UC San Diego

UC San Diego Electronic Theses and Dissertations

Title

MicroRNAs in the Making: Post-transcriptional Regulation of MicroRNA Biogenesis in Caenorhabditis elegans

Permalink

https://escholarship.org/uc/item/3f93318b

Author

Mondol, Vanessa

Publication Date

2015

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA, SAN DIEGO

MicroRNAs in the Making: Post-transcriptional Regulation of MicroRNA Biogenesis in *Caenorhabditis elegans*

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy

in

Biology

by

Vanessa Mondol

Committee in charge:

Professor Amy E. Pasquinelli, Chair
Professor Tracy Johnson
Professor Jens Lykke-Andersen
Professor Emily Troemel
Professor Miles Wilkinson

2015

Copyright

Vanessa Mondol, 2015

All rights reserved

The Dissertation of Vanessa Mondol is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____

_____

Chair

University of California, San Diego

2015

DEDICATION

I dedicate this dissertation to my mother, Lorsy Maria Mondol, and grandmother, Ruth Fernandez, who are strong women and excellent role models. They work hard for their goals, always speak their minds, and do not allow anyone to disrespect them. To my dad, Amaury Mondol, who, for as long as I can remember, was always taking some sort of class. Besides a great respect for education, he taught me to read everything and question everything I read. Finally, to my late cousin Mirta Isabel Viaña Pantoja (1980 – 2013); your warm welcoming smile could light up the whole room; you are truly missed.

EPIGRAPH

Nothing in life is to be feared, it is only to be understood.

Marie Curie

TABLE OF CONTENTS

Signature Page..... iii

Dedication..... iv

Epigraph..... v

Table of Contents..... vi

List of Figures..... ix

List of Tables..... xi

Acknowledgments..... xii

Vita..... xiv

Abstract of the Dissertation..... xvi

Chapter I Introduction..... 1

Chapter II Let's Make It Happen: The Role of let-7 MicroRNA in Development..... 11

A. Abstract..... 11
B. Introduction..... 11
C. The discovery of miRNAs..... 17
C.1. lin-4 miRNA..... 17
C.2. let-7 miRNA..... 19
C.3. let-7 miRNA in other organisms..... 20
D. The role of let-7 in *C. elegans* development..... 22
D.1. Biogenesis of let-7..... 22
D.2. Identification of let-7 targets..... 28
D.3. Developmental role of let-7 and its targets..... 32
E. The role of let-7 in development and disease across species..... 33
E.1. Stem cells and differentiation..... 33
E.2. Cancer..... 35
E.3. Neuronal functions..... 36
F. Conclusions..... 37
G. Acknowledgments..... 38

Chapter III Regulation of lin-4 microRNA expression by a conserved RNA binding protein in *Caenorhabditis elegans* 46
A. Abstract 46
B. Introduction 47
C. Results and Discussion 49
C.1. Two independently synthesized pri-lin-4 RNA transcripts solely contribute to lin-4 biogenesis 49
C.2. RBM-28 is an essential conserved RNA recognition motif protein ubiquitously expressed in *C. elegans* 54
C.3. RBM-28 specifically stabilizes mature lin-4 expression 55
D. Materials and Methods 57
D.1. Worm strains and culture conditions 57
D.2. Expression constructs and transgenic strains 58
E. Acknowledgments 58

Chapter IV Splicing remodels the let-7 primary microRNA to facilitate Drosha processing in *Caenorhabditis elegans* 65
A. Abstract 65
B. Introduction 66
C. Results and Discussion 69
C.1. Splicing remodels the pri-let-7 secondary structure 69
C.2. The splicing-induced structural changes facilitate pri-let-7 processing 70
C.3. The Microprocessor prefers spliced pri-let-7 *in vivo* 73
D. Materials and Methods 76
D.1. *In vitro* transcription of RNA 76
D.2. RNase secondary structure analysis 77
D.3. RNA immunoprecipitations 77
D.4. Sequences and Primers used in this study 78
E. Acknowledgments 80
F. Addendum 80

Chapter V Characterizing clustered microRNAs in *Caenorhabditis elegans* 91
A. Abstract 91
B. Introduction 91
C. Results 94
C.1. Identification of clustered pri-miRNAs in *C. elegans* 94
C.2. Confirming differential expression of miRNAs in clusters 98

C.2.a. Optimizing Firefly BioWorks multiplex cellular assay for miRNA detection..... 98
C.2.b. Mature miRNA expression over development measured by Northern Blot analysis..... 101
C.3. Identifying factors that affect regulation..... 102
D. Experimental Procedures..... 106
D.1. *In vitro* transcription of radiolabeled pri-miRNAs..... 107
D.2. *In vitro* Drosha processing..... 107
E. Discussion..... 109
F. Acknowledgments..... 110

Chapter VI Discussion..... 120
A. Insights on lin-4 and RBM-28 regulation..... 120
B. Implications of splicing and secondary structure on miRNA regulation..... 122
C. Perspectives on regulation of clustered miRNA biogenesis..... 127
D. Concluding remarks..... 130

References..... 133

LIST OF FIGURES

FIGURE 2.1 General miRNA Biogenesis Pathway..... 39

FIGURE 2.2 Mechanisms of miRNA target regulation..... 40

FIGURE 2.3 *C. elegans let-7* mutant phenotypes..... 41

FIGURE 2.4 Examples of let-7 target site interactions..... 42

FIGURE 2.5 Homologs of let-7 in worms and humans..... 42

FIGURE 2.6 Expression of let-7 throughout *C. elegans* development..... 43

FIGURE 2.7 Regulation of let-7 biogenesis and function in *C. elegans*..... 44

FIGURE 2.8 Regulation of cellular differentiation by let-7..... 45

FIGURE 3.1 Expression of lin-4 throughout early development..... 61

FIGURE 3.2 Validating pri-lin-4 transcripts..... 62

FIGURE 3.3 Characteriznig RBM-28 in *C. elegans*..... 63

FIGURE 3.4 RBM-28 regulates mature lin-4..... 64

FIGURE 4.1 Secondary structures of unspliced and spliced pri-let-7 RNAs.. 84

FIGURE 4.2 Splicing is important for let-7 biogenesis *in vivo*..... 86

FIGURE 4.3 The Microprocessor prefers SL1-spliced pri-let-7 for processing *in vivo*..... 88

SUPPLEMENTARY FIGURE 4.S1 RNase secondary structure assay of pri-let-7 isoforms..... 89

FIGURE 5.1 Expression of primary clustered microRNAs over development 115

FIGURE 5.2 Pri-miR-241-48 cluster..... 115

FIGURE 5.3 Differential miRNA expression measured by multiplex assay.... 116

FIGURE 5.4 Clustered miRNA expression in wild-type worms..... 117

FIGURE 5.5 Clustered miRNA expression over development..... 118

FIGURE 5.6 Potential models of clustered miRNA regulation..... 119

LIST OF TABLES

TABLE 3.1 Oligos used for this study..... 60

TABLE 4.1 RNA immunoprecipitation troubleshooting..... 82

TABLE 5.1 Predicted clustered pri-miRNAs..... 112

TABLE 5.2 Temporal expression of validated clustered pri-miRNAs..... 112

TABLE 5.3 Multiplex miRNA assay probe sequences..... 113

TABLE 5.4 Oligos used for this study..... 114

ACKNOWLEDGMENTS

Chapter II, in full, is a reprint of the material as it appears in Current Topics in Developmental Biology, Mondol, Vanessa & Pasquinelli, Amy E., Academic Press, 2012. I was the primary author. Chapter III, in part, contains material from “Regulation of lin-4 miRNA expression, organismal growth, and development by a conserved RNA binding protein in *C. elegans*”. Bracht, John R., Van Wynsberghe, Priscilla M., Mondol, Vanessa & Pasquinelli, Amy E., Developmental Biology, 2010. John and Priscilla were the primary authors; I was the second author. Chapter IV, in full, is a reprint of “Splicing remodels the let-7 primary microRNA to facilitate Drosha processing in *Caenorhabditis elegans*.” Mondol, Vanessa, Ahn, Byoung Chan. & Pasquinelli, Amy E., RNA, 2015. I was the primary author.

I would like to thank Professor Amy E. Pasquinelli for her support and guidance throughout my graduate school career. Amy has been a better mentor than I could have ever asked for. Her patience, brilliance, and passion for research and teaching are something I aspire to. I would also like to acknowledge my committee members, Professor Tracy Johnson, Professor Jens Lykke-Andersen, Professor Emily Troemel, and Professor Miles Wilkinson, for thoughtfully shaping my scientific thinking and challenging me towards excellence. In addition, thanks to my colleagues in the Pasquinelli lab, who made every day enjoyable, including Roger, James, Sarah, Antti and Ian. It is rare to see labs hang out and leave the bench to have lunch together, but we did it every day. In particular, I want to thank Dr. Priscilla Van-Wynsberghe for training me

when I came in as a rotation student and Emily Finnegan for always going to the bar with me after another failed experiment.

I have the upmost gratitude for my former PI, Professor Virginia W. Cornish. Without her, I would not be here; she encouraged me to pursue my graduate degree and gave me the confidence to try. She taught me so much more than just laboratory work. Virginia is outspoken, thoughtful, generous, and a fierce competitor; she is a great role model for any female scientist. I especially have to thank all the former Cornish lab members who I still call my very close friends: Phil, Laura, Monnica, Heather, and Nili. I enjoy catching up on our "journal club". Dr. Pamela Peralta-Yahya was a second year graduate student when she began training me in the Cornish lab. I was intimidated when I joined the lab, but Pamela taught me not to doubt myself and reminded me that I knew more about my research than anyone else. I love that I can still count on Pamela to help me with scientific advice and career decisions.

Outside of science, I thank my strong network of amazingly supportive and non-judgmental friends that allow me to be myself and remind me that experiments do not define me, especially Dillon, Amy, Dana, Maya, Rachel, Endi, Kimmy, Fuzz, and Maribell. I also want to thank the Ladies who Lunch and the 94th aerosquad brunch crew.

Last, but not least, I would like to thank my family, including my two Chihuahuas for their never-ending love and support. In particular, I would like to thank my husband Peder Pace, for putting up with me the last 16 years, even when I said my worms have more priority.

VITA

2005	Bachelor of Science in Biology, Stony Brook University
2009	Post-baccalaureate Bridge to Ph.D. Program in the Natural Sciences, Columbia University
2015	Doctor of Philosophy in Biology, University of California, San Diego

HONORS AND AWARDS

2010	Keystone Symposia Travel Scholarship
2012	RNA Society Travel Fellowship
2012 – 2014	National Institutes of Health Cell, Molecular and Genetics (CMG) Training Grant
2014	Best Graduate Student Talk, UCSD/Salk Division of Biological Sciences Retreat
2014 – 2015	National Institute of Health Ruth L. Kirschstein National Research Service Award for Individual Predoctoral Fellowship to Promote Diversity in Health-Related Research (F31)
2015	Elkes Foundation Keystone Symposia Travel Scholarship

PUBLICATIONS

Mondol V, Ahn BC, Pasquinelli AE. “Splicing remodels the let-7 primary miRNA to facilitate Drosha processing in *Caenorhabditis elegans*” RNA, 2015; 21(8)

Romanini DW, Peralta-Yahya P, **Mondol V**, Cornish VW. “A Heritable Recombination System for Synthetic Darwinian Evolution in Yeast”. ACS Synthetic Biology, 2012; 1(12):602-609

Massirer KB, Perez SG., **Mondol V**, Pasquinelli AE. “The miR-35-41 family of microRNAs regulates RNAi sensitivity in *Caenorhabditis elegans*”. PLOS Genetics, 2012; 8(3):e1002536

Mondol V, Pasquinelli AE. 2012. “Let’s make it happen: The role of the let-7 microRNA in development.” In *Current topics in developmental biology: microRNAs in development* (ed. Hornstein E.), pp. 1-30. Academic Press

Bracht JR., Van Wynsberghe PM., **Mondol V**, Pasquinelli AE. “Regulation of lin-4 miRNA expression, organismal growth and development by a conserved RNA binding protein in *C. elegans*”. Developmental Biology, 2010; 348(2):210-21

ABSTRACT OF THE DISSERTATION

MicroRNAs in the Making: Post-transcriptional Regulation of MicroRNA Biogenesis in *Caenorhabditis elegans*

by

Vanessa Mondol

Doctor of Philosophy in Biology

University of California, San Diego, 2015

Professor Amy E. Pasquinelli, Chair

MicroRNAs (miRNAs) are small non-coding RNAs, ~22 nucleotides (nt) long, with major roles in gene regulation. MiRNAs bind imperfectly to complementary sequences in the 3' untranslated region of target messenger RNAs (mRNAs) causing translational repression and destabilization. A single miRNA has the potential to regulate hundreds of different mRNA targets, highlighting the importance of miRNAs in almost all cellular pathways. Originally

discovered as part of the *Caenorhabditis elegans* (*C. elegans*) developmental timing pathway, miRNAs were soon found in a multitude of other organisms, including humans. MiRBase, an online database for miRNAs, now lists >35,000 miRNAs, in >200 different species, including viruses, though many of their roles remain to be characterized. Because misregulation is often associated with disease, especially cancer, exploring miRNA biogenesis is critical for understanding the intricacies of disease development. Furthermore, the conserved temporal expression of various miRNAs highlights the importance of these regulators in pluripotency and development. Understanding how miRNAs are produced and regulated has been a major topic of study over the past 15 years. While the basic mechanisms of miRNA biogenesis and function have been uncovered, how these processes are regulated remains an outstanding problem.

There are multiple instances of transcriptional and post-transcriptional regulation during miRNA biogenesis. In **Chapter I**, I introduce much of the latest understanding about the mechanisms of miRNA biogenesis and regulation. Details about the discovery of miRNAs, *C. elegans* as a model, as well as general information on biogenesis and targeting, can be found in **Chapter II**. I worked on several projects investigating post-transcriptional regulation of miRNA biogenesis in *C. elegans*. In **Chapter III**, I identify and characterize the primary lin-4 transcripts, and demonstrate how a conserved RNA binding protein, RBM-28, regulates the expression of mature lin-4, but not the primary or precursor forms. My investigation in **Chapter IV** led to notable insights on how splicing pri-let-7 leads to a secondary structure rearrangement that facilities recognition by

the Microprocessor. My survey of polycistronic worm miRNAs discussed in **Chapter V** indicates that there are many more examples awaiting further study. Overall, this research describes novel examples of post-transcriptional regulation of miRNAs.

Chapter I

Introduction

The central dogma of molecular biology describes how DNA genes are transcribed to RNA and subsequently translated to proteins (Crick, 1958, 1970). At the dawn of the millennium, several technological advances, including the full genome sequencing of *Caenorhabditis elegans*, *Drosophila melanogaster*, and *Homo sapiens*, eased the hunt for conserved essential genes and brought about the hope that a full human sequence would hold the key to understanding various genetic disorders and mechanisms of development (Consortium, 1998; Adams et al., 2000; Consortium, 2001, 2004). Unexpectedly, the number of protein-coding genes was much smaller than previously estimated, accounting for only about 1.5% of the DNA genome. This, along with recently discovered mechanisms of gene silencing by small non-coding RNAs (ncRNAs), RNA interference (RNAi) and microRNAs (miRNAs), meant the human genome was much more complex than previously imagined (Fire et al., 1998; Hammond et al., 2000; Pasquinelli et al., 2000; Zamore et al., 2000; Lagos-Quintana et al., 2001; Lau et al., 2001; Lee & Ambros, 2001; Lagos-Quintana et al., 2002; Reinhart et al., 2002). Many questions were raised about how ncRNAs are encoded, how to identify them, and what their many possible functions could be (Storz, 2002). In this introductory Chapter, I discuss the latest information on mechanisms regulating miRNA expression, including other ncRNAs. Further details regarding the history of miRNAs, the use of *Caenorhabditis elegans* as a model, and the role of

miRNAs in disease and development are discussed in Chapter II.

MiRNAs are ~22 nt ncRNAs capable of translational repression and/or deadenylation of messenger RNAs (mRNAs), by imperfectly binding to complementary sequences in target 3' untranslated regions (UTR) (Mondol & Pasquinelli, 2012). Thousands of miRNAs have been annotated, each with the potential to bind up to hundreds of target mRNAs (Griffiths-Jones et al., 2006; Kozomara & Griffiths-Jones, 2014; Li et al., 2014; Agarwal et al., 2015; Bastian et al., 2015; Chiu et al., 2015; Liu et al., 2015; Vlachos et al., 2015). Many miRNAs are highly conserved, in not only sequence, but also developmental expression. In addition, several components of the biogenesis pathway are also highly conserved. While I go into greater detail in Chapter II about the discovery and biogenesis of miRNAs, much has been learned since that work was published. Many more miRNA regulators have been studied, and several important complexes have been structurally and enzymatically characterized (Macias et al., 2012; Schirle & MacRae, 2012; Finnegan & Pasquinelli, 2013; Ha & Kim, 2014; Quick-Cleveland et al., 2014; Schirle et al., 2014; Chang et al., 2015; Lin & Gregory, 2015; Nguyen et al., 2015; Shen & Hung, 2015).

Canonical miRNA biogenesis begins with transcription of capped and polyadenylated primary miRNAs (pri-miRNAs) by RNA polymerase II (Pol II), although there are rare incidences of RNA Pol III transcribed miRNAs, typically after Alu repeats (Finnegan & Pasquinelli, 2013; Ha & Kim, 2014). Pri-miRNAs were originally thought to be up to several kilobases (kb) long, however, a recent genome-wide analysis of pri-miRNAs in humans and mice revealed that some

pri-miRNAs can be as large as hundreds of kilobases (Chang et al., 2015). A transcript of that size can potentially contain many regions of conservation and regulation; easily overlooked because regions so far away from the mature ~22 nucleotide (nt) miRNA sequence were not examined. Most research regarding pri-miRNA processing has focused on the ~100-200 nt flanking the precursor miRNA (pre-miRNA) hairpin (Denli et al., 2004; Han et al., 2004; Zeng & Cullen, 2005; Han et al., 2006; Lee et al., 2007; Morlando et al., 2008; Ballarino et al., 2009). Most pri-miRNAs are found as intergenically transcribed independent units. However, several pri-miRNAs are contiguous with downstream protein coding genes, suggesting possible coordinated expression or some other Drosha-mediated regulation of the downstream protein (Ballarino et al., 2009; Burger & Gullerova, 2015; Chang et al., 2015). In addition, some pri-miRNAs are found contiguous to previously annotated ncRNAs, suggesting that those ncRNAs may actually be degradation products of processing. Of the pri-miRNAs found in protein-coding genes, ~75-80% are intronic, and of these, about a third has been predicted to be independently transcribed (Ozsolak et al., 2008; Corcoran et al., 2009; Chang et al., 2015). Finally, while instances of polycistronically transcribed “clustered” miRNAs have been annotated, a recent study finds some are instead transcribed independently or as small subsets of the multi-miRNA cluster (Chang et al., 2015).

Regarding transcriptional regulation, alternative promoters are used frequently and they can control temporal and tissue-specific expression (Kai et al., 2012; Chang et al., 2015). Recent studies have found that many transcribed

genes, including pri-miRNAs, undergo oscillations of expression during development again suggesting regulation (William et al., 2007; Kim et al., 2013; Hendriks et al., 2014). Several well-studied transcription factors like Myc, Nanog, and Oct4, have been associated with miRNA accumulation or diminished transcription (Finnegan & Pasquinelli, 2013; Ha & Kim, 2014). Interestingly, many transcription factors are themselves targets of miRNA regulation, highlighting how abruptly misexpression of a miRNA can lead to a cascade of signaling problems throughout a cell (Lin & Gregory, 2015). Alternative splicing is also a common way to regulate addition of an alternative exon or processing of an overlapping pri-miRNA (Mattioli et al., 2013; Melamed et al., 2013; Chang et al., 2015).

Pri-miRNAs are co-transcriptionally processed to ~65 nt pre-miRNA hairpins by the Microprocessor complex, made up minimally of one RNase III enzyme, Drosha, and two copies of the RNA-binding partner Pasha (DGCR8 in vertebrates)(Finnegan & Pasquinelli, 2013; Ha & Kim, 2014; Nguyen et al., 2015). This processing relies mainly on the RNA secondary structure of the pri-miRNA, however, an enriched binding motif (CNNC) was found downstream of most pri-miRNA hairpins in Bilaterian animals, but not nematodes (Auyeung et al., 2013). Regarding secondary structure, unpaired basal segments at the end of the pri-miRNA seem to be the most important for processing, with assistance by internal bulges around the Drosha and Dicer cleavage sites (Finnegan & Pasquinelli, 2013; Ha & Kim, 2014). Many positive and negative regulators of pri-miRNA processing by the Microprocessor have been described, including KSRP,

TDP43, LIN-28, and hnRNPA1. In addition, Adenosine deaminases acting on RNA (ADARs) are involved in editing of adenosine residues to inosine in double-stranded RNA (dsRNA), disrupting secondary structure of pri- and pre-miRNAs. Interestingly, a high-throughput sequencing and cross-linking immunoprecipitation (HITS-CLIP) analysis of DGCR8 identified bound snoRNAs and rRNAs, suggesting other roles for this RNA binding protein (Macias et al., 2012). Not all pri-miRNAs are subject to Drosha cleavage. A number of ~65 nt miRNA-containing introns, known as mirtrons, bypass the Microprocessor and instead are processed by the spliceosome as part of canonical splicing that occurs during mRNA processing (Berezikov et al., 2007; Chan & Slack, 2007; Okamura et al., 2007; Chung et al., 2011; Westholm & Lai, 2011).

After processing, pre-miRNAs are exported to the cytoplasm through Exportin-5, however a similar exporting protein has not been determined in *C. elegans* (Finnegan & Pasquinelli, 2013; Ha & Kim, 2014). Once in the cytoplasm, RNase III enzyme Dicer further cleaves the pre-miRNA, to form the ~22 nt miRNA and the miRNA star (*) strand. As with the Microprocessor, several positive and negative regulators of Dicer processing have been described. For example, in *Drosophila,* the Loquacious RNA-binding proteins are essential for Dicer processing. In addition editing of the pre-miRNA, either by ADARs or uridylases like TUT4, has been shown to block Dicer processing. In mammals, pre-miRNA processing is enhanced by TRBP and PACT in mammals. A recent study finds Dicer that cannot bind TRBP or PACT has defects in guide strand selection (Wilson et al., 2015).

After Dicer cleavage, the remaining RNA duplex is unwound and one strand is preferentially loaded onto particular Argonaute proteins (ALG-1-2 in *C. elegans*, AGO1-4 in humans, AGO1 in flies), while the opposing passenger strand is ejected and degraded (Finnegan & Pasquinelli, 2013; Ha & Kim, 2014). Argonaute is the primary component of the miRNA induced silencing complex (miRISC), which seeks to find complementary sequences in target mRNAs to cause translational repression and deadenylation. Elucidation of the human Ago2 crystal structure, as well as single molecule fluorescence resonance energy transfer (FRET) experiments reveal how Ago2 rapidly scans mRNA targets with nucleotides 2-4, slowing down for perfect matches to the miRNA “seed”, nucleotides 2-8 (Chandradoss et al., 2015; Salomon et al., 2015). Once paired, AGO2 rearranges to allow base pairing by downstream nucleotides.

MicroRNA turnover is less understood. While its accepted that miRNA* strands rapidly degrade, there are cases where they are loaded onto Argonaute and direct mRNA target repression (Okamura et al., 2008; Guo & Lu, 2010). Mature miRNAs can also be degraded in the absence of either Argonaute or their targets. In *C. elegans*, the exonucleases XRN-1 and XRN-2 have been implicated in this form of miRNA decay (Grosshans & Chatterjee, 2010; Chatterjee et al., 2011). In *Drosophila*, the opposite seems true. The degree of target pairing increases miRNA decay rates and mature miRNAs can be tailed and trimmed (Finnegan & Pasquinelli, 2013; Ha & Kim, 2014).

As a part of RISC, miRNAs function to destabilize mRNA targets, through translational repression and/or deadenylation, although whether both events

occur, the order in which they operate, or whether they are mutually exclusive, remains up for debate as there is evidence for each scenario (Huntzinger & Izaurralde, 2011; Fabian & Sonenberg, 2012). The strength of complementarity between guide and target determines RISC dissociation from the target; whether RISC simply binds or binds and cleaves the target (Salomon et al., 2015). RISC can also be loaded with endogenous small interfering RNAs (siRNAs), which arise from double-stranded RNAs, and work to silence mRNA targets through perfect base pairing. In humans, when RISC binds to target sites with perfect complementarity, it triggers endonucleolytic cleavage of the mRNA by Ago2 (Liu et al., 2004). MiRNA-mediated RISC silencing occurs using imperfect complementary binding of miRNA and target (Tritschler et al., 2010; Huntzinger & Izaurralde, 2011; Fabian & Sonenberg, 2012). Ago and GW182 work as part of the RISC complex to inhibit translation initiation and to block protein accumulation post-initiation. RISC can inhibit ribosome scanning by recruiting eIF4AII through an interaction with the NOT1 subunit of the CCR4-NOT deadenylase complex, blocking initiation (Mathonnet et al., 2007; Meijer et al., 2013). Deadenylation of mRNA targets occurs when GW182 proteins recruit deadenylase complexes, namely PAN2-PAN3 and CCR4-NOT (Wu et al., 2010; Huntzinger & Izaurralde, 2011; Fabian & Sonenberg, 2012). Deadenylated transcripts are decapped by decapping complexes and degraded by exonucleases.

The balance between miRNA and target abundance and how they may regulate each other has been further complicated by recent findings of competing

endogenous RNAs (ceRNAs), also known as miRNA sponges (Wang et al., 2015). A quantitative study on miRNA and targets per cell finds the ratio of miRNA:target abundance can greatly influence the effect of ceRNAs. For example, low-abundant miRNAs with highly abundant targets predominantly bind high-affinity sites, due to the high competition for that miRNA. Thus, the effect of ceRNA is low, unless that ceRNA also contain high-affinity binding sites. High-abundance miRNAs are not susceptible to ceRNA titration because there are plentiful amounts of that miRNA (Bosson et al., 2014; Guil & Esteller, 2015).

Any RNA harboring a miRNA target can be considered a ceRNA for that miRNA, including pseudogenes, circular RNAs (circRNAs), and long non-coding RNAs (lncRNAs). A pseudogene is a copy of a gene that is non-functional, due to accumulated mutations. A recent study finds the PTEN pseudogene PTENP1 is actively targeted by the same miRNAs as the PTEN gene, due to perfectly conserved seed matches for several PTEN-targeting miRNA families (Poliseno et al., 2010). A more recent study in mice finds overexpression of a conserved BRAF pseudogene, Braf-rs1, leads to ceRNA-mediated microRNA sequestration and development of aggressive malignancy (Karreth et al., 2015). These examples show how miRNA target abundance and competition for miRNA binding can be important for maintaining cellular homeostasis.

CircRNAs arise as products of backsplicing; the 3’ splice site of an upstream intron, interacts with the 5’ splice site of a downstream intron, causing the exons in between to ligate and circularize. Recent studies show circRNAs can regulate miRNA abundance by acting as a sponge for that miRNA, lowering

the effects of repression on the targets of that miRNA (Hansen et al., 2013; Guil & Esteller, 2015). CircRNAs can contain many imperfect binding sites for the miRNAs they sponge, however circRNAs are only degraded when miRNAs bind to sites with near-perfect complementarity. For example, a circRNA identified in humans and mice, ciRS-7, contains more than 70 conserved miR-7 miRNA target sites and is associated with AGO in a miR-7 dependent manner. However, miR-671 can bind perfectly to ciRS-7 and subject it to endonucleolytic cleavage (Hansen et al., 2013; Guil & Esteller, 2015). Interestingly, since circRNAs arise from coding regions, many of the annotated miRNA binding sites in coding regions may actually represent miRNA-circRNA binding instead (Memczak et al., 2013; Guil & Esteller, 2015).

Besides competing, some ncRNAs, including miRNAs can actually bind to and block or enhance processing of pri-miRNAs. In our lab, we found that *C. elegans* let-7 binds to pri-let-7 and enhances processing in an ALG-1 dependent manner (Zisoulis et al., 2012). In mice, miR-709 binds to a site in pri-miR-15a/16-1 and blocks processing (Tang et al., 2012). LncRNAs can also regulate expression of miRNAs. The human lncRNA Uc.283+A has been shown to block pri-miR-195 processing by binding to the stem-loop and directly inhibiting Microprocessor binding (Liz et al., 2014). In another example, in humans, pri-miR-484 processing is blocked when bound by miR-361 (Wang et al., 2014). Levels of miR-361 are also regulated by the lncRNA MDRL, which acts as a miR-361 sponge. When MDRL is highly expressed, miR-361-mediated silencing of pri-miR-484 is relieved allowing continued Microprocessor processing.

Overall, this summary of recent studies demonstrating miRNA regulation is just the tip of the iceberg. For example, the recent genome-wide annotation of pri-miRNAs in humans will surely lead to a better understanding of transcriptional and post-transcriptional regulation of pri-miRNA expression and processing (Chang et al., 2015). MiRNAs primarily function in the cytoplasm to regulate mRNA targets. However, examples of miRNAs regulating pri-miRNAs, which are nuclear, begs the question, what determines whether a miRNA is brought into nucleus? In addition, studies on newly described ncRNAs, like circRNAs and lncRNAs, may have regulatory themes that overlap with miRNA regulation. Finally, Chapters III-V describe my research identifying novel miRNA regulators in *C. elegans.* My work highlights the importance of understanding pri-, pre-, and mature forms of miRNAs in order to pinpoint mechanisms of regulation. In addition, I show that secondary structure, and its rearrangements due to the act of splicing, can have a large impact on the processing of pri-miRNAs.

Chapter II

Let's Make It Happen: The Role of let-7 MicroRNA in Development

A. Abstract

Non-coding RNAs have emerged as an integral part of post-transcriptional gene regulation. Among that class of RNAs are the microRNAs (miRNAs), which post-transcriptionally regulate target mRNAs containing complementary sequences. The broad presence of miRNAs in lower eukaryotes, plants, and mammals highlights their importance throughout evolution. MiRNAs have been shown to regulate many pathways, including development, and disruption of miRNA function can lead to disease (Jiang et al., 2009; Ivey & Srivastava, 2010). Although the first miRNA genes were discovered in the nematode, *Caenorhabditis elegans*, almost twenty years ago, the field of miRNA research began when they were found in multiple organisms a little over a decade ago. (Lee et al., 1993; Wightman et al., 1993; Pasquinelli et al., 2000; Lagos-Quintana et al., 2001; Lau et al., 2001; Lee & Ambros, 2001) Here we review one of the first characterized miRNAs, let-7, and describe its role in development and the intricacies of its biogenesis and function.

B. Introduction

MiRNAs distinguish themselves from other small non-coding RNAs by several unique features. Since miRNAs and the small interfering RNAs (siRNAs) that function in RNA interference (RNAi) associate with common Argonaute

proteins and are functionally indistinct in some species, these RNAs are primarily classified based upon their biogenesis pathways. MiRNA genes are scattered throughout the genome, in intra- and inter-genic positions, and are transcribed as single stranded RNAs capable of forming stem loops that contain mismatches and bulges (FIG 2.1) (Kim et al., 2009b). The stem loop structures are processed from the nascent transcripts into ~65 nt hairpin shaped RNAs that undergo final processing to the ~21 nt mature miRNAs (Krol et al., 2010). In contrast, siRNAs are usually produced from convergent transcripts forming long double-stranded RNAs that serve as substrates for RNase processing to the eventual 20-25 nt forms (Czech & Hannon, 2011). Perfect double stranded RNAs from exogenous and endogenous sources generate exo- and endo-siRNAs, respectively. In animals, miRNAs and siRNAs are also often characterized by distinct mechanisms for regulating gene expression. The miRNAs typically bind to the 3' untranslated regions (UTRs) of target messenger RNAs (mRNAs). Binding is usually imperfect, and target mRNAs are either translationally repressed or deadenylated and degraded (Bartel, 2009; Fabian et al., 2010). This is in contrast to siRNAs, which form perfectly complementary bonds to any region in target mRNAs and cause degradation by cleavage at the binding site. This distinction means that one miRNA can regulate multiple mRNAs with non-identical target sites, while siRNAs would be limited to targeting mRNAs with sites of perfect complementarity (Lim et al., 2005). However, off-target RNAi effects are often attributable to partial pairing between siRNAs and unintended target sites, and miRNAs can direct mRNA cleavage when presented with perfectly

complementary target sequences (Doench et al., 2003; Zeng et al., 2003; Alemán et al., 2007).

Typically, miRNA biogenesis begins with transcription by RNA Polymerase II, either through an independent promoter, or as part of a host gene where the miRNA is embedded within an intron of a protein-coding gene (FIG 2.1) (Kim et al., 2009b). Some miRNAs are closely arranged in the genome and considered a "cluster" when they are synthesized as part of a common transcript (Lau et al., 2001; Lee & Ambros, 2001; Lim et al., 2003). The capped and polyadenylated transcripts, known as primary miRNAs (pri-miRNAs), contain the stem-loop structure that houses the functional ~21 nt miRNA sequence (Kim et al., 2009b). The RNase III enzyme Drosha, working in a complex with Pasha (also known as DGCR8) excises the ~65 nt hairpin, forming what is known as the precursor miRNA (pre-miRNA). A non-canonical Drosha independent mechanism for producing pre-miRNAs was recently discovered in mammalian cells, *Drosophila melanogaster,* and *C. elegans* (Berezikov et al., 2007; Okamura et al., 2007; Ruby et al., 2007a). In this pathway, pre-miRNAs derive from debranching of short introns excised from pre-mRNAs. The spliced introns have characteristics of miRNA precursors: they are ~65 nt and can fold into hairpins that contain 5' monophosphate and 3' hydroxyl residues. Computational and experimental analyses of short RNA transcripts in flies, nematodes, and mammals have so far revealed few mirtrons, with 14 found in *C. elegans* (Berezikov et al., 2007; Chung et al., 2011). In vertebrates and flies, pre-miRNAs are transported from the nucleus to the cytoplasm by a nucleocytoplasmic transport factor, Exportin-5,

where the RNase III enzyme Dicer removes the loop region, leaving ~21 nt double-stranded RNA fragments with 2 nt 3' end overhangs (Kim et al., 2009b; Krol et al., 2010). Curiously, *C. elegans* lacks an obvious homolog of Exportin-5 and the cellular location of miRNA processing events is yet to be determined in nematodes (Bohnsack et al., 2004; Murphy et al., 2008).

The mature miRNA strand, also known as the guide strand, is then loaded onto an Argonaute (Ago) protein, a key factor in the RNA-induced silencing complex (RISC) (FIG 2.1) (Kim et al., 2009b; Krol et al., 2010). The leftover strand, called the star strand, is degraded by an unclear mechanism. Historically, the more abundant mature miRNA species has been named the guide and the less frequently detected one the star strand. However, the abundance of one strand versus the other has been found to reverse under different conditions and, thus, the term star strand does not always indicate a non-functional processing byproduct. In fact, both halves of the Dicer cleavage product, albeit at unequal levels, are often incorporated into Ago complexes (Okamura et al., 2008; Zisoulis et al., 2010). The bias for loading one strand versus the other is affected by the thermodynamic stability of the duplex termini, with the strand containing the weaker paired 5’ end being favored, and the identity of the first nucleotide (Czech & Hannon, 2011).

The 27 Ago family proteins in *C. elegans* exert different roles in small RNA pathways (Yigit et al., 2006). The Argonaute-Like-Genes 1 and 2 (alg-1, alg-2) are loaded with miRNAs and direct translational silencing or deadenylation of target mRNAs (Steiner et al., 2007; Zhang et al., 2007; Ding & Groszhans, 2009;

Wu et al., 2010; Zisoulis et al., 2010). Most of the other Ago proteins function in the exo- and endo-siRNA pathways, which involve target mRNA cleavage (Yigit et al., 2006). The choice of which Ago is loaded with particular small RNA species is not entirely clear but seems to depend on the origin and structure of the small RNA duplex (Steiner et al., 2007; Jannot et al., 2008).

MiRNAs serve as guides to lead the RISC complex to regulatory targets. The idea that miRNAs use limited base-pairing complementarity to recognize target sites was originally described for the first discovered miRNA, lin-4, and its target, *lin-14* (Lee et al., 1993; Wightman et al., 1993). Several partially complementary sites in well-conserved regions of the *lin-14* 3'UTR were predicted to pair with the lin-4 miRNA. The next identified miRNA, let-7, and 3'UTR sequences in its target, *lin-41*, also exhibited partial base-pairing capacity, further implicating the functionality of imperfect duplexes between miRNAs and target sites (Reinhart et al., 2000). This pattern holds true for the hundreds of animal miRNAs now recognized, but differs from the capacity of plant miRNAs to typically base pair perfectly with target mRNAs (Bartel, 2009). Although the inability of animal miRNAs to form antisense pairs with mRNAs complicates the assignment of miRNAs to specific targets, several parameters have emerged as useful predictors of these interactions. A common motif is the ability of the first 2-7 nt from the 5' end of the miRNA to perfectly pair with mRNA sequences (Lewis et al., 2003; Brennecke et al., 2005; Krek et al., 2005; Lewis et al., 2005). This part of the miRNA is known as the "seed" sequence, and the complementary site in the mRNA is called the "seed match" (Lewis et al., 2003). The requirements for

conservation, 3’ UTR position, and structural accessibility are also often used to restrict predicted target sites (Bartel, 2009). There are examples of targets that lack seed matches and instead utilize 3’ supplementary, where extensive complementarity with the 3’ end of the miRNA compensates, or centralized pairing conformations (Bartel, 2009; Shin et al., 2010).

The first characterized miRNA target sites were all found to reside in 3’UTR sequences (Lee et al., 1993; Wightman et al., 1993; Moss et al., 1997; Reinhart et al., 2000; Slack et al., 2000). Computational analyses often focus on conserved regions in 3’UTRs and take into account features like the number of target sites, their position relative to the stop codon and polyA signal, structural accessibility and whether they exist in A/U rich areas, to better predict miRNA target sites (Bartel, 2009). Surprisingly, an experimentally based genome-wide analysis of *C. elegans* Argonaute binding sites revealed that ~50% occur in coding exons and ~35% in 3’UTRs of bound mRNAs (Zisoulis et al., 2010). This is consistent with similar analyses in mammalian cells, where about half of the Argonaute binding sites were found in coding exon regions (Chi et al., 2009; Hafner et al., 2010). While some target sites in coding exons have been shown to confer miRNA function, the general efficacy of regulation in translated regions of mRNAs may be reduced compared to that in 3’UTRs (Kloosterman et al., 2004; Gu et al., 2009). Recent studies trying to understand the functionality of miRNA target sites in coding regions reveal a synergistic effect, where there are stronger effects on mRNAs bound by miRNAs in the 3’UTR and coding exons, though they are not as strong as the effects when there are two target sites in the 3’UTR

(Fang & Rajewsky, 2011).

Association of the miRNA complex with a target mRNA results in downregulation of the protein expression through mechanisms that are not entirely understood. There is an ongoing debate about whether the primary mode of regulation is translational repression of the mRNA, or deadenylation followed by degradation of the mRNA (FIG 2.2) (Krol et al., 2010; Djuranovic et al., 2011; Huntzinger & Izaurralde, 2011). Original studies of *lin-14* repression by lin-4 miRNA detected little change in target mRNA levels or polysome loading and concluded that regulation involved a post-transcriptional mechanism acting at a stage after translational initiation (Wightman et al., 1993; Olsen & Ambros, 1999). Subsequent work showed that mRNA levels for the *lin-14*, *lin-28* and *lin-41* target mRNAs are down-regulated in response to miRNA regulation, and there is also a shift on polysomes indicative of inhibition of translation initiation (Bagga et al., 2005; Ding & Groszhans, 2009). It remains to be determined if target mRNA degradation is a cause or consequence of halted translation. Argonaute interacts with GW182 proteins (AIN-1 and AIN-2 in *C. elegans*), which recruit factors capable of interfering with translation or promoting deadenylation of the target mRNA (Tritschler et al., 2010). Thus, depending on the cell type or conditions, different factors may be available to regulate miRNA targets through diverse mechanisms, ultimately resulting in diminished protein production.

C. The discovery of miRNAs

C.1 lin-4 miRNA

Development of *C. elegans* progresses through four larval stages before reaching adulthood. The temporal fate of each cell division has been mapped and genes that regulate the timing of these events are part of the heterochronic pathway (Sulston & Horvitz, 1977). Mutations in heterochronic genes cause either precocious development, where later larval cell fates occur prematurely, or retarded development, where early larval cell fates reoccur in later stages (Chalfie et al., 1981; Ambros & Horvitz, 1984; Ambros, 1989). These type of mutations cause various developmental abnormalities, some of which result in lethality. While investigating the genes important for larval development, researchers found that *lin-4* loss of function (lf) mutations recapitulated phenotypes observed in worms with *lin-14* gain of function (gf) mutations (Ambros, 1989). These mutant worms repeat L1 cell fates, causing the absence of adult structures, like the vulva and differentiated hypodermal seam cells (Chalfie et al., 1981). Loss of *lin-14* activity has the opposite phenotype, where later larval fates are precociously expressed in the first larval stage (Ambros & Horvitz, 1984). Additional genetic and molecular experiments led to the model that LIN-14 protein promotes the first larval stage cell fates and then is downregulated by *lin-4* activity so that later larval fates can proceed (Ruvkun & Giusto, 1989; Arasu et al., 1991).

