Regulation of RNA interference pathways in *C. elegans*

Ahilya Nirvana Sawh

Department of Biochemistry

McGill University

Montreal, Quebec, Canada

June, 2015

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Doctor of Philosophy

Table of Contents

Abstract

RNA interference (RNAi) regulates gene expression through overlapping pathways in a multitude of organisms. These pathways are driven by small RNAs that can target virtually any expressed RNA in the cell to mediate transcriptional or post-transcriptional gene silencing. The biogenesis of these small RNAs and their downstream silencing effects are mediated by highly conserved genes of the Dicer and Argonaute nuclease families, among others. The research presented here specifically investigates how Dicer is regulated, the interplay of RNAi pathways, and the mechanism of a novel branch of RNAi called nuclear RNAi, using *C. elegans* as a model system.

Dicer (*C. elegans* DCR-1) is the RNase III enzyme required for the production of small RNAs in the miRNA, exoRNAi and ERI endoRNAi pathways. We have discovered that DCR-1 is proteolytically cleaved *in vivo* to produce a stable, developmentally regulated, and abundant C-terminal fragment (sDCR-1). sDCR-1 actively promotes the exoRNAi pathway activity, and inhibits the miRNA pathway through competitive inhibition of miRNA Argonautes. We have also uncovered a cluster of phosphorylation sites on DCR-1 that affects its activity in exoRNAi, and its protein interactions with exoRNAi and ERI endoRNAi partners. This cluster is also important to maintain proper cell fate specification, developmental timing, and the production of viable progeny. Post-translational modification of DCR-1 therefore plays a major role in the function of the enzyme in multiple RNAi pathways.

Additionally, we pursued the elucidation of nuclear RNAi pathways. In order to assess the impact of the ERI endoRNAi pathway on the transcription of target genes, we developed and optimized a robust nuclear run-on assay in *C. elegans* embryos. Using this assay, we determined that the ERI endoRNAi pathway transcriptionally inhibits its targets, and established a basis for future genome-wide transcriptional analysis. To further clarify the mechanism of RNAi-induced transcriptional inhibition, we used proteomics to establish a network of nuclear RNAi factors. We discovered the first physical links between a nuclear Argonaute, NRDE-3, and chromatin-modifying machinery in animals. Moreover, we discovered a novel multi-Argonaute interacting protein in the ERI endoRNAi and exoRNAi pathways, which we propose acts at the step of Argonaute loading.

These studies provide biochemical insight into small RNA biogenesis through regulation of DCR-1, and the mechanisms of nuclear RNAi.

Résumé

L'interférence à l'ARN (ARNi) régule l'expression génique par des voies apparentées, au travers une grande diversité d'organismes. Ces voies de régulation génique sont guidées par de petits ARNs non-codants qui contrôlent l'expression d'une myriade de gènes par des mécanismes transcriptionnel et post-transcriptionnel. La biogenèse de ces petits ARNs, de même que les mécanismes conséquents mettent en œuvre une machinerie moléculaire très largement conservée, telles les nucléases Dicer et Argonaute. Les travaux présentés dans cette thèse, principalement chez le nématode modèle *C. elegans*, explorent la régulation de Dicer, les interactions fonctionnelles entre les voies ARNi, et les mécanismes sous-jacents à l'ARNi nucléaire, une sous-branche du ARNi qui fut récemment identifiée.

Dicer (nommé DCR-1 chez *C. elegans*) est une RNase type III essentielle pour la biogenèse des petits ARNs qui guident les voies des microARNs, de l'exoARNi et de l'ERI endoARNi. Nous avons découvert que DCR-1 est tronquée *in vivo*, par une activité endo-protéolytique modulée au cours du développement, produisant ainsi sDCR-1, un fragment C-terminal stable. sDCR-1 promeut l'activité exoARNi d'une part, et d'autre part inhibe la voie des microARNs en inhibant de façon compétitive les Argonautes qui y sont dédiés. Nous avons de plus identifié un agrégat d'acides aminés phosphorylés dans la séquence de DCR-1 qui régule son activité dans la voie exoARNi, de même que ses interactions protéine-protéine impliquées dans l'exoARNi et l'ERI

endoARNi. Ces acides aminés sont cruciaux pour la progression temporelle normale du développement, affectent les destinées cellulaires, de même que la viabilité de la descendance de l'animal.

Afin de comprendre les mécanismes transcriptionnel et post-transcriptionnel de régulation génique par la voie ERI endoARNi, nous avons développé et optimisé des essais robustes de Run-ON nucléaire. Ce faisant, nous avons démontré que l'ERI endoARNi conduit à une extinction génique au niveau transcriptionnel de ses gènes-cibles. Nous avons de plus adapté cet essai pour un examen à l'échelle génomique des cibles ERI endoARNi en utilisant la méthode de séquençage génomique à haut débit.

Finalement, afin d'identifier les composantes fonctionnelles impliquées dans l'ARNi nucléaire, nous avons amorcé une analyse protéomique ciblée des protéines nucléaires impliquées dans l'ERI endoRNAi. L'analyse des interactions avec l'argonaute nucléaire NRDE-3 a révélé entre autres une association avec des protéines qui modifient la chromatine, de même que NIP-1, une protéine intimement associée à plusieurs argonautes nucléaires.

Les travaux présentés dans cette thèse élucident la régulation des mécanismes de biogenèse des petits ARNs de même que de nouveaux aspects des mécanismes de l'extinction génique qu'ils guident.

List of Figures

List of Tables

List of Abbreviations

A – adenosine
AGO – Argonaute
ATP – adenosine triphosphate
C – cytosine
CTP – cytoosine triphosphate
DCL – Dicer-like
DCR – Dicer related
DNA – deoxyribonucleic acid
dsRNA – double stranded ribonucleic acid
RBD – RNA binding domain
DTT – dithiothreitol
EDTA – ethylenediaminetetraacetic acid
endoRNAi – endogenous RNAi
ERI- enhancer of RNAi
exoRNAi – exogenous RNAi
G – guanosine
GTP– guanosine triphosphate
HDAC – histone deacetylase
HMT – histone methyltransferase
HRDE – heritable RNAi deficient
IP - immunoprecipitation
kDa – kilodalton
m7G – 7-methylguanosine
MID domain – middle domain
miRNA – microRNA
miRISC – microRNA-induced silencing complex
MUT - mutator
NIP – NRDE-3 interacting protein

nt – nucleotide

NRDE – nuclear RNAi deficient

NTP – nucleoside triphosphate

PACT - protein activator of PKR

PAGE – polyacrylamide gel electrophoresis

PAZ – PIWI/Argonaute/Zwille

piRNA – PIWI interacting RNA

PIWI – P-element induced wimpi testes

PRO-seq – precise run on sequencing

PTGS – post-transcriptional gene silencing

RDE – RNAi defective

RdRP– RNA-dependent RNA polymerase

RISC – RNA-induced silencing complex

RNA – ribonucleic acids

RNAi – RNA interference

SDS – sodium dodecyl sulfate

siRNA – small interfering RNA

ssRNA – single stranded ribonucleic acid

TGS – transcriptional gene silencing

TRBP – TAR RNA-binding protein

U – uracil

UTP – uridine triphosphate

UTR – untranslated region

Preface

This thesis consists of one published research article and three manuscripts in preparation.

Chapter 1: **Sawh A. N**. and Duchaine T. F. A Truncated Dicer Tilts the Balance of RNAi Pathways. *Cell Reports*, August 2013, Volume 4, Issue 3

Chapter 2: **Sawh A. N**., Lewis A., Wohlschlegel J. and Duchaine T. F., Phosphorylation of Dicer in the substrate-recognition domain modulates its activity in RNAi pathways in *C. elegans*. (in preparation)

Chapter 3: **Sawh A. N**. and Duchaine T. F., Nuclear run-on and PRO-seq in *C. elegans* embryos reveal the impact of endogenous RNAi pathways on the transcriptional landscape. (in preparation)

Chapter 4: **Sawh A. N**., Vashisht A., Flamand M., Lewis A., Wohlschlegel J., and Duchaine T. F., Comparative proteomics physically links nuclear RNAi factors with histone modifying machinery and novel RNAi components to mediate transcriptional gene silencing. (in preparation)

Contribution of Authors

Chapter 1: ANS and TFD conceived the study and designed the experiments. ANS performed and analyzed data from all experiments. ANS and TFD wrote the manuscript.

Chapter 2: ANS and TFD conceived the study and designed the experiments. TFD performed initial DCR-1 IPs and JW performed mass spectrometry experiments to identify phosphorylation sites. AL performed *unc-22* RNAi assays, assisted in some IPs, and in cloning and UV integration of some transgenes. ANS performed all other experiments.

Chapter 3: ANS and TFD conceived the study and designed the experiments. ANS performed and analyzed data from all experiments.

Chapter 4: ANS and TFD conceived the study and designed the experiments. ANS performed all protein purification submitted to MudPIT. AV in the lab of JW performed MudPIT, ANS analyzed the data. MF produced recombinant protein used to generate antibodies. AL assisted with IPs. ANS performed all other experiments.

Original Contributions to Knowledge

Chapter 1: A Truncated Dicer Tilts the Balance of RNAi Pathways.

- Endogenous DCR-1 is proteolytically cleaved to produce a stable, abundant C-terminal product: sDCR-1.
- sDCR-1 enhances the exoRNAi pathway.
- sDCR-1 inhibits the biogenesis of miRNAs by competing for ALG-1 and ALG-2.
- sDCR-1 causes the exoRNAi pathway to compete with miRNA biogenesis.

Chapter 2: Phosphorylation of Dicer in the substrate-recognition domain modulates its activity in RNAi pathways in *C. elegans*.

- Endogenous DCR-1 is phosphorylated in a cluster of sites 3' to the PAZ domain termed the ST cluster.
- ST cluster phospho-mimetic mutant DCR-1 animals are similar to wild-type DCR-1 animals in exoRNAi activity and physiology.
- ST cluster phospho-null mutant DCR-1 animals are defective in exoRNAi activity, and have strengthened endoRNAi protein complexes.
- ST cluster phospho-null mutant DCR-1 animals display dominant heterochronic defects, reduced viability and severe phenotypes.

Chapter 3: Nuclear run-on and PRO-seq in *C. elegans* embryos reveal the impact of endogenous RNAi pathways on the transcriptional landscape.

- Development and optimization of a robust nuclear run-on assay using *C. elegans* embryos.
- The endoRNAi pathway transcriptionally inhibits its target genes.

Chapter 4: Comparative proteomics physically links nuclear RNAi factors with histone modifying machinery and novel RNAi components to mediate transcriptional gene silencing.

- First physical links established between a nuclear Argonaute, NRDE-3, and chromatin-modifying machinery in animals.
- Discovery of a novel multi-Argonaute interacting protein, NIP-1, in the ERI endoRNAi and exoRNAi pathways.
- Establishment of a proteomic network of nuclear RNAi factors.

Acknowledgements

I am sincerely grateful to my supervisor, Dr. Thomas Duchaine, for providing a creative, invigorating, and challenging environment to pursue these studies. Thank you for granting me the intellectual freedom to ask the questions I was most interested in. Thank you also for being an extremely generous, enthusiastic, and dedicated teacher.

To my research advisory committee, Dr. Xiang-Jiao Yang, Dr. Richard Roy and Dr. Paul Lasko, thank you for your support, critical advice and guidance. To all the past and present members of the Duchaine Lab, thank you for the lively discussions and collaborations. To Edlyn Wu, Mathieu Flamand and Vinay Mayya in particular, thank you for not only being excellent colleagues, but stalwart friends.

My research career began during my undergraduate studies at McMaster University and Sanofi Pasteur, and I would like to give heartfelt thanks to my first research supervisors Dr. Ashok K. Grover and Dr. Bruce Carpick. Your combined guidance taught me how to ask pertinent questions, and inspired in me a fervent interest in basic science.

To my life-long friends near and far, and my family members scattered throughout the continent, thank you for believing in my success and supporting me though this path. To my parents Nesha Sawh and Virendra Sawh, and my brother Shakti Sawh, I dedicate this thesis to you. Your sacrifices, love and understanding have made this possible.

Finally, to Maxwell Shafer, thank you for sharing with me your friendship, love and laughter. You constantly inspire and motivate me to be a better scientist and a better person. Together we have shared numerous career- and life-defining moments, and I look forward to many more.

Introduction to small RNA-mediated pathways

I.I Key discoveries and introduction to the exoRNAi pathways

The discovery that small non-coding RNAs can repress gene expression in a potent and sequence-dependent manner through a multitude of mechanisms has rapidly revolutionized our understanding of RNA biology. RNA-driven silencing phenomena, or RNA interference (RNAi), of endogenous gene expression was first observed in the petunia plant (Napoli et al., 1990; van der Krol et al., 1990) and the fungi *Neurospora crassa* (Romano and Macino, 1992) where these phenomena were termed “co-suppression” and “quelling”, respectively. In these studies, the introduction of an exogenous additional copy or homologous RNA to an endogenous gene led to the silencing of both gene products. Soon after, evidence for this type of phenomenon in animals was observed in *Caenorhabditis elegans* (Guo and Kemphues, 1995), where the addition of RNA homologous to an endogenous gene resulted in its destruction. This work laid the groundwork for an entire new field of study in genetics, biology, and biochemistry as researchers strove to understand the mechanisms of RNAi. A monumental study by Fire and Mello in *C. elegans* then clearly demonstrated that

double stranded (ds)RNA was in fact the trigger responsible for gene silencing (Fire et al., 1998), a finding for which they were awarded the Nobel prize in Physiology or Medicine in 2006. The repertoire of model organisms in which RNAi was found soon expanded to include yeast, flies, trypanosomes, planaria, zebrafish and mice (Kennerdell and Carthew, 1998; Misquitta and Paterson, 1999; Ngo et al., 1998; Sanchez Alvarado and Newmark, 1999; Volpe et al., 2002; Wargelius et al., 1999; Wianny and Zernicka-Goetz, 2000). Importantly, multiple lines of evidence supported the theory that RNAi functioned at the post-transcriptional level, since promoter or intronic sequences did not trigger silencing (Fire et al., 1998), individual mRNAs arising from a polycistron could be targeted separately (Montgomery et al., 1998) and the transcription of a target gene was unaffected by dsRNA-induced silencing (Jones et al., 2001).

On the heels of the seminal Fire et al. (1998) paper, it was shown in multiple systems that the initial trigger long dsRNA was the precursor to a shorter species of RNAs termed small interfering (si) RNAs of 21-25 nucleotides (nt) in length, which were the

ultimate RNA targeting molecules (Elbashir et al., 2001; Hamilton and Baulcombe, 1999; Hammond et al., 2000; Zamore et al., 2000). Elbashir et al. further determined that siRNAs of 21-22nt length with 2nt overhangs at the termini were most effective at silencing complementary mRNAs, through endonucleolytic cleavage within the siRNA-complementary region (Elbashir et al., 2001). This target cleavage or "slicer" activity was attributed to a complex of as yet unidentified proteins termed the RNA-induced Silencing Complex (RISC) (Elbashir et al., 2001; Hammond et al., 2000; Martinez et al., 2002; Nykanen et al., 2001).

Delving further into the mechanism of RNAi, researchers identified key genes involved in the biogenesis of siRNAs and in the RISC-mediated effector steps of target silencing. Powerful forward genetic screens performed in *C. elegans*, *Neurospora*, and *Arabidopsis* were able to identify a host of genes required for RNAi (Cogoni and Macino, 1997; Elmayan et al., 1998; Fagard et al., 2000; Tabara et al., 1999) , and further genetic and biochemical studies were able to flesh out the mechanism.

A landmark study by Bernstein *et al*. in *Drosophila* cell extracts detailed the discovery of the enzyme responsible for the generation of siRNAs from long dsRNA precursors, Dicer. Dicer was identified by a candidate-based approach testing predicted ribonuclease (RNase) III enzymes in their ability to specifically cleave long dsRNA into discrete populations of 21-23nt siRNAs, and was biochemically separable from RISC-mediated target destruction activities (Bernstein et al., 2001). Dicer was quickly found to be highly conserved in numerous other model organisms including *C. elegans* (Grishok et al., 2001; Ketting et al., 2001; Knight and Bass, 2001), plants (Schauer et al., 2002), zebrafish (Wienholds et al., 2003), and mice (Bernstein et al., 2003). The literature on Dicer mechanism of action is more thoroughly examined below in Section I.V.

Finding the slicer enzyme of the RISC provided another leap forward in our understanding of RNAi processes. The Argonaute (AGO) proteins had been implicated in both genetic and biochemical studies as core RISC components across species (Catalanotto et al., 2000; Fagard et al., 2000; Hammond et al., 2001; Martinez et al., 2002; Tabara et al., 1999). Structural studies next revealed that AGO proteins bound 3’

ends of single-stranded nucleic acids in their PAZ domain, implicating them in binding the 2nt 3' overhangs present in siRNAs (Lingel et al., 2003, 2004; Ma et al., 2004; Song et al., 2003; Yan et al., 2003). Thorough structural analysis of another AGO domain (the PIWI) uncovered a protein fold highly similar to RNase H enzymes, which were known to cleave RNA present in RNA-DNA hybrids and generate products bearing 3'OH and 5' phosphate groups - the same type of products generated by the RISC (Song et al., 2004). Several groups were then able to convincingly demonstrate that AGO proteins contained slicer activity using *in vitro* target cleavage assays (Liu et al., 2004; Meister et al., 2004; Song et al., 2004). The literature on AGOs, with an emphasis on *C. elegans* AGOs and alternate mechanisms of silencing, is more thoroughly examined below in Section I.VI.

A critical step in the mechanism of RNAi is the process by which siRNA duplexes generated by Dicer are passed on to the AGO-containing RISC (termed RISC loading), following which one strand of the duplex is discarded (passenger strand) and one strand remains embedded in the AGO (guide strand) to pair via homologous sequence to a

target mRNA. In fact, it was found that the AGO plays a key role in the strand selection process by preferentially using the RNA with lowest 5' end thermodynamic stability to serve as the guide, and cleaving the passenger strand of the siRNA duplex (Khvorova et al., 2003; Matranga et al., 2005; Rand et al., 2005). Despite these critical functions, AGO does not always carry out silencing alone. Other key members of the core RISC or "holo-RISC" include Dicer itself and dsRNA binding domain (dsRBD)-containing proteins R2D2 (*Drosophila*), RDE-4 (*C. elegans*), and TRBP or PACT (*H. sapiens*) (Chendrimada et al., 2005; Gregory et al., 2005; Pham et al., 2004; Tabara et al., 2002). Importantly, TRBP was found to be necessary to recruit AGO to Dicer (Chendrimada et al., 2005) and it was shown that the Dicer-TRBP-AGO complex is sufficient to carry out RNAi *in vitro* (Gregory et al., 2005).

Described above are snapshots of cellular responses to the addition of exogenous dsRNA. This pathway is what is commonly referred to as the "classical RNAi" or the "exoRNAi" pathway. But what, if any, biological roles does the exoRNAi pathway serve? Evidence for this came from early studies in which genes required for exoRNAi were

also shown to protect the integrity of the genome against transposons and repetitive elements (Ketting et al., 1999; Tabara et al., 1999). The exoRNAi pathway was also shown to serve as an antiviral defence mechanism in many species (Ashe et al., 2013; Felix et al., 2011; Galiana-Arnoux et al., 2006; Ratcliff et al., 1999; Saleh et al., 2009; van Rij et al., 2006; Wilkins et al., 2005). In *C. elegans*, in addition to a specialized AGO (RDE-1) and its dsRBD partner RDE-4, the function of the exoRNAi pathway requires many other genes (See Appendix 4).

I.II Introduction to the miRNA pathways

Soon after the initial discovery of RNAi, it became clear that the exoRNAi pathway was the tip of the iceberg in the magnitude of small-RNA directed gene silencing pathways. Multiple related pathways driven by endogenous small RNAs were found, and classified based on their origin and mechanism of silencing. First, an abundant class of genome-encoded small RNAs called micro(mi)RNAs were discovered and are now widely recognized to play important roles in virtually all gene regulatory networks, in species ranging from plants to humans. The first miRNA, lin-4, was in fact discovered prior to the Fire and Mello breakthrough, in *C. elegans* by the Ambros and Ruvkun groups (Lee

et al., 1993; Wightman et al., 1993). lin-4 was shown to be a short non-protein coding gene specific to nematodes which generated two species of RNA called the long (61nt) and short (22nt), later renamed the precursor and mature, which negatively regulated the level of LIN-14 protein expression, contributing to proper developmental timing of the animal. Curiously, the lin-4 small RNA displayed homology to multiple sites in the 3' untranslated region (UTR) of the *lin-14* mRNA, leading the groups to propose a model by which lin-4 downregulated LIN-14 expression by binding through complementary sequence to its mRNA and preventing its translation (Lee et al., 1993; Wightman et al., 1993). The identification of another small endogenously-derived RNA conserved in many species called let-7 (Pasquinelli et al., 2000; Reinhart et al., 2000), combined with the mechanistic strides made in the exoRNAi field, and the advent of next-generation sequencing technologies next spurred a wave of discovery. Hundreds of miRNAs were identified in *C. elegans*, plants, *Drosophila*, mice, and human cells (Lagos-Quintana et al., 2001; Lagos-Quintana et al., 2003; Lagos-Quintana et al., 2002; Lau et al., 2001; Lee and Ambros, 2001; Lim et al., 2003a; Lim et al., 2003b; Llave et al., 2002a;

Mourelatos et al., 2002; Park et al., 2002; Reinhart et al., 2002). These novel endogenous small RNAs were found to arise from primary transcripts by RNA Polymerase II, regions of which formed hairpins via intramolecular base-pairing (Lee et al., 2004a). The hairpins (precursor (pre-) miRNAs) are excised from the primary transcript by the Microprocessor complex containing the RNase III Drosha in the nucleus before being exported to the cytoplasm, where Dicer removes the terminal loop (with rare exceptions: (Cheloufi et al., 2010; Cifuentes et al., 2010)) and loads the mature miRNA into the miRNA-induced silencing complex (miRISC) (Figure I.I, (He and Hannon, 2004; Lai, 2003; Lee et al., 2003; Lee et al., 2002)).

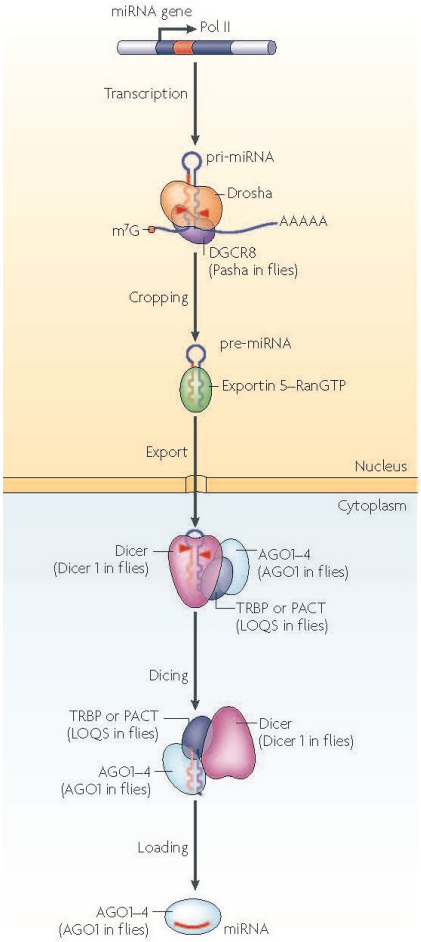


Figure I.I: Biogenesis of miRNAs. Sourced with permission from (Kim et al., 2009). Typically, miRNA genes are transcribed by RNA Polymerase II into primary transcripts (pri-miRNA), which are bound and processed by the Microprocessor complex containing the RNase III Drosha into the pre-miRNA. The pre-miRNA is exported from the nucleus where Dicer removes the terminal loop and loads the mature miRNA onto the AGO-containing RISC. The RISC then goes on to target and silence mRNAs.

miRNAs are similar in size to siRNAs (~22nt), but most miRNAs base-pair imperfectly and within a short “seed” region (nucleotides 2-8 from the 5’ end) in the 3’ untranslated region (UTR) of the target mRNA, with the exception of plant miRNAs which base-pair perfectly to their targets (Figure I.II, (Llave et al., 2002b; Olsen and Ambros, 1999; Rhoades et al., 2002; Seggerson et al., 2002)).

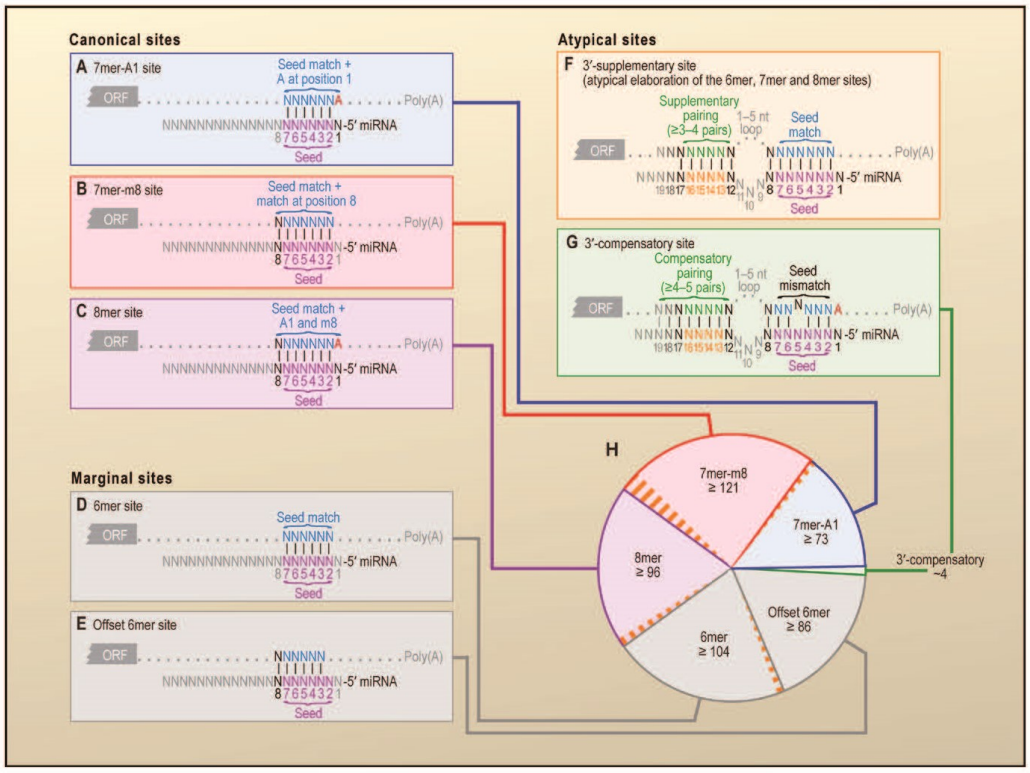


Figure I.II: Types of miRNA-target interactions. Sourced with permission from (Bartel, 2009). Shown are canonical sites (A-C) with full seed-only matches, marginal sites (D-E) with reduced seed matches, and atypical sites (F-G) that contain 3’ pairing. (H) From 87 conserved vertebrate miRNA families, the number of miRNA sites that fall into the

above categories (Friedman et al., 2009). Orange hatched regions indicate those with 3' supplementary pairing.

Therefore, due to the short region of homology required between the guide and target, individual miRNAs can regulate many more different genes and individual genes may be regulated by many different miRNAs. For example, let-7, whose seed sequence is invariable from *C. elegans* to humans, has been shown to target *lin-41* (Kanamoto et al., 2006; Maller Schulman et al., 2008; Pasquinelli et al., 2000; Reinhart et al., 2000; Slack et al., 2000; Vella et al., 2004), *daf-12* (Grosshans et al., 2005), *HMGA2* (Boyerinas et al., 2008), *MYC* (Sampson et al., 2007), and *let-60/RAS* (Johnson et al., 2005) among many others. On the other hand, *lin-41* can be targeted by not only let-7 at two sites in its 3' UTR, but by lin-4 as well (Slack et al., 2000). Several complementary computational methods have been developed to predict which mRNAs can be targeted by a specific miRNA, and vice versa (Enright et al., 2003; John et al., 2004; Kiriakidou et al., 2004; Lewis et al., 2003; Rajewsky and Socci, 2004; Stark et al., 2003). Using such methods, researchers estimate that more than 60% of human protein-coding genes are

under the direct control of miRNAs (Friedman et al., 2009). This is supported by quantitative proteomic studies (Baek et al., 2008; Selbach et al., 2008). Loss of some miRNAs, like lin-4 and let-7 who have cell-type specific and temporal expression patterns, have profound effects on the development and viability of the organism through dramatic mis-regulation of their targets (Lee et al., 1993; Moss et al., 1997; Reinhart et al., 2000; Wightman et al., 1993). Others have been shown to play pivotal roles in cell proliferation, death, metabolism, and fate determination (Brennecke et al., 2003; Chen et al., 2004; Johnston and Hobert, 2003; Xu et al., 2003). On the other hand, miRNAs do not always act as on-off switches, many miRNAs act as fine-tuners of gene expression, eliciting subtle effects on target genes at the RNA and protein level (Baek et al., 2008; Mourelatos, 2008; Selbach et al., 2008). These seemingly contradictory functions emphasize the complexity of miRNA-mediated gene silencing. As each target gene has its distinct expression profile and tissue specificity, so does each miRNA gene. Furthermore, each miRNA-target relationship is unique in terms of the degree of silencing that is exerted to achieve its physiological purpose. These

considerations amount to highly context-dependent regulatory functions of miRNAs in gene expression.

I.III Introduction to the endoRNAi pathways

To add another layer of complexity to the rapidly emerging field of non-coding small regulatory RNAs, another class of endogenous small RNAs distinct from miRNAs called endogenous (endo) siRNAs were discovered in parallel large-scale sequencing efforts. endo-siRNAs are not derived from hairpin precursors like miRNAs and were first found in the fission yeast *Schizosaccharomyces pombe*, *Arabidopsis thaliana*, and *C. elegans*. *S. pombe* does not carry any detectable miRNAs, but does have Dicer and AGO homologs. In the fission yeast, the RNAi machinery is almost exclusively nuclear and functions in maintaining heterochromatin at peri-centromeric regions and at a select group of coding loci. Heterochromatin regions are marked by histone H3K9 methylation, which is recognized by the chromodomain protein HP1/SWI6 (a homolog of HPL-2 in *C. elegans*), which establishes and maintains its heterochromatin state (Volpe et al., 2002). When heterochromatin regions are abnormally transcribed, an Argonaute complex termed the RNA induced transcriptional silencing (RITS) (Buker et al., 2007; Verdel et

al., 2004) complex associates with nascent transcripts through base-pairing of its small RNA, and recruits an RdRP complex (RDRC) to generate dsRNA (Colmenares et al., 2007). This, in turn, is a substrate for *S. pombe* Dicer, and generates more siRNAs specifically targeting the leaking locus (Motamedi et al., 2004). There, the RITS recruits the histone methyltransferase Clr4, the catalytic subunit of the CLRC complex, to the target locus to specifically methylate more H3K9 sites (Zhang et al., 2008). The RITS complex is not only responsible for directing histone marks by recruiting CLRC, it also displays affinity for the accumulated H3K9Me modifications through the chromodomain protein Chp1. The dual interaction of the RITS and heterochromatin loci through both siRNAs and Chp1 allows continuous association with H3K9me marks, to survey heterochromatin domains for faulty transcription. This mechanism thus acts as a "self-enforcing loop" to ensure heterochromatin maintenance in *S. pombe.* Importantly, the self-enforcing loop model explains maintenance of heterochromatin domains, and not how these domains were determined and initiated in the first place, or if RNAi was involved in those steps.

In plants, many non-miRNA classes of endo-siRNAs have been found, with the largest group composed of RNAs that are required similarly to methylate histones and DNA, and drive TGS of transposons and repeat elements (Bologna and Voinnet, 2014; Kasschau et al., 2007; Sunkar and Zhu, 2004). In *C. elegans*, hundreds of small antisense RNAs with exact complementarity to over 500 protein coding genes were discovered soon after the initial identification of endo-siRNAs in yeast, and this class was later expanded (Ambros et al., 2003; Ruby et al., 2006). The exact functions of many of these endo-siRNAs remain elusive, as they map to both silent and actively transcribed regions of the genome. One common element in the yeast, plant and worm systems is the general requirement for RdRP activity to generate the dsRNA triggers to be processed by Dicer proteins. This feature precluded the immediate identification of similar pathways in mammals, since no RdRP homologs have been found. However, not all endo-siRNAs require the activity of an RdRP to produce trigger dsRNA. These triggers could arise from a region within an RNA with extended complementarity, distinct RNAs with extended complementarity, or RNAs generated by convergent transcription,

as in the case of one class of plant siRNAs, the *cis* natural antisense transcript RNAs (Borsani et al., 2005). In fact, convergent overlapping gene pairs occur quite frequently in humans (4-9% of all genes), *Drosophila* (22%) and *Arabidopsis* (10%) (Boi et al., 2004; Jen et al., 2005; Wang et al., 2005). Thus, it is possible for mammals to produce endo-siRNAs from these types of RNAs, and in 2008, multiple groups demonstrated this in *Drosophila* and mice (Czech et al., 2008; Ghildiyal et al., 2008; Kawamura et al., 2008; Okamura et al., 2008a; Okamura et al., 2008b; Tam et al., 2008; Watanabe et al., 2008). These endo-siRNAs were found to silence mobile genetic elements and some protein-coding mRNAs in the germline and somatic cells through sequencing of AGO-bound RNAs (Czech et al., 2008; Kawamura et al., 2008), by computational prediction followed by detection (Okamura et al., 2008a; Okamura et al., 2008b), by searching for endogenous siRNAs with 3' methylation (a modification also seen on siRNAs downstream of the exoRNAi pathway) (Ghildiyal et al., 2008), or by deep sequencing to discover siRNA populations distinct from miRNAs and piRNAs (See Section I.IV for more information on piRNAs) (Tam et al., 2008; Watanabe et al., 2008). One clear

function for endo-siRNAs in multiple systems is to silence transposons, but the mechanisms by which they accomplish this in animals and their role in the regulation of protein-coding genes is currently ambiguous.

I.IV Introduction to the piRNA pathways

Another class of small RNAs found in animals and expressed in germline cells are the Piwi-interacting (pi)RNAs. piRNAs are so named because of their association with a specific type of AGO of the Piwi clade (discussed further in Section I.VI). piRNAs are generally longer than miRNAs (between 24 and 29nt), with the notable exception of *C. elegans* piRNAs which are 21nt, and have important functions in genome surveillance and fertility (Aravin et al., 2006; Batista et al., 2008; Das et al., 2008; Girard et al., 2006; Grivna et al., 2006; Lau et al., 2006). piRNAs act through both transcriptional and post-transcriptional gene silencing mechanisms on transposons and some protein-coding genes (Brennecke et al., 2007; Le Thomas et al., 2013; Sienski et al., 2012). The origins of piRNAs are not fully understood or consistent between organisms, but one important common feature is that piRNA biogenesis is independent of Dicer. In *Drosophila*, piRNAs are derived from long transcripts of piRNA clusters in genomic regions devoid

of protein-coding genes, which are matured by sequential cleavages by an endonuclease to generate the 5' end, loaded onto a specialized AGO, and trimmed and modified to produce the 3' end in specialized cellular compartments called nuage granules (Haase et al., 2010; Ipsaro et al., 2012; Kawaoka et al., 2011; Nishimasu et al., 2012; Olivieri et al., 2010). The piRNA population is then maintained and amplified in a cycle of target cleavages by the piRNA-loaded AGOs termed the ping-pong model (Aravin et al., 2006; Brennecke et al., 2007; Gunawardane et al., 2007).

Mammalian piRNAs have been identified and originate in genomic clusters similar to *Drosophila*. Mouse piRNAs require the Piwi Argonaute MIWI for their biogenesis, accumulate at the onset of meiosis in male germ cells and are essential for spermatogenesis (Aravin et al., 2006; Girard et al., 2006; Grivna et al., 2006).

In *C. elegans*, piRNAs are called 21U RNAs due to their size (21nt) and bias for a uracil in the first position of the small RNA. They arise from thousands of unique and independent genomic loci, typically intergenic or intronic, are distributed on a single chromosome and share a characteristic upstream sequence motif (Batista et al., 2008;

Ruby et al., 2006). *C. elegans* piRNAs were also shown to associate specifically with the Piwi Argonaute PRG-1, loss of which leads to reduction of germ nuclei and defects in fertility (Batista et al., 2008; Cox et al., 1998). Similar to other systems, piRNAs target transposon expression and mobility, but through a unique mechanism of inducing the expression of DCR-1 (*C. elegans* Dicer)-independent endo-siRNAs called 22Gs which amplify the silencing signal in a distinct but analogous manner to the *Drosophila* ping-pong model (Batista et al., 2008; Das et al., 2008). The *C. elegans* piRNA pathway is also required for the expression of some spermatogenesis-specific mRNAs in the male gonad, and therefore may play a larger role in germ cell gene expression in addition to transposon silencing (Wang and Reinke, 2008).

I.V Dicer features and mechanisms

Since the initial discovery in 2001 (Bernstein et al., 2001), Dicer and its mechanisms of dsRNA cleavage have been the subject of intense study in many systems. Dicer proteins are highly conserved and specialized members of the RNase III superfamily, a group of enzymes that specifically cleave dsRNA substrates from bacteria to humans. The unifying feature of Dicers is the presence of at least one RNase III domain, which is

responsible for dsRNA cleavage. The majority of Dicers possess two RNAse III domains, arranged in tandem, followed by a dsRNA binding domain (dsRBD). Dicers from the parasite *Giardia intestinalis* up to humans also have a PAZ domain, named for homology to domains also found in Piwi/Argonaute/Zwille proteins, and important for dsRNA end recognition. Increased domain complexity is seen in Dicers from plants to humans with the presence of N-terminal large helicase domains and DUF283 (conserved but domain of unknown function) domains (Figure I.III).

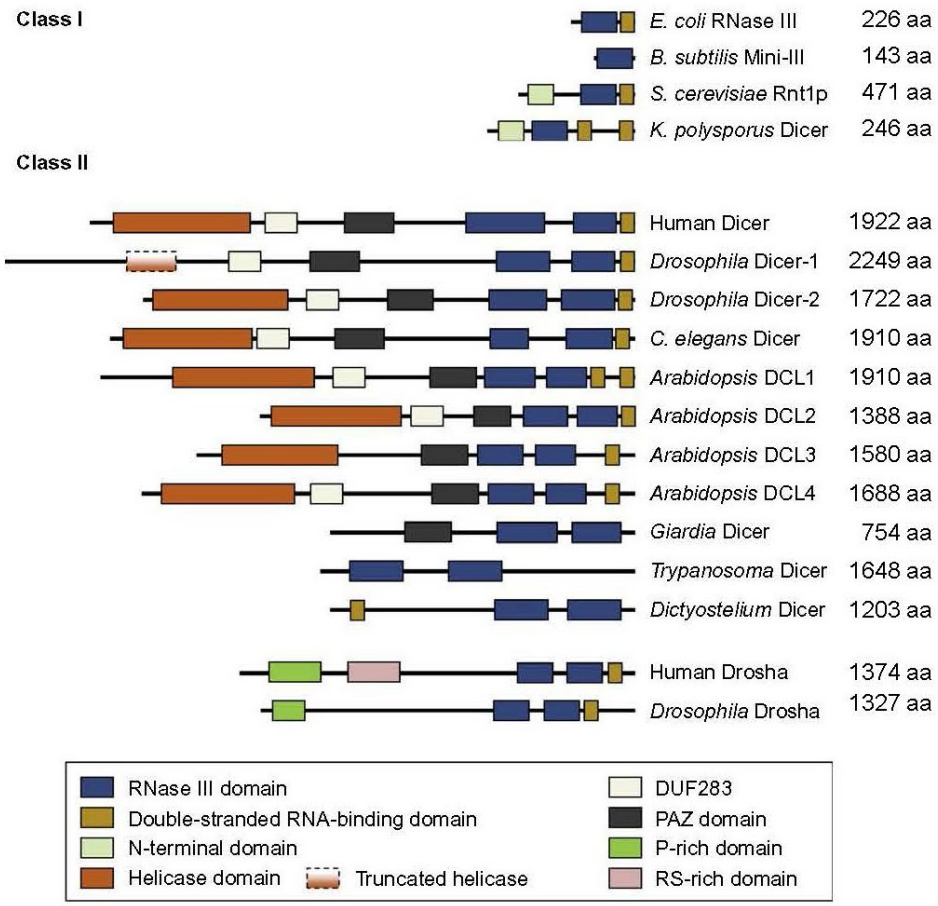


Figure I.III: Schematic representations of RNase III superfamily members. Sourced with permission from (Doyle, 2012), adapted to include protein sizes in amino acids (aa). RNase III enzymes are divided into two classes, those with a single RNase III domain versus those with multiple. Shown are domain composition and arrangement.

Loss of Dicer is lethal in *C. elegans*, *Drosophila*, and mice. This has been largely attributed to global defects in miRNA biogenesis, resulting in pleiotropic phenotypes, arrested development, and ultimately death (Bernstein et al., 2003; Grishok et al., 2001;

Hutvagner et al., 2001; Ketting et al., 2001; Knight and Bass, 2001; Lee et al., 2004b; Wienholds et al., 2003). This assumption is supported by the fact that loss of the Dicer-dependent exoRNAi and endoRNAi pathway specialized components are generally tolerated, while loss of core miRNA pathway components can also be lethal.

Some organisms have evolved multiple Dicer genes with altered domain composition, which ostensibly reflect different functions for Dicers in the different RNAi pathways (Figure I.IX). For example, *Arabidopsis* carry four Dicers (DCL1-4) and *Drosophila* carry two (Dicer-1 and Dicer-2), while *C. elegans* (DCR-1), *S. pombe* (Dicer) and humans (Dicer) possess only one gene. Single gene scenarios, however, do not preclude the possibility of diversification in Dicer proteins through alternative splicing or post-translational modification. In plants, DCL-1 is dedicated to the miRNA pathway and produces 21nt RNAs, while DCL2-4 are dedicated to the exo- and endoRNAi pathways and generate 22nt, 24nt and 21nt products (Qi et al., 2005). Similarly, *Drosophila* Dicer-1 is dedicated to the miRNA pathway and Dicer-2 to exo- and endoRNAi pathways (Lee et al., 2004b). How this pathway dedication arises is often through specific protein-

protein interactions. For example, in *Drosophila*, both Dicers are capable of processing pre-miRNA substrates *in vitro*, however, *in vivo* the dsRBD protein R2D2 binds specifically to Dicer-2 and collaborates with physiological concentrations of inorganic phosphate to restrict Dicer-2 from participating in the miRNA pathway (Cenik et al., 2011). Similarly, plant DCL-4 has its own dsRBD protein partner in DRB4, while DCL-1 specifically interacts with another dsRBD protein, HYL1. These specific interactions take place due to sequence divergence between the DUF domains of DCL1 and DCL-4 (Qin et al., 2010). It is therefore critical to understand the functions and contributions of each Dicer domain in order to gain insight into its mechanism of small RNA biogenesis in the different RNAi pathways.

Dual RNase III domains

Early biochemical studies on recombinant human Dicer revealed that the enzyme preferentially cleaves dsRNA from its termini in a Mg^{2+}-dependent reaction (Provost et al., 2002; Zhang et al., 2002). Subsequently, it was found that the RNase III (RIII) domains form an intramolecular dimer with slightly staggered active sites containing

acidic residues to co-ordinate Mg^{2+} and sever the phosphodiester bond, and each domain is responsible for the cleavage of one strand of the dsRNA substrate. This results in the generation of siRNAs with the 2nt 3' overhangs, 5' phosphates and 3' OH groups characteristic of RNase III products (Gan et al., 2008; Macrae et al., 2006; Takeshita et al., 2007; Zhang et al., 2004). It was also found that the substrate strand carrying the 3' OH (3p strand) is cleaved by the RIIIa domain, whilst the strand carrying the 5' phosphate (5p strand) is cleaved by the RIIIb. This feature is especially important in the case of pre-miRNA substrates, since defects in the RIIIa causes the selective loss of miRNAs originating from the 3' OH strand, and defects in the RIIIb causes loss of miRNAs from the 5' phosphate strand (Gurtan et al., 2012; Ye et al., 2007; Zhang et al., 2004). The RIIIa domain is also a site of direct interaction between Dicer and the PIWI domain of Argonautes *in vitro* (Sasaki and Shimizu, 2007), and this interaction likely plays a key role in the transfer of Dicer products into the Argonaute for downstream targeting (RISC loading).

Helicase domain

The large N-terminal helicase domain of Dicer was so named due to its homology with DExD/H-box helicases, which play roles in RNA metabolism by unwinding duplexes and translocating along a substrate (Fuller-Pace, 2006; Lee and Hurwitz, 1992). To date, no helicase activity has been described for Dicer, however it has been shown that the helicase domain is key to substrate recognition and Dicer's ability to cleave long dsRNA substrates in a processive manner. Using *C. elegans* extracts and purified *Drosophila* Dicer-2, the helicase was found to promote processive ATP-dependent cleavage of substrates with blunt or 5' overhanging termini, but not those with the 3' overhangs characteristic of pre-miRNAs (Cenik et al., 2011; Sinha et al., 2015; Welker et al., 2011). In line with this, mutations which abrogate ATP hydrolysis or helicase activity in related enzymes resulted in the loss of a subset of endo-siRNAs called 26Gs (26nt long with 5'G), but not miRNAs or exo-siRNAs when introduced into *C. elegans* Dicer (DCR-1). Moreover, the helicase of *Drosophila* Dicer-1 is incapable of ATP hydrolysis, and this allows it to make specific contacts with its substrate pre-miRNA loop, whereas Dicer-2 can hydrolyze ATP and cleave long dsRNAs processively (Tsutsumi et al., 2011). The

helicase has also been shown to reduce the catalytic activity of the enzyme using recombinant human Dicer *in vitro* cleavage assays. This auto-inhibition was somewhat alleviated in the presence of the dsRBD protein partner TAR-RNA binding protein (TRBP) which binds to the helicase (Daniels et al., 2009; Ma et al., 2008). Thus, the autoinhibition seen may reflect the need for cofactors to achieve optimal cleavage activity, perhaps through inducing conformational changes in Dicer's helicase. The helicase can be seen as a protein scaffolding module in other organisms as well. In *C. elegans*, there is a specialized complex of proteins whose presence and interaction within the DCR-1 helicase is required for the generation of primary endo-siRNAs called the ERI complex (Duchaine et al., 2006; Thivierge et al., 2012).

DUF283 domain

The DUF283 domain is specific to Dicers but has not been extensively studied. However, one group in *Arabidopdis* has shown the DUF283 adopts an atypical dsRNA binding fold and displays weak RNA binding activity. Further, the DUFs of different plant Dicers (DCL-1 and DCL-4) mediate protein-protein interactions with their respective

dsRBD protein partners (Qin et al., 2010). Thus, the DUF283 may act to mediate protein-protein interactions in Dicers of other species as well.

The ability of Dicer proteins to generate products of discrete sizes is crucial to the function of the RNAi pathways, as changes in the register of the si/miRNAs could have profound impacts on the repertoire of mRNAs that are subsequently targeted. How Dicer measures its substrates is therefore key to the understanding of its role. The landmark finding on this topic came from structural and biochemical studies of *Giardia* Dicer, which is composed of a PAZ, a largely helical linker region, and tandem RIII domains. Modeling a dsRNA substrate into the crystal structure revealed that the distance between the substrate 3' end docking at the PAZ and the sites of RNA cleavage in the active site was "measured" by the intervening linker, establishing the Dicer protein itself as a molecular ruler (Figure I.IV, (MacRae et al., 2007; Macrae et al., 2006)). How the substrate is docked thereby became an important consideration in size determination of the products generated, and from early on the PAZ domain was shown to be critical to this function.

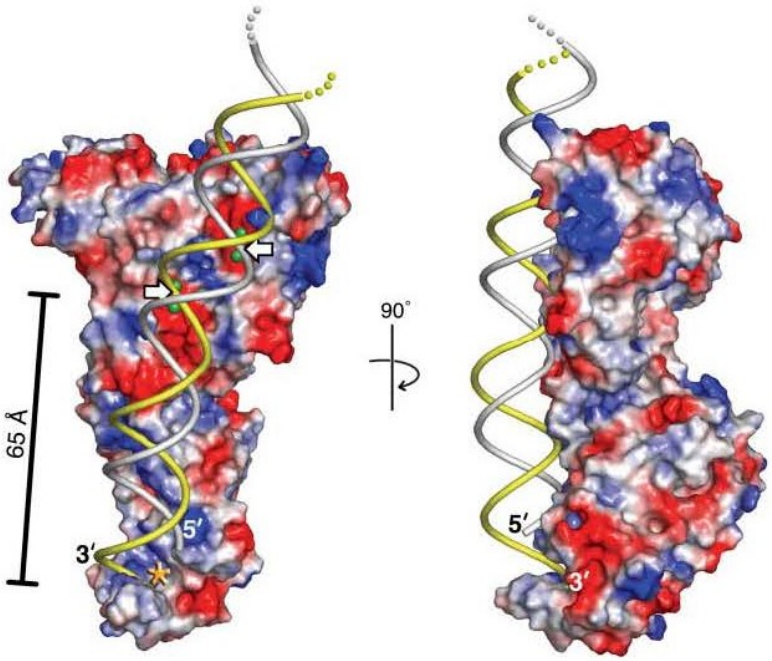


Figure I.IV: Structure-based dsRNA cleavage model of *Giardia* Dicer. Sourced with permission from (Macrae et al., 2006). In these orientations, the PAZ is at the bottom of the crystal binding the 3' end of a modeled dsRNA substrate (asterisk), and the dual RNase III domains at the top, cleaving both strands of the substrate at their active sites (white arrows). The distance between the PAZ and RNase III contacts corresponds to the size of siRNA produced.

PAZ domain

The PAZ domain of Dicer is one of the most well-studied of its domains. The PAZ is also common to Argonaute proteins (further discussed in Section I.VI below) and is responsible for binding the 3' terminus of the dsRNA substrate. Mutations of the conserved residues in human Dicer's PAZ result in the loss of substrate binding, and

subsequent processing (Zhang et al., 2004). Furthermore, initial crystal structures like those of *Giardia* Dicer have allowed RNA substrates to be modeled into the domain's folds, establishing a putative pocket that binds the 2nt overhanging residues typical of pre-miRNA substrates (Lingel et al., 2004; Ma et al., 2004; Macrae et al., 2006). More recently, a crystal structure was completed of the RNA-bound human PAZ domain flanked at the 5' by a structure termed the "platform" and at the 3' by a helix connecting the PAZ to the RNase IIIa (Tian et al., 2014). RNA substrates with 2nt 3' overhangs (like pre-miRNAs) were found to contact 7 residues called the 3' pocket of the PAZ through intermolecular hydrogen bonds, and mutations of these residues significantly diminished substrate binding and cleavage (Park et al., 2011; Tian et al., 2014). The platform has also been implicated in substrate end recognition using *in vitro* cleavage assays, specifically with the phosphate group on the end of the 5' strand, in cases where the substrate ends differ from the optimal 2nt 3' overhang (Park et al., 2011). The 5' phosphate-binding pocket responsible for this was found to be a patch of basic residues in the crystal structure adjacent to the 3' pocket and separated by short helical segment

of the PAZ (Tian et al., 2014). It is unclear at this point what the physiological significance of this 5' pocket is, since Dicer crystals with 5' phosphate ends bound in the pocket were grown in the presence of atypical blunt substrates with the terminal 2nt unpaired, and no crystals of 5' pocket mutants were produced. A thorough cataloguing and analysis of natural Dicer substrates with such a structure has not been completed, and it remains a possibility that *in vivo* in the presence of canonical 2nt 3' overhang substrates, this 5' pocket co-ordinates binding to an inorganic phosphate molecule, as found in some crystals. The short helical segment found to separate the 5' and 3' pockets was found to be disordered or melted in some cases, changing the angle of the substrate with respect to the protein, while having little effect on substrate binding or cleavage. The authors postulate that the disordered state, which positions the substrate closer to the plane of the protein could represent a cleavage competent state, while the helical form, which angles the substrate farther away, could represent a product release or transfer state (Tian et al., 2014).

Building on the initial work in *Drosophila* that found the helicase of Dicer-1 contacts the pre-miRNA loop (Tsutsumi et al., 2011), one group found that using human Dicer, the loop position relative to the active site of RNA cleavage also affected the size of miRNAs produced (Gu et al., 2012a). Of note is the fact that atypical substrates with 4nt 3' overhangs were used in this study.

Nevertheless, what has emerged from these studies is a more thorough understanding of substrate measurement and a set of criteria that govern the size of products. These were termed the "5' counting", "3' counting" and "loop counting" rules, and it is clear that substrate structure plays an important role in determining which rule or rules are followed.

dsRBD domain

Finally, the C-terminal dsRBD of Dicer has been shown to adopt the canonical dsRBD structure of α-β-β-β-α (Du et al., 2008). Early work showed that the dsRBD of human recombinant Dicer was dispensible for its activity in cleavage assays (Zhang et al., 2004). Further *in vitro* work showed that the dsRBD is only required for stabilizing the

Dicer-RNA interaction in the absence of a PAZ domain, and does not contribute to the size determination of products (Ma et al., 2012b). What then is the purpose *in vivo* of the dsRBD? Clues can be found in the studies of *S. pombe* Dicer, which is predominantly nuclear and required for the production of heterochromatic siRNAs. It was found that this nuclear localization was a function of a nuclear retention signal within the dsRBD of *S. pombe* Dicer (Emmerth et al., 2010). Recently, human Dicer was also found to contain a nuclear localization signal in its dsRBD (Doyle et al., 2013), suggesting a possibly conserved nuclear function similar to *S. pombe*, which was later substantiated (White et al., 2014).

