Analysis of lin-4 microRNA biogenesis and function in C. elegans

Author
Bracht, John Russell

Publication Date
2009

Peer reviewed|Thesis/dissertation
Analysis of \textit{lin}-4 microRNA biogenesis and function in \textit{C. elegans}

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

in

Biology

by

John Russell Bracht

Committee in charge:

Professor Amy E. Pasquinelli Chair
Professor Andrew Dillin
Professor Terry Gaasterland
Professor Tracy Johnson
Professor William McGinnis

2009
The Dissertation of John Russell Bracht is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

______________________________________________________________

______________________________________________________________

______________________________________________________________

______________________________________________________________

______________________________________________________________

Chair

University of California, San Diego

2009
Dedication

I dedicate this dissertation to my lovely wife Kate who is in every way my partner and best friend.
Epigraph

It is the glory of God to conceal a matter;
to search out a matter is the glory of kings.

Proverbs 25:2
# Table of Contents

Signature Page........................................................................................................ iii
Dedication................................................................................................................ iv
Epigraph.................................................................................................................. v
Table of Contents.................................................................................................... vi
List of Figures.......................................................................................................... vii
List of Tables.......................................................................................................... ix
Acknowledgements............................................................................................... x
Vita........................................................................................................................... xii
Abstract of the Dissertation.................................................................................... xiv
Introduction............................................................................................................. 1

Chapter 1: Regulation of \textit{lin-4} miRNA Processing, Organismal Growth and Development by a Conserved RNA Binding Protein Gene in \textit{C. elegans}........................................................................................................... 30

Chapter 2: Analysis of the Mechanism of Target Regulation by \textit{lin-4}..... 71

Conclusions and Future Directions....................................................................... 89

Appendix 1: Isolation and Characterization of a Novel \textit{Alg-1} Mutant...... 108

Appendix 2: Characterization and Genotyping of the \textit{lin-4(e192)} Rearrangement........................................................................................................ 119

Appendix 3: Plasmids and Strains Generated..................................................... 125

References............................................................................................................. 139
List of Figures

Figure 1: The microRNA biogenesis pathway ......................................... 6
Figure 2: Predicted structures of primary and SL1-spliced let-7 transcripts 8
Figure 3: Post-transcriptional control of microRNA biogenesis .......... 14
Figure 4: Analysis of lin-4 transcription ........................................... 62
Figure 5: Evidence that lin-4 biogenesis is regulated post-transcriptionally 64
Figure 6: A candidate-based RNAi screen for factors involved in lin-4 biogenesis ................................................................. 65
Figure 7: Characterization of the role of R05H10.2 in microRNA biogenesis ................................................................. 66
Figure 8: Interaction of R05H10.2 with the heterochronic pathway ...... 67
Figure 9: A model for the role of lin-14 and lin-28 in modulating germline and somatic growth ................................................ 68
Figure 10: Developmentally staged RNA samples analyzed for lin-4, lin-14, and lin-28 RNA expression........................................ 87
Figure 11: lin-4 miRNA developmentally regulates target mRNA expression ................................................................. 88
Figure 12: Quantitation of the bursting phenotype observed in alg-1(ap45) mutant animals after 2-generation RNAi against vector or dcr-1.. 114
Figure 13: Genetic mapping of ap45 mutant (originally called dre-2)...... 115
Figure 14: Altered DNA and protein sequence in ap45 allele and verification of the snp by genotyping PCR ............................ 115
Figure 15: The ap45 amino acid change (A756V) occurs in a highly conserved residue in the PIWI domain of ALG-1...................... 116
Figure 16: Molecular analysis of the ap45 allele ................................ 117
Figure 17: Attempts to rescue alg-1(ap45) with a wild-type alg-1 rescue construct and an alg-1(ap45) rescue construct ................. 118
Figure 18: Duplication and deletion in the \textit{lin-4(e912)} mutant
List of Tables

Table 1A: Vulva is absent or defective in adult N2 on R05H10.2 (RNAi)... 69

Table 1B: R05H10.2 (RNAi) induces alae phenotypes in N2 adults........ 69

Table 2: Suppression of the precocious alae of *lin-14* and *lin-28* mutants by R05H10.2 (RNAi). .......................................................... 70
It is probably not possible to adequately acknowledge everyone who contributes toward a Ph.D journey, but I would like to give thanks to my lovely wife, Kate Smoot Bracht, who has endured my many hours spent in the lab and writing as I've finished this degree. I also want to acknowledge my parents, Roger and Patty Bracht, my sister Laura Bracht, and my brother Stephen Bracht as well as his wife, Ginny Bracht. I also want to acknowledge the support of my father-in-law, Jon Smoot, and his wife, Mary Hobbie. All have been very encouraging and supportive along the way.

Friends who have made a difference along the way include Kyle Beardsley, David Minh, Miguel and Jillian LaSerna, Jeff Wong, Laurie Rice, Magda Chechlacz, Peter Gent, Kristin Irwin, Cynthia Yagi, Chris and Jen Eugley, Jean and Peter Shen, Valerie and Andrew Thatcher, Rob Gaebler, and Jaeson Ma.

I would like to acknowledge Professor Amy Pasquinelli for her guidance and insights along the way. She truly lives her passion for science and conveys that excitement in every meeting with her that I have had. I also want to acknowledge many interesting and helpful discussions with many present and past members of the Pasquinelli lab including Shuan Hunter, Katlin Massirer, Zoya Kai, Katya Melnik-Martinez, Priscilla Van Wynsberghe, Dimitri
Rotation students I have had the honor of mentoring include Christie Chang, Nicole Orazio, Siler Panowski, and Tiffany Dunbar.

Chapter 1, in full, has been submitted for publication of the material as it may appear in Development, 2009, Bracht, John and Pasquinelli, Amy E. The dissertation author was the primary investigator and author of the paper.

Chapter 2 describes work performed by the dissertation author which was published in Cell, 2005 Aug 26;122(4):553-63., with contributions from Shveta Bagga, Shaun Hunter, Katlin Massirer, Janette Holtz, Rachel Eachus, and Amy E. Pasquinelli. The dissertation author was second author on the paper. Only those figures for which the dissertation author was primary investigator are reprinted in chapter 2, and the accompanying text has been rewritten and updated by the dissertation author, who is the primary author of the text presented in the chapter.
Vita

EDUCATION

2002 – 2009
University of California, San Diego
Ph.D in Biology
Research: studying the biogenesis of micro-RNAs let-7 and lin-4 and the mechanism of their target gene regulation in C. elegans.

1998 – 2001
New Mexico Tech
Bachelor of Science in Biology
GPA 3.96 (highest honors)

1997 – 1998
University of New Mexico, Los Alamos
GPA 4.0

RESEARCH EXPERIENCE

2002 – 2009
University of California, San Diego
Graduate Student
- Research on microRNA biogenesis and function in C. elegans

2001 – 2002
Baylor University
Research Assistant
- Performed intensive literature research on evolution, complex systems, and self-organization
- Wrote an analysis paper based on this work that was published in Complexity
- Co-wrote a genetic algorithm to study the evolution of biological adaptations (www.iscid.org/mesa)

1996 – 2000
Los Alamos National Laboratory
Undergraduate Research Assistant, High-school Co-op student
- Worked in the biophysics group (P21), which utilizes ultra-sensitive magnetic field sensors (known as SQUIDs) to study magnetic fields emitted by the brain and heart
- Co-designed and built several different test coils which were used to calibrate the SQUID devices
- Wrote a graphical user interface for control of SQUIDs which was widely utilized by other researchers in the laboratory
- Assisted in writing MatLAB software for analyzing and displaying data obtained from SQUIDs
- Functioned for a summer as the sole AutoCAD designer and machinist liaison
TEACHING EXPERIENCE
2003-2005 Teaching Assistant, UCSD. Classes: Biochemical Techniques, Genetics, and Structural Biochemistry.

AWARDS AND SCHOLARSHIPS
2003-2006 NIH Cellular and Molecular Genetics (CMG) Training Grant
2002-2003 Graduate Assistance in Areas of National Need (GAANN) Fellowship
1998-2001 New Mexico Tech Silver Scholar

PUBLICATIONS


PRESENTATIONS

Platform – West Coast Worm Meeting, Santa Barbara CA.