The surprising identity of the *lin-4* gene product revealed a novel mechanism for regulation of *lin-14*. Through arduous mapping and genetic rescue experiments, researchers in the Ambros lab narrowed down the location of *lin-4* to a 693 nt region, which was later shown to contain the primary lin-4

transcripts that lacked apparent protein-coding potential (Lee et al., 1993; Bracht et al., 2010). Instead, this region was found to express 22 nt and 61 nt RNA products that are now recognized as the mature and precursor forms of lin-4 miRNA, respectively. Since the Ruvkun lab had established that *lin-14* is regulated at the post-transcriptional level through elements in its 3'UTR, the possibility of lin-4 base-pairing to these regions was realized by both labs (Lee et al., 1993; Wightman et al., 1993). There are seven potential lin-4 target sites in the *lin-14* 3'UTR that are predicted to form non-identical partial duplexes (Wightman et al., 1993). Using genetic mutants, molecular experiments and reporter assays, the Ambros and Ruvkun labs defined a novel form of gene regulation whereby the lin-4 RNA base-pairs to specific sites in the 3'UTR of the *lin-14* mRNA, resulting in down-regulation of LIN-14 protein expression (Lee et al., 1993; Wightman et al., 1993). Whether or not this mechanism of gene regulation would be an isolated example was an open question since the small RNA, lin-4, and its target gene, *lin-14*, were not obviously conserved in species beyond *Caenorhabditae*.

C.2 let-7 miRNA

Four years after the discovery of lin-4 and its target, another protein-coding gene in the heterochronic pathway was found to be under the direct control of *lin-4*. The *lin-28* 3'UTR contains a single *lin-4* binding site that mediates repression of protein expression at the post-transcriptional level (Moss et al., 1997). With two different genes regulated in a similar manner by the same RNA, researchers suspected that other small RNA regulators might exist in *C. elegans*

to control developmental timing. Characterization of mutations that cause delayed temporal fates in the last larval stages led to the *let-7* gene (FIG 2.3a) (Reinhart et al., 2000). The *let-7* gene was originally named based on the lethal phenotype where worms rupture through the vulva and die before reaching adulthood (FIG 2.3b) (Meneely & Herman, 1979). Genetic mapping, rescue experiments, and RNA expression analyses established that *let-7* was the second example of a tiny RNA gene in the developmental timing pathway (Reinhart et al., 2000). Like *lin-4*, the *let-7* gene expressed a ~21 nt RNA that could potentially derive from a processed intramolecular hairpin structure (Lee et al., 1993; Reinhart et al., 2000). Coincidentally, studies of the *lin-41* gene in the same lab showed that mutations in this heterochronic gene resulted in precocious adoption of differentiated cells fates, a phenotype opposite to that of *let-7* mutants (FIG 2.3a) (Reinhart et al., 2000; Slack et al., 2000). Using a variety of biochemical and genetic techniques, the let-7 RNA was demonstrated to regulate expression of *lin-41* through two sites of partial complementarity in its 3'UTR (FIG 2.4) (Reinhart et al., 2000; Slack et al., 2000; Vella et al., 2004a). Thus, the accumulation of mature let-7 RNA at the end of the third larval stage causes down-regulation of LIN-41 protein expression, allowing the adoption of later larval and adult fates.

C.3 let-7 miRNA in other organisms

Unlike *lin-14*, homologs of *lin-41* were found to exist in Drosophila, mouse and other animals, raising the question of whether its small RNA regulator would also be conserved (Slack et al., 2000). With whole-genome databases becoming

available, the Ruvkun lab identified sequences in the Drosophila and human genomes that matched the mature let-7 sequence (Pasquinelli et al., 2000). Northern blot analyses of a wide range of animal species showed that let-7 RNAs are present broadly in bilaterians, but apparently not in basal metazoans or non-animal organisms. Remarkably, potential let-7 binding sites exist in the 3'UTRs of *lin-41* homologs in several species and, in some cases, regulation by let-7 has been demonstrated (FIG 2.4) (Pasquinelli et al., 2000; Kanamoto et al., 2006; Maller-Schulman et al., 2008). The discovery of let-7 across species fueled the search for other such RNA genes, and soon many small RNAs were identified in worms, flies, human cell lines, mice, and plants through cloning methods (Lagos-Quintana et al., 2001; Lau et al., 2001; Lee & Ambros, 2001; Lagos-Quintana et al., 2002; Reinhart et al., 2002). A signature of these ~21 nt RNAs is that they are processed from ~65 nt stem-loop precursors and this feature distinguishes the class of miRNAs from other tiny ncRNAs (Ambros et al., 2003).

While there is one *let-7* gene in *C. elegans*, multiple let-7 genes exist in most vertebrates. The human genome includes three let-7 genes that produce mature RNAs of identical sequence (let-7a-1, let-a-2, and let-7a-3) as well as nine others that differ by one or more nucleotides (FIG 2.5). Since the seed sequence, nucleotides 2-7, of the miRNA plays a key role in target recognition, miRNAs with identical seeds are often considered part of a family (Lim et al., 2005). MiRNAs within the same family potentially regulate common targets, thus explaining why mutation of one member sometimes results in no discernible phenotypes (Miska et al., 2007). In *C. elegans*, *let-7* has six "sister" genes that

share 5'-end sequences (FIG 2.5). Nonetheless, mutation of *let-7* alone is sufficient to cause developmental abnormalities and lethality in *C. elegans* (FIG 2.3) (Reinhart et al., 2000). Thus, members of a miRNA family do not necessarily compensate for each other. In mammals, some of the let-7 family members exhibit specific expression patterns and have distinct targets, indicating that they may play different roles in diverse biological pathways (Boyerinas et al., 2010; Chiang et al., 2010).

D. The role of let-7 in *C. elegans* development

D.1 Biogenesis of let-7

Consistent with its role in promoting later larval cell fates, mature let-7 miRNA starts to accumulate midway through the third larval stage of development (FIG 2.6) (Reinhart et al., 2000; Van Wynsberghe et al., 2011). The *let-7* gene expresses three let-7 primary transcripts, two unspliced variants with differing transcriptional start sites and one SL1 *trans*-spliced isoform (Bracht et al., 2004). The transcripts are capped and polyadenylated via a conserved polyA signal located ~670 nt downstream of the mature let-7 sequence. Surprisingly, expression of the let-7 primary transcripts initiates at the end of the first larval stage and oscillates during each subsequent stage (FIG 2.6) (Van Wynsberghe et al., 2011). This expression pattern is regulated at the transcriptional level, as indicated by reporter genes containing the let-7 promoter fused to GFP (Johnson et al., 2003; Esquela-Kerscher et al., 2005; Martinez et al., 2008; Van Wynsberghe et al., 2011). These reporter studies also show that let-7 appears to

be transcribed in most somatic tissues, consistent with the broad expression pattern of let-7 in differentiated cell types across animal species.

The search for *let*-7 transcription elements identified a *cis*-acting sequence called the temporal regulatory element, TRE, located 223 nt and 1064 nt upstream from the two transcriptional start sites (FIG 2.7) (Johnson et al., 2003; Bracht et al., 2004). Internal deletions of the let-7 promoter in a GFP reporter assay identified this 116 nt region, which contains a 9 nt inverted repeat and is conserved in *C. briggsae*, as necessary for expression of GFP in the hypodermal seam cells (Johnson et al., 2003). Compared to wild-type, *let-7* transgenes lacking the TRE exhibit reduced rescue activity. The proteins that regulate transcription of let-7 through the TRE, as well as the elements responsible for the oscillating transcription pattern and expression in other tissues are yet to be identified.

So far, two transcription factors have been found to regulate the expression of *let-7* in *C. elegans*. The *hbl-1* gene encodes a zinc-finger transcription factor that shares homology with the *Drosophila* Hunchback gene (Fay et al., 1999; Abrahante et al., 2003; Lin et al., 2003). HBL-1 is predicted to bind to an A-rich sequence 18 nt downstream of the TRE and repress transcription of *let-7* in the hypodermal seam cells (FIG 2.7) (Roush & Slack, 2009). The nuclear hormone receptor DAF-12 regulates the transcription of *let-7* in a hormone dependent manner (Bethke et al., 2009; Hammell et al., 2009a). In the absence of ligand, DAF-12 represses the expression of *let-7* and several of its sister miRNAs. When bound to ligand, DAF-12 activates expression of some

let-7 family members. DAF-12 response elements have been identified in the promoters of *mir-241* and *mir-84* but direct interaction with the *let-7* promoter remains to be demonstrated (Bethke et al., 2009). Interestingly, both *hbl-1* and *daf-12* are targets of regulation by *let-7* family miRNAs (Abrahante et al., 2003; Lin et al., 2003; Abbott et al., 2005; Großhans et al., 2005). Through multiple 3'UTR complementary sites, the let-7 sisters, mir-48, mir-84, and mir-241, initiate repression of *hbl-1* expression during the transition from the second to the third larval stage (Abbott et al., 2005). These miRNAs are expressed by the second larval stage, one stage earlier than let-7, thus, providing a mechanism to reduce HBL-1 levels and allow for transcription of let-7 in the seam cells (Abbott et al., 2005; Roush & Slack, 2009). Likewise, the let-7 sisters also target *daf-12* for down-regulation at the L3 stage, which may promote transcription of let-7 in some tissues (Bethke et al., 2009; Hammell et al., 2009a). Expression of mature let-7 in L3 augments repression of *hbl-1* and *daf-12*, adding to the feedback loop of transcriptional and miRNA-mediated control (Abrahante et al., 2003; Lin et al., 2003; Abbott et al., 2005; Großhans et al., 2005; Bethke et al., 2009; Hammell et al., 2009a).

In *C. elegans*, ~70% of all mRNAs are trans-spliced to one of two 22 nt splice leader (SL) sequences, SL1 or SL2 (Blumenthal, 2012). The *trans*-splicing reaction is carried out by a spliceosome that functions similarly to the *cis*-splicing complex, which removes introns, except that the 5' end of the SL RNA acts as the 5' exon and is ligated to a 3' splice site downstream of the 5' cap in the acceptor mRNA (Blumenthal, 2012). The role of *trans*-splicing is not fully

understood, but is thought to aid in nuclear export and translation. Thus, it was surprising to find SL1-spliced let-7 primary transcripts (Bracht et al., 2004). The 3' splice site required for *trans*-splicing is conserved in other nematodes, and disruption of this signal impairs Drosha processing of the primary transcripts (Bracht et al., 2004). Secondary structure predictions by mfold suggest that splicing alters the context around the precursor hairpin, perhaps making it a better substrate for subsequent processing. Although SL-mediated *trans*-splicing is not found in most vertebrates, the primary transcripts of several human let-7 genes are part of spliced transcripts. The mechanistic role of splicing in let-7 biogenesis in worms, and perhaps other species, is yet to be resolved.

In vertebrates and Drosophila, Exportin-5 delivers miRNA precursors from the nucleus to the cytoplasm for Dicer processing (FIG 2.1) (Yi et al., 2003; Bohnsack et al., 2004; Lund et al., 2004). The apparent absence of an Exportin-5 homolog in *C. elegans* suggests that other cellular transport factors are involved in miRNA biogenesis in this organism (Bohnsack et al., 2004; Murphy et al., 2008). The nuclear export receptor XPO-1, as well as the nuclear cap-binding complex (CBC), has been implicated in the *let-7* biogenesis pathway (FIG 2.7) (Bussing et al., 2010). Depletion of either XPO-1 or either of the two subunits of the CBC results in reduced levels of precursor and mature let-7 and accumulated levels of pri-let-7. Considering the role of XPO-1 and the CBC in mediating nuclear export of m^7G-capped U snRNAs (Hutten & Kehlenbach, 2007), one possibility is that capped primary miRNA transcripts are also substrates for transport to the cytoplasm. Since the cellular location of Drosha is not known in

worms, nuclear export of pri-let-7 transcripts by XPO-1 and CBC could be important for processing (Bussing et al., 2010).

The temporal expression of pri-let-7 suggests complex regulation at both the transcriptional and post-transcriptional levels (FIG 2.7). Production of primary let-7 in the L1 and L2 stages is not coupled to accumulation of precursor and mature miRNA (FIG 2.6) (Van Wynsberghe et al., 2011). The RNA binding protein LIN-28 mediates this phase of post-transcriptional regulation (Lehrbach et al., 2009; Van Wynsberghe et al., 2011). LIN-28 binds endogenous let-7 pri-miRNAs co-transcriptionally and blocks Drosha processing (FIG 2.7) (Van Wynsberghe et al., 2011). Since *lin-28* is controlled by lin-4 miRNA, expression of this miRNA at the end of L1 results in a steady decline of LIN-28 protein and, thus, relief of let-7 processing inhibition by the third larval stage (Moss et al., 1997; Lehrbach et al., 2009; Van Wynsberghe et al., 2011). Post-transcriptional regulation of let-7 by LIN-28 was originally discovered in mammalian cells (Wulczyn et al., 2007; Heo et al., 2008; Newman et al., 2008; Viswanathan et al., 2008; Viswanathan & Daley, 2010). In addition to preventing Drosha processing of let-7 pri-miRNAs, LIN-28 also recruits TUT4 (PUP-2 in *C. elegans*) to add a 3' end U-tail to let-7 precursors, which blocks Dicer processing and promotes destabilization of the RNAs (Heo et al., 2008; Hagan et al., 2009; Heo et al., 2009). Comparable to the role of LIN-28 in blocking expression of mature let-7 early in worm development, high levels of Lin28 in mammalian stem cells also prevents accumulation of let-7 miRNAs.

For many miRNAs, including let-7, only one-half of the duplex that results

from Dicer processing accumulates as the mature miRNA, while the other strand is presumably degraded. One factor that appears to influence miRNA stability is the association with Argonaute proteins (Kai & Pasquinelli, 2010). The ability to base pair with target sequences influences the accumulation of mature miRNAs through a mechanism dubbed target-mediated miRNA protection (TMMP) (Chatterjee et al., 2011). Thus, miRNA passenger strands that lack target sites are released from Argonaute and subject to degradation. The 5'→3' exonucleases XRN-1 and XRN-2 degrade miRNAs that lose association with Argonaute (FIG 2.7) (Chatterjee & Grosshans, 2009; Chatterjee et al., 2011). Although the accumulation of let-7 miRNA can be regulated by the availability of target site interactions in *C. elegans*, whether TMMP functions in other organisms is yet to be determined.

While Argonaute and the AIN-1/-2 (GW182-related) proteins are required for miRNA function, several cofactors have been found to modulate miRNA activity. The TRIM-NHL family of proteins, which contain TRIM (tripartite-containing motif; RING, B-Box, coiled-coil) and NHL domains (named after the first three proteins discovered to contain this motif; NCL-1, Ht2A, and LIN-41), includes a broadly conserved class of proteins involved in diverse biological pathways (Slack & Ruvkun, 1998). Two members of this class, NHL-2 in *C. elegans* and TRIM32 in mice, enhance the ability of let-7 miRNA to regulate target genes (Hammell et al., 2009b; Schwamborn et al., 2009). These proteins associate with Argonaute complexes and stimulate the repressive activity of miRNAs on certain targets through an unknown mechanism (FIG 2.7). The small

ribosome subunit protein, RPS-14, also co-precipitates with ALG-1 but, instead, seems to negatively regulate let-7 function in *C. elegans* (FIG 2.7) (Chan & Slack, 2009). Thus, the effectiveness of let-7 in target regulation is influenced not only by the level of the miRNA but also by the presence of specific miRNA complex accessory proteins.

D.2 Identification of let-7 targets

The first miRNA targets were identified as genetic suppressors of miRNA mutant phenotypes (Lee et al., 1993; Wightman et al., 1993; Moss et al., 1997; Reinhart et al., 2000; Slack et al., 2000). Since miRNAs negatively regulate target gene expression, loss of function mutations that reduce miRNA target levels can compensate for the absence of the miRNA. For example, the reiteration of hypodermal seam cell divisions and rupturing vulva phenotypes displayed by *let-7* mutants are suppressed in worms that also contain mutations in *lin-41*, a direct target of *let-7* regulation (Reinhart et al., 2000; Slack et al., 2000). In addition to the isolation of miRNA suppressors through traditional forward genetic screens, RNAi has enabled high throughput screening to identify potential miRNA targets. Several transcription factors, including *daf-12,* emerged as *let-7* targets from an RNAi screen of candidate genes that contain predicted binding sites (Großhans et al., 2005). The ability to detect single miRNA targets through genetic approaches suggests that miRNAs have a few key targets that are responsible for phenotypes associated with loss of the miRNA or that the targets are part of interconnected pathways that are sensitive to changes in expression of individual genes.

Various computational approaches have predicted dozens to hundreds of let-7 targets in *C. elegans* (Enright et al., 2003; John et al., 2004; Lall et al., 2006; Miranda et al., 2006; Ruby et al., 2006; Kertesz et al., 2007; Ruby et al., 2007b; Betel et al., 2008; Hammell et al., 2008). While most programs incorporate seed pairing, RNA structure, free energy, and conservation of target sites, the stringency of these parameters accounts for some of the variation among the predictions (Bartel, 2009). Computational predictions are often tested by fusing the target 3'UTR to a reporter gene and determining if it is regulated in the presence of the miRNA. In many cases disruption of seed pairing between the miRNA and target is sufficient and necessary to block regulation (Doench & Sharp, 2004; Kloosterman et al., 2004; Brennecke et al., 2005; Lai, 2005). However, some target sites appear more complicated. For example, systematic analysis of the *lin-41* 3'UTR revealed that multiple elements within and around the two let-7 binding sites are required for regulation (Vella et al., 2004a; Vella et al., 2004b).

Genome wide experimental approaches that analyze changes in gene expression in response to specific miRNAs have proven effective for identifying miRNA targets. After the recognition that established miRNA targets are regulated at the level of mRNA stability (Bagga et al., 2005; Wu & Belasco, 2005; Giraldez et al., 2006), microarray technology was embraced as a high-throughput method to detect potential targets (Lim et al., 2005; Schmitter et al., 2006; Sood et al., 2006; Huang et al., 2007). In the first of these studies, introduction of miR-1 or miR-124 into HeLa cells resulted in down-regulation of genes that were

enriched for complementary binding sites to the appropriate miRNA (Lim et al., 2005). Microarray analysis of gene expression changes during early mouse development, as let-7 miRNA starts to accumulate, coupled with target prediction programs, indicate that several early embryonic genes, including HMGA2 and IMP-1, are directly repressed by this miRNA (Boyerinas et al., 2008). Some of these targets are re-expressed in cancer cells when regulation by let-7 is lost (Boyerinas et al., 2010).

Monitoring global changes in protein expression has also provided strong evidence for target regulation by specific miRNAs. Two studies that employed stable isotope labeling with amino acids in cell culture (SILAC) in conjunction with mass spectrometry identified hundreds of targets regulated by specific miRNAs (Baek et al., 2008; Selbach et al., 2008). A comparable proteomics approach performed with extracts from wild-type and *let-7* mutant worms focused on a set of predicted let-7 targets to analyze for changes in protein levels (Jovanovic et al., 2010). Each of these proteomic studies also noted varying degrees of target mRNA destabilization associated with the changes in protein levels. Recently, ribosome profiling was used to compare changes in mRNA levels to polysome association, as an indicator of translational efficiency (Guo et al., 2010). This study showed that changes in translation largely reflected the changes in target mRNA levels, leading to the conclusion that miRNA regulation is predominantly through target mRNA destabilization at least in some cell types (Guo et al., 2010).

Since miRNAs form regulatory complexes with Argonaute and AIN-1/2

(GW182) proteins, target mRNAs can be detected in association with these factors (Beitzinger et al., 2007; Easow et al., 2007; Karginov et al., 2007; Zhang et al., 2007; Hendrickson et al., 2008; Landthaler et al., 2008; Zhang et al., 2009). For example, AIN-1/2 co-precipitates with the majority of established miRNA targets in *C. elegans*, suggesting that many of the other isolated mRNAs are good candidates for regulation by the miRNA pathway (Zhang et al., 2007). An advancement over the isolation of entire mRNA targets is the ability to detect the sequence of the mRNA fragment directly bound by the miRNA complex through a method called CLIP (Cross-linking immunoprecipitation). This technique uses UV-irradiation to bond proteins covalently to nucleic acids, followed by immunoprecipitation of Argonaute complexes and deep sequencing of directly associated sequences (Zisoulis et al., 2010). CLIP studies in mouse brain, mammalian tissue culture cells, and whole worms have revealed miRISC (miRNA induced silencing complex) binding sites on a genome-wide scale (Chi et al., 2009; Hafner et al., 2010; Leung et al., 2010; Zisoulis et al., 2010). This method narrows the miRNA binding site to 50-100 nt and provides biochemical evidence that an mRNA is bound by miRISC. In *C. elegans,* many of the well-established let-7 target sites, including those in *lin-41, daf-12,* and *hbl-1*, were detected by the CLIP method (Zisoulis et al., 2010). While the exact miRNA recognition site is yet to be determined in most of the Argonaute bound sequences, several general features of miRNA targeting emerged from these studies. Sites are predominantly in coding exons and 3’UTRs, seed pairing capacity is enriched within Ago bound regions, and association of Ago with

3’UTR sites is associated with target mRNA destabilization (Chi et al., 2009; Hafner et al., 2010; Zisoulis et al., 2010).

D.3 Developmental role of let-7 and its targets

A common theme for targets of *let-7* regulation is a role in promoting cellular division and self-renewal (Büssing et al., 2008). This is consistent with the phenotype of *let-7* mutants where the seam cells fail to terminally differentiate at the appropriate time and instead continue dividing (FIG 2.3) (Reinhart et al., 2000). In *C. elegans* loss of *lin-41*, a direct let-7 target, results in premature cellular differentiation (FIG 2.3) (Reinhart et al., 2000; Slack et al., 2000). Consistent with its role in promoting the undifferentiated state in worms, the mouse homolog is concentrated in stem and early embryonic cells (Rybak et al., 2009). Although the molecular function of the *C. elegans* LIN-41 protein is not yet known, recent work in mouse cells shows that mLin41 acts as an E3 ubiquitin ligase that modifies Ago2, reducing its stability (Rybak et al., 2009). Thus, mLin41 can indirectly regulate the levels and function of mature miRNAs by antagonizing Argonaute (FIG 2.7).

As mentioned above, several targets of let-7 are transcription factors. Loss of *hbl-1* or *daf-12* activity results in precious expression of adult fates in larval stage worms (Abrahante et al., 2003; Lin et al., 2003; Großhans et al., 2005). In addition to regulation of *let-7* family members, several protein-coding genes have been identified as potential direct targets of DAF-12 and HBL-1 transcriptional control (Shostak et al., 2004; Niwa et al., 2009). How these transcriptional networks contribute to the maintenance of undifferentiated cell fates prior to

expression of *let-7* is yet to be fully explored.

Mis-regulation of *let-60/RAS* contributes to the lethal phenotype of *let-7* mutants that rupture through the vulva (FIG 2.3) (Esquela-Kerscher et al., 2005). The 3'UTR of *let-60/RAS* contains multiple conserved sites of complementarity to let-7 family miRNAs (Esquela-Kerscher et al., 2005). These miRNAs act to repress expression of *let-60/RAS* in specific vulval precursor cells to restrict RAS signaling and promote adoption of appropriate cell fates. Remarkably, functional let-7 sites were also detected in the 3'UTRs of human RAS genes. Reduced expression of mature let-7 in several types of cancers, especially lung, is often linked to increased RAS protein (Esquela-Kerscher et al., 2005; Boyerinas et al., 2010). By regulating key signaling molecules and transcription factors, let-7 family miRNAs control a large network of genes that determine the correct timing of cell fates during worm development, making this regulatory pathway indispensable.

E. The role of let-7 in development and disease across species

E.1 Stem cells and differentiation

Since several recent reviews detail the roles of mammalian let-7 miRNAs in regulating development and disease in vertebrates, only a few select highlights are presented here (Nimmo & Slack, 2009; Boyerinas et al., 2010; Ivey & Srivastava, 2010; Osada & Takahashi, 2011). Consistent with its role in promoting differentiated cell fates in *C. elegans,* let-7 miRNAs are depleted from mammalian embryonic stem cells (ESCs) (FIG 2.8) (Suh et al., 2004; Bar et al.,

2008; Laurent, 2008; Hinton et al., 2010). This expression pattern seems to be largely accomplished by LIN28, which is abundant in ESCs (Viswanathan & Daley, 2010). As described above, the LIN28 protein inhibits the accumulation of mature let-7, thus preventing cellular differentiation pathways. During stem cell differentiation, miR-125 and let-7 down-regulate expression of LIN28 through complementary sites in its 3'UTR (Wu & Belasco, 2005; Rybak et al., 2008). Thus, let-7 promotes its own expression by targeting its negative regulator LIN28. Given that let-7 is widely expressed in most somatic tissues, the role of LIN28 in inducing pluripotent stem (iPS) cells from fibroblasts may depend largely on its ability to repress let-7 and, thus, differentiation pathways. Additionally, aberrant activation of LIN28 occurs in some tumors, resulting in down-regulation of let-7 and its ability to repress targets that promote cell division (Viswanathan & Daley, 2010).

The default pathway for ESCs depleted of miRNA processing factors is self-renewal (Kanellopoulou et al., 2005; Murchison et al., 2005; Wang et al., 2007). However, introduction of mature let-7 miRNAs into such cells results in silencing of the self-renewal program, demonstrating the potent ability of these miRNAs to inhibit stem cell properties (Melton & Blelloch, 2010). This effect is not observed in normal ESCs because an opposing set of miRNAs is sufficient to maintain the self-renewal properties in the presence of let-7. Given the fundamental role of let-7 in promoting differentiation pathways, it is perhaps not surprising that inhibition of these miRNAs in mouse fibroblasts enhances the ability of reprogramming factors to produce iPS cells (Melton & Blelloch, 2010).

E.2 Cancer

Cancer results when cells fail to either differentiate properly or acquire the ability to self-renew, like stem cells. In many cases, this requires the activation of genes repressed by let-7 miRNAs. This is accomplished by down-regulation of mature let-7 miRNA levels or loss of target regulation through deletion of 3'UTR target sites, for example. Some or all of the let-7 family miRNAs are depleted in tumor compared to normal tissues for many types of cancer (Boyerinas et al., 2010). While re-expression of LIN28 explains some cases of decreased let-7 levels, other mechanisms for reducing let-7 in cancer cells are yet to be determined. The HMGA2 gene is an example of a let-7 target that escapes regulation in some types of tumors through truncation of its 3'UTR, which removes the let-7 binding sites (Lee & Dutta, 2007; Mayr et al., 2007; Shell et al., 2007). HMGA2 encodes a chromatin-associated non-histone protein that is highly expressed in embryonic tissues and undetectable is most differentiated adult cells. Chromosomal rearrangements can separate the HMGA2 3'UTR from its open reading frame, allowing for unregulated expression of the protein (Lee & Dutta, 2007; Mayr et al., 2007; Shell et al., 2007). The truncated HMGA2 gene has been shown to promote oncogenesis, underscoring the importance of regulation through its 3'UTR (Mayr et al., 2007).

The first indication that let-7 may act as a tumor suppressor came with the discovery that it has a conserved role in regulating the expression of RAS (Esquela-Kerscher et al., 2005). The three human RAS oncogenes each contain multiple let-7 binding sites in their 3'UTRs. The RAS proteins are frequently

upregulated in lung tumors, where let-7 family miRNAs tend to be downregulated (Takamizawa et al., 2004; Esquela-Kerscher et al., 2005). A direct role for let-7 in regulating RAS and preventing lung tumor growth has been demonstrated in cell culture and mouse model systems (Takamizawa et al., 2004; Esquela-Kerscher et al., 2005; Johnson et al., 2007; Esquela-Kerscher et al., 2008; Kumar et al., 2008). In addition to RAS, several oncogenes, such as IMP and MYC, have been identified as direct let-7 targets, further supporting its classification as a tumor suppressor (Boyerinas et al., 2008; Kim et al., 2009a). Although down-regulation of let-7 family miRNAs has been associated with many types of cancers, there are a few examples of increased let-7 expression in certain tumors (Boyerinas et al., 2010). Thus, the role of let-7 as a tumor suppressor will likely depend on the cellular environment and targets available for regulation.

E.3 Neuronal functions

The let-7 family miRNAs are highly expressed in neuronal cells across species, suggesting important roles in neurologic development and function (Johnson et al., 2003; Sempere et al., 2004; Smirnova et al., 2005; Caygill & Johnston, 2008; Sokol et al., 2008). Consistent with increasing expression of let-7 miRNAs during mammalian brain development, these miRNAs repress neuronal proliferation and promote differentiation programs (Nishino et al., 2008; Rybak et al., 2008; Sokol et al., 2008; Schwamborn et al., 2009; Zhao et al., 2010). In Drosophila, let-7 regulates the timing of neuromuscular remodeling events and adult behaviors (Caygill & Johnston, 2008; Sokol et al., 2008). Additional work in Drosophila has revealed a potential link between disruption of

let-7 function and Parkinson's disease. The Parkinson's disease related protein LRRK2 antagonizes the ability of let-7 to repress target genes (Gehrke et al., 2010). In particular, mis-regulation of the let-7 target *dp* results in loss of fly dopaminergic neurons, exemplifying the importance of let-7 activity in neuronal health. It is likely that other targets of let-7 that promote cell division, like *dp*, also become mis-regulated by pathogenic LRRK2, which results in inappropriate activation of cell division programs in post-mitotic neurons, leading to cell death.

F. Conclusions

Despite only about a decade's worth of research, miRNAs have emerged as key regulators of development. In particular, the let-7 miRNAs have been found to play pivotal roles in cell differentiation pathways. The strict conservation of let-7 sequences across animal species suggests an ancient function in regulating core genes that control cell division programs. The essential role of let-7 in regulating temporal development in model organisms relates to the importance of this miRNA in human disease phenotypes. In some cases, not only the general pathways but also the targets of let-7 are conserved across phylogeny. The let-7 miRNA has also been a model for understanding the complex transcriptional and post-transcriptional mechanisms that control miRNA biogenesis. The past decade introduced us to miRNAs and the basic steps for their synthesis and control of target expression. Future challenges include the elucidation of mechanisms controlling the levels of specific miRNAs, determination of features that faithfully predict miRNA target recognition in the

endogenous context, and analysis of gene networks under normal and perturbed miRNA conditions. Given its conserved and essential role in animal development, it is likely that let-7 will continue to be a focus for understanding the complexities of miRNA expression and function.

G. Acknowledgements

We thank members of the Pasquinelli lab for critical reading of this manuscript. The Pasquinelli lab is supported by the US National Institutes of Health (GM071654), Keck, Searle, V. Emerald, and Peter Gruber Foundations.

Chapter II, in full, is a reprint of the material as it appears in Current Topics in Developmental Biology, Mondol, Vanessa & Pasquinelli, Amy E., Academic Press, 2012. I was the primary author.

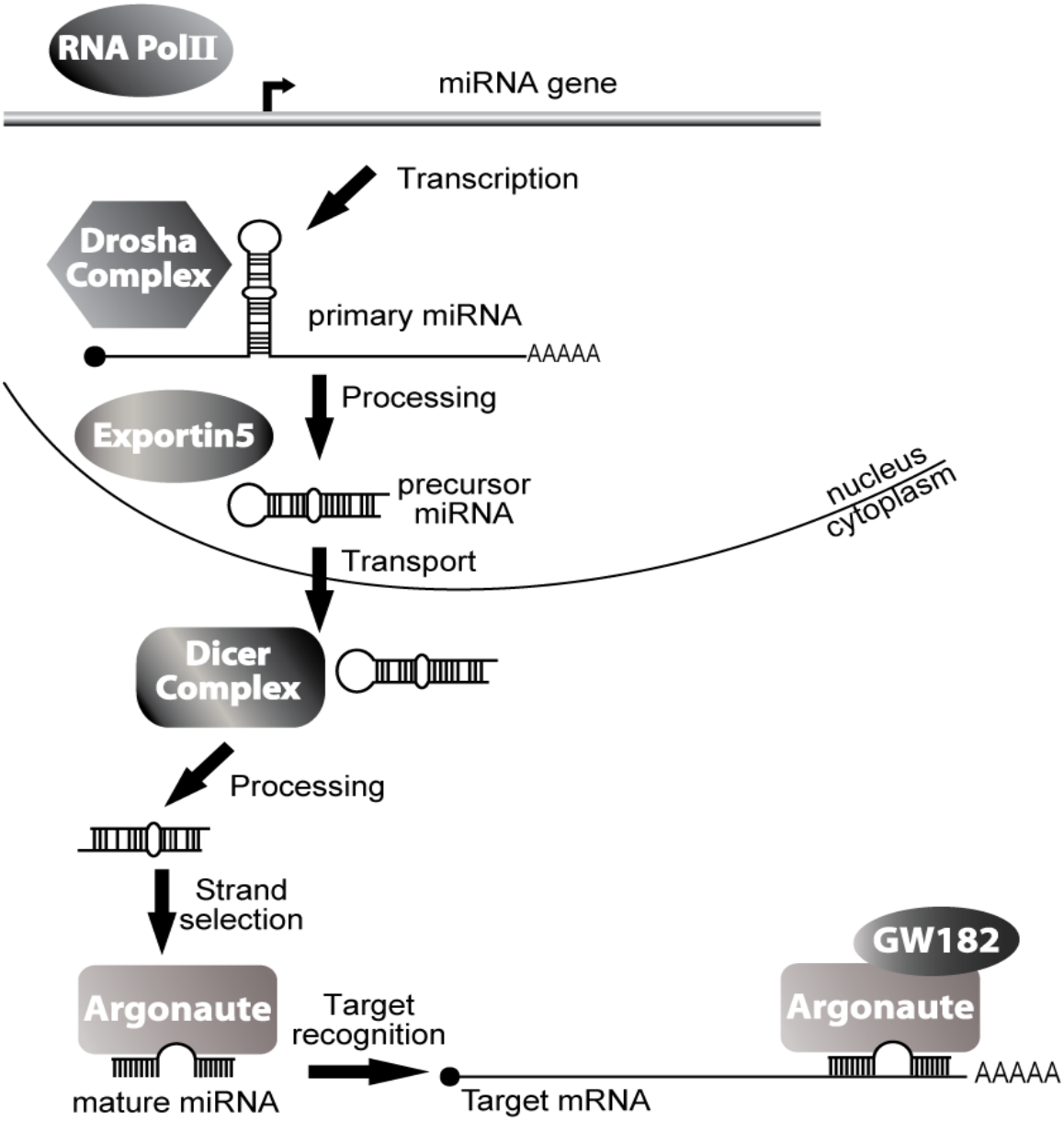


FIGURE 2.1 General miRNA Biogenesis Pathway. RNA Polymerase II (RNA PolII) transcribes miRNA genes to produce a capped and polyadenylated primary miRNA (pri-miRNA). A complex containing the RNase III enzyme Drosha excises the ~65 nt precursor miRNA (pre-miRNA), leaving a 2 nt overhang at the 3' end. In flies and vertebrates, nuclear transport protein Exportin-5, exports the pre-miRNA to the cytoplasm, where it is subsequently cleaved by the RNase III enzyme Dicer to produce a ~21nt double-stranded product that features 2 nt overhangs at both ends. The mature miRNA strand is then loaded into Argonaute (Ago), where it works as part of a complex with GW182 and other proteins to target mRNAs for repression.

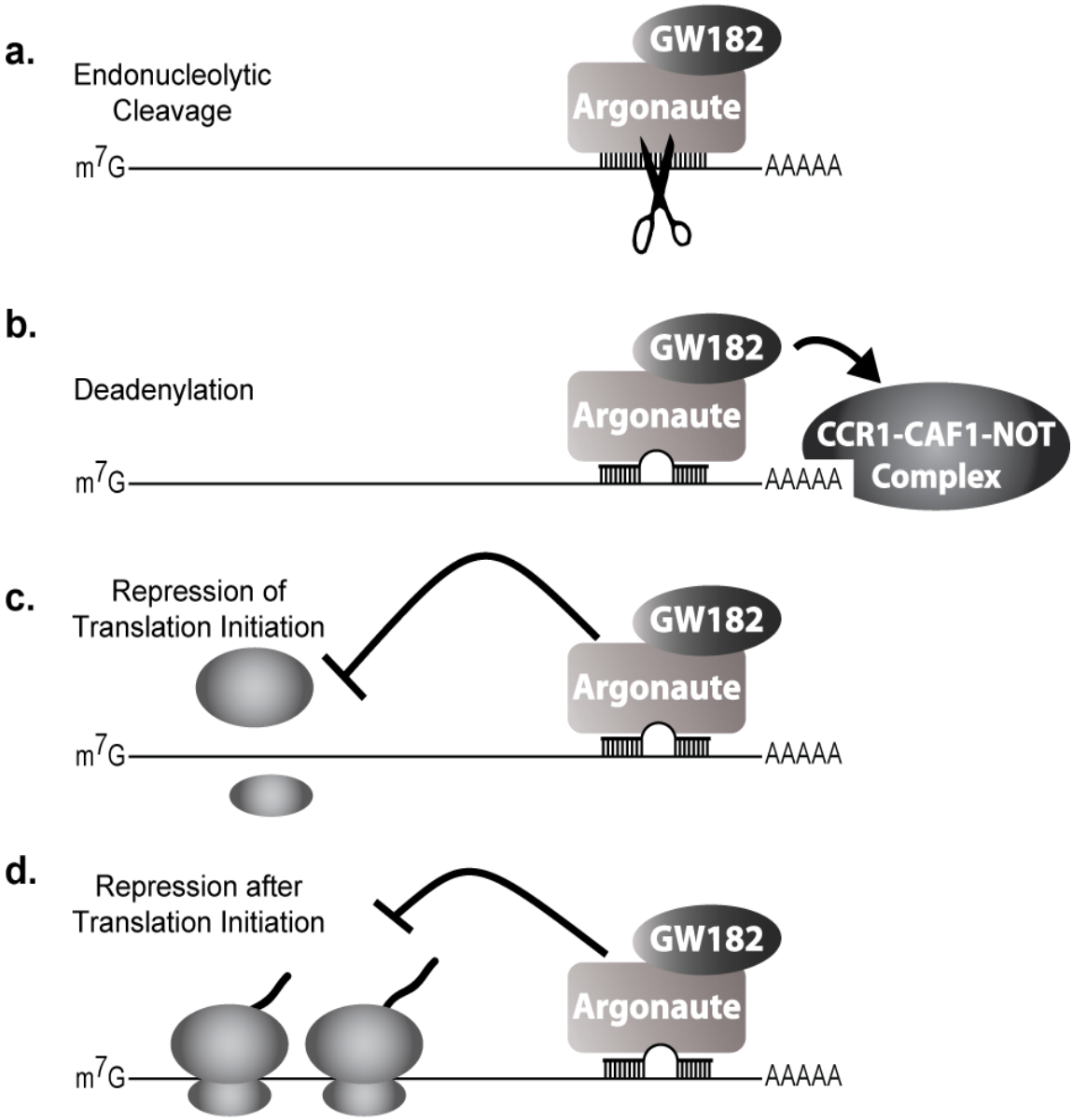


FIGURE 2.2 Mechanisms of miRNA target regulation. (a) When miRNAs bind to target sites with a high level of complementarity, it triggers endonucleolytic cleavage of the mRNA by Argonaute (Ago). (b) Ago interacts with GW182 proteins, which bind to the CCR1-CAF1-NOT complex to trigger deadenylation of bound target mRNAs. (c-d) Through unclear mechanisms, the Ago-GW182 complex represses translation (c) by inhibiting translation initiation and (d) by blocking protein accumulation after translation has initiated.