Could the dsRBD also contribute to other, non-RNAi, functions of the protein? Both human and *C. elegans* Dicer have recently been shown to bind a wide variety of RNA substrates: miRNAs, tRNAs, snoRNAs, mRNAs, and promoter RNAs. Surprisingly, Dicer does not generate siRNAs from some targets and therefore regulates their expression passively (Rybak-Wolf et al., 2014). It would be interesting to determine how

Dicer binds but does not cleave these passive targets, and which domains are key to this function.

Domain arrangement

Since 2006 brought the first crystal structure of a primitive Dicer (Macrae et al., 2006), structural studies have been instrumental in defining the mechanisms of the enzyme across species. The structural models of full-length complex Dicers, like that of human Dicer, have evolved over time to shed light into how the domains are arranged as well as reconciled many of Dicer's biochemical attributes with its structure ((Sawh and Duchaine, 2012), Appendix 2: Turning Dicer on its head). Since the large size of complex metazoan Dicers (>200 kDa) precluded crystallography, researchers turned to electron microscopy (EM) techniques combined with crystal docking of individual domains (Lau et al., 2009; Wang et al., 2009). EM structures defined the full-length human Dicer as an "L"-shaped molecule (Figure I.V), but the placement of the individual domains was debatable. It was not until recently that one group uprooted conventional views of Dicer's architecture through domain mapping using antibodies followed by EM

(Lau et al., 2012). The current model places the PAZ domain in the head of the “L”, the helicase making up the bulk of the base, and the RIII core at the vertex (Figure I.V, I.VI).

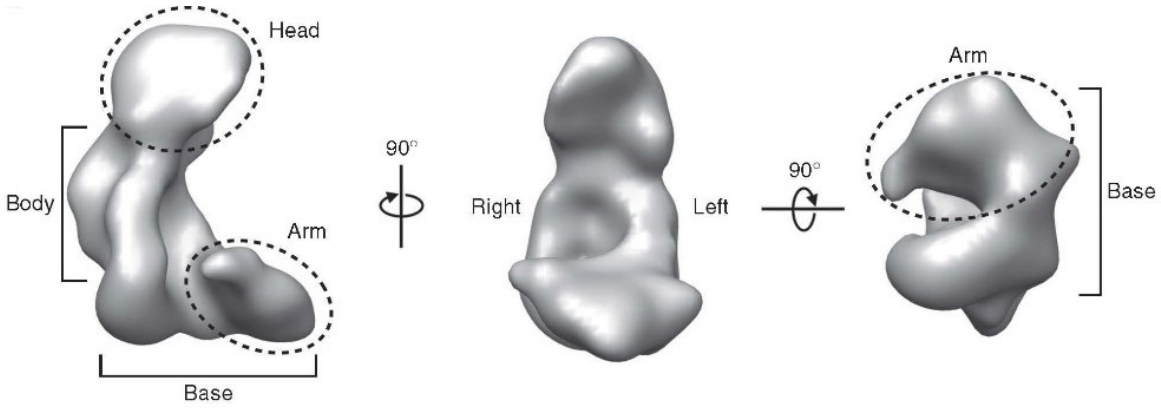


Figure I.V: EM reconstruction of full-length human Dicer. Sourced with permission from (Lau et al., 2012). Dicer is an “L”-shaped molecule with head, body, base and arm regions. There is a central channel that can accommodate dsRNA through the clamp-like base and along the body.

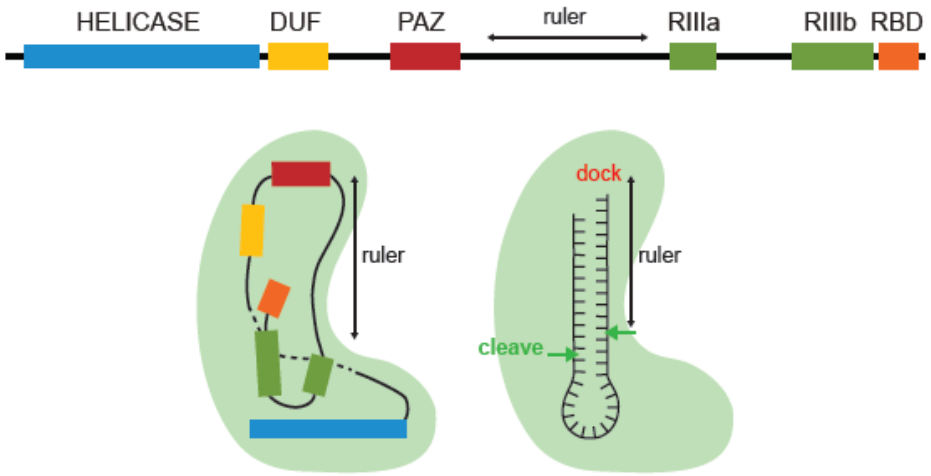


Figure I.VI: Model of domain arrangement in metazoan Dicers. Adapted from (Sawh and Duchaine, 2012), See Appendix 2. Domains are defined by sequence conservation, and a pre-miRNA substrate is shown. The PAZ, which docks the 3' end of substrates is in the head region of the enzyme, the helicase makes up most of the base, and the RNase III core is at the vertex. The ruler separates the PAZ and RNase III domains in space, effectively measuring the size of si/miRNA produced.

This domain arrangement fit well with much of the previous biochemical data. Of particular note is the placement of the clamp-shaped helicase in close proximity to the RIII core and opposite to the PAZ, which explains well the previously described autoinhibitory function (Ma et al., 2008), processive mode (Cenik et al., 2011; Welker et al., 2011), and loop-contacting (Tsutsumi et al., 2011) features for this domain (Figure I.VI, See also Appendix 2). Docking of the PAZ and RIII core in the latest EM structure also provided a structural basis for the size disparity seen in products of human and *Giardia* Dicer. Human Dicer products are likely shorter since the distance in three-dimensional space between the docking at the PAZ and the site of cleavage is smaller than in *Giardia* (Lau et al., 2012). This raises the possibility that Dicers from different

organisms can produce slightly different sized products based on the three dimensional organization of their domains. Taking this data together with EM reconstructions of the human RISC-loading complex (Dicer, TRBP and Ago2) showing interactions between Ago2 and Dicer's head and base regions (Wang et al., 2009), it seems probable that Argonautes make extensive contacts to the helicase and PAZ domains in addition to the RIIIa domain, as previously reported (Sasaki and Shimizu, 2007).

The ruler distance is not the only consideration when determining product size, however. In addition to known roles of dsRBD protein partners in altering substrate specificity, like in the case of *Drosophila* Dicer-2 and R2D2 (Cenik et al., 2011), dsRBD protein partners can alter the size of siRNAs or miRNAs produced. *Drosophila* Dicer-1 can bind to three different dsRBD partners (Loqs-PA, Loqs-PB, or Loqs-PD), and interaction with Loq-PB was found to alter miRNA size, register and subsequent target regulation compared to Dicer-1 alone or Dicer-1 in complex with Loqs-PA (Fukunaga et al., 2012). Similarly, human Dicer partners with dsRBD protein partners TRBP or PACT,

and it was recently shown that the Dicer-TRBP and Dicer-PACT complexes produce miRNAs of different sizes (Lee and Doudna, 2012; Lee et al., 2013).

I.VI Argonaute features and mechanisms

Argonautes are a conserved group of proteins common to the effector steps of all RNAi pathways. Typically, they receive the small RNA processed by Dicer through physical interaction, then go on to target mRNAs via complementary base-pairing with their embedded small RNA, and finally recruit other proteins involved in silencing. They are composed of four major domains: N-terminal, PAZ, MID and PIWI (Figure I.VII). Extensive structural studies of full-length AGO or individual domains have revealed important features of these effector proteins and their multi-functional roles in uptake of small RNAs and target recognition. The PAZ domain binds the 3' 2nt of an si/miRNA strand, making primary contacts with the RNA backbone (Lingel et al., 2004; Ma et al., 2004; Yan et al., 2003). The opposite end of the RNA is bound at the 5' phosphate by the MID domain, and some AGOs display a preference for a particular nucleotide at the first position by making specific contact with its base (Boland et al., 2011; Boland et al., 2010; Frank et al., 2010). It was also found that the first nucleotide could not play a role

in target recognition since it is buried within the MID, thereby providing structural insight into the seed recognition beginning at position 2 of the guide RNA (Lewis et al., 2005; Lewis et al., 2003).

Slicer activity is attributable to the PIWI domain, which contains the catalytic tetrad of residues DEDX (where X=D or H) that co-ordinate Mg^{2+} and cleave the passenger/target RNA leaving 5' phosphate and 3' OH groups on the products (Nakanishi et al., 2012; Song et al., 2004; Wang et al., 2008).

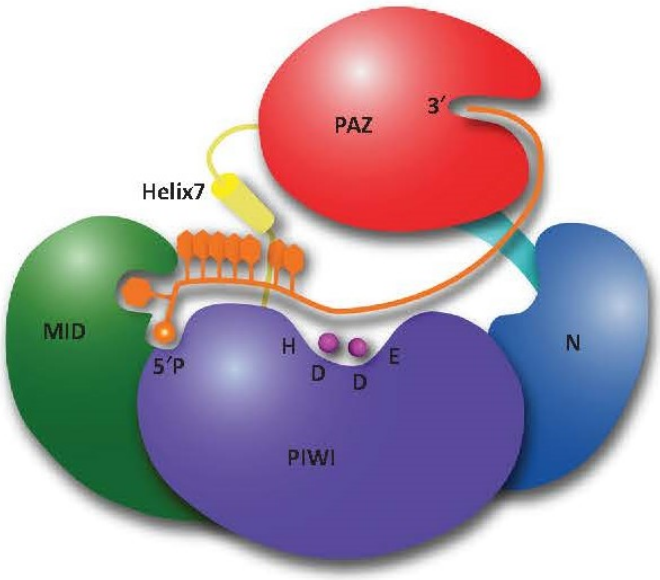


Figure I.VII: Schematic model of Argonaute domain arrangement. Sourced with permission from (Kuhn and Joshua-Tor, 2013). The PAZ domains bind the 3' end of the small RNA, the MID domain bind the 5' end, and the PIWI, with key catalytic residues

indicated, is responsible for passenger strand or target mRNA cleavage. Nucleotides 2-8 are exposed and able to base-pair to target sequences.

Among species, the number of AGO genes varies greatly (Figure I.VIII), and not all AGOs contain this catalytic tetrad in their PIWI domains. In *C. elegans*, for example, of the 27 AGOs, only 10 conserve the catalytic residues (Yigit et al., 2006). Moreover, conservation of the motif is not always sufficient to confer slicer activity. Human Ago1, 3 and 4 do not exhibit endonucleolytic activity *in vitro* even though they carry the tetrad, and it was found that structural rearrangements in the N-terminus of Ago3, compared to the catalytically competent Ago2, block slicing activity (Hauptmann et al., 2013). Slicer activity is therefore clearly not a prerequisite for gene silencing. In fact, many AGOs recruit other proteins to accomplish this. In miRNA-mediated silencing for example, the RISC bound to a target will recruit a glycine and tryptophan (GW) repeat-rich protein partner, which in turn recruits poly-A binding proteins (PABPs) and the CCR4/NOT deadenylation complex, leading to translational repression, deadenylation and destabilization of the mRNA (Baek et al., 2008; Bazzini et al., 2012; Behm-Ansmant et

al., 2006; Ding and Grosshans, 2009; Djuranovic et al., 2012; Eulalio et al., 2009; Giraldez et al., 2006; Pillai et al., 2005; Selbach et al., 2008; Wu et al., 2010; Wu et al., 2006). In *S. pombe*, the AGO is part of a specialized effector complex that mediates RNA-induced Transcriptional Silencing (RITS) which recruits a histone methyltransferase to place repressive chromatin modifications on target genomic loci, repress transcription, and to establish and maintain heterochromatin (Moazed, 2009).

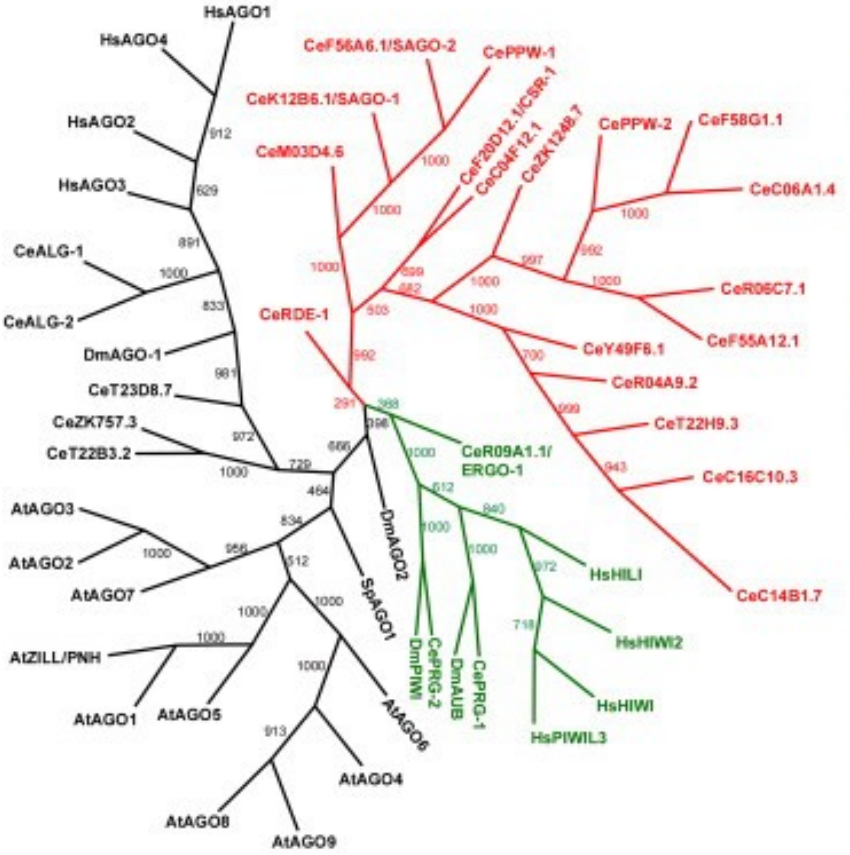


Figure I.VIII: Phylogenetic tree of AGOs among species. Sourced with permission from (Yigit et al., 2006). Genes are divided into three clades: PIWI (green), Argonaute (black) and Worm-specific WAGO (red).

In *C. elegans*, the 27 AGOs are functionally heterogeneous, and this expansion may translate to subtler, uncharacterized functions of the fewer AGOs in other systems. Two *C. elegans* AGOs are dedicated to the miRNA pathway (ALG-1/2), two to the piRNA pathway (PRG-1/2), and the remaining are members of the exoRNAi and endoRNAi pathways. Within the DCR-1-dependent exo- and endoRNAi pathways, there is a hierarchal relationship among the AGOs. RDE-1 binds primary siRNAs derived from long, exogenous dsRNA (exoRNAi) and viral triggers. RDE-1 slicer activity is required for passenger strand removal in the RISC, but not silencing of the target mRNA (Steiner et al., 2009). ERGO-1 binds 26-nucleotide long small RNAs generated by the ERI endoRNAi complex (26Gs or primary ERI endo-siRNAs) (Thivierge et al., 2012; Vasale et al., 2010), ALG-3/4 associate with a similar class of small RNAs that are sperm-specific (Conine et al., 2013), and PRG-1/2 associate with piRNAs (21Us) (Batista et al.,

2008; Wang and Reinke, 2008). Successful target recognition by primary Argonaute/siRNA results in an amplification of the signal. Secondary siRNAs are generated by the RNA-dependent RNA polymerases (RdRP) RRF-1 and/or EGO-1 and their co-factor DRH-3. Secondary siRNAs (also named 22G-siRNAs) are 22nt-long, bear a 5' tri-phosphate, and have a bias for a G in the 5'-most position (Gent et al., 2010; Vasale et al., 2010). 22G-siRNAs program the WAGO clade of AGOs (secondary AGOs) and effect gene silencing through a variety of post-transcriptional and transcriptional mechanisms (Figures I.VIII, I.IX, (Pak and Fire, 2007; Sijen et al., 2001; Tsai et al., 2015; Yigit et al., 2006)). These secondary AGOs are shared and competed for by the exoRNAi and endoRNAi pathways (Duchaine et al., 2006; Gu et al., 2009; Lee et al., 2006; Yigit et al., 2006). Since most secondary AGOs lack the key catalytic residues necessary for slicer activity, their mechanisms of target silencing are inherently non-endonucleolytic and occur in both the cytoplasm (Yang et al., 2014; Yang et al., 2012) and the nucleus (see below). Recently, genetic approaches in *C. elegans* revealed that piRNA, exoRNAi, endoRNAi, and even virus-induced RNAi all mediate

transcriptional gene silencing (TGS) through a specialized group of secondary AGOs that are nuclear, including NRDE-3 and HRDE-1 (Alcazar et al., 2008; Ashe et al., 2012; Rechavi et al., 2011; Vastenhouw et al., 2006). Another nuclear AGO, CSR-1, was even recently proposed to alleviate silencing through gene expression licensing (Seth et al., 2013).

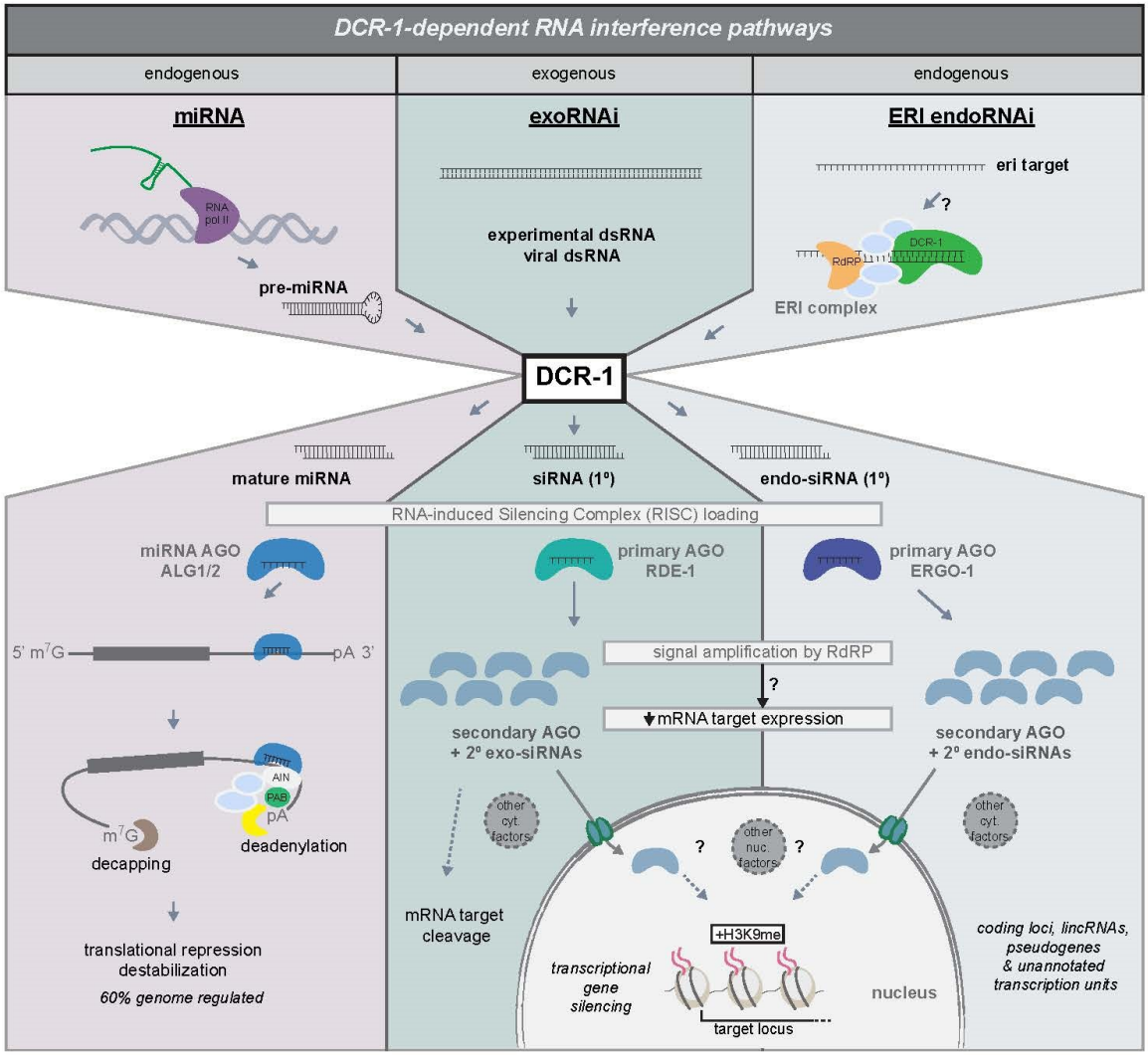


Figure I.IX: Model of DCR-1-dependent RNAi pathways in *C. elegans.* DCR-1 processes trigger dsRNAs from both exogenous (exoRNAi) and endogenous (miRNA and endoRNAi) sources and participates in the loading of miRNA-specific AGOs (ALG-1/2) and primary AGOs of the exoRNAi (RDE-1) and ERI endoRNAi (ERGO-1) pathways. Generally, miRISCs go on to target protein coding genes in their 3' UTR, leading to translational repression, deadenylation and destabilization of targets. The exoRNAi and ERI endoRNAi pathways share a pool of secondary AGOs which are loaded with secondary siRNA products of RdRPs. Secondary AGOs go on to silence targets by diverse and not fully characterized mechanisms.

I.VII Nuclear and Inheritable RNAi in *C. elegans*

TGS in *C. elegans* was recently linked to a specialized sub-set of nuclear Argonautes. An elegant genetic screen for factors required to silence nuclear RNA targets in somatic tissues identified the *n*uclear *R*NAi *de*fective (*nrde*) mutants (Guang et al., 2008). This class of genes encodes several conserved proteins, including the nuclear Argonaute NRDE-3, which contributes to both exoRNAi and the ERI endoRNAi pathways. NRDE-3 is a secondary Argonaute and binds secondary (22G) siRNAs of these pathways in the cytoplasm and subsequently translocates to the nucleus, where it associates with

nascent chains of pre-mRNAs and recruits at least one of the other NRDE proteins. Soon after targeting, H3K9 tri-methyl modifications accumulate on histone tails at genomic locations matching the dsRNA trigger (Burton et al., 2011; Gu et al., 2012b). It was also claimed that following induction of exoRNAi, these events lead to transcriptional pause at the step of RNA Pol-II elongation (Guang et al., 2010). Since this pioneering genetic screen, another nuclear Argonaute (HRDE-1/WAGO-9) was found to be involved in TGS in the germline, where it is also required for the inheritance of silencing phenotypes triggered by the exoRNAi and piRNA branches of RNAi (Ashe et al., 2012; Buckley et al., 2012; Burton et al., 2011; Lee et al., 2012; Shirayama et al., 2012). Because primary piRNAs target a wide range of coding and non-coding loci in *C. elegans*, their secondary (22G) siRNAs, which also program HRDE-1, map to a large variety of genomic targets. As both nuclear Argonautes genetically require some of the same co-factors, and induce the same histone modifications, the current state of literature supports the view that NRDE-3 and HRDE-1 initiate paralogous nuclear RNAi mechanisms in the soma and germline, respectively (Figure I.X).

The nuclear NRDE-1, 2, and 4 proteins are required for both NRDE-3 and HRDE-1-directed nuclear RNAi, and act both in the soma and germline (Burkhart et al., 2011; Guang et al., 2010). Whereas NRDE-1 and 4 do not encode obviously recognizable domains, NRDE-2 is a 148 kDa protein, conserved from fungi to mammals, which encodes a domain of unknown function (DUF 1740), a serine/arginine-rich domain, and a half-a-tetratricopeptide (HAT)-like domain, often found in proteins involved in RNA metabolism. The nuclear RNAi machinery interacts with target loci through two types of physical interactions. First, nuclear Argonaute NRDE-3 and its associated small RNAs are genetically required for the recruitment of NRDE-2, NRDE-1 and NRDE-4 to nascent pre-mRNA transcripts. Second, through chromatin immunoprecipitation (ChIP), NRDE-1 was proposed to interact with chromatin. As loss of NRDE-4 impairs NRDE-1 association with chromatin without affecting its association with pre-mRNAs (nor NRDE-2's or -3's), this chromatin interaction is genetically dissociable, and is distinct from nascent pre-mRNA association. Silenced loci and transgenes targeted by nuclear RNAi through NRDE-3 or HRDE-1 enrich histone H3K9 tri-methyl (H3K9me3) marks, a

modification closely correlated with stable chromatin silencing (Ashe et al., 2012; Burton et al., 2011), and the NRDE proteins are required for this accumulation.

Curiously, when TGS is established through exoRNAi or piRNA pathways, the silenced state of the target locus is heritable. If the silencing phenotype is selected for a few generations, its penetrance decreases at first, and then increases to become stable without further selective pressure (Vastenhouw et al., 2006). This prompted two groups to screen for genes required for inheritance of RNAi-induced epimutations. In a candidate screen by Miska and colleagues, knock-down of an ortholog of heterochromatin protein 1 (HPL-2), a methylated H3K9-binding protein, and putative histone methyltransferases (HMTs) SET-25 and SET-32 reversed the inherited phenotype (Ashe et al., 2012). SET-25 was recently characterized as an H3K9me3 methyl-transferase. SET-25 acts sequentially after MET-2, an ortholog of SETDB1, which deposits H3K9me and H3K9me2 modifications. Interestingly, SET-25 remains associated with its product marks on chromatin, independently of its activity (Towbin et al., 2012). Even though it encodes the telltale SET domain of a histone

methyltransferase, the specificity and activity of SET-32 have yet to be characterized. Surprisingly, genes identified in an inheritable RNAi reversion screen performed by Plasterk and colleagues did not overlap with the screen by the Miska group (Vastenhouw et al., 2006). Instead, they identified genes encoding chromodomain protein MRG-1, a class II histone deacetylase HDA-4, the MYST-domain K03D10.3 putative histone acetyltransferase, and chromatin-remodeling ATPase ISW-1. An important difference in their respective designs however, is that the Miska group conducted their reversion screen *early* after triggering the RNAi epimutation, whereas the screen by the Plasterk group was carried out much later, at a time when the epimutation was stabilized. It is therefore possible that they reflect distinct requirements, and perhaps distinct mechanisms involved at different steps of the stabilization of the RNAi-induced epimutation (Figure I.X).

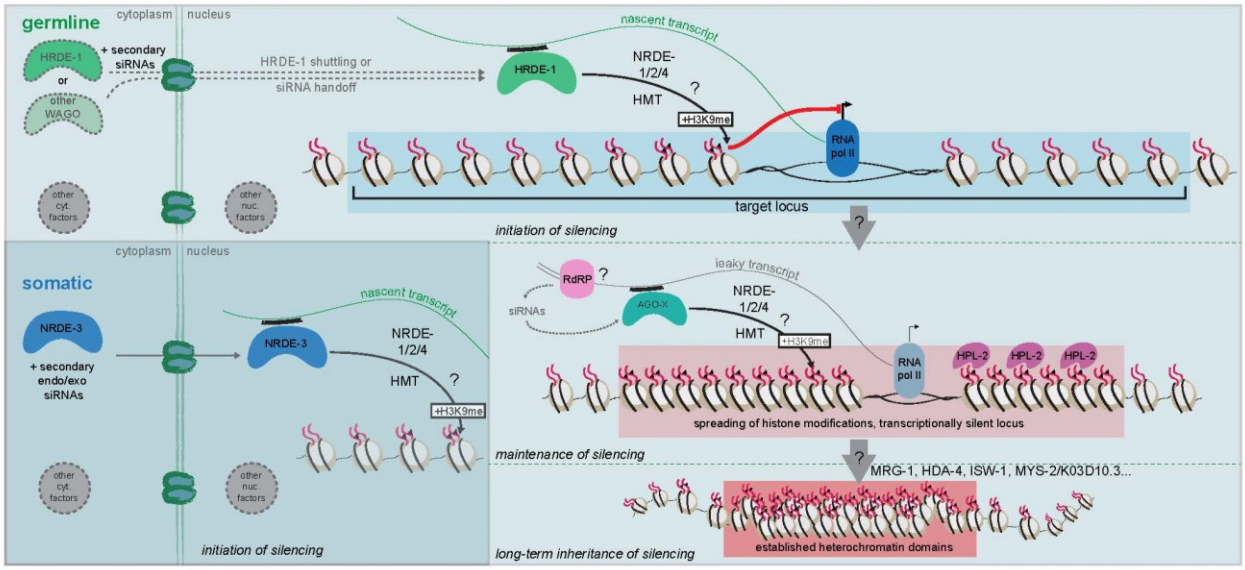


Figure I.X: Model of nuclear and inheritable RNAi. Nuclear Argonaute paralogs NRDE-3 and HRDE-1 direct silencing of small RNA targets in the soma and the germline, respectively. The NRDE-1/2/4 proteins are required for this process, which results in histone methyltransferases (HMTs) modifying histone tails at the target genomic site. Upon transmission to progeny and selection for a few generations, a self-enforcing loop mediated by nuclear Argonaute HRDE-1 or another Argonaute (AGO-X) could enforce and spread H3K9me3 marks on the locus. HPL-2 recognizes this mark, consolidates heterochromatin assembly on the locus, which results in its overall transcriptional shutdown. Distinct genetic requirements for longer-term inheritance of the RNAi-induced epimutation suggest that a subsequent shift towards a different set of histone modifications. This subsequent mechanism may or may not involve RNAi and a distinct nuclear Argonaute.

The studies outlined above have identified a novel branch of RNAi termed “nuclear RNAi”, and several genes involved in this process. However, many mechanistic details remain unclear. What is the exact sequence of events that dictates an Argonaute to translocate into the nucleus to mediate TGS? Why do some Argonautes enter the nucleus but others remain cytoplasmic and mediate PTGS? In the grand scheme, how do the stages of RNAi-induced silencing progress from transient PTGS, to TGS, and finally to an indefinitely silenced state? Further mechanistic studies are needed to answer these unresolved questions.

I.VIII Thesis Rationale and Objectives

The work presented in this thesis revolves around the DCR-1-dependent small RNA pathways of gene silencing, using *C. elegans* as a model system. Using biochemical methods, we investigated the regulation of key protein components in detail, as well as ultimate mechanisms of target silencing in the ERI endoRNAi and exoRNAi pathways.

In Chapters 1 and 2, we sought to understand how DCR-1 could operate in multiple specialized and overlapping RNAi pathways since, as in vertebrates, it is present as a single gene. How DCR-1 is directed to different protein complexes, which are reflective of the different RNAi pathways is also a mystery. We found that endogenous DCR-1 is proteolytically cleaved to produce small DCR-1 (sDCR-1), a stable and active C-terminal fragment (Chapter 1). sDCR-1 acts positively in the exoRNAi pathway and negatively in the miRNA pathway, and its high expression causes these two pathways to compete with one another. Abberant expression of sDCR-1 in early stages of development negatively affected developmental progression and viability. We propose that sDCR-1 could promote anti-viral defence due to its enhancement of the exoRNAi pathway, at stages where it is not detrimental to development.

We also pursued the hypothesis that DCR-1 could be modified by phosphorylation to alter its activity in the different RNAi pathways. We identified a cluster of phosphorylated sites in the PAZ domain and close to the sites of dsRNA terminus binding (Chapter 2). Mutation of these sites led to impaired exoRNAi activity, altered protein interactions with ERI endoRNAi protein partners, and severe developmental defects. The results indicate that phosphorylation of this cluster is important for the function of all the DCR-1-dependent RNAi pathways and the overall health and viability of the organism.

In chapter 3, we investigated the mechanism of silencing exerted by the ERI endoRNAi pathway on its targets. We hypothesized that the ERI pathway transcriptionally inhibits its target genes, similarly to the *S. pombe* RNAi pathway, and this is supported by the fact that secondary ERI endo-siRNAs are loaded onto a nuclear Argonaute. In order to test this, we developed and optimized a robust nuclear run-on assay using *C. elegans* embryos. In *eri* mutants, target gene transcription was de-repressed, confiming our hypothesis. We tested whether the exoRNAi pathway could mediate transcriptional gene silencing as well, and found that knockdown of an endogenous target curiously

leads to an increase in its transcription. Ultimately, this work forms a basis for future precise transcriptome-wide analysis in our model system.

Finally, in chapter 4, we sought to further elucidate the mechanism of nuclear RNAi using extensive proteomics to complement the existing genetic data in the literature. We focused on the identification of protein partners of the nuclear Argonaute NRDE-3, which acts downstream of the ERI endoRNAi and exoRNAi pathways. We validated and characterized the top candidate in our screen, a previously uncharacterized protein which we named NRDE-3-interacting protein 1 (NIP-1). We found that NIP-1 is not only a partner of NRDE-3, but also multiple other Argonautes of the ERI endoRNAi and exoRNAi pathways. Additionally, NIP-1 is present in Argonaute complexes at multiple stages in RNAi: from the DCR-1 and primary Argonaute steps to the secondary Argonaute steps, and we propose that NIP-1 plays a role in the integral process of loading the Argonautes with siRNAs.

Preface to Chapter 1

Dicer (DCR-1 in *C. elegans*) is the highly conserved and essential RNase III enzyme that is required for the biogenesis of small RNAs in multiple RNA interference pathways. Previous work from our group determined that DCR-1 is a member of multiple and distinct protein complexes which correspond to its functions in the different pathways. How this pathway dedication occurs in *C. elegans*, and if DCR-1 protein is differentially modified within these pathways was unknown. We therefore hypothesized that DCR-1 could be modified at the post-translational level to generate this diversity of function. Our studies then led to the fortuitous discovery of a truncated DCR-1 protein, which is conserved in humans, and performs opposing functions in two RNAi pathways.

Chapter 1: A Truncated Dicer Tilts the Balance of RNAi Pathways

Originally published in the journal Cell Reports, August 2013, Volume 4, Issue 3

Sawh A. N. and Duchaine T. F.

Received: September 4, 2012; Received in revised form: June 25, 2013; Accepted: July 12, 2013; Published Online: August 08, 2013

1.1 Introduction

As master regulatory processes, RNA interference (RNAi) phenomena orchestrate complex and evolutionary conserved programs of gene silencing. At the crux of these events is the highly conserved RNase III enzyme Dicer (DCR-1 in *C. elegans*). Dicer is essential for the biogenesis of the small RNA species which direct the sequence specificity of silencing (Bernstein et al., 2001) as part of the miRNA, exogenous (exo)RNAi and ERI endogenous (endo)RNAi pathways. miRNA hairpin precursors are processed via the sequential actions of Drosha and Dicer, along with their cofactors. miRNAs then associate with the Argonaute proteins into an effector complex termed RISC, and pair via complementary sequences to their target mRNAs to instigate silencing post-transcriptionally (Ambros, 2004; Bartel, 2009). exoRNAi is characterized by targeted silencing through the action of exogenous long double stranded RNA (dsRNA), which is processed into siRNAs, and incorporated into the RISC to direct negative transcriptional, and post-transcriptional target regulation (Fire et al., 1998; Hannon, 2002; Mello and Conte, 2004; Song et al., 2004). The ERI endoRNAi pathway

involves a separate class of small interfering RNAs (26G siRNAs) and a specialized DCR-1 multi-protein complex required for their production (Duchaine et al., 2006; Lee et al., 2006; Pavelec et al., 2009). In addition to trigger dsRNA cleavage, Dicer is required for the loading of small RNAs onto the RISC, and is a stable component of the holo-RISC in many species (Kim et al., 2007; Lee et al., 2004b; Maniataki and Mourelatos, 2005; Pham et al., 2004). Although several organisms have evolved distinct Dicer-like genes exclusive to specific pathways, both humans and *C. elegans* maintain one Dicer gene; therefore it stands to reason that its activities are differentially regulated in each RNAi pathway.

DCR-1 is a large, multi-domain enzyme that contains an N-terminal DExD/H box family helicase domain and substrate recognition domain (PAZ), a domain of unknown function (DUF), as well as C-terminal catalytic RNase III (RIII) domains and a dsRNA-binding domain (dsRBD). The dual RNase III domains of DCR-1 dimerize to form the catalytic core of the enzyme, with each domain responsible for the cleavage of one strand of the dsRNA substrate (Macrae et al., 2006; Takeshita et al., 2007; Zhang et al.,

2004). The N-terminal domains are joined to the catalytic core by a flexible linker region of the protein, which is proposed to act as a molecular ruler to determine the size of the small RNAs generated (Macrae et al., 2006; Park et al., 2011).

Here we report that the proteolytic processing of DCR-1 generates a truncated C-terminal form, which directs intricate mechanisms of regulation on the RNAi phenomena.

1.2 Results

1.2.1 Small DCR-1 (sDCR-1): an abundant, developmentally regulated C-terminal fragment of DCR-1.

In the course of analysing full-length (FL) endogenous DCR-1 expression, we noticed that a polyclonal antibody directed against the C-terminal linker (Figure 1.1A) also detected an abundant, and previously unidentified band migrating at ~95kDa. Both bands were lost in DCR-1-depleted animals (DCR-1 m+/z-) (Figure 1.1B). A strain expressing an 8HA-tagged FL DCR-1 also expressed the shorter polypeptide with a shift in size corresponding to the epitope tag (Figure 1.1C).

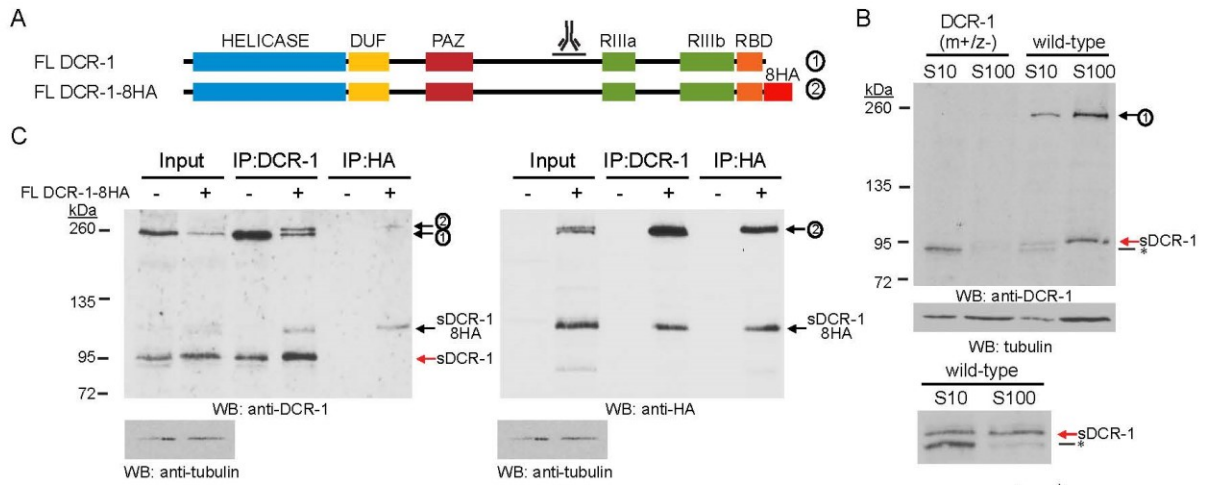


Figure 1.1: Endogenous DCR-1 is expressed as full-length and small DCR-1. (A) Schematic representation of full-length DCR-1, with the anti-DCR-1 polyclonal antibody epitope indicated. (B) Western blot of WT extracts compared to DCR-1-depleted extracts. Non-specific proteins detected by the polyclonal antibody are indicated with *. (C) Western blots of total lysate and IPs of WT endogenous DCR-1 alone (lanes with (-)) or with additional C-terminally tagged DCR-1-8HA (lanes with (+)).

Furthermore, both FL DCR-1 and the shorter species remained stable in extracts incubated on ice or at room temperature, and even in the absence of protease inhibitors, arguing against its production during extract preparation (Figure 1.2A, and A1.1). Thus, *C. elegans* DCR-1 is also expressed as a truncated, stable, C-terminal fragment. We named this polypeptide small- or sDCR-1. sDCR-1 expression markedly

increased in later larval stages (Figure 1.2B), and reached approximately a 1:1 ratio with FL DCR-1 in the adult (Figure 1.2C). Finally, comparison with adults lacking germline tissue or sperm cells indicated an enrichment of sDCR-1 over FL DCR-1 in the somatic tissues (Figure 1.2D).

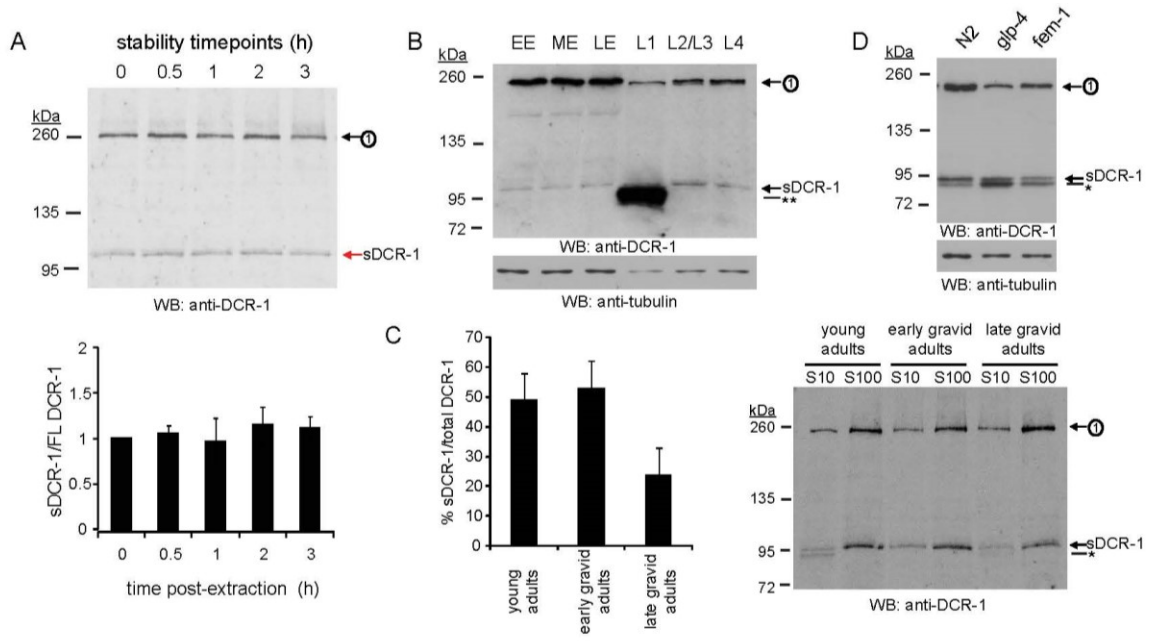


Figure 1.2: sDCR-1 is stable and developmentally regulated. (A) Western blot of DCR-1 and sDCR-1 over 3 hours post-extract preparation. The stability of sDCR-1 (normalized to FL DCR-1 with the initial levels set to 1) is quantified (lower panel), n = 3 biological replicates. (B) Western blot of endogenous DCR-1 in stage-synchronized WT animals in early embryos (EE), mid-stage embryos (ME), late embryos (LE), and the larval stages L1-L4. The non-specific band in L1 indicated by **, was confirmed to be not sDCR-1 (Figure A1.1). FL DCR-1 consistently migrates as a doublet, indicating possible post-

translational modification. (C) Quantification of sDCR-1, done by western blot using the intensity of sDCR-1 relative to total DCR-1 (sDCR-1 + FL DCR-1) in S100 extracts, n = 3 biological replicates. Data represented as mean +/- standard deviation. A representative blot is shown (right), and no significant difference in the ratio was found between S10 and S100 extracts. (D) Western blot of adult extracts from WT animals (late gravid) compared to two temperature-sensitive germline-deficient alleles of glp-4 and fem-1 grown at the non-permissive temperature. See also Figure A1.1.

Based on the migration of sDCR-1 on SDS PAGE, and antibody mapping (Figure 1.1C), we estimated that sDCR-1 is composed of an intact C-terminus containing both RNase III domains and the dsRBD, and may lack the N-terminal Helicase, DUF and PAZ domains. To map what is encoded within the sDCR-1 polypeptide, we constructed a transgene encompassing residues 957 to 1910, which includes the linker region, the RNase IIIa and b, and the dsRBD. The transgenic protein was tagged with N-terminal MYC and C-terminal 3FLAG epitope tags (Figure 1.3A). Since sDCR-1 is enriched in somatic tissues (Figure 1.2D), we elected to drive the transgene from the muscle-specific *myo-3* promoter (Fire and Waterston, 1989), and independent lines were

generated. In extracts of the strain expressing MYC-pre-sDCR-1-3FLAG, western blotting with an anti-MYC antibody revealed that the transgene-derived protein migrated at ~135kDa (Figure 1.3B). IPs with anti-MYC and western with the DCR-1 antibody detected only one polypeptide, corresponding to the entire transgene-driven fusion (Figure 1.3B). However, IPs with anti-FLAG detected two proteins originating from the MYC-pre-sDCR-1-3FLAG transgene (Figure 1.3C). The smaller polypeptide detected is consistent with the size of sDCR-1 with the intact 3FLAG C-terminal tag (Figure 1.3C). Similar results were obtained from let-858 promoter-driven transgenes (Below, see also Figure A1.3B). Hence, both the endogenous DCR-1 and the transgenic MYC-pre-sDCR-1-3FLAG produce a C-terminal truncated protein *in vivo*.

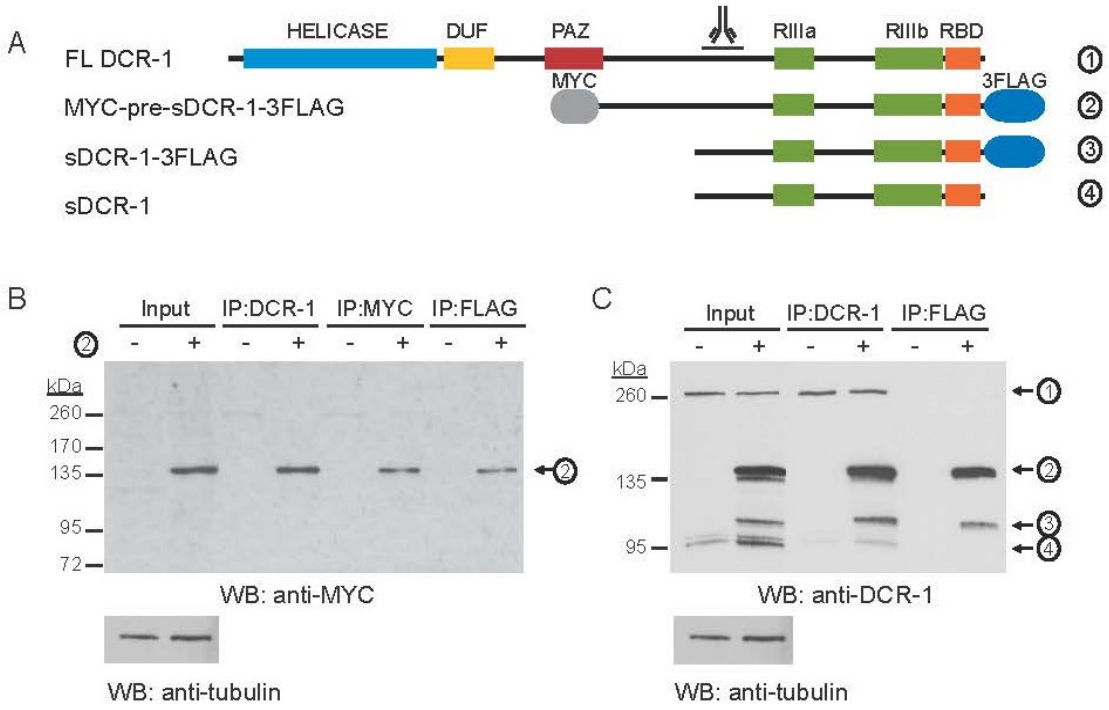


Figure 1.3: Endogenous and transgenic *dcr-1* genes produce sDCR-1. (A) Schematic representation of endogenous (① and ④) and transgenic DCR-1 proteins (② and ③). (B and C) Western blots of total lysates and IPs from WT animals or a transgenic line carrying the myo-3-driven MYC-pre-sDCR-1-3FLAG.

1.2.2 sDCR-1 is generated through proteolytic cleavage.

Expression of FL DCR-1 from either genomic DNA or cDNA in transgenic animals produced indistinguishable proteins for both FL DCR-1 and sDCR-1 (Figure 1.4), ruling out expression of sDCR-1 by alternative splicing and suggesting a proteolytic biogenesis. In order to map the N-terminus of sDCR-1, we took advantage of the fact

that MYC-pre-sDCR-1-3FLAG produces sDCR-1-3FLAG *in vivo* (Figure 1.3C). From large-scale FLAG purifications, the band corresponding to sDCR-1-3FLAG was submitted to mass spectrometry. We found a drop-off in the peptide coverage near position 1200 in the protein sequence of DCR-1 (Figure A1.2), and subsequently designed deletions on the MYC-pre-sDCR-1-3FLAG transgene in order to abolish a putative cleavage event. Deletions were designed to encode an internal HA tag, to orient the fragments of DCR-1 produced relative to putative cleavage site(s) (See Materials and Methods). Deletion of amino acids 1163 to 1184 (deletion 2) most severely altered the generation of sDCR-1 and abolished the major band detected by western blot (Figure 1.5), indicating that these residues are integral to proteolytic cleavage. Deletion of amino acids 1200 to 1221 (deletion 3) diminished, but did not abolish, sDCR-1, and is located C-terminal to cleavage (Figure 1.5, HA blot). Deletion of the intervening sequence between deletion 2 and 3 had no effect on sDCR-1 production (not shown). These results indicate that the determinants of cleavage may span an extended or complex sequence.

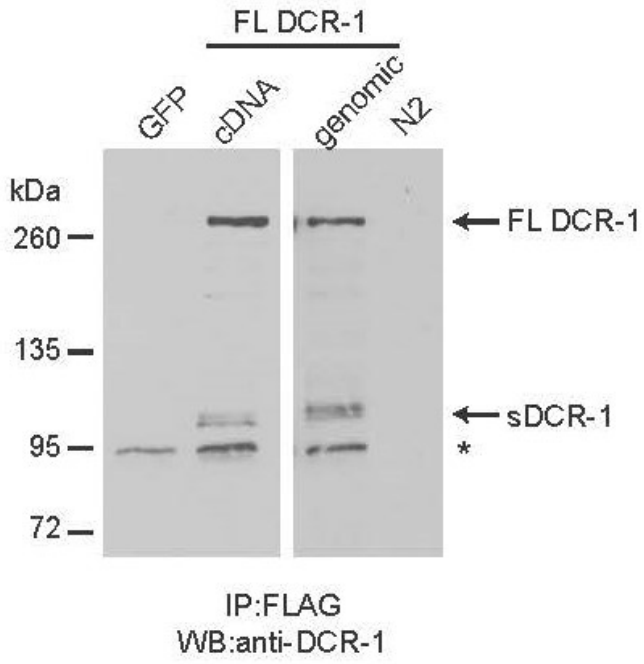


Figure 1.4: sDCR-1 is generated *in vivo* via proteolytic cleavage. Western blot of animals expressing full-length DCR-1 constructs from either genomic DNA (C-terminal 3FLAG) or cDNA (C-terminal 2FLAG). Samples were run on the same gel, and different exposures are shown to normalize for expression level. Negative controls are WT (N2) and sur-5::GFP (co-injection marker) alone. Non-specific bands marked with *. See also Figure S2.

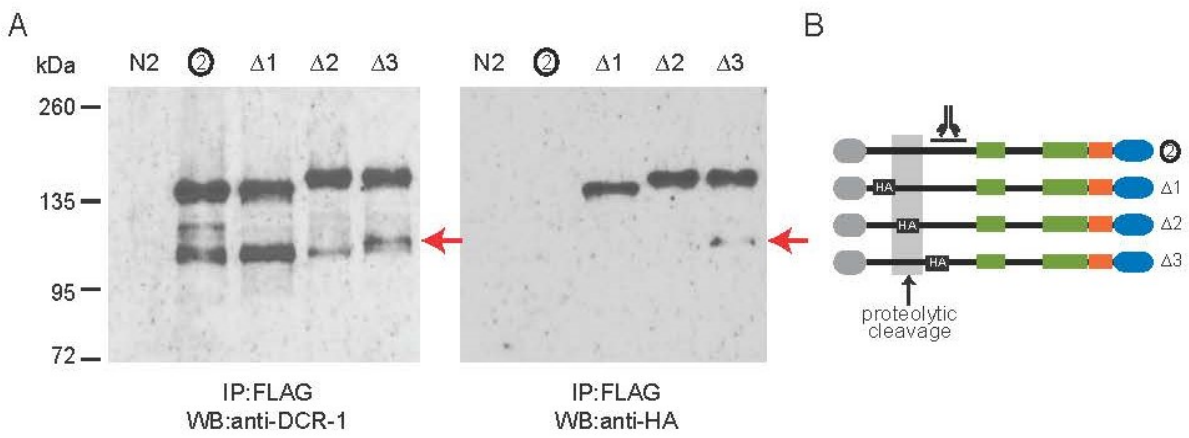


Figure 1.5: sDCR-1 is generated by cleavage between the PAZ and RNase IIIa domains. (A) Deletions within the pre-sDCR-1 transgene were generated with and HA tag insertion at the indicated regions. Western blots on FLAG IPs of mixed stage animals using the anti-DCR-1 or anti-HA antibodies to monitor the generation of sDCR-1 and the orientation of potential proteolytic cleavage. The arrow indicates the major band of the sDCR-1 doublet, indicating modification or major and minor cleavage sites. (B) Schematic representation of proteolytic cleavage site in DCR-1.

1.2.3 sDCR-1 enhances exoRNAi and impinges on miRNA biogenesis.