Platform – International Worm Meeting, Los Angeles, CA

Poster – International RNA Meeting, Madison, WI

Poster– International RNA Meeting, Banff, Canada
MicroRNAs are a newly-discovered class of short regulatory RNAs that function as sequence-specific guides for the down-regulation of target gene expression. *lin-4* was the first microRNA to be discovered, and it has two known targets: *lin-14* and *lin-28*, which are regulated by partial base-pairing of the microRNA to target regions in their 3' untranslated regions (3'UTRs).

Chapter 1 of this thesis is devoted to analyzing the regulation of *lin-4* biogenesis. MicroRNAs are initially transcribed as long primary transcripts, which are processed by RNaseIII enzymes Drosha and Dicer to liberate the
mature microRNA. The expression of mature lin-4 is developmentally controlled such that worms do not express detectable levels of the microRNA, until ~12 hours of development on food. However, RT-PCR techniques revealed that transcription of lin-4 primary transcripts occurs during the early period when mature microRNA is undetectable, suggesting the existence of a post-transcriptional regulatory mechanism preventing premature processing of these transcripts. A GFP-sensor-based RNAi screen for factors involved in this processing block identified a conserved RRM-domain protein, R05H10.2. Inactivation of R05H10.2 by RNAi causes a post-transcriptional defect in the accumulation of mature lin-4 and induces a dramatic growth defect. However, this growth defect is substantially suppressed in lin-14 and lin-28 mutant worms, revealing a previously undescribed linkage between the heterochronic pathway and organismal growth.

Chapter 2 of this dissertation is an analysis of the mechanism of target regulation for lin-4. When my dissertation work commenced, the prevailing model for microRNA function held that target regulation occurred by translational control which left target mRNA levels unaffected. The work detailed in chapter 2 is an analysis of the levels of lin-14 and lin-28 mRNAs undergoing lin-4 mediated regulation in wild-type and lin-4 mutant worms, showing that the target mRNAs are degraded as a result of their interaction with the microRNA. The Pasquinelli lab was among the first groups to report degradation of target mRNAs by microRNA regulation of endogenous targets,
and my work with *lin-4* was combined with analysis of *let-7* target regulation performed by other workers in the Pasquinelli lab, and published in *Cell* in 2005.
Introduction
The discovery of microRNAs and their functions

The existence of microRNAs was heralded in 1993 by a paper published by Victor Ambros and coworkers describing the mapping and characterization of the developmental lineage mutant \textit{lin-4}, which was found to have a defect in a small RNA rather than a protein product. (Lee 1993). In the same issue, Gary Ruvkun and coworkers published an analysis of a \textit{lin-4} target gene, \textit{lin-14}, which is post-transcriptionally regulated by base-pairing interactions between \textit{lin-4} and target sites in the \textit{lin-14} 3'UTR (Wightman 1993). Discovery of the second, highly conserved microRNA \textit{let-7} in \textit{C. elegans} in 2000 (Reinhart 2000, Pasquinelli 2001, Pasquinelli 2003) sparked a frenzy of microRNA identification primarily through cloning and sequencing size-selected small RNAs (Lee 2001)(Lau 2001)(Lagos-Quintana 2001), which uncovered hundreds of examples of similar short regulatory RNAs, some of which are conserved to insects and humans.

\textit{lin-4} and the heterochronic pathway

\textit{lin-4} and its targets are embedded in a pathway that controls developmental timing in \textit{C. elegans}, known as the heterochronic pathway (Slack 1997). The defining feature of members of this class of genes is that when mutated, they cause certain tissues to exhibit precocious expression of later cell fates at earlier stages, or reiteration of earlier cell fates at later
stages. Indeed, the \textit{lin-4} target gene \textit{lin-14} was initially characterized as a gene which, when mutated to loss-of-function, caused the hypodermis and seam cell lineages to skip the first larval stage (L1) cell fate, instead manifesting second-stage (L2) division patterns (Ambros 1984). Interestingly, 3'UTR mutations in \textit{lin-14} were found to be gain-of-function alleles, and they had the opposite effect: reiteration of the L1-stage division patterns indefinitely (Ambros 1984, Wightman 1991, Ruvkun 1991). \textit{lin-4} mutants display a similar phenotype to \textit{lin-14} gain-of-function mutants in which L1-stage divisions are reiterated, which is coherent with the discovery that the \textit{lin-4} microRNA is a developmental timer that switches \textit{lin-14} off, utilizing elements contained its 3'UTR (Wightman 1991, Wightman 1993). Since that initial discovery, the cold-shock-domain protein \textit{lin-28} has been characterized as another target of \textit{lin-4} regulation, and \textit{lin-28} also controls stage-specific developmental patterns, in this case of the second larval stage (L2) (Moss 1997). \textit{lin-28} loss-of-function mutants go directly from L1-stage to L3-stage developmental patterns, skipping L2 fates; gain-of-function alleles have not yet been identified but a rescuing \textit{lin-28} transgene with \textit{lin-4} binding sites deleted from its 3'UTR displayed reiterations of L2 fates (Moss 1997).

\textit{lin-28} as a marker of stemness and undifferentiated tissues

The \textit{lin-4} target gene \textit{lin-28} is highly conserved across bilaterian phylogeny (Moss 2003). Furthermore, its down-regulation during development
and regulation by microRNAs is also conserved (Moss 2003, Wu 2005).
Interestingly, neuronal differentiation was found in Wu and Belasco 2005 to
require the down-regulation of lin-28 by the human lin-4 microRNA homologs,
mir-125a and miR-125b. Consistent with this data, C. elegans lin-4 expression
has been reported in neurons based on promoter-GFP constructs (Esquela-
Kerscher 2005, Ow 2008, Martinez 2008, and my work reported in chapter 1)
which suggests that the tissues in which lin-28 is downregulated by lin-4 (or
lin-4 homologs) are also conserved across phylogeny, and that lin-28 is likely
to play a role in neuronal differentiation in C. elegans also. In addition to the
conservation of the lin-4-lin-28 linkage and role in neuronal development,
several recent publications have identified a role for lin-28 mediating post-
transcriptional processing of let-7 microRNA primary transcripts and/or
Piskounova 2008, Viswanathan 2008); ongoing work from the Pasquinelli lab
has shown that lin-28 plays a similar role in C. elegans. For a more detailed
discussion of the literature reporting the role of lin-28 in mediating let-7
biogenesis, see below, “Post-Transcriptional Regulation of MicroRNA
Processing.” Perhaps most intriguingly, lin-28 was recently characterized as
one of four genes that when overexpressed, can reprogram human somatic
cells into stem cells (Yu 2007). Furthermore, the lin-28-regulated microRNA
let-7 plays important conserved roles in driving cell differentiation, regulates
oncogenes RAS and HMGA2 and is often reduced in cancers of the lung and
breast (for review, see Nimmo 2009). These findings highlight the importance and conservation of the lin-4 genetic network for directing cell differentiation, with roles in stem cells and cancer. My studies on lin-4 in C. elegans have implications for all these aspects of the lin-4 network and may contribute ultimately toward our understanding of human diseases.

**Chapter 1: mechanisms of lin-4 biogenesis**

The first chapter of this thesis is devoted to an analysis of the biogenesis of the lin-4 microRNA. In general, microRNAs begin life as pol-II or pol-III-transcribed primary transcripts which can be multiple kb long (Bracht 2004, Corcoran 2009) (Fig. 1). Endonucleolytic cleavage, performed by the microprocessor complex (Denli 2004, Gregory 2004) comprised of Drosha (an RNAselIII nuclease) and Pasha/DGCR8 (an RNA-binding protein and guide) (Han 2004, Gregory 2004, Landthaler 2004, Han 2006), releases a hairpin structure containing the microRNA on one side and the mir-star (or mir*) on the other (Fig. 1). This microRNA-containing hairpin structure, known as a precursor, is generally thought to be exported from the nucleus to the cytoplasm, where it encounters the RNAselIII nuclease Dicer which cleaves the hairpin loop to liberate the microRNA-mir* duplex (Fig. 1). An unknown activity unwinds this duplex and selects the microRNA for loading into a protein complex known as the miRISC complex (microRNA Induced Silencing Complex), whose identity and role is defined by the presence of Argonaute
protein family members. The miRISC complex utilizes the microRNA as a sequence-specific guide to mediate degradation or translational control of target mRNAs (Bartel 2004, Chekulaeva 2009b) (Fig. 1).

**Figure 1.** The microRNA biogenesis pathway. The microRNA is transcribed as a Pol-II or Pol-III primary transcript (shown here is Pol-II) from DNA in upper left. The primary transcript is processed by the Drosha nuclease and its RNA binding partner Pasha, also known as the microprocessor complex, to liberate the precursor hairpin species. Dicer processing produces a microRNA duplex which is unwound and only the microRNA is retained to be incorporated into Argonaute containing complexes to serve as sequence-specific guides mediating target regulation via translational repression and/or mRNA degradation.