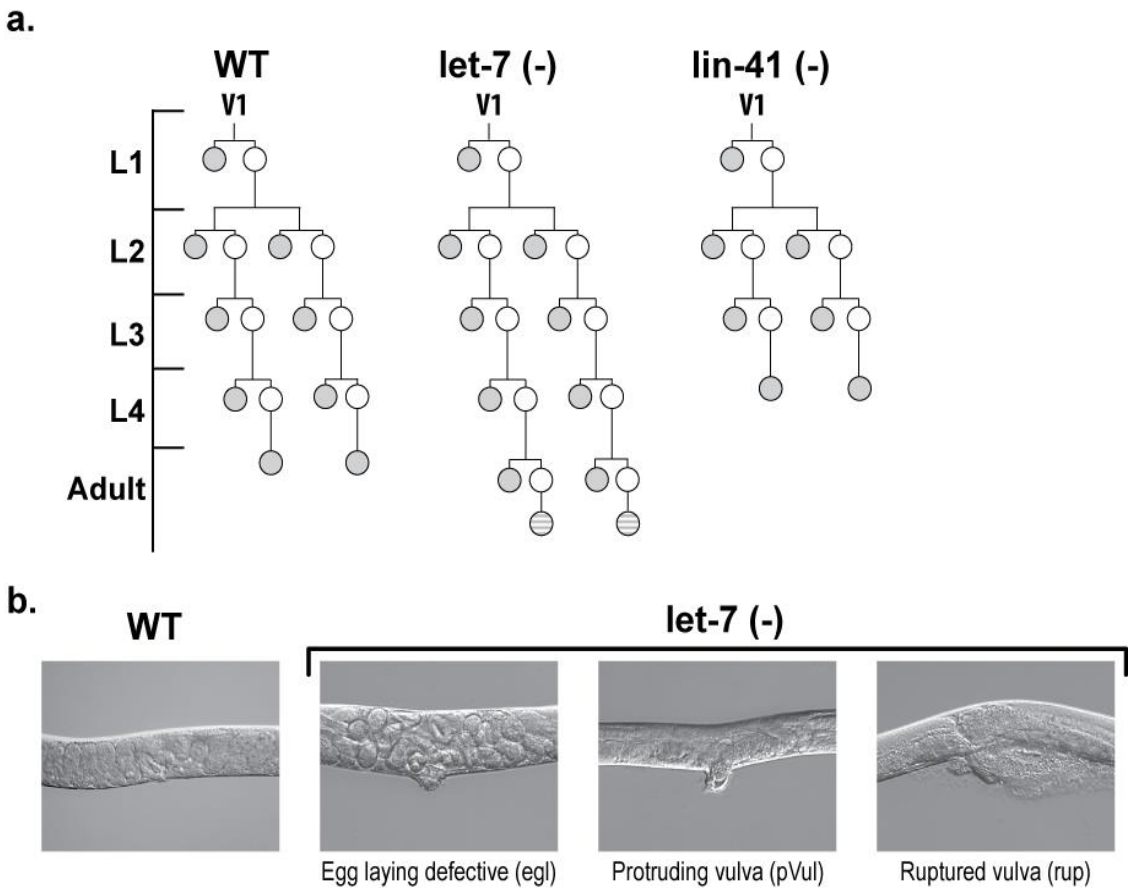


FIGURE 2.3 *C. elegans* let-7 mutant phenotypes. (a) The lineage division patterns of wild-type (WT) V1 hypodermal seam cells are shown for the larval (L1-L4) and adult stages. In WT worms, V1 cells divide at each molt, with one cell differentiating (grey) and the other continuing to divide at the next stage (white), until the cells terminally differentiate at adulthood (grey). Worms with loss of function mutations in let-7 (let-7(-)) have retarded heterochronic phenotypes where V1 cells can undergo one or more additional divisions after L4, sometimes failing to execute the adult differentiated fate (striped). The opposite phenotype, precocious, is expressed by worms with loss of *lin-41* (lin-41(-)), where the V1 cells terminate dividing and take on the differentiated adult fate at the L3 to L4 molt. (b) Pictures of adult worms at 200x magnification; the anterior is to the left and the vulva is on the ventral side. WT adult worms form eggs which they lay through their vulva, a narrow opening that is smoothly in line with the ventral side of the worm. Worms with mutations in let-7 display a variety of easily identifiable vulval phenotypes. Egg laying defective (egl) is characterized by a bloated, egg-filled worm. Protruding vulva (pVul) is defined by a protrusion due to abnormal development of the vulva. Ruptured vulva (rup) occurs during the L3 to L4 molt, when the worm innards burst through the weak unstructured vulva, resulting in lethality.

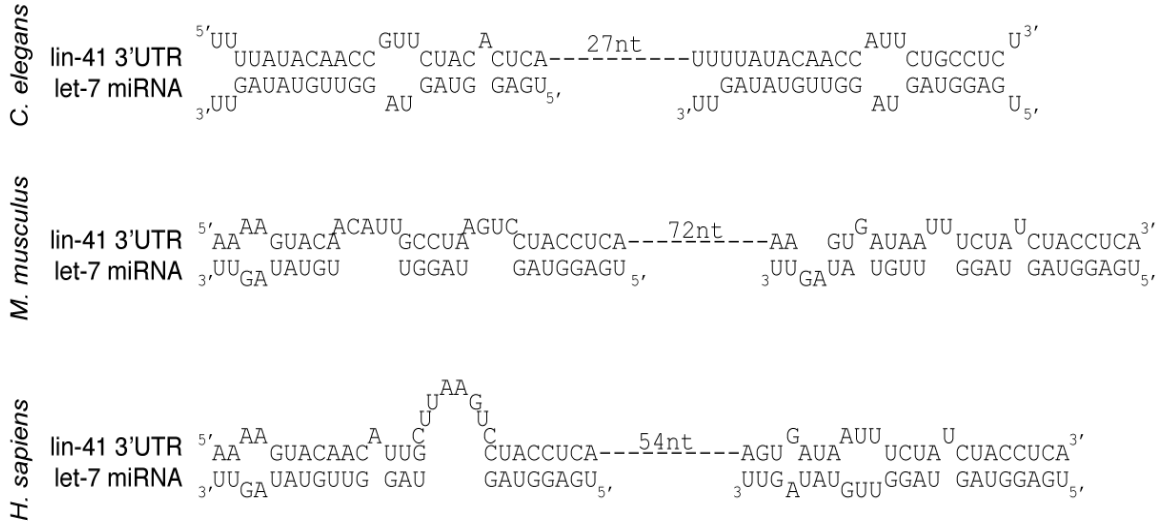


FIGURE 2.4 Examples of let-7 target site interactions. Illustrated here is the imperfect pairing between let-7 and 3' UTR sequences in lin-41 (also called Trim71 in mammals) genes in *C. elegans*, *M. musculus*, and *H. sapiens*.

C. elegans		***H. sapiens***	
let-7	UGAGGUAGUAGGUUGUAUAGUU	let-7a (3)	UGAGGUAGUAGGUUGUAUAGUU
mir-48	UGAGGUAGGCUCAGUAGAUGCGA	let-7b	UGAGGUAGUAGGUUGUGUGGUU
mir-84	UGAGGUAGUAUGUAAUAUUGUA	let-7c	UGAGGUAGUAGGUUGUAUGGUU
mir-241	UGAGGUAGGUGCGAGAAAUGA	let-7d	UGAGGUAGUAGGUUGCAUAGUU
mir-793	UGAGGUAUCUUAGUUAGACAGA	let-7e	UGAGGUAGGAGGUUGUAUAGUU
mir-794	UGAGGUAUUCAUCGUUGUCACU	let-7f (2)	UGAGGUAGUAGAUUGUAUAGUU
mir-795	UGAGGUAGAUUGAUCAGCGAGCUU	let-7g	UGAGGUAGUAGUUUGUACAGUU
		let-7i	UGAGGUAGUAGUUUGUGCUGUU
		miR-98	UGAGGUAGUAAGUUGUAUUGUU
		miR-202	AGAGGUAUAGGGCAUGGGAAAA

FIGURE 2.5 Homologs of let-7 in worms and humans. The let-7 family of genes is defined by the conserved seed sequence (nt 2-7). In C. elegans, let-7 has six non-identical "sister" miRNAs. The human genome contains 3 genes that produce mature miRNAs identical to the worm let-7 (let-a-1, let-7a-2, and let-a-3), 2 copies of let-7f, and 8 other miRNAs that share the let-7 seed sequence.

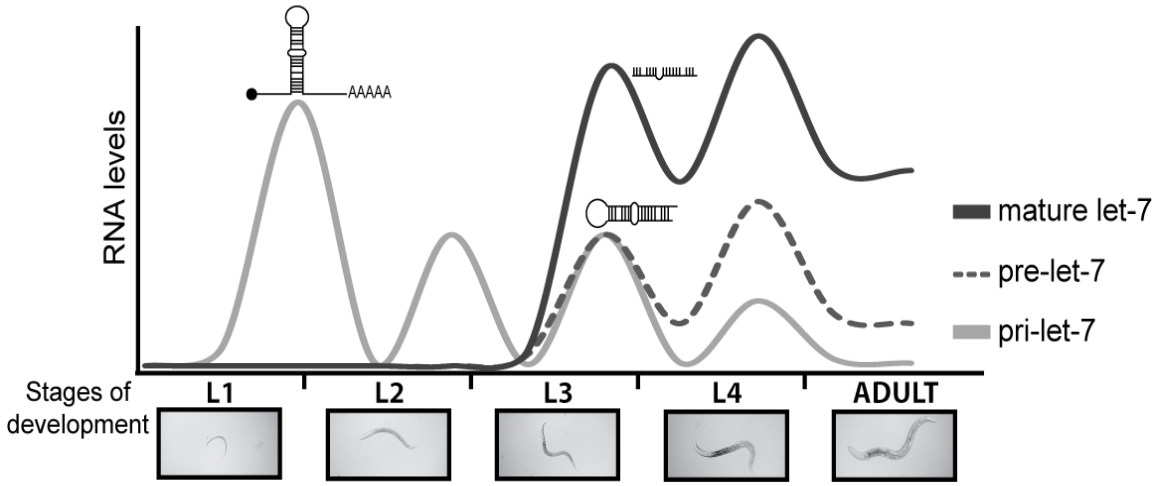


FIGURE 2.6 Expression of let-7 throughout *C. elegans* development. Graphical depiction of the accumulation of primary (pri-let-7), precursor (pre-let-7), and mature let-7 RNAs during the larval (L1-L4) and adult stages of worm development based on data from (Van Wynsberghe *et al.*, 2011). Worm pictures were captured at 100X magnification.

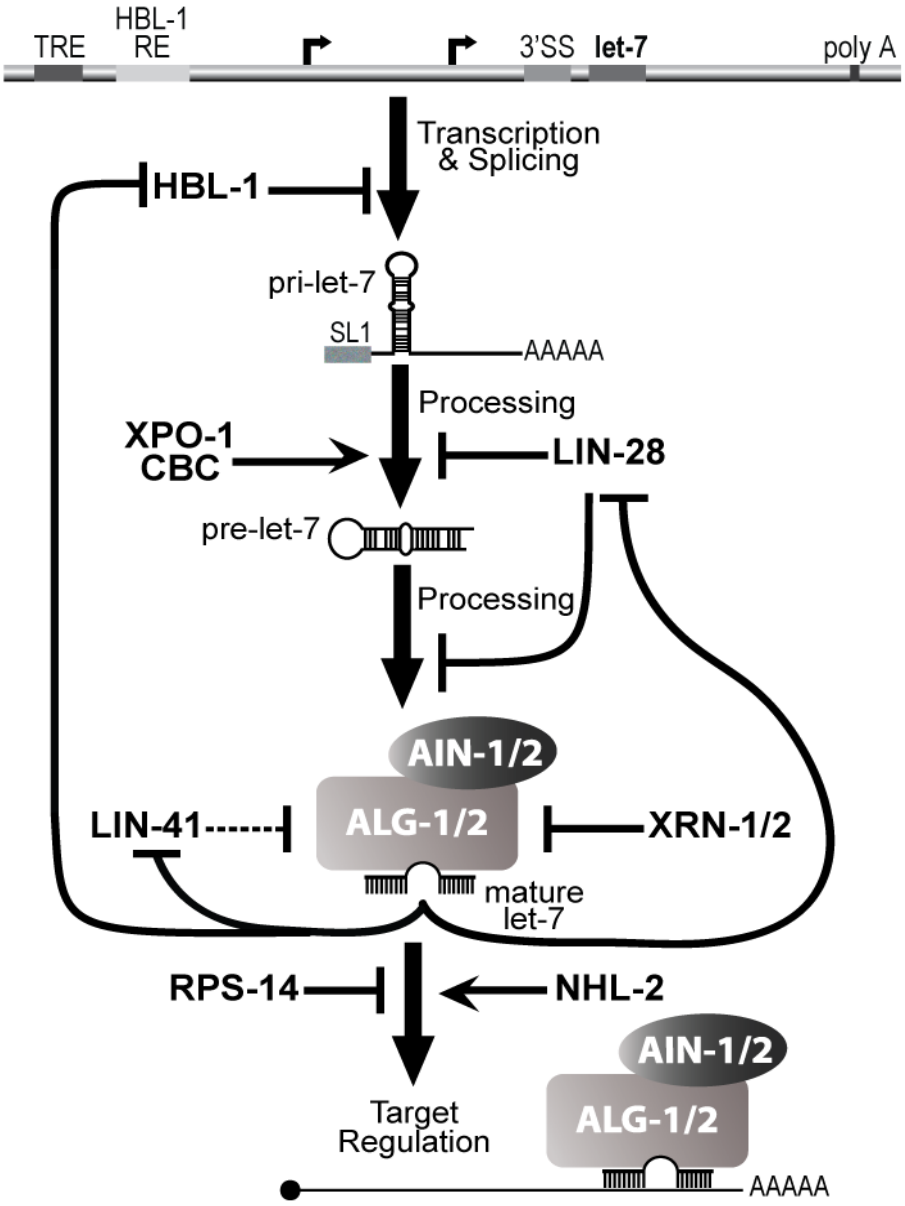


FIGURE 2.7 Regulation of let-7 biogenesis and function in *C. elegans*. Two cis-acting elements in the let-7 promoter, the temporal regulatory element (TRE) and putative HBL-1 response element, repress transcription in hypodermal seam cells. Two transcription start sites produce primary transcripts that are polyadenylated and subject to trans-splicing by the spliced leader 1 RNA (SL1) at the 3' splice site (3'ss) found upstream of the mature let-7 sequence. The nuclear transport factors XPO-1 and CBC promote conversion of primary to precursor and mature, while LIN-28 blocks processing of primary and precursor RNAs. The 3'→5' exonucleases XRN-1/2 degrade mature miRNAs upon release from the Ago complex. LIN-41, a target of let-7, indirectly regulates mature let-7 levels by targeting Argonaute for degradation in mouse cells (a dashed line represents this step since this LIN-41 activity has not yet been demonstrated in worms). As interactors with the Ago complex, NHL-2 enhances repression of let-7 targets, while RPS-14 antagonizes let-7 function.

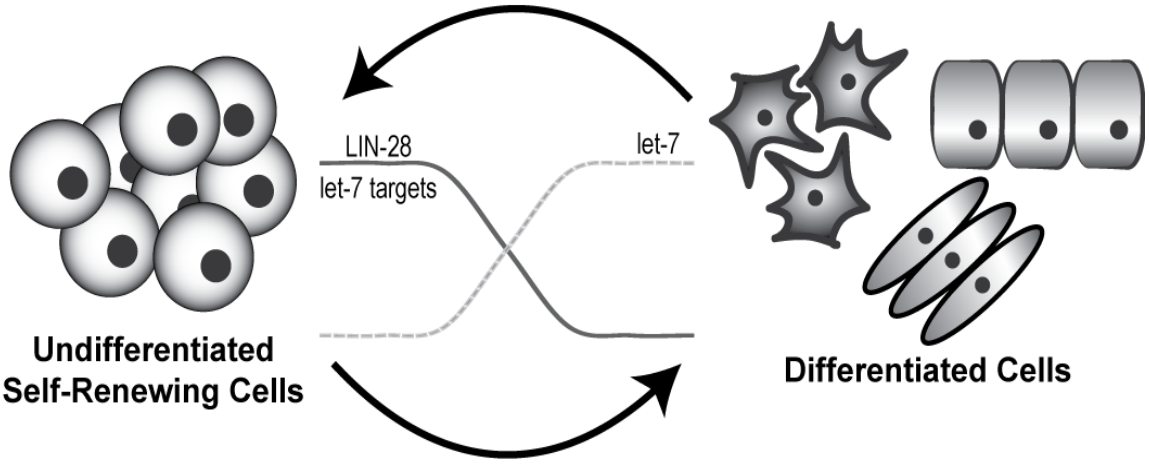


FIGURE 2.8 Regulation of cellular differentiation by let-7. Many types of self-renewing cells, such as embryonic stem cells and tumor cells, express high levels of LIN-28 and other targets of let-7 regulation that promote cell division. When cells are induced to differentiate, levels of these genes drop as let-7 expression increases, resulting in relatively high levels of let-7 in most differentiated cell types and repressed cell division programs.

Chapter III

Regulation of lin-4 microRNA expression by a conserved RNA binding protein in *Caenorhabditis elegans*

A. Abstract

MicroRNAs (miRNAs) are small non-coding regulatory RNAs that promote translational repression and/or deadenylation of their messenger RNA (mRNA) targets. The first miRNA discovered, lin-4, was named for its role in determining the lineage or temporal fate, of certain embryonic cells in the *Caenorhabditis elegans* (*C. elegans*) worms (Chalfie et al., 1981; Lee et al., 1993). In search of lin-4 regulators, a former graduate student in the lab, John Bracht, performed an RNAi screen of ~250 RNA-binding proteins (Bracht et al., 2010). As part of this project to investigate how the expression of lin-4 is regulated, I characterized the lin-4 primary transcripts and found that they are independently transcribed at the late larval 1 (L1) stage, despite the lin-4 sequence being present in the intron of a constitutively expressed overlapping host gene. We identified and characterized RBM-28, a ubiquitously expressed conserved RNA-binding protein, as a regulator of lin-4 expression. We also demonstrated that knockdown of RBM-28 affects expression of primary, precursor, and mature lin-4, as well as expression of the lin-4 targets, LIN-14 and LIN-28. Our work supports a model where RBM-28 genetically interacts with the lin-4 miRNA biogenesis pathway and promotes stabilization of mature miRNA expression.

B. Introduction

MiRNAs are found in a wide variety of species including humans, flies, and plants. Proper timing and expression levels of these small regulators is important for managing expression of their targets, which are often factors related to cellular differentiation and growth. As such, much work has been done to describe a general miRNA biogenesis pathway (Finnegan & Pasquinelli, 2013; Ha & Kim, 2014). Briefly, RNA Polymerase II synthesizes capped and polyadenylated primary miRNA (pri-miRNA) transcripts. These pri-miRNAs can be hundreds and even thousands of nucleotides (nt) long, and feature a hairpin-like stem-loop containing the mature miRNA sequence. The Microprocessor complex, made up of RNase III enzyme Drosha and RNA binding protein Pasha (DGCR8 in mammals), binds to the pri-miRNA, and cleaves out the ~65 nt precursor miRNA (pre-miRNA) stem-loop. After export to the cytoplasm, the pre-miRNA is further processed by the RNase III enzyme Dicer, leaving the mature miRNA sequence and its complementary star sequence (miRNA*). The miRNA* is degraded, and the mature miRNA is loaded onto an Argonaute protein which serves as the main functional component in the miRNA induced silencing complex (miRISC). As part of miRISC, the miRNA sequence serves as a guide to find imperfect complementary sites in the 3’ untranslated (UTR) of target mRNA, and promote either translational repression or deadenylation and degradation of that target. This stepwise process allows for transcriptional and post-transcriptional mechanisms of regulation.

Since its discovery in *C. elegans*, several insights have been made on biogenesis, regulation, and function of the lin-4 miRNA. After hatching from self-fertilized eggs, *C. elegans* undergo four larval stages (L1-L4), before becoming fully differentiated egg-laying adults. The temporal fate of each embryonic cell has been mapped along this heterochronic pathway (Sulston & Horvitz, 1977). The *lin-4* gene and its targets, *lin-14* and *lin28*, were among several gain-of-function and loss-of-function lineage mutants identified that caused misexpression of cell fate timing (Chalfie et al., 1981; Ambros & Horvitz, 1984; Ambros, 1989). Loss of *lin-4* or over expression of *lin-14* causes vulval and hypodermal seam cells to fail to differentiate, due to repeated L1 cell fates during development. The opposite is true of *lin-14* and *lin-28* loss of function mutants; instead, differentiation occurs too early. Researchers found that *lin-4* acts antagonistically to downregulate *lin-14* and *lin-28* after the L1 larval stage through interactions with target 3'UTR (Arasu et al., 1991; Moss et al., 1997).

Transcriptional GFP reporters of lin-4 detect measurable fluorescence at L1, when mature lin-4 begins to accumulate (Esquela-Kerscher et al., 2005; Baugh & Sternberg, 2006; Martinez et al., 2008; Ow et al., 2008; Kim et al., 2013). In addition, a conserved FLYWCH transcription factor binding site was found upstream of *lin-4* and negatively regulates lin-4 transcription in early larval development (Ow et al., 2008). A more recent study released after publication of the work presented in this chapter, found the lin-4 mature miRNA oscillates between larval stages, during the larval molts, and that this oscillation provides *lin-14* with a smooth temporal gradient of expression, which otherwise changes

to an oscillating pattern in the absence of lin-4(Kim et al., 2013).

This collaborative work was published in 2010. As an author, here I present the theory, troubleshooting, and results behind experiments I carried out towards this research (Bracht et al., 2010). In this study, I identified two independently transcribed primary transcripts of lin-4 (pri-lin-4), which begin to accumulate at late L1 stage. I also show evidence that the overlapping host gene does not contribute to the production of lin-4. This was an interesting finding because most intronic miRNAs were believed to be regulated by their host gene, and there were few examples showing otherwise. Furthermore, we found that a ubiquitously expressed conserved RNA binding protein, RBM-28, has a role in stabilizing mature lin-4, since knockdown of this protein greatly decreased lin-4 expression with no effect on either pre-lin-4 accumulation, or pri-lin-4 transcription and accumulation.

C. Results and Discussion

C.1 Two independently synthesized pri-lin-4 RNA transcripts solely contribute to lin-4 biogenesis

Pri-lin-4 had yet to be characterized, likely because lin-4 resides in the sense direction within intron 9 of the F59G1.4 gene (referred to from here on as its host gene) (Lee et al., 1993). Most studies suggested that many intronic miRNAs are transcribed as part of their host gene (Ying & Lin, 2005; Saini et al., 2007; Lutter et al., 2010). This particular host gene is constitutively expressed, confounding efforts to identify a potential pri-lin-4 individual transcript. However,

there was evidence that an individual lin-4 transcript derived from an intronic promoter could exist; a ~700 nt DNA fragment is sufficient to rescue the *lin-4(e912)* null mutation that contains ~5 kilobase (kb) deletion in intron 9 of the host gene (Lee et al., 1993). Additionally, while expression of the host gene mRNA begins early in development, mature lin-4 does not appear until the late L1. Finally, studies that used reporters which fused sequences in intron 9 upstream of lin-4 to GFP demonstrate transcriptional activation at late L1; one study even found a *cis* promoter element which negatively regulated GFP activation, suggesting an independent lin-4 promoter (Esquela-Kerscher et al., 2005; Baugh & Sternberg, 2006; Martinez et al., 2008; Ow et al., 2008).

To analyze temporal expression of pri-lin-4 during early development, I extracted total RNA from three biological replicates of synchronized N2 wild-type worms grown at +20°C and collected in two-hour (hr) intervals until early L2 (18 hr). Complementary DNA (cDNA) was prepared by reverse transcription (RT) using random oligonucleotides. I quantified the expression of the host gene and independent pri-lin-4 transcripts using quantitative RT-PCR (qRT-PCR) with oligos flanking either the lin-4 hairpin, or a similarly sized region upstream in host gene intron 9. To distinguish between expression of host gene and pri-lin-4, I calculated the ratio of pri-lin-4 over host gene, such that a ratio of one would mean pri-lin-4 detection is actually due to host gene intron 9, but a higher ratio would indicate separate independent pri-lin-4 expression. This analysis revealed that pri-lin-4 begins to accumulate during late L1, peaking at nearly 6-fold higher than host gene expression at ~12 hours, before processing likely leads to a drop

in detection at early L2 (FIG 3.1A). In addition, the 12-hour peak in pri-lin-4 transcription coincides with the accumulation of mature lin-4 (FIG 3.1B).

To confirm the expression of independent lin-4 transcripts at late L1, I performed an agarose Northern Blot comparing three biological replicates of synchronized wild-type N2 worms grown to 4 hr (early L1), 12 hr (late L1), and 16 hr (L2) time points (FIG 3.2A). As a negative control, I included a 12 hr timepoint of the *lin-4(e912)* null allele. Using a lin-4 probe flanking the hairpin, I was able to detect two distinct bands in the 12-hr timepoint that were not present in the null, or any other time points. I also probed for 18S ribosomal RNA (rRNA) as a loading control. This result correlated well with the previous qRT-PCR results.

To test whether these bands were indeed pri-lin-4 and subject to Microprocessor processing, I set up another agarose Northern Blot using three biological replicates of 12-hr staged RNAi sensitive worms *rrf-3(pk1426)* treated with empty vector (vector), Drosha (*drsh-1*), or Pasha (*pash-1*) RNAi (FIG 3.2B). Because only pri-miRNAs are subject to processing, knockdown of these crucial Microprocessor proteins should lead to an accumulation of pri-miRNA transcripts. I also included RNA size markers, Century Plus and Millennium (Ambion), which range from 100 nt – 9 kb, in order to measure the size of the pri-lin-4 transcripts. During the first attempt, RNAs smaller than ~200-300 nt migrated off the gel. For the subsequent run, I applied 80V for 135 minutes and another 45 minutes at 100V, and stopped the electrophoresis when I began to see distinct separation of the RNA size markers. This explains why the two bands previously seen on the developmental Northern Blot probed for pri-lin-4 are not as well-defined on this

Northern (FIG 3.2B). Nevertheless, the pri-lin-4 transcripts did accumulate in both Drosha and Pasha knockdown experiments indicating that the bands are indeed substrates of the Microprocessor. In addition, the RNA size markers helped determine that pri-lin-4 transcripts were ~300 – 400 nt in size.

From the above results, it was not clear whether the host gene contributed to the production of lin-4. Given that pri-miRNAs are processed by the Microprocessor, RNAi knockdown of *drsh-1* and *pash-1* should lead to an accumulation of only pri-miRNA transcripts and not of transcripts that do not contribute to the miRNA pathway. Thus, I used this approach to examine the effect of Microprocessor knowndown on the host gene by qRT-PCR. I prepared cDNA using three biological replicates of 12-hr staged RNAi sensitive worms *rrf-3(pk1426)* treated with empty vector (vector), Drosha (*drsh-1*), or Pasha (*pash-1*) RNAi. I used oligo-dT to prepare the cDNA by RT in order to capture full mRNA sequences and reduce detection of degradation products. I compared expression of pri-lin-4 and host gene intron 9 measured by qRT-PCR and normalized to actin, an abundant housekeeping gene (FIG 3.2C). While pri-lin-4 transcripts accumulated upon Microprocessor knockdown, expression of the host gene remained unchanged at levels lower than vector-treated pri-lin-4. This confirms that pri-lin-4 is being processed by the Microprocessor and any contribution of the host gene to lin-4 expression is negligible.

After detecting ~300 – 400 nt pri-lin-4 transcripts at late L1, we used two biological replicates of synchronized wild-type N2 12-hr worms in order to define transcriptional start and end sites by Rapid Amplification of cDNA ends (RACE).

This technique captures capped mRNAs by ligating a unique sequence to the 5' end or 3' polyA, followed by RT and nested PCR. From these experiments we detected two transcriptional start sites, A and B, that corresponded to the pri-lin-4 transcript sizes detected by Northern Blot analysis (FIG 3.2D). Thus, we are confident about the transcriptional start sites we identified. In addition, I performed qRT-PCR using oligos that start before and after the A and B sites at 4, 12, and 16 hr timepoints, and compared that to host gene intron 9 expression to ensure we can see the pulse of pri-lin-4 expression at late L1, as previously detected (data not shown).

Establishing the 3' end of pri-lin-4 was more challenging and required multiple approaches. Given the mapped start sites, and pri-lin-4 RNAs detected by Northern blotting, we predicted the polyadenylation (polyA) signal was 64 nt downstream of pre-lin-4. PCR conditions including oligos and extension times were adjusted to optimize detection and all PCR experiments included a no-template control. To ensure that the designed oligos would amplify our region of interest, they were aligned to the *C. elegans* genome using the BLAST tool available on WormBase (Harris et al., 2014). I also examined the possibility that pri-lin-4 is *trans*-spliced. In *C. elegans*, 70% of all mRNA is *trans*-spliced with a splice-leader (SL) sequence. The function of *trans*-splicing is not well understood, but one role seems to be distinguishing between two genes in an operon. I carried out RT-PCR with amplified 12-hr staged wild-type cDNA but did not detect evidence of SL-splicing.

Nested 3’ RACE reactions using three biological replicates of 12-hr staged wild-type worms detected the Drosha cleavage site after pre-lin-4 and downstream A-rich regions in intron 10 and exon 11. To examine the validity of this putative 3’ end, I tested the developmental Northern Blot setup in Figure 3.2A with a probe for sequences that begin in exon 10, (~60 nt after the putative polyA signal) and extend near the 3’ end of exon 11. This probe detected a band in the 12-hr staged *lin-4(e912)* null and positive control *in vitro* transcript, but not in any of the 4, 12, or 16 hr staged wild-type worms (data not shown). Thus, our previously detected pri-lin-4 transcripts are likely upstream of exon 10. In addition, I performed 3’ end-walking, using qRT-PCR to measure pri-lin-4 expression in 12-hr staged worms using one common forward primer and sequential reverse primers (FIG 3.2E). Reverse oligos around the polyA site demonstrated robust expression of pri-lin-4; however, this expression dropped significantly using a reverse oligo that resides in exon 10. Taken together, these experiments suggest that pri-lin-4 3’ end probably occurs after the poly A signal just before exon 10 (FIG 3.2D).

C.2 RBM-28 is an essential conserved RNA recognition motif protein ubiquitously expressed in *C. elegans*

John Bracht performed an RNAi screen of ~250 RNA-binding proteins to look for regulators of lin-4 and found that knockdown of a conserved protein, RBM-28, decreased expression of mature lin-4 by nearly four-fold (Bracht et al., 2010). RBM-28 encodes a protein containing multiple RNA recognition motifs (RRM) as well as acidic regions, with homology to human RBM28 and yeast

Nop4/Nop77 protein (FIG3.3A) (Sun & Woolford, 1994; Sun & Woolford, 1997; Damianov et al., 2006).

I attempted to measure temporal expression of RBM-28 in wild-type worms using several strategies. I tested a human RBM-28 polyclonal antibody under a number of different conditions with the hope that sequence similarity with the worm sequence would be sufficient for detection. Unfortunately, despite a number of Western Blots including wild-type and RNAi treated worms, I was unable to detect RBM-28. I also attempted to create a GFP-tagged transgene to insert into worms; however, RBM-28, including promoter and 3'UTR, is ~10 kb long, making construction of such a plasmid difficult. Finally, Priscilla Van Wynsberghe a former post-doc who co-first authored the paper, obtained a transgenic worm expressing a transcriptional GFP reporter of RBM-28. This animal expresses GFP in all tissues at all time points (FIG 3.3B).

C.3 RBM-28 specifically stabilizes mature lin-4 expression

To better understand how RBM-28 affects lin-4 biogenesis, levels of primary, precursor, and mature lin-4 were measured by PAGE Northern blotting and RT-PCR (FIG 3.4). Three biological replicates of L2-staged *rrf3(pk1426)* worms were grown under vector or *rbm-28* RNAi conditions, and total RNA was extracted. Because knockdown of RBM-28 slowed developmental growth, worms were staged according to size and gonad development. Knockdown in RNAi conditions was confirmed by comparing RBM-28 expression by qRT-PCR (data not shown). Northern Blot analysis indicated that while pre-lin-4 expression is not significantly changed under decreased RBM-28 levels, expression of mature lin-4

is greatly reduced (FIG 3.4A). Furthermore, RT-PCR of pri-lin-4 and host gene revealed no change in expression, while lin-4 targets, *lin-14* and *lin-28*,were upregulated (FIG 3.4B).

Corroborating this data, I found that RBM-28 RNAi treated pri-lin-4 reporter worms (plin-4::GFP) did not demonstrate alternative transcriptional spatial or temporal patterns (FIG 3.5B). John Bracht constructed the plin-4::GFP strain, a transgenic worm expressing the promoter region, ~500 nt upstream of lin-4, fused to GFP. I detected no differences in expression when compared to GFP transcriptional reporters published by the Ambros lab (Ow et al., 2008) (data not shown). I also examined RBM-28 effects on LIN-28 by treating GFP-tagged lin-28 transgenic worms (PQ272) with either empty vector or RBM-28 RNAi and observing whether there were differences in spatial and/or temporal fluorescence (FIG 3.4D). I analyzed fluorescence in 34 worms from duplicate biological samples of PQ272 worms treated with RNAi collected at L3 by gonad staging. As a target of lin-4, we expected expression of LIN-28 to be low at L3 when mature lin-4 is expressed. Instead, we observed robust LIN-28 expression at L3 in worms treated with RBM-28 RNAi, again indicating that lin-4 is greatly reduced without RBM-28.

Taken together, my data supported a model in which two pri-lin-4 RNA are independently transcribed and processed during late L1, followed by stabilization of their mature form by the RNA-binding protein RBM-28. As levels of lin-4 decrease in *rbm-28* knockdown worms, *lin-14* and *lin-28* become misexpressed, leading to developmental challenges. Interestingly, RBM-28 knockdown worms

have a delayed-growth phenotype not seen in *lin-4(e912)* null mutants (Bracht et al., 2010). This delayed development is suppressed in *lin-14* and *lin-28* mutant worms, implicating *lin-28* and *lin-14* as important organismal growth factors. Surprisingly, measurements of other early expressed miRNAs, miR-2, and miR-47, showed an increase upon *rbm-28* knockdown, due to changes in the pri-miRNA abundance. Whether these changes somehow contribute to the overall stability of lin-4 or these miRNAs affect lin-4 biogenesis is unclear. This research also raises questions about how independently transcribed intronic miRNAs can avoid processing when transcribed as part of the host gene, and how the cellular machinery, specifically the Microprocessor, can distinguish between pri-miRNAs that are independently transcribed versus pri-miRNAs under host gene transcriptional control.

D. Materials and Methods

Several methods including worm staging, RNA extraction, RNAi treatment, Western blot analysis, Agarose and PAGE Northern Blot analysis, and reverse transcriptase PCR assays have been previously described (Van Wynsberghe et al., 2011; Zisoulis et al., 2012). The Invitrogen GeneRACER kit was used for 5' and 3' RACE as previously described (Bracht et al., 2004). Wild-type worms were N2 Bristol. Oligos used for this study are listed in Table 3.1

D.1 Worm strains and culture conditions

The following *C. elegans* strains were used: *lin-4(e912)* (DR721), *lin-14(n179)* (GR1106), *lin-14(n179);lin-4(e912)* (IH008), *lin-28(n719)* (GR1115), *rrf-*

3(pk1426) (NL2099]. pRBM28::GFP (BC14466) expresses GFP from the rbm-28 promoter. Worms were cultured at 20°C, unless otherwise indicated. Gravid adult worms are vigorously shaken in hypochlorite solution (20% bleach, 0.5 M KOH) for six minutes, washed in M9 solution [22 mM KH2PO4, 42 mM Na2HPO4, 85.5 mM NaCl, 1mM MgSO4], and incubated in M9 solution overnight at 20°C to synchronize. Development was initiated by plating starved L1 hatchlings on plates seeded with bacteria.

D.2 Expression constructs and transgenic strains

For the lin-4 promoter fusion construct, the plin4::GFP (pJRB3) plasmid was created by PCR amplification of ~400bp of sequence 100nt upstream of mature lin-4. This insert and plasmid pPD95.75 (A. Fire laboratory) were digested with Age-I and ligated by standard subcloning techniques. pJRB3 was linearized with Spe-I and injected at 50 ng/µl per plasmid, along with the pha-1 rescuing plasmid pBX (Granato et al., 1994) at 50 ng/µl into *pha-1(e2123)* mutant animals, and transgenic animals were identified by rescue of pharynx development in F1 animals at 25°C. Construction of PQ272 is previously described (Van Wynsberghe et al., 2011).

E. Acknowledgments

This chapter in part, contains material from “Regulation of lin-4 miRNA expression, organismal growth, and development by a conserved RNA binding protein in C. elegans”. Bracht, John R., Van Wynsberghe, Priscilla M., Mondol,

Vanessa & Pasquinelli, Amy E., Developmental Biology, 2010. John and Priscilla were the primary authors; I was the second author.

Table 3.1 Oligos used for this study

Probe Primers for PAGE and Agarose Northern Blotting		
PRIMER	**LAB DESIGNATION**	**SEQUENCE 5' → 3'**
lin-4 starfire	A1916	TCACACTTGAGGTCTCAGGGA
5.8s rRNA	A479	CTAGCTTCAGCGATGGATCGGTTGC
	A480	GAACCAGACGTACCAACTGGAGGCCC
pri-lin-4	A64	GTCGACGAGACGCCGAGTCTCCC
	A65	CGTTTGACCCTTTTCCCCGAATACC
18s rRNA	A839	GCGTACGGCTCATTAGAGCAGATATCAC
	A840	GGTCAGAACTAGGGCGGTATCTAATCG
downstream of pri-lin-4	A2466	CACTAACGCCAACTGATGGAGA
	A2467	TGTAGATGGCTCTGAGAGC
Primers for RT-PCR		
PRIMER	**LAB DESIGNATION**	**SEQUENCE 5' → 3'**
pri-lin-4	A65	CGTTTGACCCTTTTCCCCGAATACC
	A2015	TTTATGCTTCCGGCCTGTTC
F59G1.4 intron 9	A463	GAGCGAAGCTTAAGGGGATGATGTGTC
	A464	GTATACCGGTCATGAAAAGTGACACTTGC
actin	A810	GTGTTCCCATCCATTGTCGGAAGAC
	A811	GCACTTGCGGTGAACGATGGATGGG
18s rRNA	A839	GCGTACGGCTCATTAGAGCAGATATCAC
	A840	GGTCAGAACTAGGGCGGTATCTAATCG
lin-14	A425	GTTACGAGGTAAGCGAAGCAAGATTCAG
	A426	GGGGCCAAATATCTATATCCAAAGTAG
lin-28	A427	GCGTTCGCCCGCAATAGCGGAACTTACG
	A424	GAGCTTTCTGATAGTTTTTCAG
SL1	A90	GGTTTAATTACCCAAGTTTGAG
SL2	A91	GGTTTTAACCCAGTTACTCAAG
exon 11 rev	A2467	TGTAGATGGCTCTGAGAGC
Primers for qRT-PCR		
PRIMER	**LAB DESIGNATION**	**SEQUENCE 5' → 3'**
pri-lin-4	A2275	CTAGACAATTTCTAGAGTTTTGGTTGGT
	A2276	GGAACTAGCTCCCAGTGTGAAAA
F59G1.4 intron 9	A2277	ACTCCGTCGTAGTAACCCATAAC
	A2278	TTGGCTCTCTGTAATCCAACAATTCAA
actin	A1755	GTTCACCGCAAGTGCTTCTAAATG
	A1756	GCAAATGAGTGAAAGGACAATAAGG
rbm-28	A2033	GATTCAGGAGTTATGGGTATAATTCTTC
	A2330	GTGCTCAAATTGCTCTGTTTCAG
pri-lin-4 3' end walking	A65	CGTTTGACCCTTTTCCCCGAATACC
	A1146	TACCAGGACGGTTTGAGCAGATC
	A1873	TCTCCATCAGTTGGCGTTAGTG
	A2276	GGAACTAGCTCCCAGTGTGAAAA
	A2489	TTCCCCGAATACCATTTTATTG
Primers for RACE analysis		
PRIMER	**LAB DESIGNATION**	**SEQUENCE 5' → 3'**
RT	*Oligo dT	TGTCAACGATACGCTACGTAACGGCATGACAGTG$(T)_{24}$
RACE PCR	*Generacer 5'	CGACTGGAGCACGAGGACACTGA
	A274	ATTTTATTGGAACTAGCTCCCAGTGTG
RACE nested PCR	*Generacer nest	GGACACTGGACATGGACTGAAGGAGTA
	A276	GATCTGCTCAAACCGTCCTGGTACCCGG
3' RACE PCR	A2423	CGAAGCGACCGAATGACCCA
	*Generacer	GCTGTCAACGATACGCTACGTAACG
3' RACE nested PCR	A1146	TACCAGGACGGTTTGAGCAGATC
	*Generacer nest	CGCTACGTAACGGCATGACAGTG

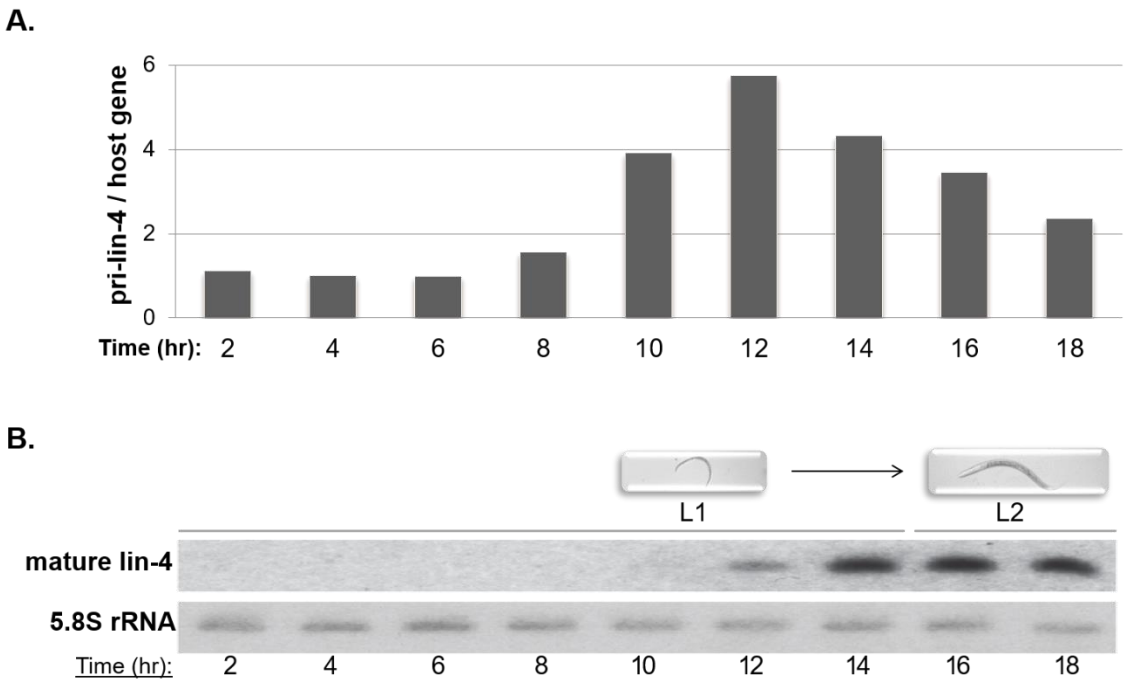


Figure 3.1 Expression of lin-4 throughout early development. Total RNA was extracted from synchronized wild-type N2 worms at the indicated time points in L1 and early L2, and used for the following analyses. Representative of three independent experiments shown. (**A**) The qRT-PCR measured ratio of expression of pri-lin-4 over host gene, normalized to actin, relative to early L1. (**B**) PAGE Northern Blot analysis of mature lin-4 and 5.8S ribosomal RNA (rRNA) loading control. Worm images taken at 200X magnification at the indicated developmental stage, anterior end left.

