tube.
3. Wash with M9 three times, or until all the bacteria has been removed.
4. Resuspend worms in 1mL Fixation Solution (made fresh) and incubate at room temperature for 45 minutes.
5. Spin down worms and remove Fixation Solution. Wash twice with 1XPBS in DEPC-H_2O.
6. Resuspend in 1mL 70% EtOH in DEPC- H_2O and incubate for 1 hour to overnight at 4°C (can be stored at this stage for several weeks).

Day 4- Hybridize with OrV probe

3. Spin down and remove 70% EtOH. Wash twice with Wash Buffer.
4. Resuspend worms in 100µL Hybridization Buffer and 1µL Orsay probe (aliquots of 500µL Hybridization Buffer and 5µL orsay probe are stored in the -20°C in the Maduro Lab). Incubate at 30°C overnight.
5. Spin down and remove hybridization buffer. Wash twice with Wash Buffer.
6. If desired, incubate with 50-100µL DAPI at 30°C for 30 minutes. Spin down and wash twice with Wash Buffer. (DAPI stock is 1000x, dilute in wash buffer for hybridization)

7. Wash twice with 2X SSC. Resuspend worms in 100-500µL 2X SSC. At this point worms can be stored at room temperature long term.

Day ?- Score OrV infected animals

1. To mount worms on slides, add 40µL to a slide, place coverslip over worms and wic away excess SSC (worms should not have enough liquid under the slide to move). Seal with clear nail polish. Once the worms have been mounted, they are good for about a day, after which they begin to desiccate.
2. Obtain the proper training to use the upright microscope. Observe the slide under the FITC filter. OrV will infect the intestinal cells; a few examples are given in Figure J-3.
3. Record your observations in your lab notebook.
4. Starting in one corner of the slide, begin scanning the slide and scoring each worm as "infected" or "not infected". Record your results.

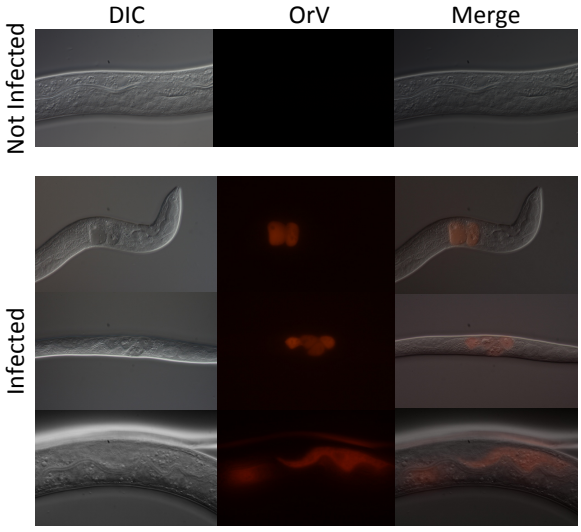

Figure J-3. OrV detection by smFISH

Fixation Solution	Wash Buffer
5mL 10x PBS	5mL 20X SSC
5mL Formaldehyde	5mL Formamide
40mL DEPC- H_2O	40mL DEPC- H_2O

smFISH protocol is from Gina Broitman-Maduro in the Maduro Lab and was Adapted from: Arjun Raj and Sanjay Tyagi. Detection of Individual Endogenous RNA Transcripts In Situ Using Multiple Singly Labeled Probes. *Methods in Enzymology*, Volume 472. 2010

4. Determine if the mutants are capable of exogenous RNAi

Day 1- Inoculate cultures

1. Add 100μL of carbenicillin and tetracycline to every 100mL of LB.
2. Add 50mL LB+carb+tet to three100mL bottles. Label bottles "L4440", "*unc-22*", and "*pop-1*". L4440 is a no-insert control.
3. Under the hood, use a sterile pipette tip to pick up a colony of *E. coli* for each of the feeding RNAi clones and drop it into the corresponding bottle.
4. Shake at 37°C O/N

Day 2- Seed plates

1. Spin down bacteria cultures at 3,000rpm for 10 minutes. While the cultures are spinning, label small feeding plates.
2. Pour off the supernatant and vortex each sample

3. Resuspend bacterial cultures in 10mL of LB Broth.
4. Pipette 200μL onto small plates. Repeat steps 2-4 for each clone.
5. Let plates dry for at least two days at room temperature.

Day 4- Set up Feeding RNAi experiments

1. For each mutant, place 5 adult worms to (1) an L4440 plate and (2) three *unc-22* plates.
2. For each mutant, place 3 young L4 worms onto (1) three L4440 plates and (2) three *pop-1* plates.
3. Repeat step 1-2 with wild-type worms and a RDE control, such as *rde-1* or *rde-4*.
4. Put all the plates at 20°C.

Day 6- Score phenotypes

1. For *pop-1* plates, score the viability of the strain. Using the dissecting scope, look for dead eggs and viable L1 worms. Wild-type animals will lay dead eggs on *pop-1* plates. Compare to wild-type animals on the L4440 plates, which will lay live eggs and will have many L1 worms. RDE animals will lay live eggs on the *pop-1* plates. There will not be a difference between RDE animals on the L4440 plates and the *pop-1* plates. Observe each mutant plate and write down your observations in your lab notebook. See Figure J-3.
2. For *unc-22* plates it is best to score L4 stages worms. Using the dissecting score, observe the worms and look for twitching or paralyzed worms. The wild-type worms will be twitching. The RDE worms should continue to move around the plate normally. Observe each mutant plate and compare it to the L4440 plate to

ensure that the unc phenotype is caused by the RNAi and not the EMS-induced mutation. Write down your observations in your lab notebook.

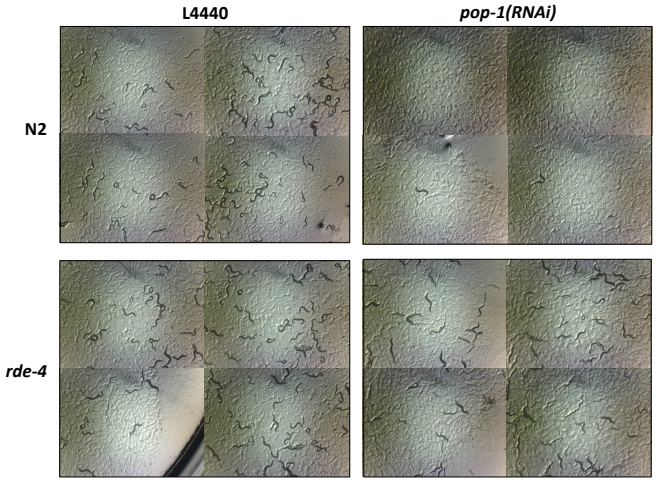

Figure J-4. Phenotype of *pop-1(RNAi)*