**Post-Transcriptional Regulation of MicroRNA Processing**

**Trans-splicing of let-7 primary transcripts**

In the first publication from the Pasquinelli lab, we examined the primary transcripts arising from the let-7 microRNA locus (Bracht 2004). In that work we reported a novel trans-splicing event which appears to be required for
Drosha-mediated processing of the *let-7* primary transcript. In nematodes, ~70% of mRNAs are appended at the 5’ end with a short (22nt) leader sequence which arises from a separate noncoding RNA (termed SL1 or SL2) (Blumenthal 1995). *C. elegans* genes are often arranged into operons in which one promoter drives expression of multiple coding regions. The SL1 splice leader is nearly always appended to the 5’ end of the first coding region, while SL2 is appended to downstream genes within operons, separating the coding sequences into individual mRNAs (Blumenthal 1995). Our work in 2004 identified the first (and to date, only) example of trans-splicing of microRNA primary transcripts and demonstrated that this splicing step is essential for Drosha processing of primary *let-7* transcripts.

It is unclear whether the trans-splicing event recruits proteins that promote Drosha processing, or whether the spliced transcript takes on a structure that is amenable to Drosha processing. However, we did observe that the splice-leader sequence added by the trans-splicing reaction is capable of base-pairing to a downstream region in the primary transcript, inducing a conformational change, which may be important to facilitate Drosha processing (Fig. 2).

Bracht 2004 laid the groundwork for the analysis presented in Chapter 1 of this thesis, in which a similar analysis of endogenous microRNA primary transcripts was undertaken for the developmental regulator *lin-4*.
Figure 2. Predicted structures of primary and SL1-spliced let-7 transcripts. (A) Potential secondary structures of the 890-nt primary let-7 transcript and (B) the SL1-spliced RNA were generated using the mfold program (Zuker 2003). The structures shown are the lowest free energy predictions. The SL1 splice site, SL1 donor sequence, and mature let-7 RNA are outlined and labeled. The boxes indicate predicted structural rearrangements surrounding the let-7 stem loop region resulting from SL1 splicing.

p68, p72, and Drosha processing of microRNAs

Recent work has uncovered several instances of protein cofactors that regulate microRNA processing. In the initial report describing the microprocessor complex, which contains Drosha and Pasha and mediates cropping of primary transcripts into precursors, several other associated proteins were also described as interacting with Drosha in human cells (Gregory 2004). These Drosha interacting proteins include two DEAD-box helicases, DDX17 and DDX5, also known as p72 and p68. These helicases were disrupted by homologous knockout in mice in Fukuda 2007, and found to
be essential for embryonic development and early postnatal survival. Analysis of global microRNA levels revealed that these helicases are important for efficient expression of a subset of microRNAs. p72 deficiency affected 94 out of 266 microRNAs tested; most of these, with one exception, were dependent on p68 as well, and were reduced to approximately half their normal levels when either helicase was mutated or depleted (Fukuda 2007). The regulated microRNAs showed no enrichment for motifs that might mediate helicase interaction, suggesting that other cues such as tertiary structure of the primary transcripts are specifically recognized by p68 and p72, and that these helicases may play a role in presenting them to the Drosha nuclease (Fukuda 2007) (Fig. 3).

**SMAD proteins and microRNA biogenesis**

Human smooth muscle differentiation is at least partly regulated by miR-21 (Davis 2008), and processing of miR-21 primary transcripts to precursor forms was found to be controlled by TGF-β and BMP signaling pathways. Both these pathways use SMAD proteins as signal transducers, and Davis and coworkers demonstrated that SMAD proteins can directly interact with the p68-Drosha complex as well as microRNA primary transcripts. They propose that TGF-β and BMP signaling pathways cause nuclear localization of SMAD proteins, which interact directly with specific microRNA primary transcripts (at least for miR-21 and miR-199a), stabilizing their association with the p68-Drosha complex and facilitating their processing into
precursor forms (Fig. 3). In turn, these microRNAs drive the differentiation of vascular smooth muscle cells by down-regulation of target genes such as $PCD4$, a tumor suppressor and known miR-21 target (Frankel 2008)

**lin-28 and let-7**

In 2008 a flurry of papers announced an unprecedented discovery: *lin-28* is a direct regulator of *let-7* biogenesis, although the mechanisms reported were varied and are still under investigation. Viswanathan et al. in February 2008 reported that *lin-28* prevents microprocessor-mediated cropping of *let-7* family primary transcripts in human embryonic stem cells, and a later report from the same lab (Piskounova 2008) defined the loop region as critical for direct *lin-28* binding, and identified specific residues of both the *let-7* loop and *lin-28* cold-shock domain as important for the interaction. In August of 2008 two more papers were published on *lin-28*. A report from the Hammond lab (Newman 2008) reported that *lin-28* regulates primary *let-7* processing in undifferentiated embryonic P19 cells in mouse. The block in *let-7* primary transcript processing was found to be due to interference with Drosha function, and specific conserved residues in the hairpin loop were found to mediate direct binding to *lin-28* and consequent inhibition of Drosha processing.

A report from the Wulczyn lab (Rybak 2008) examined the role of *lin-28* in mouse neural stem cells and presented a different model than that proposed by the Hammond and Gregory labs. In their report, Rybak et al.
show that *lin-28* regulates the processing of *let-7* precursors by Dicer and that down-regulation of *lin-28* in neuronal stem cells is mediated by *miR-125* (the mammalian homolog of *lin-4* from *C. elegans*) and *let-7*, implicating an autofeedback loop in differentiation of neuronal stem-cell lineages. Both *miR-125* and *let-7* specifically interact with the *lin-28* 3’UTR through partially complementary sequences, and down-regulate protein expression. Furthermore, the direct interaction between the *let-7* precursor and *lin-28* was shown to inhibit Dicer processing *in vitro* (Rybak 2008).

Finally, in October of 2008 a report from the Kim lab (Heo 2008) suggested yet another model for *lin-28* activity. This report analyzed the role of *lin-28* in human cell lines, and observed a striking block of Drosha processing of *let-7* primary transcripts by *lin-28* protein; however there was also a strong destabilization of precursor isoforms accompanied by a shift in mobility of the targeted RNAs. Sequence analysis of these longer isoforms revealed that they were uridylated forms of the precursor, and Heo and coworkers propose a model in which *lin-28* antagonizes the stability of precursor *let-7* isoforms by recruiting a terminal uridylyl transferase (TUTase) which appends a series of uridines to the precursor, marking it for degradation by undefined nucleases. A follow-up study identified this uridylyl transferase as TUT4 (Heo 2009).

Intriguingly, a very recent paper defined a uridylyl transferase activity that acts on mature miR-26, which regulates immune signaling (Jones 2009), and found that mature miR-26 is not able to properly repress a reporter gene
when uridylated. These three papers (Heo 2008, Heo 2009, and Jones 2009) provide insight into the hitherto unknown pathways that mediate microRNA degradation and function.

Taken together, these papers point to the possibility that *lin-28* may be regulating a variety of steps in the microRNA biogenesis pathway in different tissues at different times, and suggests that there may not be one single mechanism by which *let-7* biogenesis is controlled by this RNA-binding protein (Fig. 3). Alternatively, multiple points of repression may all act in concert to contribute toward a failsafe post-transcriptional block on expression of *let-7* at early times when its expression would be harmful to the developing organism or tissue.

**hnRNP A1 and miR-18a**

A recent study showed that the RRM-domain protein *hnRNP A1* is important for the Drosha-mediated processing of *miR-18a*, but not for processing of other microRNA hairpins processed from the same primary transcript (Guil 2007). *hnRNP A1* has been well characterized as a mediator of alternative splicing and mRNA export (Piñol-Roma 1992, Mayeda 1992, Cáceres 1994, Dreyfuss 2002) so it was surprising to find that it also plays a very specific role in the biogenesis of one particular microRNA. Mechanistic analysis in Michlewski et al. 2008 revealed that *hnRNP A1* binding to *miR-18a* alters its structure such that it becomes accessible to the microprocessor (Fig.
The binding of *hnRNP A1* to *miR-18a* primary transcript occurs at the terminal loop, but structural changes propagate to the Drosha cleavage site near the base of the hairpin (Guil 2007, Michlewski 2008). Michlewski et al. further noted that evolutionary conservation of terminal loop sequences is a signal of regulation by RNA-binding proteins, and that ~14% of human microRNA precursors have highly conserved loop regions which may interact with regulatory proteins. They tested a subset of these microRNAs with *in vitro* processing assays, which demonstrated post-transcriptional regulation mediated by stem-loop sequences (Michlewski 2008). These data demonstrate that well-studied RNA-binding proteins may in fact have important roles in regulating the microRNA pathway, and the work presented in Chapter 1 of this thesis presents my contribution to understanding the post-transcriptional regulation of microRNA biogenesis.
Figure 3. Post-transcriptional control of microRNA biogenesis. p68 & p72, SMADs, and hnRNP A1 all function to promote Drosha processing of one or more microRNAs, while \textit{lin-28} serves as a repressor of \textit{let-7} biogenesis by a variety of mechanisms reported in the literature.