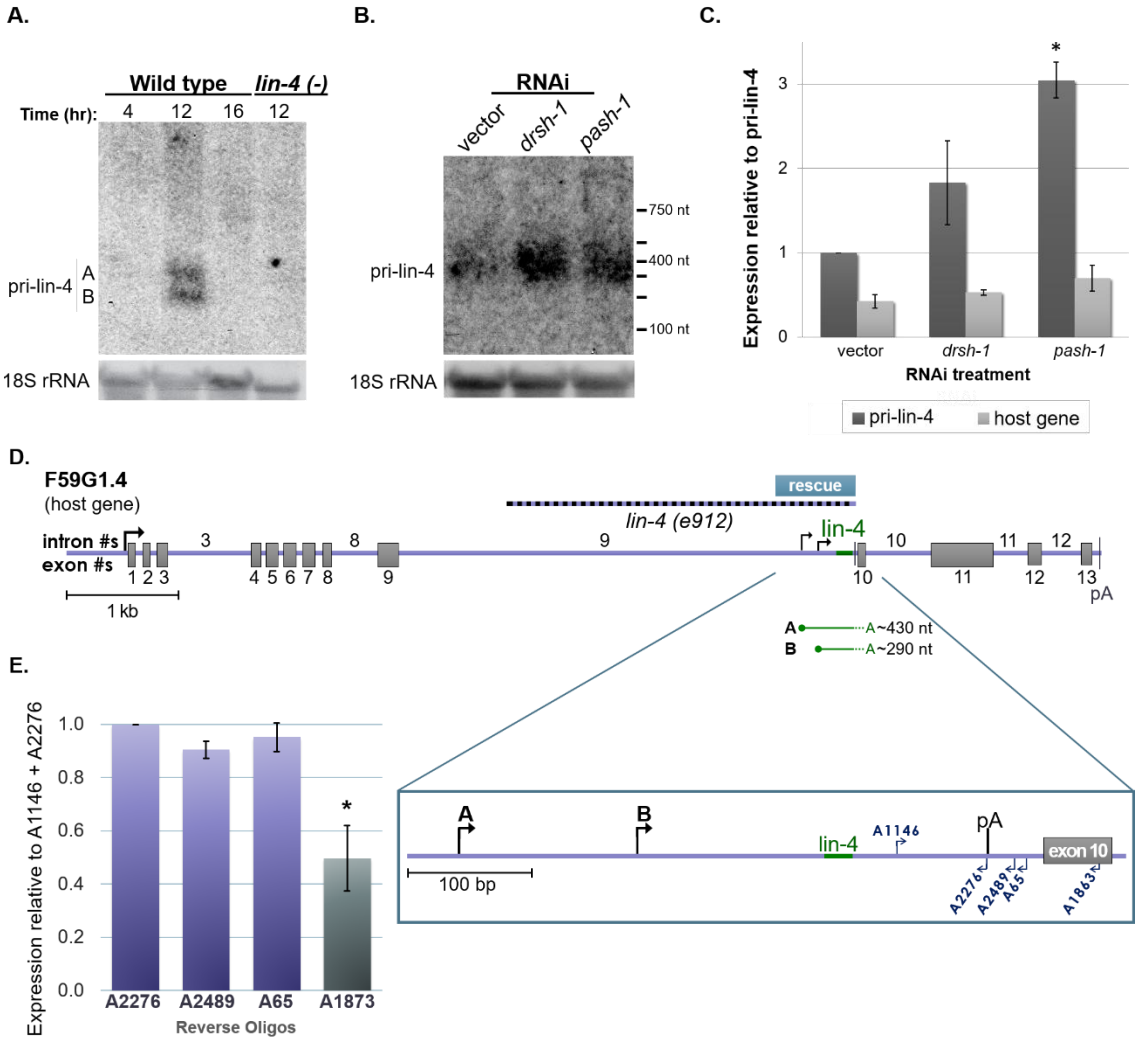


Figure 3.2 Validating pri-lin-4 transcripts. All experiments based on three biological replicates. Error Bars represent standard error. *p < 0.05 (**A**) Total RNA was extracted from synchronized wild-type or *lin-4(e912)* null worms at the indicated time points. Representative Northern Blot analysis of pri-lin-4 and 18S rRNA loading control. (**B-C**) Total RNA was extracted from 12 hr staged *rrf-3(pk1426)* RNAi-sensitive worms, treated with empty vector (vector), Drosha (*drsh-1*), or Pasha (*pash-1*) RNAi. (**B**) Representative Northern Blot analysis of pri-lin-4 and 18S rRNA. Sizes based on an RNA marker are indicated. (**C**) Oligo dt was used for cDNA preparation. qRT-PCR measurements of pri-lin-4 or host gene normalized to actin, relative to vector-treated pri-lin-4 expression. (**D-E**) Total RNA was extracted from 12-hr staged wild-type worms. (**D**) Model depicting F59G1.4 host gene with intronic lin-4 (green), introns and exons (numbered), transcriptional start sites (black arrow), and polyadenylation signal (pA). Pri-lin-4 A and B transcripts were mapped by 5' RACE, and are illustrated below (green) with approximate sizes. Above the gene model, *lin-4(e912)* deletion (dashed line) and ~700 nt rescue fragment is depicted. Close up of pri-lin-4 region maps oligos used for 3' end determination. (**E**) qRT-PCR average expression of a common forward oligo (A1146) and subsequent reverse oligos (A2276, A2489, A65, A1873), normalized to actin, relative to A1146 + A2275.

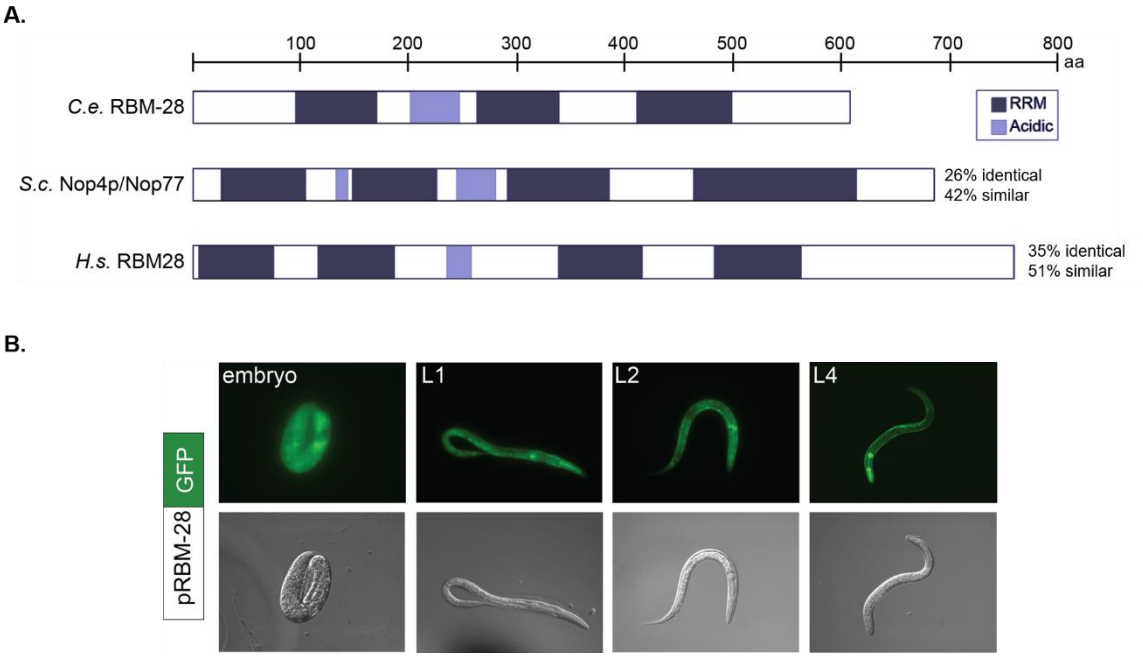


Figure 3.3: Characteriznig RBM-28 in *C. elegans.* (**A**) The RBM-28 protein is homologous the human RBM28 protein and the yeast Nop4 protein. All proteins contain multiple RNA recognition motifs (RRM) (dark purple) and one or more acidic regions (light purple). (**B**) Worms expressing GFP driven by the R05H10.2 promoter were analyzed by fluoescent microscopy throughout development. Representative images are shown

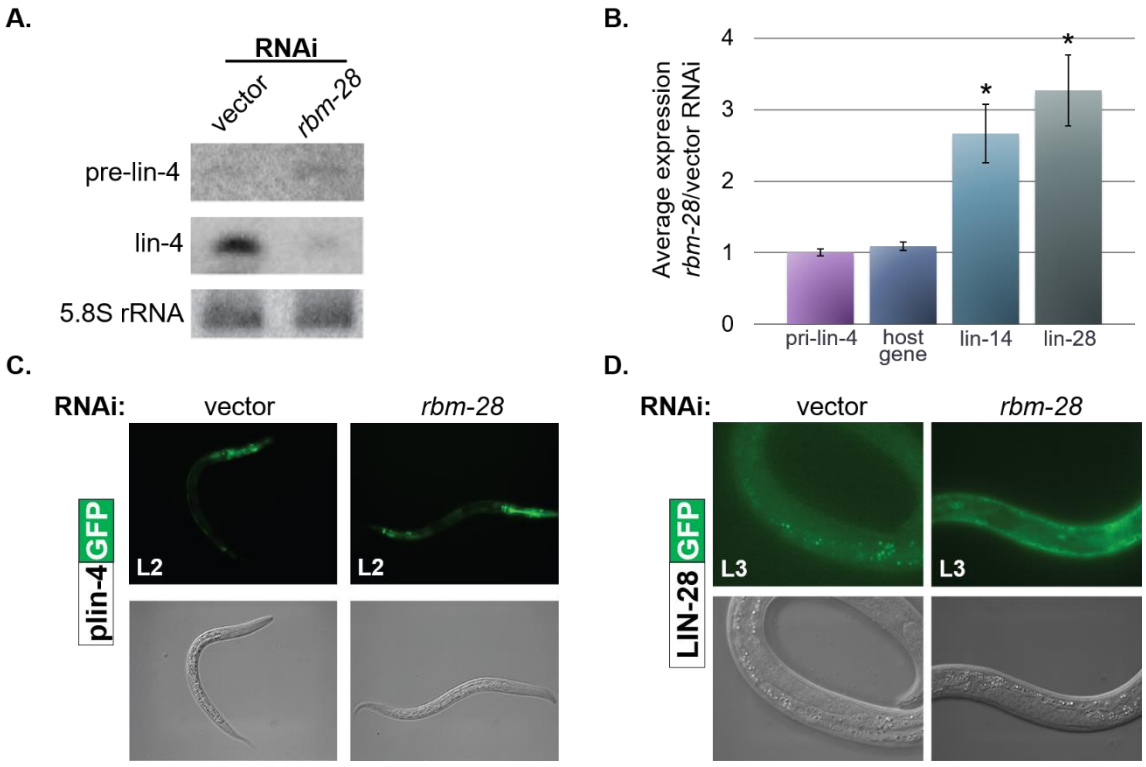


Figure 3.4: RBM-28 regulates mature lin-4. (**A-B)** Three biological replicates of synchronized L2-staged, *rrf3(pk1426)* worms were treated with either empty vector or *rbm-28* RNAi. (**A**) Representative PAGE Northern Blot of pre-lin-4, mature lin-4, and 5.8S rRNA. (**B**) Average RT-PCR measured ratio of expression of pri-lin-4, host gene, lin-14 and lin-28 in actin normalized, *rbm-28*/vector RNAi treated worms. Bars indicate standard error. *p<0.05 (**C**) Representative images of L2-staged transgenic worms expressing pri-lin-4 transcriptional GFP reporter under empty vector or *rbm-28* RNAi conditons, examined by fluorscent microscopy. (**D**) Representative images of L3-staged transgenic worms expressing GFP-tagged LIN-28 under empty vector or *rbm-28* RNAi conditions, examined by fluorscent microscopy.

Chapter IV

Splicing remodels the let-7 primary microRNA to facilitate Drosha processing in *Caenorhabditis elegans*

A. Abstract

MicroRNAs (miRNAs) are a class of small non-coding RNAs that use partial base pairing to recognize and regulate the expression of messenger RNAs (mRNAs). Mature miRNAs arise from longer primary transcripts (pri-miRNAs) that are processed to a shorter hairpin precursor miRNA (pre-miRNA) by the Microprocessor complex. In *Caenorhabditis elegans,* the primary let-7 (pri-let-7) transcript undergoes *trans*-splicing, where pri-let-7 is cleaved at a 3' splice site, and the splice-leader-1 (SL1) sequence is appended at the 5' end. Here we investigate the role of this splicing event in the biogenesis of let-7 miRNA. We hypothesized that splicing changes the secondary structure of the pri-let-7 transcript, creating a more favorable substrate for recognition by the Microprocessor. Supporting this idea, we detected conspicuous structural differences between unspliced and SL1-spliced pri-let-7 transcripts using *in vitro* ribonuclease (RNase) assays. Through the generation of transgenic worm strains, we found that the RNA secondary structure produced by splicing, as opposed to the act of splicing itself, optimizes processing of pri-let-7 by the Microprocessor *in vivo*. We also observed that the endogenous spliced, but not the unspliced, pri-let-7 transcripts bind to the Microprocessor and accumulate upon its depletion. We conclude that splicing is a key step in generating pri-let-7

transcripts with a structure that enables downstream processing events to produce appropriate levels of mature let-7.

B. Introduction

MiRNAs are ~22 nucleotide (nt) non-coding RNAs that use imperfect base pairing to target mRNAs for down-regulated expression (Pasquinelli, 2012). Thousands of miRNAs have been discovered in a wide variety of organisms, including plants, flies, worms, and humans (Kozomara & Griffiths-Jones, 2014). Moreover, a single miRNA has the potential to regulate hundreds of different targets (Ha & Kim, 2014). Consequently, miRNAs are implicated in almost all biological pathways and their misexpression can lead to developmental impairment and disease. The let-7 miRNA in particular is abnormally expressed in various human cancers and the consequent misregulation of its protein-coding targets has been directly implicated in the disease state (Mondol & Pasquinelli, 2012; Gurtan & Sharp, 2013). Additionally, the mature sequence of let-7 is perfectly conserved across Bilaterian species (Pasquinelli et al., 2000), making it a relevant model for studying the regulation of miRNA biogenesis.

The general miRNA biogenesis pathway is well understood (Finnegan & Pasquinelli, 2013; Ha & Kim, 2014). Briefly, long primary miRNAs (pri-miRNAs) are transcribed by RNA Polymerase II (Pol II) and, like mRNAs, they are capped and polyadenylated. The Microprocessor, which contains the RNase III enzyme Drosha and the RNA binding protein Pasha (DGCR8 in mammals), cleaves the pri-miRNA into a ~65 nt imperfect hairpin known as the precursor miRNA (pre-

miRNA). A non-canonical pathway exists for a class of miRNAs known as mirtrons, which bypasses the need for Drosha processing (Westholm & Lai, 2011). Instead, the pre-miRNA is a ~65 nt intron that is excised by the spliceosome. In mammals and Drosophila, Exportin5 then transports the pre-miRNA from the nucleus to the cytoplasm where it is subsequently cleaved by Dicer to produce the mature ~22 nt miRNA. The mature miRNA sequence, or guide strand, is then loaded onto an Argonaute protein, which is the principle protein in the miRNA Induced Silencing Complex (miRISC). The loaded miRISC targets mRNAs with partial complementarity to induce deadenylation and/or translational repression (Ha & Kim, 2014).

The substrate requirements for the Microprocessor have been studied extensively (Lee et al., 2003; Denli et al., 2004; Gregory et al., 2004; Zeng & Cullen, 2005; Han et al., 2006; Kim & Kim, 2007; Kataoka et al., 2009; Flynt et al., 2010; Janas et al., 2011; Warf et al., 2011; Macias et al., 2012; Auyeung et al., 2013; Conrad et al., 2014; Quick-Cleveland et al., 2014). Through deep sequencing and *in vitro* processing assays, a general model of the ideal Drosha substrate has emerged. The archetype metazoan pri-miRNA contains a pre-miRNA hairpin ~65 nt long, or ~3 helical turns, which features a ~10 nt terminal loop, internal bulges in the stem every ~11 nt or one helical turn, and flanking unstructured RNA sequences (known as the basal segment) (Lee et al., 2003; Zeng & Cullen, 2005; Han et al., 2006). Sequences adjacent to Drosha cleavage sites tend to fold into 2-4 nt symmetrical internal loops. These sites are found on average ~11 nt away from the unpaired basal segment. Recent work suggests

that processing of human pri-miRNAs also involves a sequence motif that is not seen in worms (Auyeung et al., 2013; Conrad et al., 2014). Across species, though, the secondary structure is critical for determining the efficiency of pri-miRNA processing.

While investigating the transcription of primary let-7 (pri-let-7), our lab discovered two transcriptional start sites (A and B) that produce nascent pri-let-7 transcripts and one or both give rise to an SL1 *trans*-spliced isoform (Bracht et al., 2004). Notably, the 3' splice site (ss) required for *trans*-splicing is conserved in sequence and position in let-7 genes in other nematode species (Bracht et al., 2004). *Trans*-splicing is a common event in *C. elegans* as up to of 70% of mRNAs have their 5' ends replaced by one of two, 22 nt trimethylguanosine-capped RNA leader sequences (SL1 or SL2) (Blumenthal, 2012). In some cases, *trans*-splicing is essential for separating mRNAs in an operon. The function of this event for mRNAs from non-operonic genes remains elusive, but it is thought to aid in their nuclear export or translation (Wallace et al., 2010). None of these functions seems relevant for pri-miRNAs, as they are non-coding RNAs processed in the nucleus, eliminating the SL1 sequence before export to the cytoplasm. Instead, we predicted that the SL1 sequence, the act of splicing, and/or a resulting structural change in the primary let-7 RNA is important for downstream miRNA processing events. Here, these possibilities were tested through a series of *in vitro* structural studies, *in vivo* rescue experiments, and RNA immunoprecipitation assays. Altogether, our results support the conclusion that SL1 *trans*-splicing of pri-let-7 in *C. elegans* facilitates structural

rearrangements that promote Microprocessor binding and cleavage. This example raises the possibility that splicing may be broadly employed to enable structural changes that regulate Microprocessor activity of the many other pri-miRNAs embedded in transcripts subject to *cis*- or *trans*-splicing.

C. Results and Discussion

C.1 Splicing remodels the pri-let-7 secondary structure

Trans-splicing of pri-let-7 occurs when the 3' ss, which is found 38 nt upstream from the pre-let-7 sequence, is cleaved, and the 22 nt SL1 sequence is appended, replacing the cap and the region between the 3' ss and transcriptional start sites (Bracht et al., 2004). Secondary structure predictions using lowest free energy thermodynamic computations on the mfold server provided initial evidence that there were differences between unspliced and SL1-spliced pri-let-7 RNAs (Zuker, 2003; Bracht et al., 2004). However, previous research on structural features of pre-miRNAs found eight out of ten experimentally determined structures differed from those predicted by mfold (Krol et al., 2004), primarily in the terminal loop region and other secondary bulges in the stem. Therefore, we sought to interrogate the potential pri-let-7 secondary structures using biochemical methods. We subjected ~200 nt of *in vitro* transcribed spliced and unspliced primary transcripts to RNase secondary structure analysis to identify single-stranded and double-stranded regions in the folded RNA molecules (FIG 4.4.S1). Figure 4.1 depicts the structures best supported by RNase structure probing coupled with mfold predictions. Several bases in

potential loop regions produced variable patterns of paired and unpaired structures, indicating transient interactions in these regions. These results indicate that the unspliced pri-let-7 transcript features structured regions at the base of the precursor, disrupting what should be unpaired areas critical for Drosha slicing activity (Zeng & Cullen, 2005; Han et al., 2006; Warf et al., 2011; Quick-Cleveland et al., 2014). In contrast, the SL1-spliced model takes on a structure that better resembles a canonical pri-miRNA substrate, including a terminal loop, internal loops in the upper and lower stem, and flanking unpaired basal segments.

C.2 The splicing-induced structural changes facilitate pri-let-7 processing

We investigated the role of SL1 splicing of pri-let-7 *in vivo* by introducing let-7 transgenes using Mos-1 transposon-mediated Single Copy Insertion (MosSCI) and crossing the resulting transgenic animals to a let-7 null background, *let-7(mn112)* (Reinhart et al., 2002; Frokjaer-Jensen et al., 2008) (FIG 4.4.2A). The transgenes were engineered to test the splicing, sequence, and structural requirements of pri-let-7 processing. As part of the heterochronic pathway, let-7 is crucial for developmental timing of cell fates during the fourth larval stage (L4) to adult transition (Mondol & Pasquinelli, 2012). Pri-let-7 expression begins at the first larval stage (L1) and oscillates throughout development with peaks preceding each larval molt (Van Wynsberghe et al., 2011). During early larval stages, LIN-28 protein co-transcriptionally binds to pri-let-7 and blocks its processing (Van Wynsberghe et al., 2011; Stefani et al.,

2015). By the L4 stage, LIN-28 levels have decreased over ten-fold, allowing processing of pri-let-7 to proceed. Worms with mutant alleles of let-7 display various phenotypes ranging from lethality to aberrations in cuticular structures known as alae (Reinhart et al., 2000) (FIG 4.4.2B). Accordingly, we examined whether our transgenic animals exhibited phenotypes associated with loss of let-7 (Reinhart et al., 2000). In addition, we analyzed primary and mature let-7 RNA expression from the transgenes at the L4 stage of development using Northern blot and quantitative reverse transcriptase PCR assays (qRT-PCR) (FIG 4.4.2C-F).

Although the transgenic worms lacking a 3' ss (FIG 4.4.2A Δ3'SS) were viable, about 25% of the adults exhibited patchy or no alae, indicative of incomplete rescue activity (FIG 4.4.2B). This degree of abnormal alae development is similar to that observed in *let-7(mg279)* mutants, which produce 2-fold less mature let-7 miRNA (Reinhart et al., 2000; Bracht et al., 2004). We confirmed that the pri-let-7 RNAs generated in the Δ3'SS transgenic animals do not undergo *trans*-splicing by RT-PCR assays (data not shown). Consistent with the defective alae phenotype, the Δ3'SS strain exhibited reduced processing of let-7 with a five-fold accumulation of pri-let-7 and a greater than two-fold decrease in the level of mature let-7, compared to the amount of these RNAs detected in the wild-type (WT) strain (FIG 4.4.2C-F). These results demonstrate that SL1-splicing is important for let-7 biogenesis *in vivo*.

Because the Microprocessor has been shown to associate with spliceosomal proteins, we asked whether the act of splicing is needed for let-7

biogenesis (Gregory et al., 2004; Kim & Kim, 2007; Kataoka et al., 2009; Janas et al., 2011). The SL1 “pre-spliced” construct replaces the 3’ss with the 22 nt SL1 sequence (FIG 4.4.2A SL1). These worms still use the annotated A and B transcriptional start sites to produce pri-let-7 transcripts that encode the SL1 sequence, without undergoing splicing; the resulting RNAs are predicted to fold similarly to the WT *trans*-spliced version in the vicinity of the precursor sequence. These transgenic worms have almost no discernible alae phenotypes, indicating that the “pre-spliced” SL1 pri-let-7 transgene retains near WT levels of rescue activity (FIG 4.2B). Consistent with these observations, “pre-spliced” SL1 worms accumulate modest levels of pri-let-7 and produce amounts of mature let-7 that are comparable to those expressed from the WT construct (FIG 4.2C-F). Thus, any contribution of the SL1-spliceosome to let-7 biogenesis is minor.

We also examined the necessity of the SL1 sequence by replacing it with a random 22 nt sequence (RDM) that is similar in GC content to the SL1 sequence but unable to support the same structure (FIG 4.2A). The transgenic worms expressing the RDM construct displayed defective alae and let-7 expression phenotypes similar to those of the Δ3’SS transgenic animals, indicating limited rescue activity (FIG 4.2B-F).

To differentiate further whether the structural change provided by the SL1 sequence, as opposed to the sequence itself, is crucial for facilitating adequate let-7 expression, we generated a “misfolded” SL1-spliced construct (MSF). This transgene includes the 22 nt SL1 sequence in place of the 3’ss and a 48 nt mutagenized region 24 bases downstream of the 3’ end of the pre let-7 (FIG

4.2A). Despite the inclusion of the SL1 sequence, this construct was completely incapable of rescuing the null let-7 mutant, implying that the presence of the SL1 sequence is not sufficient for facilitating pri-let-7 processing. When analyzed by mfold, this transgene creates a highly structured hairpin at the base of pre-let-7, similar to that found in unspliced pri-let-7. Because of the high complementarity in the sequence, this molecule is predicted to take on a very rigid structure, with little room for flexibility. The likely inability of the MSF RNA secondary structure to "breathe" to the same degree observed for the unspliced form suggests that the potential for RNA molecules to be dynamic contributes to the processing of pri-let-7 in transgenic worms lacking a 3'ss. Furthermore, it points to structural remodeling as the primary role for *trans*-splicing of pri-let-7 transcripts.

C.3 The Microprocessor prefers spliced pri-let-7 *in vivo*

As part of the Microprocessor, Pasha is responsible for recognizing appropriate pri-miRNA substrates and guiding Drosha to cleave them. To test if Pasha differentially binds endogenous spliced versus unspliced pri-let-7 transcripts, we analyzed RNAs that co-immunoprecipitated with a rescuing GFP-tagged Pasha protein (Lehrbach et al., 2012). Immunoprecipitation of extracts from L4-stage worms revealed a ~15-fold enrichment of SL1-pri-let-7 in the Pasha:GFP expressing strain (+) compared to the non-transgenic control (-) (FIG 4.3A). In contrast, association of the unspliced isoform with Pasha:GFP was not detected above background levels. Thus, the enrichment of total pri-let-7 in the Pasha-GFP IP seems to be entirely reflective of SL1-pri-let-7 binding to this Microprocessor factor. This association is specific for the miRNA-containing

transcript, as abundant SL1-spliced mRNAs, such as tba-1 and rbm-28, were not detected in the Pasha-GFP IPs.

To further test if there is a difference in the pri-let-7 isoform utilized by the Microprocessor *in vivo*, we examined the accumulation of pri-miRNA transcripts in Pasha mutant worms. The temperature sensitive *pash-1(mj100)* worms exhibit embryonic lethality when grown at 25°C but develop normally at 15°C (Lehrbach et al., 2012). An extrachromosomal transgene driving ubiquitous expression of a PASH-1:GFP fusion protein is able to rescue the temperature sensitive allele (Lehrbach et al., 2012). Compared to control worms expressing the PASH-1:GFP transgene (+), there was a four-fold increase in the level of SL1- and total pri-let-7 in the pasha mutants (-) (FIG 4.3B and C). The unspliced pri-let-7 transcripts were unaffected by the loss of Pasha, as was pri-mirtron-62, which bypasses the canonical miRNA biogenesis pathway (Ruby et al., 2007a). These results demonstrate that the spliced form of pri-let-7 is specifically sensitive to the loss of Pasha and, thus, is the preferred substrate for processing *in vivo*. This conclusion is also consistent with previous observations that the SL1 spliced pri-let-7 transcript specifically accumulates upon disruption of a let-7 and Argonaute-Like-Gene 1 (ALG-1) auto-regulatory loop that promotes pri-let-7 processing and is the preferred substrate for processing when LIN-28 mediated repression is absent from the first larval stage (Van Wynsberghe et al., 2011; Zisoulis et al., 2012).

Taken together our results demonstrate that a role of *trans*-splicing in let-7 biogenesis is to remodel the secondary structure of pri-let-7 to promote favorable

Drosha processing. Removing the spliceosome requirement by replacing the 3' ss with the SL1 sequence in our transgene did not have a substantial effect on let-7 biogenesis. Although it is possible that the SL1 sequence itself could help recruit miRNA biogenesis factors, its presence is not sufficient for supporting processing when it is uncoupled from seeding structural remodeling of pri-let-7. Since *trans*-splicing was not an absolute requirement for let-7 rescue activity and production of the mature miRNA, it is likely that the structure of the unspliced transcript is somewhat flexible or that the Microprocessor has some degree of accessibility to suboptimal processing substrates *in vivo*. Considering the multiple *cis*-acting elements and *trans*-acting factors that regulate pri-let-7 transcription (Johnson et al., 2003; Roush & Slack, 2009; Kai et al., 2012), it is also possible that processing of the spliced versus unspliced pri-let-7 transcripts may be under differential control depending on the timing or location of expression.

Although *trans*-splicing is not common across animal species, pri-miRNAs in other organisms are often found in transcripts subject to *cis*-splicing. The majority of human miRNAs, including several let-7 family members, are located in introns of host genes (Kim & Kim, 2007; Roush & Slack, 2008). In some cases, Drosha excises the pre-miRNAs rapidly and independently of pre-mRNA splicing (Kim & Kim, 2007). In other studies, the spliceosome and Microprocessor associate together on the pre-mRNA and work mutually to coordinate miRNA processing and pre-mRNA splicing (Kataoka et al., 2009; Janas et al., 2011). In addition, exon-intron junction spanning pre-miRNAs have been found to be subject to regulation by alternative splicing through changes that occur in the pre-

miRNA secondary structure (Melamed et al., 2013). Finally, in plants, some pri-miRNAs contain introns that enhance biogenesis when spliced, although the reason for this effect was not clear (Bielewicz et al., 2013; Schwab et al., 2013). Considering the importance of RNA structure for pri-miRNA processing by Drosha, splicing may serve as a mechanism to regulate the folding and, hence, the efficiency of recognition by the Microprocessor for many pri-miRNAs. Thus, the disruption of splicing that accompanies some stress and disease conditions could contribute to the reduced levels of certain miRNAs associated with these states.

D. Materials and Methods

Several methods including worm staging, MosSCI integration, RNA extraction, RNAi treatment, Western blot analysis, Agarose and PAGE Northern Blot analysis, and reverse transcriptase PCR assays have been previously described (Van Wynsberghe et al., 2011; Zisoulis et al., 2012). Wild-type worms were N2 Bristol.

D.1 *In vitro* transcription of RNA

DNA templates for *in vitro* transcription of unspliced and SL1-spliced pri-let-7 were amplified from wild-type genomic DNA and SL1-spliced pri-let-7 plasmid DNA using the primers listed below. DNA templates were then purified with the QIAquick Gel Extraction Kit (Qiagen). The MEGA Shortscript T7 transcription kit (Ambion) was used to transcribe RNA from these DNA templates and the resulting RNA was phenol chloroform extracted and analyzed by

spectrophotometry for purity and quantity. 50 µg of transcribed RNA was treated with Calf Intestinal Alkaline Phosphatase (CIP) to remove the 5' phosphates and then isolated and purified through 5% TBE-UREA polyacrylamide gel electrophoresis (PAGE) by cutting out the band of interest, crushing it and rocking in 500 µL 0.3 M NaCl at 4°C. CoStar Spin-X columns containing a 0.45 µM filter were used to isolate the supernatant, which was then alcohol precipitated and eluted to 10 µL. The purified transcripts were 5' end labeled with ɣ-32P in a T4 Polynucleotide Kinase (PNK) reaction and spun through an Illustra Microspin G-50 column (GE Healthcare) to remove unincorporated nucleotides. Afterwards, transcripts were run on a test PAGE to ensure homogeneity.

D.2 RNase secondary structure analysis

As individual mixes, unspliced and spliced labeled RNA was mixed with 1 µg of Yeast RNA (Ambion) boiled and allowed to refold at room temperature. In separate tubes, samples were treated with either 0.02 Units (U) of RNase V1 (Ambion), 0.02 U of RNase T1 (Ambion), 0.002 U of RNase A (Ambion) or buffer for 15 minutes at room temperature. An all nucleotide ladder was produced by boiling the RNA mix in an alkaline hydrolysis buffer for 5 minutes, and Decade Markers (Ambion) were labeled and prepared per instructions. Reactions were stopped by adding formamide loading buffer and analyzed by 20 cm x 40 cm x 0.4 mm 8% TBE-UREA-PAGE. Gels were dried, exposed to phosphor screens, and scanned on a Typhoon phosphorimager.

D.3 RNA immunoprecipitations

RNA immunoprecipitations were performed as previously described (Van Wynsberghe et al., 2011; Zisoulis et al., 2012; Broughton & Pasquinelli, 2013). Briefly, L4-staged Pasha:GFP and wild-type, nontransgenic, worms were resuspended in Lysis Buffer [100 mM NaCl, 25 mM HEPES pH7.5, 250 µM EDTA pH8.0, 0.1% SDS, 0.1% NP-40, Complete Mini Protease Inhibitor cocktail (Roche), 2 mM DTT, 25 U/mL rRNasin (Promega)] and sonicated in ten second pulses, five times, resting on ice for one minute in between pulses. Extracts were cleared by spinning at 16,000 x g for 15 minutes, 4°C, then snap frozen on dry-ice ethanol and stored at -80°C. Lysates were thawed by rocking at 4°C and protein concentration measured with a Qubit Fluorometer. Equal lysate amounts were precleared with Protein G Dynabeads (Invitrogen), followed by incubation at 4°C either with pre-conjugated GFP magnetic beads (Clonotek) for two hours or with GFP polyclonal antibody overnight and 1 hour with washed Protein G Dynabeads. Beads were washed twice with high salt wash buffer [50 mM Tris-HCl, pH 7.4, 1 M NaCl, 1 mM EDTA, 1% NP-40, 0.1% SDS, 0.5% sodium deoxycholate], low salt wash buffer [20 mM Tris-HCl pH 7.4, 10 mM MgCl2, 0.2% Tween-20], and proteinase K buffer [100 mM TrisCl pH 7.4, 50 mM NaCl, and 10 mM EDTA], before treatment with proteinase K (Invitrogen) and urea. RNA was Trizol (Invitrogen) extracted, treated with RQ1 DNase (Promega) and re-extracted before cDNA synthesis with random primers and Superscript III (Invitrogen). PCR was performed with the below listed oligos (IDT).

D.4 Sequences and Primers used in this study

Sequences inserted/deleted in this study: SL1 5’ GGTTTAATTACCCAAGTTTGAG 3’, 3’ss 5’ TTTTCAG-3’, RDM 5’ GTATCCGTAAAGCTCATTAAGC 3’, MSF 5’ AAGACGACGCAGCTTCGAAGAGTTCTGTCTCCTCTACTAATCGCCTGC 3’

Primers for T7 transcription: unspliced pri-let-7 fwd 5’ TAATACGACTCACTATAGGTTTTTCAGGCAAGCAGGCGAT 3’ and rev 5’ GAAAAACAAAGAGGTGAAAGTAAG 3’, SL1-spliced let-7 fwd 5’ TAATACGACTCACTATAGGTTTAATTACCCAAGTTTGA 3’ and rev 5’ AAAGAAAGTTGTGAGAGCAAGACG 3’. Primers for RT-PCR: unspliced pri-let-7 fwd 5’ GTCTAATTTAACAACAAGTACTAATCCATT 3’, SL1-sequence fwd 5’ GGTTTAATTACCCAAGTTTGAG 3’, total pri-let-7 fwd 5’ CAAGCAGGCGATTGGTGGA 3’ and pri-let-7 rev 5’ GTAAGGTAGAAAATTGCATAGTTC 3’, pri-mir-58 fwd 5’ GGCTTCAGTGGCTCCTCT 3’ and pri-mir-58 rev 5’ CGTTTAGTGCGCACATTCGGCAA 3’, mirtron-62 fwd 5’ CCATGTACTCCGGCTATAGTGAG 3’ and mirtron-62 rev 5’ GATGTTGAACAACCTGTAAGCTAGATT 3’, actin fwd 5’ GTGTTCCCATCCATTGTCGGAAGAC 3’ and actin rev 5’ GTGAGGAGGACTGGGTGCTCTT 3’, tba-1 fwd 5’ ATGCGTGAGGTCATCTCCAT 3’ and tba-1 rev 5’ TGATGGCATAGTTCCATCGG 3’, rbm-28 fed 5’ GATTCAGGAGTTATGGGTATAATTCTTC 3’ and rbm-28 rev 5’ GGCTGTATCCGCCGTAGC 3’.

E. Acknowledgments

We thank E. Finnegan and members of the Pasquinelli lab for their suggestions and critical reading of this manuscript. We thank undergraduate helpers A. Weaver and K. Nassiri for helping with phenotype analysis, RNA extractions, and PCRs. We thank Dr. Miska for the Pasha mutant and GFP-tagged strains, Dr. David for use of a Real-Time PCR machine, and Dr. Oegema and Dr. Desai for the polyclonal GFP antibody. V.M. was supported in part by the University of California, San Diego, Cellular and Molecular Genetics Training Program through an institutional training grant from the National Institute of General Medical Sciences (T32 GM007240) and an NIH Ruth L. Kirschstein National Research Service Award Individual Predoctoral Fellowship to Promote Diversity in Health-Related Research (F3194041A). This work was funded by grants from the NIH (GM071654) the Keck, and Peter Gruber Foundations (A.E.P.).

Chapter IV is a reprint of “Splicing remodels the let-7 primary microRNA to facilitate Drosha processing in *Caenorhabditis elegans*.” Mondol, Vanessa, Ahn, Byoung Chan. & Pasquinelli, Amy E., RNA, 2015. I was the primary author.

F. Addendum

RNA immunoprecipitations of the Microprocessor were challenging. Table 4.1 outlines the many troubleshooting efforts taken in order to pull down primary miRNA. Ultimately, a tagged Drosha construct was not useful for this purpose

due to its RNase activity, so we used a tagged Pasha construct instead. In addition, sonicating fresh worms with fewer pulses minimized RNA degradation.

TABLE 4.1 RNA Immunoprecipitation troubleshooting. A number of different strategies were employed to pull down primary let-7 from the Microprocessor. Two constitutively expressed Drosha transgenic strains were constructed via Mos-1 transposon mediated homologous recombination (MosSCI) in a Drosha *null* background such that one contained a basic FLAG tag, while the other contained an enhanced FLAG tag as well as a GFP tag. (FLAG:Drosha and 3XFLAG:GFP:Drosha). A temperature sensitive Pasha strain, obtained from the Miska lab, was grown at non-permissive temperatures (25°C) in the presence of constitutively expressed a GFP-tagged Pasha chromosomal array (Pasha-ts + GFP:Pasha). Worms were grown until the fourth larval stage (L4) when expression of let-7 begins. For some experiments, L4 time courses were collected every two hours from 30-34 hours. UV crosslinking and Sonication were carried out as described in the Methods section. Different antibodies were used including pre-conjugated FLAG (Sigma) and GFP beads (ChromoTek), as well as Protein G Dynabeads with FLAG antibody (Sigma) and GFP antibody (Oegama lab). Ultimately, pri-let-7 was amplified from RIP-cDNA using a reverse oligo that begins before the Drosha cleavage site.