We next tested the effect of sDCR-1 on the exoRNAi pathway by performing *unc-22* RNAi on WT and animals bearing enforced sDCR-1 expression (sDCR-1+). The *unc-22* phenotype, ranging from varying degrees of body wall muscle twitching to complete paralysis, can be used as a quantitative readout of the potency of exoRNAi (Fire et al.,

1998; Yigit et al., 2006). Three independent lines of muscle-restricted MYC-pre-sDCR-1-3FLAG (pmyo-3::sDCR-1+) displayed a 5- to 14-fold enhancement of paralysis in progeny (F1) (Figure 1.6 left), which demonstrates that sDCR-1 enhances the exoRNAi pathway. Furthermore, a ubiquitous let-858-driven sDCR-1+ (two lines displaying mild 4x and 8x expression of sDCR-1 over the endogenous FL DCR-1, see Figure A1.3B) resulted in a 18 to 31-fold enhancement of the exoRNAi response over WT levels (Figure 1.6 right). Importantly, when the conserved catalytic residues in the RNase III domains were inactivated in the MYC-pre-sDCR-1-1FLAG transgene, the resulting protein did not enhance exoRNAi (Figure 1.6 left, myo-3::catsDCR-1+) despite comparable expression levels (Figure A1.3B). Thus, the catalytic activity of sDCR-1 is required for the enhancement of the exoRNAi response.

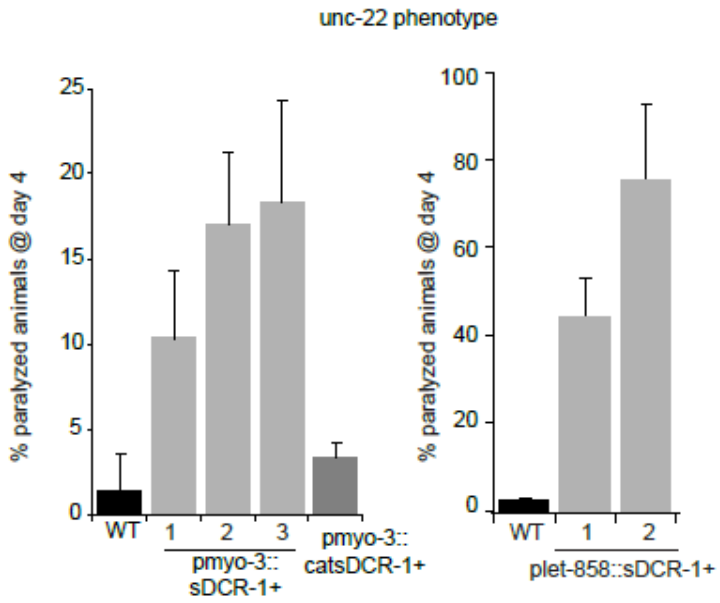


Figure 1.6: sDCR-1 expression enhances exoRNAi in an activity dependent manner. Complete paralysis of the animals exposed to unc-22 RNAi was quantified (n ≥ 3 biological replicates of P0 animals) in WT adults compared to strains carrying an additional myo-3, let-858-driven pre-sDCR-1, or a catalytically inactive version of pre-sDCR-1 (catsDCR-1). See also Fig. A1.3A,B.

We also tested the effect of an early enforcement of sDCR-1 expression on a battery of conserved and developmentally important miRNAs (Figure 1.7). In embryos, sDCR-1 caused a substantial (4 to 8-fold) accumulation of the precursor (pre-)miR-58 (bantam) and pre-miR-35, as well as a decrease in the level of mature miR-35 (22 to 41%), which is an abundant miRNA key to embryonic development (Heo et al., 2008). In L2-L3 stage

animals, sDCR-1 caused an accumulation of pre-lin-4 (2.4 to 3.2 fold), a key regulator of developmental timing and fate of neuronal and hypodermal cell lineages (Heo et al., 2012; Ma et al., 2012b; Piskounova et al., 2011; Viswanathan et al., 2008). In contrast, sDCR-1+ caused a reduction in the levels of both pre-let-7 (24 to 59%) and mature let-7 (36 to 52%), a miRNA required for the transition between L4 and adult cell fates (Alvarez-Saavedra and Horvitz, 2010; Napoli et al., 1990), and a member of a larger miRNA family that also controls early developmental decisions (Montgomery et al., 1998; Volpe et al., 2002). Thus, heightened levels of sDCR-1 earlier in development result in broad defects in pre- and mature miRNA expression with varying severity, the most robust of which being the accumulation of pre-miRNAs.

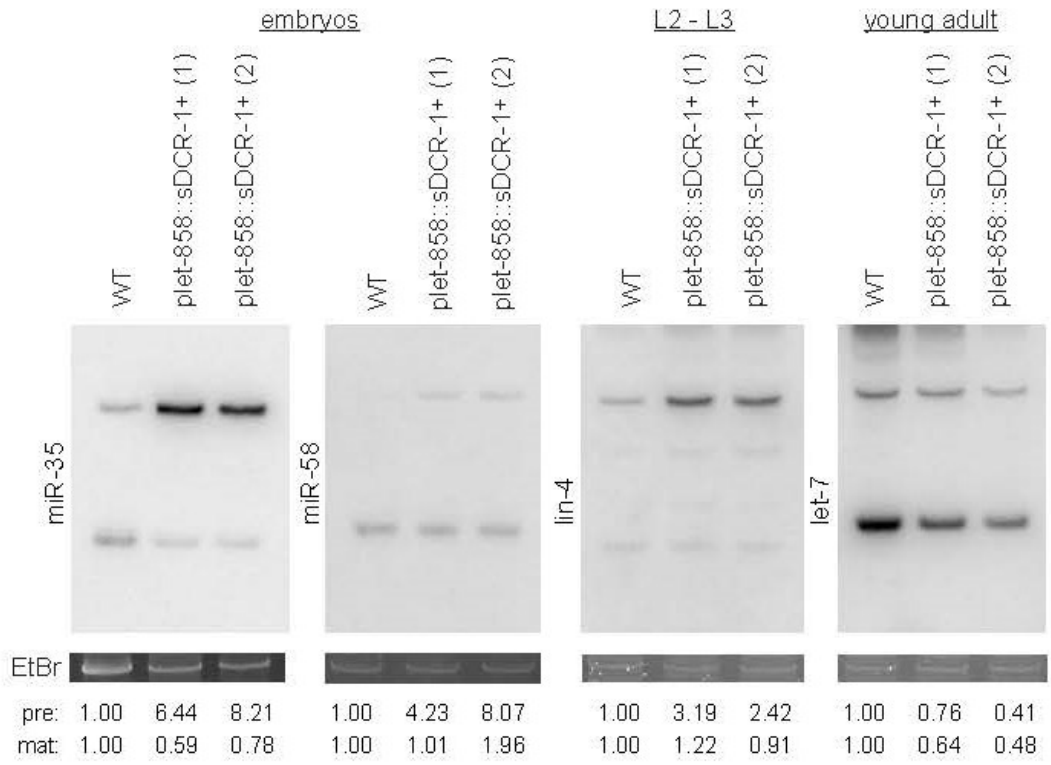


Figure 1.7: sDCR-1 represses miRNA biogenesis. Northern blots of miR-35, miR-58, lin-4 and let-7 on RNA from animals at the indicated developmental stages.

To examine the interplay between the miRNA and exoRNAi functions of sDCR-1, we took advantage of the pmyo-3::sDCR-1 transgene, the expression of which overlaps with the muscle specific *unc-22* gene and mir-1. miR-1 is a key miRNA for muscle differentiation and integrity (Guo and Kemphues, 1995; Kennerdell and Carthew, 1998; Misquitta and Paterson, 1999; Romano and Macino, 1992; van der Krol et al., 1990).

Pre-miR-1, while low in abundance in WT animals, accumulated in the sDCR-1+ strain without a detectable change in mature miRNA (Figure 1.8A), similar to the changes seen in lin-4. The ratio of miR-1 to pre-miR-1 in sDCR-1+ animals was reduced approximately 10-fold (Figure 1.8B). In addition, pre-miR-1 accumulation was still visible in animals expressing catalytically inactive sDCR-1, although to a lesser level (Figure 1.8A). Thus, in contrast with its role in enhancement of exoRNAi, the catalytic activity of sDCR-1 is not required for the accumulation of pre-miRNAs. Surprisingly, inhibition of pre-miR-1 processing was exacerbated when triggering exoRNAi. In pmyo-3::sDCR-1+ or catsDCR-1+ animals, triggering *unc-22* RNAi significantly reduced the ratio of mature to pre-miR-1 (Figure 1.8B, right). These results indicate that high levels of sDCR-1 allow exoRNAi to compete with pre-miRNA processing, and suggest that expression of sDCR-1 alters the mechanistic boundaries between the exoRNAi and miRNA pathways.

We next tested whether the molecular defects induced by the early expression of sDCR-1 could alter the physiology of the animal. To compare WT to plet-858::sDCR-1+ animals, synchronized P0 L4 animals were isolated and their progeny (F1) were scored

for developmental progression and brood size. Strikingly, sDCR-1+ animals displayed delayed developmental progression (Figure 1.9A), and animals reaching maturity had significantly reduced brood size (Figure 1.9B).

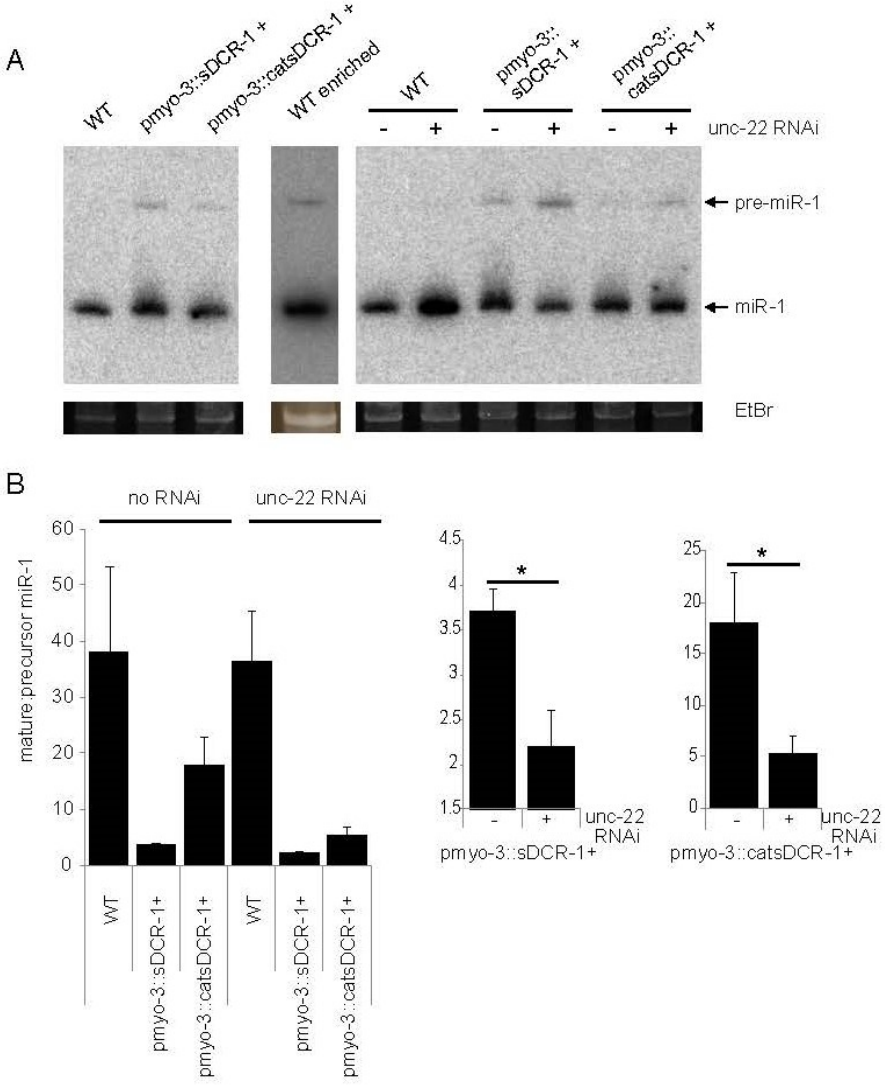


Figure 1.8: sDCR-1 represses miRNA processing independent of catalytic activity and causes the exo- and miRNA pathways to compete with each other. (A) Northern blot of

the muscle-specific miR-1 in RNA of WT adults compared to pmyo-3::sDCR-1+ , or catsDCR-1+ lines, with or without unc-22 RNAi. RNA from WT animals was enriched for small RNAs to reach detectable levels of pre-miR-1 (WT enriched lane). (B) Quantification of the ratio of mature miR-1 to precursor miR-1, with or without unc-22 RNAi triggered. Loading is normalized internally as mature miRNA/precursor miRNA for each lane. n ≥ 3 biological replicates.

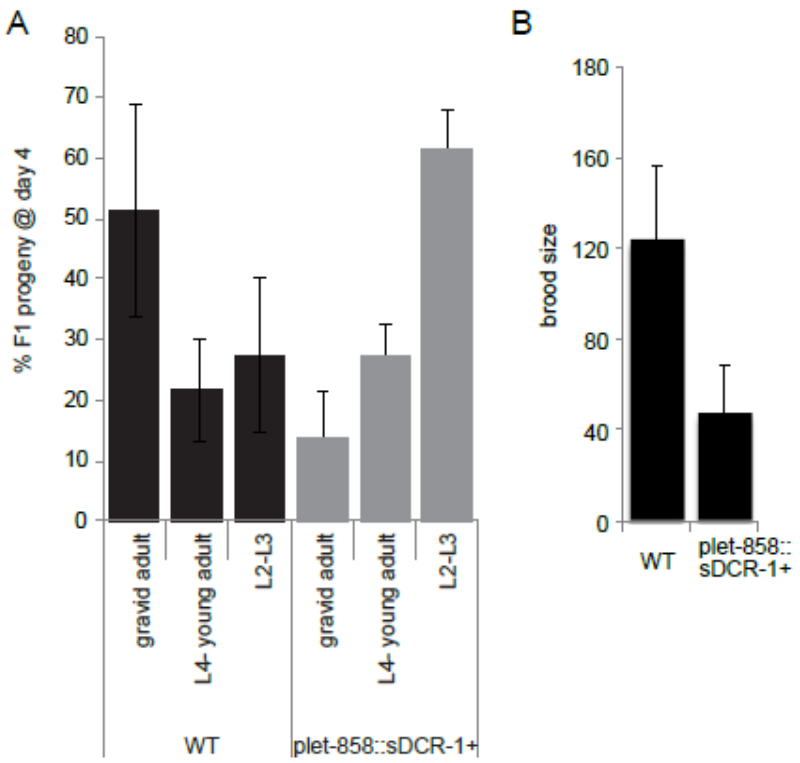


Figure 1.9: Early sDCR-1 expression causes developmental defects. (A) Developmental staging of WT or plet-858::sDCR-1+ F1 animals after P0 L4 animals are isolated and propagated at room temperature for 4 days. (B) Brood sizes of WT and plet-858::sDCR-

1+ animals. Data represented as mean +/- standard deviation. n ≥ 3 biological replicates.

To further test the physiological implications of early sDCR-1+ expression, and since we observed a marked reduction in mature let-7 expression (Figure 1.7), we asked whether sDCR-1+ could genetically interact with the temperature-sensitive allele *let-7(n2853)*. This mutant displays temperature-sensitive adult lethality due to bursting vulva, a phenotype visible at ~20% penetrance at permissive 16^{o}C (Alvarez-Saavedra and Horvitz, 2010; Napoli et al., 1990). Transgenic *let-7(n2853)*; plet-858::sDCR-1+ F1 animals displayed several phenotypes including 47% embryonic or L1 arrest, 14% lethality before L4, and 2% sterility. Of those animals that survived to adulthood, 53% displayed bursting vulvas (1.10), a *let-7* signature phenotype which was exacerbated compared to a control GFP transgene alone (Figure A1.3C). plet-858::sDCR-1+ transgenesis in *let-7(n2853)* failed to produce stable transmitting lines, whereas GFP injection produced several, suggesting that sDCR-1 expression may also have a negative effect on the fertility or viability of the F1 mosaic-expressing animals. This data

shows that early enforcement of sDCR-1 expression exacerbates *let-7* phenotypes, and leads to severe pleiotropic developmental defects.

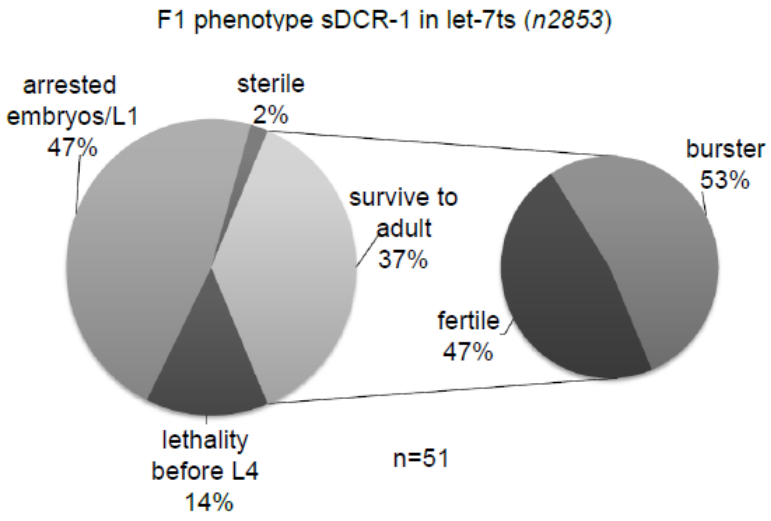


Figure 1.10: sDCR-1 expression exacerbates *let-7* phenotypes. Phenotype of F1 transgenic let-7ts (*n2853*) animals after microinjection of plet-858::sDCR-1. See also Figure A1.3.

1.2.4 sDCR-1 selectively interacts with the miRNA Argonautes ALG-1 and ALG-2.

Given the abundance of sDCR-1 relative to FL DCR-1, and since this function does not require catalytic activity, we reasoned that it may act as a competitor by sequestering interacting RNAi co-factors. To test this idea, IPs of endogenous DCR-1, transgenic

MYC-pre-sDCR-1-3FLAG, and sDCR-1-3FLAG were probed for components of three known DCR-1-dependent RNAi pathways (Figure 1.11). While they co-IP with FL DCR-1, the ERIC components ERI-1b and ERI-5 did not interact with sDCR-1-3FLAG. Such results concur with the mapping of the ERI proteins to the N-terminus of FL DCR-1 (Thivierge et al., 2012). Furthermore, sDCR-1 did not interact with the dsRNA-binding protein RDE-4, an essential cofactor of DCR-1 in exoRNAi (Tabara et al., 2002). This finding demonstrates that RDE-4 interacts with the N-terminus of FL DCR-1, and indicates that sDCR-1 can initiate and/or enhance exoRNAi without a stable physical interaction with RDE-4. In contrast, sDCR-1 did co-IP the miRNA-specific Argonautes ALG-1 and ALG-2 (Grishok et al., 2001) (Figure 1.11A). In a reciprocal experiment, ALG-2 co-immunoprecipitated endogenous FL DCR-1, as well as sDCR-1 (Figure 1.11B). To determine whether sDCR-1 contacts miRNA species *in vivo*, we probed the IPs by northern blot. Neither mature, nor pre-miRNAs could be detected in sDCR-1 IP (Figure 1.12A). In contrast, IP of ALG-2 under the same conditions enriched mature and pre-let-7 (Figure 1.12B). Finally, affinity pull-down of mature miRISC failed to capture FL

or sDCR-1, while effectively capturing ALG-1 and ALG-2 (Figure A1.4). Thus, sDCR-1 does not stably associate with mature or pre-miRNAs. These results support a model wherein sDCR-1 sequesters the Argonautes from the functional FL DCR-1 complex or the mature miRISC, thus acting as a competitive inhibitor in miRNA processing (Figure 1.11C, and Figure 1.13, Model).

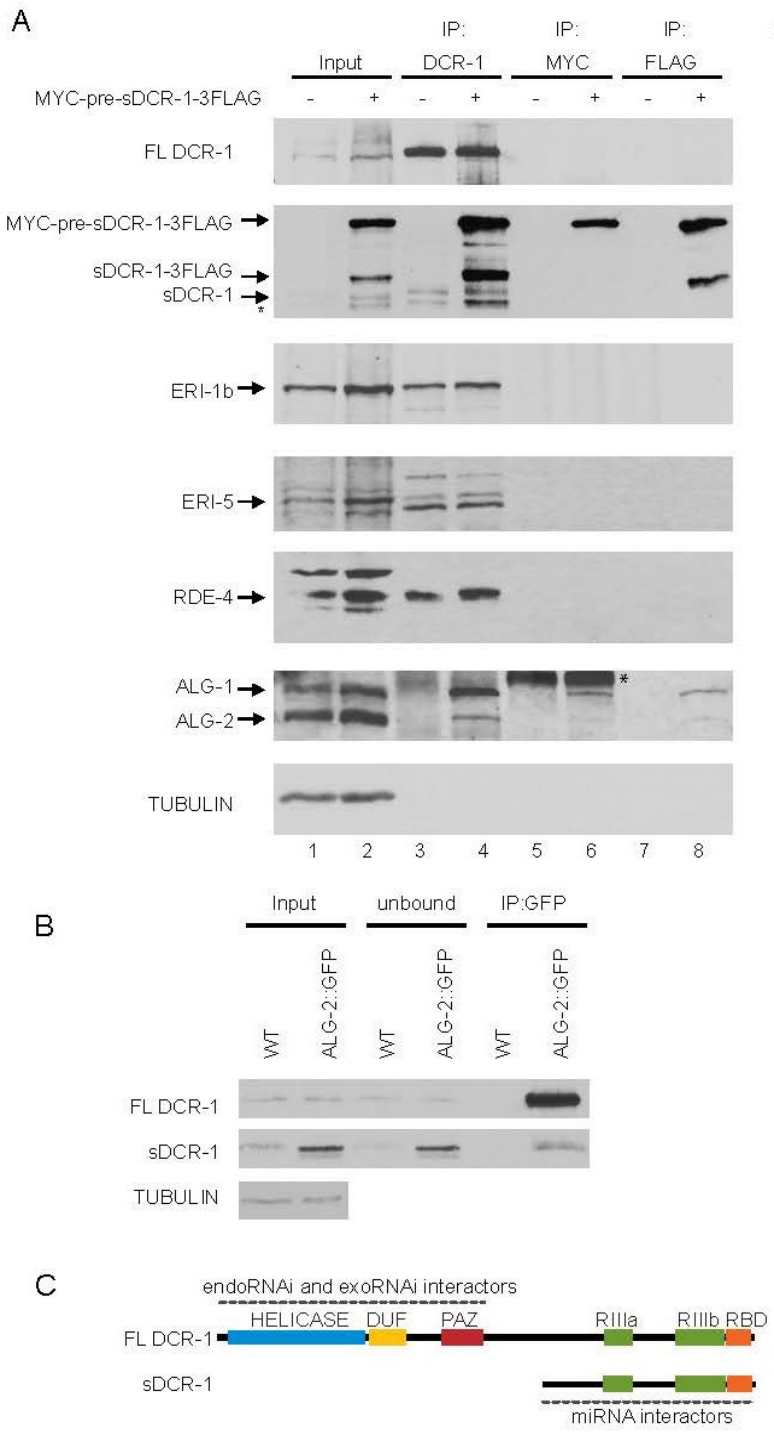


Figure 1.11: sDCR-1 specifically interacts with the miRNA Argonautes ALG-1 and ALG-2. (A) IPs of total DCR-1 proteins (DCR-1) compared to sDCR-1 specifically (MYC and

FLAG) in WT and pmyo-3::sDCR-1+ animals, probed for known FL DCR-1 cofactors. (B) IPs of ALG-2 in WT (negative control) and a strain carrying a rescuing ALG-2::GFP transgene, probed for DCR-1. (C) Schematic representation of FL DCR-1 and sDCR-1 interacting proteins. See also Figure S4.

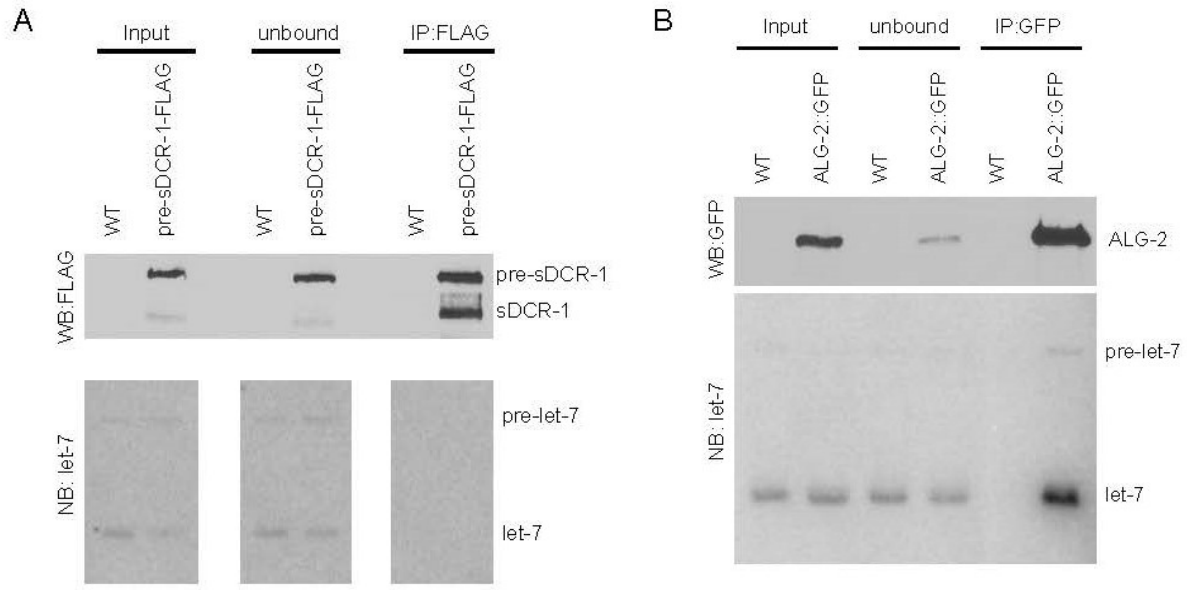


Figure 1.12: sDCR-1 does not stably interact with small RNA species. RNA extracted from sDCR-1 or ALG-2 IPs and probed by northern blot for let-7. (E) Schematic representation of FL DCR-1 and sDCR-1 interacting proteins. See also Figure A1.4.

1.3 Discussion

sDCR-1 reaches an exceptionally high level relative to FL DCR-1; nearing 1:1 in the somatic tissue of adults. This prompts the possibility of competition between sDCR-1 and FL DCR-1 for interacting co-factors and substrates. Our data indicate that sDCR-1 indeed interacts with the ALG-1 and ALG-2 Argonautes, but not with pre-miRNAs or mature miRNAs. In several systems including *C. elegans,* Argonautes are required not only for miRISC function, but also for pre-miRNA processing (Grishok et al., 2001). Furthermore, Argonaute abundance is often limiting for RNAi mechanisms (Diederichs et al., 2008; Lund et al., 2011; Yigit et al., 2006). Consistently, enforced early expression of sDCR-1 mimics the pre-miRNA accumulation observed in ALG-1/2 or DCR-1 depletion (Grishok et al., 2001). Together with the fact that catalytically inactive sDCR-1 causes similar defects, our observations are consistent with a model wherein sDCR-1 functions by sequestering the Argonaute proteins, preventing pre-miRNA recognition and loading of mature miRNAs. Curiously, the severity of the defect observed in sDCR-1+ transgenics is miRNA-specific. This may be due to differences in

pre-miRNA structure, affinity for the complexes, onset of expression, stability, and the incidence of alternate routes of control by degradation.

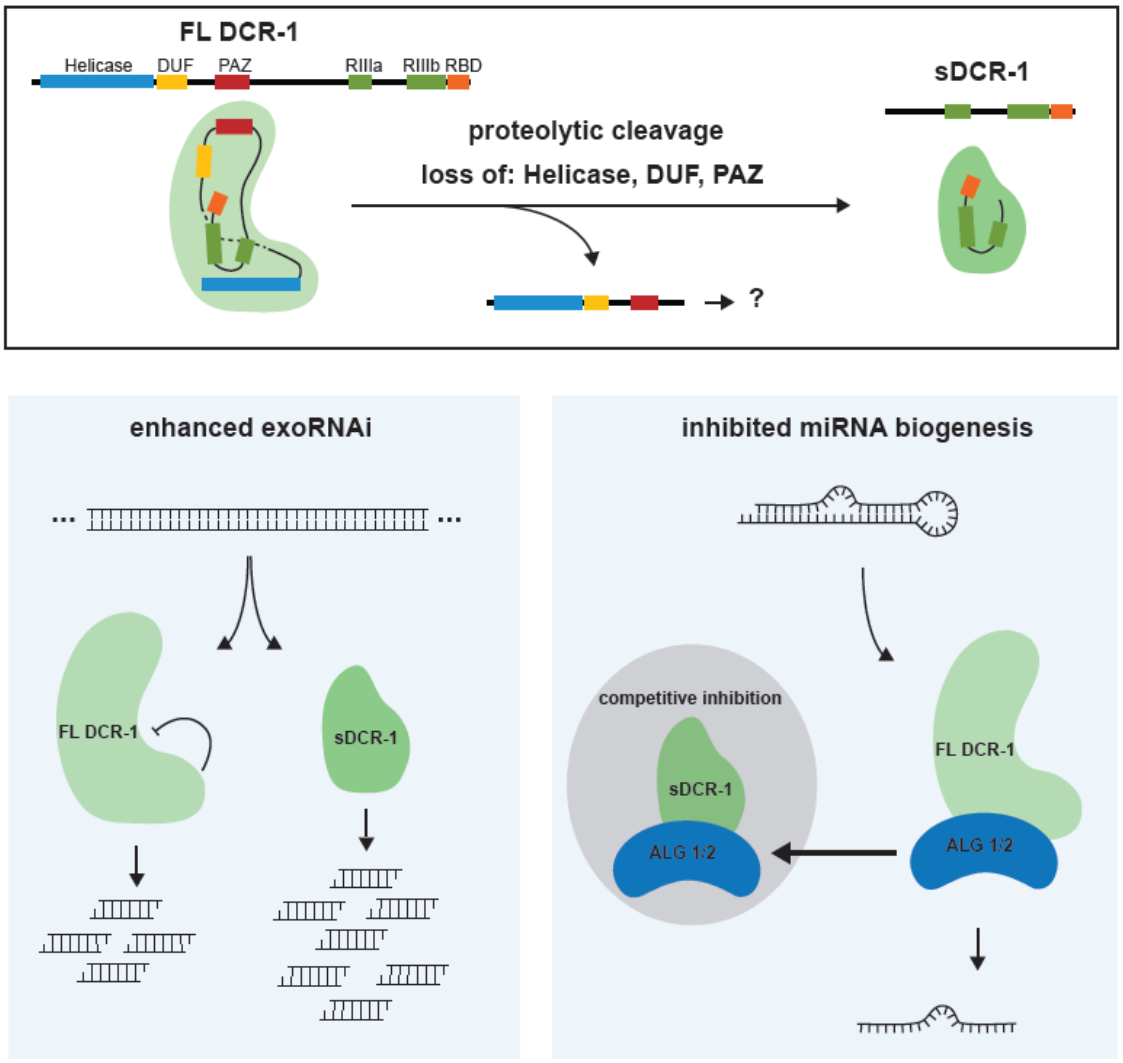


Figure 1.13: Model of sDCR-1 function in the RNAi pathways. Full-length DCR-1 is proteolytically cleaved to produce sDCR-1, a C-terminal fragment containing the catalytic domains and lacking many regulatory domains: helicase, DUF and PAZ.

sDCR-1 enhances exoRNAi, possibly through inherent higher catalytic activity due to the loss of the autoinhibitory helicase domain, and inhibits the miRNA maturation pathway by sequestering miRNA-dedicated Argonautes, causing an accumulation of precursor miRNA and reduction of mature miRNA levels.

DCR-1 cleavage occurs in a poorly conserved region between the PAZ and RNase III domains, and therefore excludes the N-terminal DExD/H RNA helicase, DUF283, and PAZ domains. Interaction analyses further indicate severance from RDE-4, and ERIC protein interactions. Possible mechanistic implications for sDCR-1 functions can be suggested based on the well-established properties of the excluded domains. Enhancement of exoRNAi could, at least in part, be explained by the loss of the helicase domain, which inhibits the catalytic efficiency of human Dicer (Ma et al., 2008). The helicase also directs a processive mode of dsRNA cleavage in FL DCR-1 (Cenik et al., 2011; Welker et al., 2011), thus it could be predicted that sDCR-1 functions as a single-turnover enzyme. Furthermore, since sDCR-1 lacks a PAZ domain, it lacks the 'molecular ruler' responsible to precisely determine the size of the small RNA products

(Macrae et al., 2006; Park et al., 2011). In line with this, a recent report shows that a recombinant human Dicer C-terminal fragment encoding the same domains as sDCR-1 generates a slightly broader size range of products that are still consistent with siRNA lengths (~20-23 nt) (Ma et al., 2012a).

Using mildly enforced and early sDCR-1 expression, we were able to assess the physiological significance of the *timing and extent* of sDCR-1 production. We reason that the regulated expression of sDCR-1 and its system-oriented role in permitting competition between pathways reflect different physiological requirements for RNAi pathways at distinct developmental stages. A low level of sDCR-1 in the embryo is necessary to permit important functions of miRNAs in early developmental decisions and to avoid severe pleiotropic phenotypes. Nearing completion of development, higher levels of sDCR-1 can provide heightened exoRNAi defence against RNA-based pathogens. High levels of sDCR-1 may then serve to permeate the functional insulation between miRNA and exoRNAi pathways, thus tilting the balance of the RNAi mechanisms. Determining the full biological implications of sDCR-1 regulatory functions

will require complete prevention of its production. Thus far, FL DCR-1 constructs bearing the large deletions required to impair sDCR-1 processing (Figure 1.5) failed to rescue *dcr-1* animals (not shown). We reason that the linker region is important for generating miRNA products with fidelity, and it is likely that such broad deletions result in fatal defects in FL DCR-1 functions, and obscure the phenotypes resulting from the loss of sDCR-1.

Interestingly, recent reports on alternative splicing of human DICER revealed isoform 'e', which encodes a protein sharing the same domains as sDCR-1, and is expressed in differentiated epithelial cells as well as a number of breast cancer cell lines (Grelier et al., 2009; Hinkal et al., 2011). When we ectopically expressed isoform "e" in HEK293T cells, we observed a mild, but significant reduction in the ratio of mature to precursor for let-7 (Figure A1.3), but not miR-19b, indicating that similar to sDCR-1, this function is miRNA-specific. Possible conservation of short Dicer forms bears important implications, as Dicer acts as a tumor suppressor in specific cancers (Grelier et al., 2009; Karube et al., 2005; Valastyan and Weinberg, 2010). Ultimately, the impact of

aberrant DICER expression in cancer should depend on the specific gene lesions, the balance of the isoforms represented, and the precise portfolio of the miRNAs expressed.

1.4 Materials and Methods

1.4.1 *C. elegans* Strains and RNAi Assays

All strains were cultured as in (Brenner 1974). N2 was used as the wild-type strain. Alleles used were *glp-4(bn2)*, and *fem-1(hc17)*. RNAi was performed as in (Fire et al., 1998; Timmons et al., 2001) on L4 stage P0 animals and harvested/scored at gravid adult F1 animals. DCR-1 (m+/z-) samples were obtained as in (Duchaine et al., 2006). The let-7 temperature-sensitive allele (*n2853*) and the ALG-2::GFP transgenic rescue strain (MJS26)(Jones et al., 2001) were kind gifts from Dr. Martin Simard.

1.4.2 Sample Preparation

Pellets were homogenized in 50mM Tris-HCl pH8/150mM NaCl/1mM EDTA with Complete EDTA-free protease inhibitors (Roche), and cleared by 10 000xg (S10) or 100

000xg (S100) centrifugation. For IP, extracts were supplemented to 1% Triton X100 prior to antibody incubation. For IP-northern blot experiments, the lysis buffer used was 30mM HEPES-KOH pH7.4/150mM KOAc/5mM $Mg(OAc)_2$/0.1% Igepal, with RiboLock RNase inhibitor (Thermo Scientific), and RNA was isolated from S10 extracts, unbound samples and IPs with phenol/chloroform extraction.

1.4.3 Northern Blotting

Northern blotting was performed as in (Thivierge et al., 2012). Data was quantified using Image J. Statistical significance was calculated using independent 2-tailed Student's *t*-test.

1.4.4 Transgenics

For myo-3-driven constructs, DCR-1 amplicons were cloned into L2534 (Addgene plasmid 1608). For let-858-driven constructs, DCR-1 amplicons with 3'UTR were cloned into L2865 (Addgene plasmid 1522). Transgenic animals were obtained by microinjection of wild-type animals with DCR-1 constructs at 5-30ng/μl, mixed with

pTG96 (sur-5::GFP) at 70ng/μl as a selectable marker. Injection of *n2865* worms was done with pTG96 at 70ng/μl ± the plet-858::sDCR-1 construct at 10ng/μl.

Transgene	Strain Name
dcr-1::DCR-1-8HA	N2HD
pmyo-3::MYC-pre-sDCR-1-3FLAG	qeIs2(M9), qeIs3(M16)
pmyo-3::MYC-pre-sDCR-1-intFLAG	J
pmyo-3::MYC-cat4pre-sDCR-1FLAG	qeIs5(R25)
pmyo-3::MYC-DCR-1-3FLAG (genomic)	FL gen1, FL gen2
pmyo-3::MYC-DCR-1-2FLAG (cDNA)	C5A, C5B, C5D, C30A, C30B
pmyo-3::MYC-pre-sDCR-1-3FLAG Δ1	Del1B
pmyo-3::MYC-pre-sDCR-1-3FLAG Δ2	Del2B
pmyo-3::MYC-pre-sDCR-1-3FLAG Δ3	Del3B
plet-858:: MYC-pre-sDCR-1-3FLAG	qeEx1, qeIs9 (LDU4), qeIs11 (LDU6)

Table 1.1: Transgenic strains generated in Chapter 1.

1.4.5 IP and Western Blotting

IPs and western blots were performed on extracts prepared from young-gravid adults, unless otherwise indicated. Antibodies used were: rabbit polyclonals against DCR-1, RDE-4, ALG1/2, and ERI-1 (Duchaine et al., 2006); mouse polyclonal against ERI-5 (Thivierge et al., 2012), mouse monoclonals against alpha tubulin (Abcam), c-MYC (Abcam), HA (Bioshop), GFP (Roche) and FLAG (Sigma). HRP-conjugated rabbit and mouse TrueBlots were used as secondary antibodies (eBioscience). Protein A

Sepharose CL 4B (GE Healthcare) or Protein G Sepharose 4 Fast Flow (GE Healthcare) were used in IP. Human Dicer was detected with a polyclonal rabbit antibody raised to residues within aa1200-1300 (Abcam). For IP-northern blot experiments, GFP-Trap (Chromotek) and ANTI-FLAG M2 (Sigma) beads were used according to the manufacturer's instructions.

1.4.6 Protein purification & mass spectrometry sample preparation

FLAG-tagged proteins were purified using the ANTI-FLAG M2 Affinity Gel (Sigma A2220) following extract preparation, and eluted using the 3XFLAG peptide (Sigma F4799). Bands corresponding to sDCR-1-3FLAG on silver-stained gel (Invitrogen SilverQuest LC6070) were excised and submitted to the Taplin Biological Mass Spectrometry Facility (Harvard Medical School) for LC/MS/MS analysis after trypsin and chymotrypsin/trypsin digests.

1.4.7 2'O-Methyl Pull-down

N2 gravid adults were homogenized in lysis buffer (25 mM Hepes-KOH pH7.4, 150 mM NaCl, 1 mM EDTA, 1 mM DTT, 10 % glycerol, 0.5 % Triton X-100 and protease

inhibitors) using a stainless steel homogenizer. S10 lysate was pre-cleared with 25μL of m-280 streptavidin beads (Invitrogen) and an unrelated 2'O-methylated oligonucleotide (0.1μmol anti-miR-35) for 1h at 4°C with rotation. The supernatant was incubated with biotinylated 2'-O-Me oligonucleotides (0.1μmol) complementary to miR-1 or let-7 (or human miR-16 as a negative control) for 1 hour at 22°C. After centrifugation for 5min at 13000rpm, the supernatant is incubated with 25μL of m-280 streptadividin beads for 30 minutes at 4°C. Beads were washed three times using ice-cold lysis buffer containing 0.1% Triton X-100 and 2mM DTT, followed by a wash without detergent and 2mM DTT. Beads were resuspended in 50μL of 2X SDS loading buffer and eluted by heating at 95°C for 5 minutes. One tenth of the eluate is analyzed for ALG-1/2 interaction, and three fifths of the eluate for DCR-1.

1.4.8 Cell Culture

HEK293 cells were grown in Dulbecco's modified Eagle's medium supplemented with 10% FBS, 2mM L-Glutamine, 10mM HEPES, and Penicillin-Streptomycin (Wisent). Cells were transfected with FLAG-eDicer-MYC at ~90% confluence using Lipofectamine

2000 (Invitrogen 11668-027) according to the manufacturer's instructions, and were harvested 50 hours post-transfection for protein and RNA analysis.

1.4.9 Plasmid Construction

FL DCR-1 and sDCR-1 were amplified from *C. elegans* genomic DNA or cDNA using primers TDO577&578 (for FL) and TDO581&578 (for pre-sDCR-1). Forward primers contain a MYC tag and reverse primers contain a NotI site upstream of a stop codon. Amplicons were cloned first in pSC-A-amp/kan (Stratagene), sequenced fully, then subcloned into L2534 (Addgene plasmid 1608 from the Fire Lab Vector Kit) at the NheI-AgeI sites. The FLAG tag was added by ligating annealed primers TDO696&697 into the NotI site. Size and orientation of tags were verified by sequencing. An internal FLAG tag was generated by all-around PCR at aa1292 to 1299 within the linker region using primers TDO720&721. The catalytically dead pre-sDCR-1 transgene was generated by sequential all-around PCR using primers TDO859-886 (D1420A, E1578A, D1686A, E1804A), and the final construct was sequenced to confirm the substitutions were made and no additional mutations were present. The deletions to abolish sDCR-1 generation

and insert an HA tag were made by PCR using primers TDO1382-1387, followed by blunt ligation and sequencing to confirm. Residues deleted were: 1143 to 1162 (deletion1), 1163 to 1184 (deletion2), and 1200 to 1221 (deletion 3). The endogenous DCR-1 3' UTR was amplified from genomic DNA with primers TDO1079&1080 and replaced the *unc-54* 3' UTR in the pmyo-3::MYC-pre-sDCR-1-3FLAG transgene. The sDCR-1-DCR-1 3' UTR fragment was then subcloned into the let-858 expression vector L2865 for ubiquitous protein expression.

Human eDicer was amplified from UACC-812 total cDNA (gift from Caroline Moyret-Lalle) using TDO1597 &1239 and cloned into pSC-A-amp/kan. All clones originated from longer isoforms of the gene, and the frame was corrected by deletion PCR using TDO1683&1684 followed by blunt ligation. The eDicer amplicon was then subcloned into pcDNA3.1 for HEK293 cell expression.

Name	Description	Sequence
TDO577	FL DCR-1 L	GCTAGCATGGAACAGAAACTGATTAGCGAAGA AGATCTGGTCAGGGTAAGAGCTGATTTAC
TDO578	FL DCR-1 R	ACCGGTCTAGCGGCCGCAAACAGTTGTTAATG ATGGGC
TDO581	sDCR-1 L	GCTAGCATGGAACAGAAACTGATTAGCGAAGA

		AGATCTGCGTCGTTCAAGAACTGTGAGTAAC
TDO696	1FLAGFOR	ggc cgc GAC TAC AAG GAC GAC GAT GAC AAG t gc
TDO697	1FLAGREV	G GCC GC A CTT GTC ATC GTC GTC CTT GTA GTC GC
TDO720	intFLAGsdcrF	GATTATAAAGATGATGATGATAAACAAGAGGAT GAGGTAAAT
TDO721	intFLAGsdcrR	CTCTTGTTTATCATCATCATCTTTATAATCTTCT AAATTCAC
TDO859	sDCR D1420A F	GGAGCCTCTTTCTTGAAG
TDO860	sDCR D1420A R	AGAGGCTCCGATTGTTTC
TDO861	sDCR E1578A F	GTCGCAGCTTTGATCGGA
TDO862	sDCR E1578A R	AGCTGCGACTGCGTCAGC
TDO863	sDCR D1686A F	GGAGCCGCTGTTCTCGAC
TDO864	sDCR D1686A R	AGCGGCTCCGAGGAATTC
TDO865	sDCR E1804A F	TTTGCATCAGTAGCTGGC
TDO866	sDCR E1804A R	TGATGCAAATATGTCACC
TDO1079	xhoI dcr3utr sense	GGGCCCTAGTAAGATGTTCCAATTTTG
TDO1080	apaI dcr3utr antisense	CAGCTGATAGAATTAGATCATCATAAATACCAG
TDO1382	DEL1Fb	GTC CCT GAT TAC GCT GGT GCC AGG CTC ACT TCT AAC
TDO1383	DEL1Rb	ATC GTA AGG GTA ACC ACC AAG TCC TGA AAT TCA GAT GAT TAT TA
TDO1384	DEL2Fb	GTC CCT GAT TAC GCT GGT GGA TGG GGA GAT TGG GAT G
TDO1385	DEL2Rb	ATC GTA AGG GTA ACC ACC TGT GTA ATC ATG CAT TAC AAA CGG
TDO1386	DEL3Fb	GTC CCT GAT TAC GCT GGT GTT TTT GAT

		CCA TCT ACT GCT TCG TC
TDO1387	DEL3Rb	ATC GTA AGG GTA ACC ACC CAT GGG ACT ATT ATC TGG TTC AGG T
TDO1597	pcdnaedicer F2	AAGCTTATGGACTACAAGGACGACGATGACAA GCTGGCGTGGGAGTCAGATCAC
TDO1239	pcdnaedicer R	CTCGAGGCTATTGGGAACCTGAGGTTGATTAG CTTTGAGGCTTCGGAGG
TDO1683	edicer gen f	AGC TGA AAA TGA TAA TTA CTG
TDO1684	edicer gen r	AAA ATC CGC AGG AAG TGA TCT

Table 1.2: List of primers used in Chapter 1.

1.5 Acknowledgements

We thank Luc DesGroseillers for suggesting the cDNA expression experiment, and Mathieu Flamand for advice on the pull-downs. We also thank Darryl Conte Jr. and Martin Simard for helpful comments on the manuscript. Strains were obtained from the Caenorhabditis Genetics Center (CGC). This work was supported by the Canadian Institute of Health Research (CIHR) MOP 86577 (T.F.D.), the Canada Foundation for Innovation (CFI), the Fonds de la Rercherche en Santé du Québec, Chercheur-Boursier Salary Award J2 (T.F.D.). A.N.S. is the recipient of the CIHR Frederick Banting and Charles Best Canada Graduate Scholarship.

Preface to Chapter 2

Our work in Chapter 1 alerted us to the existence of a proteolytic processing event on full-length DCR-1 to generate a truncated form (sDCR-1) with distinct functional consequences on the RNAi pathways. This work did not, however, answer the question of if or how full-length DCR-1 could function differently in the RNAi pathways. We hypothesized that additional post-translational events could regulate DCR-1 function on the molecular level, and consequently the activity of the RNAi pathways. To this end, in the following chapter, we identified a dense cluster of phosphorylation sites near the key DCR-1 dsRNA substrate recognition domain, which plays a major role in the efficacy of the exoRNAi pathway, ERI endoRNAi protein interactions, and animal development.

Chapter 2: Phosphorylation-dependent regulation of DCR-1

Sawh A. N., Lewis A., Wohlschlegel J. and Duchaine T. F., Phosphorylation of Dicer in the substrate-recognition domain modulates its activity in RNAi pathways in *C. elegans*. (manuscript in preparation)

2.1 Introduction

Dicer (DCR-1 in *C. elegans*) is the essential RNase III enzyme central to the biogenesis of small RNAs in the miRNA, exoRNAi and ERI endoRNAi pathways. miRNAs are genome-encoded and undergo sequential processing events before being loaded into Argonaute (AGO) proteins within the RNA induced silencing complex (RISC). They then initiate silencing of a wide variety of protein coding genes post-transcriptionally, generally by base-pairing with imperfect complementarity to sequences in the 3' UTR of target mRNAs (Ambros, 2004; Bartel, 2009). The exoRNAi pathway is triggered by dsRNA from an exogenous source, which is processed into siRNAs, and loaded onto a distinct RISC, to direct negative transcriptional and post-transcriptional silencing (Fire et al., 1998; Hannon, 2002; Mello and Conte, 2004; Song et al., 2004; Tsai et al., 2015). In ERI endoRNAi, another endogenous class of siRNAs (26G siRNAs) is generated by DCR-1 and the ERI complex to regulate target genes through little known mechanisms (Duchaine et al., 2006; Lee et al., 2006; Pavelec et al., 2009).

Within these pathways, DCR-1 not only cleaves the trigger dsRNA, but participates in RISC loading, and further remains in complex with the AGO in many species during downstream effector steps of silencing as part of the “holo-RISC” (Kim et al., 2007; Lee et al., 2004b; Maniataki and Mourelatos, 2005; Pham et al., 2004). In *C. elegans*, DCR-1 is a member of distinct protein complexes that correspond to its activities in the three pathways. In the miRNA pathway, DCR-1 associates with the miRNA-specific AGOs ALG-1 and ALG-2. In the exoRNAi pathway, DCR-1 associates with the RDE complex, which contains its dsRNA binding domain (dsRBD) protein partner RDE-4 and the exoRNAi-specific primary AGO RDE-1. Finally, in ERI endoRNAi, DCR-1 associates with the ERI complex, a key member of which is the RNA-dependent RNA polymerase (RdRP) RRF-3 which is thought to produce the trigger dsRNA specific to this pathway (Duchaine et al., 2006; Thivierge et al., 2012).

Molecularly, DCR-1 is made up of a large N-terminal DExD/H box helicase domain, a domain of unknown function (DUF), a structure known as the “platform” which is implicated in phosphate-binding, a substrate recognition domain (PAZ), a long linker

(also known as the ruler), tandem RNase III (RIII) domains, and a dsRBD. Substrates are recognized at their 3' ends by the PAZ and sometimes the platform, which are separated in space by the linker/ruler from the RIII domains. The RIII domains form an intramolecular dimer and cleave the substrate (Macrae et al., 2006; Park et al., 2011; Takeshita et al., 2007; Zhang et al., 2004). Since *C. elegans* only carries one *dcr-1* gene, like all characterized vertebrates, and no alternative splicing variants have been identified thus far, we reasoned that post-translational modification (PTM) of the DCR-1 protein could play a pivotal role in directing the activities of the enzyme in the three RNAi pathways described above. In fact, we have previously found that DCR-1 is proteolytically cleaved within the linker, severing the N-terminal domains and producing a stable C-terminal fragment which acts in multiple RNAi pathways (Sawh and Duchaine, 2013). Very recently, another group has found that DCR-1 is phosphorylated at conserved residues of the RIIIb and dsRBD in the *C. elegans* germline prior to fertilization, which leads to suppression of DCR-1 activity and to its nuclear translocation (Drake et al., 2014). Our own studies have identified several additional phosphorylation

sites on DCR-1, and here we present evidence supporting the role of PTMs in the PAZ domain influencing DCR-1 activity in exoRNAi, DCR-1's protein-protein interactions, and the overall physiology of the animal.

2.2 Results

2.2.1 DCR-1 is phosphorylated at multiple sites *in vivo*

In order to determine if DCR-1 is phosphorylated *in vivo*, DCR-1 immunoprecipitates (IPs) from *C. elegans* embryos were submitted for mass spectrometry (MS) to identify phosphorylated peptides. A number of phosphorylated peptides were detected, mapping to 32 amino acid residues in DCR-1. We decided to focus on sites that were 1) conserved between *C. elegans* DCR-1 and human Dicer, and 2) located in regions known to be critical to DCR-1 function. In one such region, the 3' end of the PAZ domain (amino acids 959-989), phosphorylation in a cluster of closely-spaced serines and threonines was detected (Figure 2.1A). We chose these residues, dubbed the "ST cluster", for functional characterization. The ST cluster spans 31 amino acids and contains 8 sites of phosphorylation detected by MS and an additional 7 sites predicted to be phosphorylated by two or more algorithms (See Appendix 3). The first half of this

cluster aligns to DCR-1 protein homologs in other species, while the second half could represent a nematode-specific expansion (Figure 2.1B). This region of human Dicer (amino acids 755-1055) has been crystallized recently (Tian et al., 2014), and was shown to be in close proximity to the dsRNA 3' overhang binding site in the PAZ domain and a region immediately N-terminal to the PAZ called the "platform". In fact, two crystal structures of the human ST cluster were shown to adopt different conformations – one partially helical and one completely disordered/melted (Figure 2.1C). We therefore hypothesized that the ST cluster could adopt multiple conformations *in vivo*, and that phosphorylation in this region could affect dsRNA binding, protein-protein interactions, and ultimately DCR-1 activity.