**Post-transcriptional regulation of \textit{lin-4} processing**

In addition to its status as the first microRNA to be discovered, the \textit{lin-4} microRNA is intriguing because its expression is developmentally controlled. Prior work identified regulatory regions upstream of the \textit{lin-4} microRNA capable of driving GFP expression that mirrors the expression of mature microRNA (Esquela-Kerscher 2005, Ow 2008, Martinez 2008). My contribution to this field identified more distal upstream regions capable of driving GFP expression ubiquitously during the postembryonic lifecycle, suggesting that some primary transcripts are post-transcriptionally regulated. Testing of endogenous primary transcripts confirmed the post-transcriptional regulation
of lin-4. I followed up on this observation by creating a GFP-sensor for lin-4 microRNA function, which was used in a candidate-based RNAi screen to search for post-transcriptional regulatory factors. The identification of R05H10.2 as a conserved RRM-domain protein which plays important roles in mediating post-transcriptional processing of lin-4 primary transcripts is presented in Chapter 1. This work also identifies a previously uncharacterized function of lin-14 in mediating germline development. This finding situates the heterochronic pathway within whole-organism growth networks, with R05H10.2 as a master regulator for both somatic and germline growth pathways.

Chapter 2: Mechanism of microRNA target regulation

microRNAs and the RNA interference (RNAi) pathway

When I began work on my thesis in 2004, the prevailing model for microRNA regulation of target mRNAs held that the fate of targeted mRNAs was determined strictly by the degree of base-pairing to the regulating microRNA. This model was based on findings from the RNA interference field that originated with studies in C. elegans, in which exogenously introduced double-stranded RNAs induced degradation of homologous targets (Fire 1998). The mechanism of RNA interference was found to involve cleavage of the double-stranded RNA into many short interfering RNAs (siRNAs) of approximately 22nt by a RNaseIII enzyme that was consequently dubbed
Dicer (Bernstein 2001). These siRNAs function as sequence-specific guides for endonucleolytic cleavage of homologous target RNAs through the activity of the RNA-Induced Silencing Complex (RISC). It is now known that the nuclease responsible for target cleavage is an Argonaute protein, a defining member of the RISC (Liu 2004).

The apparent similarity in size between siRNAs and microRNAs did not go unnoticed, and Mello and coworkers elucidated a remarkable degree of similarity in the biogenesis pathways for siRNAs and microRNAs (Grishok 2001). Due to their partial self-complementarity, microRNA hairpin precursor transcripts act somewhat like endogenous double-stranded RNAi “triggers” and get processed by Dicer into a 22nt mature form. The main difference between microRNAs and RNAi triggers is in their origin (microRNAs are endogenously produced while RNAi triggers are exogenously introduced) and their early biogenesis (Drosha is not utilized in the RNAi pathway but is essential for most microRNAs).

**Target regulation by perfect versus imperfect small RNA base-pairing**

Given the high degree of similarity between microRNA and RNAi pathways, it was most natural to assume that mechanisms of target regulation would essentially be mechanistically very similar as well. However, microRNA base-pairing to targets incorporates mismatches between the guide RNA and target mRNA around the site of endonucleolytic cleavage (which, with perfect
matches, occurs in the target RNA across from nucleotides 10 and 11 of the guide RNA) (Elbashier 2001). Therefore, it was proposed that targets with a high degree of complementarity with the guide RNA would undergo mRNA degradation through an RNAi-like mechanism of endonucleolytic cleavage, but targets with mismatched interactions with microRNAs would be subjected to translational repression, leaving target mRNA unaffected (Bartel 2004, Doench 2003, Zeng 2003, Hutvágner 2002). In these early papers, microRNAs were found to be able to guide cleavage of perfectly matched transgenic targets (Hutvágner 2002) and artificial siRNAs with bulged matches induced translational repression of reporter transgenes without degradation (Doench 2003). However, these studies were based on exogenously introduced regulatory RNAs and target constructs, not from endogenous microRNA-target interactions.

**Endogenous microRNAs and endogenous target regulation**

Evidence supporting this paradigm for endogenous microRNAs and targets came primarily from *lin-4-lin-14* and *lin-4-lin-28* interactions in *C. elegans* (Wightman 1993, Olson 1999, Seggerson 2002). The earliest reports of the mechanism by which microRNAs regulate target protein expression examined the *lin-4-lin-14* microRNA-target model and observed little or no decrease in target mRNA abundance when *lin-4* was present (Wightman 1993, Olsen 1999), suggesting that a translational control mechanism was
repressing *LIN-14* protein accumulation. However, my experiments with the *lin-4-lin-14* microRNA-target pair led to a different conclusion: that *lin-4* does indeed induce degradation of its targets, including *lin-14* and *lin-28*. In light of these findings, it is worth exploring how the translation-control-only paradigm of microRNA-target regulation was established.

Olsen et al. 1999 performed polysome analysis of *lin-14* mRNA in wild-type worms, comparing stages before and after *lin-4* expression, and reported no differences in polysome profiles between the stages. They concluded that *lin-4* regulates post-initiation steps of *lin-14* translation and postulated a proteolytic mechanism to prevent protein accumulation after translation (Olsen 1999). However, they do report a modest decrease in *lin-14* mRNA (from total RNA) of 2.3-fold (table 1, Olsen 1999) and in their conclusion they state that “… we found that the levels of *lin-14* mRNA in the L1 and L2 differ by a factor of less than 3 ….” Implicit in their argument here is that Wightman et al. 1993 reported that *LIN-14* protein decreases over 10-fold between early (L1) and later (L2-L3) stages, while mRNA levels do not change. Intriguingly, however, Wightman et al. 1993 also report that *LIN-14* protein is still downregulated in the *lin-4(e912)* null mutant by 2.3-fold. This implies that the *lin-4* microRNA is responsible for a 5-fold down-regulation in protein levels. This correlates well with my analysis published in Bagga 2005, which showed that *lin-14* mRNA decreases approximately 5.5-fold from L1 to L2 (with 1.5-fold standard deviation).
It is possible to partially reconcile our findings with the data published in the Olsen and Wightman papers. The 2-3 fold down-regulation of \textit{lin-14} mRNA in L2 versus L1 stage worms reported in Olsen 1999 is nearly within the range we reported. Examination of Fig. 4 from Olsen et al. also reveals that there appears to be distinctly less total \textit{lin-14} mRNA in the L2 stage polysome profiles relative to L1 stage profiles (approximately 2-fold). There is still a difference between Olsen et al. 1999 and my findings in Chapter 2, but this may be due in part to our RNA-detecting methodology. My study utilized Northern blots whereas Olsen et al. used RNAse protection assays, which do not detect the full length transcripts—they can only detect the total RNA covering the region defined by the probe-target overlap, and would therefore be expected to be less sensitive to variations in isoform-specific RNA levels. In Bagga et al. 2005 we analyzed \textit{lin-41}, a target of the microRNA \textit{let-7}, and some relatively stable truncated mRNA fragments were observed on Northern blots which could potentially contribute to total signal for RNAse protection assays (Bagga et al. 2005); however, we did not observe stable fragments for either \textit{lin-14} or \textit{lin-28} (data not shown). The basis for the different observations remains unclear, but another methodological difference in our study was that we compared equivalent time-points in wild-type and microRNA mutant animals when target regulation by the microRNA should be in effect. Olsen et al. 1999 compared target mRNA levels in stages before and after \textit{lin-4} accumulates; however, Wightman et al. 1993 did perform wild-type versus
mutant analysis and reported no differences in mRNA levels. Two potential problems should caution our interpretation of the Wightman 1993 experiments: 1. the staged worms were not synchronized populations but rather were size-selected and thus not homogenous populations; and any younger worms mixed into later-stage populations would cause the mRNA levels to be high, and 2. RNAse protection assays were utilized rather than Northern blots.