Strain Used	Lysate Prep	IP time	Result	Course of action
FLAG:Drosha	L4 staged, UV crosslink, Sonicate on ice	1hr preclear, 1hr IP	IP unsuccessful	Replace pre-conjugated FLAG beads with FLAG mouse antibody
FLAG:Drosha	L4 staged, Sonicate on ice	1.5hr preclear, 1hr IP	IP unsuccessful	Try again
FLAG:Drosha	L4 staged, UV crosslink, Sonicate on ice	1hr IP	IP unsuccessful	Use Miska strain
Pasha-ts + GFP:Pasha	L4 staged, UV crosslink, Sonicate on ice	30m preclear, 1.5hr IP	Western Blot unclear.	Use more antibody. Compare pre-conjugated GFP beads to Oegema rabbit GFP antibody
Pasha-ts + GFP:Pasha	L4 staged, UV crosslink, Sonicate on ice	30m preclear, 1.5hr IP	Clonotech GFP beads gives very clean result on Western. cDNA made does not amplify pri-miRNAs	Do not crosslink
Pasha-ts + GFP:Pasha	L4 staged, Sonicate on ice	30m preclear, 3hr IP	Oegema GFP antibody works well for IP. Water contamination in RT-PCR. cDNA not amplifying pri-miRNAs	Repeat
Pasha-ts + GFP:Pasha	L4 staged, Sonicate on ice	1hr preclear, 1hr IP	IP worked.	
Pasha-ts + GFP:Pasha	L4 staged, Sonicate on ice	1hr IP	IP worked. cDNA does not amplify pri-let-7	Try other timepoints. Run Northern Blot instead of RT-PCR
Pasha-ts + GFP:Pasha	L4 vs. mixed staged, UV crosslink, Sonicate on ice	1hr IP	IP worked well. Input and Sup Total RNA on agarose blot shows signs of degradation, Northern does not detect pri-let-7	Try 30-34 hr 2hr timecourse
Pasha-ts + GFP:Pasha	30-34hr staged, Sonicate on ice	45m preclear, 1hr IP	IP failed	repeat
Pasha-ts + GFP:Pasha	30-34hr staged, Sonicate on ice	1hr preclear, O/N IP	IP- bead sample not very enriched. cDNA does not amplify pri-let-7	repeat with crosslinking
Pasha-ts + GFP:Pasha	30-34hr staged, UV Crosslink, Sonicate on ice	1hr preclear, 1hr IP	did not extract RNA or test Western because decided to repeat without sonication	
Pasha-ts + GFP:Pasha	30-34hr staged, UV Crosslink, dounce homogenizer on ice	1hr preclear, 1hr IP	Western failed while testing IP. IP did not work very well. cDNA did not amplify pri-let-7	new aliqot of Ogema GFP antibody not working as well
3XFLAG:GFP:Drosha	L4 staged, UV crosslink, liquid nitogen crush	overnight IP	Used FLAG conjugated beads, but they didn't work	repeat with FLAG mouse antibody
3XFLAG:GFP:Drosha	L4 staged, UV crosslink, dounce homogenizer	1hr preclear, O/N IP	can barely detect protein enrichment on beads. Used RNA for Northern but saw Supernatent samples degraded	repeat, doubling protease inhibitor cocktail and RNasin in buffer
3XFLAG:GFP:Drosha	L4 staged, UV crosslink, dounce homogenizer	45m preclear, 1hr vs. O/N IP	after 1hr IP protein no longer in Sup. After O/N IP, protein very faint on beads. 1hr cDNA is unable to amplify pri-let-7 and pri-miR-58 on beads	repeat, doubling protease inhibitor cocktail and RNasin in buffer, reducing DTT
3XFLAG:GFP:Drosha	L4 staged, UV crosslink, snap-freeze, dounce homogenizer	1hr preclear, 1hr IP	Protein detectable in Input, not Beads or Sup. Input and Sup RNA appear slightly degraded.	repeat, doubling protease inhibitor cocktail and RNasin in buffer
3XFLAG:GFP:Drosha	L4 staged, UV crosslink, snap-freeze, dounce homogenizer	45m preclear, 1hr IP	Protein detectable in Input, not Beads or Sup.	repeat using rabbit α-FLAG
3XFLAG:GFP:Drosha	L4 staged, UV crosslink, snap-freeze, dounce homogenizer	45m preclear, 1hr IP	Protein detectable in Input, not Beads or Sup.	increase Input protein concentration to 2mg/mL, try pre-conjugated GFP beads and FLAG mouse antibodoies
3XFLAG:GFP:Drosha	L4 staged, UV crosslink, snap-freeze, dounce homogenizer	45m preclear, 1hr IP	Protein detectable in Input, not Beads or Sup.	Increase Input protein to 4mg/mL, reduce washes, change wash buffers: test High Salt, Mild,and Tris buffers in separate reactions
3XFLAG:GFP:Drosha	L4 staged, UV crosslink, snap-freeze, dounce homogenizer	45m preclear, 1hr IP	Protein detectable in Input, not Beads or Sup.	repeat with GFP-IP kit vs. FLAG mouse, reduce protein to .4mg/mL, collect all washes
3XFLAG:GFP:Drosha	L4 staged, UV crosslink, snap-freeze, dounce homogenizer	45m preclear, 1hr IP	No protein detectable in washes. FLAG-IP: no protein on beads or sup. GFP-IP: very little protein on Beads some left on Sup. cDNA is able to amplify pri-let-7 and pri-miR-58 in Input and Sup but not beads.	GFP-IP kit seems to work minimally. Drosha is either degraded or sequestered by FLAG antibody, Dynabeads, or solution? Test 3 new Lysate Buffers and 5mg/mL protein with FLAG-mouse. Repeat GFP-IP kit
3XFLAG:GFP:Drosha	L4 staged, UV crosslink, snap-freeze, dounce homogenizer	45m preclear, 2hr GFP-IP, 1hr FLAG-IP	FLAG-IP: no protein on beads, little in sup no matter which buffer. GFP-IP: very little protein on beads. Where is FLAG-Drosha going?	Repeat testing with/without DNase, and with/without mild sonication (tested 3 different conditions) to break up DNA that maybe interfereing since Microprocessor works co-transcriptionally. 8 IPs total
3XFLAG:GFP:Drosha	L4 staged, UV crosslink, snap-freeze, dounce homogenizer	1hr preclear, 40m IP	still no Drosha in beads or sup.	repeat GFP-IP (Allele vs. Chromotek) skip pre-clear. Test FLAG-IP with/without crosslinking and use Trizol extracted protein instead of boiled SDS protocol
3XFLAG:GFP:Drosha	L4 staged, snap-freeze, dounce homogenizer	no preclear, 1hr IP	GFP-IP: Very faint band in Chromotek beads, but not Allele. FLAG-IP: protein not detectble by Western	Repeat. Switch from Dynabeads/Magnetic Beads to Sepharose/Agarose. Do not snap-freeze worms before making lysates. Use mild sonication
3XFLAG:GFP:Drosha	mixed stage, UV crosslink, sonicate 5s 10X, w/ 1min ice intervals	no pre-clear, 2hr IP	IP successful! prepared RNA extractions with Glyco-Blue to ensure bead pellet is not lost. No signs of RNA degradation. cDNA amplifies pri-miR-58 in Input, Beads and Sup, but pri-let-7 only detectable in Input and Sup.	Drosha may be cleaving too fast for pri-miRNA pulldown. Repeat RT-PCR using 3' oligo upstream from initial Drosha cleavage site. Test cDNA from ALL previous successful IP's with intact RNA (~70 cDNA samples)
re-tested cDNA from previous succesful IP's			RIPs from Pasha-ts strain were successful at pulling down pri-let-7. Both Dynabeads and Chromotek pre-conjugated beads worked well, with/without crosslinking. SL1-pri-let-7 also detectable on N2 bead control.	amplified actin or 18S controls as well as other pri-miRNAs, and SL1-spliced controls. Optimized # of PCR cycles.

FIGURE 4.1. Secondary structures of unspliced and spliced pri-let-7 RNAs. RNA secondary structure analysis of *in vitro* transcribed unspliced and SL1-spliced model pri-let-7 transcripts. Numbers indicate position of base in model transcript. The composite results from four independent experiments are shown. Here we show a close-up view of the area in the red box subject to processing by the Microprocessor. Three shades of color are used to represent “high”, “medium” and “low” sensitivity of each nucleotide to the RNase enzyme, such that the darker the color the more reactive to cleavage and the more “structured” (blue) or “unpaired” (red) the base appears to be. Nucleotides in shades of blue circles were sensitive to cleavage by RNase V1, denoting double-strandedness. Shades of red circles represent single-stranded nucleotides that were cleaved in response to RNase T1 or RNase A treatment. Nucleotides labeled both red and blue indicate evidence for single and double-stranded nature, which is expected from G-U pairs and regions of the transcript that are more dynamic. Unlabeled (white) nucleotides were inconclusive. Open circles between nucleotides represent regions of expected pairing that were not detectable by the RNase structure probing. The mature let-7 miRNA sequence is highlighted in yellow. The SL1 sequence is highlighted in green and labeled. Representative PAGE used for data analyses are shown in Supplementary Figure 4.1.

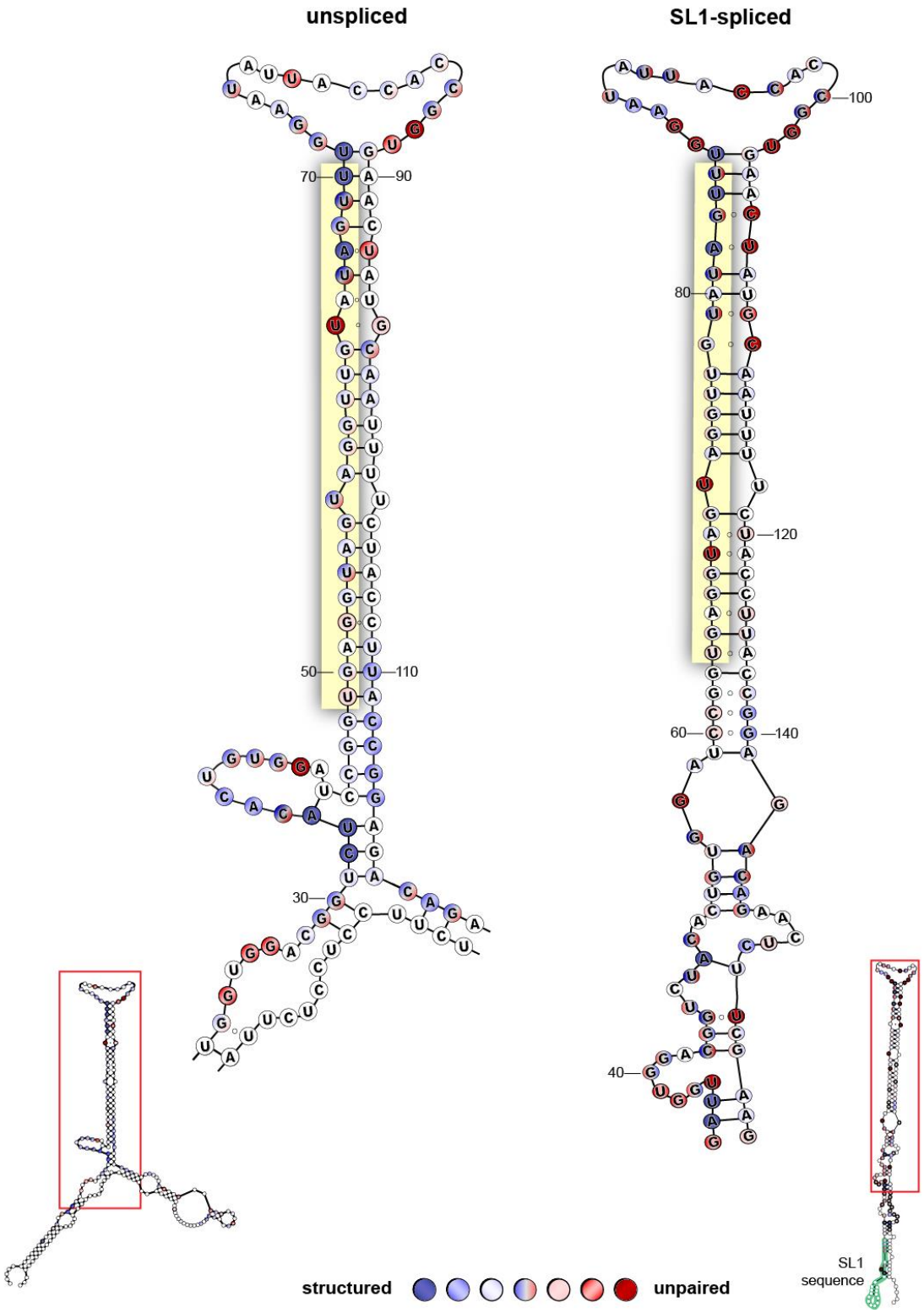
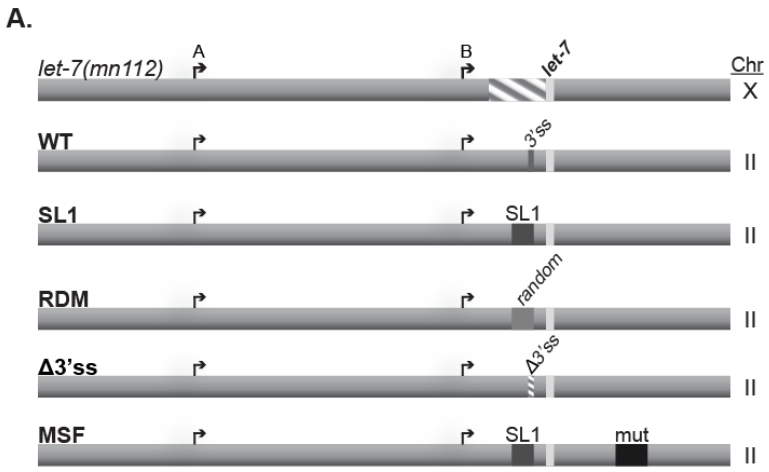
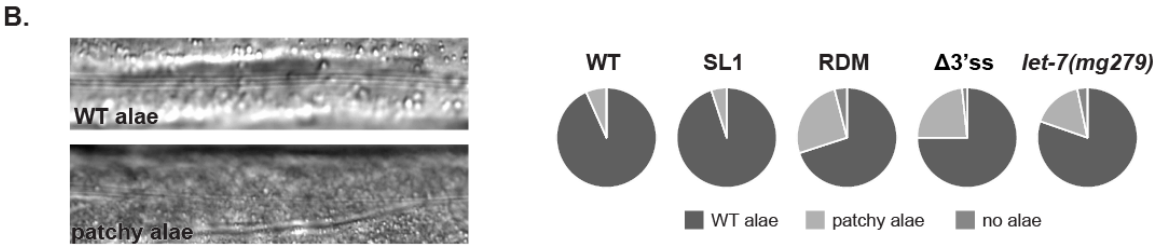
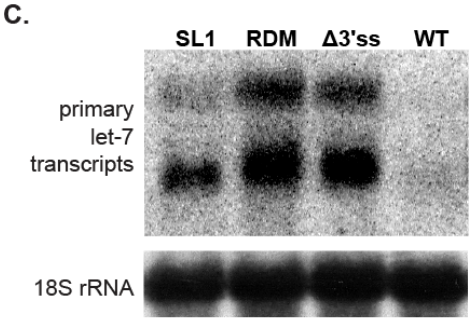
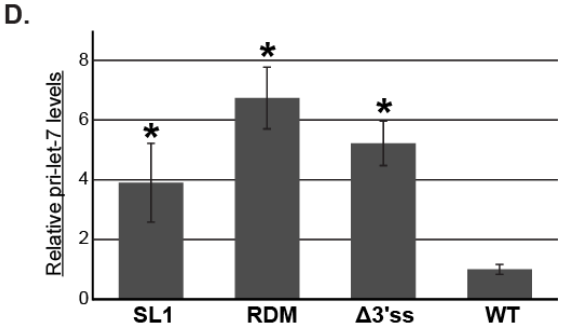

unspliced
SL1-spliced
70
90
50
110
30
100
80
120
60
140
40
structured
unpaired
SL1
sequence

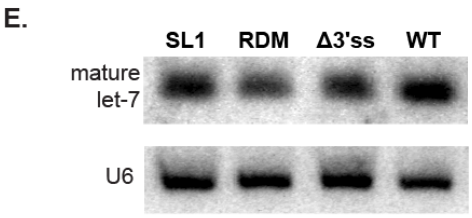
FIGURE 4.2. Splicing is important for let-7 biogenesis *in vivo*. (A) Single copy transgenes with altered versions of the *trans*-splice site were inserted in Chromosome (Chr) II and expressed in the *let-7(Δmn112)* (Chr X) null mutant background to test for effects on rescue activity and let-7 biogenesis. The arrows represent let-7 A and B transcriptional start sites, the striped area represents the *mn112* deletion, and the light rectangle is indicative of the mature let-7 sequence. In addition to the wild-type (WT) rescue construct, which contains the let-7 promoter regions, transcriptional start sites, and the 3' splice site (3'ss) recognized by the spliceosome, four mutated versions of the let-7 transgene were generated. The "pre-spliced" construct (SL1) replaces the 3' ss with the 22 nt splice leader sequence. The "random" construct (RDM) replaces the 3' ss with a 22 nt sequence that is similar in GC content to the SL1 sequence but is predicted to support a different secondary structure. The splice site knockout (Δ3'ss) is missing 7 nt essential for splicing recognition. The "misfolded" construct (MSF) contains the SL1 sequence in place of the 3' ss but also includes a 48-nt mutation that creates a hairpin at the base of the pre-let-7 hairpin, similar to the unspliced secondary structure. (B) Analysis of the alae formation phenotype. Alae are a group of three cuticular ridges that form along the length of the adult *C. elegans* worm. Examples of alae in WT and worms that have insufficient let-7 activity, which results in "patchy" and gapped alae, are shown. Alae were analyzed by high-powered microscopy in transgenic worms and in a mutant that expresses 2-fold reduced levels of mature let-7 (*let-7(mg279)*). Results for RDM are presented as the average of two independent experiments where n= 25, while all other strains are presented as the average of three independent experiments where n=20 in each. (C-F) Total RNA from triplicate L4 staged SL1, RDM, Δ3'ss, and WT transgenic worms was used for Northern Blot and qRT-PCR analyses. Worms with the MSF transgene were inviable and, thus, could not be analyzed for let-7 expression. (C) Representative agarose Northern Blot of primary let-7 expression. Detection of 18S ribosomal RNA serves as a loading control. (D) Quantitative RT-PCR of transgenic worms where pri-let-7 (all isoforms) levels were normalized to 18S rRNA and relative expression is compared to average WT expression. Standard error is depicted. *$p < 0.05$. (E) Representative PAGE Northern Blot of mature let-7 expression. U6 snRNA serves as a loading control (F) TaqMan qRT-PCR of mature let- 7 normalized to 18S rRNA levels relative to WT. Standard error is depicted. *p value < 0.05.

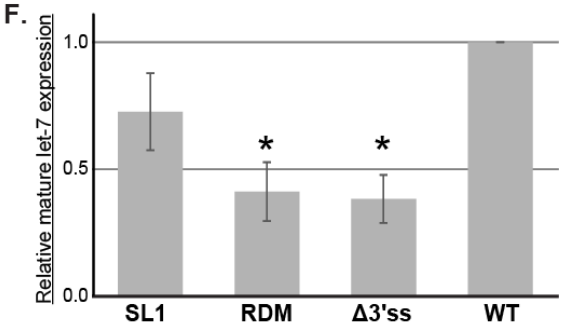

A.
let-7(mn112)
A
B
let-7
Chr
X
WT
3'ss
II
SL1
SL1
II
RDM
random
II
Δ3'ss
Δ3'ss
II
MSF
SL1
mut
II


B.
WT alae
patchy alae
WT
SL1
RDM
Δ3'ss
let-7(mg279)
WT alae
patchy alae
no alae


C.
SL1
RDM
Δ3'ss
WT
primary let-7 transcripts
18S rRNA


D.
Relative pri-let-7 levels
0
2
4
6
8
*
*
*
SL1
RDM
Δ3'ss
WT


E.
SL1
RDM
Δ3'ss
WT
mature let-7
U6


F.
Relative mature let-7 expression
0.0
0.5
1.0
*
*
SL1
RDM
Δ3'ss
WT

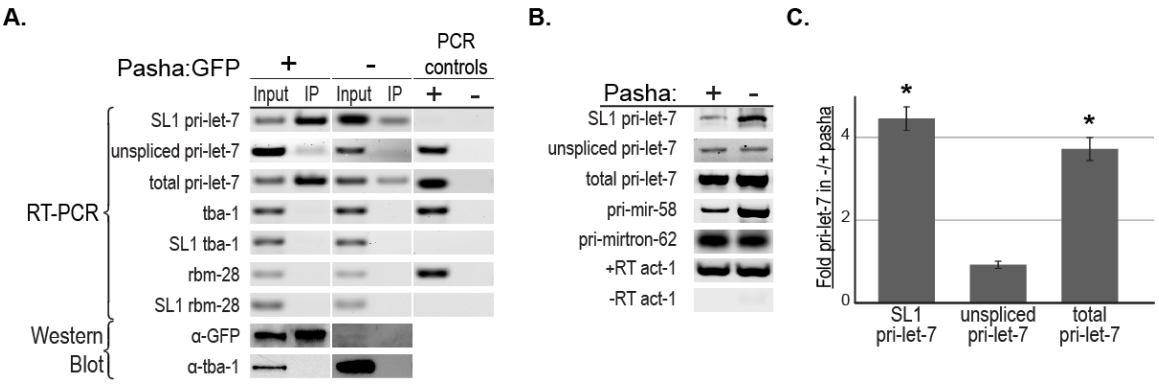
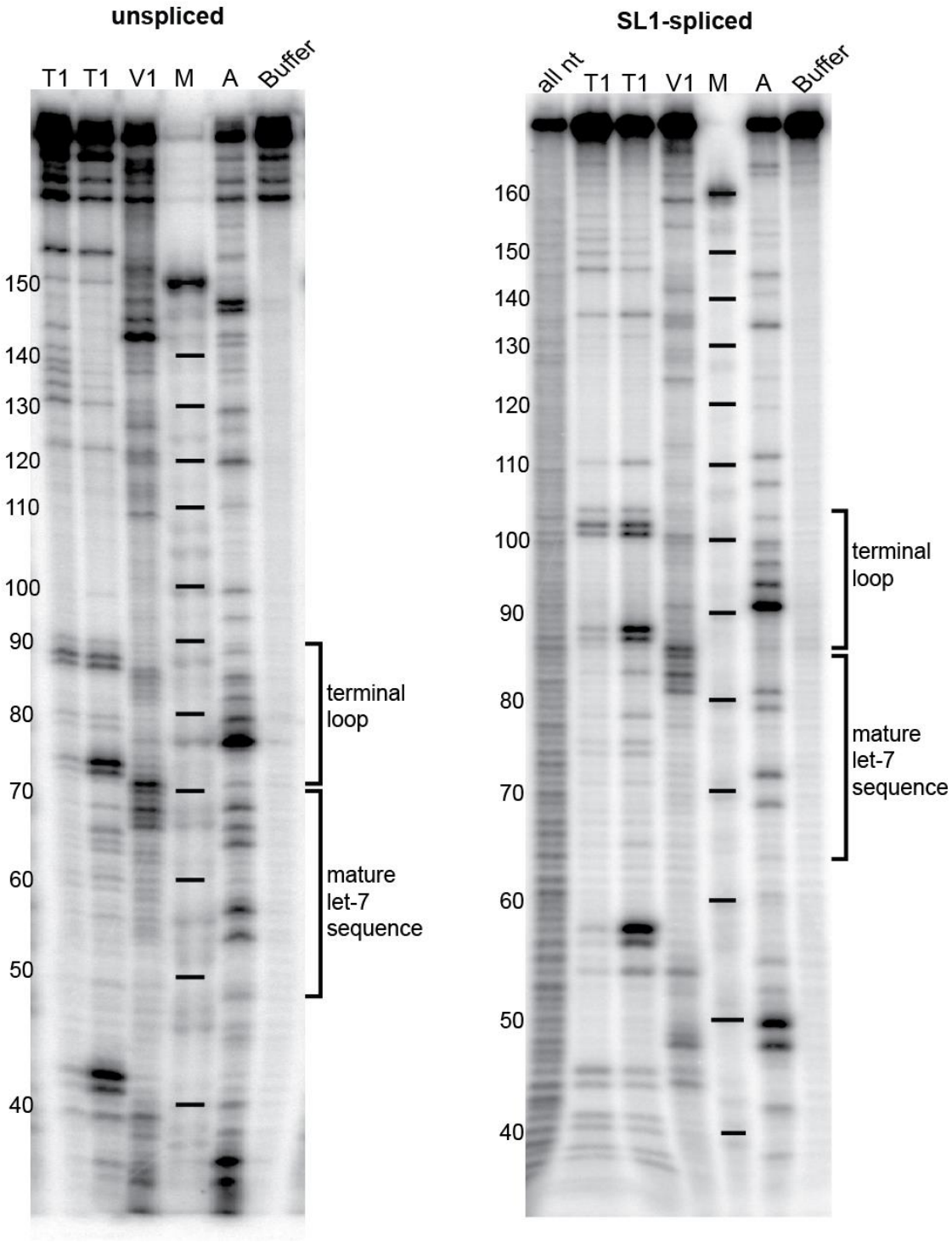


FIGURE 4.3. The Microprocessor prefers SL1-spliced pri-let-7 for processing *in vivo*. (A) Detection of transcripts associated with the Microprocessor *in vivo*. Extracts from L4 stage worm strains with (+) and without (-) the rescuing Pasha:GFP transgene were subjected to Immunoprecipitation (IP) using an anti-GFP antibody. Total (Input) and Immunoprecipitated (IP) RNAs were detected by RT-PCR followed by Agarose gel electrophoresis. PCR controls show the results using genomic DNA (+), which should not detect SL1-spliced sequences, or water (-) as templates. The bottom panels show the results of Western Blot analysis with a GFP antibody to monitor the IP of Pasha:GFP and tubulin (TBA-1) as a loading and specificity control. Results are representative of two independent IP experiments. (B) Total RNA from pasha mutant (-) and rescued (+) worms was collected from L4 staged animals and used for RT-PCR analysis of the indicated transcripts. Results were analyzed by agarose gel electrophoresis and represent three independent experiments. Actin (act-1) served both as a loading control and to control for genomic DNA contamination in –RT samples. Pri-miR-58 is a constitutively expressed miRNA subject to Pasha regulation. Pri-mirtron-62 is not part of the canonical miRNA biogenesis pathway. (C) Average fold increase of pri-let-7 isoforms in pasha mutants (-) relative to rescued (+) worms collected at L4 and analyzed by qRT-PCR. Each pri-let-7 isoform was first normalized to the control transcript Y45F10D.4. The error bars represent standard error. (*) $P<0.05$

SUPPLEMENTARY FIGURE 4.S1. RNase secondary structure assay of pri-let-7 isoforms. Polyacrylamide gel electrophoresis (PAGE) of T7 transcribed pri-let-7 isoforms (unspliced and SL1-spliced) that were 5’ end labeled, folded, and subject to treatment by RNase T1, V1, A, or buffer only. All nucleotide ladder (all nt) represents every base in the sequence, T1 indicates unpaired guanine bases, V1 indicates paired bases, and A indicates unpaired cytosine and uracil bases. Labeled RNA marker (M) serves as a reference and is a combination of labeled Decade Marker and all nucleotide ladder. Gels are representative of at least four independent experiments.


unspliced
T1 T1 V1 M A Buffer
150
140
130
120
110
100
90
80
70
60
50
40
terminal loop
mature let-7 sequence
SL1-spliced
all nt T1 T1 V1 M A Buffer
160
150
140
130
120
110
100
90
80
70
60
50
40
terminal loop
mature let-7 sequence

Chapter V

Characterizing clustered microRNAs in *Caenorhabditis elegans*

A. Abstract

MicroRNAs (miRNAs) are a class of small non-coding RNAs that regulate the expression of target messenger RNAs (mRNAs) post-transcriptionally. Investigations of miRNA regulation have revealed an intricate biogenesis pathway whereby longer primary miRNA (pri-miRNA) transcripts are processed by Drosha and Dicer to form mature miRNAs that serve as guides for the miRNA induced silencing complex (miRISC) to target mRNAs for degradation or deadenylation. When two or more miRNAs are transcribed in tandem on one primary transcript, this arrangement is known as a cluster. Deep sequencing and computational studies have predicted a number of miRNA clusters wherein often only one miRNA is dominantly expressed and this bias can change throughout development. However, the mechanisms controlling differential expression of clustered miRNAs are largely unknown. The goal of this work is to characterize clustered miRNAs in *Caenorhabditis elegans* (*C. elegans*) and to identify factors responsible for the regulation of differential expression of clustered miRNAs.

B. Introduction

Primary miRNAs (pri-miRNAs) are long non-coding RNAs transcribed by RNA Polymerase II and co-transcriptionally processed by the Microprocessor complex (Finnegan & Pasquinelli, 2013). Pri-miRNAs are capped and

polyadenylated and can range in size from a few hundred to a few thousand nucleotides. A standard pri-miRNA features a ~65 nt precursor miRNA (pre-miRNA), which is a hairpin-like stem loop that contains the mature miRNA sequence. The hairpin measures at about 3 helical turns, and features a ~10 nt terminal loop, with internal 2-4 nt bulges in the stem every ~11 nt or one helical turn, and flanking unstructured RNA sequences (known as the basal segment) (Lee et al., 2003; Denli et al., 2004; Han et al., 2006). Drosha cleaves around the internal bulges, on average ~11 nt away from the unpaired basal segment. Functional studies of the Microprocessor reveal that it works as a trimeric complex with one Drosha and two Pasha proteins (Nguyen et al., 2015). Drosha recognizes the basal segment at the base of the primary hairpin and Pasha serves as a guide to ensure fidelity in cleavage. Recent work suggests that processing of human pri-miRNAs also involves a sequence motif that is not detected in worms (Auyeung et al., 2013; Conrad et al., 2014).

Two or more miRNAs transcribed polycistronically on one primary transcript form a cluster. Clustered miRNAs are usually part of the same miRNA family. MiRNA families often share the same “seed” sequence, nucleotides 2-8, which is the main determinate of target interactions. While miRNA clusters are abundant, few studies have characterized their processing. One study suggests that the first miRNA transcribed will be expressed at higher levels than subsequent miRNAs due to co-transcriptional degradation of the latter by the exonuclease, XRN2 (Ballarino et al., 2009). While this may be true in some cases, there is evidence that this model is not universal.

A global analysis of human miRNAs found that of 326 genes, 148 are found in 51 clusters, some of which are differentially expressed (Yu et al., 2007). On the other hand, a more recent genome-wide analysis of primary miRNAs in humans and mice found that several annotated clustered miRNAs were actually transcribed from alternative promoters and independent transcriptional start sites (Chang et al., 2015). Indeed, this study revealed examples where potentially clustered miRNAs are actually transcribed independently to form distinct pri-miRNAs, and other examples where pri-miRNA transcripts produced subsets of clustered miRNAs. However, these are not the only explanations for differential expression of clustered miRNAs. Another recent study of the human miR-17-92 cluster found that during embryonic development the last miRNA in the cluster, miR-92, is steadily expressed, while levels of the other five members of the cluster increase and peak after three days (Du et al., 2015). They found that the pri-miR-17-92 transcript undergoes a splicing step to produce a “progenitor” miRNA or “pro-miRNA” which is a more favorable substrate for Drosha processing and allows for the expression of all the miRNAs within the cluster. This is consistent with current literature regarding the structural requirements of processing by the Microprocessor as well as the results in Chapter IV showing that splicing can act as a novel regulator of primary miRNA biogenesis(Mondol et al., 2015).

Considering the widespread existence of clustered miRNAs and the indications of post-transcriptional mechanisms to produce different levels of such co-transcribed miRNAs, my studies to elucidate how the expression of clustered

miRNAs is regulated in *C. elegans* will likely provide new insights for miRNA biogenesis across species.

C. Results

C.1 Identification of clustered pri-miRNAs in *C. elegans*

To search for potential clustered miRNAs, I first needed to set a definition for miRNAs predicted to reside in a cluster as two or more miRNAs residing within 2.1 kilobases (kb), because primary miRNAs are capable of being that size (Bracht et al., 2004; Bracht et al., 2010; Massirer et al., 2012). Of the 207 *C. elegans* miRNAs annotated in miRBase, I found that 53 are in 19 potential clusters (Kozomara & Griffiths-Jones, 2014). I then examined the number of "reads" from deep sequencing data from three different studies, to determine expression patterns of these miRNAs (Kato et al., 2009; Zisoulis et al., 2010; Warf et al., 2011). Generally, the number of "reads" correlates with miRNA abundance. In all cases, one miRNA from each cluster has 2 to 100 times more reads than the others have. I carefully looked at the expression patterns and structural features of each hairpin and narrowed down my analysis to six intergenic clusters that appeared to be post-transcriptionally regulated. In three of these clusters, the dominant miRNA (in terms of expression) is not the first one in the transcript. In the other three clusters, the most abundant miRNA exhibit changes in expression during development (TABLE 5.1). Of the six clusters, three have <400 nt between the individual miRNAs within them increasing the likelihood that they are expressed from one promoter. Careful inspection of *C.*

elegans RNA Polymerase II (Pol II) Chromatin Immunoprecipitation (ChIP) data available through Wormbase as part of the modENCODE project shows Pol II occupancy along the loci for all clustered miRNAs of interest, supporting the likelihood that they are co-transcribed as clusters (Harris et al., 2014).

I validated that the potential clustered pri-miRNAs are expressed and processed as part of the miRNA pathway by measuring expression in wild-type worms compared to worms lacking a functional Microprocessor. I first performed RT-PCR analysis of wild-type worms collected over two-hour intervals using oligos that flank the first and last miRNA in the cluster. This allowed me to identify time points of interest where the potential miRNA clusters can be studied in early (larval stage 1 or L1), mid (larval stage 2 or L2), and late (larval stage 4 or L4) larval stages in the lifecycle of *C. elegans* (TABLE 5.2). I next looked at whether the clustered pri-miRNA transcripts were processed *in vivo* at early, mid, and late developmental stages by depleting functional PASH-1 and comparing against expression in wild-type worms (FIG 5.1). Early and mid-larval stages represent two biological repeats where wild-type N2 worms were treated with empty vector or Pash-1 RNAi. Late larval stage results represent three biological repeats using temperature sensitive *pash-1(mj100*) worms that exhibit embryonic lethality when grown at 25°C but develop normally at 15°C. Worms were collected at the developmental time of interest, total RNA extracted, cDNA prepared, and analyzed by RT-PCR and Agarose gel electrophoresis. From these experiments, five of the six proposed potential miRNA clusters were

detectable by RT-PCR and were processed in at least one of the developmental stages measured (FIG 5.1).

For evidence of binding to the Microprocessor *in vivo*, I worked exhaustively to optimize RNA immunoprecipitations (RIP) of tagged Microprocessor proteins, DRSH-1 and PASH-1 (TABLE 4.1). After variable immunoprecipitation (IP) results using a FLAG:DRSH-1 transgenic worm for RIPs, I constructed another transgenic line that ubiquitously expresses a single insertion of 3XFLAG:GFP:DRSH-1 and rescues the sterility of a *drsh-1(ok369)* null mutant. While I was able to fully optimize the protein IP, I could not detect any pri-miRNAs in the pull-down. I also tried IPs with the temperature sensitive *pash-1(mj100*) worms containing an extrachromosomal transgene driving ubiquitous expression of a PASH-1:GFP fusion protein which is able to rescue the temperature sensitive allele. Again, I was able to fully optimize the protein IP, however after many attempts, I was only able to recover partially cleaved pri-let-7, where the first Drosha cleavage at the 3' end of the precursor miRNA has already occurred. Indeed, processing is co-transcriptional, which is why it is difficult to detect most pri-miRNA transcripts at stages where mature expression is high. I believe that I was able to only pull down pri-let-7 because it is a very abundant pri-miRNA. While this outcome was somewhat of a setback, the transgenic worms can still be used for *in vitro* assays in future experiments to test how mutations in the primary transcript can effect processing of clustered miRNAs.

The Pasquinelli lab has confirmed that at least one of the predicted clusters expresses co-transcribed miRNAs that significantly differ in mature levels. We found that the let-7 “sisters” miR-241 and miR-48 are embedded in one ~3 kb primary transcript through 5’ and 3’ rapid-amplification of cDNA ends (RACE) mapping (FIG 5.2). This is in contrast to an original report suggesting that these miRNAs are transcribed individually (Li et al., 2005). Through Northern blotting, we have confirmed that pri-miR-241-48 is expressed (data not shown), and that the transcript is a substrate for processing by the Microprocessor through RT-PCR (FIG 5.1).

Agarose Northern blot analysis was performed to determine whether the predicted miRNA clusters are actually transcribed as one unit. Probing for primary transcripts can prove challenging as they are usually observable before the mature stage and drop below detectable expression levels as processing progresses. However, by examining a developmental time course via RT-PCR, I was able to pinpoint developmental time points where the pri-miRNAs should be detectable (TABLE 5.2). In addition, by including Pasha-depleted RNA on my Northern blot, I could confirm that the transcript detected is an actual primary miRNA, as it should accumulate in the mutants. I set up two Agarose Northern Blots, one at the L4 time point, one at gravid adult time point. Each experiment contained total RNA from three biological replicates of wild-type vs. *pash-1(mj100)* worms, as well as RNA markers to be able to measure detected transcripts, and DNA positive controls for the five clustered pri-miRNAs to validate that the Northern probes are working correctly. Unfortunately, probes for

pri-miR-61-250 and pri-miR-53-51 failed to detect the positive control. This experiment should be repeated with fresh radioactivity. Once the Agarose Northern Blots are complete, they can serve to guide the RACE-mapping experiments to give defined start and polyadenylation sites for each clustered pri-miRNA

C.2 Confirming differential expression of miRNAs in clusters

High-throughput RNA deep-sequencing experiments from several labs, including ours, show differential expression of miRNAs found in clusters (TABLE 5.1) (Kato et al., 2009; Zisoulis et al., 2010; Warf et al., 2011). To confirm and analyze mature clustered miRNA expression levels, I performed PAGE Northern blot analysis using oligo standards for quantification. In addition, I worked to optimize a miRNA multiplex assay by Firefly BioWorks for use in *C. elegans.* Ultimately, we would like to be able to quantify all of the differentially expressed clustered miRNAs and their miRNA* strands from 1 µg of total RNA, quickly and easily with the precision of other less high-throughput friendly assays like TaqMan or Northern Blot analysis. With such a small volume of RNA required, one could setup a multi-well RNAi screen of RNA-binding proteins to look for factors that determine the differential expression of miRNAs in a cluster.

C.2.a Optimizing Firefly BioWorks multiplex cellular assay for miRNA detection

The multiplex cellular miRNA assay works without reverse transcription, library preparation, or amplification. Instead, barcoded hydrogel microparticles in each well of a 96-well plate are used to detect up to 68 different miRNAs and

small RNA controls. Total RNA is added to a well and allowed to hybridize to the DNA particle probes, a fluorescent universal label is ligated to each miRNA, and a fluorescent reporter added for detection via a benchtop flow cytometer.

I worked to determine whether whole worms would be sufficient for accurate detection instead of purified RNA. Worms were treated with empty vector, ALG-1, DCR-1, LIN-42, PASH-1, or XRN-2 RNAi, and grown to the gravid adult stage in three biological replicates. These genes were chosen because they are known factors in the miRNA biogenesis pathway and we wanted to ensure that the multiplex assay was sensitive enough to detect changes in RNAi treated worms. Compared to empty vector, worms treated with ALG-1, DCR-1, or PASH-1 RNAi should have lower miRNA expression, while worms treated with LIN-42 or XRN-2 RNAi should exhibit miRNA accumulation.

Over several experimental attempts, I tested 1, 5, 10, 50, ~500, ~1000, ~5000 whole worms treated with six different RNAi conditions and their corresponding RNA. I extracted total RNA and sent equal concentrations of each condition to be assayed against 25 miRNAs and their star-strands, for 50 miRNA probes in total. I also collected a sample from the RNAi treated worms from each biological replicate, in order to perform TaqMan qRT-PCR and compare the results of mature let-7 expression to the multiplex assay. Purified RNA was more efficient at miRNA detection than whole worms. Of the 50 probes, 22 gave a detectable signal above background in all RNAi-treated worm samples. The let-7 miRNA was also detectable in each sample at similar levels to those measured by TaqMan qRT-PCR, making optimizing the multiplex assay more promising.

Many miRNAs isoforms generally vary at the 3’ end, where the miRNA is trimmed by one or more nucleotides (Ameres & Zamore, 2013). The Firefly protocol generally ligates the fluorescent reporter to the 3’ end of the miRNA, suggesting that various isoforms would not be detected. To improve miRNA detection, I redesigned each probe such that the ligation would instead occur at the less variable 5’ end of the miRNA. I used miRBase *C. elegans* deep sequencing data to choose the probe sequence that corresponds to the most abundant isoform of the miRNA (Kozomara & Griffiths-Jones, 2014). I also reduced the number of miRNAs probed to only the clusters of interest and included at least three Drosha-independent, small non-coding RNA controls (Ambros et al., 2003). In addition, I optimized the amount of RNA to use by testing seven concentrations (150 ng, 300 ng, 600 ng, 1.25 µg, 2.5 µg, 5 µg, and 10 µg) in three biological replicates of L4-grown wild-type and *pash-1(mj100)* worms grown at 25°C.

These optimizations greatly improved the output from the Firefly multiplex assay. All N2 wild-type concentrations tested were able to detect levels of each miRNA and miRNA* probed. However, at least 1 µg of RNA was needed to detect discernable miRNA* in *pash-1(mj100)* mutant worms. When compared to TaqMan qRT-PCR results of let-7 expression, 5 µg of total RNA (or a small pellet of ~200 adult worms) had the most correlative results (FIG 5.3). Overall, the Firefly assay confirmed that in most cases, the first miRNA in the cluster is not always the most highly expressed (FIG 5.4). While the miR-53-51 cluster seems to be the exception to that rule at the L4 stage, Northern Blot measurements at

different time points suggests that miR-51 peaks at early developmental stages and decreases over time.