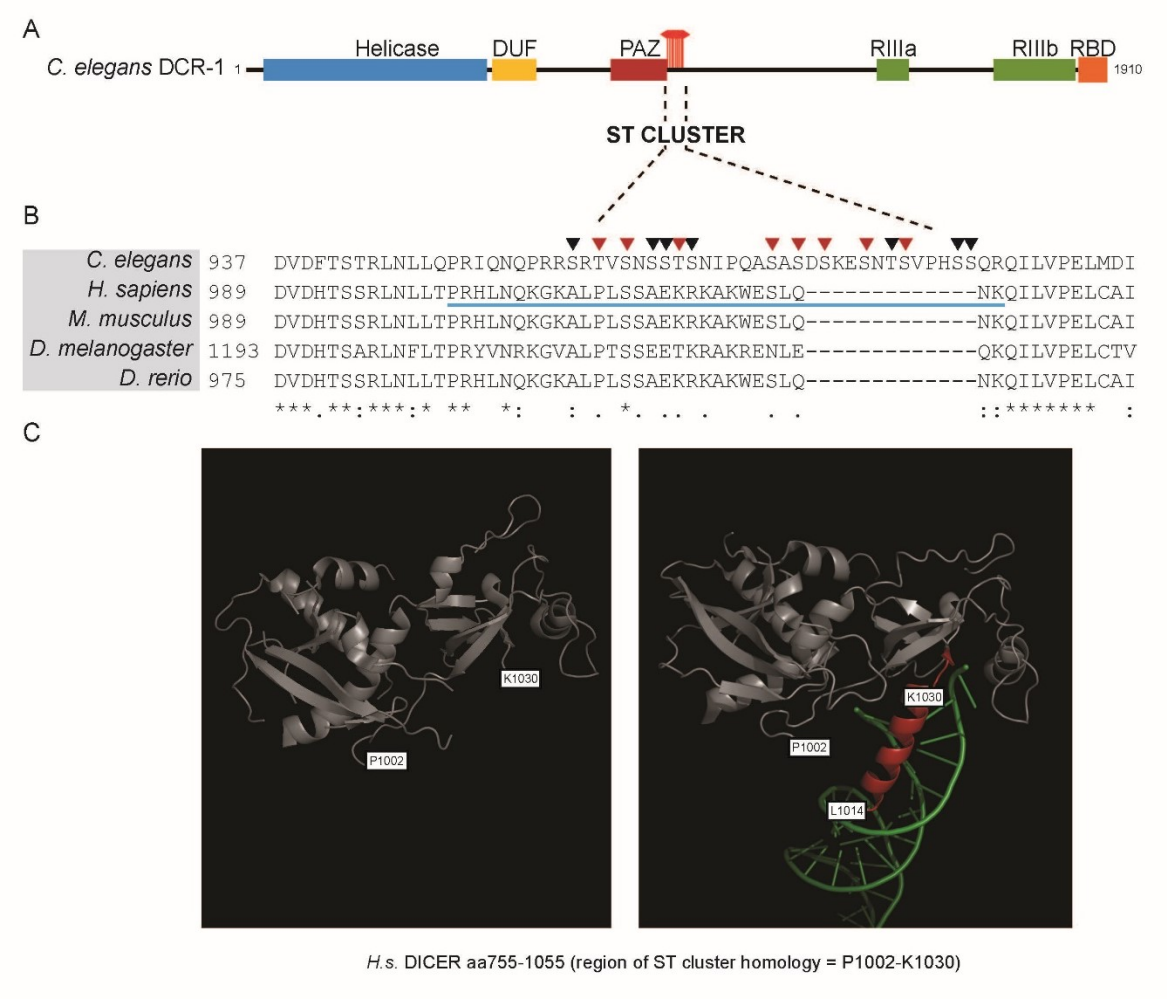


Figure 2.1: Phosphorylation of DCR-1 in the dsRNA substrate-recognition domain. (A) Schematic representation of DCR-1 with the phosphorylation positions of the "ST cluster" indicated with red flags in the 3' end of the PAZ domain. (B) Clustal W alignment of the protein sequence surrounding the ST cluster of DCR-1 homologs from the species indicated. Arrowheads point to the 15 residues in *C. elegans* DCR-1 identified (red) or predicted (black) to be phosphorylated. (C) Crystal structures of a human Dicer fragment, PDB files 4NGF and 4NHA (Tian et al., 2014)) containing the region homologous to the ST cluster (P1002-K1030, blue underline in (B)). The first

crystal (left) shows that the ST cluster residues completely lack electron density, indicating this sequence exists in a flexible and disordered state, and the second crystal (right) shows the ST cluster adopts a partially helical structure (red) in close proximity to dsRNA substrate.

2.2.2 Predicted kinases of the ST cluster

We next used both *in silico* and experimental methods to test which kinases could be responsible for ST cluster phosphorylation. Using 4 predictive programs, many kinases were identified as candidates in this region (See Appendix 3). To test six of these experimentally, endogenous DCR-1 from wild-type gravid adult extracts was immunoprecipitated, and the recovered fraction was probed with a panel of antibodies specific to phosphorylated substrates of the kinases AMPK, AKT, PKA, ATM/ATR, PKC and CDK (Figure 2.2). The AMPK phospho-substrate antibody strongly recognized full-length DCR-1 in these experiments, and those of PKC and CDK produced weaker signals. Next, the region of DCR-1 recognized by the AMPK phospho-substrate antibody was narrowed down. To do this, DCR-1 IPs were carried out in strains expressing full-length DCR-1 and truncated DCR-1 proteins. IPs were also performed

from wild-type animals exposed to dsRNA against the muscle gene *unc-22*, animals with no germline (*glp-4(bn2)*), and animals with no sperm (*fem-1(hc17)*), and subsequently probed with the AMPK phospho-substrate antibody (Figure 2.3). Endogenous full-length DCR-1 and transgenic pre-sDCR-1 are proteolytically cleaved *in vivo* to produce the stable and active sDCR-1, which acts in multiple RNAi pathways ((Sawh and Duchaine, 2013), See Chapter 1). We found that the AMPK phospho-substrate antibody recognized full-length DCR-1 and pre-sDCR-1, but not sDCR-1. This narrows down at least one region of antibody recognition to the first half of the DCR-1 ruler/linker, which coincides with the location of the ST cluster. We note that based on these results, additional sites in the N-terminus of full-length DCR-1 cannot be ruled out. Furthermore, the AMPK phospho-substrate antibody could detect DCR-1 in extracts of germline-loss (*glp-4(bn2)*) and sperm-loss (*fem-1(hc17)*) mutants, indicating that phosphorylation of DCR-1 is at least partially somatic (Figure 2.3). Altogether, these experiments suggest that AMPK phosphorylates DCR-1 in the ST cluster, where it is predicted to do so at position 961 in the amino acid sequence.

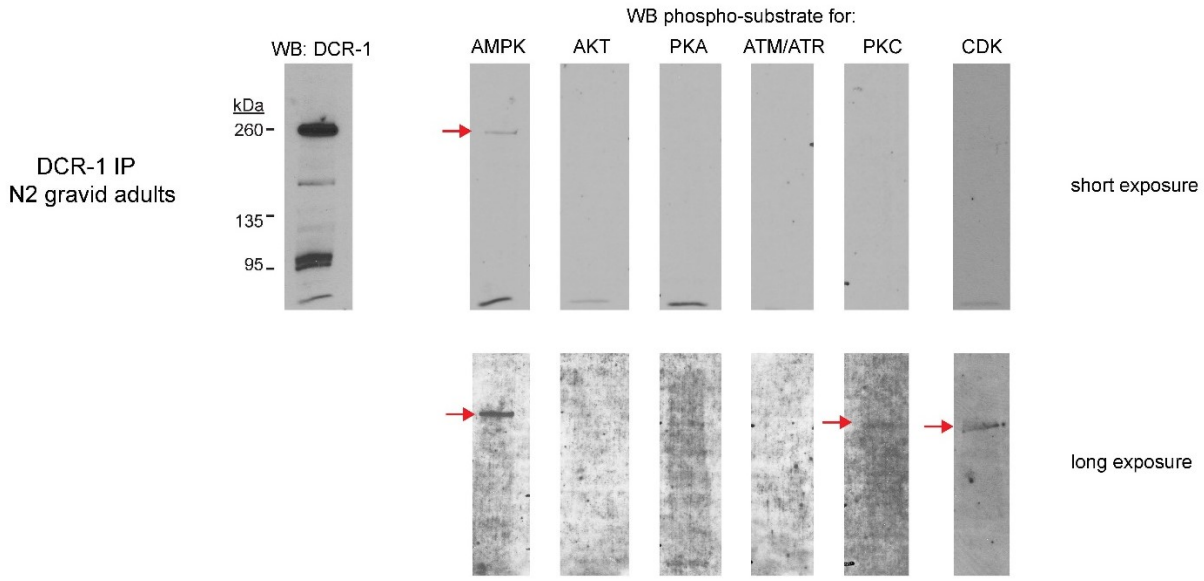


Figure 2.2: DCR-1 encodes phosphorylated consensus sites for multiple kinases. DCR-1 IPs from wild-type gravid adults were probed with an antibody to endogenous DCR-1 and antibodies that specifically recognize the phosphorylated consensus sites of different kinases. The AMPK phospho-substrate antibody produced the strongest signal for full-length DCR-1, while weaker signals were seen using the CDK and PKC phospho-substrate antibodies (red arrows).

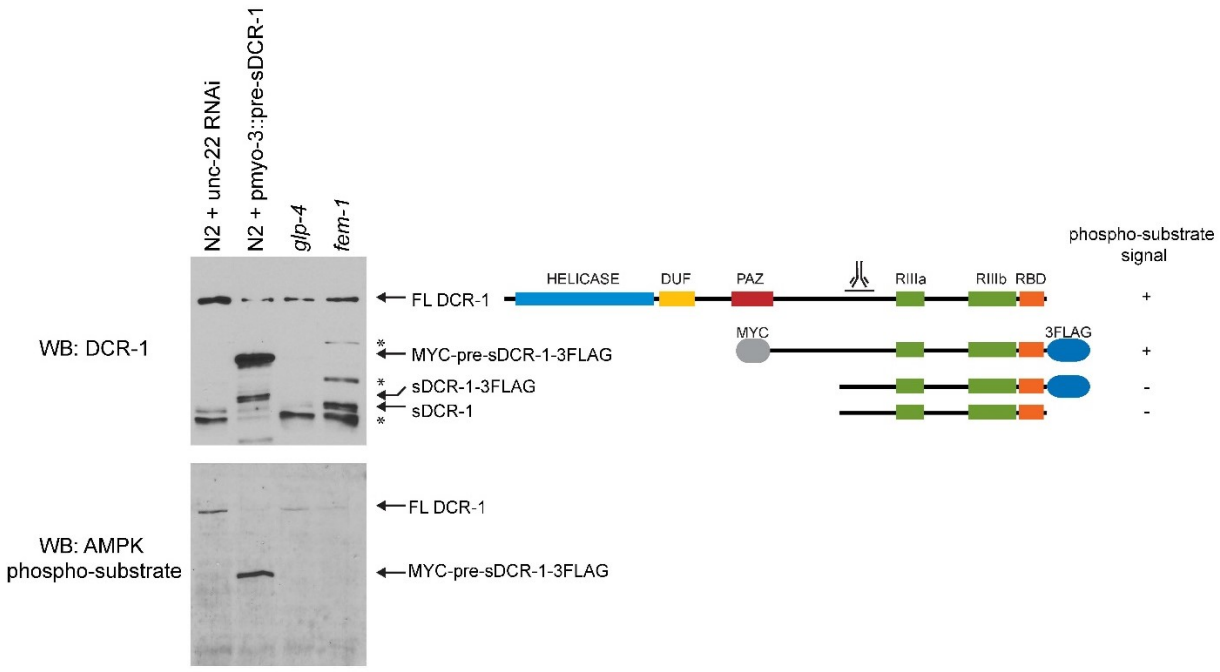


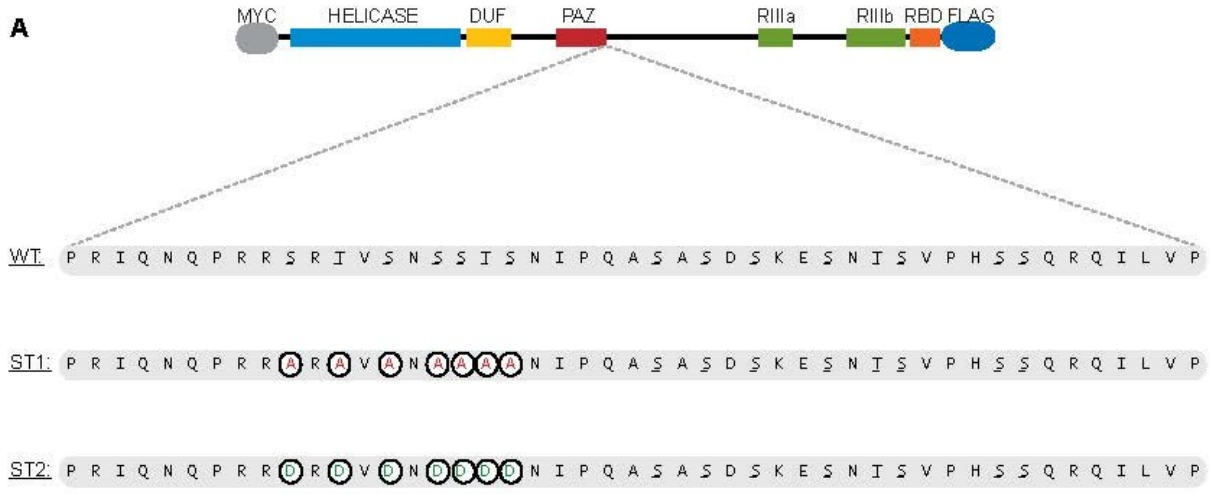
Figure 2.3: AMPK putatively phosphorylates DCR-1 within the first half of the linker. The AMPK phospho-substrate antibody was used to probe DCR-1 IPs. The following conditions were tested: wild-type animals undergoing *unc-22* RNAi, wild-type animals expressing a muscle-driven pre-sDCR-1 transgene, germline-loss mutant *glp-4(bn2)* and sperm-loss mutant *fem-1(hc17)*. The AMPK phospho-substrate antibody recognizes full-length DCR-1 and pre-sDCR-1, but not sDCR-1. This indicates that the region of antibody recognition and putative AMPK phosphorylation is between the end of the PAZ domain and the middle of the linker, where the ST cluster phosphorylation lies.

2.2.3 Generation of strains carrying DCR-1 with mutations in the ST cluster

We next tested the function of the ST cluster phosphorylation through the use of phospho-mimetic (serine/threonine to aspartic acid) and phospho-null (serine/threonine

to alanine) mutations. First, we elected to test the effect of multi-site mutations, instead of single site mutations (for example, the AMPK site alone), on DCR-1 function to eliminate the possibility of contributions from redundant or compensatory phosphorylation on nearby residues in the ST cluster. Mutations were generated for the conserved region of the cluster (amino acids 959 - 968) in an N- and C-terminally tagged full-length DCR-1 construct ("multi-site mutants", Figure 2.4A). Simple transgenic arrays were then expressed in wild-type and *dcr-1* null mutant backgrounds (Figure 2.4B). Since complete loss of DCR-1 is lethal, *dcr-1(ok247)* null mutants maintain viability through a balancing segment of extrachromosomal DNA (the free duplication sDp3) carrying a wild-type copy of the *dcr-1* gene. Attempts to rescue the null allele (i.e. replacement of the balancer with tagged mutant or wild-type constructs) by traditional transgenic array methods did not produce any viable progeny. However, it seemed unlikely that the entire pool of DCR-1 would be present in a phosphorylated or unphosphorylated state in the entire animal. Thus, we opted to analyze the function of the ST cluster mutants by expression from an extrachromosomal array over

endogenous wild-type *dcr-1* loci. This strategy generates a mixed population of DCR-1 proteins that are either wild-type or mutated in the ST cluster in the entire animal. We confirmed that transgenic extrachromosomal arrays drove the expression of DCR-1 to a similar level as seen in wild-type animals (Figure 2.5).



B

Genetic background	*Transgene*	*# Lines*
N2 (wild-type)	MYC-DCR-1-3FLAG (WT)	13
N2 (wild-type)	MYC-DCR-1 ST1 (A)-3FLAG	3
N2 (wild-type)	MYC-DCR-1 ST2 (D)-3FLAG	7
dcr-1 -/- (+ balancer)	MYC-DCR-1-3FLAG (WT)	5
dcr-1 -/- (+ balancer)	MYC-DCR-1 ST1 (A)-3FLAG	1
dcr-1 -/- (+ balancer)	MYC-DCR-1 ST2 (D)-3FLAG	3

Figure 2.4: Transgenic strategy of phospho-null and phospho-mimetic mutations on full-length DCR-1. (A) Schematic representation of DCR-1 with multi-site mutations

generated in the conserved region of the ST cluster to abolish (ST1) or mimic (ST2) complete phosphorylation in the region. (B) Transgenic lines obtained by microinjection of constructs expressing wild-type or mutant DCR-1 from the ubiquitous promoter *let-858*.

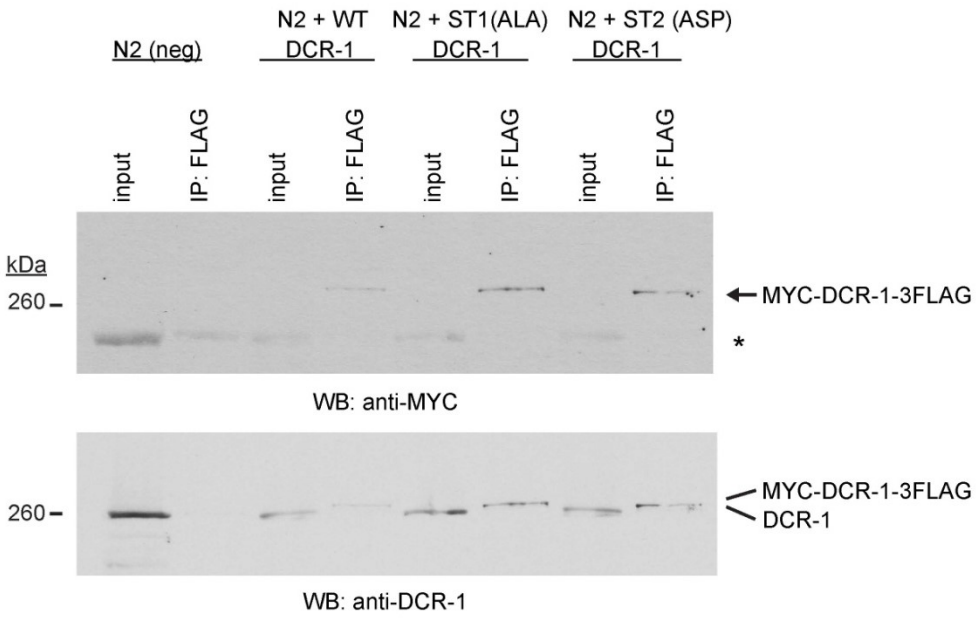


Figure 2.5: Expression of transgenic DCR-1 constructs. FLAG IP and western blots on mixed stage animals of the wild-type (N2) background expressing transgenic MYC-DCR-1-3FLAG from extrachromosomal arrays. N2 was used as negative control. Transgenic DCR-1 with the wild-type sequence (WT), phospho-null mutations (ST1(ALA)), and phospho-mimetic mutations (ST2(ASP)) are detected specifically with the MYC antibody (top). Total DCR-1 protein is detected in all samples with the DCR-1 antibody (bottom).

2.2.4 The ST cluster is required for exoRNAi activity

We next assessed the effect of phospho-mimetic and null mutations on the efficacy of the exoRNAi pathway. To do so, the exoRNAi pathway was triggered by feeding P0 animals bacteria expressing *unc-22* dsRNA, and the resulting phenotype in the F1 generation was quantified (Fire et al., 1998; Yigit et al., 2006). By scoring the phenotype associated with *unc-22* knockdown, body wall muscle twitching, we showed that animals expressing phospho-mimetic DCR-1 (33% twitching) behaved similarly to those expressing WT DCR-1 (35% twitching) (Figure 2.6). On the other hand, animals expressing phospho-null DCR-1 had significantly less twitching (13%) (Figure 2.6), indicating that loss of phosphorylation in the ST cluster negatively impacts exoRNAi activity.

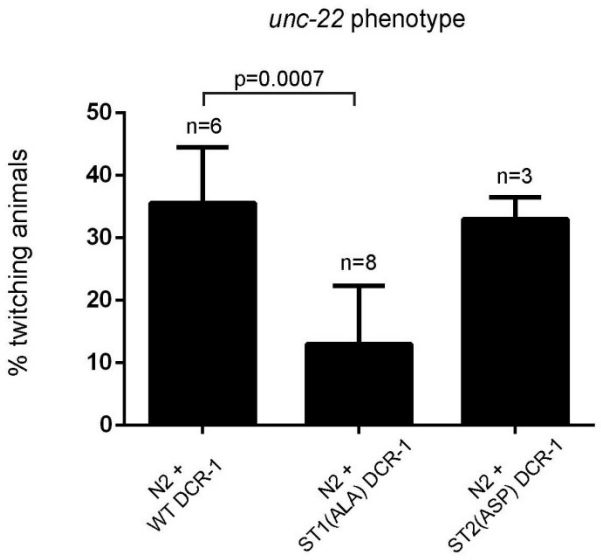


Figure 2.6: The DCR-1 ST cluster is required for exoRNAi activity. RNAi against muscle gene *unc-22* was triggered in P0 animals expressing WT/mutant constructs from extrachromosomal arrays, and the resulting twitching phenotype in F1 animals was quantified. ST cluster phospho-null mutants display significantly reduced exoRNAi response compared to wild-type or phospho-mimetic, calculated by two-tailed *t*- test. Data are plotted as mean ± standard deviation. n= number of biological replicates indicated for each strain.

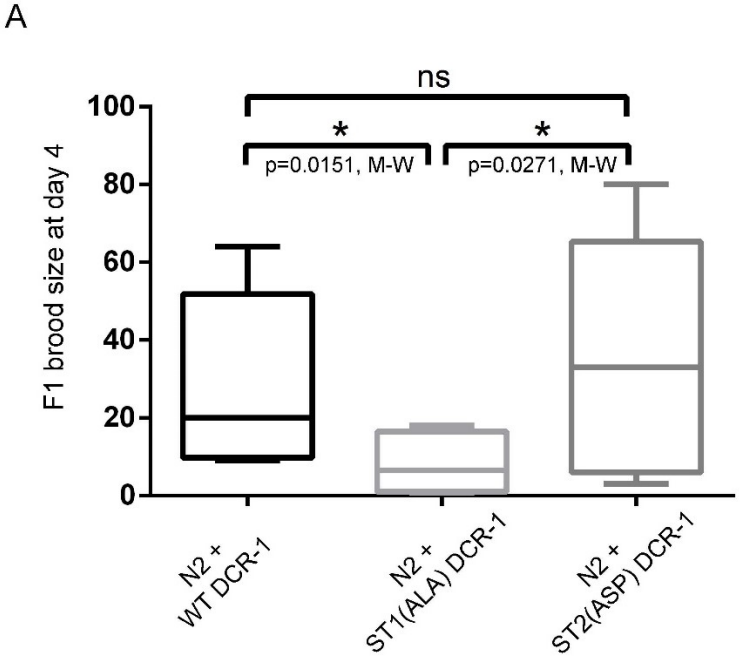
2.2.5 Mutation of the ST cluster leads to severe developmental defects

We noticed that following genome integration of the extrachromosomal arrays, strains expressing the phospho-null and –mimetic mutants displayed gross physiological defects compared to the strain expressing the WT DCR-1. ST1 (ALA) mutants showed the most severe defects, with a high degree of variability from animal to animal. The

most consistent phenotypes observed were short and chubby (dumpy), protruding vulva, and rare bursting vulva. ST2 (ASP) mutants were close to WT in appearance, but also occasionally were dumpy or showed protruding vulva (Figure 2.7). To assess the effect of these mutants on fertility and developmental timing, synchronized P0 L4 animals were isolated and their progeny (F1) were scored for developmental progression and brood size after four days at room temperature. While the staging of ST2 (ASP) mutants (53% adult, 36% L3-L4) have patterns similar to WT (71% adult, 22% L3-L4), we found that ST1 (ALA) mutant animals display delayed developmental progression (28% adult, 51% L3-L4) (Figure 2.8). Moreover, ST2 (ASP) animals have similar brood size to WT animals (means of 35 and 28 F1s, respectively), while ST1 (ALA) animals have low amounts of viable progeny (mean of 8 F1s). We consistently observed phenotypic variation in the ST1 (ALA) and ST2 (ASP) strains, and even the strain expressing WT DCR-1 displayed variable brood size. Nevertheless, these data demonstrate that the ST cluster plays an important role in the overall health and development of the animal.

Figure 2.7: ST cluster phosphorylation mutants display severe developmental defects. Brightfield microscopy images of gravid adult animals of wild-type genetic background

(N2) with additional integrated WT (panels A-D), ST1 (ALA) (panels E-L) or ST2 (ASP) (panels M-T) DCR-1 transgenes. ST1 phospho-null mutants display severe developmental defects including protruding vulva (black arrows), dumpy (white arrows) and rare bursting vulva (red arrow), with high variability. ST2 phospho-mimetic mutants are closer to WT in phenotype, but are sometimes dumpy (white arrows), with occasional protruding vulva (black arrow).



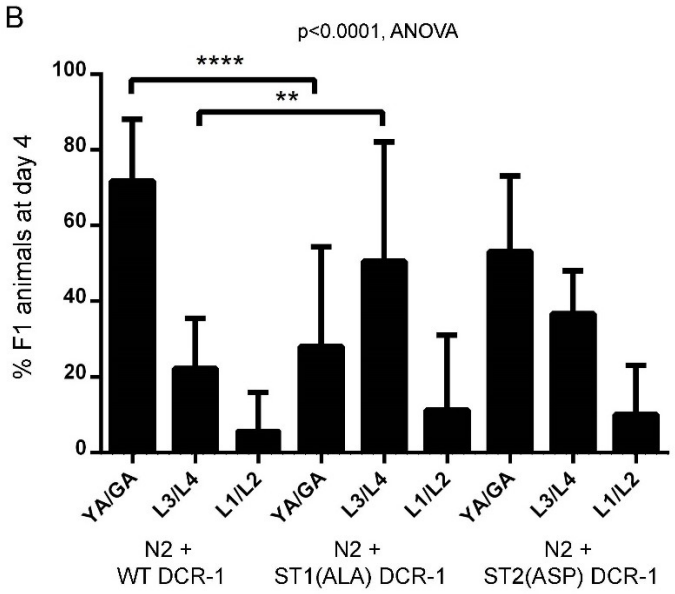


Figure 2.8: ST cluster phosphorylation mutants display reduced brood size and delayed development. Strains of wild-type genetic background (N2) with additional integrated WT, ST1 (ALA) or ST2 (ASP) DCR-1 transgenes were synchronized in the P0 generation by picking gravid adults onto individual plates, then scoring the brood size (A) and staging (B) of the F1 generation after 4 days at room temperature. Data are grouped by stage: young adult and gravid adult (YA/GA), larval stage 3 and larval stage 4 (L3/L4), and larval stage 1 and larval stage 2 (L1/L2). Data for each strain was variable between biological replicates, but ST2 brood size and staging patterns resemble that of WT, while ST1 had reduced brood size and a peak of L3/L4 animals compared to YA/GA. Data are plotted as box and whisker plot in (A) and as mean ± standard deviation in (B). Statistical significance was calculated using Mann-Whitney test (A) and ANOVA (B). n=10 biological replicates for each strain.

2.2.6 Phospho-null DCR-1 is directed towards the ERI endoRNAi pathway protein complex

Since phospho-null mutations resulted in diminished exoRNAi activity and abnormal development in animals, we hypothesized that this effect may be due to altered protein complexes dependent on phosphorylation state. Thus, to determine if ST cluster phosphorylation could affect protein interactions with DCR-1, we probed DCR-1 FLAG IPs with key members of the DCR-1-dependent RNAi pathways.

By analyzing the proteins bound to WT and ST1 (ALA) DCR-1, we found that phospho-null mutations resulted in an enhanced interaction between DCR-1 and RDE-4. RDE-4 is a dsRNA binding protein (dsRBD) partner and member of both the exoRNAi and ERI endoRNAi pathways. Furthermore, we observed a greater amount of endoRNAi components ERI-1b and ERI-5 in IPs of the phospho-null DCR-1 compared to wild-type DCR-1. This data indicates that unphosphorylated DCR-1 favours interaction with the ERI endoRNAi complex. The enhanced RDE-4 interaction may be as part of the ERI endoRNAi complex and/or the RDE exoRNAi complex. Interaction between DCR-1 and

the miRNA-specific Argonautes ALG-1/2 were not detectable under these conditions. The effect on the miRNA pathway will therefore be assessed through small RNA analysis.

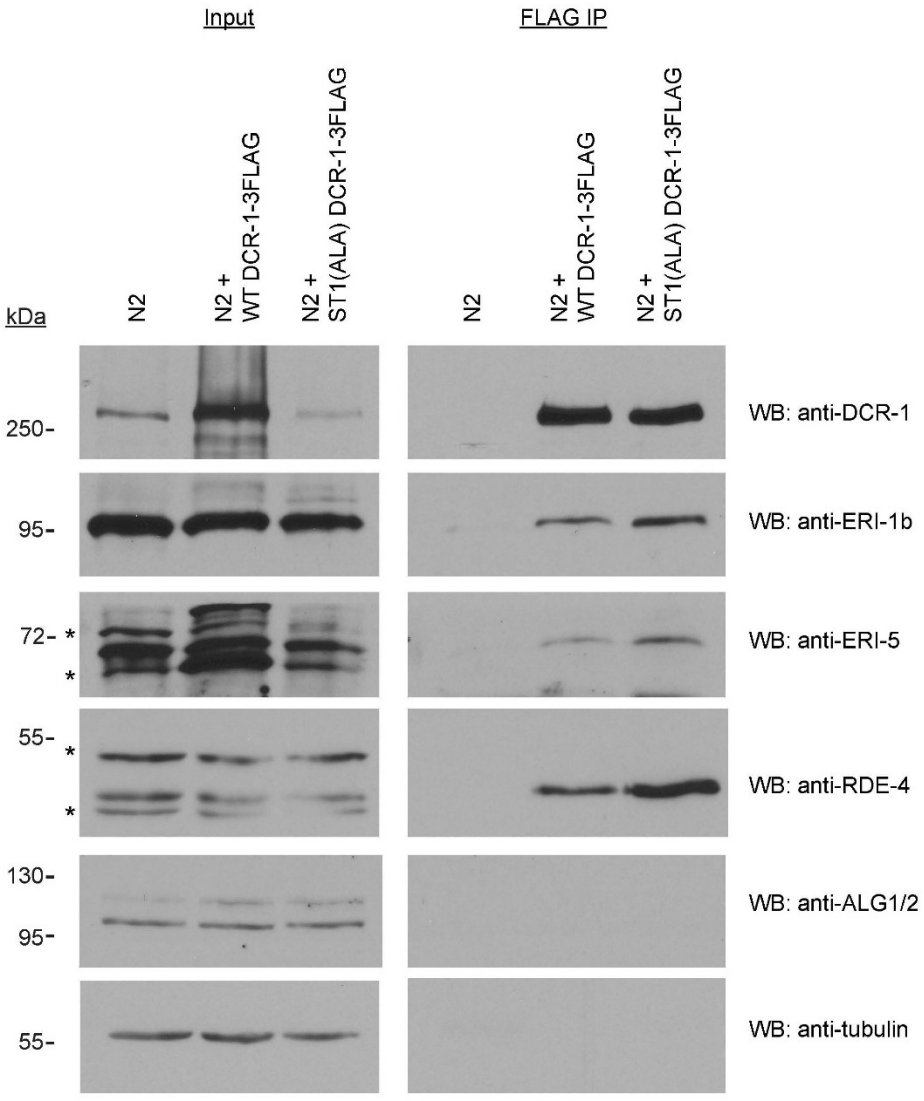


Figure 2.9: Phospho-null DCR-1 reinforces specific protein interactions. FLAG IPs of integrated transgenic tagged DCR-1 constructs expressing the wild-type DCR-1

sequence (WT) or phospho-null mutant in the ST cluster (ST1(ALA)) were performed on extracts from gravid adults. In large-scale culture conditions, animals expressing the integrated ST2 (ASP) DCR-1 lost expression of the transgene, hence protein complex analysis was not possible at this point. N2 extracts were used as the negative control as they are the same genetic background and carry no FLAG-tagged protein. Representative blots from two biological replicates are shown.

2.3 Discussion

We have identified a cluster of closely-spaced phosphorylation sites on endogenous DCR-1 protein, in a conserved and functionally important region for the activity of the enzyme. Its location at the end of the PAZ domain places the ST cluster in close proximity to the substrate dsRNA terminus, as predicted by analogy to the human sequence. We hypothesized that phosphorylation in this sequence would affect RNA binding, subsequent production of si/miRNAs, RISC-loading functions, and/or protein interactions of DCR-1. In a recent crystal structure of a truncated PAZ-containing protein fragment of human Dicer, three residues of the region homologous to the ST cluster were shown to make intermolecular contacts with the dsRNA substrate: S1015,

S1016 and W1024 (referred to as S1005, S1006 and W1014 mistakenly in (Tian et al., 2014)). Mutation of the serines to alanine did not significantly alter substrate binding in the context of the truncated protein fragment using surface plasmon resonance. However, mutation of the tryptophan to alanine reduced substrate binding ~2.5-fold, while the tryptophan mutated to the positively charged arginine increased substrate binding ~2-fold (Tian et al., 2014). A related study from the same group found that the tryptophan to alanine mutation did not alter full-length Dicer's ability to produce siRNAs in an *in vitro* cleavage reaction on a 35bp dsRNA substrate, and the tryptophan to arginine mutation was not tested (Park et al., 2011). Tian et al. (2014) also showed that hDicer with a deletion of amino acids S1016 to N1029 (referred to as S1006-N1019 mistakenly) produced miR-16 and miR-21 at similar levels to wild-type Dicer when transfected into Dicer-null mouse embryonic stem cells. This span of residues corresponds to S963-Q990 in *C. elegans* DCR-1. It is important to note that the robustness and reproducibility of this result is unclear at this point, as only the results of a single experiment were presented in the publication, and the total levels of Dicer

proteins were not quantified. We therefore argue that a comprehensive examination of the effect of mutations in this region on full-length Dicer is needed *in vivo*, in a context where Dicer is normally active (instead of Dicer-null cells), and on a variety of physiologically relevant Dicer substrates.

To address the role of this region in *C. elegans*, we engineered multi-site phospho-null and phospho-mimetic ST cluster mutations in full-length DCR-1 and expressed them over the wild-type endogenous locus. Surprisingly, doing so led to dramatic phenotypes in strains expressing the phospho-null (ST1(ALA)) DCR-1 protein, but not those expressing additional wild-type (WT) and phospho-mimetic (ST2(ASP)) DCR-1 copies. ST1 (ALA) mutants display reduced fertility, delayed development, and severely dumpy, protruding vulva, and bursting vulva phenotypes, while ST2 (ASP) mutants are only slightly dumpy. The phospho-null ST cluster is therefore more deleterious to the health of the animal than the phospho-mimetic. We hypothesize that these phenotypes may be the result of a dominant negative effect of the mutant DCR-1 protein, since both wild-type endogenous copies of *dcr-1* are present in these strains. We further hypothesize

that these phenotypes are at least in part due to defects in the miRNA pathway, since loss of the exoRNAi and endoRNAi pathway functions do not produce similar phenotypes. exoRNAi mutants are defective in gene silencing initiated from exogenous triggers and display some transposon activation, but closely resemble wild-type animals in growth and development (Tabara et al., 1999). ERI endoRNAi mutants show temperature-sensitive sterility and a high incidence of males (Duchaine et al., 2006; Kennedy et al., 2004; Simmer et al., 2002), phenotypes we have not observed in the ST cluster mutants. In contrast, miRNA pathway mutants have similar phenotypic abnormalities as the ST1 (ALA) animals. Defects in miRISC components as well as specific miRNAs (lin-4 and let-7) lead to defects in vulval and hypodermal cell differentiation to produce protruding and bursting vulva (Abrahante et al., 2003; Ding et al., 2005; Euling and Ambros, 1996; Lin et al., 2003). The reduction of viable progeny seen in ST1 (ALA) mutants may be due to defects in specific embryonic miRNA families (miR-35-42, miR-51-56, and miR-58/bantam) that are essential for early development (Alvarez-Saavedra and Horvitz, 2010). The underlying cause of the dumpy phenotype is

more difficult to predict and pinpoint, as it has not been previously linked to defects in any RNAi pathway. The dumpy phenotype has resulted from mutations in genes in a wide range of cellular processes, from X-chromosome dosage compensation and sex determination (Hodgkin, 1983; Hsu and Meyer, 1994; Meneely and Wood, 1984; Plenefisch et al., 1989) to collagen biosynthesis and extracellular matrix cuticle production (McMahon et al., 2003). To corroborate our predictions, it will be important to analyze the effects of the ST cluster mutants of DCR-1 on the small RNA populations. This will be done by small RNA sequencing to detect changes in endo-siRNAs and miRNAs, as well as exo-siRNAs following addition of exogenous dsRNA trigger. Northern blots on representative RNAs from the different DCR-1 dependent pathways will be used to support the sequencing data. It may then be possible to identify downstream targets of one or more mis-regulated mi/siRNA that would account for the observed phenotypes. It is also possible that some of the observed phenotypes are due to defects in non-RNAi roles of DCR-1. DCR-1 has recently been shown to bind miRNAs, tRNAs, snoRNAs, mRNAs, and promoter RNAs. Surprisingly, DCR-1 does not

generate siRNAs from some targets and therefore regulates their expression passively (Rybak-Wolf et al., 2014). It will therefore be informative to analyze the effects of DCR-1 ST cluster mutation on all RNA populations.

Through a quantitative assay for the exoRNAi pathway activity, we found that phospho-null DCR-1-carrying animals were deficient compared to wild-type or the phospho-mimetic mutants. We speculate, based on this data, that a significant fraction of the pool of DCR-1 protein in the animal dedicated to the exoRNAi pathway is phosphorylated in the ST cluster. In order to test this, we will immunoprecipitate the RDE exoRNAi and ERI endoRNAi complexes specifically, and test the phosphorylation state of DCR-1 either by mass spectrometry or by using the AMPK substrate antibody. ST cluster phosphorylation, due to its location close to the substrate binding determinants, could also alter the cleavage activity of DCR-1. DCR-1 is known to cleave dsRNA processively in the presence of a long dsRNA substrate and conversely, in a single turnover manner in the presence of short hairpin substrates (Cenik et al., 2011; Sinha et al., 2015; Welker et al., 2011). Since the substrate for the exoRNAi pathway is long

dsRNA, and we show that wild-type and phospho-mimetic DCR-1 function similarly in response to this substrate, it is possible that phosphorylation in the ST cluster is beneficial, or neutral, to a processive mode of action. This could be directly tested in *in vitro* cleavage assays using *C. elegans* extracts containing wild-type or ST cluster mutant DCR-1 as shown previously (Welker et al., 2011).

We also reasoned that phosphorylation in the ST cluster could alter important interactions with DCR-1 protein partners. Proteins known to bind directly to DCR-1 in its N-terminus include those of the ERI complex (Thivierge et al., 2012). We have also observed that RDE-4 interacts with the N-terminal half of DCR-1 (Sawh and Duchaine, 2013). These very interactions are strengthened in the ST cluster phospho-null mutant DCR-1. Our data indicate that the ST cluster may be an important site of ERI complex and RDE-4 direct interaction. Alternatively, it is possible that phosphorylation of the cluster could induce a conformational change in DCR-1 at a distal site required for these protein interactions. Another possibility is that ST cluster phosphorylation state could reinforce or inhibit an interaction with an unknown protein which bridges the DCR-1 –

ERI complex/RDE-4 interaction. We propose that the pool of DCR-1 dedicated to the ERI endoRNAi pathway is in the unphosphorylated state. RDE-4 is an important member of both the exoRNAi and endoRNAi pathways (Thivierge et al., 2012), thus it will be important to determine if a separate population of RDE-4 is bound to unphosphorylated DCR-1 versus phosphorylated DCR-1.

Finally, to determine which residue(s) of the ST cluster is critical to the functions described above, we opt to engineer mutations on the endogenous *dcr-1* locus by CRISPR/Cas9 gene editing technology using a recently described co-conversion strategy with which we have had preliminary success (Ward, 2015). We elect to mutate the putative AMPK target site (T961) to alanine or aspartic acid to generate phospho-null and –mimetic mutations respectively, as in the multi-site mutants (Figure 2.10). If the mutants are viable, we will assess their phenotypes, protein interactions and small RNA production. However, if the putative AMPK site is the key to ST cluster function, we predict that the phospho-null mutant on the endogenous *dcr-1* gene will not be viable, based on the severe developmental defects observed by expressing the multi-

site phospho-null mutant (ST1(ALA)). Alternatively, it is possible that multiple sites in the ST cluster contribute to its function. We will therefore perform small RNA analysis on multi-site and single-site ST cluster mutants in parallel.

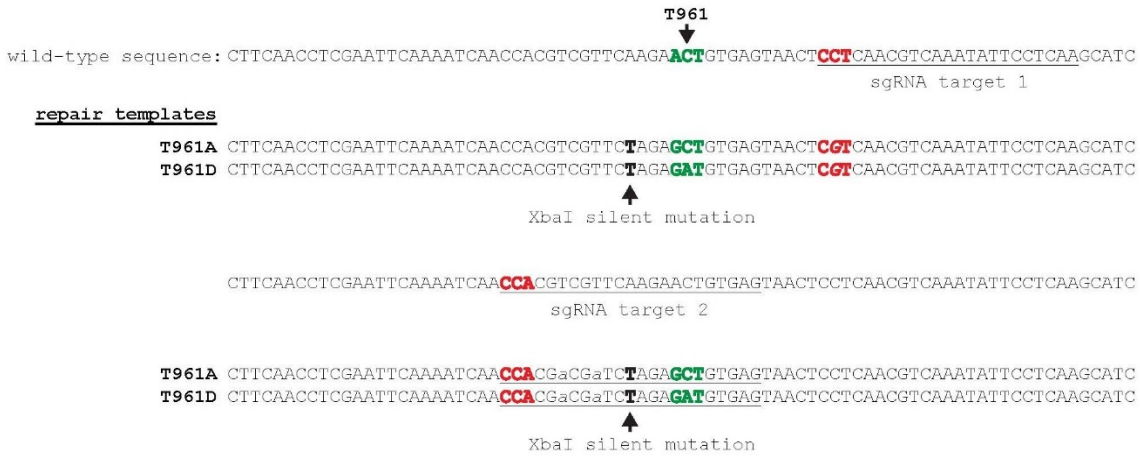


Figure 2.10: Gene editing strategy for T961 in the DCR-1 ST cluster. sgRNAs and repair templates were designed to mutate T961 to A or D (green), and also to introduce a silent mutation nearby to generate an XbaI restriction site (bold). The PAM site for each target is highlighted in red, and is also silently mutated in the repair template to prevent re-targeting of the edited locus (italics). For target site 2, mutation of the PAM would change the amino acid sequence, so silent mutations downstream are made (lowercase italics). Note that sgRNA sequences are the reverse complements of the underlined regions and do not include the PAM. This method was designed according to the protocol of (Ward, 2015).

Determining the kinase responsible for DCR-1 ST cluster phosphorylation will link the regulation of DCR-1 activity to an upstream signalling cascade, which could be a result of specific developmental programs within the organism. Our data indicates that DCR-1 phosphorylation is not exclusive to the germline, and it will be very interesting in the future to determine in which cells specifically the events of phosphorylation and dephosphorylation occur.

2.4 Materials and methods

2.4.1 *C. elegans* Strains and RNAi Assays

All strains were cultured as in (Brenner 1974). N2 was used as the wild-type strain. Alleles used were *glp-4(bn2)*, and *fem-1(hc17)*. RNAi was performed as in (Fire et al., 1998; Timmons et al., 2001). *unc-22* RNAi was performed on L4 P0 animals and scored at gravid adult F1 animals. Under standard conditions, diminished function of the *unc-22* gene leads to a severe twitching phenotype. *unc-22* dsRNA-expressing bacteria was cultured at 37°C in LB and 100µg/ml ampicillin. 100µg/ml ampicillin and 1mM IPTG were included in the standard NGM agar plates for the RNAi assay. Plates were seeded with a 1:1 mixture of OP50 bacteria and *unc-22* dsRNA-expressing bacteria.

Seeded plates were left at room temperature overnight to induce expression of *unc-22* dsRNA. For each strain, between 5 and 10 L4 P0 animals expressing the transgene (GFP+) were picked on each plate and were left at room temperature for 4-5 days until F1 progeny reached the gravid adult stage and GFP+ animals were scored. In total, 74 F1 WT-1 animals (n=6 P0), 78 F1 ST1-1 animals (n=8 P0), and 54 F1 ST2-1 animals (n=3 P0) were scored for the twitching phenotype. A two-tailed *t*-test was used to calculate significance.

2.4.2 Transgenics

Transgenic animals were obtained by microinjection of wild-type animals or *dcr-1* null background (BC4264) with DCR-1 constructs at 4ng/μl, mixed with pTG96 (sur-5::GFP) at 80ng/μl as a selectable marker to form extra-chromosomal arrays. Transgenes were integrated into the genome by UV irradiation followed by three outcrosses with wild-type animals to remove any additional mutations.

2.4.3 Sample Preparation

Pellets were homogenized in 50mM Tris-HCl pH8/150mM NaCl/1mM EDTA/0.1% Igepal with Complete EDTA-free protease inhibitors (Roche), and cleared by 17 000xg centrifugation.

2.4.4 IP and Western Blotting

IPs and western blots were performed on extracts prepared from embryos or young-gravid adults, unless otherwise indicated. Antibodies used were: rabbit polyclonals against DCR-1, RDE-4, ALG1/2, and ERI-1 (Duchaine et al., 2006); mouse monoclonals against ERI-5 (Thivierge et al., 2012), alpha tubulin (Abcam), c-MYC (Abcam), GFP (Roche), Phospho-(Ser/Thr) Kinase Substrate Antibody Sampler (Cell Signaling), and FLAG (Sigma). HRP-conjugated rabbit and mouse TrueBlots were used as secondary antibodies (eBioscience). Protein A Sepharose CL 4B (GE Healthcare) or Protein G Sepharose 4 Fast Flow (GE Healthcare) and ANTI-FLAG M2 (Sigma) beads were used according to the manufacturer's instructions for IP.

2.4.5 Plasmid Construction

DCR-1 genomic amplicons with 3'UTR were cloned into L2865 under the control of plet-858 (Addgene plasmid 1522) with N-terminal MYC and C-terminal 3XFLAG tags (Sawh

and Duchaine 2013). To generate multi-site ST cluster mutations, PCR was performed with end-to-end primers containing the desired mutations, PNK-treated and ligated. Correct mutations and frame was verified by sequencing. To generate ST1 (A) mutants, TDO2466 (AAC GCT GCT GCT GCT AAT ATT CCT CAA GCA TCT G) and TDO2468 (AGC CAC AGC TCT AGC ACG ACG TGG TTG ATT T) were used. To generate ST2 (D) mutants, TDO2467 (AAC GAT GAT GAC GAC AAT ATT CCT CAA GCA TCT G) and TDO2469 (GTC CAC ATC TCT ATC ACG ACG TGG TTG ATT T) were used.

2.4.6 Microscopy

Brightfield images of live animals were taken on a Zeiss SteREO Lumar.V12 microscope fitted with AxioCam HRc at 80X magnification. Images were then processed using AxioVision Rel 4.8 software.

2.4.7 CRISPR/Cas9 genome editing

Genome editing using a co-conversion strategy was designed according to (Ward, 2015). Two sgRNAs were designed to target DCR-1 near T961 and to have no off-target sites in known genes (crispr.mit.edu).

DCR-1 ssDNA repair templates were designed to mutate T961 to A or D (bold), and the PAM (red) silently to prevent re-targeting (italics). If this was not possible, as in target site 2, several silent mutations were made upstream of the PAM. Repair templates also included a silent mutation to the XbaI restriction site (underlined) to enable screening.

sgRNA1, score=96 (crispr.mit.edu): CAACGTCAAATATTCCTCAA

repair template 1 T961A:
cttcaacctcgaattcaaaatcaaccacgtcgttctaga**G**ctgtgagtaactC*GT*TCAACGTCAAATATTCCTCAAgcatc

repair template 1 T961D:
cttcaacctcgaattcaaaatcaaccacgtcgttctaga**GA**tgtgagtaactC*GT*CAACGTCAAATATTCCTCAAgcatc

sgRNA2, score=98: CGTCGTTCTAGAACTGTGAG

repair template 2 T961A:
cttcaacctcgaattcaaaatcaaCCACG*a*CG*a*TCTAGA**G**CTGTGAGtaactcctcaacgtcaaatattcctcaagcatc

repair template 2 T961D:
cttcaacctcgaattcaaaatcaaCCACG*a*CG*a*TCTAGA**GA**TGTGAGtaactcctcaacgtcaaatattcctcaagcatc

2.5 Acknowledgements

We thank the members of the Duchaine lab for helpful and supportive discussions. We are also grateful to Maxwell Shafer and Katherine Stewart for microscopy assistance. This work was supported by the Canadian Institute of Health Research (CIHR) MOP123352, the Human Frontiers Science Program, and the Fonds de la Rercherche en Santé du Québec (FRQS), Chercheur-Boursier Salary Award J2 to (T.F.D.). A.L. is supported by the Faculty of Medicine NSERC Undergraduate Student Research Award and the Rose Mamelak Johnstone Research Bursary. A.N.S. is supported by the CIHR Frederick Banting and Charles Best Canada Graduate Scholarship, the Defi Corporatif Canderel Award, the Alexander McFee Memorial Fellowship, and the Ruth and Alex Dworkin Scholarship.

Preface to Chapter 3

Previous work from our group and others have identified a class of endogenous siRNAs in *C. elegans* called the ERI endo-siRNAs with largely unknown functions, and the complex of proteins required for their biogenesis, the ERI complex (ERIC), which includes DCR-1. Both their biogenesis and downstream effector steps are largely uncharacterized, and they are predicted to target a wide variety of gene products. Recently, work from the Kennedy lab unearthed novel nuclear functions for some RNAi components in *C. elegans*, and notably characterized the nuclear Argonaute NRDE-3. NRDE-3 is loaded with ERI endo-siRNAs and shuttles to and from the nucleus. In the next two chapters we build upon this work and attempt to elucidate the mechanism of nuclear RNAi. In Chapter 3, we hypothesized that the ERI endo-siRNAs direct transcriptional gene silencing of their targets and sought to develop a nuclear run-on assay to allow us to directly test this hypothesis. Additionally, we sought to establish a platform for future genome-wide transcriptome analysis in *C. elegans*.

Chapter 3: Mechanistic insights into RNAi-induced transcriptional gene regulation

Sawh A. N. and Duchaine T. F., Nuclear run-on and PRO-seq in *C. elegans* embryos reveal the impact of endogenous RNAi pathways on the transcriptional landscape. (manuscript in preparation)

3.1 Introduction

Of all the DCR-1-dependent RNAi pathways in *C. elegans*, the mechanisms of small RNA biogenesis and target silencing in the ERI endoRNAi pathway remain the least well elucidated. This pathway is important for fertility and embryogenesis (Conine et al., 2010; Conine et al., 2013; Duchaine et al., 2006; Pavelec et al., 2009), and is characterized by a specific subset of endo-siRNAs that are mono-phosphorylated, 26nt long and display a bias for guanosine in the first position (26Gs or primary endo-siRNAs) (Batista et al., 2008; Ruby et al., 2006). Endo-siRNAs are predicted to target hundreds of loci, both protein-coding and non-coding (Ambros et al., 2003; Gu et al., 2009; Ruby et al., 2006). The biogenesis of 26Gs is unclear, but genetically requires DCR-1 and the ERI proteins, which form a complex with DCR-1. The ERI complex includes the SAP-exonuclease ERI-1b, the novel protein ERI-3, the RNA-dependent RNA polymerase (RdRP) RRF-3, tandem Tudor domain protein ERI-5, and the Dicer-related helicase DRH-3 (Duchaine et al., 2006; Gent et al., 2010; Pavelec et al., 2009; Welker et al., 2010). The ERI complex consists of two modules that interact with the

helicase domain of DCR-1: the RdRP module made up of RRF-3, ERI-5 and DRH-3, and the ERI-1-ERI-3 module (Thivierge et al., 2012).

The primary endo-siRNAs generated by the ERI complex are loaded onto specific primary AGOs: ALG-3/4 in sperm and ERGO-1 in embryos (Conine et al., 2010; Vasale et al., 2010). The silencing signal is then amplified to produce an abundant class of secondary endo-siRNAs (22Gs) in a process that genetically requires the RdRPs EGO-1 and RRF-1, and is independent of DCR-1. Secondary endo-siRNAs are exclusively antisense to target mRNA sequences, indicating that the RdRPs use the mRNAs targeted by primary endo-siRNAs as a template to amplify the pool of secondary siRNAs (Aoki et al., 2007; Pak and Fire, 2007; Sijen et al., 2001). Secondary endo-siRNAs are loaded onto AGOs of the WAGO clade, some of which (NRDE-3 and HRDE-1) have been recently shown to have nuclear localization (Buckley et al., 2012; Guang et al., 2008). The steady state mRNA levels of some predicted targets increase in *eri* and *nrde-3* mutants (Duchaine et al., 2006; Gent et al., 2010; Zhuang et al., 2013),

but whether gene silencing is a result of transcriptional or post-transcriptional regulation, and the extent to which silencing is mediated by NRDE-3 remain unknown.

To definitively answer this question, we decided to employ an assay that would directly measure target gene transcription instead of RNA steady state or proxy readouts like histone modification and polymerase occupancy, and use it to compare wild-type and *eri* mutants. Total RNA sequencing or quantitative RT-PCR of specific genes measures the steady state level of RNA species, which is influenced by transcription as well as RNA turnover. Specific histone tail modification patterns can be correlated with the transcriptional status of a genomic region in many cases, but not all, and important contradictions exist in the literature (Berger, 2007). It is for these reasons that we turned to nuclear run-on. The nuclear run-on assay essentially freezes polymerases in place within intact nuclei, and then pulses them with labeled ribonucleoside triphosphates, which are incorporated into nascent transcripts. Since transcription initiation is blocked in this process, the detection of the specifically labeled RNAs provides a measure of the amount of transcripts being produced from a specific genomic region (Derman et al.,

1981; Greenberg and Ziff, 1984; Powell et al., 1984; Smale, 2009; Ucker and Yamamoto, 1984).