Whatever the reasons for the discrepancies between studies, degradation of target mRNAs by cognate microRNAs should not alter translation or ribosome behavior. The fact that Olsen et al. report no alterations in the structure of polysome profiles upon lin-4 microRNA expression is quite consistent, therefore, with the model that lin-4-mediated downregulation of LIN-14 protein levels is carried out primarily at the level of mRNA stability rather than translation or protein stability.

Soon after Olsen et al., Eric Moss and co-workers published a paper (Seggerson 2002) in which they analyzed lin-28 regulation by lin-4 in much the same way as Olsen et al. with lin-14. They also performed RNAse protection assays from L1 and L2 staged wild-type worms and observed no decrease in lin-28 mRNA abundance; they also reported that polysome profiles do not change between these stages. However, they performed no direct comparison of lin-28 mRNA levels in wild-type versus lin-4 mutant animals. Furthermore, their polysome profiles show an overall lower level of both lin-14 and lin-28 in L2 than in L1 stage animals. In spite of this hint of target mRNA degradation,
they also concluded that post-initiation translational steps were the most likely the target of *lin-4* mediated repression, and supported the proteolytic mechanism in their discussion of the data.

When the work described here commenced, we had adopted the prevailing model of translational repression after initiation of translation, and were not looking to challenge it. However when Dr. Pasquinelli performed preliminary analysis of the *let-7* microRNA target gene *lin-41* in wild-type and *let-7* mutant worms, she observed that mRNA levels were much higher in the *let-7* mutant than in wild-type animals. With this initial observation to spark our search, we began testing different microRNA target mRNAs in the corresponding mutant animals by Northern blot, RT-PCR, and qRT-PCR techniques. By this time I had been working with *lin-4* to analyze its biogenesis, so it made sense for me to analyze its regulation of *lin-14* and *lin-28* as well. Intriguingly, I found that both *lin-14* and *lin-28* mRNA levels were strongly reduced by *lin-4*-mediated repression and that this repression was both dependent on the microRNA and (in the case of *lin-14*) the presence of *lin-4* complementary elements in the 3'UTR of the target gene. RNA interference destabilizes target mRNAs via endonucleolytic cleavage of target mRNAs between nucleotides 10 and 11 of the siRNA (Elbashier 2001), and the cleavage event leaves a 5' phosphorylated product, which can be detected by ligation and cloning techniques (5' RACE). We attempted to perform this technique for *lin-4* and *let-7* targets in *C. elegans*, but did not observe the
products expected from RNAi-type endonucleolytic cleavage. Because of this, we concluded that general mRNA degradation pathways were responsible for microRNA-mediated target down-regulation, most likely through deadenylation and decapping followed by exonucleolytic shortening from the 5’ end (explaining the role of 5’-3’ exonucleases in target regulation that we observed in Bagga et al.), and/or the 3’ deadenylated end. Several months before our work was published a report was released in Nature demonstrating that transfection of an exogenous microRNA into tissue culture cells caused global down-regulation of mRNAs bearing partial complementarity to the transfected microRNA (Lim 2005), which was broadly consistent with our findings.

**microRNAs and deadenylation**

Further work has since confirmed that microRNA-based regulation can induce deadenylation of targets (Wu 2006, Wakiyama 2007, Eulalio 2009a), and one mechanism for deadenylation is the association of argonaute proteins with GW182 proteins (rich in glycine and tryptophan residues), (review: Eulalio 2009b) which localize to sites of mRNA decay called P-bodies (Parker 2007). Intriguingly, P-bodies contain deadenylases as well as decapping enzymes and RNA exonucleases involved in degradation. Deadenylation might be expected to both repress translation and also lead to mRNA destruction through decapping and exonucleolytic degradation. However, it has become clear that at least some target mRNAs are not degraded but are stable and
translationally repressed even though deadenylated (Wakiyama 2007, Eulalio 2008, Takimoto 2009, Beilharz 2009). Indeed, some targets appear to be mainly “translational repression” targets and some are “degradation targets” (Eulalio 2009a). In general, two broad classes of mechanisms of microRNA-target regulation have been proposed: GW182-mediated (hence, deadenylation-mediated) translation repression and/or decay, and GW182-independent translational regulatory mechanisms impinging upon post-initiation steps or protein stability (Olsen 1999, Seggerson 2002, Nottrott et al. 2006, Peterson 2006, Maroney 2006).

Another model, which fits into the category of GW182-independent mechanisms, posits that a novel cap-binding motif in Argonaute proteins competes for cap-binding complex eIF4E and thereby directly represses translation initiation (Kiriakidou 2007). Recent work has contested these findings, however, as Eulalio et al., 2008 demonstrate that the residues initially thought to mediate cap binding by argonaute proteins actually are necessary for interaction with GW182 proteins and microRNAs.

The apparent contradictory nature of the myriad models for microRNA-target regulation is an indication that perhaps the exact process by which microRNAs regulate their targets is actually variable based on organism, cell type, and microRNA-target interaction. Additionally, it is tempting to speculate that a given microRNA might regulate its targets via different mechanisms depending upon associated cofactors which might be regulated in temporal or
spatial ways to increase the options for developmental control available to an evolving organism.

**GW182 proteins in *C. elegans***

The worm model system has proven useful in uncovering new phenomena with broad-ranging applicability (eg, microRNAs). However, the *C. elegans* GW182 proteins appear to be structurally divergent from those in Humans and Drosophila, which have an ubiquitin-associating-like (UBA) domain, followed by a GW-repeat domain, followed by an RRM motif. Several studies implicate the C-terminal end of GW182 proteins, including the RRM motif, in target gene regulation (Eulalio 2009c, Lazzaretti 2009, Zipprich 2009); however, the *C. elegans* GW182 proteins, called *Agl-1 Interacting Protein-1* and 2 (*ain-1* and *ain-2*) contain only the GW182 repeat domain, lacking the potentially important RRM domain (Eulalio 2009b). Reassuringly, recent work has demonstrated striking similarities in functionality between the *C. elegans* protein and the more complex human and Drosophila proteins. Ding et al. 2009 analyzed microRNA target regulation in wild-type versus *ain-1;ain-2* double mutant worms and found that polysome profiles were slightly shifted toward heavier fractions (ie, polysomes containing greater numbers of ribosomes) in the double mutant, consistent with a role of these proteins in repressing translation. However, Ding et al. 2009 also showed that target degradation was substantially impaired in the *ain-1;ain-2* double mutant,
implicating *C. elegans* GW182 proteins in both translation repression and target degradation pathways. The same study also demonstrated that the *lin-4* microRNA mediates both degradation and translational inhibition of *lin-14* and *lin-28* mRNAs. Since mutant polysome profiles shift toward heavier fractions relative to wild-type profiles, and because translation initiation machinery genetically interacts with *let-7* (Ding 2008), the authors conclude that microRNAs repress translation initiation in addition to mediating target degradation. This finding contrasts with early studies on *lin-4*, *lin-14*, and *lin-28* (Wightman 1993, Olsen 1999, Seggerson 2002), which concluded that steps after translation initiation were affected, but it does nicely harmonize the worm GW182 field with findings in Drosophila and human systems, where translational repression and degradation are common features of GW182 and microRNA function (Eualio 2009c, Eualio 2008, Takimoto 2008, Lazzaretti 2009, Lian 2009, Zipprich 2009, Chekulaeva 2009). It is tempting to speculate that both translational repression and target degradation are initiated through deadenylation.

**Evidence to support post-initiation translational control**

While the majority of the microRNA field has begun to coalesce around mechanisms of target regulation different from the initial reports of translation repression after initiation, several papers have been published which present data supporting the original model. In Nottrott et al. 2006, the *C. elegans lin-41*
3'UTR, targeted by \textit{let-7}, was attached to a luciferase reporter gene which was tested for repression in HeLa cells. Intriguingly, the authors report that repressed reporter mRNA was associated with actively translating polyribosomes, which displayed sensitivity toward puromycin (a terminator of peptide elongation). Furthermore, nascent peptide epitopes can be immunoprecipitated from control mRNAs but not from microRNA-repressed target mRNAs, leading the authors to conclude that nascent peptides are destroyed. Therefore, the microRNA field is circling back to some of the original ideas presented in the Olsen 1999 and Seggerson 2002 papers.