C.2.b. Mature miRNA expression over development measured by Northern Blot analysis

To further confirm the validity of the Firefly BioWorks assay, and to ensure that the clustered miRNAs are expressed differentially, I carried out several PAGE Northern Blots, testing Embryo, L1, L2, L4, and gravid adult stages of the worm. In addition, I included an oligo DNA standard to control for proper hybridization of the Northern probe, and to be able to accurately quantify and compare expression of the various clustered miRNAs tested. To not over-strip the Northern membrane between probes, RNA samples were split such that Figure 5.5 is representative of four different Northern Blots that tested two biological replicates of Embryo, L2, and L4 staged wild-type worms. The oligo standards I used were not PAGE purified, however, and for some probes I was able to detect a larger-sized sequence in my miRNA control wells (not shown).

As expected given the multiplex assay results, miR-53 is more abundantly expressed than miR-51 at the L4 stage, while the opposite is true in worm embryos (FIG 5.5). Mature miR-48 is expressed more than miR-241, and miR-61 is expressed more than miR-250. In addition, several of the miRNAs were not expressed in sufficient amounts for detection, though miR-44 and miR-56 remain to be tested. Also not represented are the results from two biological replicates of L1, L2, L4, and gravid adult collected worms, split to three membranes. So far, the three membranes have been probed with miR-55, miR-250, and miR-43.

Those results indicate that miR-55 peaks at L1 and then stabilizes, while the miR-250 and miR-43 probes were unable to detect any miRNAs.

Overall, the inability to detect some miRNAs by Northern Blot analysis is consistent with the expression levels detected by the multiplex assay, which is very low compared to other robustly expressed miRNAs like let-7 and lin-4 (FIG 5.4). Future attempts should include XRN-1 and XRN-2 RNAi-treated worms. XRN-1 and XRN-2 are exoribonucleases that have known roles in mature miRNA stability, therefore knockdown of these factors should help increase the amount of mature miRNA detected (Grosshans & Chatterjee, 2010; Chatterjee et al., 2011). Precursor miRNAs were detectable when probing for miR-53, mir-51, miR-48, miR-61, miR-42, and miR-55 (not shown). In addition, blots probed with miR-61, miR-42, miR-54, and miR-55 detected an RNA transcript larger than the precursor, which could possibly be the primary transcript (not shown).

C.3 Identifying factors that affect regulation

The stepwise nature of miRNA biogenesis presents multiple possible mechanisms that could result in different levels of mature miRNAs that originated from a cluster (FIG 5.6). Differences in mature miRNA accumulation suggest diminished stability at the mature or precursor level (FIG 5.6A&D), preferential processing of pre-miRNAs to mature by Dicer (FIG 5.6B), or uneven processing of primary to precursor by Drosha (FIG 5.6C)

To discriminate between the various possible mechanisms, one can analyze pre-miRNA levels using PAGE Northern Blot analysis. Given the differences in mature miRNA expression, equal pre-miRNA levels suggest that

an unknown factor is preferentially degrading one or stabilizing the other mature miRNA (Fig. 5.6A). Unfortunately, not all of the clusters had detectable precursors. To get around this, one can probe for pre-miRNAs using Dicer RNAi-treated worms. This treatment blocks pre-miRNA processing and allows accumulation. If Dicer is preferentially processing one pre-miRNA over the other, knocking it down should equalize the uneven levels. Alternatively, if one pre-miRNA continues to accumulate in Dicer-knockdown worms, this alludes to either unequal primary processing by Drosha (FIG 5.6C) or preferential stabilization or degradation of one pre-miRNA over the other (FIG 5.6D). To test these possibilities further, one can perform *in vitro* Dicer processing reactions using the two different precursors and worm extracts. This method has been successfully used to study precursor processing in our lab and in other published work (Chatterjee & Grosshans, 2009). Together, these assays can determine whether Dicer preferentially processes one pre-miRNA over another (FIG 5.6B).

As my results in Chapter IV have shown, pri-miRNA remodeling can serve to regulate processing. I therefore asked whether secondary structures of the clustered pri-miRNAs played a role in preferential Drosha processing by performing a preliminary scan of predicted secondary structures using lowest free energy thermodynamics on the mfold server (Zuker, 2003). This analysis is still preliminary because I have not completed the RACE experiments, nor do I have Northern Blot data that define the 5’ and 3’ ends of the pri-miRNA clusters. Instead, I included 200 – 300 nucleotides around the miRNA cluster in the mfold analysis. From this initial survey, three pri-miRNA clusters, pri-miR-53-51,

pri-miR-54-56, and pri-miR-241-48, revealed at least one potential structural arrangement that buried one mature miRNA sequence in a way that it was no longer in the canonical hairpin required for Microprocessor cleavage.

To test the possibility of pri-miRNA processing preferences, I have optimized IP of a 3XFLAG:GFP:Drosha transgenic worm. These worms express tagged-Drosha ubiquitously, and rescue the *drsh-1(ok369)* null allele. Drosha cleavage assays have been successful using IP extracts of FLAG-tagged Drosha enzyme overexpressed in human cell lines (Lee et al., 2007). Thus, we expect to be able to introduce this method to *C. elegans* to mimic primary to precursor processing *in vitro*. If there is a structural reason for Drosha preferential processing, it should be detected by this assay. One can also use RACE to examine Drosha cleavage products *in vivo* and see whether processing is occurring for both miRNAs at the corresponding expression points, as we have previously done in the analysis of lin-4 miRNA processing (Bracht et al., 2010). Furthermore, one can examine clustered miRNA expression *in vivo* using CRISPR-Cas9 mutant worms that disrupt as well as reestablish structure through compensatory mutations. If the results of these assays demonstrate that similar levels of pre-miRNAs are produced *in vitro*, but Dicer-knockdown results in unequal levels *in vivo*, it suggests that one of the pre-miRNAs is targeted for degradation by some unknown factor absent from a Drosha IP (FIG 5.6D).

A thorough analysis of the sequences in and around the primary transcript of clustered miRNAs can serve to identify regions that may be responsible for favorable miRNA processing because regulatory elements are often highly

conserved. For example, several *cis* regulatory elements discovered in pri-let-7 are conserved among nematode species, including the 3' splice site, polyadenylation site, let-7 complementary site, and let-7 transcription element (Bracht et al., 2004; Kai et al., 2012; Zisoulis et al., 2012). Using targeted CRISPR-Cas9 mutagenesis, one can create mutations in conserved regions around miRNA clusters and test for effects on miRNA biogenesis. Mature and precursor miRNA levels produced by these mutant worms can then be analyzed by multiplex miRNA assay, Northern blot analysis, and RT-PCR. If primary to precursor processing is the point of regulation, one can switch the precursor stem-loop positions to examine whether local context is responsible for the regulation. If this swap does not alter biogenesis, then it implies that something intrinsic to the hairpins controls processing and thus narrows the search for *cis* elements. If precursor to mature processing is the point of regulation, we could swap the sequences that make up the loop of the hairpin, for example, to test how important they are for Dicer processing.

Finally, the multiplex miRNA assay can be used to measure the effects of differential expression of all five miRNA clusters under RNAi conditions in a high-throughput manner. One can conduct an RNAi screen using a library of RNA-binding proteins, previously used in our lab to discover the role of RBM-28 in lin-4 expression (Bracht et al., 2010). One can prioritize the RNAi candidates based on prior evidence for the factor in the miRNA pathway and then based on broad conservation. *C. elegans* has proven to be an excellent model system for this, as RNAi knockdowns can be done simply through feeding the worms bacteria that

express a double-stranded RNA corresponding to the gene one wishes to downregulate. Because the multiplex assay requires very little RNA, one can use multi-well plates for the RNAi knockdown, facilitating the screen. In addition, because the multiplex assay is done on a 96 well plate, we can screen through at least 30 different proteins in triplicate at any time. When screening through an RNAi library, one would expect to find proteins whose downregulation cause differences in the endogenous expression ratio of a miRNA cluster. For example, RNAi of a positive regulator of miR-48 will cause its expression to decrease. A negative regulator of miR-241 under RNAi will cause an increase in expression. If both miR-48 levels fall and miR-241 levels rise under RNAi conditions, this may indicate a factor that can negatively and positively regulate miR-48 and miR-241, respectively. A caveat to this approach is that it assumes that annotated RNA binding proteins will be involved. Over 800 RNA-binding proteins have been cataloged in *C. elegans*, however, this relatively high throughput screen can be expanded to include other types of candidates if the initial candidates do not show effects (Tamburino et al., 2013).

D. Experimental Procedures

Several methods including worm staging, MosSCI integration, RNA extraction, RNAi treatment, Western blot analysis, Agarose and PAGE Northern Blot analysis, and reverse transcriptase PCR assays have been previously described (Van Wynsberghe et al., 2011; Zisoulis et al., 2012). Wild-type worms were N2 Bristol. *pash-1(mj100)* worms were obtained from the Miska lab.

3XFLAG:GFP:Drosha was constructed using USER cloning (Nour-Eldin et al., 2010) to stitch together several PCR products into a MosSCI compatible vector: the let-858 constitutive promoter, the 3XFLAG:GFP tag, and Drosha gene including introns and 3'UTR. Sequences used for the optimized multiplex assay probes are listed in TABLE 5.3, while oligos used in this study is found listed in TABLE 5.4. In all cases, data from at least three independent experiments are used for quantification and appropriate statistical methods employed to determine significance.

D.1 *In vitro* transcription of radiolabeled pri-miRNAs

Mix 1 microliter (µL) template DNA (30-100 ng), 4 µL 5X T7 Buffer (Promega), 2 µL 10 mM rDTP + 1 mM CTP, 0.5 µL RNasin Plus (Promega), 1.5 µL α-^{32}P-CTP (20mCi/ml) (Perkin Elmer), 1 µL T7 Polymerase (Promega), and 10 µL DEPC Water. Incubate reaction 37°C for 3 hours (hrs) – overnight. Add 1 µL DNase and incubate at 37°C for an additional 20-60 minutes. Bring volume up to 50 µL with DEPC Water and clear unincorporated nucleotides using a GE G50 spin column. Isolate and purify RNA through 5% TBE-UREA polyacrylamide gel electrophoresis (PAGE) by cutting out the band of interest, crushing it and rocking in 500 µL 0.3 M NaCl at 4°C overnight. Add sludge to CoStar Spin-X columns containing a 0.45 µM filter, isolate the supernatant, alcohol precipitate, and elute to 20 µL.

D.2 *In vitro* Drosha processing

Prepare Lysis Buffer (50 mM HEPES pH 7.5, 100 mM KCl, 1 mM $MgCl_2$, 1mM EGTA, 10% Glycerol, 0.05% NP-40, 1 mM DTT, protease inhibitor cocktail)

and Wash Buffer (20 mM Tris-HCl pH 8.0, 100 mM KCl, 0.2 mM EDTA) on ice, and chill glass tissue homogenizer. Collect 12 gravid adult plates of Drosha FLAG-tagged worms. Egg-prep via standard bleach treatment and transfer eggs to low-adherence microcentrifuge tubes using 1 mL Lysis Buffer. Resuspend egg pellet in equal volume Lysis Buffer, transfer to tissue homogenizer and dounce 60 times on ice. Use Lysis Buffer to rinse homogenizer and bring egg-extract volume to 1 mL. Clear lysate by spinning top-speed for ten minutes in a 4°C centrifuge. Take 20 µL Anti-FLAG M2 Affinity Gel resin (Sigma), spin at 4°C, 8200xg, 30 seconds, rest 2 minutes, and then remove supernatant. Wash resin with 0.5 mL cold Lysis Buffer twice, spinning as described. Take 50 µL of egg lysate for protein prep for Western (add 50 µL 2x SDS Sample Buffer). Take 50 µL of worm lysate for processing reaction (leave on ice). Transfer remaining worm lysate to washed beads. Rock at 4°C for 1-2 hours. Centrifuge 8200g, 4°C, 30”, rest 2’ on ice. Take 50 µL of supernatant for protein prep (add 50 µL 2x SDS Sample Buffer), discard the rest. Wash beads w/ 1ml Wash Buffer, invert tube 6-7X, Spin 8200g, 4°C, 30”, rest 2’ on ice. Repeat washes 4 more times - Change tube for last wash. Add 500 µL Wash Buffer to beads, take 50 µL of slurry for protein prep (add 50 µL 2x SDS Sample Buffer). Spin 8200g, 4°C, 30”, rest 2’ on ice, discard sup leaving ~40 µL total beads & buffer – Leave on ice while RNA is prepared. Mix 9 µL labelled pri-miRNA, 9 µL 64 mM $MgCl_2$, 2.2 µL RNasin Plus (Promega) and bring volume to 45 µL with nuclease-free water. Aliquot to three tubes of 15 µL, which will be mixed with either 15 µL FLAG-IP beads, egg lysate, or lysis buffer. Incubate at 37° 90 minutes, raise volume to 200 µL, phenol-

chloroform extract, and ethanol precipitate. Run a Western Blot to test IP using input, bead, and supernatant samples collected during the experiment. Run 15 µL eluted processed RNA on a thin 20x20 cm PAGE 8% UREAgel. Dry gel with 60°C heat for 1 hour and expose to film or phosphorscreen.

E. Discussion

Considering the widespread role of miRNAs in gene regulation, and the abundance and importance of many clustered miRNAs, understanding the mechanisms controlling their differential expression will further studies attempting to utilize miRNAs as therapeutics for disease. It is clear from the data I have gathered so far, that these clustered miRNAs are not only differentially expressed, but some seem to also be regulated through development in some way. For example, in the miR-53-51 cluster, miR-51 is expressed at much higher rates during the embryo stage than miR-53. While miR-53 levels remain somewhat steady during development, levels of miR-51 drop, such that by the L4 stage miR-53 is expressed higher than miR-51. This was seen both by Northern Blot and by multiplex miRNA assay. The timing and expression levels of miRNAs are very important. That two miRNAs coming from one primary transcript can have such different levels through development speaks to a highly coordinated and regulated system.

Secondary structure rearrangements due to splicing may be one way some of these clustered pri-miRNAs are regulated, but it may not be for others. This is why I propose a multifaceted approach to look at structure and other *cis*

elements as well as an RNAi screen to search for *trans*-factors. However, I recognize that positive hits are not always immediately obvious from a screen, which might limit my ability to characterize the *trans* factors for functions in regulating clustered miRNA expression. In addition, some of these miRNA clusters may prove more amenable to study than others.

Overall, given the multiple ways with which I planned to search for regulators, this project is very likely to yield fruitful results. The Pasquinelli lab has extensive experience of discovering primary transcripts, describing regulation of miRNA expression during biogenesis, and characterizing *cis* elements (Bracht et al., 2010; Kai et al., 2012; Massirer et al., 2012; Mondol et al., 2015). In addition the lab has been successful in discovering RBPs and mechanisms that regulate miRNA biogenesis (Bracht et al., 2004; Bracht et al., 2010; Van Wynsberghe et al., 2011; Kai et al., ; Massirer et al., 2012; Zisoulis et al., 2012). Hence, I am proposing methods to identify both *cis* and *trans* factors so that one can gain insights towards differential regulation of miRNA expression either through methodical analysis of miRNA sequences and structures and/ or through more open-ended screening methods.

F. Acknowledgments

I thank undergraduate helpers A. Weaver, K. Nassiri, and J. Olivas for helping with RNAi knockdown of Microprocessor proteins, Western Blots, RNA extractions, and RT-PCRs. J. Olivas also helped generate the 3XFLAG:GFP:Drosha-1 transgenic worm. My research was supported in by an

NIH Ruth L. Kirschstein National Research Service Award Individual Predoctoral Fellowship to Promote Diversity in Health-Related Research (F3194041A). I would also like to thank Jessica Tytell at Firefly Bioworks for use of the multiplex miRNA assay.

TABLE 5.1 Predicted clustered pri-miRNAs**.** Temporal patterns and minimum size for clustered miRNAs are indicated. Stage of peak expression and fold increase are from deep sequencing experiments (Kato et al., 2009; Zisoulis et al., 2010; Warf et al., 2011)

temporal expression	cluster (5' - 3')	Number of nucleotides between miRNAs	larval stage of peak expression	~ fold increase compared to least abundant miRNA
last miRNA consistantly expressed more than the other	miR-241	1871	L4	10 - 400X
	miR-48			
	miR-238	2155	L4	6 - 9X
	miR-80/227			
	miR-61	239	L3	2 - 10X
	miR-250			
most abundant miRNA changes during larval development	miR-53	1678	Embryo	2X
	miR-51		L1	4X
	miR-54	403	Embryo	5 - 4,000X
	miR-55		L1	2 - 4X
	miR-56			
	miR-42	307	Embryo	30 - 2,000X
	miR-43			
	miR-44		L4	80 - 10,000X

TABLE 5.2 Temporal expression of validated clustered pri-miRNAs. Measured by a 2-hour developmental RT-PCR analysis in wild-type worms.

pri-miRNA clusters	**nucleotides between first and last clustered miRNAs**	**pri-miRNA temporal expression pattern**
pri-miR-61-250	239	ubiquitously expressed
pri-miR-42-44	307	L1, L2, L4
pri-miR-54-56	403	L1, L2, L4, young adult
pri-miR-53-51	1678	early L1
pri-miR-241-48	1871	oscillates every six hours
pri-miR-238-80	2155	not detectable by RT-PCR

TABLE 5.3 Multiplex miRNA assay probe sequences

Probe Name	Sequence
miR-241	UGAGGUAGGUGCGAGAAAUGA
miR-241*	AUUGUCUCUCAGCUGCUUCAUC
miR-48	UGAGGUAGGCUCAGUAGAUGCGA
miR-48*	ACAUCCACCAGCCUAGCUCGCA
miR-61*	UGGGUUACGGGGCUUAGUCCUU
miR-61	UGACUAGAACCGUUACUCAUC
miR-250*	CCUUCAGUUGCCUCGUGAUCCG
miR-250	AAUCACAGUCAACUGUUGGCA
miR-53	CACCCGUACAUUUGUUUCCGUGCU
miR-53*	CACGGCACAAUAUAUGGGUCGC
miR-51	UACCCGUAGCUCCUAUCCAUGUU
miR-51*	CAUGGAAGCAGGUACAGGUGCA
miR-54	UACCCGUAAUCUUCAUAAUCCGAG
miR-54*	AGGAUAUGAGACGACGAGAACA
miR-55	UACCCGUAUAAGUUUCUGCUGAG
miR-55*	CGGCAGAAACCUAUCGGUUAUA
miR-56	UACCCGUAAUGUUUCCGCU
miR-56*	UGGCGGAUCCAUUUUGGGUUGUACC
miR-42	UCACCGGGUUAACAUCUACAGA
miR-42*	GUGGGUGUUUGCUUUUUCGGUGAAG
miR-43	UAUCACAGUUUACUUGCUGUCGC
miR-43*	GACAUCAAGAAACUAGUGAUUAUG
miR-44	UGACUAGAGACACAUUCAGCU
miR-44*	CUGGAUGUGCUCGUUGGUCAUA
let-7	UGAGGUAGUAGGUUGUAUAGUU
let-7*	CUAUGCAAUUUUCUACCUUACC
lin-4	UCCCUGAGACCUCAAGUGUGA
lin-4*	ACACCUGGGCUCUCCGGGUACC
mirtron-62	UGAUAUGUAAUCUAGCUUACAG
mirtron-1018	AGAGAGAUCAUUGGACUUACAG
mirtron-1020	AUUAUUCUGUGACACUUUCAG
mirtron-1020*	GUAAGUGUUACAGAAUAAUCUU
tncR7	GACAACCAUUCCGUAGGCUG
tncR22	GACAAUCAGCCGAUCAUAGUC
U18 snoRNA	TGGCAGTGATGATCACAAATCCGTGTTTCTGACAAGCGAT TGACGATAGAAAACCGGCTGAGCCA

TABLE 5.4 Oligos used in this study

miRNA	Oligo Number	Sequence (5' -> 3')	Notes
miR-241	A1134	TCATTTCTCGCACCTACCTCA/3StarFire/	Oligo used for starfire Northern probes
miR-241	A2796	TGAGGTAGGTGCGAGAAATGA	Oligo used for DNA standard
miR-48	A1132	TCGCATCTACTGAGCCTACCTCA/3StarFire/	Oligo used for starfire Northern probes
miR-48	A2795	TGAGGTAGGCTCAGTAGATGCGA	Oligo used for DNA standard
pri-miR-241 F	A2417	GACACCATCCTTCCGCTTGTTGT	forward before miR-241
pri-miR-241 F	A3526	CCTTATTGAAGGTAAGTGTTGG	forward before miR-241
pri-miR-48 R	A2422	GTACCTTGAAACTTGGCTTTAACCTGA	reverse miR-48
pri-miR-48 R	A3527	GTAGCACCACGTTATTGAATG	reverse miR-48
miR-61	A2800	TGACTAGAACCGTTACTCATC	Oligo used for DNA standard
miR-61	A2812	GATGAGTAACGGTTCTAGTCA/3StarFire/	Oligo used for starfire Northern probes
miR-250	A2799	AATCACAGTCAACTGTTGGC	Oligo used for DNA standard
miR-250	A2811	TGCCAACAGTTGACTGTGATT/3StarFire/	Oligo used for starfire Northern probes
pri-miR-61 F	A2981	CTGTCCCCTTATTCTGCTAC	forward before miR-61
pri-miR-61 F	A3571	GCCACTTCGCCATTCTCCT	forward pri-miR-61-250 (414nt with A3570)
pri-miR-61 F	A3572	CGTCTCTTCACATCACTGCG	forward pri-miR-61-250 (498nt with A3570)
pri-miR-250 R	A2982	GAAATCGTCTTTGAAACTTTG	reverse after miR-250
pri-miR-250 R	A3570	CCACTTTCTCTCAAGTCCAATCC	reverse pri-miR-61-250
miR-53	A2802	CACCCGTACATTTGTTTCCGTGCT	Oligo used for DNA standard
miR-53	A2814	AGCACGGAAACAAATGTACGGGTG/3StarFire/	Oligo used for starfire Northern probes
miR-51	A2801	TACCCGTAGCTCCTATCCATGTT	Oligo used for DNA standard
miR-51	A2813	AACATGGATAGGAGCTACGGGTA/3StarFire/	Oligo used for starfire Northern probes
pri-miR-53 F	A2977	ACAACTCGACATCATTCGTC	forward before miR-53
pri-miR-53 F	A3577	CTACGGCTGTCTCGGTTTCTC	forward pri-miR-53-51 (1850nt with A3578)
pri-miR-53 R	A3580	CTATAACAACGGTGAAGCCGTT	reverse pri-miR-53 (435nt with A3577)
pri-miR-51 F	A3579	CTCAACCGCAGTGAATCTAGAG	forward pri-miR-51 (479nt with A3578)
pri-miR-51 R	A2978	TCTGCATGAAGAATTGAAAA	reverse after miR-51
pri-miR-51 R	A3578	AATTCTGAACCACCCACGC	reverse pri-miR-53-51 (1850nt with A3577)
miR-54	A2803	AGGATATGAGACGACGAGAACA	Oligo used for DNA standard
miR-54	A2815	CTCGGATTATGAAGATTACGGGTA/3StarFire/	Oligo used for starfire Northern probes
miR-55	A2804	TACCCGTATAAGTTTCTGCTGAG	Oligo used for DNA standard
miR-55	A2816	CTCAGCAGAAACTTATACGGGTA/3StarFire/	Oligo used for starfire Northern probes
miR-56	A2805	TACCCGTAATGTTTCCGCTGAG	Oligo used for DNA standard
miR-56	A2817	CTCAGCGGAAACATTACGGGTA/3StarFire/	Oligo used for starfire Northern probes
pri-miR-54 F	A2983	TCTCTACCCCTCTTGTTCTG	forward before miR-54
pri-miR-54 F	A3573	CTCCGCTTCGTCCTATGCC	forward pri-miR-54-56 (566nt with A3574)
pri-miR-56 R	A2984	GGTATCCAATGGGCAAAG	reverse after miR-56
pri-miR-56 R	A3574	ACTGCCAAGATATCCTCTGGG	reverse pri--miR-54-56 (566nt with A3573)
miR-42	A2806	TCACCGGGTTAACATCTACAGA	Oligo used for DNA standard
miR-42	A2818	TCTGTAGATGTTAACCCGGTGA/3StarFire/	Oligo used for starfire Northern probes
miR-43	A2807	TATCACAGTTTACTTGCTGTCGC	Oligo used for DNA standard
miR-43	A2819	GCGACAGCAAGTAAACTGTGATA/3StarFire/	Oligo used for starfire Northern probes
miR-44	A2808	CTGGATGTGCTCGTTGGTCATA	Oligo used for DNA standard
miR-44	A2820	AGCTGAATGTGTCTCTAGTCA/3StarFire/	Oligo used for starfire Northern probes
pri-miR-42 F	A2985	CCTTATTCCTTCAAACTTCTG	forward before miR-42
pri-miR-42 F	A3575	GTCCTCGATGTAAATACCGCC	forward pri-miR-42-44 (603nt with A3576)
pri-miR-44 R	A2986	TTTGTTAGTTAATCACTTACTGTTG	reverse after miR-44
pri-miR-44 R	A3576	GAATCTTGAGCGGTCATATCTCTA	reverse pri-miR-42-44 (603nt with A3575)
pri-miR-238 F	A2979	TTTGTCAGTTAACCCACTCC	forward before miR-238
pri-miR-80 R	A2980	ACTTCAGGTTGTGAATGTGG	reverse after miR-80
pri-miR-35 F	A3522	CCCTCTCCTAATTTCCATTCC	forward miR-35
pri-miR-41 R	A3523	AAGTCTTCCGGCAGATCAGA	reverse miR-41
pri-mirtron-62 F	A3524	CATGTACTCCGGCTATAGTG	forward miR-62
pri-mirtron-62 R	A3525	GCTGGAGAGTAGCACTGTC	reverse miR-62
pri-let-7 F	A41	CAAGCAGGCGATTGGTGGA	total pri-let-7 forward
pri-let-7 R	A42	GACGCAGCTTCGAAGAGTTCTGTC	pri-let-7 reverse
Plet-858 F	A3015	GGCTTAAUAAGCTTCTGGAGATTGCTCCAGC	for USER cloning let-858 promoter to MosSCI USER
Plet-858 R	A3016	AGTCCAUACCGGTATCTTACTATAAAAAAGTTTGAAT	for USER cloning let-858 promoter to 3X FLAG GFP
drsh-1 F	A3017	ACTATACAAAAUGTCGGACGAAAAGATTTCAATGACGC	for USER cloning GFP to Drsh-1
drsh-1 R	A2856	CAA CGA CTG ATC CAA AGG AC	downstream of *drsh-1(ok369)* deletion on Chr I
drsh-1 R	A3018	GGTTTAAUGGAGAGAAATTGTAATTTATTGCAGCAATTCAAAC	for USER cloning Drsh-1 to MosSCI USER

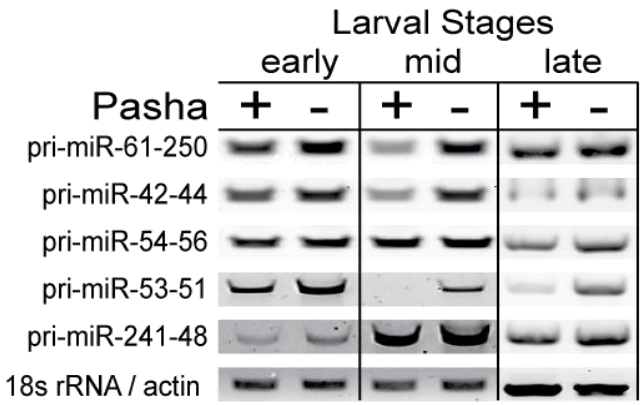


Figure 5.1 Expression of primary transcripts over *C. elegans* development. RT-PCR analysis show primary transcripts from clustered miRNAs accumulate when Pasha is depleted. 18S ribosomal RNA (rRNA) and actin serve as controls both for loading and to show mRNA transcripts are not regulated by the miRNA pathway.

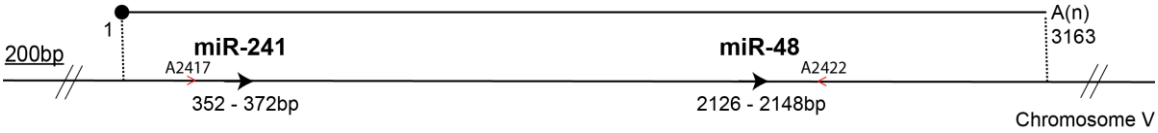


Figure 5.2 Pri-miR-241-48 cluster. The 3162 nt primary transcript of the miR-241-48 cluster is capped and polyadenylated. The depicted transcriptional start site and polyadenylation site was discovered in our lab by RACE and confirmed by Northern Blot and qRT-PCR analysis (data not shown).

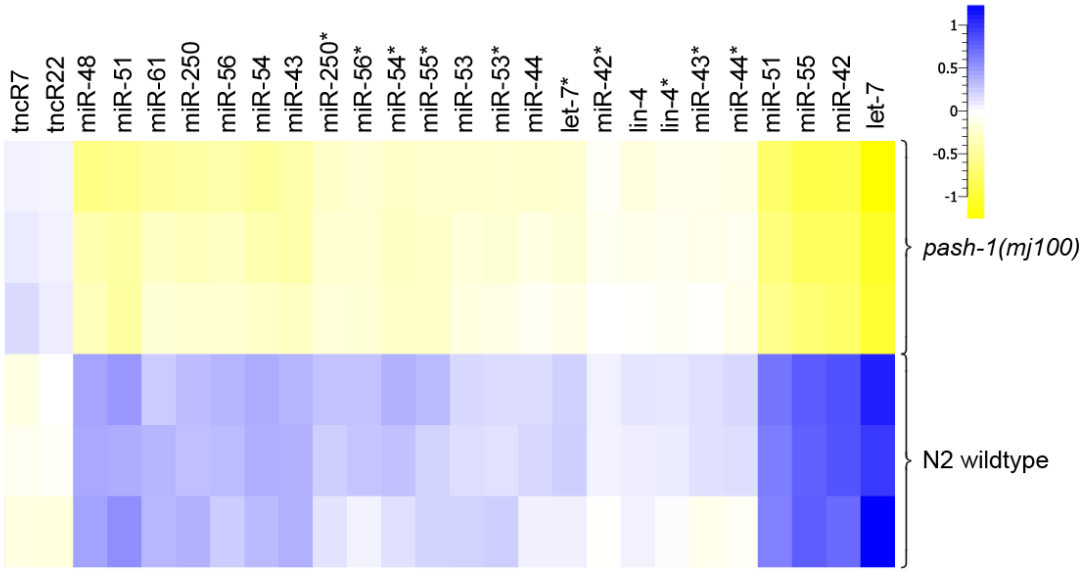


Figure 5.3 Differential miRNA expression measured by multiplex assay. Expression of two Drosha independent small non-coding RNAs, 14 mature miRNAs and star strands in three biological samples of L4 staged wild-type or *pash-1(mj100)* worms grown at non-permissive temperatures. The heat map was generated using Firefly Analysis Workbench software.

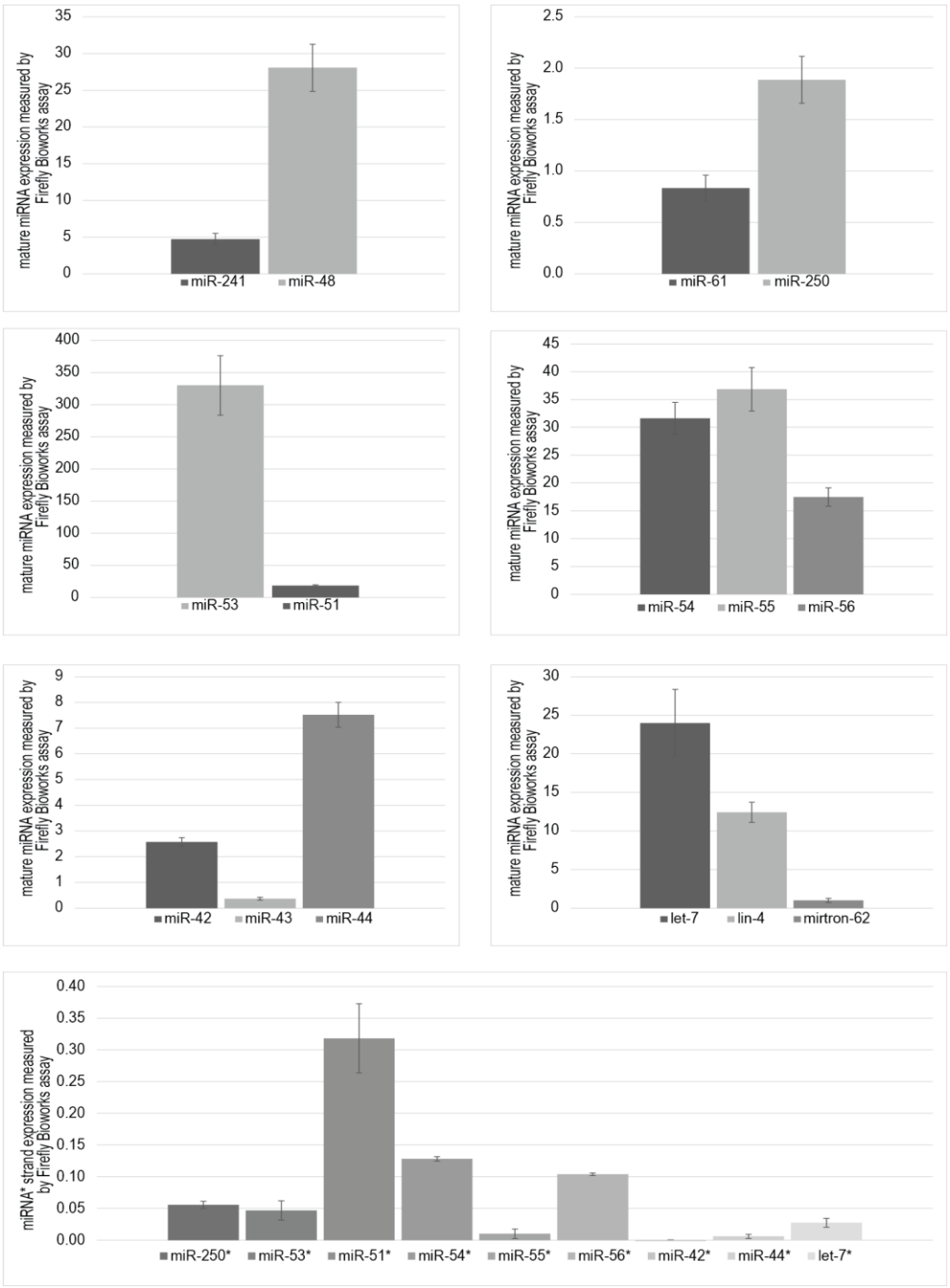


Figure 5.4 Clustered miRNA expression in wild-type worms. From top left to bottom right: Five clusters, miR-241-48, miR-61-250, miR-53-51, miR-54-56, and miR-42-44, along with control miRNAs let-7, lin-4, mirtron-62, and their corresponding miRNA* strands were measured in N2 wild-type worms at the L4 larval stage via multiplex assay. Three biological replicates were measured using two technical replicates. Standard error bars.

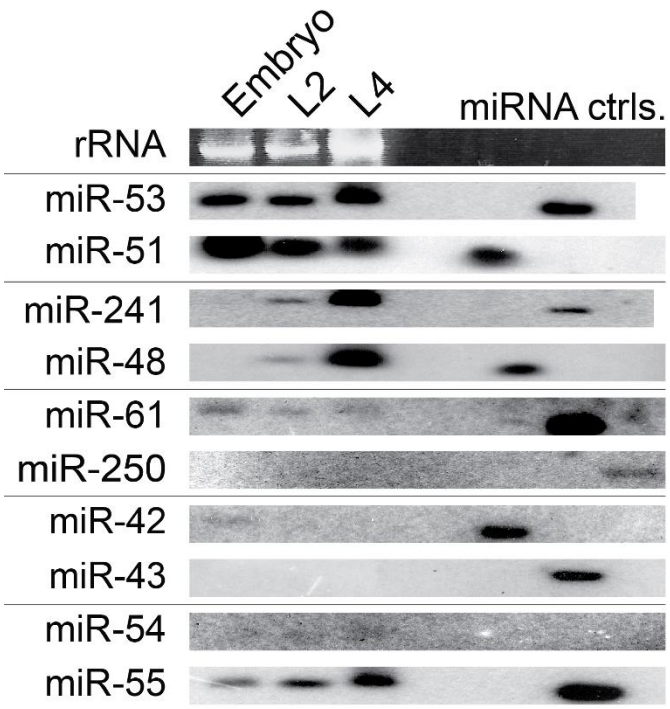


Figure 5.5 Clustered miRNA expression over development. Representative Northern blot of two biological replicates using total RNA from wild-type N2 worms grown to different timepoints. MiRNA controls are standardized DNA oligos. Each row is indicative of two of the miRNAs in a cluster from 5' to 3'.

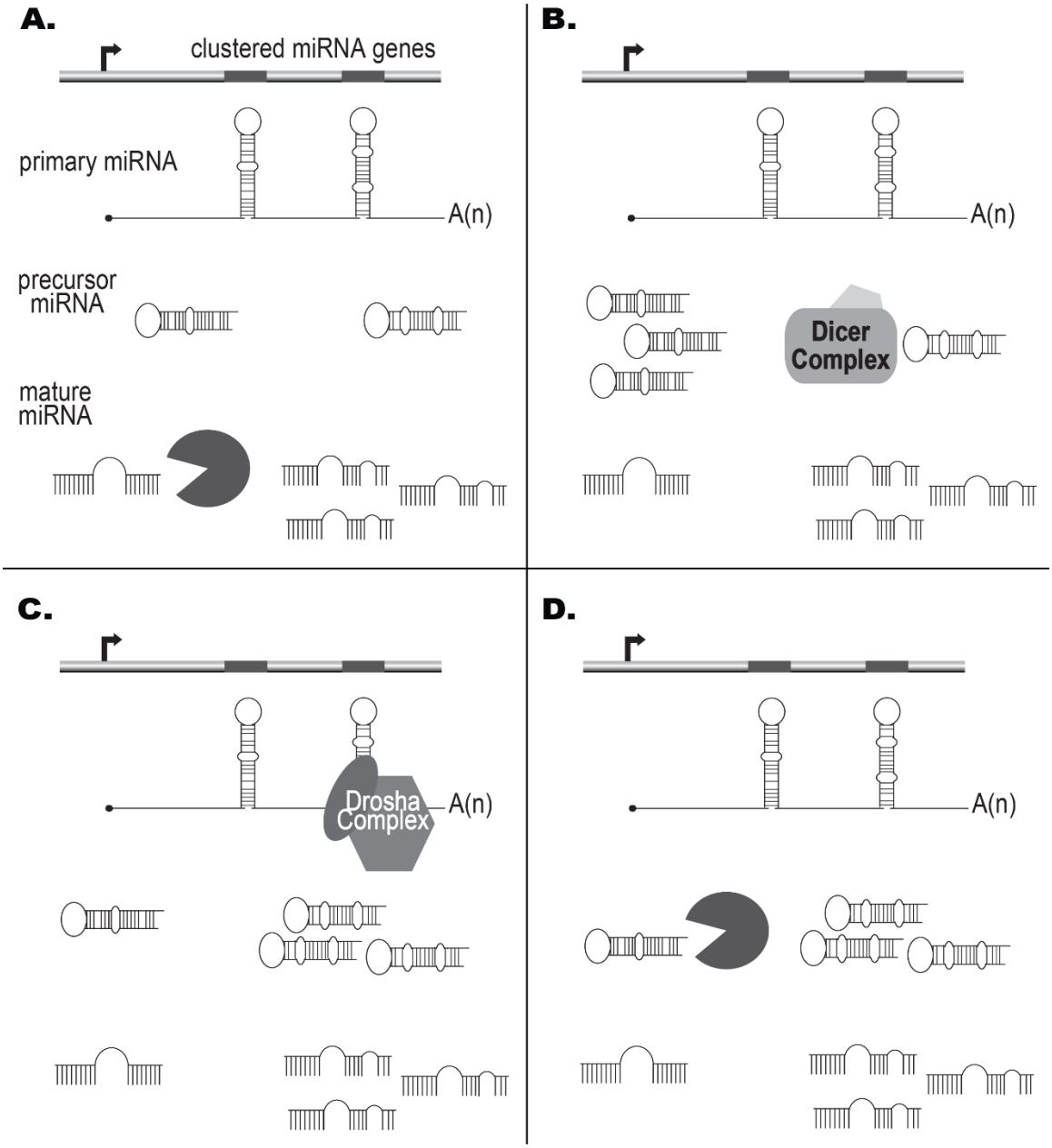


Figure 5.6 MiRNA biogenesis in post-transcriptionally regulated clusters begins with transcription of two or more miRNA genes from one promoter to form a pri-miRNA. The pre-miRNA hairpins are produced by Drosha cleavage and processed to the mature by Dicer. When a miRNA from a cluster is overrepresented, it suggests four possible mechanisms of regulation. (**A**) One of the mature miRNA species may be favorably degraded compared to the other. (**B**) When pre-miRNAs are also unevenly expressed this indicates preferential processing of one miRNA by Dicer, (**C**) unequal processing of the hairpins by Drosha, or (**D**) increased degradation of one of the pre-miRNAs.