We designed and optimized a robust nuclear run-on assay in *C. elegans* embryos which directly measures nascent RNAs. In order to assess the role of the ERI endoRNAi pathway in the transcription of target genes, we compared the transcription in wild-type and *eri* mutant embryos. Target gene transcription was increased in *eri* mutants, demonstrating that the ERI endoRNAi pathway transcriptionally inhibits its targets. Using the run-on assay, we further determined that triggering knock-down of a gene by exogenous dsRNA (exoRNAi) leads to an increase in the transcription of the target. These studies provide a basis to expand our understanding of the *C. elegans* embryonic transcriptome.

3.2 Results

3.2.1 Development of nuclear run-on in *C. elegans*

We set out to directly test the effect of the ERI endoRNAi pathway on the transcription of target genes by nuclear run-on assay. In order to sensitively detect changes in transcription, we opted to develop the assay using embryos, the most homogeneous

population of cells that we could obtain in large quantities from the animal. Typically, nuclear run-on assay is performed using purified nuclei (Smale, 2009). However, early attempts determined that the process of purifying nuclei from *C. elegans* embryos resulted in breakage of the nuclear membrane and significant loss of the organelle's nucleoplasmic contents (Figure 3.1).

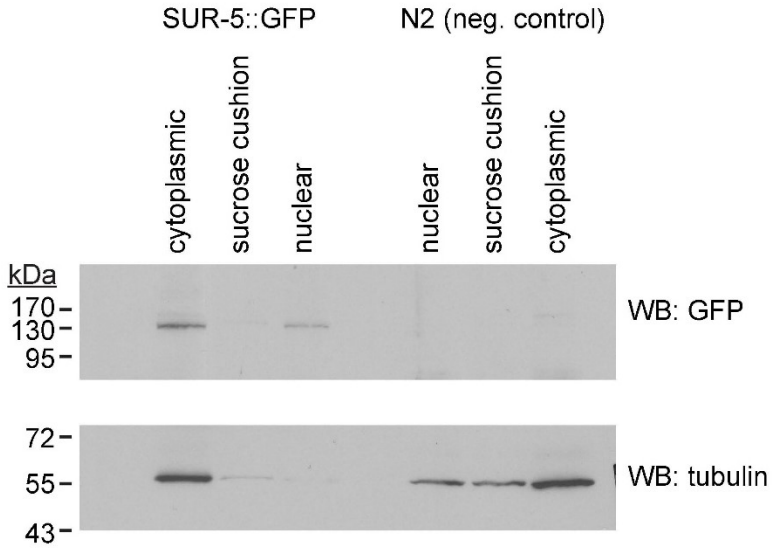


Figure 3.1: Nuclei of *C. elegans* embryos lyse during cellular fractionation. Embryos of wild-type animals (N2) or those expressing a GFP-tagged protein localized to the nucleoplasm (SUR-5::GFP) were homogenized and loaded onto a sucrose cushion as in (So and Rosbash, 1997). Following centrifugation, the cytoplasmic, cushion and nuclear fractions were separated and analyzed by western blot. GFP signal was used as a marker to judge the integrity of the purified nuclei. GFP presence in the

cytoplasmic fraction indicated nuclei were lysed during this procedure. Tubulin is used as a cytoplasmic marker.

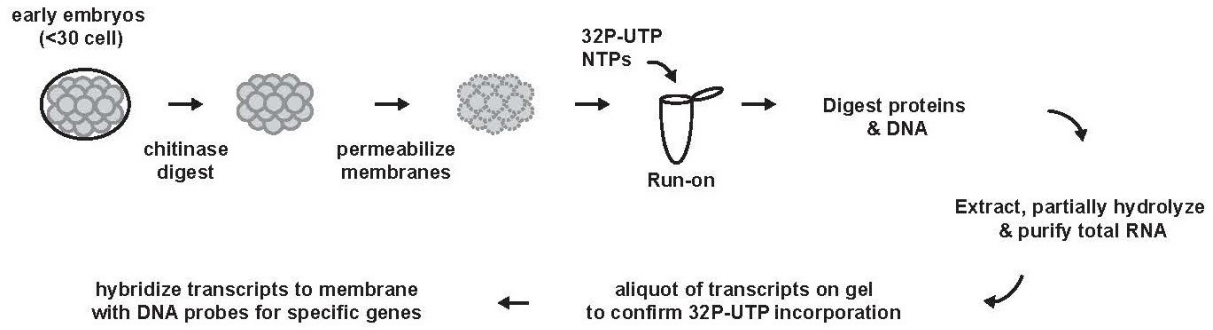


Figure 3.2: Schematic representation of 32P-UTP nuclear run on in *C. elegans* embryos.

Subsequently, we optimized the run-on reaction conditions based on an earlier method that uses whole, permeabilized embryos ((Schauer and Wood, 1990), Figure 3.2). By monitoring the incorporation of labeled nucleotides into RNA transcripts by gel electrophoresis, we demonstrated that early-stage embryos (most harvested in utero from hermaphrodite animals, generally categorized as <30 cell) could uptake and use

exogenous nucleotides more effectively than late-stage embryos (most harvested after embryos are ejected from P0 animals) (Figure 3.3).

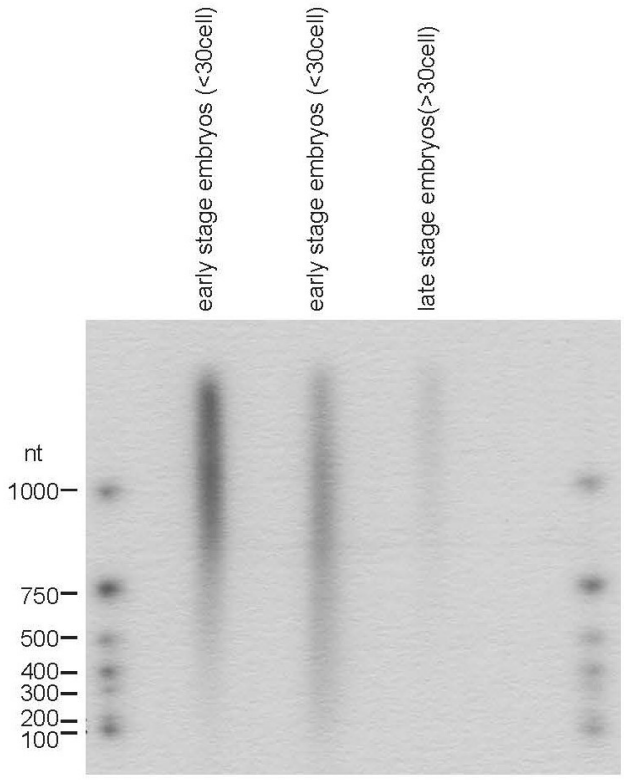


Figure 3.3: Early embryos display optimal incorporation of exogenous labeled nucleotides. Aliquots of total RNA extracted from representative biologically independent run-on reactions were run on a denaturing gel, followed by autoradiography to detect and compare 32P-UTP incorporation in newly formed transcripts.

3.2.2 Assessing the effect of the ERI endoRNAi pathway on target transcription

Next, to detect changes in transcription of specific genes, we hybridized the total RNA from the run-on reaction to membranes cross-linked with denatured DNA probes to ERI target genes and housekeeping genes (Figures 3.4).

Performing this experiment using both wild-type embryos and *eri-3* null mutant embryos then allowed us to assess the effect of loss of the ERI endoRNAi pathway function on transcription. *eri* mutants display temperature-sensitive sterility above 20°C, and thus all strains were grown at the permissive temperature 16°C. We chose to analyze three previously characterized ERI endoRNAi target genes: K02E2.6, X-cluster and E01G4.5. In ERI mutants, siRNAs mapping to these genes are lost, and their steady state mRNA abundance increases (Duchaine et al., 2006; Lee et al., 2006). Following run-on, signals for these ERI target genes were normalized to housekeeping genes within each extract, then these normalized values are compared between wild-type and *eri-3* mutant backgrounds.

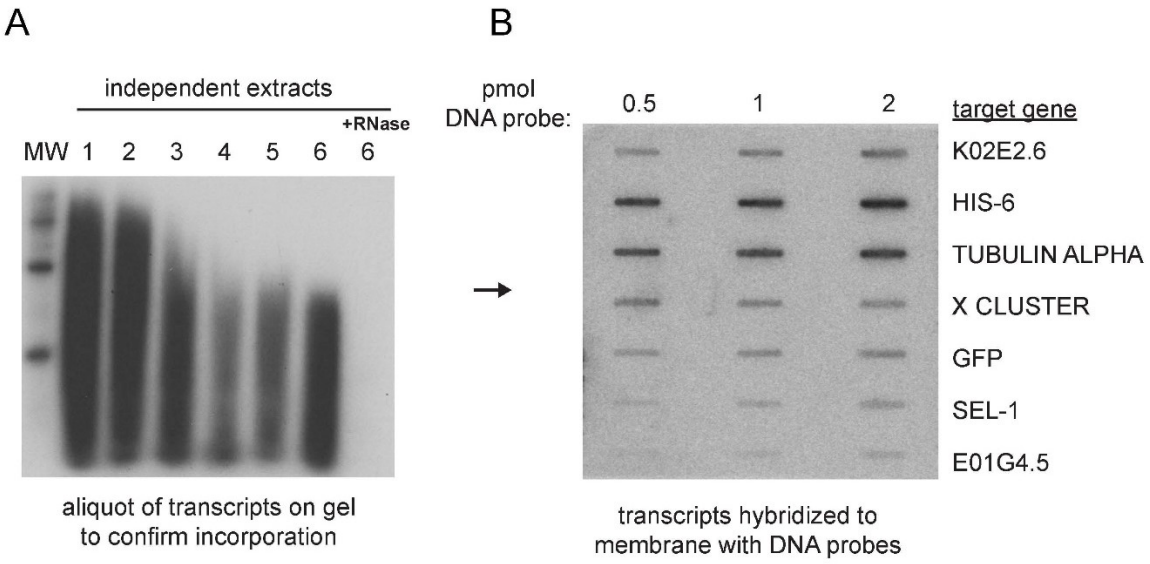


Figure 3.4: Hybridization of total RNA to DNA probes detects transcripts of individual genes. (A) Autoradiography of 4% denaturing polyacrylamide gel of nascent RNAs incorporated with 32P-UTP. A sample treated with RNase A is loaded to confirm 32P-UTP is incorporated into longer RNA. (B) Representative blot of RNA from an independent extract in (A) hybridized with a positively charged nylon membrane slot-blotted and cross-linked with DNA probes for specific genes of interest (ERI targets) and housekeeping genes. K02E2.6, X-cluster and E01G4.5 are representative ERI endoRNAi targets and HIS-6, Tubulin Alpha, GFP and SEL-1 are representative housekeeping genes. Experiments were carried out in wild-type or *eri-3* mutants, both carrying the translational fusion SUR-5::GFP transgene.

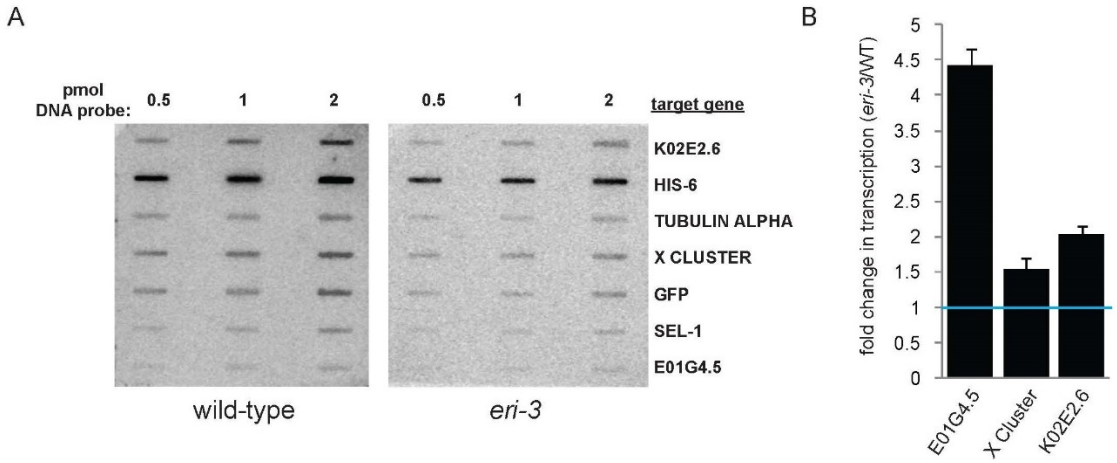


Figure 3.5: The ERI endoRNAi pathway transcriptionally inhibits its target genes. Intensity of signals for ERI targets (A) is quantified using Image J and normalized to signals for housekeeping genes (B). The values obtained from run-on in the *eri-3 -/-* extracts are presented relative to those from wild-type extracts (set to 1). Data are plotted as mean ± standard deviation. n > 3 biological replicates.

Our analysis of transcription on these genes further show that in ERI mutants, transcription is increased relative to wild-type levels (Figure 3.5). Transcription on the X-cluster locus increased by 1.5-fold, K02E2.6 increased 2-fold, and E01G4.5 increased 4.3-fold. These data therefore demonstrate a transcriptionally repressive role for the ERI endoRNAi pathway in the regulation of target genes.

3.2.3 Assessing the effect of the exoRNAi pathway on target transcription

We next asked whether triggering exoRNAi could induce a transcriptional response on the target gene. For this, we chose as a target the *lin-15* locus, which encodes the *lin-15B* and *lin-15A* pre-mRNAs sequentially on a single polycistron, before they are separated by trans-splicing (Clark et al., 1994; Huang et al., 1994). We designed run-on probes in order to discriminate between different modes of transcription along the locus: initiation, elongation and termination. exoRNAi was triggered by feeding bacteria expressing dsRNA against a region in *lin-15B* in P0 animals and the F1 generation was harvested for analysis. Nuclear run-on was then performed and transcription along the *lin-15B* and *lin-15A* sequence was measured with a probe upstream to the dsRNA trigger, and four probes downstream of the trigger (Figure 3.6). The data for each *lin-15* probe was normalized internally to *his-6* (as a representative housekeeping gene), and then the fold change between +RNAi and no RNAi samples was calculated (Figure 3.6C). Despite some variation between biological replicates, the data demonstrate that transcription of the exoRNAi target locus *lin-15* is not inhibited following knock-down. Instead, transcription is slightly increased (1.5-2-fold) in regions before and after the

trigger dsRNA, and remained increased downstream throughout the *lin-15A* sequence. These data suggest that exoRNAi may trigger a feedback increase in target transcription.

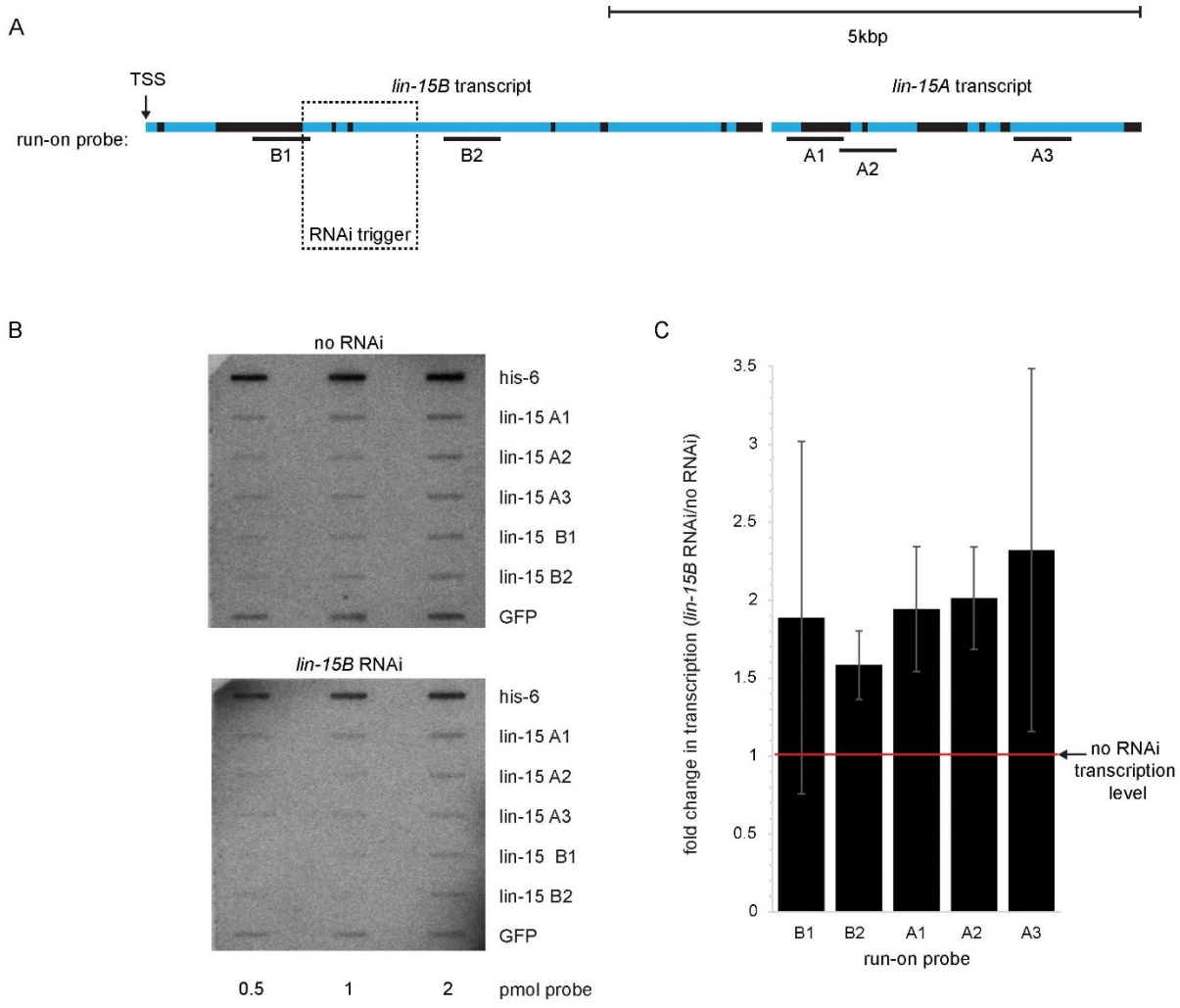


Figure 3.6: Triggering exoRNAi against *lin-15* for a single generation results in transcriptional induction of the target locus. exoRNAi against the *lin-15* locus (A) was triggered by feeding, and transcription across the *lin-15B* and *lin-15A* polycistron was

measured by nuclear run on. Schematic of the *lin-15* transcripts are shown to scale, with the position of the transcription start site (TSS), exons (blue), RNAi trigger sequence (dotted box) and run-on probes indicated. Signals obtained in (B) were normalized internally to *his-6*, and the fold change between +RNAi/no RNAi is plotted in (C). Experiments were carried out in *eri-1* mutants carrying an embryonic GFP transgene (Ppes-10::GFP). Data are plotted as mean ± standard deviation. n= 3 biological replicates.

The results of the experiments described in Figure 3.6 directly contradict the findings of another study investigating the function of genes involved in nuclear RNAi. Guang et al. (2010) also triggered exoRNAi against *lin-15* and subsequently monitored transcription on the locus. Instead of directly measuring the transcripts (as in Figure 3.6), Guang et al. incorporated Br-UTP into nascent RNAs, immunoprecipitated RNA with an anti-BrdU antibody, and quantified the recovered RNA using qRT-PCR. Their experiments yielded a decrease in transcription downstream of the dsRNA trigger, leading them to propose a model whereby RNAi directs the silencing of a target gene during the elongation phase of transcription (Guang et al., 2010). We aim to address this contradiction and

definitively conclude what effect exoRNAi has on transcription. Since our 32P-UTP run-on showed variability between biological replicates, and to greatly improve the resolution of the measurements, we decided to adapt our run-on protocol to globally and sensitively assess transcriptional changes using RNA sequencing.

3.2.4 Development of PRO-seq in *C. elegans*

We decided expand and refine the transcriptional changes seen in endoRNAi mutants (Figure 3.5). To this end, we adapted a recently described run-on method that incorporates biotin-labeled nucleotides into nascent transcripts followed by affinity purification with streptavidin-coated beads, library preparation and deep sequencing called Precise Run-on sequencing (PRO-seq) (Kwak et al., 2013). The incorporation of biotin-labeled nucleotides into transcripts via this method leads to a termination of the transcription reaction. This feature allows precise mapping within gene bodies of the site of transcription, and will allow us to differentiate between initiation and elongation modes of silencing. We modified the method by substituting our previously optimized *C. elegans* run-on reaction conditions in place of the published protocol (Fig. 3.7).

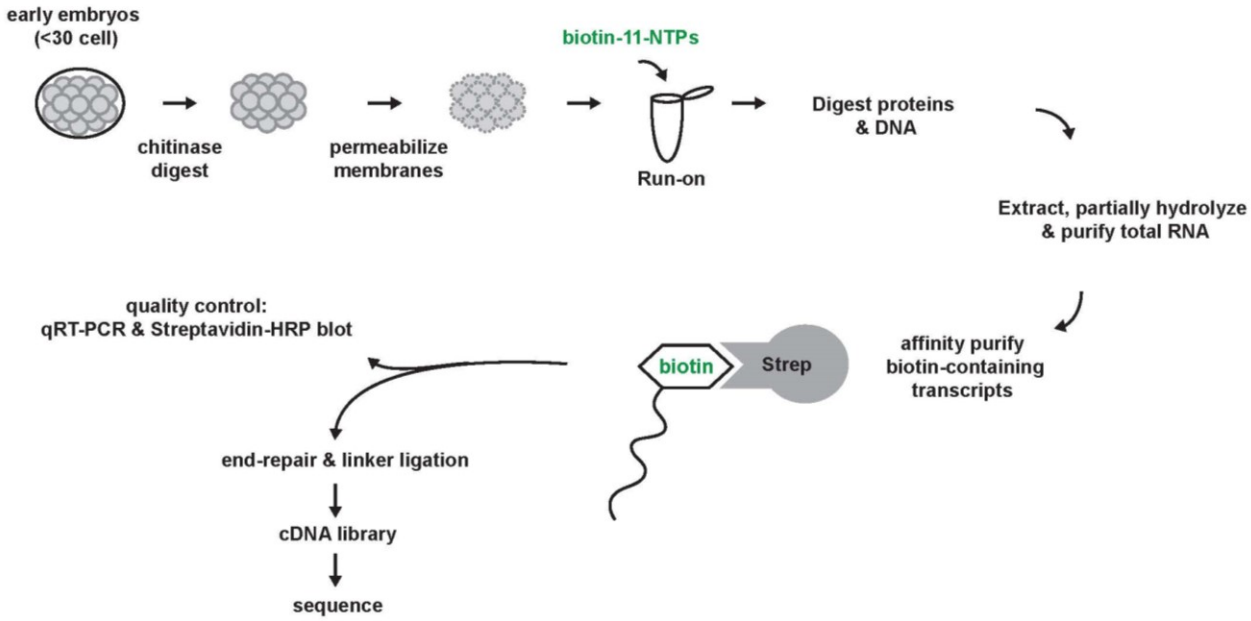


Figure 3.7: PRO-seq in *C. elegans* embryos. Schematic representation of PRO-seq protocol. The nuclear run-on is modified to perform PRO-seq, incorporating biotin-11-NTPs instead of 32P-UTP to track nascent transcripts.

Since the original protocol did not incorporate quality control steps to verify labeled RNAs were efficiently purified (Kwak et al., 2013), we further developed a method to track the presence of labeled transcripts throughout purification steps (Figure 3.8). To do this, RNA fractions recovered at steps following run-on and throughout biotin affinity purification were slot-blotted and cross-linked onto a nylon membrane, and then detected with a streptavidin-HRP antibody (Figure 3.8).

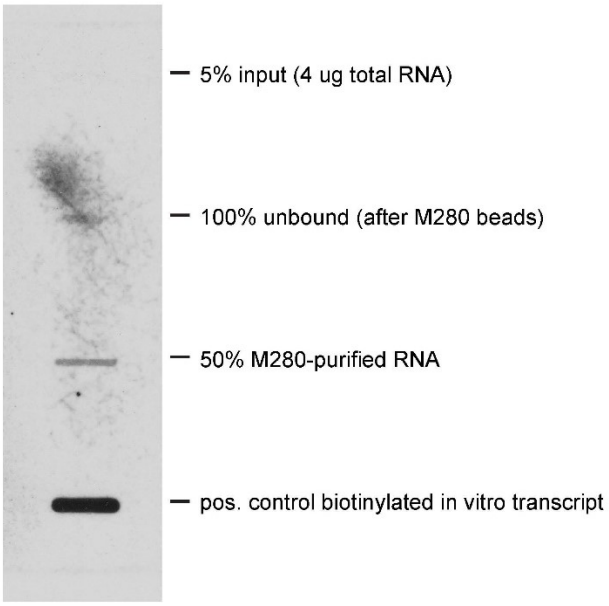


Figure 3.8: Incorporation of biotinylated nucleotides and purification of labeled transcripts. Biotin-labeled RNAs are purified from total RNA using Streptavidin M280 beads. Recovered fractions are then slot-blotted onto nylon membrane and detected with Streptavidin-HRP to confirm incorporation of biotin-labeled nucleotides into RNAs, and their presence in final samples. An unrelated RNA produced by *in vitro* transcription in the presence of biotinylated nucleotides is used as the positive control.

Future directions for this project will include sequencing and analysis of libraries made from wild-type and multiple *eri* genetic backgrounds to generate transcriptome profiles

and to further dissect the mechanism of ERI endoRNAi-mediated transcriptional silencing. We will also perform the precise run-on analysis on genes that have been targeted by exoRNAi, to resolve the conflicting data between our study and another (Guang et al., 2010). We will measure the effect on target gene transcription following a single generation of exoRNAi (as in Figure 3.6 and (Guang et al., 2010)), and after multiple generations of knockdown, in order to temporally define the transcriptional response.

3.3 Discussion

Previous work has implied that the exoRNAi pathway induces transcriptional gene silencing of its target genes (Guang et al., 2010; Vastenhouw et al., 2006). The mechanism of target silencing downstream of the ERI endoRNAi pathway, however, remained elusive. We have developed and optimized a robust nuclear run-on in *C. elegans* embryos to directly measure transcription of endogenous genes in order to address this matter. Using the run-on, we were able to demonstrate for the first time that the ERI endoRNAi pathway transcriptionally inhibits its target genes. In this manner, the ERI pathway can be likened to the sole RNAi pathway in the fission yeast *S. pombe*. In

S. pombe, the RNAi pathway targets heterochromatin at peri-centromeric regions and some coding loci. Heterochromatin regions are marked by histone H3K9 methylation, which is recognized by the chromodomain protein HP1/SWI6, to establish and maintains its heterochromatin state (Volpe et al., 2002). When target regions produce leaky transcripts, an AGO complex known as the RNA-induced transcriptional silencing (RITS) complex (Buker et al., 2007; Verdel et al., 2004) base-pairs with the leaky transcript through the siRNA, and recruits an RdRP complex (RDRC) to generate dsRNA (Colmenares et al., 2007). *S. pombe* Dicer uses this dsRNA as a substrate and generates more siRNAs specifically targeting the leaking locus (Motamedi et al., 2004). At the transcription site, the RITS recruits the histone methyltransferase Clr4, the catalytic subunit of the CLRC complex, to the target locus to specifically methylate more H3K9 sites (Zhang et al., 2008). The RITS complex is not only responsible for directing histone marks by recruiting CLRC, it also displays affinity for the accumulated H3K9Me modifications through the chromodomain protein Chp1. The dual interaction of the RITS and heterochromatin loci through both siRNAs and Chp1 allows continuous association

with H3K9me marks, to survey heterochromatin domains for transcription. This mechanism thus acts as a "self-enforcing loop" to ensure heterochromatin maintenance in *S. pombe.* In *C. elegans*, the ERI endo-siRNAs are loaded onto the Argonaute NRDE-3, which then translocates to the nucleus (Guang et al., 2008). We have shown that the ERI pathway ultimately transcriptionally inhibits its target genes, but the sequence of events between NRDE-3 nuclear translocation and transcriptional silencing of *eri* targets are unclear. Proteomic studies of the ERI complex (ERIC) components have not identified nuclear protein partners (Thivierge et al., 2012), so we do not predict that the ERIC adopts nuclear localization. We hypothesize that NRDE-3 locates targets similarly to the RITS and proceeds to recruit as yet unknown chromatin modifying factors and induce transcriptional silencing. In order to elucidate this mechanism, a proteomic study of NRDE-3 will be informative (See Chapter 4).

In our studies, we surprisingly found that knock-down of an endogenous mRNA (*lin-15*) causes an increase in its transcription all along the gene body, not an inhibition downstream of the trigger dsRNA as claimed in another report using a qRT-PCR-based

method (Guang et al., 2010). A substantial advantage of 32P-UTP run on that we have developed is that the intensity of signal generated by nascent transcripts is measured directly by autoradiography. At no point is the RNA generated in the embryos subject to possible amplification biases introduced by qRT-PCR or deep sequencing protocols (Aird et al., 2011; Roberts et al., 2011). Since transcription of the target gene was increased 1.5-2-fold both upstream and downstream of the trigger, we conclude that exoRNAi targeting leads to an increase in transcription initiation of the *lin-15* locus. However, is a feedback response such as the one seen with *lin-15* a general response following exoRNAi targeting? It would be surprising if this were true for all genes. It will be very interesting to test whether qualitatively dissimilar genes behave likewise. For example, one may predict that in order to compensate for the loss of its mRNA, knock-down of an essential gene would induce a stronger transcriptional upregulation than a non-essential gene. Along the same lines, if a gene is only required at a specific stage in development, its knockdown at another stage may be tolerated and result in a weaker transcriptional upregulation. The *lin-15* gene is important for the determination of vulval

cell fate (Clark et al., 1994; Huang et al., 1994), but its role in the embryo is unknown. The relative abundance and specific expression domain may also play a role in the transcriptional response of different genes to exoRNAi. Therefore, it will be interesting to test the transcriptional response of a battery of differentially expressed, essential, non-essential and redundant genes following knockdown using the nuclear run-on assay developed in this study.

Earlier studies in *C. elegans* have shown that the phenotype associated with knockdown of a GFP transgene, and other differentially expressed genes, can be inherited and persist in the progeny for many generations, implicating regulatory changes at the transcriptional level (Vastenhouw et al., 2006). It should be noted however, that progeny displaying the knockdown phenotype were selected in these cases, and the total level of knockdown in the F1 generation was not assessed. Thus, the inheritance of silencing triggered by exoRNAi may be a result of such sufficiently rare events that they escape detection when considering an entire population. When assessing target transcription of the ERI endoRNAi pathway, we employed stable

genotypes of the wild-type and *eri* null mutant backgrounds, as opposed to single-generation knockdown, and determined that ERI targets are transcriptionally repressed. Considering this, it is possible that one generation of knockdown is insufficient to trigger transcriptional inhibition on a specific gene, and that the feedback response we have observed for *lin-15* precedes repression in later generations. One intriguing possibility is that a high level of transcription may compete with nuclear RNAi processes for control of a locus. To test this, the selective pressure applied by exoRNAi can be maintained for many generations, and the nuclear run on can be used to test if at any point knockdown induces transcriptional repression.

Finally, while our aim is to expand the nuclear run-on to a global scale using the PRO-seq technology originally developed by the Lis lab, we affirm that any transcriptional changes identified in such analyses should be further validated by 32P-UTP run on. Our overarching goal is to establish reference transcriptomes for wild-type and *eri* mutant embryos at different stages of development that will be of use to the *C. elegans* community at large.

3.4 Materials and methods

3.4.1 *C. elegans* Strains and RNAi Assays

All strains were cultured as in (Brenner 1974). N2 was used as the wild-type strain. Alleles used were *eri-3(mg366)* and *eri-1(tm1361)*. RNAi was performed as in (Fire et al., 1998; Timmons et al., 2001). *eri* mutants were grown at 16ºC due to temperature sensitive sterility seen above 20ºC.

3.4.2 Preparation of gene specific probes on membranes

dsDNA probes were immobilized on Hybond XL membranes by slot-blotting. Purified PCR products ~500bp in length were denatured with 0.1 volumes of 3M NaOH at 60ºC for 1h, quenched on ice, diluted with 1 volume 6X SSC, and blotted at amounts: 0.5pmol, 1pmol & 2pmol for each transcript of interest. The membrane is allowed to dry on filter paper then UV cross-linked 2X.

3.4.3 Preparation of embryonic extract

Chitinase digestion:

200ul of freshly harvested embryos were added to 1ml of chitinase buffer (5mM HEPES pH8/110mM NaCl/4mM KCl/5mM MgCl2/1U/ml Chitinase) and incubated at room

temperature rotating for 15min. Digestion was monitored by microscope, by observation for for grape bunch-like appearance and “hatched” larvae. The mixture is centifuged and washed 2x with ice cold embryo buffer (5mM HEPES pH8/110mM NaCl/4mM KCl/5mM MgCl2), and resuspended in 2 volumes of nuclear storage buffer (20mM Tris-HCl, pH8/75mM NaCl/0.5mM EDTA/1mM DTT/0.125mM PMSF/50% glycerol). The extract was frozen in liquid nitrogen and stored at -80ºC.

Nuclei purification:

Embryos were defrosted on ice and homogenized in 2ml homogenization buffer (10mM HEPES-KOH, pH7.4/10mM KCl/0.8M sucrose/0.1mM EDTA/0.5mM PMSF/1mM DTT) using a stainless steel homogenizer. The homogenate was then loaded onto 3ml sucrose cushion (10mM HEPES-KOH, pH7.4/10mM KCl/1M sucrose/0.1mM EDTA/2mM MgOAc/3mM CaCl2/10% glycerol/ 1mM DTT/Complete EDTA-free protease inhibitor tablet (Roche)), and centrifuged for 10min in SW 55Ti rotor @12000rpm (Beckman L-90K Ultracentrifuge). Cytoplasmic and cushion fractions were

removed and the nuclear pellet was resuspended in nuclear resuspension buffer (50mM HEPES-KOH, pH7.4/0.1mM EDTA/5mM MgCl2/40% glycerol).

3.4.4 Nuclear Run-on

The DNA blots were pre-hybridized in a bottle with pre-heated 10ml Ambion Ultrahyb buffer for at least 1h at 42ºC. 32.4ul of nuclear extract was added to 10ul of 5X reaction buffer to obtain final concentrations: 100mM Tris-HCl, pH8/50mM NaCl/100mM KCl/0.1mM PMSF/1.2mM DTT/5mM $MgCl_2$/28mM ammonium sulfate/10mM creatine phosphate/30% glycerol/0.5% sarkosyl/2mM each ATP, CTP, GTP/10uM UTP/200u/ml RNasin. 7.6ul ^{32}P-UTP (~80uCi) was added to the reaction and incubated at room temperature for 1h. The run-on was terminated with 1 unit Turbo DNase, incubated at 37 º C for 15min, followed by 1 volume of termination buffer (20mM Tris-HCl, pH7.5/10mM EDTA/2% SDS/0.2mg/ml proteinase K/0.1mg/ml yeast tRNA) at 42ºC for 30min. RNA was extracted once with phenol:chloroform, and precipitated with GlycoBlue (Ambion). The pellet was resuspended pellet in 75ul RNase-free H2O and purified on Roche RNA G50 spin column. The RNA (100ul) was then partially hydrolyzed with 60ul 0.2M Na_2CO_3 and 40ul 0.2M $NaHCO_3$ and incubated at 60ºC for

1h, then neutralized w/ 6ul 3M NaAc pH5.2 and 10ul 10% Acetic acid. 5-10ul of RNA was run on 4% acrylamide/7M urea gel (in Ambion gel loading buffer II) in 0.5X TBE, and exposed to film overnight. The rest of the RNA was hybridized to the DNA blots overnight at 42ºC. The blots were washed blots 2x 5min at 42C in 2xSSC/0.1%SDS, 2x 15min at 42C in 0.1xSSC/0.1%SDS, dried and exposed to an imaging plate. Signals were quantified using a PhosphorImager and Image J.

3.4.5 PRO-seq in *C. elegans* embryos.

Run on was performed as in 3.4.4, except biotin-11-NTPs (Perkin Elmer) were substituted for NTPs to a final concentration of 50uM in the final reaction. Trizol LS was used to extract the RNA. Biotin-labeled RNAs were purified with Streptavidin-coated M280 magnetic beads (Invitrogen) as in (Kwak et al., 2013).

3.4.6 Detection of biotinylated transcripts.

Aliquots of biotinylated RNA before, during and after purification using Streptavidin-coated M280 magnetic beads (Invitrogen) were slot-blotted onto Zeta-probe membrane (Biorad) and UV crosslinked. The RNAs were then detected using a protocol adapted from the BrightStar BioDetect Kit (Ambion-discontinued). Membranes were washed 2x

5min in Wash Buffer (5X Wash buffer: 290mM Na2HPO4/ 85mM NaH2PO4/ 340mM NaCl/ 0.5% SDS), 2x 5min in Blocking Buffer (58mM Na2HPO4/ 17mM NaH2PO4/68mM NaCl/ 0.3% Casien/ 2% SDS), and incubated 30min in Blocking Buffer. Membranes were then probed with High sensitivity Streptavidin-HRP (Pierce) in Blocking Buffer (1/40 000) for 30min, washed 1x 10min in Blocking Buffer, 3x 5min in Wash Buffer, and 2x 2min in Assay Buffer (10X Assay buffer: 1M Tris-HCl pH9.5/ 1M NaCl). Signals were developed using Western Lighting Plus ECL (PerkinElmer).

3.5 Acknowledgements

We thank Leighton Core, Hojoong Kwak and John Lis for very helpful advice on PRO-seq. This work was supported by the Canadian Institute of Health Research (CIHR) MOP123352, the Human Frontiers Science Program, and the Fonds de la Rercherche en Santé du Québec (FRQS), Chercheur-Boursier Salary Award J2 to (T.F.D.). A.N.S. is supported by the CIHR Frederick Banting and Charles Best Canada Graduate Scholarship, the Defi Corporatif Canderel Award, the Alexander McFee Memorial Fellowship, and the Ruth and Alex Dworkin Scholarship.

Preface to Chapter 4

In Chapter 3, we determined that the ERI endoRNAi pathway transcriptionally silences its target genes through the development of nuclear run-on. This silencing however, is accomplished through unknown mechanisms. Building upon previous work that identified NRDE-3 as a nuclear Argonaute loaded with ERI endo-siRNAs, and genetic screens completed by the Kennedy, Plasterk and Miska groups on related nuclear processes, we sought to clarify this mechanism. We opted to use an unbiased and extensive proteomics-based approach to identify and refine the protein complexes involved in nuclear RNAi. Using such methods, we identified a list of high-confidence novel Argonaute interactors and proteins with known roles in chromatin modification or association. We further identified and characterized a novel multi-Argonaute interacting protein, shared by the ERI endoRNAi and exoRNAi pathways. Finally, we establish the basis to build a proteomic network of nuclear RNAi factors and further elucidate the mechanism of RNAi- induced transcriptional silencing in animals.

Chapter 4: Establishing a Proteomic Network of Nuclear RNAi Factors

Sawh A. N., Vashisht A., Flamand M., Lewis A., Wohlschlegel J., and Duchaine T. F., Comparative proteomics physically links nuclear RNAi factors with histone modifying machinery and novel RNAi components to mediate transcriptional gene silencing. (manuscript in preparation)

4.1 Introduction

The ultimate mechanisms of target silencing initiated by the ERI endoRNAi and the exoRNAi pathways have not been fully resolved. These two pathways share and compete for the WAGO clade of Argonautes which are loaded with secondary siRNAs and are responsible for the majority of target silencing. Since the WAGOs are incapable of slicer activity, endonucleolytic target cleavage by the Argonaute can be ruled out as the primary mode of silencing (Yigit et al., 2006). Instead, the WAGOs are thought to recruit accessory proteins which would then lead to mRNA turnover via post-transcriptional gene silencing (PTGS) or transcriptional gene silencing (TGS). Some WAGOs, like WAGO-1, display clear cytoplasmic localization and are thought to mediate PTGS (Gu et al., 2009; Tsai et al., 2015). On the other hand, several members of the WAGOs (NRDE-3 and HRDE-1) were recently found to localize to the nucleus and trigger H3K9 trimethylation and transcriptional gene silencing of target loci downstream of the exoRNAi and piRNA pathways (Ashe et al., 2012; Buckley et al.,

2012; Burton et al., 2011; Guang et al., 2008; Lee et al., 2012; Shirayama et al., 2012). This branch of RNAi was subsequently termed “nuclear RNAi”.

Several genetic screens have recently been completed to identify factors that participate in nuclear RNAi downstream of both exoRNAi and piRNA triggers. The first screen identified the *n*uclear *R*NAi *de*fective (*nrde*) mutants, including the somatic nuclear Argonaute NRDE-3 (Guang et al., 2008). NRDE-3 is loaded with secondary siRNAs from both the ERI endoRNAi and exoRNAi pathways in the cytoplasm, translocates to the nucleus, where it associates with nascent chains of pre-mRNAs and recruits at least one of the other NRDE proteins. Following exoRNAi induction, H3K9 trimethylation occurs on loci corresponding to the dsRNA trigger, in a process dependent on NRDE-3 (Burton et al., 2011; Gu et al., 2012b). Transcription of an exoRNAi target was suggested to be down-regulated at the step of RNA Polymerase II elongation (Guang et al., 2010). Similarly, another nuclear Argonaute (HRDE-1/WAGO-9) was found to be involved in TGS in the germline, where it is also required for the inheritance of silencing phenotypes triggered by the exoRNAi and piRNA pathways (Ashe et al., 2012; Buckley

et al., 2012; Burton et al., 2011; Lee et al., 2012; Shirayama et al., 2012). Furthermore, NRDE-3 and HRDE-1 share a genetic requirement for the NRDE-1, 2, and 4 proteins to direct nuclear RNAi. NRDE-1, 2, and 4 proteins associate with pre-mRNAs, and NRDE-1 also interacts with chromatin (Burkhart et al., 2011; Guang et al., 2010). When TGS is established through exoRNAi or piRNA pathways, the silenced state of the target locus can be inherited for many generations. This inheritance of silencing in early generations requires a protein which binds methylated H3K9 (HPL-2), and histone methyltransferases (HMTs) SET-25 and SET-32 (Ashe et al., 2012). An independent genetic screen conducted in animals that have silenced targets indefinitely following exoRNAi induction identified genes encoding a chromodomain protein (MRG-1), a class II histone deacetylase (HDA-4), the MYST-domain putative histone acetyltransferase (K03D10.3), and a chromatin-remodeling ATPase (ISW-1) as key to the maintenance of long-term silencing ((Vastenhouw et al., 2006), See Table 4.1).

Protein	Domains/Activity	Alleles	Transgenic Strains
HRDE-1	Nuclear Argonaute	*mj278,tm1200*	3FLAG::GFP::HRDE-1
NRDE-3	Nuclear Argonaute	*gg66,gg74,gg86,gg54,gg56*	3FLAG::GFP::NRDE-3

NRDE-2	DUF 1740, serine/arginine-rich domain, half-a-tetratricopeptide (HAT)-like domain, pre-mRNA interacting	*gg91, gg95*	3FLAG::GFP::NRDE-2
NRDE-1	Novel, chromatin interacting, pre-mRNA interacting	*gg88*	
NRDE-4	Novel, pre-mRNA interacting	*gg129,gg131,gg132, gg194*	
SET-25	Histone methyltransferase H3K9me3	*n5021*	mCherry::SET-25
SET-32	Putative Histone methyltransferase	*ok1457*	
MET-2	Histone methyltransferase H3K9me, H3K9me2	*n4256*	mCherry::MET-2
HPL-2	Heterochromatin protein 1 (HP1) ortholog	*tm1489*	
MRG-1	Chromodomain	*tm1227*	
HDA-4	Class II histone deacetylase	*ok518*	
ISW-1	Chromatin-remodeling ATPase	*n4066*	
MYS-2/ K03D10.3	MYST domain, putative histone deacetylase	*ok2429*	

Table 4.1: List of key proteins involved in nuclear and inheritable RNAi. Key proteins, their characterized domains/activity, along with related characterized alleles, and strains are indicated.

Importantly, NRDE-3 is loaded with secondary ERI endo-siRNAs in wild-type conditions. In situations where the ERI siRNAs are lost, as in *eri-1* mutants, NRDE-3 remains in the cytoplasm (Guang et al., 2008). Similar to *eri* mutants, *nrde-3* mutants display increased

steady state levels of ERI endoRNAi target mRNAs (Duchaine et al., 2006; Gent et al., 2010; Zhuang et al., 2013). Taken together with our evidence that transcription of ERI targets are increased in *eri* mutants (Chapter 3), these findings suggest that NRDE-3 directs transcriptional silencing of ERI endoRNAi targets.

Despite the great strides made by these recent genetic studies, the exact sequence of events that direct the Argonautes to enter the nucleus and initiate TGS remain unclear. Therefore, in order to biochemically elucidate the mechanisms of nuclear RNAi, we have undertaken a proteomics-based approach. We aim to establish a proteomic network of factors involved in these processes, through the use of multi-dimensional protein identification technology (MudPIT). MudPIT employs separation of tryptic peptides by 2-D liquid chromatography followed by tandem mass spectrometry, and is uniquely suited to the identification of proteins from complex mixtures (Washburn et al., 2001). We have used this strategy successfully in the past to define distinct DCR-1 complexes, refine the miRNA RISC, and refine the ERI endoRNAi complex (Duchaine et al., 2006; Thivierge et al., 2012; Wu et al., 2010). By applying this method to the

nuclear RNAi branch, we uncover the first physical interactions between the nuclear Argonaute NRDE-3 and chromatin-modifying factors, and identify a previously unknown pan-Argonaute interacting protein.

4.2 Results

4.2.1 A proteomic network of nuclear RNAi factors

We initiated the proteomics survey of nuclear RNAi factors with the Argonautes (NRDE-3 and HRDE-1), a conserved component genetically required for their function (NRDE-2), a histone methyl transferase implicated in nuclear RNAi (SET-25), and a histone methyl transferase known to act upstream of SET-25 (MET-2). The transgenic strains used were generated by microparticle bombardment (Praitis et al., 2001) to produce low copy numbers of the tagged proteins of interest. We next purified the nuclear RNAi factors under native conditions from extracts of their respective strains. Negative control mock-purifications were performed under identical conditions, but using extracts of wild-type animals (no transgene) or wild-type animals expressing an unrelated nuclear-localized GFP (SUR-5::GFP). Any protein identified in a negative control purification was regarded as a non-specific interaction and subsequently excluded from further

analysis. NRDE-3, NRDE-2, and HRDE-1 were tagged with GFP and 3XFLAG, while SET-25 and MET-2 were tagged with mCherry. Both tags were used in immunopurifications, but FLAG proteins were eluted with 3XFLAG peptide, while mCherry proteins were eluted with SDS-PAGE loading buffer. We determined the FLAG purifications of NRDE-3 and NRDE-2 to have sufficient purity and scale to be submitted to MudPIT (Figure 4.1). In the case of NRDE-3 and NRDE-2 purifications, bands on a silver stained gel corresponding in size to the tagged proteins were visible in the eluates. The same bands appear in the silver stained gel in the “beads” fraction, which represents the proteins eluted with SDS-PAGE loading buffer after FLAG elution. Western blot with a GFP antibody was used to monitor the purifications. These experiments demonstrate that elution with the FLAG peptide typically released at least 50% of the captured FLAG-tagged protein from the matrix. The advantage of this strategy, despite the loss of almost half the target protein, is that the peptide elution specifically disrupts the antibody – antigen interaction, leaving most non-specifically

bound proteins to remain on the beads. We then submitted such samples to MudPIT. Similar results were obtained for 3XFLAG::GFP::HRDE-1.

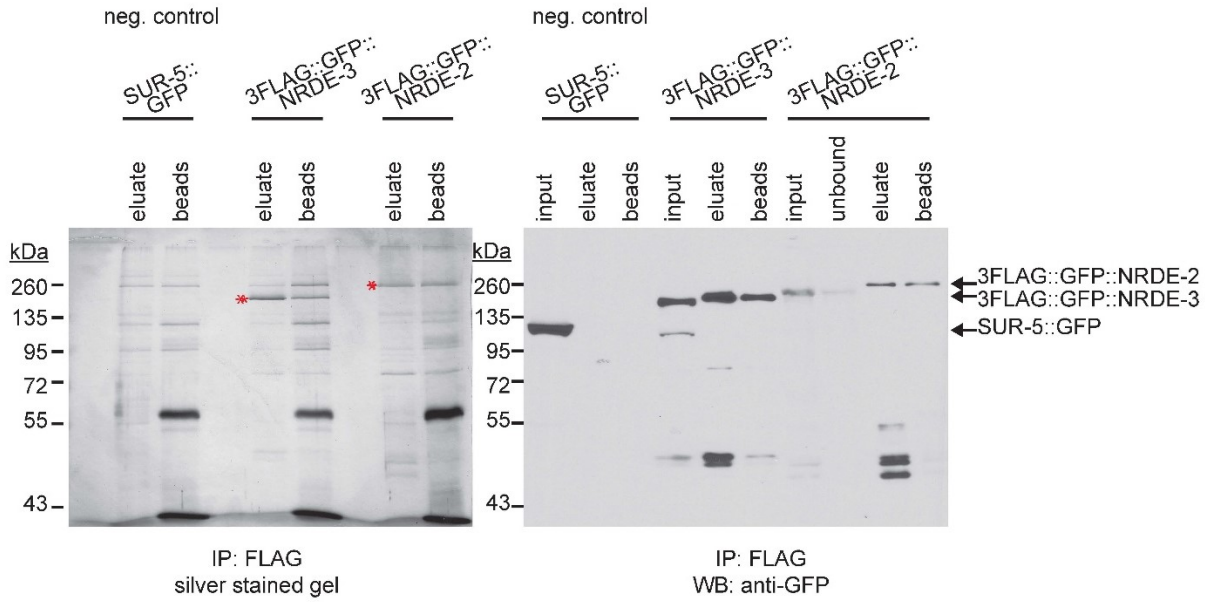


Figure 4.1: Purification of Nuclear RNAi factors NRDE-3 and NRDE-2. Representative gel and western blot from FLAG purifications performed on extracts from strains expressing tagged versions of NRDE-3 and NRDE-2 compared to a strain expressing an unrelated nuclear GFP as the negative control. Proteins were eluted with FLAG peptide and analyzed by silver stained gel and western blot to confirm efficacy of the purifications before submission to MudPIT. Red asterisks mark the bands corresponding to 3FLAG::GFP::NRDE-3 and 3FLAG::GFP::NRDE-2.

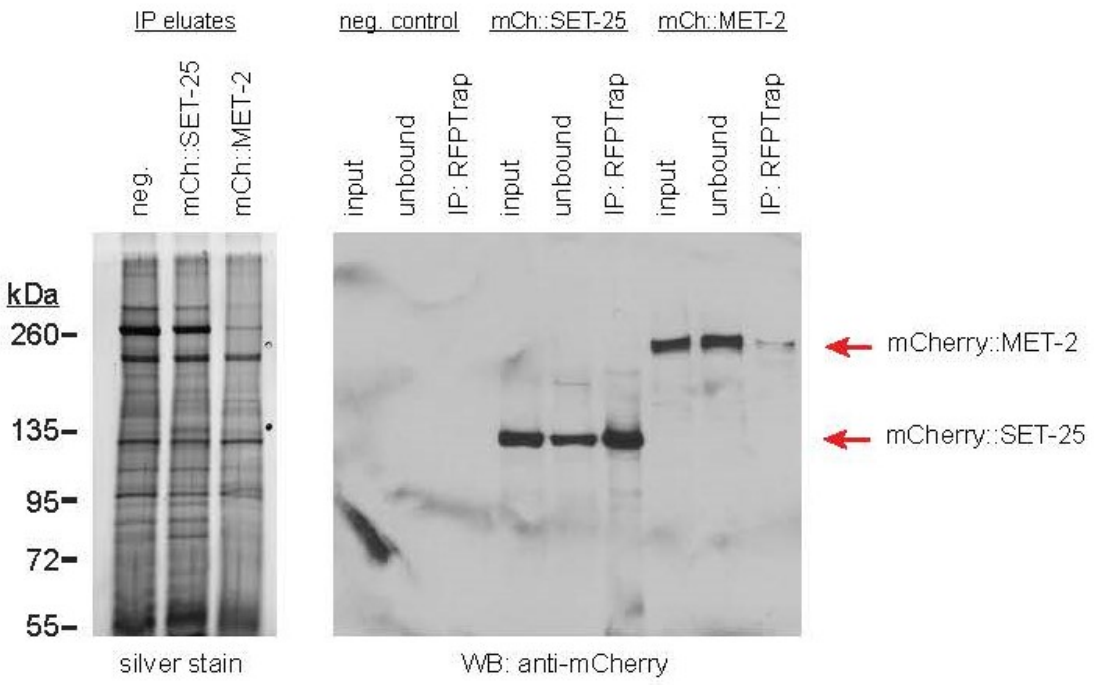


Figure 4.2: Purification of histone methyltransferases SET-25 and MET-2. Representative gel and western blot from RFPTrap IPs performed on extracts from strains expressing tagged versions of SET-25 or MET-2 compared to wild-type animals as the negative control. SET-25 IPs were effective, but of insufficient purity for MudPIT analysis, while the mCherry tag on MET-2 could not be immunoprecipitated efficiently.