Another paper from the Sharp lab (Peterson 2006) supports the idea that microRNAs function at post-initiation steps of translation, but suggests an alternative model. This study utilized an artificial siRNA with bulged target sites in the 3'UTR of a luciferase reporter. When regulatory RNA and target are coexpressed in human 293T cells, translational repression occurs and this repression does not affect polysomes, which were also found to be sensitive to puromycin. Repression of target gene expression in the presence of the short regulatory RNAs was still evident when Internal Ribosome Entry Sites (IRES) were utilized in the target mRNA, supporting the idea that post-initiation steps are being regulated in this case. Furthermore, nascent peptides were not observed from targeted mRNAs under repressed conditions, leading the authors to suggest a “ribosome drop-off” model in which ribosomes fail to elongate the length of the coding sequence of the mRNA. While these data do
support post-initiation mechanisms of translation repression, the artificiality of
the system raises questions regarding the applicability of these results to
endogenous microRNAs and their targets. An alternative approach was
adopted by the Nilsen lab (Maroney 2006), who evaluated the polysome
association of several endogenous microRNAs (including let-7) and their
target mRNAs in HeLa cells and found that both tend to associate with
polysomes, and in-vitro ribosome run-off experiments demonstrated that
translation termination and elongation appear normal. These data also support
a post-initiation mechanism of target regulation, and demonstrate that
microRNA-mediated repression of target genes occurs by a multifaceted
mechanism which may be different for different organisms, tissues, and even
target mRNAs.

Does mechanism matter?

Given that either translational repression or mRNA degradation
eliminates protein expression, one might wonder whether this issue is worth
the research money being devoted to it. However, the mechanism of
microRNA-target regulation has important implications for experimental
determination of direct microRNA targets. Since animal microRNAs only
imperfectly basepair with their targets, the potential interactome is very large—
and computationally as much as 1/3 to 2/3 of a given organism’s genome has
been predicted to be targeted by microRNAs (Friedman 2009, Lewis 2005).
However, only a subset of those interactions actually occur under physiological conditions in real organisms, and experimental validation of the interactions is critical. If microRNAs mediate degradation of their targets, then validation of microRNA-target interactions are straightforward—microarray or deep sequencing analysis of wild-type versus microRNA mutant organisms should highlight mRNAs with different expression levels. Specifically, genuine targets should be present at higher levels in the mutant organism than in wild-type. While the differential regulation of targets in the microRNA mutant organism is a first step, there is still potential for differences in gene expression to be due to downstream effects and thus, indirect targets. Directness can be secondarily verified by reporter assays with transgenes containing the putative microRNA-targeted 3'UTRs, both with and without presumptive target sites.

However, if targets are repressed at the level of translation, then more labor-intensive and technically challenging protein expression analysis may be required for large-scale target gene discovery. Interestingly, two global analyses of the effect of microRNAs on both protein and mRNA levels in human cells concluded that target degradation is a significant component of microRNA-target interactions (Baek 2008, Selbach 2008).

While ongoing research is focused on determining how microRNAs regulate their targets, my work with *lin-4, lin-14*, and *lin-28* challenged the original paradigm in the field by analyzing endogenous microRNAs and their
targets. Our paper ultimately proved catalytic in prompting work into the mechanistic basis of microRNA target regulation that is ongoing in multiple organisms.
Chapter 1: Regulation of *lin-4* miRNA processing, organismal growth and development by a conserved RNA binding protein gene in *C. elegans*
Introduction

MicroRNAs are a large class of short (~22nt) regulatory RNAs that typically direct down-regulation of target genes. The pioneering work of Victor Ambros, Gary Ruvkun, and co-workers described the first microRNA gene, *lin-4* and its target, *lin-14* in *C. elegans* in 1993 (Lee 1993, Wightman 1993), but since then hundreds of examples have been uncovered and they appear to be present in all multicellular organisms, often with high levels of conservation across phylogeny (Pasquinelli 2000, Pasquinelli 2003, Ambros 2003, Bartel 2004, Jones-Rhoades 2006). The biological functions carried out by these small regulators are diverse, but the majority of work to date has characterized their roles in development, including developmental timing (Ambros 1984, Feinbaum 1999, Moss 1997, Wightman 1993, Reinhardt 2000), defining left/right asymmetry in neurons (Johnston 2003, Chang 2004), differentiation of specific tissues including neurons (Wu 2005, Wulczyn 2007) and immune cells (Chen 2004, Rodriguez 2007 Thai 2007), control of hox-gene expression (Stark 2008, Woltering 2008, Yekta 2004), defining leaf polarity in plants (Kidner 2004, Garcia 2006, Mallory 2008), and fat storage and apoptosis in flies (Brennecke 2003, Xu 2003). Recent work has implicated many microRNAs in cancer, either as tumor suppressors or as oncogenes (for review see Visone 2009), roles which are coherent with their ability to drive differentiation of tissues.
The biogenesis of microRNAs is generally understood but specific cases often reveal complexities, especially since various RNA processing events such as splicing can interact with microRNA processing and afford opportunities for regulated expression (Bracht 2004, Kim 2007, Kim 2009, Kataoga 2009). The biogenesis of a 22nt microRNA begins with transcription (by pol-II or pol-III) of a primary transcript, which can be multiple kilobases long (Bracht 2004, Corcoran 2009). Endonucleolytic cleavage, performed by the microprocessor complex (Denli 2004, Gregory 2004) comprised of Drosha (an RNAseIII nuclease)(Lee 2003) and Pasha (an RNA-binding protein and guide) (Han 2004, Gregory 2004, Landthaler 2004, Han 2006), releases a hairpin structure containing the microRNA. The microRNA-containing hairpin structure, known as a precursor, is generally thought to be exported from the nucleus to the cytoplasm where it encounters the RNAseIII nuclease Dicer, which in turn excises the mature microRNA and facilitates its loading into a protein complex (miRISC), whose identity and role is defined by the presence of Argonaute protein family members. The miRISC complex utilizes the microRNA as a sequence-specific guide to mediate degradation (Chekulaeva 2009b) or translational control of target mRNAs (Bartel 2004).

As the defining member of the microRNA family of small RNAs, the biological function of lin-4 has been extensively characterized. It was one of several genes which initially defined the heterochronic class (Ambros 1984, Ambros 1989): mutations in these genes affect developmental timing of
certain tissues relative to germline development, which is largely unaffected. In particular, *lin-4* loss-of-function mutations caused reiteration of earlier cell fates at later times, a condition known as retarded development, which is particularly noticeable in the hypodermis and vulval lineages of *C. elegans*. In contrast, the two most widely studied targets of *lin-4*, *lin-14* and *lin-28*, cause the opposite effect when inactivated: some early stages are skipped altogether, in essence shifting the entire developmental cascade forward, which causes later developmental events to occur earlier than normal—a condition known as precocious development (Ambros 1984, Ambros 1989).

In this work we describe a new member of the heterochronic pathway that functions as a positive regulator of *lin-4* biogenesis and, as might be expected, causes classically retarded developmental phenotypes when inactivated. A recent paper from the Ambros lab described novel flywch-domain transcriptional repressors of *lin-4* in embryos, but surprisingly failed to observe any heterochronic phenotypes when these genes were inactivated (Ow 2008). The conclusion of that work was that an undefined mechanism was responsible for post-embryonic regulation of *lin-4* biogenesis. In this work we identify a post-transcriptional *lin-4* regulatory factor, the triple-RRM motif protein R05H10.2, which promotes the accumulation of mature *lin-4* in late L1 stage animals. Post-transcriptional regulation of *lin-4* confers robustness to the timing of expression of this microRNA, which is necessary because *lin-4* primary transcripts accumulate in early L1 (this study).
The work presented here was initiated as a study of *lin-4* primary transcripts arising endogenously. *lin-4* is interesting because the mature microRNA only accumulates after a defined temporal delay (11-12h at 20ºC) post-onset of feeding and reproductive growth (Feinbaum 1999, Chendrimanda 2007, Holtz 2009, this work). The mechanism by which *lin-4* is activated after the requisite developmental delay is largely unknown, although several previous studies have identified enhancer elements upstream of the microRNA that drive green fluorescent protein (GFP) at times consistent with accumulation of the mature RNA, suggesting a model in which the mature microRNA is activated at the transcriptional level (Esquela-Kerscher 2005, Ow 2008, Martinez 2008).