Chapter VI

Discussion

In the past two decades, miRNAs have emerged as significant genetic regulators of almost all cellular pathways in a large variety of eukaryotes. Multiple companies are currently pursuing the use of miRNAs as drugs to downregulate overexpressed oncogenes, or as biomarkers for disease diagnosis. Our understanding of how miRNAs function and what regulates miRNA expression has evolved rapidly. The research described here highlights the importance of characterizing pri-miRNA transcripts, as they can offer clues to regulatory mechanisms operating on particular miRNAs.

A. Insights on lin-4 and RBM-28 regulation

After a screen of ~250 candidate RNA binding proteins, RBM-28 emerged as a regulator of lin-4 expression, however it was unclear which step of biogenesis was affected. We demonstrate that RBM-28 is an essential, conserved RNA recognition-motif protein ubiquitously expressed in *C. elegans*. Moreover, we found that RBM-28 stabilizes mature lin-4 expression without affecting primary or precursor levels. This required identifying and characterizing the expression of pri-lin-4. The lin-4 miRNA is located in the intron of a larger, ubiquitously expressed host gene. Notably, we found two independent pri-lin-4 transcripts that derive from intronic promoters, not the host gene promoter. This was a significant finding, since host gene promoters were believed to be the

primary, if not exclusive contributor, to the expression of intronic miRNAs (Ying & Lin, 2006; Kim & Kim, 2007; Saini et al., 2007). Only recent studies have found instances of intronic promoters providing independent transcriptional regulation of miRNAs (Ramalingam et al., 2014; Chang et al., 2015). We also detected a large peak of pri-lin-4 expression at the end of the L1 larval stage, which diminished at the L2 larval stage. It is now understood that pri-lin-4 is one of many genes that oscillate, and oscillation is important to maintain a balanced expression of lin-4 target, LIN-14 (Kim et al., 2013; Hendriks et al., 2014). As more research has emerged on pri-miRNA expression, we now know that alternative promoters are commonly used to transcribe miRNAs in mice and humans (Chang et al., 2015). These multiple promoters can have distinct *cis* regulatory elements that allow both spatial and temporal control over the pri-miRNA isoforms (Kai et al., 2012). It is still unclear how intronic miRNAs transcribed as part of a host gene avoids processing.

Both RBM-28 and lin-4 have homologs in humans. Lin-4 is a homolog of the human miR-125, and homologs sharing the same seed sequence can be found in several species (Griffiths-Jones et al., 2006; Kozomara & Griffiths-Jones, 2014). Because they are found in both nematodes and vertebrates, lin-4 is believed to have appeared at least 550 million years ago (Potenza & Russo, 2013). Interestingly, in all organisms, lin-4 homologs play crucial roles in development and maintenance of adult tissues. For example, in *D. melanogaster*, miR-125 is expressed at the juvenile to adult transition, which is similar to expression in *C. elegans* where lin-4 is expressed at the L1/L2 larval transition

and regulates proteins important for the L2/L3 transition (Caygill & Johnston, 2008). Lin-4 targeting is also conserved: in humans and mice miR-125 downregulates LIN-28 (Wu & Belasco, 2005; Zhong et al., 2010).

It is unknown whether human RBM28 has an effect on miR-125 processing or expression. In humans, decreased expression of RBM28 had been linked to a rare hair disorder, Alopecia-neurological defects-endocrinopathy (ANE) syndrome (Damianov et al., 2006; Nousbeck et al., 2008). A more recent study exploring miRNA targets in keratinocytes shows that RBM28 regulates expression of miR-203, and miR-203 in turn regulates a transcription factor necessary for expression of the stem cell protein p63 (Warshauer et al., 2015). In our study, we found that knockdown of RBM-28 led to increased expression of primary and mature mir-2 and mir-47, which are also embryonically expressed miRNAs. Whether RBM-28 regulates other miRNAs in humans or other organisms is still unclear. Also unclear is how RBM-28, a nucleolar protein with roles in splicing, is affecting the expression of select miRNAs, and whether this is a direct or indirect form of regulation.

B. Implications of splicing and secondary structure on miRNA regulation

In 2004, the Pasquinelli lab identified the pri-let-7 transcripts in *C. elegans* and found two promoters, as well as an SL1 *trans*-spliced isoform (Bracht et al., 2004). Pri-let-7 undergoes a *trans*-splicing event whereby a 3’ss, ~30 nt sequence upstream of mature let-7, is cleaved and a 22 nt splice leader 1 (SL1)

sequence is appended at the 5’ end. The splice site is conserved in other *Caenorhabditis* species and mutant worms lacking this site produced reduced levels of mature let-7, often causing lethality. My research uncovered a link between splicing and processing of let-7. RNase assays indicate that the unspliced pri-let 7 transcript features structured regions at the base of the precursor, disrupting what should be unpaired areas critical for Drosha slicing activity. In contrast, the SL1-spliced pri-let-7 molecule takes on a structure that better resembles a canonical pri-miRNA substrate, including a terminal loop, internal loops in the upper and lower stem, and flanking unpaired basal segments. The Microprocessor preferentially binds and processes the SL1-spliced pri-let-7 *in vivo.* In addition, we found that worms expressing single-copy mutated transgenes that disrupted splicing and the SL1 sequence were deficient in let-7 processing. Finally, a transgenic worm expressing the SL1 sequence as well as a downstream mutation that forces a secondary structure predicted to fold like unspliced pri-let-7 was unable to rescue the *let-7* null lethality. This indicates that the structural rearrangements and not the SL1 sequence is the primary determinant of pri-let-7 processing.

My work shows that the genomic sequence does not necessarily serve as the miRNA primary transcript and suggests that splicing in general could be a mechanism for remodeling some miRNA primary transcripts to enhance their recognition by Drosha. While *trans*-splicing does not normally occur in higher organisms, several human pri-let-7 isoforms are expressed as introns, and thus are subject to splicing. In addition, instances of *trans*-splicing in tumors are

common, and this may have implications on how miRNAs are misregulated in cancer (Li et al., 2008; Zaphiropoulos, 2011). A recent study in humans described that the pri-miR-17-92 transcript undergoes a splicing step during embryonic development to produce a progenitor miRNA (pro-miRNA) that is a more structurally favorable substrate for Drosha processing, enhancing expression of all the miRNAs within the cluster (Du et al., 2015). Thus, my finding that pri-miRNAs can be regulated by splicing-induced structural rearrangements may be more broadly applied. As another potential example of this, in plants, intron adjacent miRNAs are enhanced when those introns are spliced out, though the mechanism remains unclear (Schwab et al., 2013). In addition, multiple studies have shown that secondary structure can influence alternative splicing, and there are multiple instances of competition between pri-miRNA processing and exon inclusion (Janas et al., 2011; Melamed et al., 2013; Chang et al., 2015; Sohail & Xie, 2015). New insights on circRNAs, which are formed through back-splicing, reveals that splicing is even more complex than previously thought, and in fact may regulate expression of miRNAs by providing competing binding sites for that particular miRNA (Hansen et al., 2013; Memczak et al., 2013).

Future studies will reveal the generality of RNA splicing as a means to regulate expression through structural rearrangements. A major breakthrough that may shed light on this are advances in genome-wide RNA structural analysis. In particular, the recent investigations of secondary structures of the transcriptome have recently been investigated in yeast, humans, mouse embryonic cells, bacteria, and plants (Kertesz et al., 2010; Wan et al., 2011;

Incarnato et al., 2014; Wan et al., 2014; Del Campo et al., 2015; Foley et al., 2015; Wan et al., 2016). Interestingly, plant and yeast mRNAs tend to be more structured in coding regions (CDS), while untranslated regions (UTRs) in mammals tend to be most structured. In contrast, in bacteria structures were found similarly in CDS and UTRs. Messenger RNAs subject to post-transcriptional regulation had unique secondary structures and the most structured RNAs tend to be bound by RNA-binding proteins, emphasizing the importance of structure on protein-RNA interactions. All of these studies have discovered potential mechanisms for structure-based gene regulation; especially given the vast number of RNA modifications observed that alter secondary structure (Kertesz et al., 2010; Wan et al., 2011; Incarnato et al., 2014; Wan et al., 2014; Del Campo et al., 2015; Foley et al., 2015).

In plants, researchers have found examples of RNA structures that can inhibit translation but these structures destabilize at higher temperatures, releasing the translational inhibition, thus allowing a heat-shock response (Foley et al., 2015). In *Escherichia coli*, RNA structures are highly dynamic, taking on multiple isoforms, and changing over development (Del Campo et al., 2015). In humans, AGO bound-sites in 3'UTRs tend to be more structurally available underscoring the relationship between mRNA structure and the potential for miRNA regulation (Wan et al., 2014). This study, however, was unlikely to capture pri-miRNAs, which are co-transcriptionally processed rapidly and efficiently by the Microprocessor. It would be interesting to perform genome-wide analysis of RNA secondary structure in worms using the *pash-1(mj100)* mutant

worms, or by creating a dominant-negative Drosha mutant, which was used to detect pri-miRNAs through RNA-seq in humans (Chang et al., 2015; Wan et al., 2016).

In *C. elegans*, SL-splicing has roles in separating polycistronic mRNAs, though most SL-spliced mRNAs are monocistronic. A recent study examining the evolution of operons and the function of SL-splicing in metazoans found a conserved enrichment of operon expression in the germline, particularly in maternally deposited mRNAs (Danks et al., 2015). However, this seems irrelevant for pri-let-7, which is not expressed in embryos. Another interesting aspect of SL-splicing in *C. elegans* is the fact that the majority of unspliced transcripts are difficult to detect due to the proficiency of splicing (Blumenthal, 2012). My study found that the unspliced transcripts do not interact with the Microprocessor *in vivo*. We also know that A and B transcripts of pri-let-7 are transcribed in different tissues and at different times during development, and that this is transcriptionally regulated, although at early timepoints processing is blocked by LIN-28 (Van Wynsberghe et al., 2011; Kai et al., 2012). These findings beg the questions: What is the purpose of pri-let-7 transcripts that are not subject to processing early in development, and what purpose do these unspliced transcripts have later in development? There are rare examples showing direct RNA regulation by pri-miRNAs (Trujillo et al., 2010; Yue et al., 2011; Roy-Chaudhuri et al., 2014). It is possible that unspliced pri-let-7 may serve different roles as lncRNAs, and may even act as a ceRNA.

C. Perspectives on regulation of clustered miRNA biogenesis

When two or more miRNAs are transcribed in tandem on one primary transcript, it is known as a cluster. There are multiple instances of miRNAs arranged in an operon-like cluster throughout the genome (Lau et al., 2001; Lee & Ambros, 2001; Lim et al., 2003). Despite coming from the same primary transcript, oftentimes only one miRNA will be dominantly expressed, and this dominance can change during larval development. In many diseases, including cancer, miRNA levels change through undefined post-transcriptional mechanisms. Thus, factors discovered in these studies may help elucidate how miRNA expression is altered in disease states.

Examining RNA-seq data of miRNA clusters in *C. elegans* shows evidence for differential regulation of 53 miRNAs in 19 potential clusters, where there is over-representation of one of the mature miRNAs compared to the others in the cluster, suggesting post-transcriptional regulation. One confirmed cluster includes the let-7 family members, miR-241, and miR-48. Interestingly, high-throughput RNA deep-sequencing experiments from other labs as well as Northern analysis and multiplex miRNA assays in our lab demonstrate that mature miR-241 is expressed at a significantly lower level than miR-48 (Kato et al., 2009; Zisoulis et al., 2010; Warf et al., 2011). According to miRBase, let-7 family members are found clustered in flies and humans (Kozomara & Griffiths-Jones, 2014). Therefore, analyzing regulation of biogenesis of these clustered miRNAs in *C. elegans* can have implications in overall miRNA biogenesis regulation across species.

I narrowed down a group of clusters in *C*. *elegans* and made progress characterizing the differential expression pattern of the miRNAs in each cluster. The data I have gathered so far indicate that these clustered miRNAs are differentially expressed and the mechanism that regulates this appears to change as development progresses. The timing and expression levels of miRNAs are very important; that two or more miRNAs coming from one primary transcript can have such different levels during development speaks to a highly coordinated and regulated system.

Studies on miRNA clusters have revealed distinct mechanisms of regulation. One study suggests that the first miRNA transcribed from a cluster will be expressed at a higher level due to co-transcriptional degradation of the other by the exonuclease XRN2 (Ballarino et al., 2009). Contrary to this, the clusters I examined showed evidence of the last miRNA in the cluster being most highly expressed.

Several recent studies indicate the importance of structure concerning individual processing of clustered pri-miRNAs. A cluster of 12 miRNAs transcribed from one pri-miRNA transcript in Kaposi's sarcoma herpesvirus are expressed differentially depending on the structural context around the pre-miRNA (Contrant et al., 2014). The ~3 kb pri-miRNA secondary structure was analyzed biochemically and found to contain many structural elements; viral microRNAs with optimal secondary structures, such as loose basal stems, were highly expressed, while miRNAs flanked with other stem-loops were less likely to be processed.

A more recent study in *Drosophila* found that structural rearrangements due to processing and/or transcription of one miRNA in a cluster may enhance the processing of the other miRNA. The pri-miR-11-998 cluster is expressed as part of host gene. However, the expression of miR-998 is dependent on expression of miR-11, such that a deletion of miR-11 interrupts Microprocessor recognition of pri-miR-998 (Truscott et al., 2015). Swapping the position of miR-11 and miR-998 in the cluster did not change the dependence of miR-998 on miR-11. However, replacing miR-11 with an inert short hairpin miRNA, miR-1/mCherry, rescued the miR-998 expression. In addition, a construct altering the secondary structure of the pri-miR-998 hairpin was processed without the need for a nearby miRNA. Taken together, these results suggest that the presence or processing of a nearby miRNA may be responsible for structural rearrangements that enhance Drosha recognition of pri-miR-998.

Another recent paper found that expression of miRNAs in the human pri-miR-17-92 cluster are regulated by a splicing event that modifies the secondary structure in order to facilitate processing of all the miRNAs in this human cluster (Du et al., 2015). Interestingly, mechanisms regulating pri-miR-17-92 processing had previously been reported. However, these researchers did most of their studies *in vitro* and failed to characterize the full pri-miR-17-92 sequence, which is up ~ 8 kb long. This is why they most likely overlooked regulatory elements that contributed to the differential expression of the miRNAs in the cluster (Michlewski et al., 2008; Chaulk et al., 2014).

These recent studies imply that secondary structure is a common

regulatory mechanism for clustered miRNA biogenesis. Examining secondary structure predictions of the clustered miRNAs in my *C. elegans* study identified three pri-miRNA clusters that are potentially regulated through this mechanism. Notably, the pri-miRNA clusters have not all been fully characterized, and this will be important for determining secondary structure contributions. The data I have gathered so far indicates that these clusters are transcribed and subject to Microprocessor processing; determining transcriptional start and end sites may find alternative promoters regulating independent transcription of one or more miRNA in the cluster, which is another recently found mechanism of clustered miRNA biogenesis (Chang et al., 2015).

D. Concluding remarks

The advances made at the turn of the millennium have unlocked a great appreciation of non-coding RNAs (ncRNAs). Originally, these areas of the genome were described as “junk” or “transcriptional noise” but in less than two decades, research on ncRNA regulation and function has challenged this view and greatly enriched our understanding on how cellular networks operate. My research was able to help characterize a conserved RNA-binding protein in the lin-4 biogenesis pathway, and found a role for splicing in let-7 biogenesis. In addition, the data I gathered on clustered miRNAs has uncovered examples of differentially expressed miRNAs transcribed from a common pri-miRNA. Yet this work leaves several open questions.

One question pertains to the specificity of processing. In addition to *C.*

elegans pri-lin-4, there are human examples of miRNAs transcribed from intronic promoters. How do they avoid being processed out of the host pre-mRNA? Perhaps the secondary structure is not conducive to processing or other *trans* acting factors may prohibit binding of the Microprocessor. Another question pertains to the splicing regulator RBM-28. How does it regulate expression of lin-4? For example, does it act directly or indirectly? Does it regulate other miRNAs or in the same manner or is its ability to regulate lin-4 unique? How does it fit overall amongst the RNA regulatory network?

After SL1 *trans*-splicing, unspliced pri-let-7 transcripts remain highly expressed, which is unexpected. Could unspliced pri-let-7 have additional functions? RNA regulation by a pri-miRNA has been shown in a few rare examples (Trujillo et al., 2010; Yue et al., 2011; Roy-Chaudhuri et al., 2014). In addition, several *cis* regulatory elements exist in these unspliced molecules; perhaps they serve to sponge away factors that may otherwise hinder SL1-pri-let-7 processing.

Finally, considering the widespread existence of clustered miRNAs and the indications of post-transcriptional mechanisms to produce different levels of such co-transcribed miRNAs, my studies that elucidated how the expression of clustered miRNAs is regulated in *C. elegans* will likely provide new insights for general mechanisms that control miRNA biogenesis.

In the past 15 years, advances in molecular and bioinformatics tools have improved exponentially, allowing almost daily discoveries in the ncRNA field. With those discoveries come a lot of unanswered questions and unknown

complexities. The progress made researching miRNA regulation will be valuable for understanding organismal development, diagnosing disease and advancing cancer therapeutics.

REFERENCES

Abbott AL, Alvarez-Saavedra E, Miska EA, Lau NC, Bartel DP, Horvitz HR, Ambros V. 2005. The let-7 MicroRNA family members mir-48, mir-84, and mir-241 function together to regulate developmental timing in Caenorhabditis elegans. *Dev Cell 9*:403 - 414.

Abrahante JE, Daul AL, Li M, Volk ML, Tennessen JM, Miller EA, Rougvie AE. 2003. The Caenorhabditis elegans hunchback-like Gene lin-57/hbl-1 Controls Developmental Time and Is Regulated by MicroRNAs. *Developmental Cell 4*:625-637.

Adams MD, Celniker SE, Holt RA, Evans CA, Gocayne JD, Amanatides PG, Scherer SE, Li PW, Hoskins RA, Galle RF, George RA, Lewis SE, Richards S, Ashburner M, Henderson SN, Sutton GG, Wortman JR, Yandell MD, Zhang Q, Chen LX, Brandon RC, Rogers Y-HC, Blazej RG, Champe M, Pfeiffer BD, Wan KH, Doyle C, Baxter EG, Helt G, Nelson CR, Gabor GL, Miklos, Abril JF, Agbayani A, An H-J, Andrews-Pfannkoch C, Baldwin D, Ballew RM, Basu A, Baxendale J, Bayraktaroglu L, Beasley EM, Beeson KY, Benos PV, Berman BP, Bhandari D, Bolshakov S, Borkova D, Botchan MR, Bouck J, Brokstein P, Brottier P, Burtis KC, Busam DA, Butler H, Cadieu E, Center A, Chandra I, Cherry JM, Cawley S, Dahlke C, Davenport LB, Davies P, Pablos Bd, Delcher A, Deng Z, Mays AD, Dew I, Dietz SM, Dodson K, Doup LE, Downes M, Dugan-Rocha S, Dunkov BC, Dunn P, Durbin KJ, Evangelista CC, Ferraz C, Ferriera S, Fleischmann W, Fosler C, Gabrielian AE, Garg NS, Gelbart WM, Glasser K, Glodek A, Gong F, Gorrell JH, Gu Z, Guan P, Harris M, Harris NL, Harvey D, Heiman TJ, Hernandez JR, Houck J, Hostin D, Houston KA, Howland TJ, Wei M-H, Ibegwam C, Jalali M, Kalush F, Karpen GH, Ke Z, Kennison JA, Ketchum KA, Kimmel BE, Kodira CD, Kraft C, Kravitz S, Kulp D, Lai Z, Lasko P, Lei Y, Levitsky AA, Li J, Li Z, Liang Y, Lin X, Liu X, Mattei B, McIntosh TC, McLeod MP, McPherson D, Merkulov G, Milshina NV, Mobarry C, Morris J, Moshrefi A, Mount SM, Moy M, Murphy B, Murphy L, Muzny DM, Nelson DL, Nelson DR, Nelson KA, Nixon K, Nusskern DR, Pacleb JM, Palazzolo M, Pittman GS, Pan S, Pollard J, Puri V, Reese MG, Reinert K, Remington K, Saunders RDC, Scheeler F, Shen H, Shue BC, Sidén-Kiamos I, Simpson M, Skupski MP, Smith T, Spier E, Spradling AC, Stapleton M, Strong R, Sun E, Svirskas

R, Tector C, Turner R, Venter E, Wang AH, Wang X, Wang Z-Y, Wassarman DA, Weinstock GM, Weissenbach J, Williams SM, Woodage T, Worley KC, Wu D, Yang S, Yao QA, Ye J, Yeh R-F, Zaveri JS, Zhan M, Zhang G, Zhao Q, Zheng L, Zheng XH, Zhong FN, Zhong W, Zhou X, Zhu S, Zhu X, Smith HO, Gibbs RA, Myers EW, Rubin GM, Venter JC. 2000. The Genome Sequence of Drosophila melanogaster. *Science 287*:2185-2195.

Agarwal V, Bell GW, Nam J-W, Bartel DP. 2015. Predicting effective microRNA target sites in mammalian mRNAs. *eLife 4*:e05005.

Alemán LM, Doench J, Sharp PA. 2007. Comparison of siRNA-induced off-target RNA and protein effects. *RNA 13*:385-395.

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

Ambros V, Horvitz HR. 1984. Heterochronic Mutants of the Nematode *Caenorhabditis elegans*. *Science 226*:409-416.

Ambros V, Lee RC, Lavanway A, Williams PT, Jewell D. 2003. MicroRNAs and Other Tiny Endogenous RNAs in C. elegans. *Current biology : CB 13*:807-818.

Ameres SL, Zamore PD. 2013. Diversifying microRNA sequence and function. *Nat Rev Mol Cell Biol 14*:475-488.

Arasu P, Wightman B, Ruvkun G. 1991. Temporal regulation of lin-14 by the antagonistic action of two other heterochronic genes, lin-4 and lin-28. *Genes & Development 5*:1825-1833.

Auyeung VC, Ulitsky I, McGeary SE, Bartel DP. 2013. Beyond Secondary Structure: Primary-Sequence Determinants License Pri-miRNA Hairpins for Processing. *Cell 152*:844-858.

Baek D, Villen J, Shin C, Camargo F, Gygi S, Bartel D. 2008. The impact of microRNAs on protein output. *Nature 455*:64 - 71.

Bagga S, Bracht J, Hunter S, Massirer K, Holtz J, Eachus R, Pasquinelli AE. 2005. Regulation by let-7 and lin-4 miRNAs Results in Target mRNA Degradation. *Cell 122*:553-563.

Ballarino M, Pagano F, Girardi E, Morlando M, Cacchiarelli D, Marchioni M, Proudfoot NJ, Bozzoni I. 2009. Coupled RNA Processing and Transcription of Intergenic Primary MicroRNAs. *Mol Cell Biol 29*:5632-5638.

Bar M, Wyman SK, Fritz BR, Qi J, Garg KS, Parkin RK, Kroh EM, Bendoraite A, Mitchell PS, Nelson AM, Ruzzo WL, Ware C, Radich JP, Gentleman R, Ruohola-Baker H, Tewari M. 2008. MicroRNA Discovery and Profiling in Human Embryonic Stem Cells by Deep Sequencing of Small RNA Libraries. *STEM CELLS 26*:2496-2505.

Bartel DP. 2009. MicroRNAs: Target Recognition and Regulatory Functions. *Cell 136*:215-233.

Bastian F, Tyler B, Liam EP, Morten J, James ET, Benjamin LK, James MN, Lorenzo FS, Kjersti F, Eivind H, Kevin JP. 2015. A Uniform System for the Annotation of Vertebrate microRNA Genes and the Evolution of the Human microRNAome. *Annual Review of Genetics 49*.

Baugh LR, Sternberg PW. 2006. DAF-16/FOXO Regulates Transcription of cki-1/Cip/Kip and Repression of lin-4 during C. elegans L1 Arrest. *Current Biology 16*:780-785.

Beitzinger M, Peters L, Zhu JY, Kremmer E, Meister G. 2007. Identification of Human microRNA Targets From Isolated Argonaute Protein Complexes. *RNA Biology 4*:76-84.

Berezikov E, Chung W-J, Willis J, Cuppen E, Lai EC. 2007. Mammalian Mirtron Genes. *Molecular Cell 28*:328-336.

Betel D, Wilson M, Gabow A, Marks DS, Sander C. 2008. The microRNA.org resource: targets and expression. *Nucleic Acids Res*:D149 - D153.

Bethke A, Fielenbach N, Wang Z, Mangelsdorf DJ, Antebi A. 2009. Nuclear Hormone Receptor Regulation of MicroRNAs Controls Developmental Progression. *Science 324*:95-98.

Bielewicz D, Kalak M, Kalyna M, Windels D, Barta A, Vazquez F, Szweykowska-Kulinska Z, Jarmolowski A. 2013. Introns of plant pri-miRNAs enhance miRNA biogenesis. *EMBO Reports 14*:622-628.

Blumenthal T. 2012. Trans-splicing and operons. In: Community TCeR, ed. *WormBook*: WormBook.

Bohnsack MT, Czaplinski K, GÃ–Rlich D. 2004. Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. *RNA 10*:185-191.

Bosson Andrew D, Zamudio Jesse R, Sharp Phillip A. 2014. Endogenous miRNA and Target Concentrations Determine Susceptibility to Potential ceRNA Competition. *Molecular Cell 56*:347-359.

Boyerinas B, Park S-M, Hau A, Murmann AE, Peter ME. 2010. The role of let-7 in cell differentiation and cancer. *Endocr Relat Cancer 17*:F19-36.

Boyerinas B, Park S-M, Shomron N, Hedegaard MM, Vinther J, Andersen JS, Feig C, Xu J, Burge CB, Peter ME. 2008. Identification of Let-7-Regulated Oncofetal Genes. *Cancer Research 68*:2587-2591.

Bracht J, Hunter S, Eachus R, Weeks P, Pasquinelli AE. 2004. Trans-splicing and polyadenylation of let-7 microRNA primary transcripts. *RNA 10*:1586-1594.

Bracht JR, Van Wynsberghe PM, Mondol V, Pasquinelli AE. 2010. Regulation of lin-4 miRNA expression, organismal growth and development by a conserved RNA binding protein in C. elegans. *Developmental Biology 348*:210-221.

Brennecke J, Stark A, Russell RB, Cohen SM. 2005. Principles of MicroRNA-Target Recognition. *PLoS Biol 3*:e85.

Broughton JP, Pasquinelli AE. 2013. Identifying Argonaute binding sites in Caenorhabditis elegans using iCLIP. *Methods 63*:119-125.

Burger K, Gullerova M. 2015. Swiss army knives: non-canonical functions of nuclear Drosha and Dicer. *Nat Rev Mol Cell Biol 16*:417-430.

Büssing I, Slack FJ, Großhans H. 2008. let-7 microRNAs in development, stem cells and cancer. *Trends in Molecular Medicine 14*:400-409.

Bussing I, Yang J-S, Lai EC, Grosshans H. 2010. The nuclear export receptor XPO-1 supports primary miRNA processing in C. elegans and Drosophila. *EMBO J 29*:1830-1839.

Caygill EE, Johnston LA. 2008. Temporal Regulation of Metamorphic Processes in Drosophila by the let-7 and miR-125 Heterochronic MicroRNAs. *Current Biology 18*:943-950.

Chalfie M, Horvitz HR, Sulston JE. 1981. Mutations that lead to reiterations in the cell lineages of C. elegans. *Cell 24*:59-69.

Chan S-P, Slack FJ. 2009. Ribosomal protein RPS-14 modulates let-7 microRNA function in Caenorhabditis elegans. *Developmental Biology 334*:152-160.

Chan S, Slack F. 2007. And now introducing mammalian Mirtrons. *Developmental Cell 13*:605 - 607.

Chandradoss Stanley D, Schirle Nicole T, Szczepaniak M, MacRae Ian J, Joo C. 2015. A Dynamic Search Process Underlies MicroRNA Targeting. *Cell 162*:96-107.

Chang T-C, Pertea M, Lee S, Salzberg SL, Mendell JT. 2015. Genome-wide annotation of microRNA primary transcript structures reveals novel regulatory mechanisms. *Genome Research*.

Chatterjee S, Fasler M, Büssing I, Großhans H. 2011. Target-Mediated Protection of Endogenous MicroRNAs in C. elegans. *Developmental Cell 20*:388-396.

Chatterjee S, Grosshans H. 2009. Active turnover modulates mature microRNA activity in Caenorhabditis elegans. *Nature 461*:546-549.

Chaulk SG, Xu Z, Glover MJN, Fahlman RP. 2014. MicroRNA miR-92a-1 biogenesis and mRNA targeting is modulated by a tertiary contact within the miR-17~92 microRNA cluster. *Nucleic Acids Research 42*:5234-5244.

Chi SW, Zang JB, Mele A, Darnell RB. 2009. Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps. *Nature 460*:479 - 486.

Chiang HR, Schoenfeld LW, Ruby JG, Auyeung VC, Spies N, Baek D, Johnston WK, Russ C, Luo S, Babiarz JE, Blelloch R, Schroth GP, Nusbaum C, Bartel DP. 2010. Mammalian microRNAs: experimental evaluation of novel and previously annotated genes. *Genes & Development 24*:992-1009.

Chiu H-S, Llobet-Navas D, Yang X, Chung W-J, Ambesi-Impiombato A, Iyer A, Kim HR, Seviour EG, Luo Z, Sehgal V, Moss T, Lu Y, Ram P, Silva J, Mills GB, Califano A, Sumazin P. 2015. Cupid: simultaneous reconstruction of microRNA-target and ceRNA networks. *Genome Research 25*:257-267.

Chung W-J, Agius P, Westholm JO, Chen M, Okamura K, Robine N, Leslie CS, Lai EC. 2011. Computational and experimental identification of mirtrons in Drosophila melanogaster and Caenorhabditis elegans. *Genome Research 21*:286-300.

Conrad T, Marsico A, Gehre M, Andersson Ørom U. 2014. Microprocessor Activity Controls Differential miRNA Biogenesis In Vivo. *Cell Reports 9*:542-554.

Consortium IHGS. 2001. Initial sequencing and analysis of the human genome. *Nature 409*:860-921.

Consortium IHGS. 2004. Finishing the euchromatic sequence of the human genome. *Nature 431*:931-945.

Consortium TCeS. 1998. Genome Sequence of the Nematode C. elegans: A Platform for Investigating Biology. *Science 282*:2012-2018.

Contrant M, Fender A, Chane-Woon-Ming B, Randrianjafy R, Vivet-Boudou V, Richer D, Pfeffer S. 2014. Importance of the RNA secondary structure for the relative accumulation of clustered viral microRNAs. *Nucleic Acids Research 42*:7981-7996.

Corcoran DL, Pandit KV, Gordon B, Bhattacharjee A, Kaminski N, Benos PV. 2009. Features of mammalian microRNA promoters emerge from polymerase II chromatin immunoprecipitation data. *PLoS ONE 4*:e5279.

Crick F. 1958. The Biological Replication of Macromolecules. *Symposia of the Society for Experimental Biology*. University College London: Cambridge University Press. pp 138.

Crick F. 1970. Central Dogma of Molecular Biology. *Nature 227*:561-563.

Czech B, Hannon GJ. 2011. Small RNA sorting: matchmaking for Argonautes. *Nat Rev Genet 12*:19-31.

Damianov A, Kann M, Lane William S, Bindereif A. 2006. Human RBM28 protein is a specific nucleolar component of the spliceosomal snRNPs. *Biological Chemistry 387*:1455.

Danks GB, Raasholm M, Campsteijn C, Long AM, Manak JR, Lenhard B, Thompson EM. 2015. Trans-Splicing and Operons in Metazoans: Translational Control in Maternally Regulated Development and Recovery from Growth Arrest. *Molecular Biology and Evolution 32*:585-599.

Del Campo C, Bartholomäus A, Fedyunin I, Ignatova Z. 2015. Secondary Structure across the Bacterial Transcriptome Reveals Versatile Roles in mRNA Regulation and Function. *PLoS Genet 11*:e1005613.

Denli AM, Tops BBJ, Plasterk RHA, Ketting RF, Hannon GJ. 2004. Processing of primary microRNAs by the Microprocessor complex. *Nature 432*:231-235.

Ding XC, Groszhans H. 2009. Repression of C. elegans microRNA targets at the initiation level of translation requires GW182 proteins. *EMBO J 28*:213-222.

Djuranovic S, Nahvi A, Green R. 2011. A Parsimonious Model for Gene Regulation by miRNAs. *Science 331*:550-553.

Doench JG, Petersen CP, Sharp PA. 2003. siRNAs can function as miRNAs. *Genes & Development 17*:438-442.

Doench JG, Sharp PA. 2004. Specificity of microRNA target selection in translational repression. *Genes & Development 18*:504-511.

Du P, Wang L, Sliz P, Gregory Richard I. 2015. A Biogenesis Step Upstream of Microprocessor Controls miR-17~92 Expression. *Cell 162*:885-899.

Easow G, Teleman AA, Cohen SM. 2007. Isolation of microRNA targets by miRNP immunopurification. *RNA 13*:1198-1204.

Enright A, John B, Gaul U, Tuschl T, Sander C, Marks D. 2003. MicroRNA targets in Drosophila. *Genome Biology 5*:R1.

Esquela-Kerscher A, Johnson SM, Bai L, Saito K, Partridge J, Reinert KL, Slack FJ. 2005. Post-embryonic expression of C. elegans microRNAs belonging to the lin-4 and let-7 families in the hypodermis and the reproductive system. *Dev Dyn 234*:868 - 877.

Esquela-Kerscher A, Trang P, Wiggins JF, Patrawala L, Cheng A, Ford L, Weidhaas JB, Brown D, Bader AG, Slack FJ. 2008. The *let-7* microRNA reduces tumor growth in mouse models of lung cancer. *Cell Cycle 7*:759-764.

Fabian MR, Sonenberg N. 2012. The mechanics of miRNA-mediated gene silencing: a look under the hood of miRISC. *Nature Structural & Molecular Biology 19*:586-593.

Fabian MR, Sonenberg N, Filipowicz W. 2010. Regulation of mRNA Translation and Stability by microRNAs. *Annual Review of Biochemistry 79*:351-379.

Fang Z, Rajewsky N. 2011. The Impact of miRNA Target Sites in Coding Sequences and in 3'UTRs. *PLoS ONE 6*:e18067.

Fay DS, Stanley HM, Han M, Wood WB. 1999. A *Caenorhabditis elegans* Homologue of hunchbackIs Required for Late Stages of Development but Not Early Embryonic Patterning. *Developmental Biology 205*:240-253.

Finnegan EF, Pasquinelli AE. 2013. MicroRNA biogenesis: regulating the regulators. *Critical Reviews in Biochemistry and Molecular Biology 48*:51-68.

Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC. 1998. Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature 391*:806-811.

Flynt AS, Greimann JC, Chung W-J, Lima CD, Lai EC. 2010. MicroRNA Biogenesis via Splicing and Exosome-Mediated Trimming in Drosophila. *Molecular Cell 38*:900-907.

Foley SW, Vandivier LE, Kuksa PP, Gregory BD. 2015. Transcriptome-wide measurement of plant RNA secondary structure. *Current Opinion in Plant Biology 27*:36-43.

Frokjaer-Jensen C, Davis MW, Hopkins CE, Newman BJ, Thummel JM, Olesen SP, Grunnet M, Jorgensen EM. 2008. Single-copy insertion of transgenes in Caenorhabditis elegans. *Nat Genet 40*:1375 - 1383.

Gehrke S, Imai Y, Sokol N, Lu B. 2010. Pathogenic LRRK2 negatively regulates microRNA-mediated translational repression. *Nature 466*:637-641.

Giraldez AJ, Mishima Y, Rihel J, Grocock RJ, Van Dongen S, Inoue K, Enright AJ, Schier AF. 2006. Zebrafish MiR-430 Promotes Deadenylation and Clearance of Maternal mRNAs. *Science 312*:75-79.

Granato M, Schnabel H, Schnabel R. 1994. pha-1, a selectable marker for gene transfer in C.elegans. *Nucleic Acids Research 22*:1762-1763.

Gregory RI, Yan KP, Amuthan G, Chendrimada T, Doratotaj B, Cooch N, Shiekhattar R. 2004. The Microprocessor complex mediates the genesis of microRNAs. *Nature 432*:235 - 240.

Griffiths-Jones S, Grocock R, van Dongen S, Bateman A, Enright A. 2006. miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Research 34*:D140.

Grosshans H, Chatterjee S. 2010. MicroRNases and the Regulated Degradation of Mature Animal miRNAs. *Regulation of microRNAs*: Springer US. pp 140-155.

Großhans H, Johnson T, Reinert KL, Gerstein M, Slack FJ. 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 F, Sarnow P, Kay MA. 2009. Biological basis for restriction of microRNA targets to the 3[prime] untranslated region in mammalian mRNAs. *Nat Struct Mol Biol 16*:144-150.

Guil S, Esteller M. 2015. RNA-RNA interactions in gene regulation: the coding and noncoding players. *Trends in Biochemical Sciences 40*:248-256.

Guo H, Ingolia NT, Weissman JS, Bartel DP. 2010. Mammalian microRNAs predominantly act to decrease target mRNA levels. *Nature 466*:835-840.

Guo L, Lu Z. 2010. The Fate of miRNA* Strand through Evolutionary Analysis: Implication for Degradation As Merely Carrier Strand or Potential Regulatory Molecule? *PLoS ONE 5*:e11387.

Gurtan AM, Sharp PA. 2013. The Role of miRNAs in Regulating Gene Expression Networks. *Journal of Molecular Biology 425*:3582-3600.

Ha M, Kim VN. 2014. Regulation of microRNA biogenesis. *Nature Reviews Molecular Cell Biology 15*:509-524.

Hafner M, Landthaler M, Burger L, Khorshid M, Hausser J, Berninger P, Rothballer A, Ascano Jr M, Jungkamp A-C, Munschauer M, Ulrich A, Wardle GS, Dewell S, Zavolan M, Tuschl T. 2010. Transcriptome-wide Identification of RNA-Binding Protein and MicroRNA Target Sites by PAR-CLIP. *Cell 141*:129-141.

Hagan JP, Piskounova E, Gregory RI. 2009. Lin28 recruits the TUTase Zcchc11 to inhibit let-7 maturation in mouse embryonic stem cells. *Nat Struct Mol Biol 16*:1021-1025.

Hammell CM, Karp X, Ambros V. 2009a. A feedback circuit involving let-7-family miRNAs and DAF-12 integrates environmental signals and developmental timing in Caenorhabditis elegans. *Proceedings of the National Academy of Sciences 106*:18668-18673.

Hammell CM, Lubin I, Boag PR, Blackwell TK, Ambros V. 2009b. nhl-2 Modulates MicroRNA Activity in Caenorhabditis elegans. *Cell 136*:926-938.

Hammell M, Long D, Zhang L, Lee A, Carmack CS, Han M, Ding Y, Ambros V. 2008. mirWIP: microRNA target prediction based on microRNA-containing ribonucleoprotein-enriched transcripts. *Nat Meth 5*:813-819.

Hammond SM, Bernstein E, Beach D, Hannon GJ. 2000. An RNA-directed nuclease mediates post-transcriptional gene silencing in Drosophila cells. *Nature 404*:293-296.

Han J, Lee Y, Yeom K-H, Kim Y-K, Jin H, Kim VN. 2004. The Drosha-DGCR8 complex in primary microRNA processing. *Genes & Development 18*:3016-3027.

Han J, Lee Y, Yeom K-H, Nam J-W, Heo I, Rhee J-K, Sohn SY, Cho Y, Zhang B-T, Kim VN. 2006. Molecular Basis for the Recognition of Primary microRNAs by the Drosha-DGCR8 Complex. *Cell 125*:887-901.

Hansen TB, Jensen TI, Clausen BH, Bramsen JB, Finsen B, Damgaard CK, Kjems J. 2013. Natural RNA circles function as efficient microRNA sponges. *Nature 495*:384-388.

Harris TW, Baran J, Bieri T, Cabunoc A, Chan J, Chen WJ, Davis P, Done J, Grove C, Howe K, Kishore R, Lee R, Li Y, Muller H-M, Nakamura C, Ozersky P, Paulini M, Raciti D, Schindelman G, Tuli MA, Auken KV, Wang

D, Wang X, Williams G, Wong JD, Yook K, Schedl T, Hodgkin J, Berriman M, Kersey P, Spieth J, Stein L, Sternberg PW. 2014. WormBase 2014: new views of curated biology. *Nucleic Acids Research 42*:D789-D793.

Hendrickson DG, Hogan DJ, Herschlag D, Ferrell JE, Brown PO. 2008. Systematic Identification of mRNAs Recruited to Argonaute 2 by Specific microRNAs and Corresponding Changes in Transcript Abundance. *PLoS ONE 3*:e2126.

Hendriks G-J, Gaidatzis D, Aeschimann F, Großhans H. 2014. Extensive Oscillatory Gene Expression during *C. elegans* Larval Development. *Molecular Cell 53*:380-392.

Heo I, Joo C, Cho J, Ha M, Han J, Kim VN. 2008. Lin28 Mediates the Terminal Uridylation of let-7 Precursor MicroRNA. *Molecular Cell 32*:276-284.