In contrast, we did not succeed in achieving a similar level of purity in the mCherry purifications of SET-25 and MET-2 (Figure 4.2). The silver stained gel of the eluates showed many non-specific protein interactions in the negative control IP, and no clear band corresponding to tagged SET-25 or MET-2. A large amount of non-specific

interactions in the samples submitted to MudPIT could mask the specific interactions due to the abundance of peptides mapping to non-specific proteins, and this leads to loss of sensitivity. The western blot against mCherry did show mCherry::SET-25 in the IP fraction, but very minimal signal for mCherry::MET-2. This result implies that the region of mCherry recognized by the antibody is not readily accessible in the mCherry::MET-2 fusion, and therefore this strategy is unlikely to yield enough MET-2 target and interactors in the eluate for successful identification downstream. In the case of SET-25, more stringent washing conditions should be systematically applied to the IP in order to reduce the non-specific interactions.

Thus far, we submitted 5 biological replicate purifications of NRDE-3, NRDE-2 (1 biological replicate) and HRDE-1 (1 biological replicate) for MudPIT analysis. These samples were compared to 4 biological replicates of negative control purifications. The results of these experiments are summarized in Tables 4.2-4.6. Proteins whose peptides were present in any negative control purification were regarded as non-specific interactions, and specific interactions were ranked by the number of replicates they

appeared in. For NRDE-3 interactors (Table 4.2), we chose to analyze only proteins that appeared in at least two biological replicates to have greater confidence in the interactions. Such stringent criteria are not yet possible for NRDE-2 and HRDE-1. In total, we identified 44 novel protein interactors specific to NRDE-3, and 88 specific to HRDE-1 (Table 4.3). Furthermore, 21 interactors were shared between NRDE-3 and HRDE-1 (Table 4.4), 6 shared between NRDE-3 and NRDE-2 (Table 4.5), and 1 shared between NRDE-3, NRDE-2 and HRDE-1 (Figure 4.6). Shared interactors between NRDE-3, HRDE-1, and NRDE-2 could represent common upstream or downstream proteins involved in both somatic and germline nuclear RNAi, while interactors common to NRDE-3 and NRDE-2 alone could represent proteins exclusive to somatic nuclear RNAi. Within the groups of shared interactors, we will prioritize characterization of those found in multiple NRDE-3 datasets, which amounts to 11 interactors. None of these 11 have been previously linked to RNAi, and will be validated with multiple additional datasets for both HRDE-1 and NRDE-2. Finally, candidates will be functionally validated in nuclear silencing of target genes (See Discussion, Section 4.3).

Of note, in the NRDE-3 datasets, we identified chromodomain-containing protein MRG-1 in 2/5 biological replicates. Another chromatin-associated protein found in 2/5 NRDE-3 datasets is SIN-3, a member of histone deacetylase complexes implicated in transcriptional repression (See Discussion, Section 4.3). We also reproducibly identified several protein phosphatase components (PPFR-1, GFI-2, and PPH-4.1) in NRDE-3 datasets, indicating that NRDE-3 may be regulated by phosphorylation.

NRDE-3 INTERACTORS	*NRDE-3 (1)*	*NRDE-3 (2)*	*NRDE-3 (3)*	*NRDE-3 (4)*	*NRDE-3 (5)*	*HRDE-1*	*NRDE-2*	*% peptide coverage*	*Description*
F43C11.9	+	+	+	+	+			26 - 50	
FARS-1	+	+		+	+			4 - 16	predicted phenylalanyl-t-RNA synthetase
PPFR-1		+	+	+	+			1 - 3	protein phosphatase regulatory subunit
GFI-2	+	+		+				9 - 14	ortholog of human protein phosphatase 1
DYN-1	+	+		+				4 - 7	ortholog of the dynamin GTPase
Y54H5A.1	+	+			+			8 - 9	
ARI-1	+		+		+			6 - 14	ortholog of h.s. ariadne RBR E3 ubiquitin protein ligase 1
PPH-4.1			+	+	+			8 - 22	ortholog of human protein phosphatase 4, catalytic subunit
LGC-21			+	+	+			7 - 12	predicted extracellular ligand-gated ion channel activity

C01B10.11				+	+	14	
Y108G3AL.2				+	+	7	ortholog of human pre-mRNA processing factor 38B
MRG-1				+	+	9	chromodomain-containing protein
F33D4.6				+	+	6 - 8	ortholog of human neurofilament, heavy polypeptide
ZK1053.4				+	+	5	
SIN-3				+	+	2	ortholog of the SIN3 family of histone deacetylase subunits
CLEC-266			+	+		13 - 22	ortholog of human CD209 molecule
SKR-1		+			+	31 - 33	SKp1 related ubiquitin ligase complex component
B0495.8		+			+	8 - 9	ortholog of S. cerevisiae and human LUC7-like
STIP-1		+			+	4 - 11	ortholog of human tuftelin interacting protein 11
T24C2.2			+		+	14 - 31	
HIP-1			+		+	10 - 28	ortholog of human suppression of tumorigenicity 13
F37C4.5			+		+	4 - 17	
PAT-6			+		+	7	ortholog of alpha-parvin
C06A5.6			+		+	5 - 11	
AHCY-1			+		+	7	S-adenosylhomocysteine hydrolase (SAHH) ortholog
RSKS-1			+		+	5 - 7	putative ribosomal protein S6 kinase (S6K)
C33G3.6			+		+	3 - 4	ortholog of human dentin sialophosphoprotein
ZK688.5			+		+	2	
F30F8.9	+	+				17 - 18	ortholog of human LETM1 domain containing 1
Y54E10A.6	+	+				13 - 15	ortholog of human leucine rich repeat containing 47
DHOD-1	+	+				13 - 15	ortholog of human dihydroorotate dehydrogenase (quinone)
PSD-1	+	+				14	ortholog of human phosphatidylserine decarboxylase

C25H3.4	+	+						3 - 8	ortholog of human eukaryotic translation initiation factor 2D
C06G3.9	+	+						7 - 13	ortholog of human UFM1-specific ligase 1
DNJ-10	+	+						5 - 7	DNaJ domain (prokaryotic heat shock protein)
ULP-1	+	+						4	SUMO protease
Y55F3BR.1	+	+						6	DEAD-box helicase
WAGO-4	+	+						2 - 3	Argonaute
Y55F3AM.3	+		+					6 - 7	predicted to have nucleotide binding activity and RNA binding activity
PDI-3		+	+					11 - 13	protein disulfide isomerase (PDI)
FKB-7		+	+					8 - 10	peptidylprolyl cis/trans isomerase homologous to mammalian FK506
TNI-1	+			+				23	member of the troponin family
B0303.15		+		+				20	ortholog of human mitochondrial ribosomal protein L11
Y45F10D.7		+		+				2	WD40 repeat-containing protein, ortholog of human WDR36

Table 4.2: List of NRDE-3-specific interacting proteins.

HRDE-1 INTERACTORS	*NRDE-3 (1)*	*NRDE-3 (2)*	*NRDE-3 (3)*	*NRDE-3 (4)*	*NRDE-3 (5)*	*HRDE-1*	*NRDE-2*	*peptide coverage*	*Description*
ACP-6						+		31	ortholog of human acid phosphatase 2
F59B1.2						+		56.5	
F39H12.3						+		13.8	
ZC373.2						+		51.9	
LEV-11						+		22.9	Tropomyosin a/b/d/f, an actin-binding contractile structural protein
LEV-11						+		22.9	Tropomyosin d, an actin-binding contractile structural protein
ACT-5						+		35.7	ortholog of human cytoplasmic actin

CEY-3	+	36.6	cold-shock/Y-box domain
DHS-25	+	50	short-chain dehydrogenase predicted to be mitochondrial.
K08E3.10	+	34.6	
D1054.10	+	17	
D1054.11	+	16	
OXY-5	+	28.4	abnormal OXYgen sensitivity
F55H12.4	+	37	
W09C5.1	+	27.4	
PDE-6	+	20.5	3'-5'-cAMP phosphodiesterase similar to the mammalian PDE8 family
K12H4.7	+	11.4	
MSRA-1	+	18.8	methionine sulfoxide-S-reductase
LEC-1	+	27	tandem repeat-type galectin.
C16C10.3	+	18.6	
Y37H2A.14	+	21.5	
C50D2.7	+	8.8	
F53F10.3	+	20.3	
LACT-9	+	6.4	beta-lactamase domain-containing protein
ACS-2	+	22.5	acyl-CoA synthetase
DHS-18	+	8.5	predicted short-chain dehydrogenase
UGT-62	+	23.2	ortholog of human UDP glucuronosyltransferase 1 family, polypeptide A4
RBM-28	+	7.1	ortholog of human RNA binding motif protein 4
GLD-1	+	18.8	protein containing a K homology (KH) RNA binding domain
NGP-1	+	17.2	ortholog of human guanine nucleotide binding protein-like 2 (nucleolar)
F49E2.5	+	9.3	
C16A3.5	+	22.1	
Y97E10AL.3	+	17.9	
ZK105.1	+	16.1	
C26F1.3	+	16.7	
IDHG-2	+	18.5	homolog of the gamma subunit of an NAD+-dependent mitochondrial isocitrate dehydrogenase
F53E10.6	+	14.7	
DIM-1	+	16.1	novel protein containing three immunoglobulin-like repeats in the carboxy

			terminus
SEL-9	+	9.4	member of the *p24* family of proteins that affects growth and locomotion
F57H12.6	+	12.2	
F49E12.1	+	15.6	
H27A22.1	+	8.6	
RAL-1	+	9.1	ortholog of human v-ral simian leukemia viral oncogene homolog B
MEX-5	+	14.7	novel protein that contains two CCCH zinc finger motifs
C04C3.3	+	10.8	
UNC-27	+	10.7	troponin I isoform
F44E2.8	+	9.9	
DHS-21	+	10.8	member of the short-chain dehydrogenases/reductases family (SDR)
LET-754	+	9.6	adenylate kinase
C04F12.1	+	12.3	
CAP-2	+	8.9	beta subunit of actin capping protein that regulates actin cytoskeleton assembly and establishment of initial asymmetry in the embryo
AQP-2	+	5.6	atypical aquaglyceroporin
RNP-5	+	10	putative member of the exon-exon junction complex, orthologous to human RNPS1
LYS-1	+	7.4	putative lysozyme
SGK-1	+	10.3	serine/threonine protein kinase, orthologous to the mammalian serum- and glucocorticoid-inducible kinases (SGKs)
F33A8.4	+	6.4	
T26F2.3	+	10.4	
PQN-48	+	5.3	predicted to contain a glutamine/asparagine (Q/N)-rich ('prion') domain
ALH-4	+	6.5	ortholog of human aldehyde dehydrogenase 3B2
ACOX-1	+	9.3	orthologous to the human gene ACYL-CoA OXIDASE 1, PALMITOYL
F22G12.1	+	8.7	
TAG-173	+	12	
CEC-5	+	10.7	C.Elegans Chromodomain protein, involved in embryo development and meiotic chromosome segregation.

TNT-2	+	8.4	ortholog of human troponin T type 2 (cardiac)
SPAT-2	+	11.3	required for PAR protein-dependent cell-polarity
C37H5.5	+	7.1	
Y104H12D.3	+	5.9	
MUP-2	+	7.7	muscle contractile protein troponin T (TnT)
T01B7.5	+	9.1	
ASC-1	+	5.2	homolog of human activating signal cointegrator-1
Y110A7A.6	+	5.7	
SPTL-1	+	5.2	putative serine palmitoyltransferase
HMG-4	+	7.6	strong similarity to the highly conserved high mobility group protein SSRP1 (structure-specific DNA recognition protein), predicted to be a member of the FACT (facilitates chromatin transcription) complex
UNC-87	+	6.5	required to maintain the structure of myofilaments in body wall muscle cells
T10B11.2	+	4.7	
SAGO-1	+	5.3	Argonaute partially required for the amplification phase of RNAi responses
ZK686.2	+	5.1	
C52E12.1	+	5.2	
ACS-5	+	4.1	ortholog of human acyl-CoA synthetase long-chain family member 5
LIN-41	+	3.9	novel RBCC (Ring finger-B box-Coiled coil) protein that is a member of the NHL (NCL-1, HT2A, and LIN-41) family of proteins
PRG-1	+	3	Piwi subfamily of highly conserved Argonaute/Piwi proteins
Y94H6A.5	+	3.1	
RBD-1	+	2.8	ortholog of S. cerevisiae Mrd1p, required for processing of 18S rRNA and subsequent formation of the 40S ribosomal subunit
F13H8.2	+	4.4	
GLD-2	+	2.2	catalytic subunit of a cytoplasmic poly(A) polymerase (PAP) associated with P granules
ITR-1	+	1.9	putative inositol (1,4,5) trisphosphate receptor
TEN-1	+	1.1	type II transmembrane protein containing

									EGF-like repeats, ortholog of Drosophila Ten-m/Odz and vertebrate teneurins
ANC-1						+		0.3	orthologous to Drosophila MSP 300 and mammalian SYNE1 and SYNE2, has coiled regions, a nuclear envelope localization domain and an actin-binding domain

Table 4.3: List of HRDE-1-specific interacting proteins.

NRDE-3 + HRDE-1 INTERACTORS	*NRDE-3 (1)*	*NRDE-3 (2)*	*NRDE-3 (3)*	*NRDE-3 (4)*	*NRDE-3 (5)*	*HRDE-1*	*NRDE-2*	*peptide coverage*	*Description*
MISC-1	+	+		+		+		6 - 14	ortholog of human solute carrier family 25
C06E1.9	+	+		+		+		~4	
F53F4.11	+				+	+		6	ortholog of human ribosomal L1 domain
ECH-6	+	+				+		9 - 38	ortholog of human enoyl CoA hydratase, short chain, 1, mitochondrial
C25A8.4	+	+				+		3 - 15	ortholog of human chitinase 3-like 2
MMAA-1	+	+				+		9 - 10	ortholog of human methylmalonic aciduria (cobalamin deficiency) cblA type
B0395.3	+	+				+		6 - 9	orthologous to the human gene choline acetyltransferase isoform r
SPG-7	+	+				+		3 - 8	metalloprotease orthologous to human paraplegin, mitochondrial
C05G5.4	+					+		11 - 23	encodes an ortholog of human succinate-CoA ligase, alpha subunit
F10E7.6	+					+		9 - 21	
F25B4.1	+					+		10 - 16	orthologous to the human gene glycine cleavage system t-protein
DNJ-7	+					+		6 - 8	protein containing a DnaJ domain that is orthologous to vertebrate P58IPK/DNAJC3
RIOK-1	+					+		8 - 10	putative RIO kinase , orthologous to human RIOK1
SAD-1	+					+		2	novel serine/threonine protein kinase

IFF-2		+				+	12 - 33	ortholog of human eukaryotic translation initiation factor 5A
EIF-1		+				+	35	ortholog of human eukaryotic translation initiation factor 1B
T23D8.3		+				+	12 - 19	ortholog of yeast and human LTV1 that inhibits DHC-1
F49D11.10		+				+	6 - 8	ortholog of human WD repeat domain 75
ABTS-3			+			+	4 - 7	anion transporter
ZC247.1			+			+	1	ortholog of human mucin 17, cell surface associated

Table 4.4: List of common interacting proteins between NRDE-3 and HRDE-1.

NRDE-3 + NRDE-2 INTERACTORS	*NRDE-3 (1)*	*NRDE-3 (2)*	*NRDE-3 (3)*	*NRDE-3 (4)*	*NRDE-3 (5)*	*HRDE-1*	*NRDE-2*	*peptide coverage*	*Description*
C14H10.2			+	+	+		+	3 - 7	ortholog of human early endosome antigen
M153.1			+		+		+	8 - 21	ortholog of human pyrroline-5-carboxylate reductase 1
F44E5.4				+	+		+	5 - 8	ortholog of Hsp70 family
PBS-7			+				+	12	B-type subunit of the 26S proteasome's 20S protease core particle
PAS-7			+				+	10	proteasome alpha-type three subunit of the core 20S proteasome subcomplex
CPR-6			+				+	8 - 12	ortholog of human cathepsin B.

Table 4.5: List of common interacting proteins between NRDE-3 and NRDE-2.

NRDE-3 + NRDE-2 + HRDE-1 INTERACTORS	*NRDE-3 (1)*	*NRDE-3 (2)*	*NRDE-3 (3)*	*NRDE-3 (4)*	*NRDE-3 (5)*	*HRDE-1*	*NRDE-2*	*peptide coverage*	*Description*
F46G10.1			+			+	+	23-27	ortholog of human potassium channel tetramerization domain containing 21

Table 4.6: Common interacting protein between NRDE-3, HRDE-1 and NRDE-2.

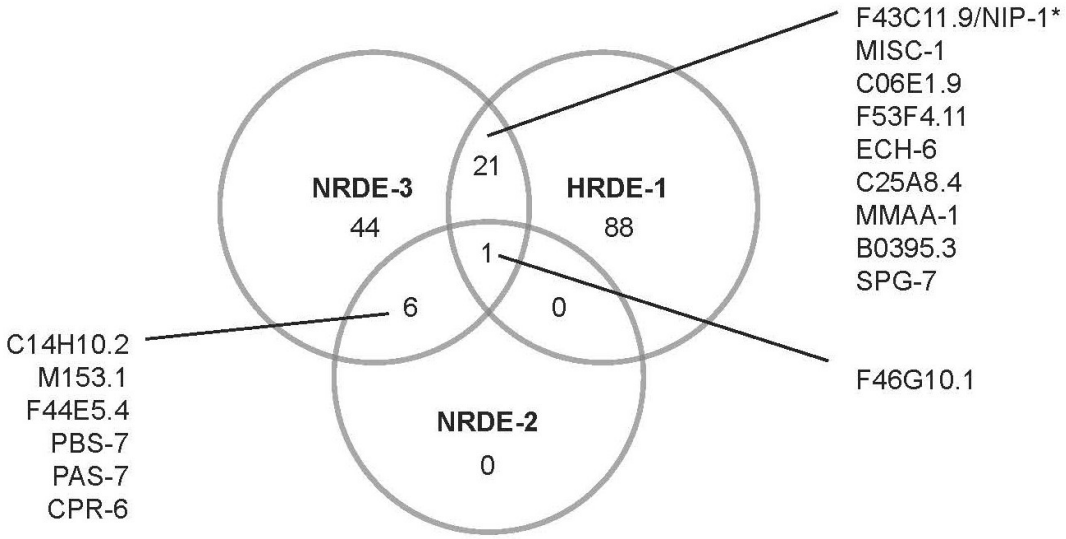


Figure 4.3: Venn diagram of interactions detected in NRDE-3, NRDE-2 and HRDE-1 purifications. For NRDE-3/HRDE-1 overlapping interactors, those present in 2 or more NRDE-3 datasets are displayed. *NIP-1 was found to interact with NRDE-3 by MudPIT and western blot, and with HRDE-1 by western blot (See Figure 4.6 below).

4.2.2 Validation and characterization of NRDE-3 interacting protein 1 (NIP-1), a novel Argonaute interactor.

Focusing on the proteins found in NRDE-3 purifications, for which we had the most biological replicates and highest confidence data, we noticed that one previously uncharacterized protein was present in 5/5 NRDE-3 datasets: F43C11.9. Furthermore, F43C11.9 was found at very high peptide coverage in all samples (26-50%) and specifically with NRDE-3. Since the peptide coverage of NRDE-3 itself ranged from 58-74% in these samples, we reasoned that F43C11.9 was a good candidate for a strong or constitutive interaction and renamed it NRDE-3 interacting protein 1(NIP-1). NIP-1 is a protein predicted to be 339 residues with no recognizable domains. Searching for proteins with homology to NIP-1, we found that NIP-1 aligns to the C-terminus of MUT-15 in *C. briggsae* (22% identity, 40% similarity), a closely related nematode to *C. elegans* (Figure 4.4). MUT-15 is a member of the ERI endoRNAi pathway and is required for the biogenesis or maintenance of primary endo-siRNAs through an unknown mechanism (Zhang et al., 2011). NIP-1 also aligns to an uncharacterized gene in *C. elegans* called T01C3.9 (23% identity, 41% similarity), whose genomic

location is immediately downstream on *C. elegans* MUT-15 (T01C3.8), which is shorter than the *C. briggsae* version (Figure 4.4). This raises the possibility that the two proteins in *C. elegans,* MUT-15 (T01C3.8) and T01C3.9, could act together to perform the same functions as *C. briggsae* MUT-15. Importantly, T01C3.9 was not found in any of the NRDE-3 MudPIT datasets, arguing against the possibility of redundancy between the two proteins in NRDE-3 association – at least in wild-type conditions. However, NIP-1 also shares 25% identity (and 46% similarity) with another uncharacterized gene in *C. elegans* called T24C2.2, which was found in 2/5 of the NRDE-3 MudPIT datasets. This data indicates that NIP-1, T01C3.9, and T24C2.2 may be members of a protein family and perform similar unknown functions.

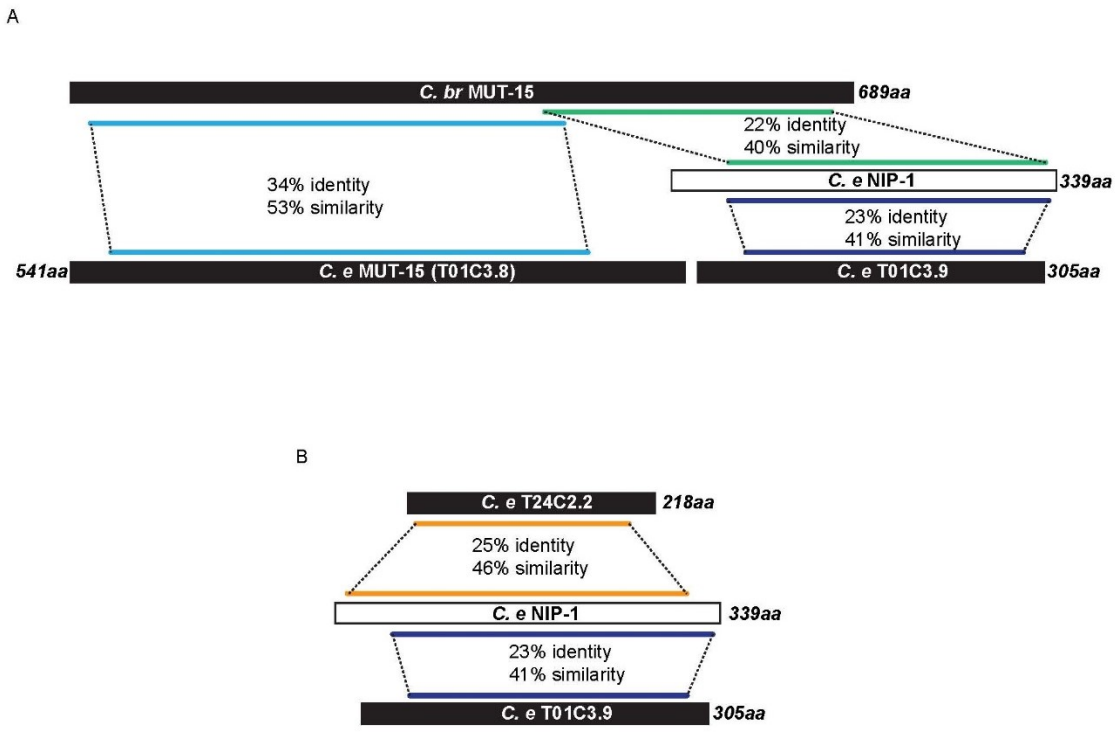


Figure 4.4: NIP-1 homologous proteins in *C. elegans* and *C. briggsae*. (A) Using NCBI Protein BLAST, sequence homology was found between NIP-1 and *C. briggsae* MUT-15, and another *C. elegans* gene T01C3.9, whose genomic location immediately follows *C. elegans* MUT-15. (B) NIP-1 shares homologous sequences with another *C. elegans* gene, T24C2.2, which is also found in NRDE-3 MudPIT datasets. T24C2.2 also shares homology with the C-terminus of *C. briggsae* MUT-15 (amino acids 418-497: 29% identity, 51% similarity; amino acids 620-668: 41% identity, 59% similarity). Protein schematics are to scale and corresponding coloured bars represent regions of homology.

In order to determine if the NIP-1 interaction with NRDE-3 is specific, and to facilitate its functional characterization within the RNAi pathways, we next developed two polyclonal rabbit antibodies to NIP-1. We performed FLAG IPs on extracts from embryos expressing 3XFLAG::GFP::NRDE-3 in the wild-type genetic background, and wild-type embryos as the negative control. A GFP antibody was used to detect tagged NRDE-3 in the input and IP fractions, and both NIP-1 antibodies (5308 and 5309) were used to probe these fractions for NIP-1 (Figure 4.5A). NIP-1 was easily detectable in the NRDE-3 IPs by western blot at the expected size of ~40kDa, confirming the MudPIT results. We also noticed that NIP-1 levels in the input fraction from the transgenic 3XFLAG::GFP::NRDE-3 strain was higher than wild-type levels (Figure 4.5, compare lanes 1 and 2 on membranes probed for NIP-1). Since we lack an antibody to the endogenous NRDE-3 protein, we are not able to calculate the exact level of over expression of NRDE-3 in the 3XFLAG::GFP::NRDE-3 strain, but it is reasonable to assume that the level is higher than wild-type. If this is true, the concomitant increase in NIP-1 levels detected in the 3XFLAG::GFP::NRDE-3 strain could reflect a coupled

expression level for these two protein partners, hence suggesting that NRDE-3 and NIP-1 are direct interactors. Coupled expression is also a feature of the ERI endoRNAi complex components RRF-3 and ERI-5, the RDE exoRNAi complex components RDE-4 and RDE-1 (Thivierge et al., 2012), human miRNA biogenesis factors Drosha and DGCR8 (Han et al., 2009), and Drosophila endo/exoRNAi pathway components R2D2 and Dicer2 (Liu et al., 2003). The NIP-1 antibodies were also used to immunoprecipitate the endogenous protein from wild-type embryo and adults extracts, where NIP-1 is expressed at detectable levels in both cases (Figure 4.5B).

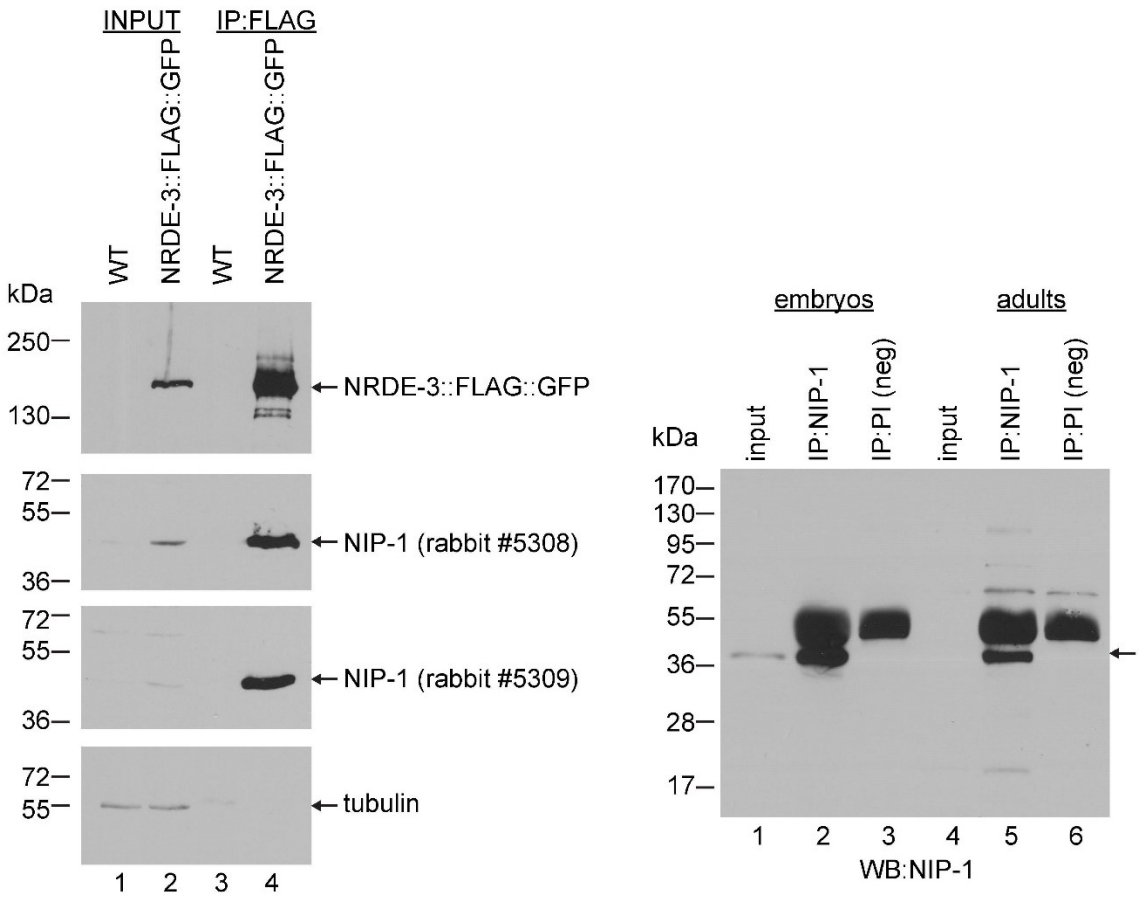


Figure 4.5: NIP-1 interacts with NRDE-3. (A) Extracts and FLAG immunoprecipitates from wild-type (negative control) or NRDE-3::3XFLAG::GFP embryos were probed with newly generated antibodies to NIP-1 (5308 and 5309) to confirm the interaction observed in MudPIT experiments. (B) Immunoprecipitation of endogenous NIP-1 with antibody 5308 in wild-type embryos and adult animals. NIP-1 is expressed in both stages, but at a higher level in embryos. Pre-immune serum was used in the IP as the negative control. The non-specific bands in both IPs close to 55kDa likely correspond to antibody heavy chains present in the rabbit serum.

Next, to test if NRDE-3 was the only Argonaute that could interact with NIP-1, we immunoprecipitated NIP-1 and immunoblotted for other AGOs or AGO-associated proteins in wild-type extracts. Surprisingly, we found that NIP-1 co-precipitated the Argonautes HRDE-1, ERGO-1 and SAGO-1. DCR-1 was also weakly detected in the NIP-1 IP (Figure 4.6). This data demonstrates that NIP-1 is not only a protein partner of the nuclear Argonautes NRDE-3 and HRDE-1, but also primary ERI endoRNAi Argonaute ERGO-1, and secondary Argonaute SAGO-1 which is shared between the exoRNAi and endoRNAi pathways. SAGO-1 is a key secondary AGO as its overexpression can rescue the phenotypes associated with loss of multiple secondary AGO genes in *C. elegans* (Yigit et al., 2006). No specific interaction was found between NIP-1 and the miRNA pathway-specific AGOs ALG-1/2 or their cofactor AIN-2, indicating that NIP-1 is excluded from the miRNA pathway (Figure 4.6).

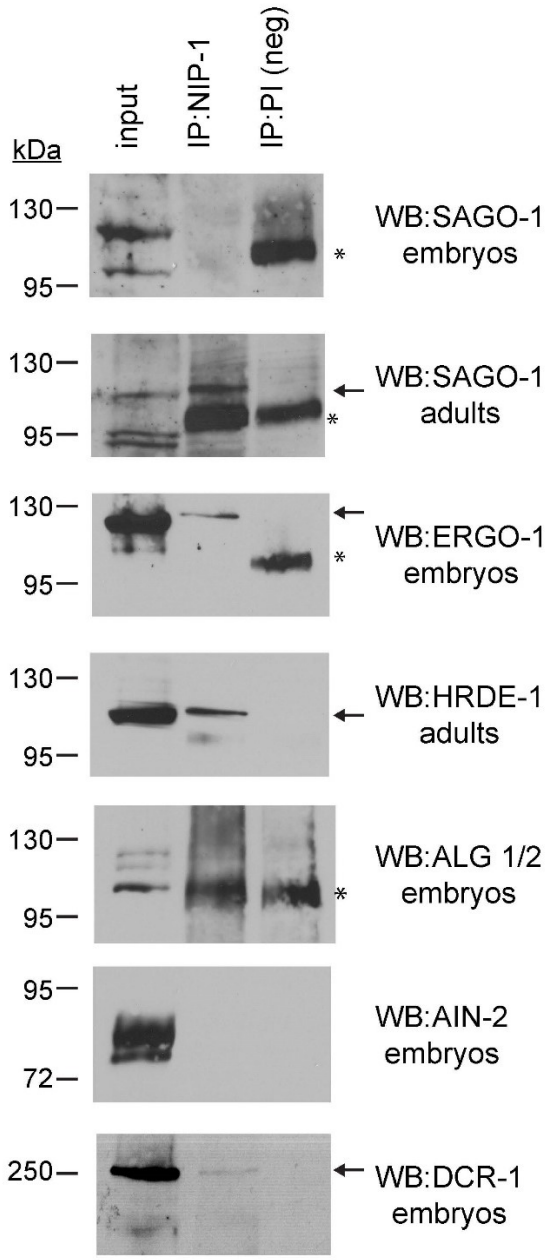


Figure 4.6: NIP-1 interacts with endoRNAi and exoRNAi Argonautes. Immunoprecipitates of NIP-1 from wild-type embryos or adults were probed with antibodies to endogenous Argonautes or known protein partners. Arrows indicate the presence of a specific interaction between NIP-1 and the protein probed for by western

blot. Asterisks indicate non-specific proteins. Pre-immune serum was used in IP as the negative control. Representative blots from 2-3 biological replicates are shown.

In sum, these experiments indicate that NIP-1 participates in multiple distinct RNAi complexes – at the DCR-1, primary Argonaute, and the secondary Argonaute level. Interestingly, all of the Argonautes that we have found in complex with NIP-1 are loaded with siRNAs downstream of RNA-dependent RNA polymerases, in a process that is poorly understood. Thus, it is possible that NIP-1 plays a role in Argonaute loading (Figure 4.7). Loss-of-function analysis and RNAi functional assays will be necessary to identify the exact function of NIP-1.

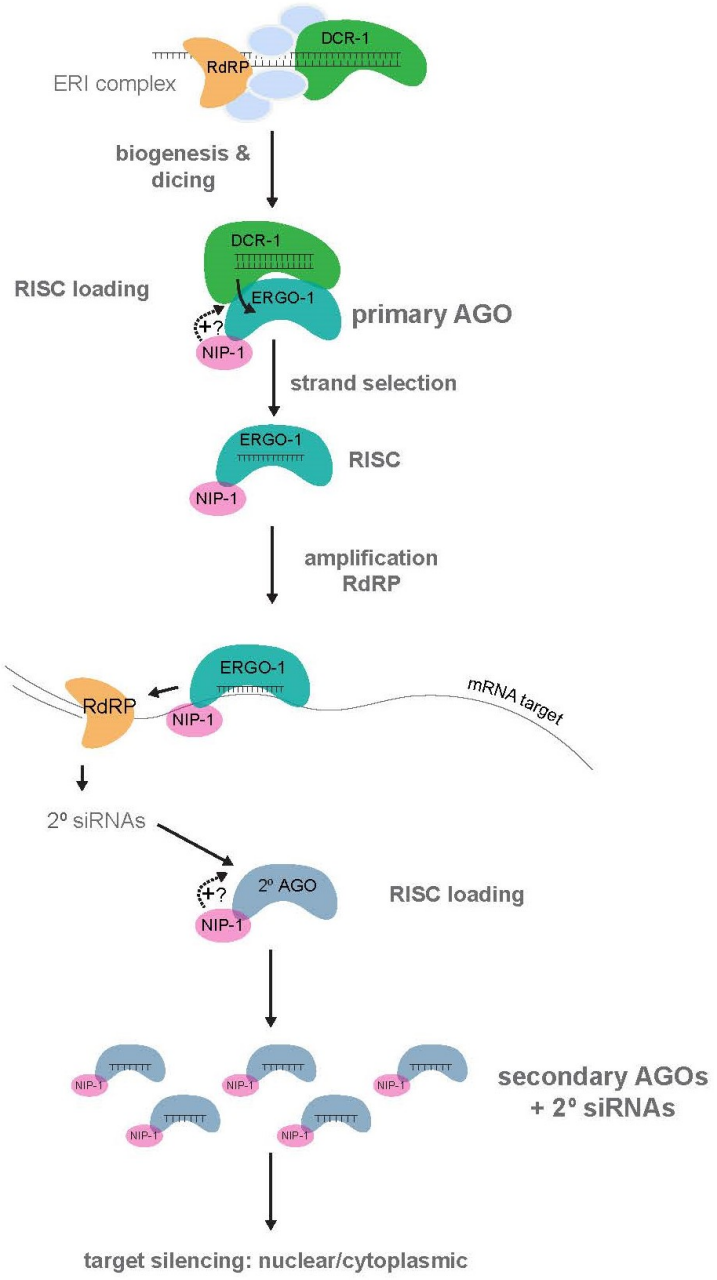


Figure 4.7: NIP-1 interactions at different levels of endo- and exoRNAi biogenesis and effector steps. Using proteomics, NIP-1 was identified as a novel interactor of secondary Argonaute NRDE-3, which shuttles to and from the nucleus. Furthermore, NIP-1 was found in complexes with multiple other Argonautes of the ERI endoRNAi and

exoRNAi pathways, as well as DCR-1. We hypothesize a direct interaction between NIP-1 and Argonaute proteins, but cannot at this point rule out the presence of other protein(s) bridging this interaction. Since NIP-1 is in complex with multiple AGOs, we propose that NIP-1 plays a role in a process that is common to all AGOs, such as RISC loading.

4.3 Discussion

We have generated a high confidence list of novel protein interactors of the nuclear Argonaute NRDE-3 using proteomics-based discovery. Moreover, we have also generated preliminary interactor lists of the germline nuclear Argonaute HRDE-1 and the conserved nuclear RNAi factor NRDE-2. The interactors of HRDE-1 and NRDE-2 will be further validated by additional MudPIT datasets from biological replicates compared to mock purifications to identify reproducible and specific interactions. Surprisingly, we did not detect any other NRDE proteins in complex with NRDE-3, arguing against a previously proposed model in which NRDE-3 in the nucleus physically recruits NRDE-2, 1 and 4 to mediate transcriptional silencing and H3K9 trimethylation (Ashe et al., 2012; Buckley et al., 2012; Burkhart et al., 2011; Burton et al., 2011; Guang

et al., 2010; Lee et al., 2012; Luteijn et al., 2012; Shirayama et al., 2012). It is important to note that these previous studies identified NRDE-1, 2, and 4 as genetically required for NRDE-3-mediated nuclear silencing, and the only physical link demonstrated between any of these proteins was between transgenic NRDE-3 and NRDE-2 (Guang et al., 2010).

We identified NIP-1 (F43C11.9) as the highest confidence protein partner of NRDE-3 found in this screen, and validated this interaction by western blot. We further demonstrated that NIP-1 co-precipitates both primary and secondary Argonautes of the endoRNAi and exoRNAi pathways, and is excluded from interactions with core miRNA pathway components. NIP-1 does not contain any recognizable domains, and its molecular function remains unknown. NIP-1 has homology to *C. briggsae* MUT-15, which in *C. elegans* is genetically required for the expression of ERGO-1 bound primary endo-siRNAs (26Gs), through an unknown mechanism (Zhang et al., 2011), and was identified as part of a gene class required for transposon silencing (Vastenhouw et al., 2003). MUT-15 is present in germline peri-nuclear structures called “mutator foci” along

with other MUT proteins and the RNA-dependent RNA polymerase RRF-1, which is required for secondary siRNA biogenesis (Phillips et al., 2012). Moreover, MUT-15 is also called RNAi defective 5 (RDE-5), and is genetically required for exoRNAi pathway function. The C-terminal portion of MUT-15 in *C. briggsae* is the region with homology to NIP-1, however this region is absent in the *C. elegans* MUT-15. NIP-1 is also homologous to the gene immediately downstream of *C. elegans* MUT-15, T01C3.9, and another *C. elegans* gene T24C2.2. These three (NIP-1, T01C3.9, and T24C2.2) are all close in size to the C-terminal region of *C. briggsae* MUT-15. Taking all this together, it is tempting to speculate that in *C. elegans*, the function of the missing C-terminal region in MUT-15 is fulfilled by a family of proteins that includes NIP-1, T01C3.9, and T24C2.2. A similar role for NIP-1 and T24C2.2 is supported by the fact that they were both reproducibly found in MudPIT datasets from NRDE-3 purifications.

If NIP-1 functions in conjunction with MUT-15, it could be involved in secondary siRNA biogenesis or stability. Additionally, since endogenous NIP-1 was found in complex with multiple Argonautes and a small amount of DCR-1, NIP-1 could also be involved in the

RISC-loading complex. Furthermore, since its interactions with the Argonautes appear to be robust and non-transitive, NIP-1 could be a stable member of the mature RISC. We will investigate if any of the interactions seen with NIP-1 are siRNA-dependent, as well as if NIP-1 has any dsRNA or ssRNA binding activity *in vitro* by EMSA. Since MudPIT does not discriminate between direct interactions and those bridged by other proteins, *in vitro* biochemical binding assays will be useful in establishing which, if any, of the identified interactors are direct. To this end, we will test if the NIP-1-Argonaute interactions are direct by GST-pulldown. It is possible that NIP-1 binds AGOs directly by interacting with conserved residues or a common structure (domain/motif) to those partners identified. In NIP-1 IPs, we noticed that SAGO-1 in adults was the only AGO that was enriched over the input levels, whereas the amount of the other AGOs (ERGO-1 and HRDE-1) in the IPs was lower than the input (Figure 4.6). This raises the possibility that NIP-1 has a greater affinity for some AGOs over others, or that the expression domains of NIP-1 and certain AGO partners have greater overlap than others. In order to gain a comprehensive view of NIP-1 protein interactions, we will

perform MudPIT analysis on NIP-1 IPs. This will be non-quantitative, but would allow us to generate a relative hierarchy of AGO partners by peptide coverage and degree of reproducibility, as well as to identify other NIP-1 interactors.

Nevertheless, we have identified a previously unknown pan-Argonaute partner in the exoRNAi and endoRNAi pathways. Current and future directions include generating a knockout strain for NIP-1 using CRISPR/Cas9-mediated gene editing and analyzing the effect of its loss of function in the RNAi pathways. Since we could not detect any specific interactions between NIP-1 and the miRNA Argonautes, we predict that NIP-1 is excluded from the miRNA pathway, and its loss may be tolerated. If *nip-1-/-* animals are viable, we will perform total small RNA sequencing at all different developmental stages in wild-type animals compared to *nip-1-/-* to determine which small RNA populations, if any, require NIP-1 function. This analysis can be further refined by sequencing the small RNA populations bound to NIP-1 partner AGOs in the wild-type versus null mutant to test if NIP-1 is involved in loading of the Argonautes. In the case of NRDE-3, it was previously shown that the Argonaute remains cytoplasmically localized in the absence

of loaded ERI endo-siRNAs (Guang et al., 2008). Therefore, the effect of NIP-1 on Argonaute loading can be tested by analyzing the localization of GFP-tagged NRDE-3 in wild-type and *nip-1-/-* animals.

In the NRDE-3 datasets, we also identified two proteins with known functions in chromatin modification: MRG-1 and SIN-3. These interactions are the first physical links between a nuclear AGO in animals and chromatin modifying machinery. MRG-1 is a chromodomain-containing protein ortholog to mammalian MRG15, which associates with histone acetyltransferases and deacetylases and has key roles in embryonic development and proliferation (Pardo et al., 2002; Tominaga et al., 2005). Additionally, in *C. elegans*, MRG-1 was found to be autosome-associated and required for repression of some X-linked genes and transgenes on extrachromosomal arrays in the germline (Takasaki et al., 2007). MRG-1 further contributes to proper homologous chromosome pairing during meiosis (Dombecki et al., 2011). On the basis of its previously characterized activities, we speculate that MRG-1 may play a key role in mediating the silencing of genes targeted by NRDE-3 and its associated small RNAs,

possibly through its interaction with histone acetyltransferases/deacetylases. MRG-1 was also identified as a genetic requirement for the maintenance of long-term silencing seen in *C. elegans* initiated by exoRNAi (Vastenhouw et al., 2006), indicating possibly convergent mechanisms of the exoRNAi and endoRNAi pathways. SIN-3 is a homolog of the yeast Switch Independent histone deacetylase (HDAC) component, and is postulated to function as a scaffolding protein in SIN-3 repressive HDAC complexes (Laherty et al., 1998; Zhang et al., 1998). Other members of this complex include two histone deacetylases (HDAC1 and HDAC2), two histone-binding proteins (RbAp46 and RbAp48), and two proteins of unknown function (SAP18 and SAP20). The SIN-3 complex has been studied well in yeast and mammalian cells, where it has been shown to mediate transcriptional repression downstream of nuclear hormone receptors (Chen and Evans, 1995; Heinzel et al., 1997; Horlein et al., 1995; Nagy et al., 1997). An appealing possibility is that NRDE-3 loaded with siRNAs, could interact with the SIN-3 complex in *C. elegans* to direct histone deacetylation and transcriptional repression of specific endoRNAi and exoRNAi target genes.

We will pursue the validation of NIP-1, MRG-1, SIN-3, and the other interactors identified by MudPIT in the mechanism of nuclear RNAi by assessing the effect of their loss on silencing of a nuclear-localized RNA, as was done to identify other components of the nuclear RNAi pathway in genetic screening (Guang et al., 2008). Interactors that are shared between NRDE-3, HRDE-1, and NRDE-2 will also be prioritized, due to the higher likelihood that they will have functional relevance in nuclear RNAi. To discriminate between nuclear RNAi and cytoplasmic RNAi, we will trigger exoRNAi against different mRNAs that generate visible phenotypes: *unc-22* in the cytoplasm, and *lir-1* in the nucleus. *unc-22* produces a non-lethal muscle twitching phenotype (Fire et al., 1998). *lir-1* exists in a polycistron with *lin-26* in the nucleus exclusively. *lir-1* loss of function is viable, while *lin-26* loss of function is not, therefore when silencing occurs in the nucleus, the polycistron is targeted and the result is sterility. If mutants of NRDE-3 interactors show no twitching under *unc-22* RNAi, it would validate their function in cytoplasmic silencing. Conversely if mutants are viable following *lir-1* RNAi, it would validate their function in nuclear RNAi. Taking NIP-1 for example, as it was found to

interact with both cytoplasmic and nuclear AGOs, we hypothesize that the null mutant would be a suppressor of both cytoplasmic and nuclear RNAi. By validating the interactors we have found specifically with NRDE-3 and by expanding the proteomic network with HRDE-1, NRDE-2 and NIP-1, we aim to achieve a greater understanding of both cytoplasmic and nuclear RNAi processes.

4.4 Materials and methods

4.4.1 *C. elegans* Strains

All strains were cultured as in (Brenner 1974). N2 was used as the wild-type strain.

4.4.2 Protein purification

Pellets were homogenized in 50mM Tris-HCl pH8/150mM NaCl/1mM EDTA/0.1% Igepal with Complete EDTA-free protease inhibitors (Roche), and cleared by 17 000xg centrifugation. FLAG purifications were performed on wild-type cytoplasmic extracts (approximately 5mg total protein each) using M2 Affinity Gel (Sigma) and eluted with 3XFLAG peptide (Sigma) according to the manufacturer's instructions.

4.4.3 MudPIT analysis

FLAG eluates were acetone precipitated before resuspension in denaturing digestion buffer, digested and analyzed by μLC/μLC-MS/MS as previously described (Duchaine et al., 2006) at UCLA in the lab of James Wohlschlegel.

4.4.4 Recombinant protein production

F43C11.9 (*nip-1*) was amplified using the forward primer ATAATAGAATTCATGAAAATAGCAAAATTCACC and reverse primer TATTATGTCGACTTACCAAACATCCAATTTATACA from N2 cDNA and cloned in pET-28a with EcoRI and SalI restriction sites (underlined). The resulting plasmid was transformed in BL-21 Rosetta DE3 cells. Bacteria was grown at 30^{o}C until the O.D. reached 0.5 and subsequently chilled on ice for 15minutes. 1mM IPTG was added and induction was performed O/N at 16^{o}C. Cells were centrifuged for 15mins at 4,000g and frozen at -80^{o}C. 6xHis-NIP-1 was purified using GE Ni-NTA matrix using batch purification. Cells were lysed (20mM sodium phosphate pH6.4, 500mM NaCl, 10mM imidazole, 10% glycerol) by emulsification and centrifuged at 25,000g for 20 minutes. The supernatant was incubated with Ni-NTA beads for 30 minutes at 4^{o}C. Beads were washed sequentially with 10 volumes of lysis buffer containing 10mM, 20mM, 40mM

and 60mM imidazole and eluted with 10 times 1 volume of elution buffer (Lysis buffer with 250mM imidazole). Eluted protein was diafiltered and concentrated in lysis buffer without imidazole and injected into rabbits for antibody production (Capralogics, Cambridge Massachussets).

4.4.5 IP and Western blotting

IPs and western blots were performed on extracts prepared from embryos or young-gravid adults as in (Sawh and Duchaine, 2013). Antibodies used were: rabbit polyclonals against NIP-1 (5308, 5309 Capralogics), DCR-1, ALG-1/2 (Duchaine et al., 2006), ERGO-1 (Vasale et al., 2009), SAGO-1 (gift from Craig C. Mello), HRDE-1 (gift from Eric Miska), AIN-2 (gift from Martin J. Simard), alpha tubulin (Abcam), GFP (Roche) and FLAG (Sigma). HRP-conjugated rabbit and mouse TrueBlots were used as secondary antibodies (eBioscience). Protein A Sepharose CL 4B (GE Healthcare) or Protein G Sepharose 4 Fast Flow (GE Healthcare) and ANTI-FLAG M2 (Sigma) beads were used according to the manufacturer's instructions for IP.

4.5 Acknowledgements

We thank Craig Mello, Eric Miska and Martin Simard for sharing antibodies. We thank Scott Kennedy for sharing the strains: YY179 [nrde-3p::3xflag::gfp::nrde-3], YY557 [hrde-1p::3xflag::gfp::hrde- 1], and YY346 [nrde-3p::3xflag::gfp::nrde-2]. We also thank Susan Gasser for the strains: GW694 [his-72p::mcherry-set-25::his-72 3'UTR; unc-119(+)];unc-119(ed3);ttTi5605 II, GW699 [his-72p::mcherry-met-2::his-72 3'UTR; unc-119(+)];unc-119(ed3);ttTi5605 II. This work was supported by the Canadian Institute of Health Research (CIHR) MOP123352, the Human Frontiers Science Program, and the Fonds de la Rercherche en Santé du Québec (FRQS), Chercheur-Boursier Salary Award J2 to (T.F.D.). A.N.S. is supported by the CIHR Frederick Banting and Charles Best Canada Graduate Scholarship, the Defi Corporatif Canderel Award, the Alexander McFee Memorial Fellowship, and the Ruth and Alex Dworkin Scholarship.

General Discussion and Summary

Complementary genetic and biochemical approaches to drive discovery in RNAi.

In these studies, we have undertaken biochemical approaches to address unanswered questions in RNA interference using *C. elegans* as a model organism. The RNAi field was primarily built on powerful forward and RNAi reverse genetic screens performed in model organisms to identify core and accessory genes involved in the multiple pathways. Of particular interest, genetic screening in *C. elegans* has identified mutants that led to the discoveries that the RNAi pathways compete with one another (Duchaine et al., 2006; Kennedy et al., 2004), and that RNAi processes are linked to regulatory changes at the chromatin level (Ashe et al., 2012; Guang et al., 2008; Vastenhouw et al., 2006). These screens were successful in identifying genes involved in RNAi processes by virtue of clearly defined phenotypes. Classical forward genetics can identify both loss-of-function and gain-of-function alleles of important genes in the particular process of interest, while RNAi reverse genetics (using exoRNAi in a genome-

wide or targeted manner) primarily identifies genes whose partial loss-of-function produces the desired phenotype.

However, essential genes and genes whose knockdown does not achieve the threshold required to produce the sought-after phenotype are excluded by these methods. Using exoRNAi as a tool to induce knockdown of genes involved in RNAi processes can also be an intrinsically limiting approach, since many of the target genes are necessary to execute knockdown, and the pathways are known to compete with one another. A substantial hurdle appears when the field reaches a point where such methods near saturation, and the discovery phase stagnates. Yet, it is very unlikely that all the genes involved in RNAi processes have now been identified. We postulate that those areas of RNAi mechanisms that remain the most elusive, like nuclear RNAi for example, involve genes that are refractory to identification by genetic screening. Not only would essential genes involved in the fundamental processes of chromatin organization be underrepresented due to inviability of their null alleles, but redundancy within gene family members would also mask the effect of a single gene loss.

Could this be the reason why NIP-1 and its putative family members have not been previously identified in RNAi? We have reached a point where it is necessary to refine the mechanisms of these RNAi pathways. Moreover, we now have the biochemical tools needed to accomplish this task. By using proteomics-based discovery, we established a platform from which we aim to build a network of nuclear RNAi factors (Chapter 4). Beginning with the Argonaute NRDE-3, which shuttles to and from the nucleus and is loaded with both secondary endo-siRNAs and exo-siRNAs, we identified chromatin modifying proteins and novel proteins as interactors. The identification and characterization of NIP-1 as a multi-Argonaute partner provides evidence that this unbiased strategy can shed light into both nuclear and cytoplasmic RNAi processes. This approach also has the potential to identify interactors that link the target protein to previously uncharacterized functions, since the readout is protein interaction rather than phenotypic assay. We have begun to expand the network to another nuclear Argonaute (HRDE-1) and a protein shared between the two Argonautes (NRDE-2). As the project

progresses, we will use validated interactors as launchpads for further proteomics-based discovery.