In our analysis multiple lines of evidence show that *lin-4* primary transcripts are produced at time-points prior to expression of the mature microRNA. In this work we show that the previously reported transcriptional up-regulation occurs in the context of ubiquitously detectable primary transcripts, and is overlaid by a mechanism of post-transcriptional control at the level of Drosha/Pasha processing. This mechanism prevents premature production of the microRNA from early-accumulating primary transcripts. Analysis of this mechanism identified the RRM-domain protein R05H10.2 as a regulator of *lin-4* biogenesis, and also revealed a previously unknown function for the heterochronic gene *lin-14* in controlling growth of the entire organism, including the germline.
Results

Characterization of endogenous lin-4 primary transcripts

MicroRNAs that are embedded within introns of protein-coding genes are often co-expressed with a common transcript giving rise to both gene products (Baskerville 2005; Kim 2007; Saini 2007). Although lin-4 is encoded in the sense direction within the 9th intron of host gene F59G1.4 (Fig. 4A), we suspected that transcription of the miRNA is regulated independently of its host gene because a relatively short (~700nt), entirely intronic DNA fragment is sufficient to rescue the lin-4(e912) null mutation (Lee 1993). To identify and characterize endogenous lin-4 containing primary transcripts, we performed 5' and 3' RACE (rapid amplification of cDNA ends). The 3' RACE results showed that the lin-4 sequence (in intron 9) was contiguous with exon 10 and that downstream splicing events matched those observed for F59G1.4, extending to the cleavage/polyadenylation site predicted for the F59G1.4 host gene (Fig. 4A). Notably, most of the lin-4 downstream sequence was not included in the original lin-4 rescue construct, indicating that these sequences and the endogenous polyadenylation site are not essential for lin-4 expression from multicopy transgenes (Lee 1993).

Using 5' RACE we identified multiple, developmentally regulated transcription start sites within intron 9 of the F59G1.4 host gene. In all, 9
putative transcription start sites were cloned (named A-I in Fig. 4A). As predicted, several start sites mapped within the *lin*-4 rescue construct but three independent sites reside upstream of the minimal rescue fragment (Fig. 4A, sites G, H, and I). In order to test whether sequences associated with the start sites are sufficient to drive transgene expression, we created promoter-GFP constructs that encompass sites C-F (p-1b) or C-I (p-1a) (Fig. 4A). Consistent with a previous report (Esquela-Kerscher 2005), we found that constructs which contained only the rescue-fragment promoter (p-1b) expressed GFP in pharynx, neurons, and seam cells coincident with when mature *lin*-4 is detected at the end of the first larval stage (L1) (Fig. 4B, C). Unexpectedly, the longer promoter-GFP construct (p-1a), which consists of p-1b (the rescue fragment promoter) extended upstream by 600nt, drove GFP expression in all stages of development (Fig. 4B, C). Furthermore, expression was maintained in L1 hatchlings deprived of food (Fig. 4C). Three start sites identified by 5’ RACE (G, H, I) are included in p-1a but excluded from p-1b and may be responsible for the early, food independent expression pattern. This complex transcriptional patterning suggests that regulation of *lin*-4 may include both transcriptional and post-transcriptional elements, and highlights the importance of studying miRNA biogenesis in the endogenous context.

**Post-transcriptional regulation of *lin*-4 miRNA biogenesis**
To test whether endogenous lin-4 primary transcripts are detectable at early timepoints, we established sensitive and reproducible semi-quantitative RT-PCR assays. We performed random-primed reverse transcription of RNA samples from staged wild-type populations at 4 (early L1) and 16 (early L2) hours of development. Oligos amplifying lin-4 primary transcripts flank the lin-4 hairpin and were designed to amplify all primary transcript isoforms; control oligos amplifying only host-gene intronic regions 5’ to all mapped lin-4 transcription start sites (F59G1.4 intron 9) were also utilized. In agreement with the 5’ RACE and p-1a promoter-GFP reporter results, the lin-4 primary transcript was readily detectable at both 4 and 16 hours of development, and in agreement with previously characterized transcriptional control (Esquela-Kerscher 2005, Baugh 2006, Martinez 2008, Ow 2008), levels were ~2-fold higher in the 16hr samples (Fig. 5A). No signal was detected for primary lin-4 from the lin-4(e912) null mutant, confirming the specificity of our lin-4 primary transcript detection methods. The mir-2 and mir-47 mature miRNAs have been reported to be constitutively expressed during development (Lim 2003), and we detected primary transcript expression for these miRNAs in the lin-4(e912) mutant and in all time points tested (Fig. 5A, B). The modest difference in lin-4 primary transcript levels in 4 versus 16 hours of development prompted us to further investigate the apparent uncoupling of lin-4 primary and mature RNA accumulation at early timepoints. It is possible that some transcripts detected by the RT-PCR assay are derived from unspliced
host gene (F59G1.4), or from intronic sequences that are inefficiently degraded after their removal by splicing. To test this hypothesis, we performed RT-PCR targeting intron 9 of F59G1.4 upstream of all cloned lin-4 start sites (Fig. 5A, B). We observed detectable levels of these intron-derived sequences at 4h, but no increase was observed at 16h (Fig. 5A).

Our observation that lin-4 primary transcripts accumulate early in L1, when mature lin-4 is undetectable, suggests that a post-transcriptional mechanism is involved in blocking the processing of these species into mature lin-4 miRNA. To further characterize the apparent post-transcriptional control of lin-4 biogenesis we performed a high-resolution temporal analysis of primary and mature microRNAs every 2hrs for the first 18hrs of C. elegans development (Fig. 5B). Consistent with the promoter-GFP (p1-a) results (Fig. 4B, C), lin-4 primary transcripts were detectable at all time points, but mature lin-4 did not accumulate in the same samples until about 12hrs of development (Fig. 5B).

The accumulation of mature lin-4 miRNA around 12 hrs of development is consistent with previous studies (Feinbaum 1999, Holtz 2009). A spike in primary transcript levels at 10hr of development marks an increase in transcriptional activity, presaging mature microRNA accumulation. However, the lin-4 primary transcript was present from 0 to 8 hours, starting at approximately 50% of the maximal levels observed at 14h, and gradually decreasing until the up-regulation at 10hr (Fig. 5B). Signal from the F59G1.4
host gene intronic sequence mirrored the early gradual decline observed in the
*lin-4* primary transcript, but did not intensify between 10 and 18hrs (Fig. 5B).
Additionally, RNAi depletion of the miRNA processing factors Drosha (*drsh-1*)
and Pasha (*pash-1*) resulted in accumulation of *lin-4*, *mir-2*, and *mir-47*
containing primary transcripts, but not the host gene intronic RNA (Fig. 5C).
Taken together, these results suggest that intronic species arising from the
host gene contribute to the early detection of *lin-4* containing transcripts, and
that up-regulation of *lin-4* primary transcript expression at the end of L1 is at
least in part due to transcripts originating close to *lin-4*. Functional analysis of
Drosha and Pasha depletion shows that host gene intronic sequences do not
accumulate, suggesting that the RNAs originating near *lin-4* (likely arising from
cloned start sites A-I in Fig. 4) are the preferred substrates for processing.
This data is consonant with a model in which host gene intronic transcripts are
ubiquitously expressed but resistant to Drosha processing, and *lin-4*-proximal
transcripts, which are efficiently processed by Drosha, are only expressed
when mature *lin-4* is needed late in L1. However, the evidence for a strict
separation in time of the different RNA species is lacking and indeed
promoter-GFP results and RACE cloning data (Fig. 4A,B,C) support the
existence of *lin-4*-proximal transcripts even in early L1-stage animals.
Therefore, we favor a model in which regulation of Drosha processing is based
on developmental timing in addition to differences in Drosha activity toward the
host-gene-intronic sequences versus the *lin-4*-proximal isoforms.
The presence of primary but not mature lin-4 miRNA in early L1 could result from failure of Drosha (DRSH-1) and its RNA binding protein partner Pasha (PASH-1) to recognize or process lin-4 primary transcripts early in worm development. In general, these biogenesis factors are present and active throughout the timecourse analyzed in Fig. 5B. By RT-PCR analyses we found steady state mRNA levels of drsh-1 and pash-1 to be unchanged through the first 18hr of worm development (Fig. 5B). Moreover, the constitutive expression of primary and mature mir-2 and mir-47 RNAs (Fig. 5B) argues that the microprocessor complex is active but is prevented from processing the early-accumulating lin-4 primary transcripts by an unknown mechanism. In order to directly test for activity of the microprocessor on early versus late-accumulating lin-4 primary transcripts, we adopted a ligation and cloning approach to detect endogenous Drosha cleavage products. As expected at 16hrs, 80% of cloned cleavage products mapped precisely to the predicted drsh-1 cleavage site (Fig. 5D, green arrowheads; site is marked with an asterisk). However, at 4hrs none of the cloned products corresponded to canonical drsh-1 cleavage sites (Fig. 5D, red arrowheads) and presumably represent random degradation products rather than drsh-1 activity on the transcript. These data directly demonstrate that early transcripts are not productively processed by drsh-1, while later transcripts are competent for microprocessor cleavage and mature miRNA biogenesis.
Regulation of lin-4 expression by the RBM28 homolog R05H10.2