Heo I, Joo C, Kim Y-K, Ha M, Yoon M-J, Cho J, Yeom K-H, Han J, Kim VN. 2009. TUT4 in Concert with Lin28 Suppresses MicroRNA Biogenesis through Pre-MicroRNA Uridylation. *Cell 138*:696-708.

Hinton A, Afrikanova I, Wilson M, King CC, Maurer B, Yeo GW, Hayek A, Pasquinelli AE, doi:10.1089/scd.2009.0224. 2010. A Distinct MicroRNA Signature for Definitive Endoderm Derived From Human Embryonic Stem Cells. *Stem Cells and Development 19*:797-807.

Huang JC, Babak T, Corson TW, Chua G, Khan S, Gallie BL, Hughes TR, Blencowe BJ, Frey BJ, Morris QD. 2007. Using expression profiling data to identify human microRNA targets. *Nat Meth 4*:1045-1049.

Huntzinger E, Izaurralde E. 2011. Gene silencing by microRNAs: contributions of translational repression and mRNA decay. *Nat Rev Genet 12*:99-110.

Hutten S, Kehlenbach RH. 2007. CRM1-mediated nuclear export: to the pore and beyond. *Trends in Cell Biology 17*:193-201.

Incarnato D, Neri F, Anselmi F, Oliviero S. 2014. Genome-wide profiling of mouse RNA secondary structures reveals key features of the mammalian transcriptome. *Genome Biology 15*:491.

Ivey KN, Srivastava D. 2010. MicroRNAs as Regulators of Differentiation and Cell Fate Decisions. *Cell Stem Cell 7*:36-41.

Janas MM, Khaled M, Schubert S, Bernstein JG, Golan D, Veguilla RA, Fisher DE, Shomron N, Levy C, Novina CD. 2011. Feed-Forward Microprocessing and Splicing Activities at a MicroRNA-Containing Intron. *PLoS Genetics 7*:e1002330.

Jannot G, Boisvert M-EL, Banville IH, Simard MJ. 2008. Two molecular features contribute to the Argonaute specificity for the microRNA and RNAi pathways in C. elegans. *RNA 14*:829-835.

Jiang Q, Wang Y, Hao Y, Juan L, Teng M, Zhang X, Li M, Wang G, Liu Y. 2009. miR2Disease: a manually curated database for microRNA deregulation in human disease. *Nucleic Acids Research 37*:D98-D104.

John B, Enright AJ, Aravin A, Tuschl T, Sander C, Marks DS. 2004. Human MicroRNA Targets. *PLoS Biol 2*:e363.

Johnson CD, Esquela-Kerscher A, Stefani G, Byrom M, Kelnar K, Ovcharenko D, Wilson M, Wang X, Shelton J, Shingara J, Chin L, Brown D, Slack FJ. 2007. The let-7 MicroRNA Represses Cell Proliferation Pathways in Human Cells. *Cancer Research 67*:7713-7722.

Johnson SM, Lin S-Y, Slack FJ. 2003. The time of appearance of the C. elegans let-7 microRNA is transcriptionally controlled utilizing a temporal regulatory element in its promoter. *Developmental Biology 259*:364-379.

Jovanovic M, Reiter L, Picotti P, Lange V, Bogan E, A Hurschler B, Blenkiron C, J Lehrbach N, C Ding X, Weiss M, P Schrimpf S, A Miska E, Groszhans H, Aebersold R, O Hengartner M. 2010. A quantitative targeted proteomics approach to validate predicted microRNA targets in C. elegans. *Nat Meth 7*:837-842.

Kai ZS, Finnegan EF, Huang S, Pasquinelli AE. 2012. Multiple cis-elements and trans-acting factors regulate dynamic spatio-temporal transcription of let-7 in C. elegans. *Developmental Biology*.

Kai ZS, Pasquinelli AE. 2010. MicroRNA assassins: factors that regulate the disappearance of miRNAs. *Nat Struct Mol Biol 17*:5-10.

Kanamoto T, Terada K, Yoshikawa H, Furukawa T. 2006. Cloning and regulation of the vertebrate homologue of lin-41 that functions as a heterochronic gene in Caenorhabditis elegans. *Developmental dynamics 235*:1142-1149.

Kanellopoulou C, Muljo SA, Kung AL, Ganesan S, Drapkin R, Jenuwein T, Livingston DM, Rajewsky K. 2005. Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. *Genes & Development 19*:489-501.

Karginov FV, Conaco C, Xuan Z, Schmidt BH, Parker JS, Mandel G, Hannon GJ. 2007. A biochemical approach to identifying microRNA targets. *Proceedings of the National Academy of Sciences 104*:19291-19296.

Karreth FA, Reschke M, Ruocco A, Ng C, Chapuy B, Léopold V, Sjoberg M, Keane TM, Verma A, Ala U, Tay Y, Wu D, Seitzer N, Velasco-Herrera Martin Del C, Bothmer A, Fung J, Langellotto F, Rodig Scott J, Elemento O, Shipp Margaret A, Adams David J, Chiarle R, Pandolfi Pier P. 2015. The BRAF Pseudogene Functions as a Competitive Endogenous RNA and Induces Lymphoma In Vivo. *Cell 161*:319-332.

Kataoka N, Fujita M, Ohno M. 2009. Functional association of the Microprocessor complex with the spliceosome. *Mol Cell Biol 29*:3243 - 3254.

Kato M, de Lencastre A, Pincus Z, Slack F. 2009. Dynamic expression of small non-coding RNAs, including novel microRNAs and piRNAs/21U-RNAs, during Caenorhabditis elegans development. *Genome Biology 10*:R54.

Kertesz M, Iovino N, Unnerstall U, Gaul U, Segal E. 2007. The role of site accessibility in microRNA target recognition. *Nature Genetics 39*:1278.

Kertesz M, Wan Y, Mazor E, Rinn JL, Nutter RC, Chang HY, Segal E. 2010. Genome-wide measurement of RNA secondary structure in yeast. *Nature 467*:103-107.

Kim Dh, Grun D, van Oudenaarden A. 2013. Dampening of expression oscillations by synchronous regulation of a microRNA and its target. *Nat Genet 45*:1337-1344.

Kim HH, Kuwano Y, Srikantan S, Lee EK, Martindale JL, Gorospe M. 2009a. HuR recruits let-7/RISC to repress c-Myc expression. *Genes & Development 23*:1743-1748.

Kim VN, Han J, Siomi MC. 2009b. Biogenesis of small RNAs in animals. *Nat Rev Mol Cell Biol 10*:126-139.

Kim YK, Kim VN. 2007. Processing of intronic microRNAs. *The EMBO Journal 26*:775-783.

Kloosterman WP, Wienholds E, Ketting RF, Plasterk RHA. 2004. Substrate requirements for let-7 function in the developing zebrafish embryo. *Nucleic Acids Research 32*:6284-6291.

Kozomara A, Griffiths-Jones S. 2014. miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Research 42*:D68-D73.

Krek A, Grun D, Poy MN, Wolf R, Rosenberg L, Epstein EJ, MacMenamin P, da Piedade I, Gunsalus KC, Stoffel M, Rajewsky N. 2005. Combinatorial microRNA target predictions. *Nat Genet 37*:495-500.

Krol J, Loedige I, Filipowicz W. 2010. The widespread regulation of microRNA biogenesis, function and decay. *Nat Rev Genet 11*:597-610.

Krol J, Sobczak K, Wilczynska U, Drath M, Jasinska A, Kaczynska D, Krzyzosiak WJ. 2004. Structural Features of MicroRNA (miRNA) Precursors and Their Relevance to miRNA Biogenesis and Small Interfering RNA/Short Hairpin RNA Design. *Journal of Biological Chemistry 279*:42230-42239.

Kumar MS, Erkeland SJ, Pester RE, Chen CY, Ebert MS, Sharp PA, Jacks T. 2008. Suppression of non-small cell lung tumor development by the let-7 microRNA family. *Proceedings of the National Academy of Sciences 105*:3903-3908.

Lagos-Quintana M, Rauhut R, Lendeckel W, Tuschl T. 2001. Identification of Novel Genes Coding for Small Expressed RNAs. *Science 294*:853-858.

Lagos-Quintana M, Rauhut R, Yalcin A, Meyer J, Lendeckel W, Tuschl T. 2002. Identification of Tissue-Specific MicroRNAs from Mouse. *Current Biology 12*:735-739.

Lai EC. 2005. miRNAs: Whys and Wherefores of miRNA-Mediated Regulation. *Current Biology 15*:R458-R460.

Lall S, Grün D, Krek A, Chen K, Wang Y-L, Dewey CN, Sood P, Colombo T, Bray N, MacMenamin P, Kao H-L, Gunsalus KC, Pachter L, Piano F, Rajewsky N. 2006. A Genome-Wide Map of Conserved MicroRNA Targets in C. elegans. *Current Biology 16*:460-471.

Landthaler M, Gaidatzis D, Rothballer A, Chen PY, Soll SJ, Dinic L, Ojo T, Hafner M, Zavolan M, Tuschl T. 2008. Molecular characterization of human Argonaute-containing ribonucleoprotein complexes and their bound target mRNAs. *RNA 14*:2580-2596.

Lau NC, Lim LP, Weinstein EG, Bartel DP. 2001. An Abundant Class of Tiny RNAs with Probable Regulatory Roles in Caenorhabditis elegans. *Science 294*:858-862.

Laurent LC. 2008. MicroRNAs in embryonic stem cells and early embryonic development. *Journal of Cellular and Molecular Medicine 12*:2181-2188.

Lee RC, Ambros V. 2001. An Extensive Class of Small RNAs in Caenorhabditis elegans. *Science 294*:862-864.

Lee RC, Feinbaum RL, Ambros V. 1993. The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell 75*:843-854.

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

Lee Y, Kim VN, John JR, Gregory JH. 2007. In Vitro and In Vivo Assays for the Activity of Drosha Complex. *Methods in Enzymology*: Academic Press. pp 87, 89-106.

Lee YS, Dutta A. 2007. The tumor suppressor microRNA let-7 represses the HMGA2 oncogene. *Genes & Development 21*:1025-1030.

Lehrbach NJ, Armisen J, Lightfoot HL, Murfitt KJ, Bugaut A, Balasubramanian S, Miska EA. 2009. LIN-28 and the poly(U) polymerase PUP-2 regulate let-7 microRNA processing in Caenorhabditis elegans. *Nat Struct Mol Biol 16*:1016-1020.

Lehrbach NJ, Castro C, Murfitt KJ, Abreu-Goodger C, Griffin JL, Miska EA. 2012. Post-developmental microRNA expression is required for normal physiology, and regulates aging in parallel to insulin/IGF-1 signaling in C. elegans. *RNA 18*:2220-2235.

Leung AKL, Young AG, Bhutkar A, Zheng GX, Bosson AD, Nielsen CB, Sharp PA. 2010. Genome-wide identification of Ago2 binding sites from mouse

embryonic stem cells with and without mature microRNAs. *Nat Struct Mol Biol 18*:237-244.

Lewis BP, Burge CB, Bartel DP. 2005. Conserved Seed Pairing, Often Flanked by Adenosines, Indicates that Thousands of Human Genes are MicroRNA Targets. *Cell 120*:15-20.

Lewis BP, Shih Ih, Jones-Rhoades MW, Bartel DP, Burge CB. 2003. Prediction of Mammalian MicroRNA Targets. *Cell 115*:787-798.

Li H, Wang J, Mor G, Sklar J. 2008. A Neoplastic Gene Fusion Mimics Trans-Splicing of RNAs in Normal Human Cells. *Science 321*:1357-1361.

Li J-H, Liu S, Zhou H, Qu L-H, Yang J-H. 2014. starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. *Nucleic Acids Research 42*:D92-D97.

Li M, Jones-Rhoades MW, Lau NC, Bartel DP, Rougvie AE. 2005. Regulatory Mutations of mir-48, a C. elegans let-7 Family MicroRNA, Cause Developmental Timing Defects. *Developmental Cell 9*:415-422.

Lim LP, Glasner ME, Yekta S, Burge CB, Bartel DP. 2003. Vertebrate MicroRNA Genes. *Science 299*:1540.

Lim LP, Lau NC, Garrett-Engele P, Grimson A, Schelter JM, Castle J, Bartel DP, Linsley PS, Johnson JM. 2005. Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. *Nature 433*:769 - 773.

Lin S-Y, Johnson SM, Abraham M, Vella MC, Pasquinelli A, Gamberi C, Gottlieb E, Slack FJ. 2003. The C. elegans hunchback Homolog, hbl-1, Controls Temporal Patterning and Is a Probable MicroRNA Target. *Developmental Cell 4*:639-650.

Lin S, Gregory RI. 2015. MicroRNA biogenesis pathways in cancer. *Nat Rev Cancer 15*:321-333.

Liu J, Carmell MA, Rivas FV, Marsden CG, Thomson JM, Song J-J, Hammond SM, Joshua-Tor L, Hannon GJ. 2004. Argonaute2 Is the Catalytic Engine of Mammalian RNAi. *Science 305*:1437-1441.

Liu S, Li J-H, Wu J, Zhou K-R, Zhou H, Yang J-H, Qu L-H. 2015. StarScan: a web server for scanning small RNA targets from degradome sequencing data. *Nucleic Acids Research 43*:W480-W486.

Liz J, Portela A, Soler M, Gómez A, Ling H, Michlewski G, Calin George A, Guil S, Esteller M. 2014. Regulation of pri-miRNA Processing by a Long Noncoding RNA Transcribed from an Ultraconserved Region. *Molecular Cell 55*:138-147.

Lund E, Guttinger S, Calado A, Dahlberg JE, Kutay U. 2004. Nuclear Export of MicroRNA Precursors. *Science 303*:95-98.

Lutter D, Marr C, Krumsiek J, Lang E, Theis F. 2010. Intronic microRNAs support their host genes by mediating synergistic and antagonistic regulatory effects. *BMC Genomics 11*:224.

Macias S, Plass M, Stajuda A, Michlewski G, Eyras E, Cáceres JF. 2012. DGCR8 HITS-CLIP reveals novel functions for the Microprocessor. *Nature Structural & Molecular Biology 19*:760-766.

Maller-Schulman BR, Liang X, Stahlhut C, DelConte C, Stefani G, Slack FJ. 2008. The let-7 microRNA target gene, Mlin41/Trim71 is required for mouse embryonic survival and neural tube closure. *Cell Cycle 7*:3935-3942.

Martinez NJ, Ow MC, Reece-Hoyes JS, Barrasa MI, Ambros VR, Walhout AJM. 2008. Genome-scale spatiotemporal analysis of Caenorhabditis elegans microRNA promoter activity. *Genome Research 18*:2005-2015.

Massirer KB, Perez SG, Mondol V, Pasquinelli AE. 2012. The miR-35-41 Family of MicroRNAs Regulates RNAi Sensitivity in *Caenorhabditis elegans*. *PLoS Genetics 8*:e1002536.

Mathonnet Gr, Fabian MR, Svitkin YV, Parsyan A, Huck L, Murata T, Biffo S, Merrick WC, Darzynkiewicz E, Pillai RS, Filipowicz W, Duchaine TF, Sonenberg N. 2007. MicroRNA Inhibition of Translation Initiation in Vitro by Targeting the Cap-Binding Complex eIF4F. *Science 317*:1764-1767.

Mattioli C, Pianigiani G, Pagani F. 2013. A competitive regulatory mechanism discriminates between juxtaposed splice sites and pri-miRNA structures. *Nucleic Acids Research 41*:8680-8691.

Mayr C, Hemann MT, Bartel DP. 2007. Disrupting the Pairing Between let-7 and Hmga2 Enhances Oncogenic Transformation. *Science 315*:1576-1579.

Meijer HA, Kong YW, Lu WT, Wilczynska A, Spriggs RV, Robinson SW, Godfrey JD, Willis AE, Bushell M. 2013. Translational Repression and eIF4A2 Activity Are Critical for MicroRNA-Mediated Gene Regulation. *Science 340*:82-85.

Melamed Ze, Levy A, Ashwal-Fluss R, Lev-Maor G, Mekahel K, Atias N, Gilad S, Sharan R, Levy C, Kadener S, Ast G. 2013. Alternative Splicing Regulates Biogenesis of miRNAs Located across Exon-Intron Junctions. *Molecular Cell 50*:869-881.

Melton C, Blelloch R. 2010. MicroRNA Regulation of Embryonic Stem Cell Self-Renewal and Differentiation. In: Meshorer E, Plath K, eds. *The Cell Biology of Stem Cells*: Springer US. pp 105-117.

Memczak S, Jens M, Elefsinioti A, Torti F, Krueger J, Rybak A, Maier L, Mackowiak SD, Gregersen LH, Munschauer M, Loewer A, Ziebold U, Landthaler M, Kocks C, le Noble F, Rajewsky N. 2013. Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature 495*:333-338.

Meneely PM, Herman RK. 1979. Lethals, steriles and deficiencies in a region of the X chromosome of Caenorhabditis elegans. *Genetics 92*:99-115.

Michlewski G, Guil S, Semple CA, Caceres JF. 2008. Posttranscriptional regulation of miRNAs harboring conserved terminal loops. *Molecular Cell 32*:383-393.

Miranda KC, Huynh T, Tay Y, Ang Y-S, Tam W-L, Thomson AM, Lim B, Rigoutsos I. 2006. A Pattern-Based Method for the Identification of MicroRNA Binding Sites and Their Corresponding Heteroduplexes. *Cell 126*:1203-1217.

Miska EA, Alvarez-Saavedra E, Abbott AL, Lau NC, Hellman AB, McGonagle SM, Bartel DP, Ambros VR, Horvitz HR. 2007. Most *Caenorhabditis elegans* microRNAs Are Individually Not Essential for Development or Viability. *PLoS Genet 3*:e215.

Mondol V, Ahn BC, Pasquinelli AE. 2015. Splicing remodels the let-7 primary microRNA to facilitate Drosha processing in *Caenorhabditis elegans*. *RNA 21*.

Mondol V, Pasquinelli AE. 2012. Let's Make It Happen: The Role of let-7 MicroRNA in Development. In: Eran H, ed. *Current Topics in Developmental Biology*: Academic Press. pp 1-30.

Morlando M, Ballarino M, Gromak N, Pagano F, Bozzoni I, Proudfoot NJ. 2008. Primary microRNA transcripts are processed co-transcriptionally. *Nat Struct Mol Biol 15*:902-909.

Moss EG, Lee RC, 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.

Murchison EP, Partridge JF, Tam OH, Cheloufi S, Hannon GJ. 2005. Characterization of Dicer-deficient murine embryonic stem cells. *Proceedings of the National Academy of Sciences of the United States of America 102*:12135-12140.

Murphy D, Dancis B, Brown J. 2008. The evolution of core proteins involved in microRNA biogenesis. *BMC Evolutionary Biology 8*:92.

Newman MA, Thomson JM, Hammond SM. 2008. Lin-28 interaction with the Let-7 precursor loop mediates regulated microRNA processing. *RNA 14*:1539-1549.

Nguyen Tuan A, Jo Myung H, Choi Y-G, Park J, Kwon CS, Hohng S, Kim NV, Woo J-S. 2015. Functional Anatomy of the Human Microprocessor. *Cell 161*:1374-1387.

Nimmo R, Slack F. 2009. An elegant miRror: microRNAs in stem cells, developmental timing and cancer. *Chromosoma 118*:405-418.

Nishino J, Kim I, Chada K, Morrison SJ. 2008. Hmga2 Promotes Neural Stem Cell Self-Renewal in Young but Not Old Mice by Reducing p16Ink4a and p19Arf Expression. *Cell 135*:227-239.

Niwa R, Hada K, Moliyama K, Ohniwa RL, Tan YM, Carter KO, Chi W, Reinke V, Slack FJ. 2009. *C. elegans sym-1* is a downstream target of the

Hunchback-like-1 developmental timing transcription factor. *Cell Cycle 8*:4147-4154.

Nour-Eldin HH, Geu-Flores F, Halkier BA. 2010. USER Cloning and USER Fusion: The Ideal Cloning Techniques for Small and Big Laboratories. In: Fett-Neto AG, ed. *Plant Secondary Metabolism Engineering*: Humana Press. pp 185-200.

Nousbeck J, Spiegel R, Ishida-Yamamoto A, Indelman M, Shani-Adir A, Adir N, Lipkin E, Bercovici S, Geiger D, van Steensel MA, Steijlen PM, Bergman R, Bindereif A, Choder M, Shalev S, Sprecher E. 2008. Alopecia, Neurological Defects, and Endocrinopathy Syndrome Caused by Decreased Expression of RBM28, a Nucleolar Protein Associated with Ribosome Biogenesis. *The American Journal of Human Genetics 82*:1114-1121.

Okamura K, Hagen JW, Duan H, Tyler DM, Lai EC. 2007. The Mirtron Pathway Generates microRNA-Class Regulatory RNAs in Drosophila. *Cell 130*:89-100.

Okamura K, Phillips MD, Tyler DM, Duan H, Chou Y-t, Lai EC. 2008. The regulatory activity of microRNA* species has substantial influence on microRNA and 3' UTR evolution. *Nat Struct Mol Biol 15*:354-363.

Olsen PH, 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.

Osada H, Takahashi T. 2011. let-7 and miR-17-92: Small-sized major players in lung cancer development. *Cancer Science 102*:9-17.

Ow MC, Martinez NJ, Olsen PH, Silverman HS, Barrasa MI, Conradt B, Walhout AJM, Ambros V. 2008. The FLYWCH transcription factors FLH-1, FLH-2, and FLH-3 repress embryonic expression of microRNA genes in C. elegans. *Genes & Development 22*:2520-2534.

Ozsolak F, Poling LL, Wang Z, Liu H, Liu XS, Roeder RG, Zhang X, Song JS, Fisher DE. 2008. Chromatin structure analyses identify miRNA promoters. *Genes Dev 22*:3172 - 3183.

Pasquinelli AE. 2012. MicroRNAs and their targets: recognition, regulation and an emerging reciprocal relationship. *Nat Rev Genet 13*:271-282.

Pasquinelli AE, Reinhart BJ, Slack F, Martindale MQ, Kuroda MI, Maller B, Hayward DC, Ball EE, Degnan B, Muller P, Spring J, Srinivasan A, Fishman M, Finnerty J, Corbo J, Levine M, Leahy P, Davidson E, Ruvkun G. 2000. Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. *Nature 408*:86-89.

Poliseno L, Salmena L, Zhang J, Carver B, Haveman WJ, Pandolfi PP. 2010. A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. *Nature 465*:1033-1038.

Potenza N, Russo A. 2013. Biogenesis, evolution and functional targets of microRNA-125a. *Molecular Genetics and Genomics 288*:381-389.

Quick-Cleveland J, Jacob Jose P, Weitz Sara H, Shoffner G, Senturia R, Guo F. 2014. The DGCR8 RNA-Binding Heme Domain Recognizes Primary MicroRNAs by Clamping the Hairpin. *Cell Reports 7*:1994-2005.

Ramalingam P, Palanichamy JK, Singh A, Das P, Bhagat M, Kassab MA, Sinha S, Chattopadhyay P. 2014. Biogenesis of intronic miRNAs located in clusters by independent transcription and alternative splicing. *RNA 20*:76-87.

Reinhart BJ, Slack FJ, Basson M, Pasquinelli AE, Bettinger JC, Rougvie AE, Horvitz HR, Ruvkun G. 2000. The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. *Nature 403*:901-906.

Reinhart BJ, Weinstein EG, Rhoades MW, Bartel B, Bartel DP. 2002. MicroRNAs in plants. *Genes & Development 16*:1616-1626.

Roush S, Slack FJ. 2008. The let-7 family of microRNAs. *Trends in Cell Biology 18*:505-516.

Roush SF, Slack FJ. 2009. Transcription of the C. elegans let-7 microRNA is temporally regulated by one of its targets, hbl-1. *Developmental Biology 334*:523-534.

Roy-Chaudhuri B, Valdmanis PN, Zhang Y, Wang Q, Luo Q-J, Kay MA. 2014. Regulation of microRNA-mediated gene silencing by microRNA precursors. *Nat Struct Mol Biol 21*:825-832.

Ruby JG, Jan C, Player C, Axtell MJ, Lee W, Nusbaum C, Ge H, Bartel DP. 2006. Large-Scale Sequencing Reveals 21U-RNAs and Additional MicroRNAs and Endogenous siRNAs in C. elegans. *Cell 127*:1193-1207.

Ruby JG, Jan CH, Bartel DP. 2007a. Intronic microRNA precursors that bypass Drosha processing. *Nature 448*:83-86.

Ruby JG, Stark A, Johnston WK, Kellis M, Bartel DP, Lai EC. 2007b. Evolution, biogenesis, expression, and target predictions of a substantially expanded set of Drosophila microRNAs. *Genome Research 17*:1850-1864.

Ruvkun G, Giusto J. 1989. The Caenorhabditis elegans heterochronic gene lin-14 encodes a nuclear protein that forms a temporal developmental switch. *Nature 338*:313-319.

Rybak A, Fuchs H, Hadian K, Smirnova L, Wulczyn EA, Michel G, Nitsch R, Krappmann D, Wulczyn FG. 2009. The let-7 target gene mouse lin-41 is a stem cell specific E3 ubiquitin ligase for the miRNA pathway protein Ago2. *Nat Cell Biol 11*:1411-1420.

Rybak A, Fuchs H, Smirnova L, Brandt C, Pohl EE, Nitsch R, Wulczyn FG. 2008. A feedback loop comprising lin-28 and let-7 controls pre-let-7 maturation during neural stem-cell commitment. *Nat Cell Biol 10*:987-993.

Saini HK, Griffiths-Jones S, Enright AJ. 2007. Genomic analysis of human microRNA transcripts. *Proceedings of the National Academy of Sciences 104*:17719-17724.

Salomon William E, Jolly Samson M, Moore Melissa J, Zamore Phillip D, Serebrov V. 2015. Single-Molecule Imaging Reveals that Argonaute Reshapes the Binding Properties of Its Nucleic Acid Guides. *Cell 162*:84-95.

Schirle NT, MacRae IJ. 2012. The Crystal Structure of Human Argonaute2. *Science 336*:1037-1040.

Schirle NT, Sheu-Gruttadauria J, MacRae IJ. 2014. Structural basis for microRNA targeting. *Science 346*:608-613.

Schmitter D, Filkowski J, Sewer A, Pillai RS, Oakeley EJ, Zavolan M, Svoboda P, Filipowicz W. 2006. Effects of Dicer and Argonaute down-regulation on

mRNA levels in human HEK293 cells. *Nucleic Acids Research 34*:4801-4815.

Schwab R, Speth C, Laubinger S, Voinnet O. 2013. Enhanced microRNA accumulation through stemloop-adjacent introns. *EMBO Reports 14*:615-621.

Schwamborn JC, Berezikov E, Knoblich JA. 2009. The TRIM-NHL Protein TRIM32 Activates MicroRNAs and Prevents Self-Renewal in Mouse Neural Progenitors. *Cell 136*:913-925.

Selbach M, Schwanhausser B, Thierfelder N, Fang Z, Khanin R, Rajewsky N. 2008. Widespread changes in protein synthesis induced by microRNAs. *Nature 455*:58 - 63.

Sempere L, Freemantle S, Pitha-Rowe I, Moss E, Dmitrovsky E, Ambros V. 2004. Expression profiling of mammalian microRNAs uncovers a subset of brain-expressed microRNAs with possible roles in murine and human neuronal differentiation. *Genome Biology 5*:R13.

Shell S, Park S-M, Radjabi AR, Schickel R, Kistner EO, Jewell DA, Feig C, Lengyel E, Peter ME. 2007. Let-7 expression defines two differentiation stages of cancer. *Proceedings of the National Academy of Sciences 104*:11400-11405.

Shen J, Hung M-C. 2015. Signaling-Mediated Regulation of MicroRNA Processing. *Cancer Research 75*:783-791.

Shin C, Nam J-W, Farh KK-H, Chiang HR, Shkumatava A, Bartel DP. 2010. Expanding the MicroRNA Targeting Code: Functional Sites with Centered Pairing. *Molecular Cell 38*:789-802.

Shostak Y, Van Gilst MR, Antebi A, Yamamoto KR. 2004. Identification of C. elegans DAF-12-binding sites, response elements, and target genes. *Genes & Development 18*:2529-2544.

Slack FJ, Basson M, Liu Z, Ambros V, Horvitz HR, 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.

Slack FJ, Ruvkun G. 1998. A novel repeat domain that is often associated with RING finger and B-box motifs. *Trends in Biochemical Sciences 23*:474-475.

Smirnova L, Gräfe A, Seiler A, Schumacher S, Nitsch R, Wulczyn FG. 2005. Regulation of miRNA expression during neural cell specification. *European Journal of Neuroscience 21*:1469-1477.

Sohail M, Xie J. 2015. Diverse regulation of 3′ splice site usage. *Cellular and Molecular Life Sciences 72*:4771-4793.

Sokol NS, Xu P, Jan Y-N, Ambros V. 2008. Drosophila let-7 microRNA is required for remodeling of the neuromusculature during metamorphosis. *Genes & Development 22*:1591-1596.

Sood P, Krek A, Zavolan M, Macino G, Rajewsky N. 2006. Cell-type-specific signatures of microRNAs on target mRNA expression. *Proceedings of the National Academy of Sciences of the United States of America 103*:2746-2751.

Stefani G, Chen X, Zhao H, Slack Fj. 2015. A novel mechanism of LIN-28 regulation of *let-7* microRNA expression revealed by in vivo HITS-CLIP in *C. elegans*. *RNA 21*.

Steiner FA, Hoogstrate SW, Okihara KL, Thijssen KL, Ketting RF, Plasterk RHA, Sijen T. 2007. Structural features of small RNA precursors determine Argonaute loading in Caenorhabditis elegans. *Nat Struct Mol Biol 14*:927-933.

Storz G. 2002. An Expanding Universe of Noncoding RNAs. *Science 296*:1260-1263.

Suh M-R, Lee Y, Kim JY, Kim S-K, Moon S-H, Lee JY, Cha K-Y, Chung HM, Yoon HS, Moon SY, Kim VN, Kim K-S. 2004. Human embryonic stem cells express a unique set of microRNAs. *Developmental Biology 270*:488-498.

Sulston JE, Horvitz HR. 1977. Post-embryonic cell lineages of the nematode, Caenorhabditis elegans. *Developmental Biology 56*:110-156.

Sun C, Woolford JL. 1994. The yeast NOP4 gene product is an essential nucleolar protein required for pre-rRNA processing and accumulation of 60S ribosomal subunits. *The EMBO Journal 13*:3127-3135.

Sun C, Woolford JL. 1997. The Yeast Nucleolar Protein Nop4p Contains Four RNA Recognition Motifs Necessary for Ribosome Biogenesis. *Journal of Biological Chemistry 272*:25345-25352.

Takamizawa J, Konishi H, Yanagisawa K, Tomida S, Osada H, Endoh H, Harano T, Yatabe Y, Nagino M, Nimura Y, Mitsudomi T, Takahashi T. 2004. Reduced Expression of the let-7 MicroRNAs in Human Lung Cancers in Association with Shortened Postoperative Survival. *Cancer Research 64*:3753-3756.

Tamburino AM, Ryder SP, Walhout AJM. 2013. A Compendium of Caenorhabditis elegans RNA Binding Proteins Predicts Extensive Regulation at Multiple Levels. *G3: Genes|Genomes|Genetics 3*:297-304.

Tang R, Li L, Zhu D, Hou D, Cao T, Gu H, Zhang J, Chen J, Zhang C-Y, Zen K. 2012. Mouse miRNA-709 directly regulates miRNA-15a/16-1 biogenesis at the posttranscriptional level in the nucleus: evidence for a microRNA hierarchy system. *Cell Research 22*:504-515.

Tritschler F, Huntzinger E, Izaurralde E. 2010. Role of GW182 proteins and PABPC1 in the miRNA pathway: a sense of deja vu. *Nat Rev Mol Cell Biol 11*:379-384.

Trujillo RD, Yue S-B, Tang Y, O'Gorman WE, Chen C-Z. 2010. The potential functions of primary microRNAs in target recognition and repression. *The EMBO Journal 29*:3272-3285.

Truscott M, Islam ABMMK, Frolov MV. 2015. Novel regulation and functional interaction of polycistronic miRNAs. *RNA 22*:1-10.

Van Wynsberghe PM, Kai ZS, Massirer KB, Burton VH, Yeo GW, Pasquinelli AE. 2011. LIN-28 co-transcriptionally binds primary let-7 to regulate miRNA maturation in Caenorhabditis elegans. *Nat Struct Mol Biol 18*:302-308.

Vella MC, Choi E-Y, Lin S-Y, Reinert K, Slack FJ. 2004a. The C. elegans microRNA let-7 binds to imperfect let-7 complementary sites from the lin-41 3'UTR. *Genes & Development 18*:132-137.

Vella MC, Reinert K, Slack FJ. 2004b. Architecture of a Validated MicroRNA::Target Interaction. *Chemistry & Biology 11*:1619-1623.

Viswanathan SR, Daley GQ. 2010. Lin28: A MicroRNA Regulator with a Macro Role. *Cell 140*:445-449.

Viswanathan SR, Daley GQ, Gregory RI. 2008. Selective Blockade of MicroRNA Processing by Lin28. *Science 320*:97-100.

Vlachos IS, Paraskevopoulou MD, Karagkouni D, Georgakilas G, Vergoulis T, Kanellos I, Anastasopoulos I-L, Maniou S, Karathanou K, Kalfakakou D, Fevgas A, Dalamagas T, Hatzigeorgiou AG. 2015. DIANA-TarBase v7.0: indexing more than half a million experimentally supported miRNA:mRNA interactions. *Nucleic Acids Research 43*:D153-D159.

Wallace A, Filbin ME, Veo B, McFarland C, Stepinski J, Jankowska-Anyszka M, Darzynkiewicz E, Davis RE. 2010. The Nematode Eukaryotic Translation Initiation Factor 4E/G Complex Works with a *trans*-Spliced Leader Stem-Loop To Enable Efficient Translation of Trimethylguanosine-Capped RNAs. *Molecular and Cellular Biology 30*:1958-1970.

Wan Y, Kertesz M, Spitale RC, Segal E, Chang HY. 2011. Understanding the transcriptome through RNA structure. *Nat Rev Genet 12*:641-655.

Wan Y, Qu K, Ouyang Z, Chang HY. 2016. Genome-Wide Probing of RNA Structures In Vitro Using Nucleases and Deep Sequencing. In: Devaux F, ed. *Yeast Functional Genomics*: Springer New York. pp 141-160.

Wan Y, Qu K, Zhang QC, Flynn RA, Manor O, Ouyang Z, Zhang J, Spitale RC, Snyder MP, Segal E, Chang HY. 2014. Landscape and variation of RNA secondary structure across the human transcriptome. *Nature 505*:706-709.

Wang K, Sun T, Li N, Wang Y, Wang J-X, Zhou L-Y, Long B, Liu C-Y, Liu F, Li P-F. 2014. MDRL lncRNA Regulates the Processing of miR-484 Primary Transcript by Targeting miR-361. *PLoS Genet 10*:e1004467.

Wang P, Zhi H, Zhang Y, Liu Y, Zhang J, Gao Y, Guo M, Ning S, Li X. 2015. miRSponge: a manually curated database for experimentally supported miRNA sponges and ceRNAs. *Database 2015*.

Wang Y, Medvid R, Melton C, Jaenisch R, Blelloch R. 2007. DGCR8 is essential for microRNA biogenesis and silencing of embryonic stem cell self-renewal. *Nat Genet 39*:380-385.

Warf MB, Johnson WE, Bass BL. 2011. Improved annotation of C. elegans microRNAs by deep sequencing reveals structures associated with processing by Drosha and Dicer. *RNA 17*:563-577.

Warshauer E, Samuelov L, Sarig O, Vodo D, Bindereif A, Kanaan M, Gat U, Fuchs-Telem D, Shomron N, Farberov L, Pasmanik-Chor M, Nardini G, Winkler E, Meilik B, Petit I, Aberdam D, Paus R, Sprecher E, Nousbeck J. 2015. RBM28, a protein deficient in ANE syndrome, regulates hair follicle growth via miR-203 and p63. *Experimental Dermatology 24*:618-622.

Westholm JO, Lai EC. 2011. Mirtrons: microRNA biogenesis via splicing. *Biochimie 93*:1897-1904.

Wightman B, Ha I, 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.

William DA, Saitta B, Gibson JD, Traas J, Markov V, Gonzalez DM, Sewell W, Anderson DM, Pratt SC, Rappaport EF, Kusumi K. 2007. Identification of oscillatory genes in somitogenesis from functional genomic analysis of a human mesenchymal stem cell model. *Dev Biol 305*:172 - 186.

Wilson Ross C, Tambe A, Kidwell Mary A, Noland Cameron L, Schneider Catherine P, Doudna Jennifer A. 2015. Dicer-TRBP Complex Formation Ensures Accurate Mammalian MicroRNA Biogenesis. *Molecular Cell 57*:397-407.

Wu E, Thivierge C, Flamand M, Mathonnet G, Vashisht AA, Wohlschlegel J, Fabian MR, Sonenberg N, Duchaine TF. 2010. Pervasive and Cooperative Deadenylation of 3'UTRs by Embryonic MicroRNA Families. *Molecular Cell 40*:558-570.

Wu L, Belasco JG. 2005. Micro-RNA Regulation of the Mammalian lin-28 Gene during Neuronal Differentiation of Embryonal Carcinoma Cells. *Mol Cell Biol 25*:9198-9208.

Wulczyn FG, Smirnova L, Rybak A, Brandt C, Kwidzinski E, Ninnemann O, Strehle M, Seiler A, Schumacher S, Nitsch R. 2007. Post-transcriptional

regulation of the let-7 microRNA during neural cell specification. *The FASEB Journal 21*:415-426.

Yi R, Qin Y, Macara IG, Cullen BR. 2003. Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes Dev 17*:3011 - 3016.

Yigit E, Batista PJ, Bei Y, Pang KM, Chen C-CG, Tolia NH, Joshua-Tor L, Mitani S, Simard MJ, Mello CC. 2006. Analysis of the C. elegans Argonaute Family Reveals that Distinct Argonautes Act Sequentially during RNAi. *Cell 127*:747-757.

Ying S-Y, Lin S-L. 2006. Current perspectives in intronic micro RNAs (miRNAs). *Journal of Biomedical Science 13*:5-15.

Ying S, Lin S. 2005. Intronic microRNAs. *Biochemical and Biophysical Research Communications 326*:515 - 520.

Yu Z, Jian Z, Shen S-H, Purisima E, Wang E. 2007. Global analysis of microRNA target gene expression reveals that miRNA targets are lower expressed in mature mouse and Drosophila tissues than in the embryos. *Nucleic Acids Research 35*:152-164.

Yue S-B, Trujillo RD, Tang Y, O'Gorman WE, Chen C-Z. 2011. Loop nucleotides control primary and mature miRNA function in target recognition and repression. *RNA Biology 8*:1115-1123.

Zamore PD, Tuschl T, Sharp PA, Bartel DP. 2000. RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. *Cell 101*:25-33.

Zaphiropoulos PG. 2011. Trans-splicing in Higher Eukaryotes: Implications for Cancer Development? *Frontiers in Genetics 2*:92.

Zeng Y, Cullen BR. 2005. Efficient Processing of Primary microRNA Hairpins by Drosha Requires Flanking Nonstructured RNA Sequences. *Journal of Biological Chemistry 280*:27595-27603.

Zeng Y, Yi R, Cullen BR. 2003. MicroRNAs and small interfering RNAs can inhibit mRNA expression by similar mechanisms. *Proceedings of the National Academy of Sciences 100*:9779-9784.

Zhang L, Ding L, Cheung TH, Dong M-Q, Chen J, Sewell AK, Liu X, Yates Iii JR, Han M. 2007. Systematic Identification of C. elegans miRISC Proteins, miRNAs, and mRNA Targets by Their Interactions with GW182 Proteins AIN-1 and AIN-2. *Molecular Cell 28*:598-613.

Zhang L, Hammell M, Kudlow BA, Ambros V, Han M. 2009. Systematic analysis of dynamic miRNA-target interactions during C. elegans development. *Development 136*:3043-3055.

Zhao C, Sun G, Li S, Lang M-F, Yang S, Li W, Shi Y. 2010. MicroRNA let-7b regulates neural stem cell proliferation and differentiation by targeting nuclear receptor TLX signaling. *Proceedings of the National Academy of Sciences 107*:1876-1881.

Zhong X, Li N, Liang S, Huang Q, Coukos G, Zhang L. 2010. Identification of MicroRNAs Regulating Reprogramming Factor LIN28 in Embryonic Stem Cells and Cancer Cells. *Journal of Biological Chemistry 285*:41961-41971.

Zisoulis DG, Kai ZS, Chang RK, Pasquinelli AE. 2012. Autoregulation of microRNA biogenesis by let-7 and Argonaute. *Nature 486*:541-544.

Zisoulis DG, Lovci MT, Wilbert ML, Hutt KR, Liang TY, Pasquinelli AE, Yeo GW. 2010. Comprehensive discovery of endogenous Argonaute binding sites in Caenorhabditis elegans. *Nat Struct Mol Biol 17*:173-179.

Zuker M. 2003. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Research 31*:3406-3415.