It would also be appealing to perform similar experiments on RNAi factors in another genetically tractable model organism, *D. rerio* for example, to ask whether any of the mechanisms we uncover are conserved, and to what extent. It is important to note that proteomic screens are not without their drawbacks as well. Interactions must be rigorously validated, and final functional significance can only be realized through genetic loss- and gain-of-function analysis. We therefore promote the use of complementary biochemical and genetic approaches to further understand the nuclear RNAi processes. The rise to prominence and ease of application of CRISPR/Cas9-mediated genome editing techniques in recent years facilitates such studies immeasurably. We now have the ability to engineer null alleles, targeted point mutations, or epitope tags on the endogenous locus, to manipulate genomes with precision. Such approaches would yield the most physiologically relevant findings in any model system.

Insights into the many mechanisms of RNAi-induced gene silencing.

In addition to the discovery of novel RNAi components, biochemical assays can also be useful in elucidating the ultimate silencing mechanisms of the pathways. Our group has previously developed a cell-free translation system to investigate the impact of miRNAs on their targets in *C. elegans* embryos, and found that miRNAs drive target mRNA deadenylation in a co-operative manner (Wu et al., 2010). Similarly, in order to precisely define the mechanism of nuclear RNAi downstream of the ERI endoRNAi pathway, we developed a robust nuclear run-on assay, and found that ERI targets are transcriptionally silenced (Chapter 3). One immediate application of this assay would be to test the genetic requirements of ERI endoRNAi-mediated transcriptional silencing. For instance, it is currently unclear if NRDE-3 is the sole nuclear Argonaute loaded with secondary ERI endo-siRNAs. Thus, the exact contribution of each secondary Argonaute (or group of Argonautes) towards the transcriptional regulation of ERI targets can be tested in the run-on, using strains bearing Argonaute null alleles.

Another open issue in the nuclear RNAi field is the exact role that histone tail modification plays in gene silencing. Does transcriptional inhibition precede histone modification or vice versa? Can the establishment of siRNA-directed silencing occur in the absence of such modifications? This is not a simple issue to tackle using conventional null alleles, since loss of chromatin modifying machinery is often deleterious to the health of the animal. Perhaps it would be possible to design a nuclear cell-free assay, similar to the embryonic run-on, where RNAi-induced transcriptional silencing and histone modification are recapitulated. The key modifying proteins can then be biochemically depleted, and transcription of an RNAi-targeted gene monitored. In such a scenario, it would also be interesting to supplement the transcriptionally-competent extract with histone modifying enzymes targeting a specific genomic location (by tethering) to monitor the causality and kinetics of transcriptional regulation.

In addition to the study of nuclear RNAi, the nuclear run-on assay we have developed can be broadly applicable to the study of any embryonic transcript. One interesting application would be to identify the exact transcriptional cascades that occur following

zygotic genome activation in the early embryo (approximately the 4-cell stage), and subsequently throughout development during the maternal to zygotic transition (MZT). The transfer of developmental control from maternally-provided factors to the initial zygotic transcripts is a fascinating process, which has largely been defined through genome-wide measurements of steady-state mRNA levels (Baugh et al., 2003; Levin et al., 2012), not transcription. Large and precisely timed embryo populations will be necessary to accurately catalog such changes, and this can be accomplished through fluorescence-activated cell sorting of embryos expressing differentially timed GFP markers as in (Stoeckius et al., 2009).

Fine-tuning of the master regulatory enzyme Dicer.

Biochemical studies are also needed to define molecular mechanisms of the important RNAi factors that have already been identified, and how they are each regulated. It is clear that in order to understand the intricate regulatory control that Dicer proteins exert on gene expression, it will be important to consider the varied roles that full-length and truncated forms of the enzyme play. To this end, we have pursued studies on the post-

translational regulation of DCR-1. We discovered that endogenous DCR-1 is proteolytically cleaved to produce a stable and active C-terminal fragment which encodes the two catalytic RNase III domains and the dsRBD domain, which we named sDCR-1 (Chapter 1). sDCR-1 expression promotes the activity of the exoRNAi pathway, while it simultaneously blocks the miRNA biogenesis pathway by competitive inhibition of the miRNA Argonautes. sDCR-1 expression is developmentally regulated, and abberant early expression leads to an inhibition of the miRNA pathway strong enough to negatively affect developmental timing and viable progeny. We propose that sDCR-1 acts to promote anti-viral defense due to its enhancement of the exoRNAi pathway, therefore it would be interesting to test the susceptibility of wild-type, sDCR-1 null and sDCR-1 overexpressed animals to viruses that were recently found to infect *C. elegans* (Felix et al., 2011).

Additionally, it may be possible that sDCR-1 has supplementary unknown functions *in vivo,* through unique targets compared to full-length DCR-1. Bacterial RNase III enzymes possess a single RNase III domain and a dsRBD, but are known to form

homodimers to cleave dsRNA (Zhang et al., 2004), therefore *C. elegans* sDCR-1 may fulfill a similar role as a general-purpose RNase in addition to its role in the RNAi pathways. The interaction between DCR-1 proteins and their substrates are largely transient by nature and therefore were historically difficult to isolate, but recent work using *in vivo* cross-link immunoprecipitation has identified many novel miRNAs and mRNA targets of full-length DCR-1 (Rybak-Wolf et al., 2014), and it would be of great interest to apply this method to assess the RNA targets of sDCR-1.

The identification of the protease responsible for sDCR-1 generation would also shed some light on the upstream cascade of events required for sDCR-1 production. Since sDCR-1 expression enhances exoRNAi, we predict that mutation of the protease would result in a suppressor of RNAi phenotype. Since our original publication (Sawh and Duchaine, 2013), another group has shown that human Dicer proteolysis to produce an sDCR-1-like fragment is stimulated by Ca^{2+} in HEK293 cells (Zimmermann et al., 2014), and it would therefore be interesting to test Ca^{2+}-dependent proteases for their activity on *C. elegans* DCR-1. An sDCR-1-like protein is also expressed in some human breast

cancer cell lines, where it is generated by alternative splicing instead of proteolytic cleavage, and is curiously down-regulated following induction of the epithelial to mesenchymal transition (EMT) (Hinkal et al., 2011). We were able to show that this protein (isoform “e” Dicer) could negatively affect the miRNA biogenesis pathway in a similar manner to sDCR-1 (Sawh and Duchaine, 2013). Another group has gone on to show that this protein is also expressed in oral cancer cells, and its loss inhibited cellular proliferation and clonogenic potential (Cantini et al., 2014). Several additional reports have identified non-sDCR-1-like truncated proteins with roles in endoRNAi (Flemr et al., 2013) and in apoptosis (Nakagawa et al., 2010).

We also identified and characterized a cluster of phosphorylation sites at the 3’ end of the DCR-1 PAZ domain (Chapter 2). Abolishment of phosphorylation in this cluster leads to impaired exoRNAi activity and severe developmental defects in animals. Importantly, these effects are observed when the phospho-null DCR-1 is expressed over the wild-type *dcr-1* loci, indicating that phospho-null DCR-1 exterts a dominant negative effect on the exoRNAi pathway and physiology. We also showed that

phospho-null DCR-1 has strengthened protein interactions with ERI endoRNAi protein components and RDE-4, which is a member of both the ERI endoRNAi pathway and exoRNAi pathways. The ERI endoRNAi and exoRNAi pathways have been shown to compete with one another for the WAGO clade Argonautes (Duchaine et al., 2006; Yigit et al., 2006). Since we predict that the developmental defects observed in phospho-null DCR-1-expressing animals are due in large part to defects in the miRNA pathway, this could be another example of RNAi pathway competition for a limiting resource. If total amounts of DCR-1 are limiting in some contexts, the phospho-null DCR-1 may be preferentially directed to the endoRNAi pathway at the detriment of the miRNA pathway. Such a scenario is reminiscent of some experiments we performed by enforcing sDCR-1 expression in the muscle (Sawh and Duchaine, 2013). sDCR-1 preferentially binds the miRNA Argonautes ALG-1/2, in a manner that is non-productive for miRNA biogenesis. This reduces the amount of functional full-length DCR-1-ALG-1/2 complex, which is required for miRNA biogenesis, and causes an accumulation of precursor miRNAs. Furthermore, when the exoRNAi pathway is triggered, the inhibition of the miRNA

pathway is exacerbated. Taken together, these findings lead us to propose that the amount of DCR-1 is limiting, and that the miRNA pathway also competes with the endo/exoRNAi pathways. Post-translational modification in the ST cluster may therefore be one way of directing DCR-1 towards the miRNA pathway. Moreover, it is likely that other modifications play a role in both the RNAi-dependent and RNAi-independent roles of DCR-1.

In summary, the work presented in this thesis helps to mechanistically refine the complex pathways of RNAi through study of DCR-1 regulation and the mechanism of nuclear RNAi. We demonstrate that the field is one still ripe for discovery, and complementary approaches in a variety of systems are needed to assess the impact of RNAi on gene regulatory networks.

Appendix 1: Supplementary Information to Chapter 1

Originally published in the journal Cell Reports, August 2013, Volume 4, Issue 3

Sawh A. N. and Duchaine T. F.

Received: September 4, 2012; Received in revised form: June 25, 2013; Accepted: July 12, 2013; Published Online: August 08, 2013

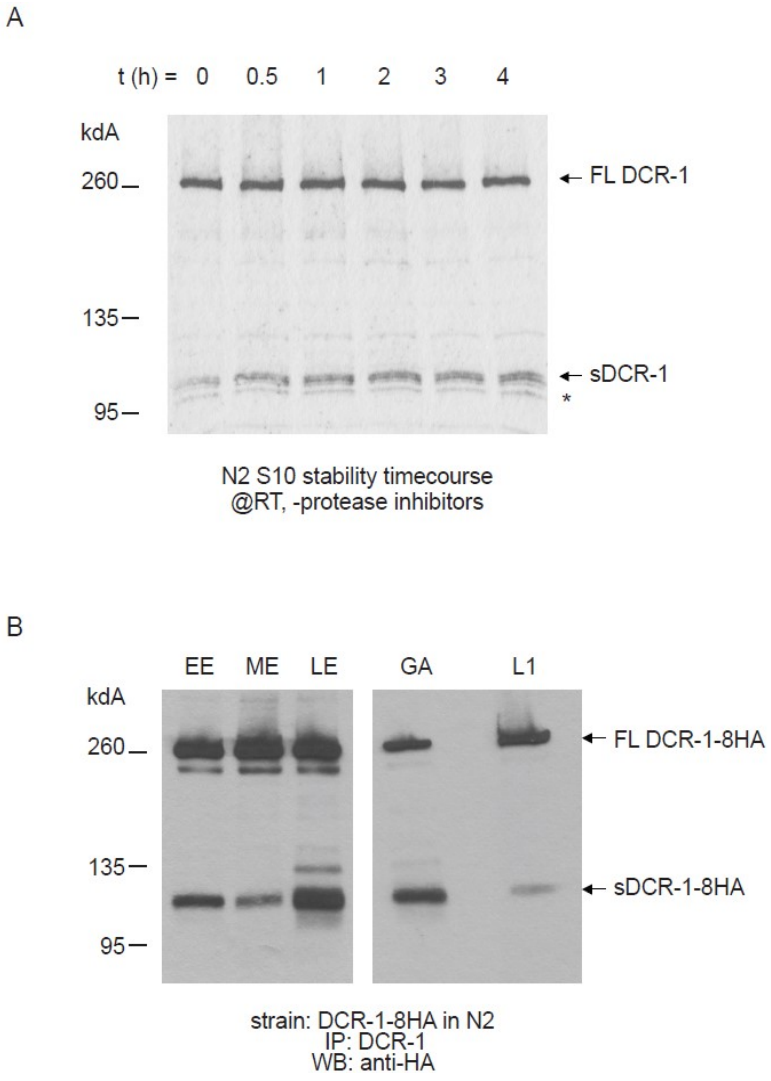


Figure A1.1: Stability and specificity of DCR-1 proteins. (A) Western Blot of endogenous DCR-1 from gravid adults. Extract was prepared in the absence of protease inhibitors, cleared by 10 000xg centrifugation, and kept at room temperature for four hours, with aliquots taken at the time points indicated. The stability of the proteins indicates that FL DCR-1 does not degrade to produce sDCR-1. (B) DCR-1 IPs were done on extracts of

the strain carrying a DCR-1-8HA transgene in the wild-type background, and blotted for HA at different developmental stages. The L1 sample contains FL DCR-1-8HA and sDCR-1-8HA, and not the ** band in Figure 1.2B, indicating that ** band is not sDCR-1, and is likely non-specific.

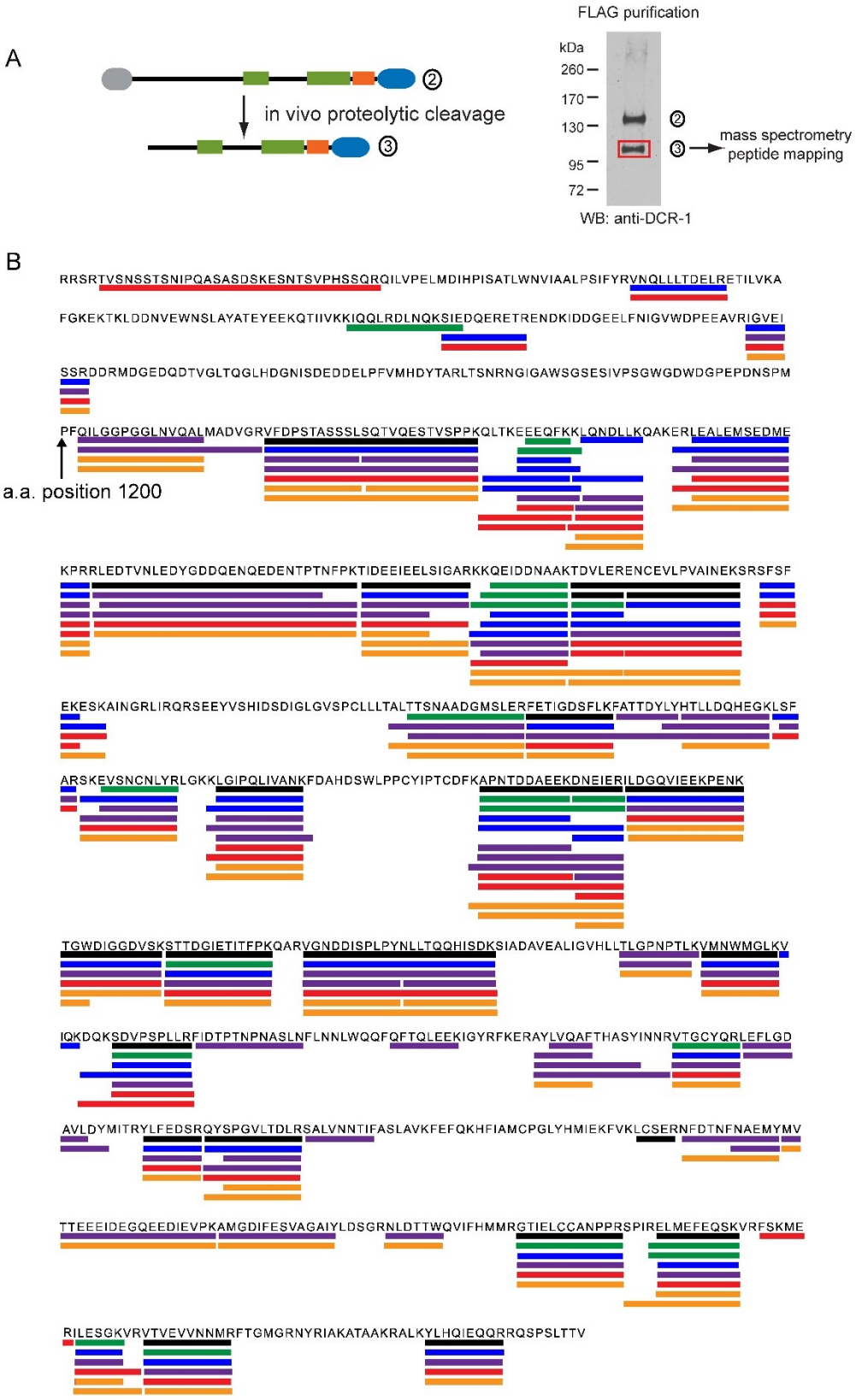

A
②
in vivo proteolytic cleavage
③
FLAG purification
kDa
260
170
130
95
72
②
③
mass spectrometry
peptide mapping
WB: anti-DCR-1
B
RRSRTVSNSSTSNIPQASASDSKESNTSVPHSSQRQILVPELMDIHPISATLWNVIAALPSIFYRVNQLLLTDELRETILVKA
FGKEKTKLDDNVEWNSLAYATEYEEKQTIIVKKIQQLRDLNQKSIEDQERETRENDKIDDGEELFNIGVWDPEEAVRIGVEI
SSRDDRMDGEDQDTVGLTQGLHDGNISDEDDELPFVMHDYTARLTSNRNGIGAWSGSESIVPSGWGDWDGPEPDNSPM
PFQILGGPGGLNVQALMADVGRVFDPSTASSSLSQTVQESTVSPPKQLTKEEEQFKKLQNDLLKQAKERLEALEMSEDME
a.a. position 1200
KPRRLEDTVNLEDYGDDQENQEDENTPTNFPKTIDEEIEELSIGARKKQEIDDNAAKTDVLERENCEVLPVAINEKSRSFSF
EKESKAINGRLIRQRSEEYVSHIDSDIGLGVSPCLLLTALTTSNAADGMSLERFETIGDSFLKFATTDYLYHTLLDQHEGKLSF
ARSKEVSNCNLYRLGKKLGIPQLIVANKFDAHDSWLPPCYIPTCDFKAPNTDDAEEKDNEIERILDGQVIEEKPENK
TGWDIGGDVSKSTTDGIETITFPKQARVGNDDISPLPYNLLTQQHISDKSIADAVEALIGVHLLTLGPNPTLKVMNWMGLKV
IQKDQKSDVPSPLLRFIDTPTNPNASLNFLNNLWQQFQFTQLEEKIGYRFKERAYLVQAFTHASYINNRVTGCYQRLEFLGD
AVLDYMITRYLFEDSRQYSPGVLTDLRSALVNNTIFASLAVKFEFQKHFIAMCPGLYHMIEKFVKLCSERNFDTNFNAEMYMV
TTEEEIDEGQEEDIEVPKAMGDIFESVAGAIYLDSGRNLDTTWQVIFHMMRGTIELCCANPPRSPIRELMEFEQSKVRFSKME
RILESGKVRVTVEVVNNMRFTGMGRNYRIAKATAAKRALKYLHQIEQQRRQSPSLTTV

Figure A1.2: Peptide Mapping of sDCR-1. Large scale FLAG purifications were performed on extracts from animals expressing myo-3::MYC-pre-sDCR-1-3FLAG. (A) The band corresponding to sDCR-1-3FLAG was excised and submitted for peptide sequencing. (B) Peptides obtained are indicated on the sequence of the pre-sDCR-1 construct (starts at amino acid position 957 in DCR-1). Different color peptide bars represent individual samples.

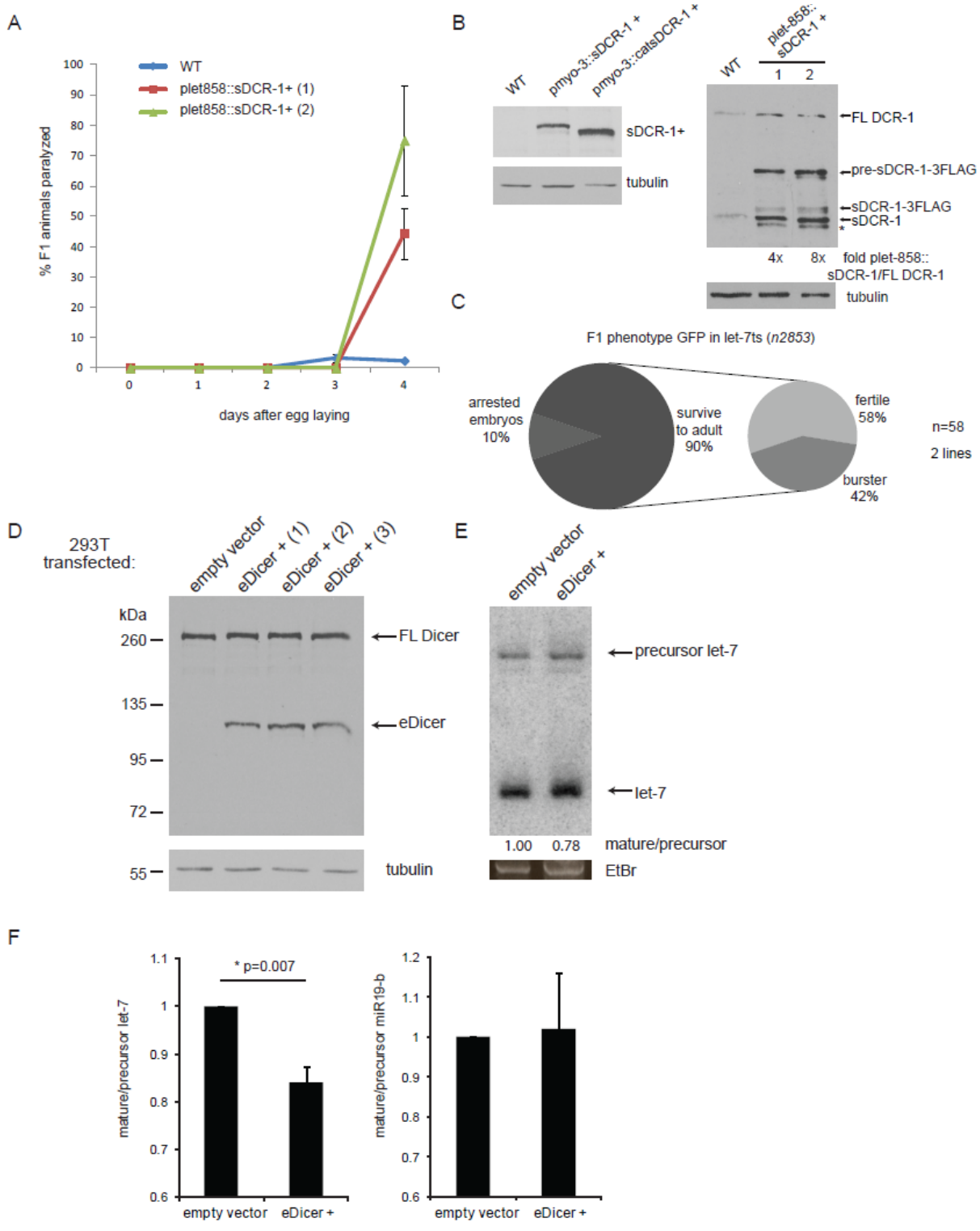

A
WT
plet858::sDCR-1+ (1)
plet858::sDCR-1+ (2)
% F1 animals paralyzed
days after egg laying
B
WT
pmyo-3::sDCR-1 +
pmyo-3::catsDCR-1 +
sDCR-1+
tubulin
plet-858::sDCR-1 +
FL DCR-1
pre-sDCR-1-3FLAG
sDCR-1-3FLAG
sDCR-1
4x
8x
fold plet-858::sDCR-1/FL DCR-1
C
F1 phenotype GFP in let-7ts (n2853)
arrested embryos 10%
survive to adult 90%
fertile 58%
burster 42%
n=58
2 lines
D
293T transfected:
empty vector
eDicer + (1)
eDicer + (2)
eDicer + (3)
kDa
260
135
95
72
55
FL Dicer
eDicer
tubulin
E
empty vector
eDicer +
precursor let-7
let-7
1.00
0.78
mature/precursor
EtBr
F
* p=0.007
mature/precursor let-7
mature/precursor miR19-b
empty vector
eDicer +

Figure A1.3: Expression and phenotype of sDCR1 and eDicer. (A) Time-course of paralysis of WT and plet-858::sDCR-1+ animals on *unc-22* RNAi. Due to a delay in development, sDCR-1+ animals do not reach the assay endpoint (gravid adult) until day 4, whereas WT animals reach gravid adult at day 3. However, since no difference is seen in the level of paralysis in WT animals between day 3 and 4, all animals are scored on day 4 for final analysis (See Figure 1.6). (B) Left, expression comparison of pmyo-3::sDCR-1 and pmyo-3::catsDCR-1 in total protein extracts. pmyo-3::catsDCR-1 migrates lower due to the 1FLAG tag on the C-terminus, as opposed to the 3FLAG tag on pmyo-3::sDCR-1. Right, protein expression of plet-858::sDCR-1 constructs compared to endogenous DCR-1 levels in adults. The fold expression of plet-858::sDCR-1/FL DCR-1 is calculated as the sum of the intensities of the pre-sDCR-1-3FLAG and sDCR-1-3FLAG, over the intensity of FL DCR-1 in the same extract. (C) Phenotype of *n2853* animals injected with GPF alone (See Figure 1.10). (D) HEK293 cells were transfected with a construct encoding eDicer, monitored for protein expression of Dicer, and small RNA analysis of let-7 and miR-19b 50h post-transfection.

Western blot shows approximately equivalent expression of FL and eDicer in transfected cells (3 biological replicates shown). (E and F) Northern blot of let-7 shows a reduction in the ratio of mature/precursor miRNA when eDicer is expressed. miR-19b is unchanged under the same conditions. Data is represented as mean +/- standard deviation, with the values from empty vector samples set to 1. Statistical significance was calculated using independent 2-tailed Student's *t*-test.

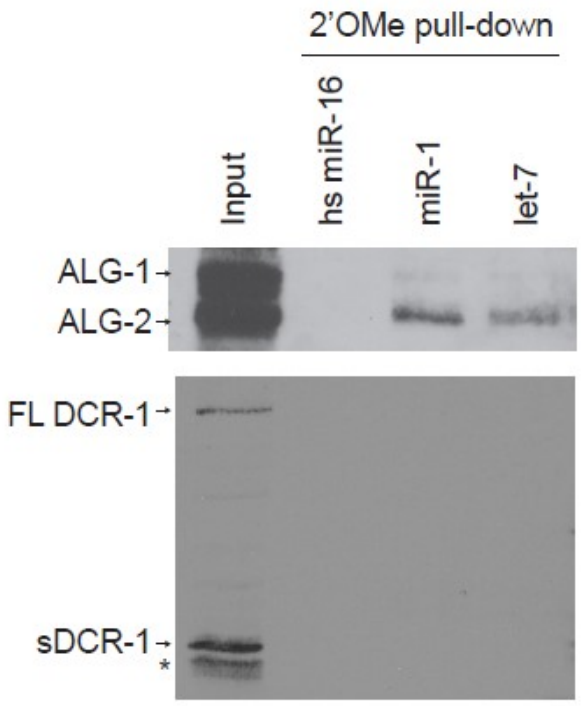


Figure A1.4: 2'O-Methyl pull-down of endogenous miR-1 and let-7 complexes in adult WT animals. Endogenous proteins in complex with miR-1 and let-7 following 2'O-Methyl pull-down of these miRNAs (with human miR-16 as a negative control) are probed for ALG-1/2 (top) and DCR-1 (bottom).

Appendix 2: Turning Dicer on its head

Originally published in the journal Nature Structural & Molecular Biology, April 2012,

Volume 19, Number 4

Sawh A. N. and Duchaine T. F.

Published Online: April 04, 2012

Turning Dicer on its head

Ahilya N Sawh & Thomas F Duchaine

Recent structural analysis of full-length human Dicer supports a new model of the enzyme's domain arrangement and provides a structural basis for many of Dicer's biochemical attributes.

RNA interference (RNAi) pathways collectively exert an extensive regulatory influence on global gene expression in a wide variety of organisms. A cornerstone of RNAi is the highly conserved and specialized eukaryotic RNase Dicer, which acts on double-stranded RNA (dsRNA) substrates from varied origins to produce small interfering RNAs (siRNAs) of discrete sizes, which in turn proceed to direct the sequence specificity of gene silencing through RNAi. The mechanism of Dicer action in a multitude of model systems has been a topic of investigation for over a decade, since its discovery as the crucial enzyme in siRNA generation[1]. Much insight has been gained through structural studies of Dicer. In this issue, Lau *et al.*[2] present a new model for the overall organization of full-length human Dicer. Although it represents a departure from currently accepted notions of Dicer's domain arrangement, the new model provides a plausible explanation for many of the protein's biochemical properties. In particular, certain properties of the helicase domain are addressed that have until now remained perplexing from a structural point of view.

Breaking Dicer down to its domain components (**Fig. 1a**), the C terminus comprises tandem RNase III domains and a dsRNA binding domain (dsRBD), whereas the N terminus contains the domains that make Dicer unique. At its very N terminus is the large helicase domain—thus named for homology over function, as no one has shown so far that it has helicase activity. However, this domain has several intriguing properties: it acts as an autoinhibitory module[3], is required for the production of certain classes of small RNAs[4], contacts precursor miRNA substrates[5] and contributes to the processivity of the enzyme[6,7]. Following the helicase domain is a small domain of unknown function (DUF283), which remains a mysterious entity, although data from plants suggest that DUF283 adopts a noncanonical dsRBD fold and could mediate protein-protein interactions with Dicer cofactors[8]. Next are a structure known as the 'platform'[9], and the PAZ domain, which is thought to play the critical role of docking the 3′ two-nucleotide overhang of the dsRNA substrate into the enzyme[10]. Presumably, it is the presence of these N-terminal domains, and their arrangement in space, that cause Dicer to be specialized to RNAi pathways rather than being a general-purpose RNase III enzyme.

Notably, the information gained from the crystal structure of *Giardia intestinalis* Dicer in 2006 represented a considerable leap forward in our understanding of Dicer function[9]. The highlight of that work was the discovery that the PAZ domain (which binds the 3′ overhang of dsRNA) is separated in space from the catalytic core of the enzyme (the tandem RNase III domains, which form an internal dimer) by a connecting helix; the separation distance corresponds well to the size of the siRNAs produced. The authors proposed that the region between the PAZ and RNase III domains acts as a 'molecular ruler' for the enzyme, thus explaining how Dicer produces small RNAs of discrete size. The *Giardia* structure thus formed the basis for the way we think of siRNA production by Dicer at its core.

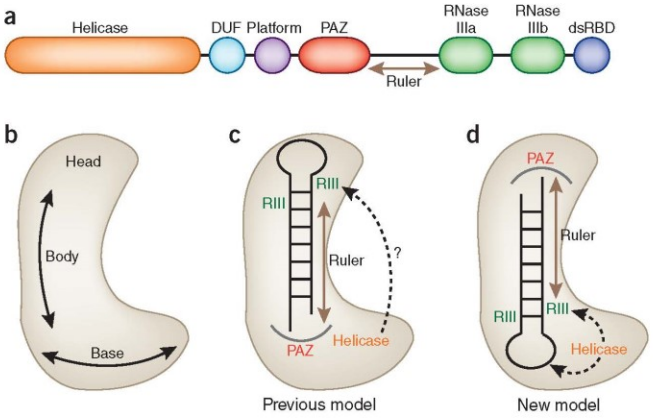


Figure 1 Human Dicer domain architecture. (**a**) Schematic of human Dicer's domain structure. (**b**) Cartoon representation of the shape of full-length human Dicer based on EM reconstructions[11,12]. (**c**) The previous model for the domain architecture presented with miRNA bound. (**d**) Model for the architecture of Dicer based on new data from Lau *et al.*[2]. Note the inversion of positions between the PAZ and RIII domains within the L shape and the proximity of the helicase domain to the pre-miRNA loop and RNase III core (arrows).

Ahilya N. Sawh and Thomas F. Duchaine are in the Goodman Cancer Research Centre and in the Department of Biochemistry, McGill University, Montreal, Quebec, Canada.
e-mail: thomas.duchaine@mcgill.ca

However, because *Giardia* Dicer lacks the domains N-terminal to the PAZ domain, the structure of the helicase domain in higher eukaryotes remained unknown.

The substantial challenge inherent to determining the structure of full-length Dicer is its size, which is upward of 200 kDa in most species. So far, crystallography studies have been limited to analyzing a few domains at a time, but negative-stain EM with single-particle analysis of full-length human Dicer has provided very useful information on the general shape of the molecule[11,12]. The collective data from these EM experiments show that Dicer is an L-shaped molecule, with the catalytic core thought to reside in the 'head' and the PAZ and helicase domains at the 'base', separated in space by the molecular ruler region (**Fig. 1b**,**c**). It is important to note that the favored positioning of the domains within the EM structures remained hypothetical and was not based on direct experimental evidence. With this scheme, we were left wondering how the helicase domain (positioned adjacent to the PAZ and opposite the RNase III domains) could undergo such large-scale rearrangements to contact the miRNA loop, auto-inhibit the activity of the RNase IIIs and instigate processivity on a long dsRNA substrate (**Fig. 1c**).

Lau *et al.*[2] present an alternative model for the arrangement of Dicer's domains, using a novel domain-mapping technique in conjunction with negative-stain EM. The authors use site-specific enzymatic biotinylation followed by streptavadin labeling on multiple points of the AviTag-Dicer protein to determine where the critical domains map within the L-shaped molecule. The labeling results are supported by EM structures of Dicer deletion mutants, which lack portions of the L corresponding to their original location on the wild-type protein. These data show that, although the separation between the PAZ domain and the RNase III core is maintained, the PAZ domain is, surprisingly, positioned in the head, and the RNase III core is positioned in the body near the base of the L (**Fig. 1d**). The authors also propose that the catalytic core in human Dicer is rotated relative to the PAZ domain, compared with *Giardia* Dicer. This rotation shortens the measurement of the ruler and accounts well for the differences in siRNA size between the two species.

Most unexpected is the new positioning of the helicase adjacent to the catalytic core and opposite to the PAZ domain. In this scheme, the previously established biochemical properties of Dicer are consistent with the structural arrangement of the domains. One can easily imagine how the helicase domain can reduce the catalytic efficiency of the enzyme in such close proximity to the RNase III core and contact the pre-miRNA loop as well. Furthermore, the refinement of the lobes of the helicase, which are shown to be shaped like a clamp, allows us to envision a 'catch and feed' motion of Dicer on a long dsRNA substrate, operating in a processive manner (**Fig. 1d**).

Moreover, this new model has important implications for Dicer's interactions with its protein partners. For example, human Argonaute 2 (AGO2), a core component of the RNA-induced silencing complex (RISC), has been shown in biochemical mapping experiments to interact with Dicer's RNase IIIa domain through its PIWI domain[13,14]. AGO2 was also predicted to contact the head and base of the L-shaped Dicer molecule in EM reconstructions of the human RISC-loading complex[12]. Taking into consideration the new data on the domain mapping of Dicer, is it possible that AGO2 forms additional contacts with the PAZ (or platform) domain at the head and the helicase at the base? Biochemically, such interactions have not been detected[13], but it is possible that these contacts may arise transiently and only in the context of the full-length proteins. An exciting next step would be to exploit the domain-mapping technique described by Lau *et al.*[2] to further refine the structure of the human RISC-loading complex. This technology also has the potential to map some of Dicer's unresolved domains, such as the DUF domain, or the interactions with its myriad other protein partners.

COMPETING FINANCIAL INTERESTS
The authors declare no competing financial interests.

1. Bernstein, E., Caudy, A.A., Hammond, S.M. & Hannon, G.J. *Nature* **409**, 363–366 (2001).
2. Lau, P.-W. *et al. Nat. Struct. Mol. Biol.* **19**, 436–440 (2012).
3. Ma, E., MacRae, I.J., Kirsch, J.F. & Doudna, J.A. *J. Mol. Biol.* **380**, 237–243 (2008).
4. Welker, N.C. *et al. RNA* **16**, 893–903 (2010).
5. Tsutsumi, A., Kawamata, T., Izumi, N., Seitz, H. & Tomari, Y. *Nat. Struct. Mol. Biol.* **18**, 1153–1158 (2011).
6. Cenik, E.S. *et al. Mol. Cell* **42**, 172–184 (2011).
7. Welker, N.C. *et al. Mol. Cell* **41**, 589–599 (2011).
8. Qin, H. *et al. RNA* **16**, 474–481 (2010).
9. Macrae, I.J. *et al. Science* **311**, 195–198 (2006).
10. Ma, J.B., Ye, K. & Patel, D.J. *Nature* **429**, 318–322 (2004).
11. Lau, P.W., Potter, C.S., Carragher, B. & MacRae, I.J. *Structure* **17**, 1326–1332 (2009).
12. Wang, H.W. *et al. Nat. Struct. Mol. Biol.* **16**, 1148–1153 (2009).
13. Tahbaz, N. *et al. EMBO Rep.* **5**, 189–194 (2004).
14. Sasaki, T. & Shimizu, N. *Gene* **396**, 312–320 (2007).

Appendix 3: Predicted kinases of the DCR-1 ST cluster

residue		Scansite	PPSP	NetPhosK	KinasPhos
959	S		AURORA-A, AURORA-B, CAK, CHK1/CHK2, DAPK, DNA-PK, GSK-3, MAPKKK, MLCK, NIMA, PAK, PKA, PKC, PKG, PLK, ROCK	PKA, CDC2, PKG	
960	R				
961	T	AMPK,PKA, AKT	AMPK, CaM-I/IV, CAM-II, CHK1/CHK2, CK1, DAPK, IKK, IPL1, MAPKAPK2, MAPKKK, PAK, PDK, PHK, PKA, PKB, PKC, PKG, ROCK, S6K	PKA	PKC, PKA
962	V				
963	S	AKT, CLK2	AMPK, CaM-I/IV, CAM-II, CHK1/CHK2, CK1, DAPK, DNA-PK, GRK, GSK3, IKK, MAPKAPK2, NIMA, PAK, PDK, PHK, PKA, PKB, PKC, PKG, ROCK, S6K	RSK,PKB,PKC,PKA	PKG,IKK,PKB
964	N				
965	S		CK1, GRK, IKK, MAPKKK, NIMA, PLK		IKK
966	S		CaM-I/IV, CK1, DNA-PK, GRK, NIMA, PLK	CDC2	
967	T		CAK, GRK	CDC2	
968	S		CAK, GRK, GSK3, NIMA, PHK	CDC2	
969	N				
970	I				
971	P				
972	Q				
973	A				
974	S		CK1, CK2, DNA-PK, GRK, NIMA, PLK		
975	A				
976	S		CK1, CK2, GRK, PKC	CDC2	ATM
977	D				
978	S		ATM, CK1, CK2, DNA-PK, GRK, IKK, NIMA, PLK	PKC	IKK, CKI
979	K				
980	E				
981	S		CK1, GRK, IKK, MLCK, PKA, PKG, PKR		
982	N				
983	T		CAK, P34CDC2, ROCK		
984	S		CK1, GRK, GSK3, IPL1, NIMA, PAK	PKC	CKI, ATM
985	V				
986	P				
987	H				
988	S		CAK, GRK, NIMA, PKC		CDK
989	S		ATM, CAK, DNA-PK, NIMA, PHK, PLK	DNAPK, ATM	ATM

Table A3: Sites within the DCR-1 ST cluster and their predicted kinases.

Scansite3.0: (Obenauer et al., 2003)

PPSP: ppsp.biocuckoo.org

NetPhosK: (Blom et al., 2004)

KinasPhos: (Huang et al., 2005)

Appendix 4: List of RNAi components in *C. elegans*

exoRNAi pathway	
Length	~22nt primary, 22G secondary
Primary biogenesis	DCR-1, RDE-4
Primary AGO	RDE-1
Secondary biogenesis	DRH-3, RRF-1 (RdRP), EGO-1 (RdRP), EKL-1
Secondary AGO	Secondary AGOs (WAGOs)
Effectors	RDE-3/8/10/11, NRDE-1/2/4
Mechanism of Action	Post-transcriptional, transcriptional?
Genes regulated	any targeted
ERI endoRNAi pathway	
Length & 5' preference	26G primary, 22G secondary
Primary biogenesis	DCR-1, ERI-1b, ERI-3, RRF-3 (RdRP), DRH-3, ERI-5, RDE-4
Primary AGO	ERGO-1, ALG-3/4 (sperm)
Primary siRNA accumulation/modification	ERI-9, ERI-6/7, MUT-16, MUT-15, MUT-2, MUT-7 HENN-1
Secondary biogenesis	DRH-3, RRF-1 (RdRP), EGO-1 (RdRP), EKL-1
Secondary AGO	Secondary AGOs (WAGOs)
Effectors	RDE-10/11, RSD-2, RSD-6, HAF-6, NRDE-1/2/4
Mechanism of Action	Post-transcriptional, transcriptional
Genes regulated	Duplicated genes, lincRNA genes, protein-coding, non-coding loci
miRNA pathway	
Length	~22nt
Primary biogenesis	DRSH-1, PASH-1, DCR-1, ALG-1/2
Primary AGO	ALG-1/2
Secondary AGO	-
Secondary biogenesis	-
Effectors	AIN-1/2, PAB-1/2, deadenylase, decapping, decay complexes
Mechanism of Action	Post-transcriptional
Genes regulated	Protein-coding

Table A4: List of key genes in *C. elegans* DCR-1-dependent RNAi pathways

References

Abrahante, J.E., Daul, A.L., Li, M., Volk, M.L., Tennessen, J.M., Miller, E.A., and Rougvie, A.E. (2003). The Caenorhabditis elegans hunchback-like gene lin-57/hbl-1 controls developmental time and is regulated by microRNAs. Developmental cell *4*, 625-637.

Aird, D., Ross, M.G., Chen, W.S., Danielsson, M., Fennell, T., Russ, C., Jaffe, D.B., Nusbaum, C., and Gnirke, A. (2011). Analyzing and minimizing PCR amplification bias in Illumina sequencing libraries. Genome biology *12*, R18.

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

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

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

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

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

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

Ashe, A., Belicard, T., Le Pen, J., Sarkies, P., Frezal, L., Lehrbach, N.J., Felix, M.A., and Miska, E.A. (2013). A deletion polymorphism in the Caenorhabditis elegans RIG-I homolog disables viral RNA dicing and antiviral immunity. eLife *2*, e00994.

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

Baek, D., Villen, J., Shin, C., Camargo, F.D., Gygi, S.P., and Bartel, D.P. (2008). The impact of microRNAs on protein output. Nature *455*, 64-71.
Bartel, D.P. (2009). MicroRNAs: target recognition and regulatory functions. Cell *136*, 215-233.
Batista, P.J., Ruby, J.G., Claycomb, J.M., Chiang, R., Fahlgren, N., Kasschau, K.D., Chaves, D.A., Gu, W., Vasale, J.J., Duan, S.*, et al.* (2008). PRG-1 and 21U-RNAs interact to form the piRNA complex required for fertility in C. elegans. Molecular cell *31*, 67-78.
Baugh, L.R., Hill, A.A., Slonim, D.K., Brown, E.L., and Hunter, C.P. (2003). Composition and dynamics of the Caenorhabditis elegans early embryonic transcriptome. Development *130*, 889-900.
Bazzini, A.A., Lee, M.T., and Giraldez, A.J. (2012). Ribosome profiling shows that miR-430 reduces translation before causing mRNA decay in zebrafish. Science *336*, 233-237.
Behm-Ansmant, I., Rehwinkel, J., Doerks, T., Stark, A., Bork, P., and Izaurralde, E. (2006). mRNA degradation by miRNAs and GW182 requires both CCR4:NOT deadenylase and DCP1:DCP2 decapping complexes. Genes & development *20*, 1885-1898.
Berger, S.L. (2007). The complex language of chromatin regulation during transcription. Nature *447*, 407-412.
Bernstein, E., Caudy, A.A., Hammond, S.M., and Hannon, G.J. (2001). Role for a bidentate ribonuclease in the initiation step of RNA interference. Nature *409*, 363-366.
Bernstein, E., Kim, S.Y., Carmell, M.A., Murchison, E.P., Alcorn, H., Li, M.Z., Mills, A.A., Elledge, S.J., Anderson, K.V., and Hannon, G.J. (2003). Dicer is essential for mouse development. Nature genetics *35*, 215-217.
Blom, N., Sicheritz-Ponten, T., Gupta, R., Gammeltoft, S., and Brunak, S. (2004). Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence. Proteomics *4*, 1633-1649.
Boi, S., Solda, G., and Tenchini, M.L. (2004). Shedding light on the dark side of the genome: Overlapping genes in higher eukaryotes. Curr Genomics *5*, 509-524.

Boland, A., Huntzinger, E., Schmidt, S., Izaurralde, E., and Weichenrieder, O. (2011). Crystal structure of the MID-PIWI lobe of a eukaryotic Argonaute protein. Proceedings of the National Academy of Sciences of the United States of America *108*, 10466-10471.
Boland, A., Tritschler, F., Heimstadt, S., Izaurralde, E., and Weichenrieder, O. (2010). Crystal structure and ligand binding of the MID domain of a eukaryotic Argonaute protein. EMBO reports *11*, 522-527.
Bologna, N.G., and Voinnet, O. (2014). The diversity, biogenesis, and activities of endogenous silencing small RNAs in Arabidopsis. Annual review of plant biology *65*, 473-503.
Borsani, O., Zhu, J., Verslues, P.E., Sunkar, R., and Zhu, J.K. (2005). Endogenous siRNAs derived from a pair of natural cis-antisense transcripts regulate salt tolerance in Arabidopsis. Cell *123*, 1279-1291.
Boyerinas, B., Park, S.M., Shomron, N., Hedegaard, M.M., Vinther, J., Andersen, J.S., Feig, C., Xu, J., Burge, C.B., and Peter, M.E. (2008). Identification of let-7-regulated oncofetal genes. Cancer research *68*, 2587-2591.
Brennecke, J., Aravin, A.A., Stark, A., Dus, M., Kellis, M., Sachidanandam, R., and Hannon, G.J. (2007). Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. Cell *128*, 1089-1103.
Brennecke, J., Hipfner, D.R., Stark, A., Russell, R.B., and Cohen, S.M. (2003). bantam encodes a developmentally regulated microRNA that controls cell proliferation and regulates the proapoptotic gene hid in Drosophila. Cell *113*, 25-36.
Buckley, B.A., Burkhart, K.B., Gu, S.G., Spracklin, G., Kershner, A., Fritz, H., Kimble, J., Fire, A., and Kennedy, S. (2012). A nuclear Argonaute promotes multigenerational epigenetic inheritance and germline immortality. Nature *489*, 447-451.
Buker, S.M., Iida, T., Buhler, M., Villen, J., Gygi, S.P., Nakayama, J., and Moazed, D. (2007). Two different Argonaute complexes are required for siRNA generation and heterochromatin assembly in fission yeast. Nature structural & molecular biology *14*, 200-207.

Burkhart, K.B., Guang, S., Buckley, B.A., Wong, L., Bochner, A.F., and Kennedy, S. (2011). A pre-mRNA-associating factor links endogenous siRNAs to chromatin regulation. PLoS genetics *7*, e1002249.
Burton, N.O., Burkhart, K.B., and Kennedy, S. (2011). Nuclear RNAi maintains heritable gene silencing in Caenorhabditis elegans. Proceedings of the National Academy of Sciences of the United States of America *108*, 19683-19688.
Cantini, L.P., Andino, L.M., Attaway, C.C., Butler, B., Dumitriu, A., Blackshaw, A., and Jakymiw, A. (2014). Identification and characterization of Dicer1e, a Dicer1 protein variant, in oral cancer cells. Molecular cancer *13*, 190.
Catalanotto, C., Azzalin, G., Macino, G., and Cogoni, C. (2000). Gene silencing in worms and fungi. Nature *404*, 245.
Cenik, E.S., Fukunaga, R., Lu, G., Dutcher, R., Wang, Y., Tanaka Hall, T.M., and Zamore, P.D. (2011). Phosphate and R2D2 restrict the substrate specificity of Dicer-2, an ATP-driven ribonuclease. Molecular cell *42*, 172-184.
Cheloufi, S., Dos Santos, C.O., Chong, M.M., and Hannon, G.J. (2010). A dicer-independent miRNA biogenesis pathway that requires Ago catalysis. Nature *465*, 584-589.
Chen, C.Z., Li, L., Lodish, H.F., and Bartel, D.P. (2004). MicroRNAs modulate hematopoietic lineage differentiation. Science *303*, 83-86.
Chen, J.D., and Evans, R.M. (1995). A transcriptional co-repressor that interacts with nuclear hormone receptors. Nature *377*, 454-457.
Chendrimada, T.P., Gregory, R.I., Kumaraswamy, E., Norman, J., Cooch, N., Nishikura, K., and Shiekhattar, R. (2005). TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. Nature *436*, 740-744.
Cifuentes, D., Xue, H., Taylor, D.W., Patnode, H., Mishima, Y., Cheloufi, S., Ma, E., Mane, S., Hannon, G.J., Lawson, N.D.*, et al.* (2010). A novel miRNA processing pathway independent of Dicer requires Argonaute2 catalytic activity. Science *328*, 1694-1698.

Clark, S.G., Lu, X., and Horvitz, H.R. (1994). The Caenorhabditis elegans locus lin-15, a negative regulator of a tyrosine kinase signaling pathway, encodes two different proteins. Genetics *137*, 987-997.
Cogoni, C., and Macino, G. (1997). Isolation of quelling-defective (qde) mutants impaired in posttranscriptional transgene-induced gene silencing in Neurospora crassa. Proceedings of the National Academy of Sciences of the United States of America *94*, 10233-10238.
Colmenares, S.U., Buker, S.M., Buhler, M., Dlakic, M., and Moazed, D. (2007). Coupling of double-stranded RNA synthesis and siRNA generation in fission yeast RNAi. Molecular cell *27*, 449-461.
Conine, C.C., Batista, P.J., Gu, W., Claycomb, J.M., Chaves, D.A., Shirayama, M., and Mello, C.C. (2010). Argonautes ALG-3 and ALG-4 are required for spermatogenesis-specific 26G-RNAs and thermotolerant sperm in Caenorhabditis elegans. Proceedings of the National Academy of Sciences of the United States of America *107*, 3588-3593.
Conine, C.C., Moresco, J.J., Gu, W., Shirayama, M., Conte, D., Jr., Yates, J.R., 3rd, and Mello, C.C. (2013). Argonautes promote male fertility and provide a paternal memory of germline gene expression in C. elegans. Cell *155*, 1532-1544.
Cox, D.N., Chao, A., Baker, J., Chang, L., Qiao, D., and Lin, H. (1998). A novel class of evolutionarily conserved genes defined by piwi are essential for stem cell self-renewal. Genes & development *12*, 3715-3727.
Czech, B., Malone, C.D., Zhou, R., Stark, A., Schlingeheyde, C., Dus, M., Perrimon, N., Kellis, M., Wohlschlegel, J.A., Sachidanandam, R.*, et al.* (2008). An endogenous small interfering RNA pathway in Drosophila. Nature *453*, 798-802.
Daniels, S.M., Melendez-Pena, C.E., Scarborough, R.J., Daher, A., Christensen, H.S., El Far, M., Purcell, D.F., Laine, S., and Gatignol, A. (2009). Characterization of the TRBP domain required for dicer interaction and function in RNA interference. BMC molecular biology *10*, 38.
Das, P.P., Bagijn, M.P., Goldstein, L.D., Woolford, J.R., Lehrbach, N.J., Sapetschnig, A., Buhecha, H.R., Gilchrist, M.J., Howe, K.L., Stark, R.*, et al.* (2008). Piwi and piRNAs

act upstream of an endogenous siRNA pathway to suppress Tc3 transposon mobility in the Caenorhabditis elegans germline. Molecular cell *31*, 79-90.
Derman, E., Krauter, K., Walling, L., Weinberger, C., Ray, M., and Darnell, J.E., Jr. (1981). Transcriptional control in the production of liver-specific mRNAs. Cell *23*, 731-739.
Diederichs, S., Jung, S., Rothenberg, S.M., Smolen, G.A., Mlody, B.G., and Haber, D.A. (2008). Coexpression of Argonaute-2 enhances RNA interference toward perfect match binding sites. Proceedings of the National Academy of Sciences of the United States of America *105*, 9284-9289.
Ding, L., Spencer, A., Morita, K., and Han, M. (2005). The developmental timing regulator AIN-1 interacts with miRISCs and may target the argonaute protein ALG-1 to cytoplasmic P bodies in C. elegans. Molecular cell *19*, 437-447.
Ding, X.C., and Grosshans, H. (2009). Repression of C. elegans microRNA targets at the initiation level of translation requires GW182 proteins. The EMBO journal *28*, 213-222.
Djuranovic, S., Nahvi, A., and Green, R. (2012). miRNA-mediated gene silencing by translational repression followed by mRNA deadenylation and decay. Science *336*, 237-240.
Dombecki, C.R., Chiang, A.C., Kang, H.J., Bilgir, C., Stefanski, N.A., Neva, B.J., Klerkx, E.P., and Nabeshima, K. (2011). The chromodomain protein MRG-1 facilitates SC-independent homologous pairing during meiosis in Caenorhabditis elegans. Developmental cell *21*, 1092-1103.
Doyle, M., Badertscher, L., Jaskiewicz, L., Guttinger, S., Jurado, S., Hugenschmidt, T., Kutay, U., and Filipowicz, W. (2013). The double-stranded RNA binding domain of human Dicer functions as a nuclear localization signal. Rna *19*, 1238-1252.
Doyle, M., Jaskiewicz, L., Filipowicz, W. (2012). Chapter One - Dicer Proteins and Their Role in Gene Silencing Pathways. In The Enzymes, Eukaryotic RNases and their Partners in RNA Degradation and Biogenesis, Part B (Elsevier), pp. 1-35.
Drake, M., Furuta, T., Suen, K.M., Gonzalez, G., Liu, B., Kalia, A., Ladbury, J.E., Fire, A.Z., Skeath, J.B., and Arur, S. (2014). A requirement for ERK-dependent Dicer

phosphorylation in coordinating oocyte-to-embryo transition in C. elegans. Developmental cell *31*, 614-628.
Du, Z., Lee, J.K., Tjhen, R., Stroud, R.M., and James, T.L. (2008). Structural and biochemical insights into the dicing mechanism of mouse Dicer: a conserved lysine is critical for dsRNA cleavage. Proceedings of the National Academy of Sciences of the United States of America *105*, 2391-2396.
Duchaine, T.F., Wohlschlegel, J.A., Kennedy, S., Bei, Y., Conte, D., Jr., Pang, K., Brownell, D.R., Harding, S., Mitani, S., Ruvkun, G.*, et al.* (2006). Functional proteomics reveals the biochemical niche of C. elegans DCR-1 in multiple small-RNA-mediated pathways. Cell *124*, 343-354.
Elbashir, S.M., Lendeckel, W., and Tuschl, T. (2001). RNA interference is mediated by 21- and 22-nucleotide RNAs. Genes & development *15*, 188-200.
Elmayan, T., Balzergue, S., Beon, F., Bourdon, V., Daubremet, J., Guenet, Y., Mourrain, P., Palauqui, J.C., Vernhettes, S., Vialle, T.*, et al.* (1998). Arabidopsis mutants impaired in cosuppression. The Plant cell *10*, 1747-1758.
Emmerth, S., Schober, H., Gaidatzis, D., Roloff, T., Jacobeit, K., and Buhler, M. (2010). Nuclear retention of fission yeast dicer is a prerequisite for RNAi-mediated heterochromatin assembly. Developmental cell *18*, 102-113.
Enright, A.J., John, B., Gaul, U., Tuschl, T., Sander, C., and Marks, D.S. (2003). MicroRNA targets in Drosophila. Genome biology *5*, R1.
Eulalio, A., Huntzinger, E., Nishihara, T., Rehwinkel, J., Fauser, M., and Izaurralde, E. (2009). Deadenylation is a widespread effect of miRNA regulation. Rna *15*, 21-32.
Euling, S., and Ambros, V. (1996). Reversal of cell fate determination in Caenorhabditis elegans vulval development. Development *122*, 2507-2515.
Fagard, M., Boutet, S., Morel, J.B., Bellini, C., and Vaucheret, H. (2000). AGO1, QDE-2, and RDE-1 are related proteins required for post-transcriptional gene silencing in plants, quelling in fungi, and RNA interference in animals. Proceedings of the National Academy of Sciences of the United States of America *97*, 11650-11654.
Felix, M.A., Ashe, A., Piffaretti, J., Wu, G., Nuez, I., Belicard, T., Jiang, Y., Zhao, G., Franz, C.J., Goldstein, L.D.*, et al.* (2011). Natural and experimental infection of

Caenorhabditis nematodes by novel viruses related to nodaviruses. PLoS biology *9*, e1000586.
Fire, A., and Waterston, R.H. (1989). Proper expression of myosin genes in transgenic nematodes. The EMBO journal *8*, 3419-3428.
Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E., and Mello, C.C. (1998). Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature *391*, 806-811.
Flemr, M., Malik, R., Franke, V., Nejepinska, J., Sedlacek, R., Vlahovicek, K., and Svoboda, P. (2013). A retrotransposon-driven dicer isoform directs endogenous small interfering RNA production in mouse oocytes. Cell *155*, 807-816.
Frank, F., Sonenberg, N., and Nagar, B. (2010). Structural basis for 5'-nucleotide base-specific recognition of guide RNA by human AGO2. Nature *465*, 818-822.
Friedman, R.C., Farh, K.K., Burge, C.B., and Bartel, D.P. (2009). Most mammalian mRNAs are conserved targets of microRNAs. Genome research *19*, 92-105.
Fukunaga, R., Han, B.W., Hung, J.H., Xu, J., Weng, Z., and Zamore, P.D. (2012). Dicer partner proteins tune the length of mature miRNAs in flies and mammals. Cell *151*, 533-546.
Fuller-Pace, F.V. (2006). DExD/H box RNA helicases: multifunctional proteins with important roles in transcriptional regulation. Nucleic acids research *34*, 4206-4215.
Galiana-Arnoux, D., Dostert, C., Schneemann, A., Hoffmann, J.A., and Imler, J.L. (2006). Essential function in vivo for Dicer-2 in host defense against RNA viruses in drosophila. Nature immunology *7*, 590-597.
Gan, J., Shaw, G., Tropea, J.E., Waugh, D.S., Court, D.L., and Ji, X. (2008). A stepwise model for double-stranded RNA processing by ribonuclease III. Molecular microbiology *67*, 143-154.
Gent, J.I., Lamm, A.T., Pavelec, D.M., Maniar, J.M., Parameswaran, P., Tao, L., Kennedy, S., and Fire, A.Z. (2010). Distinct phases of siRNA synthesis in an endogenous RNAi pathway in C. elegans soma. Molecular cell *37*, 679-689.

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

Giraldez, A.J., Mishima, Y., Rihel, J., Grocock, R.J., Van Dongen, S., Inoue, K., Enright, A.J., and Schier, A.F. (2006). Zebrafish MiR-430 promotes deadenylation and clearance of maternal mRNAs. Science *312*, 75-79.

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

Greenberg, M.E., and Ziff, E.B. (1984). Stimulation of 3T3 cells induces transcription of the c-fos proto-oncogene. Nature *311*, 433-438.

Gregory, R.I., Chendrimada, T.P., Cooch, N., and Shiekhattar, R. (2005). Human RISC couples microRNA biogenesis and posttranscriptional gene silencing. Cell *123*, 631-640.

Grelier, G., Voirin, N., Ay, A.S., Cox, D.G., Chabaud, S., Treilleux, I., Leon-Goddard, S., Rimokh, R., Mikaelian, I., Venoux, C.*, et al.* (2009). Prognostic value of Dicer expression in human breast cancers and association with the mesenchymal phenotype. British journal of cancer *101*, 673-683.

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

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

Grosshans, H., Johnson, T., Reinert, K.L., Gerstein, M., and Slack, F.J. (2005). The temporal patterning microRNA let-7 regulates several transcription factors at the larval to adult transition in C. elegans. Developmental cell *8*, 321-330.

Gu, S., Jin, L., Zhang, Y., Huang, Y., Zhang, F., Valdmanis, P.N., and Kay, M.A. (2012a). The loop position of shRNAs and pre-miRNAs is critical for the accuracy of dicer processing in vivo. Cell *151*, 900-911.

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

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

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

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

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

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

Gurtan, A.M., Lu, V., Bhutkar, A., and Sharp, P.A. (2012). In vivo structure-function analysis of human Dicer reveals directional processing of precursor miRNAs. Rna *18*, 1116-1122.

Haase, A.D., Fenoglio, S., Muerdter, F., Guzzardo, P.M., Czech, B., Pappin, D.J., Chen, C., Gordon, A., and Hannon, G.J. (2010). Probing the initiation and effector phases of the somatic piRNA pathway in Drosophila. Genes & development *24*, 2499-2504.

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

Hammond, S.M., Bernstein, E., Beach, D., and Hannon, G.J. (2000). An RNA-directed nuclease mediates post-transcriptional gene silencing in Drosophila cells. Nature *404*, 293-296.
Hammond, S.M., Boettcher, S., Caudy, A.A., Kobayashi, R., and Hannon, G.J. (2001). Argonaute2, a link between genetic and biochemical analyses of RNAi. Science *293*, 1146-1150.
Han, J., Pedersen, J.S., Kwon, S.C., Belair, C.D., Kim, Y.K., Yeom, K.H., Yang, W.Y., Haussler, D., Blelloch, R., and Kim, V.N. (2009). Posttranscriptional crossregulation between Drosha and DGCR8. Cell *136*, 75-84.
Hannon, G.J. (2002). RNA interference. Nature *418*, 244-251.
Hauptmann, J., Dueck, A., Harlander, S., Pfaff, J., Merkl, R., and Meister, G. (2013). Turning catalytically inactive human Argonaute proteins into active slicer enzymes. Nature structural & molecular biology *20*, 814-817.
He, L., and Hannon, G.J. (2004). MicroRNAs: small RNAs with a big role in gene regulation. Nature reviews. Genetics *5*, 522-531.
Heinzel, T., Lavinsky, R.M., Mullen, T.M., Soderstrom, M., Laherty, C.D., Torchia, J., Yang, W.M., Brard, G., Ngo, S.D., Davie, J.R.*, et al.* (1997). A complex containing N-CoR, mSin3 and histone deacetylase mediates transcriptional repression. Nature *387*, 43-48.
Heo, I., Ha, M., Lim, J., Yoon, M.J., Park, J.E., Kwon, S.C., Chang, H., and Kim, V.N. (2012). Mono-uridylation of pre-microRNA as a key step in the biogenesis of group II let-7 microRNAs. Cell *151*, 521-532.
Heo, I., Joo, C., Cho, J., Ha, M., Han, J., and Kim, V.N. (2008). Lin28 mediates the terminal uridylation of let-7 precursor MicroRNA. Molecular cell *32*, 276-284.
Hinkal, G.W., Grelier, G., Puisieux, A., and Moyret-Lalle, C. (2011). Complexity in the regulation of Dicer expression: Dicer variant proteins are differentially expressed in epithelial and mesenchymal breast cancer cells and decreased during EMT. British journal of cancer *104*, 387-388.
Hodgkin, J. (1983). X-Chromosome Dosage and Gene-Expression in Caenorhabditis-Elegans - 2 Unusual Dumpy Genes. Mol Gen Genet *192*, 452-458.

Horlein, A.J., Naar, A.M., Heinzel, T., Torchia, J., Gloss, B., Kurokawa, R., Ryan, A., Kamei, Y., Soderstrom, M., Glass, C.K.*, et al.* (1995). Ligand-independent repression by the thyroid hormone receptor mediated by a nuclear receptor co-repressor. Nature *377*, 397-404.

Hsu, D.R., and Meyer, B.J. (1994). The dpy-30 gene encodes an essential component of the Caenorhabditis elegans dosage compensation machinery. Genetics *137*, 999-1018.

Huang, H.D., Lee, T.Y., Tzeng, S.W., and Horng, J.T. (2005). KinasePhos: a web tool for identifying protein kinase-specific phosphorylation sites. Nucleic acids research *33*, W226-229.

Huang, L.S., Tzou, P., and Sternberg, P.W. (1994). The lin-15 locus encodes two negative regulators of Caenorhabditis elegans vulval development. Molecular biology of the cell *5*, 395-411.

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

Ipsaro, J.J., Haase, A.D., Knott, S.R., Joshua-Tor, L., and Hannon, G.J. (2012). The structural biochemistry of Zucchini implicates it as a nuclease in piRNA biogenesis. Nature *491*, 279-283.

Jen, C.H., Michalopoulos, I., Westhead, D.R., and Meyer, P. (2005). Natural antisense transcripts with coding capacity in Arabidopsis may have a regulatory role that is not linked to double-stranded RNA degradation. Genome biology *6*, R51.

John, B., Enright, A.J., Aravin, A., Tuschl, T., Sander, C., and Marks, D.S. (2004). Human MicroRNA targets. PLoS biology *2*, e363.

Johnson, S.M., Grosshans, H., Shingara, J., Byrom, M., Jarvis, R., Cheng, A., Labourier, E., Reinert, K.L., Brown, D., and Slack, F.J. (2005). RAS is regulated by the let-7 microRNA family. Cell *120*, 635-647.

Johnston, R.J., and Hobert, O. (2003). A microRNA controlling left/right neuronal asymmetry in Caenorhabditis elegans. Nature *426*, 845-849.

Jones, L., Ratcliff, F., and Baulcombe, D.C. (2001). RNA-directed transcriptional gene silencing in plants can be inherited independently of the RNA trigger and requires Met1 for maintenance. Current biology : CB *11*, 747-757.
Kanamoto, T., Terada, K., Yoshikawa, H., and Furukawa, T. (2006). Cloning and regulation of the vertebrate homologue of lin-41 that functions as a heterochronic gene in Caenorhabditis elegans. Developmental dynamics : an official publication of the American Association of Anatomists *235*, 1142-1149.
Karube, Y., Tanaka, H., Osada, H., Tomida, S., Tatematsu, Y., Yanagisawa, K., Yatabe, Y., Takamizawa, J., Miyoshi, S., Mitsudomi, T.*, et al.* (2005). Reduced expression of Dicer associated with poor prognosis in lung cancer patients. Cancer science *96*, 111-115.
Kasschau, K.D., Fahlgren, N., Chapman, E.J., Sullivan, C.M., Cumbie, J.S., Givan, S.A., and Carrington, J.C. (2007). Genome-wide profiling and analysis of Arabidopsis siRNAs. PLoS biology *5*, e57.
Kawamura, Y., Saito, K., Kin, T., Ono, Y., Asai, K., Sunohara, T., Okada, T.N., Siomi, M.C., and Siomi, H. (2008). Drosophila endogenous small RNAs bind to Argonaute 2 in somatic cells. Nature *453*, 793-797.
Kawaoka, S., Izumi, N., Katsuma, S., and Tomari, Y. (2011). 3' end formation of PIWI-interacting RNAs in vitro. Molecular cell *43*, 1015-1022.
Kennedy, S., Wang, D., and Ruvkun, G. (2004). A conserved siRNA-degrading RNase negatively regulates RNA interference in C. elegans. Nature *427*, 645-649.
Kennerdell, J.R., and Carthew, R.W. (1998). Use of dsRNA-mediated genetic interference to demonstrate that frizzled and frizzled 2 act in the wingless pathway. Cell *95*, 1017-1026.
Ketting, R.F., Fischer, S.E., Bernstein, E., Sijen, T., Hannon, G.J., and Plasterk, R.H. (2001). Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in C. elegans. Genes & development *15*, 2654-2659.
Ketting, R.F., Haverkamp, T.H., van Luenen, H.G., and Plasterk, R.H. (1999). Mut-7 of C. elegans, required for transposon silencing and RNA interference, is a homolog of Werner syndrome helicase and RNaseD. Cell *99*, 133-141.

Khvorova, A., Reynolds, A., and Jayasena, S.D. (2003). Functional siRNAs and miRNAs exhibit strand bias. Cell *115*, 209-216.
Kim, K., Lee, Y.S., and Carthew, R.W. (2007). Conversion of pre-RISC to holo-RISC by Ago2 during assembly of RNAi complexes. Rna *13*, 22-29.
Kim, V.N., Han, J., and Siomi, M.C. (2009). Biogenesis of small RNAs in animals. Nature reviews. Molecular cell biology *10*, 126-139.
Kiriakidou, M., Nelson, P.T., Kouranov, A., Fitziev, P., Bouyioukos, C., Mourelatos, Z., and Hatzigeorgiou, A. (2004). A combined computational-experimental approach predicts human microRNA targets. Genes & development *18*, 1165-1178.
Knight, S.W., and Bass, B.L. (2001). A role for the RNase III enzyme DCR-1 in RNA interference and germ line development in Caenorhabditis elegans. Science *293*, 2269-2271.
Kuhn, C.D., and Joshua-Tor, L. (2013). Eukaryotic Argonautes come into focus. Trends in biochemical sciences *38*, 263-271.
Kwak, H., Fuda, N.J., Core, L.J., and Lis, J.T. (2013). Precise maps of RNA polymerase reveal how promoters direct initiation and pausing. Science *339*, 950-953.
Lagos-Quintana, M., Rauhut, R., Lendeckel, W., and Tuschl, T. (2001). Identification of novel genes coding for small expressed RNAs. Science *294*, 853-858.
Lagos-Quintana, M., Rauhut, R., Meyer, J., Borkhardt, A., and Tuschl, T. (2003). New microRNAs from mouse and human. Rna *9*, 175-179.
Lagos-Quintana, M., Rauhut, R., Yalcin, A., Meyer, J., Lendeckel, W., and Tuschl, T. (2002). Identification of tissue-specific microRNAs from mouse. Current biology : CB *12*, 735-739.
Laherty, C.D., Billin, A.N., Lavinsky, R.M., Yochum, G.S., Bush, A.C., Sun, J.M., Mullen, T.M., Davie, J.R., Rose, D.W., Glass, C.K.*, et al.* (1998). SAP30, a component of the mSin3 corepressor complex involved in N-CoR-mediated repression by specific transcription factors. Molecular cell *2*, 33-42.
Lai, E.C. (2003). microRNAs: runts of the genome assert themselves. Current biology : CB *13*, R925-936.

Lau, N.C., Lim, L.P., Weinstein, E.G., and Bartel, D.P. (2001). An abundant class of tiny RNAs with probable regulatory roles in Caenorhabditis elegans. Science *294*, 858-862.
Lau, N.C., Seto, A.G., Kim, J., Kuramochi-Miyagawa, S., Nakano, T., Bartel, D.P., and Kingston, R.E. (2006). Characterization of the piRNA complex from rat testes. Science *313*, 363-367.
Lau, P.W., Guiley, K.Z., De, N., Potter, C.S., Carragher, B., and MacRae, I.J. (2012). The molecular architecture of human Dicer. Nature structural & molecular biology *19*, 436-440.
Lau, P.W., Potter, C.S., Carragher, B., and MacRae, I.J. (2009). Structure of the human Dicer-TRBP complex by electron microscopy. Structure *17*, 1326-1332.
Le Thomas, A., Rogers, A.K., Webster, A., Marinov, G.K., Liao, S.E., Perkins, E.M., Hur, J.K., Aravin, A.A., and Toth, K.F. (2013). Piwi induces piRNA-guided transcriptional silencing and establishment of a repressive chromatin state. Genes & development *27*, 390-399.
Lee, C.G., and Hurwitz, J. (1992). A new RNA helicase isolated from HeLa cells that catalytically translocates in the 3' to 5' direction. The Journal of biological chemistry *267*, 4398-4407.
Lee, H.C., Gu, W.F., Shirayama, M., Youngman, E., Conte, D., and Mello, C.C. (2012). C. elegans piRNAs Mediate the Genome-wide Surveillance of Germline Transcripts. Cell *150*, 78-87.
Lee, H.Y., and Doudna, J.A. (2012). TRBP alters human precursor microRNA processing in vitro. Rna *18*, 2012-2019.
Lee, H.Y., Zhou, K., Smith, A.M., Noland, C.L., and Doudna, J.A. (2013). Differential roles of human Dicer-binding proteins TRBP and PACT in small RNA processing. Nucleic acids research *41*, 6568-6576.
Lee, R.C., and Ambros, V. (2001). An extensive class of small RNAs in Caenorhabditis elegans. Science *294*, 862-864.
Lee, R.C., Feinbaum, R.L., and Ambros, V. (1993). The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell *75*, 843-854.

Lee, R.C., Hammell, C.M., and Ambros, V. (2006). Interacting endogenous and exogenous RNAi pathways in Caenorhabditis elegans. Rna *12*, 589-597.
Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., Lee, J., Provost, P., Radmark, O., Kim, S.*, et al.* (2003). The nuclear RNase III Drosha initiates microRNA processing. Nature *425*, 415-419.
Lee, Y., Jeon, K., Lee, J.T., Kim, S., and Kim, V.N. (2002). MicroRNA maturation: stepwise processing and subcellular localization. The EMBO journal *21*, 4663-4670.
Lee, Y., Kim, M., Han, J., Yeom, K.H., Lee, S., Baek, S.H., and Kim, V.N. (2004a). MicroRNA genes are transcribed by RNA polymerase II. The EMBO journal *23*, 4051-4060.
Lee, Y.S., Nakahara, K., Pham, J.W., Kim, K., He, Z., Sontheimer, E.J., and Carthew, R.W. (2004b). Distinct roles for Drosophila Dicer-1 and Dicer-2 in the siRNA/miRNA silencing pathways. Cell *117*, 69-81.
Levin, M., Hashimshony, T., Wagner, F., and Yanai, I. (2012). Developmental milestones punctuate gene expression in the Caenorhabditis embryo. Developmental cell *22*, 1101-1108.
Lewis, B.P., Burge, C.B., and Bartel, D.P. (2005). Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell *120*, 15-20.
Lewis, B.P., Shih, I.H., Jones-Rhoades, M.W., Bartel, D.P., and Burge, C.B. (2003). Prediction of mammalian microRNA targets. Cell *115*, 787-798.
Lim, L.P., Glasner, M.E., Yekta, S., Burge, C.B., and Bartel, D.P. (2003a). Vertebrate microRNA genes. Science *299*, 1540.
Lim, L.P., Lau, N.C., Weinstein, E.G., Abdelhakim, A., Yekta, S., Rhoades, M.W., Burge, C.B., and Bartel, D.P. (2003b). The microRNAs of Caenorhabditis elegans. Genes & development *17*, 991-1008.
Lin, S.Y., Johnson, S.M., Abraham, M., Vella, M.C., Pasquinelli, A., Gamberi, C., Gottlieb, E., and Slack, F.J. (2003). The C elegans hunchback homolog, hbl-1, controls temporal patterning and is a probable microRNA target. Developmental cell *4*, 639-650.

Lingel, A., Simon, B., Izaurralde, E., and Sattler, M. (2003). Structure and nucleic-acid binding of the Drosophila Argonaute 2 PAZ domain. Nature *426*, 465-469.
Lingel, A., Simon, B., Izaurralde, E., and Sattler, M. (2004). Nucleic acid 3'-end recognition by the Argonaute2 PAZ domain. Nature structural & molecular biology *11*, 576-577.
Liu, J., Carmell, M.A., Rivas, F.V., Marsden, C.G., Thomson, J.M., Song, J.J., Hammond, S.M., Joshua-Tor, L., and Hannon, G.J. (2004). Argonaute2 is the catalytic engine of mammalian RNAi. Science *305*, 1437-1441.
Liu, Q., Rand, T.A., Kalidas, S., Du, F., Kim, H.E., Smith, D.P., and Wang, X. (2003). R2D2, a bridge between the initiation and effector steps of the Drosophila RNAi pathway. Science *301*, 1921-1925.
Llave, C., Kasschau, K.D., Rector, M.A., and Carrington, J.C. (2002a). Endogenous and silencing-associated small RNAs in plants. The Plant cell *14*, 1605-1619.
Llave, C., Xie, Z., Kasschau, K.D., and Carrington, J.C. (2002b). Cleavage of Scarecrow-like mRNA targets directed by a class of Arabidopsis miRNA. Science *297*, 2053-2056.
Lund, E., Sheets, M.D., Imboden, S.B., and Dahlberg, J.E. (2011). Limiting Ago protein restricts RNAi and microRNA biogenesis during early development in Xenopus laevis. Genes & development *25*, 1121-1131.
Luteijn, M.J., van Bergeijk, P., Kaaij, L.J., Almeida, M.V., Roovers, E.F., Berezikov, E., and Ketting, R.F. (2012). Extremely stable Piwi-induced gene silencing in Caenorhabditis elegans. The EMBO journal *31*, 3422-3430.
Ma, E., MacRae, I.J., Kirsch, J.F., and Doudna, J.A. (2008). Autoinhibition of human dicer by its internal helicase domain. Journal of molecular biology *380*, 237-243.
Ma, E., Zhou, K., Kidwell, M.A., and Doudna, J.A. (2012a). Coordinated Activities of Human Dicer Domains in Regulatory RNA Processing. Journal of molecular biology.
Ma, E., Zhou, K., Kidwell, M.A., and Doudna, J.A. (2012b). Coordinated activities of human dicer domains in regulatory RNA processing. Journal of molecular biology *422*, 466-476.

Ma, J.B., Ye, K., and Patel, D.J. (2004). Structural basis for overhang-specific small interfering RNA recognition by the PAZ domain. Nature *429*, 318-322.
MacRae, I.J., Zhou, K., and Doudna, J.A. (2007). Structural determinants of RNA recognition and cleavage by Dicer. Nature structural & molecular biology *14*, 934-940.
Macrae, I.J., Zhou, K., Li, F., Repic, A., Brooks, A.N., Cande, W.Z., Adams, P.D., and Doudna, J.A. (2006). Structural basis for double-stranded RNA processing by Dicer. Science *311*, 195-198.
Maller Schulman, B.R., Liang, X., Stahlhut, C., DelConte, C., Stefani, G., and Slack, F.J. (2008). The let-7 microRNA target gene, Mlin41/Trim71 is required for mouse embryonic survival and neural tube closure. Cell cycle *7*, 3935-3942.
Maniataki, E., and Mourelatos, Z. (2005). A human, ATP-independent, RISC assembly machine fueled by pre-miRNA. Genes & development *19*, 2979-2990.
Martinez, J., Patkaniowska, A., Urlaub, H., Luhrmann, R., and Tuschl, T. (2002). Single-stranded antisense siRNAs guide target RNA cleavage in RNAi. Cell *110*, 563-574.
Matranga, C., Tomari, Y., Shin, C., Bartel, D.P., and Zamore, P.D. (2005). Passenger-strand cleavage facilitates assembly of siRNA into Ago2-containing RNAi enzyme complexes. Cell *123*, 607-620.
McMahon, L., Muriel, J.M., Roberts, B., Quinn, M., and Johnstone, I.L. (2003). Two sets of interacting collagens form functionally distinct substructures within a Caenorhabditis elegans extracellular matrix. Molecular biology of the cell *14*, 1366-1378.
Meister, G., Landthaler, M., Patkaniowska, A., Dorsett, Y., Teng, G., and Tuschl, T. (2004). Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs. Molecular cell *15*, 185-197.
Mello, C.C., and Conte, D., Jr. (2004). Revealing the world of RNA interference. Nature *431*, 338-342.
Meneely, P.M., and Wood, W.B. (1984). An autosomal gene that affects X chromosome expression and sex determination in Caenorhabditis elegans. Genetics *106*, 29-44.
Misquitta, L., and Paterson, B.M. (1999). Targeted disruption of gene function in Drosophila by RNA interference (RNA-i): a role for nautilus in embryonic somatic

muscle formation. Proceedings of the National Academy of Sciences of the United States of America *96*, 1451-1456.
Moazed, D. (2009). Small RNAs in transcriptional gene silencing and genome defence. Nature *457*, 413-420.
Montgomery, M.K., Xu, S., and Fire, A. (1998). RNA as a target of double-stranded RNA-mediated genetic interference in Caenorhabditis elegans. Proceedings of the National Academy of Sciences of the United States of America *95*, 15502-15507.
Moss, E.G., Lee, R.C., and Ambros, V. (1997). The cold shock domain protein LIN-28 controls developmental timing in C. elegans and is regulated by the lin-4 RNA. Cell *88*, 637-646.
Motamedi, M.R., Verdel, A., Colmenares, S.U., Gerber, S.A., Gygi, S.P., and Moazed, D. (2004). Two RNAi complexes, RITS and RDRC, physically interact and localize to noncoding centromeric RNAs. Cell *119*, 789-802.
Mourelatos, Z. (2008). Small RNAs: The seeds of silence. Nature *455*, 44-45.
Mourelatos, Z., Dostie, J., Paushkin, S., Sharma, A., Charroux, B., Abel, L., Rappsilber, J., Mann, M., and Dreyfuss, G. (2002). miRNPs: a novel class of ribonucleoproteins containing numerous microRNAs. Genes & development *16*, 720-728.
Nagy, L., Kao, H.Y., Chakravarti, D., Lin, R.J., Hassig, C.A., Ayer, D.E., Schreiber, S.L., and Evans, R.M. (1997). Nuclear receptor repression mediated by a complex containing SMRT, mSin3A, and histone deacetylase. Cell *89*, 373-380.
Nakagawa, A., Shi, Y., Kage-Nakadai, E., Mitani, S., and Xue, D. (2010). Caspase-dependent conversion of Dicer ribonuclease into a death-promoting deoxyribonuclease. Science *328*, 327-334.
Nakanishi, K., Weinberg, D.E., Bartel, D.P., and Patel, D.J. (2012). Structure of yeast Argonaute with guide RNA. Nature *486*, 368-374.
Napoli, C., Lemieux, C., and Jorgensen, R. (1990). Introduction of a Chimeric Chalcone Synthase Gene into Petunia Results in Reversible Co-Suppression of Homologous Genes in trans. The Plant cell *2*, 279-289.

Ngo, H., Tschudi, C., Gull, K., and Ullu, E. (1998). Double-stranded RNA induces mRNA degradation in Trypanosoma brucei. Proceedings of the National Academy of Sciences of the United States of America *95*, 14687-14692.
Nishimasu, H., Ishizu, H., Saito, K., Fukuhara, S., Kamatani, M.K., Bonnefond, L., Matsumoto, N., Nishizawa, T., Nakanaga, K., Aoki, J.*, et al.* (2012). Structure and function of Zucchini endoribonuclease in piRNA biogenesis. Nature *491*, 284-287.
Nykanen, A., Haley, B., and Zamore, P.D. (2001). ATP requirements and small interfering RNA structure in the RNA interference pathway. Cell *107*, 309-321.
Obenauer, J.C., Cantley, L.C., and Yaffe, M.B. (2003). Scansite 2.0: Proteome-wide prediction of cell signaling interactions using short sequence motifs. Nucleic acids research *31*, 3635-3641.
Okamura, K., Balla, S., Martin, R., Liu, N., and Lai, E.C. (2008a). Two distinct mechanisms generate endogenous siRNAs from bidirectional transcription in Drosophila melanogaster. Nature structural & molecular biology *15*, 581-590.
Okamura, K., Chung, W.J., Ruby, J.G., Guo, H., Bartel, D.P., and Lai, E.C. (2008b). The Drosophila hairpin RNA pathway generates endogenous short interfering RNAs. Nature *453*, 803-806.
Olivieri, D., Sykora, M.M., Sachidanandam, R., Mechtler, K., and Brennecke, J. (2010). An in vivo RNAi assay identifies major genetic and cellular requirements for primary piRNA biogenesis in Drosophila. The EMBO journal *29*, 3301-3317.
Olsen, P.H., and Ambros, V. (1999). The lin-4 regulatory RNA controls developmental timing in Caenorhabditis elegans by blocking LIN-14 protein synthesis after the initiation of translation. Developmental biology *216*, 671-680.
Pak, J., and Fire, A. (2007). Distinct populations of primary and secondary effectors during RNAi in C. elegans. Science *315*, 241-244.
Pardo, P.S., Leung, J.K., Lucchesi, J.C., and Pereira-Smith, O.M. (2002). MRG15, a novel chromodomain protein, is present in two distinct multiprotein complexes involved in transcriptional activation. The Journal of biological chemistry *277*, 50860-50866.

Park, J.E., Heo, I., Tian, Y., Simanshu, D.K., Chang, H., Jee, D., Patel, D.J., and Kim, V.N. (2011). Dicer recognizes the 5' end of RNA for efficient and accurate processing. Nature *475*, 201-205.
Park, W., Li, J., Song, R., Messing, J., and Chen, X. (2002). CARPEL FACTORY, a Dicer homolog, and HEN1, a novel protein, act in microRNA metabolism in Arabidopsis thaliana. Current biology : CB *12*, 1484-1495.
Pasquinelli, A.E., Reinhart, B.J., Slack, F., Martindale, M.Q., Kuroda, M.I., Maller, B., Hayward, D.C., Ball, E.E., Degnan, B., Muller, P.*, et al.* (2000). Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. Nature *408*, 86-89.
Pavelec, D.M., Lachowiec, J., Duchaine, T.F., Smith, H.E., and Kennedy, S. (2009). Requirement for the ERI/DICER complex in endogenous RNA interference and sperm development in Caenorhabditis elegans. Genetics *183*, 1283-1295.
Pham, J.W., Pellino, J.L., Lee, Y.S., Carthew, R.W., and Sontheimer, E.J. (2004). A Dicer-2-dependent 80s complex cleaves targeted mRNAs during RNAi in Drosophila. Cell *117*, 83-94.
Phillips, C.M., Montgomery, T.A., Breen, P.C., and Ruvkun, G. (2012). MUT-16 promotes formation of perinuclear mutator foci required for RNA silencing in the C. elegans germline. Genes & development *26*, 1433-1444.
Pillai, R.S., Bhattacharyya, S.N., Artus, C.G., Zoller, T., Cougot, N., Basyuk, E., Bertrand, E., and Filipowicz, W. (2005). Inhibition of translational initiation by Let-7 MicroRNA in human cells. Science *309*, 1573-1576.
Piskounova, E., Polytarchou, C., Thornton, J.E., LaPierre, R.J., Pothoulakis, C., Hagan, J.P., Iliopoulos, D., and Gregory, R.I. (2011). Lin28A and Lin28B inhibit let-7 microRNA biogenesis by distinct mechanisms. Cell *147*, 1066-1079.
Plenefisch, J.D., Delong, L., and Meyer, B.J. (1989). Genes That Implement the Hermaphrodite Mode of Dosage Compensation in Caenorhabditis-Elegans. Genetics *121*, 57-76.

Powell, D.J., Friedman, J.M., Oulette, A.J., Krauter, K.S., and Darnell, J.E., Jr. (1984). Transcriptional and post-transcriptional control of specific messenger RNAs in adult and embryonic liver. Journal of molecular biology *179*, 21-35.
Praitis, V., Casey, E., Collar, D., and Austin, J. (2001). Creation of low-copy integrated transgenic lines in Caenorhabditis elegans. Genetics *157*, 1217-1226.
Provost, P., Dishart, D., Doucet, J., Frendewey, D., Samuelsson, B., and Radmark, O. (2002). Ribonuclease activity and RNA binding of recombinant human Dicer. The EMBO journal *21*, 5864-5874.
Qi, Y., Denli, A.M., and Hannon, G.J. (2005). Biochemical specialization within Arabidopsis RNA silencing pathways. Molecular cell *19*, 421-428.
Qin, H., Chen, F., Huan, X., Machida, S., Song, J., and Yuan, Y.A. (2010). Structure of the Arabidopsis thaliana DCL4 DUF283 domain reveals a noncanonical double-stranded RNA-binding fold for protein-protein interaction. Rna *16*, 474-481.
Rajewsky, N., and Socci, N.D. (2004). Computational identification of microRNA targets. Developmental biology *267*, 529-535.
Rand, T.A., Petersen, S., Du, F., and Wang, X. (2005). Argonaute2 cleaves the anti-guide strand of siRNA during RISC activation. Cell *123*, 621-629.
Ratcliff, F.G., MacFarlane, S.A., and Baulcombe, D.C. (1999). Gene silencing without DNA. rna-mediated cross-protection between viruses. The Plant cell *11*, 1207-1216.
Rechavi, O., Minevich, G., and Hobert, O. (2011). Transgenerational inheritance of an acquired small RNA-based antiviral response in C. elegans. Cell *147*, 1248-1256.
Reinhart, B.J., Slack, F.J., Basson, M., Pasquinelli, A.E., Bettinger, J.C., Rougvie, A.E., Horvitz, H.R., and Ruvkun, G. (2000). The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. Nature *403*, 901-906.
Reinhart, B.J., Weinstein, E.G., Rhoades, M.W., Bartel, B., and Bartel, D.P. (2002). MicroRNAs in plants. Genes & development *16*, 1616-1626.
Rhoades, M.W., Reinhart, B.J., Lim, L.P., Burge, C.B., Bartel, B., and Bartel, D.P. (2002). Prediction of plant microRNA targets. Cell *110*, 513-520.

Roberts, A., Trapnell, C., Donaghey, J., Rinn, J.L., and Pachter, L. (2011). Improving RNA-Seq expression estimates by correcting for fragment bias. Genome biology *12*, R22.

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

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

Rybak-Wolf, A., Jens, M., Murakawa, Y., Herzog, M., Landthaler, M., and Rajewsky, N. (2014). A variety of dicer substrates in human and C. elegans. Cell *159*, 1153-1167.

Saleh, M.C., Tassetto, M., van Rij, R.P., Goic, B., Gausson, V., Berry, B., Jacquier, C., Antoniewski, C., and Andino, R. (2009). Antiviral immunity in Drosophila requires systemic RNA interference spread. Nature *458*, 346-350.

Sampson, V.B., Rong, N.H., Han, J., Yang, Q., Aris, V., Soteropoulos, P., Petrelli, N.J., Dunn, S.P., and Krueger, L.J. (2007). MicroRNA let-7a down-regulates MYC and reverts MYC-induced growth in Burkitt lymphoma cells. Cancer research *67*, 9762-9770.

Sanchez Alvarado, A., and Newmark, P.A. (1999). Double-stranded RNA specifically disrupts gene expression during planarian regeneration. Proceedings of the National Academy of Sciences of the United States of America *96*, 5049-5054.

Sasaki, T., and Shimizu, N. (2007). Evolutionary conservation of a unique amino acid sequence in human DICER protein essential for binding to Argonaute family proteins. Gene *396*, 312-320.

Sawh, A.N., and Duchaine, T.F. (2012). Turning Dicer on its head. Nature structural & molecular biology *19*, 365-366.

Sawh, A.N., and Duchaine, T.F. (2013). A truncated form of dicer tilts the balance of RNA interference pathways. Cell reports *4*, 454-463.

Schauer, I.E., and Wood, W.B. (1990). Early C. elegans embryos are transcriptionally active. Development *110*, 1303-1317.

Schauer, S.E., Jacobsen, S.E., Meinke, D.W., and Ray, A. (2002). DICER-LIKE1: blind men and elephants in Arabidopsis development. Trends in plant science *7*, 487-491.
Seggerson, K., Tang, L., and Moss, E.G. (2002). Two genetic circuits repress the Caenorhabditis elegans heterochronic gene lin-28 after translation initiation. Developmental biology *243*, 215-225.
Selbach, M., Schwanhausser, B., Thierfelder, N., Fang, Z., Khanin, R., and Rajewsky, N. (2008). Widespread changes in protein synthesis induced by microRNAs. Nature *455*, 58-63.
Seth, M., Shirayama, M., Gu, W., Ishidate, T., Conte, D., Jr., and Mello, C.C. (2013). The C. elegans CSR-1 argonaute pathway counteracts epigenetic silencing to promote germline gene expression. Developmental cell *27*, 656-663.
Shirayama, M., Seth, M., Lee, H.C., Gu, W.F., Ishidate, T., Conte, D., and Mello, C.C. (2012). piRNAs Initiate an Epigenetic Memory of Nonself RNA in the C. elegans Germline. Cell *150*, 65-77.
Sienski, G., Donertas, D., and Brennecke, J. (2012). Transcriptional silencing of transposons by Piwi and maelstrom and its impact on chromatin state and gene expression. Cell *151*, 964-980.
Sijen, T., Fleenor, J., Simmer, F., Thijssen, K.L., Parrish, S., Timmons, L., Plasterk, R.H., and Fire, A. (2001). On the role of RNA amplification in dsRNA-triggered gene silencing. Cell *107*, 465-476.
Simmer, F., Tijsterman, M., Parrish, S., Koushika, S.P., Nonet, M.L., Fire, A., Ahringer, J., and Plasterk, R.H. (2002). Loss of the putative RNA-directed RNA polymerase RRF-3 makes C. elegans hypersensitive to RNAi. Current biology : CB *12*, 1317-1319.
Sinha, N.K., Trettin, K.D., Aruscavage, P.J., and Bass, B.L. (2015). Drosophila Dicer-2 Cleavage Is Mediated by Helicase- and dsRNA Termini-Dependent States that Are Modulated by Loquacious-PD. Molecular cell *58*, 406-417.
Slack, F.J., Basson, M., Liu, Z., Ambros, V., Horvitz, H.R., and Ruvkun, G. (2000). The lin-41 RBCC gene acts in the C. elegans heterochronic pathway between the let-7 regulatory RNA and the LIN-29 transcription factor. Molecular cell *5*, 659-669.

Smale, S.T. (2009). Nuclear run-on assay. Cold Spring Harbor protocols *2009*, pdb prot5329.
So, W.V., and Rosbash, M. (1997). Post-transcriptional regulation contributes to Drosophila clock gene mRNA cycling. The EMBO journal *16*, 7146-7155.
Song, J.J., Liu, J., Tolia, N.H., Schneiderman, J., Smith, S.K., Martienssen, R.A., Hannon, G.J., and Joshua-Tor, L. (2003). The crystal structure of the Argonaute2 PAZ domain reveals an RNA binding motif in RNAi effector complexes. Nature structural biology *10*, 1026-1032.
Song, J.J., Smith, S.K., Hannon, G.J., and Joshua-Tor, L. (2004). Crystal structure of Argonaute and its implications for RISC slicer activity. Science *305*, 1434-1437.
Stark, A., Brennecke, J., Russell, R.B., and Cohen, S.M. (2003). Identification of Drosophila MicroRNA targets. PLoS biology *1*, E60.
Steiner, F.A., Okihara, K.L., Hoogstrate, S.W., Sijen, T., and Ketting, R.F. (2009). RDE-1 slicer activity is required only for passenger-strand cleavage during RNAi in Caenorhabditis elegans. Nature structural & molecular biology *16*, 207-211.
Stoeckius, M., Maaskola, J., Colombo, T., Rahn, H.P., Friedlander, M.R., Li, N., Chen, W., Piano, F., and Rajewsky, N. (2009). Large-scale sorting of C. elegans embryos reveals the dynamics of small RNA expression. Nature methods *6*, 745-751.
Sunkar, R., and Zhu, J.K. (2004). Novel and stress-regulated microRNAs and other small RNAs from Arabidopsis. The Plant cell *16*, 2001-2019.
Tabara, H., Sarkissian, M., Kelly, W.G., Fleenor, J., Grishok, A., Timmons, L., Fire, A., and Mello, C.C. (1999). The rde-1 gene, RNA interference, and transposon silencing in C. elegans. Cell *99*, 123-132.
Tabara, H., Yigit, E., Siomi, H., and Mello, C.C. (2002). The dsRNA binding protein RDE-4 interacts with RDE-1, DCR-1, and a DExH-box helicase to direct RNAi in C. elegans. Cell *109*, 861-871.
Takasaki, T., Liu, Z., Habara, Y., Nishiwaki, K., Nakayama, J., Inoue, K., Sakamoto, H., and Strome, S. (2007). MRG-1, an autosome-associated protein, silences X-linked genes and protects germline immortality in Caenorhabditis elegans. Development *134*, 757-767.

Takeshita, D., Zenno, S., Lee, W.C., Nagata, K., Saigo, K., and Tanokura, M. (2007). Homodimeric structure and double-stranded RNA cleavage activity of the C-terminal RNase III domain of human dicer. Journal of molecular biology *374*, 106-120.
Tam, O.H., Aravin, A.A., Stein, P., Girard, A., Murchison, E.P., Cheloufi, S., Hodges, E., Anger, M., Sachidanandam, R., Schultz, R.M.*, et al.* (2008). Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes. Nature *453*, 534-538.
Thivierge, C., Makil, N., Flamand, M., Vasale, J.J., Mello, C.C., Wohlschlegel, J., Conte, D., Jr., and Duchaine, T.F. (2012). Tudor domain ERI-5 tethers an RNA-dependent RNA polymerase to DCR-1 to potentiate endo-RNAi. Nature structural & molecular biology *19*, 90-97.
Tian, Y., Simanshu, D.K., Ma, J.B., Park, J.E., Heo, I., Kim, V.N., and Patel, D.J. (2014). A phosphate-binding pocket within the platform-PAZ-connector helix cassette of human Dicer. Molecular cell *53*, 606-616.
Timmons, L., Court, D.L., and Fire, A. (2001). Ingestion of bacterially expressed dsRNAs can produce specific and potent genetic interference in Caenorhabditis elegans. Gene *263*, 103-112.
Tominaga, K., Kirtane, B., Jackson, J.G., Ikeno, Y., Ikeda, T., Hawks, C., Smith, J.R., Matzuk, M.M., and Pereira-Smith, O.M. (2005). MRG15 regulates embryonic development and cell proliferation. Molecular and cellular biology *25*, 2924-2937.
Towbin, B.D., Gonzalez-Aguilera, C., Sack, R., Gaidatzis, D., Kalck, V., Meister, P., Askjaer, P., and Gasser, S.M. (2012). Step-wise methylation of histone H3K9 positions heterochromatin at the nuclear periphery. Cell *150*, 934-947.
Tsai, H.Y., Chen, C.C., Conte, D., Jr., Moresco, J.J., Chaves, D.A., Mitani, S., Yates, J.R., 3rd, Tsai, M.D., and Mello, C.C. (2015). A ribonuclease coordinates siRNA amplification and mRNA cleavage during RNAi. Cell *160*, 407-419.
Tsutsumi, A., Kawamata, T., Izumi, N., Seitz, H., and Tomari, Y. (2011). Recognition of the pre-miRNA structure by Drosophila Dicer-1. Nature structural & molecular biology *18*, 1153-1158.

Ucker, D.S., and Yamamoto, K.R. (1984). Early events in the stimulation of mammary tumor virus RNA synthesis by glucocorticoids. Novel assays of transcription rates. The Journal of biological chemistry *259*, 7416-7420.
Valastyan, S., and Weinberg, R.A. (2010). Metastasis suppression: a role of the Dice(r). Genome biology *11*, 141.
van der Krol, A.R., Mur, L.A., Beld, M., Mol, J.N., and Stuitje, A.R. (1990). Flavonoid genes in petunia: addition of a limited number of gene copies may lead to a suppression of gene expression. The Plant cell *2*, 291-299.
van Rij, R.P., Saleh, M.C., Berry, B., Foo, C., Houk, A., Antoniewski, C., and Andino, R. (2006). The RNA silencing endonuclease Argonaute 2 mediates specific antiviral immunity in Drosophila melanogaster. Genes & development *20*, 2985-2995.
Vasale, J.J., Gu, W., Thivierge, C., Batista, P.J., Claycomb, J.M., Youngman, E.M., Duchaine, T.F., Mello, C.C., and Conte, D., Jr. (2010). Sequential rounds of RNA-dependent RNA transcription drive endogenous small-RNA biogenesis in the ERGO-1/Argonaute pathway. Proceedings of the National Academy of Sciences of the United States of America *107*, 3582-3587.
Vastenhouw, N.L., Brunschwig, K., Okihara, K.L., Muller, F., Tijsterman, M., and Plasterk, R.H. (2006). Gene expression: long-term gene silencing by RNAi. Nature *442*, 882.
Vastenhouw, N.L., Fischer, S.E., Robert, V.J., Thijssen, K.L., Fraser, A.G., Kamath, R.S., Ahringer, J., and Plasterk, R.H. (2003). A genome-wide screen identifies 27 genes involved in transposon silencing in C. elegans. Current biology : CB *13*, 1311-1316.
Vella, M.C., Choi, E.Y., Lin, S.Y., Reinert, K., and Slack, F.J. (2004). The C. elegans microRNA let-7 binds to imperfect let-7 complementary sites from the lin-41 3'UTR. Genes & development *18*, 132-137.
Verdel, A., Jia, S., Gerber, S., Sugiyama, T., Gygi, S., Grewal, S.I., and Moazed, D. (2004). RNAi-mediated targeting of heterochromatin by the RITS complex. Science *303*, 672-676.
Viswanathan, S.R., Daley, G.Q., and Gregory, R.I. (2008). Selective blockade of microRNA processing by Lin28. Science *320*, 97-100.

Volpe, T.A., Kidner, C., Hall, I.M., Teng, G., Grewal, S.I., and Martienssen, R.A. (2002). Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. Science *297*, 1833-1837.
Wang, G., and Reinke, V. (2008). A C. elegans Piwi, PRG-1, regulates 21U-RNAs during spermatogenesis. Current biology : CB *18*, 861-867.
Wang, H.W., Noland, C., Siridechadilok, B., Taylor, D.W., Ma, E., Felderer, K., Doudna, J.A., and Nogales, E. (2009). Structural insights into RNA processing by the human RISC-loading complex. Nature structural & molecular biology *16*, 1148-1153.
Wang, X.J., Gaasterland, T., and Chua, N.H. (2005). Genome-wide prediction and identification of cis-natural antisense transcripts in Arabidopsis thaliana. Genome biology *6*, R30.
Wang, Y., Sheng, G., Juranek, S., Tuschl, T., and Patel, D.J. (2008). Structure of the guide-strand-containing argonaute silencing complex. Nature *456*, 209-213.
Ward, J.D. (2015). Rapid and precise engineering of the Caenorhabditis elegans genome with lethal mutation co-conversion and inactivation of NHEJ repair. Genetics *199*, 363-377.
Wargelius, A., Ellingsen, S., and Fjose, A. (1999). Double-stranded RNA induces specific developmental defects in zebrafish embryos. Biochemical and biophysical research communications *263*, 156-161.
Washburn, M.P., Wolters, D., and Yates, J.R., 3rd (2001). Large-scale analysis of the yeast proteome by multidimensional protein identification technology. Nature biotechnology *19*, 242-247.
Watanabe, T., Totoki, Y., Toyoda, A., Kaneda, M., Kuramochi-Miyagawa, S., Obata, Y., Chiba, H., Kohara, Y., Kono, T., Nakano, T.*, et al.* (2008). Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. Nature *453*, 539-543.
Welker, N.C., Maity, T.S., Ye, X., Aruscavage, P.J., Krauchuk, A.A., Liu, Q., and Bass, B.L. (2011). Dicer's helicase domain discriminates dsRNA termini to promote an altered reaction mode. Molecular cell *41*, 589-599.

Welker, N.C., Pavelec, D.M., Nix, D.A., Duchaine, T.F., Kennedy, S., and Bass, B.L. (2010). Dicer's helicase domain is required for accumulation of some, but not all, C. elegans endogenous siRNAs. Rna *16*, 893-903.
White, E., Schlackow, M., Kamieniarz-Gdula, K., Proudfoot, N.J., and Gullerova, M. (2014). Human nuclear Dicer restricts the deleterious accumulation of endogenous double-stranded RNA. Nature structural & molecular biology *21*, 552-559.
Wianny, F., and Zernicka-Goetz, M. (2000). Specific interference with gene function by double-stranded RNA in early mouse development. Nature cell biology *2*, 70-75.
Wienholds, E., Koudijs, M.J., van Eeden, F.J., Cuppen, E., and Plasterk, R.H. (2003). The microRNA-producing enzyme Dicer1 is essential for zebrafish development. Nature genetics *35*, 217-218.
Wightman, B., Ha, I., and Ruvkun, G. (1993). Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans. Cell *75*, 855-862.
Wilkins, C., Dishongh, R., Moore, S.C., Whitt, M.A., Chow, M., and Machaca, K. (2005). RNA interference is an antiviral defence mechanism in Caenorhabditis elegans. Nature *436*, 1044-1047.
Wu, E., Thivierge, C., Flamand, M., Mathonnet, G., Vashisht, A.A., Wohlschlegel, J., Fabian, M.R., Sonenberg, N., and Duchaine, T.F. (2010). Pervasive and cooperative deadenylation of 3'UTRs by embryonic microRNA families. Molecular cell *40*, 558-570.
Wu, L., Fan, J., and Belasco, J.G. (2006). MicroRNAs direct rapid deadenylation of mRNA. Proceedings of the National Academy of Sciences of the United States of America *103*, 4034-4039.
Xu, P., Vernooy, S.Y., Guo, M., and Hay, B.A. (2003). The Drosophila microRNA Mir-14 suppresses cell death and is required for normal fat metabolism. Current biology : CB *13*, 790-795.
Yan, K.S., Yan, S., Farooq, A., Han, A., Zeng, L., and Zhou, M.M. (2003). Structure and conserved RNA binding of the PAZ domain. Nature *426*, 468-474.

Yang, H., Vallandingham, J., Shiu, P., Li, H., Hunter, C.P., and Mak, H.Y. (2014). The DEAD box helicase RDE-12 promotes amplification of RNAi in cytoplasmic foci in C. elegans. Current biology : CB *24*, 832-838.
Yang, H., Zhang, Y., Vallandingham, J., Li, H., Florens, L., and Mak, H.Y. (2012). The RDE-10/RDE-11 complex triggers RNAi-induced mRNA degradation by association with target mRNA in C. elegans. Genes & development *26*, 846-856.
Ye, X., Paroo, Z., and Liu, Q. (2007). Functional anatomy of the Drosophila microRNA-generating enzyme. The Journal of biological chemistry *282*, 28373-28378.
Yigit, E., Batista, P.J., Bei, Y., Pang, K.M., Chen, C.C., Tolia, N.H., Joshua-Tor, L., Mitani, S., Simard, M.J., and Mello, C.C. (2006). Analysis of the C. elegans Argonaute family reveals that distinct Argonautes act sequentially during RNAi. Cell *127*, 747-757.
Zamore, P.D., Tuschl, T., Sharp, P.A., and Bartel, D.P. (2000). RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. Cell *101*, 25-33.
Zhang, C., Montgomery, T.A., Gabel, H.W., Fischer, S.E., Phillips, C.M., Fahlgren, N., Sullivan, C.M., Carrington, J.C., and Ruvkun, G. (2011). mut-16 and other mutator class genes modulate 22G and 26G siRNA pathways in Caenorhabditis elegans. Proceedings of the National Academy of Sciences of the United States of America *108*, 1201-1208.
Zhang, H., Kolb, F.A., Brondani, V., Billy, E., and Filipowicz, W. (2002). Human Dicer preferentially cleaves dsRNAs at their termini without a requirement for ATP. The EMBO journal *21*, 5875-5885.
Zhang, H., Kolb, F.A., Jaskiewicz, L., Westhof, E., and Filipowicz, W. (2004). Single processing center models for human Dicer and bacterial RNase III. Cell *118*, 57-68.
Zhang, K., Mosch, K., Fischle, W., and Grewal, S.I. (2008). Roles of the Clr4 methyltransferase complex in nucleation, spreading and maintenance of heterochromatin. Nature structural & molecular biology *15*, 381-388.
Zhang, Y., Sun, Z.W., Iratni, R., Erdjument-Bromage, H., Tempst, P., Hampsey, M., and Reinberg, D. (1998). SAP30, a novel protein conserved between human and yeast, is a component of a histone deacetylase complex. Molecular cell *1*, 1021-1031.

Zhuang, J.J., Banse, S.A., and Hunter, C.P. (2013). The nuclear argonaute NRDE-3 contributes to transitive RNAi in Caenorhabditis elegans. Genetics *194*, 117-131.
Zimmermann, J., Latta, L., Beck, A., Leidinger, P., Fecher-Trost, C., Schlenstedt, G., Meese, E., Wissenbach, U., and Flockerzi, V. (2014). Trans-activation response (TAR) RNA-binding protein 2 is a novel modulator of transient receptor potential canonical 4 (TRPC4) protein. The Journal of biological chemistry *289*, 9766-9780.