Given the evidence that lin-4 expression is regulated in complex transcriptional and post-transcriptional patterns, we performed a candidate-based screen for potential regulators of lin-4 biogenesis. Candidate regulators of lin-4 expression were chosen on the basis of two criteria: 1) genes annotated as nucleic-acid binding by Gene Ontology (GO), or 2) genes implicated as miRNA pathway effectors through genetic or proteomic methods (Parry 2007, Baugh 2006, Gregory 2004, Shiohama 2007). Since insufficient lin-4 miRNA causes obvious developmental abnormalities, we eliminated nucleic acid binding candidates that lacked evidence for phenotypic effects in previously reported whole-genome RNAi screens (Kamath 2003, Sönnichsen 2005), assuming either that they are not regulators of the lin-4 pathway or that they are not sufficiently inactivated by feeding RNAi. This narrowed the candidate list from 1100 to 216. The addition of the miRNA pathway effectors from the literature brought the total candidates tested to 248.

For this screen, we created a GFP-based lin-4 sensor, using the 3'UTR of its established target lin-28 (Moss 1997). Transgenic worms ubiquitously expressing GFP fused to the lin-28 3'UTR containing an intact lin-4 complementary element (+LCE) exhibited decreased fluorescence from L1 to L3 in a lin-4 dependent manner, and deletion of this site (ΔLCE) prevented down-regulation of GFP (Fig 6A). To identify genes required for expression of functional lin-4 miRNA, we screened for RNAi candidates that resulted in failed
down-regulation of GFP fused to the +LCE 3'UTR. 24 / 248 RNAi clones passed this initial screen and were then tested for effects on GFP expression from the control ΔLCE reporter. Only 6 genes passed this secondary screen, indicating that many candidates affected reporter gene expression independently of lin-4.

To identify which stage of lin-4 biogenesis and/or function was affected by the six candidates, we performed Northern blot analysis of the expression of mature lin-4 in worms undergoing RNAi against each of these genes. While pash-1(RNAi) induced a ~1.5-fold decrease in mature lin-4 (Fig. 6B), inactivation of R05H10.2 caused a nearly 4-fold decrease in mature lin-4. RNAi depletion of the other five genes resulted in non-significant or variable effects on endogenous mature lin-4 levels, and they were not pursued further.

Further analysis of R05H10.2, which encodes a conserved RNA binding protein, revealed that it differentially affects miRNA biogenesis. Since R05H10.2(RNAi) results in delayed growth (see below), we analyzed the expression of lin-4 and other miRNAs in synchronized worms collected at the L2 stage based on gonad development. While depletion of R05H10.2 resulted in ~4-fold less mature lin-4 miRNA, steady state levels of the constitutive mir-2 and mir-47 miRNAs were elevated ~2-fold (Fig. 7A). The increase in mir-2 and mir-47 correlated with increased levels of primary transcripts for these miRNAs (Fig. 7B), indicating that R05H10.2 antagonizes accumulation of these RNA species. We also observed no significant change in levels of of lin-
primary and host gene intron 9 RNA in R05H10.2 RNAi conditions compared to Vector control RNAi conditions (Fig. 7B). These results suggest that depletion of R05H10.2 inhibits the post-transcriptional processing or stability of lin-4 microRNA. In contrast, the effect of R05H10.2 inactivation on mir-2 and mir-47 is consistent with an increase in transcription or stability of primary transcripts leading to greater levels of mature miRNAs.

The diminished expression of mature lin-4 under R05H10.2(RNAi) conditions is associated with mis-regulation of its target genes. Up-regulation of lin-4 miRNA at the end of the first larval stage promotes mRNA degradation of two characterized direct targets, lin-14 and lin-28 (Bagga 2005). We detected ~2-fold higher levels of lin-14 and lin-28 mRNAs in L2 worms undergoing RNAi against R05H10.2 (Fig.7), prompting us to examine the interplay between R05H10.2 and lin-4 target genes.

Interaction of R05H10.2 with the lin-4 developmental timing pathway

failure to develop a vulva, and the absence of a series of cuticular lateral ridges known as alae, which are normally produced at the L4-to-adult transition in wild-type animals. Given the aberrant production of mature lin-4 observed under R05H10.2(RNAi) conditions, as well as defective regulation of lin-4 targets, we asked whether depletion of R05H10.2 could induce lin-4 loss of function phenotypes in developing worms. Consistent with our prediction, depletion of R05H10.2 by RNAi in wild-type worms caused 33% of animals to be vulvaless (Table 1A), and nearly all that did form vulvas displayed protruding vulva defects consistent with retarded developmental timing (Table 1A). In addition, only 30% of worms depleted of R05H10.2 produced adult cuticles with complete alae, displaying instead incompletely formed patchy alae (60%) or no alae at all (10%) (Table 1B). These incompletely penetrant phenotypes are consistent with the incomplete nature of RNAi target gene inhibition, as detectable mature lin-4 remains even when R05H10.2 is depleted (Fig. 7A). Furthermore, the Northern blot analysis presented in Fig. 7A indicates that multiple microRNAs are affected by depletion of R05H10.2, and these might partially compensate for the defects induced by lack of lin-4 miRNA.

Worms with loss of function mutations in lin-14 and lin-28 exhibit precocious alae at the L4 stage, one stage earlier than in wild-type animals. The precocious expression of alae in lin-14(n179ts) worms is fully suppressed by loss of lin-4 miRNA (Table 2). Likewise, we found that depletion of
R05H10.2 in *lin-14(n179ts)* or *lin-28(n719)* worms resulted in strong suppression of the precocious alae phenotypes associated with these mutants (Table 2). Taken together, R05H10.2(RNAi) results in diminished expression of mature *lin-4* miRNA, mis-regulation of its targets and phenotypes consistent with a role for this factor in the *lin-4* developmental timing pathway.

**Integration of Developmental Timing and Organismal Growth by R05H10.2**

In addition to the developmental timing abnormalities in worms depleted of R05H10.2, we also observed overall delayed growth of the RNAi treated animals. This phenotype is distinct from the heterochronic defects observed in *lin-4, lin-28* or *lin-14* mutants, which repeat or skip stage specific somatic cellular events without exhibiting substantial irregularities in organismal growth or germline development (Ambros 1984, Ambros 1989). Under control RNAi conditions, WT (N2), *lin-14(n179ts), lin-14(n179ts);lin-4(e912), lin-4(e912)* and *lin-28(n719)* strains reached adult size with fully formed gonads by 54hrs of development at 25°C (Fig. 8B). At this same time point in N2 worms undergoing R05H10.2 RNAi, the majority of the population had only reached the L3 stage as judged by size and gonad development (Fig. 8A&B). Strikingly, over half of the *lin-14(n179ts)* and *lin-28(n719)* mutant populations undergoing R05H10.2 RNAi were scored as adults by the 54hr time point (Fig. 8B). However, R05H10.2(RNAi) in *lin-4(e912)* or *lin-14(n179ts);lin-4(e912)*
strains resulted in stalled development, indicating that increased *lin-14* activity contributes to organism-wide growth delay in worms depleted of R05H10.2.
Discussion

The biogenesis of microRNAs is increasingly recognized as a complex process with regulation occurring at many steps (Winter 2009). Cancerous tissue has been demonstrated to have reduced levels of most microRNAs (Medina 2008), and post-transcriptional regulation of microRNA biogenesis is disrupted in at least some cancers (Thomson 2006). These data are consistent with the fact that microRNAs can drive tissue differentiation and help define tissue identity (Lim 2005), and these findings highlight the importance of understanding post-transcriptional regulation of microRNA biogenesis. In the model organism *C. elegans*, we define R05H10.2 as a post-transcriptional regulator of *lin-4* and a key player in *C. elegans* small RNA metabolism.

The prevailing model of *lin-4* control, based upon promoter-GFP reporter assays, is that its expression is regulated primarily at the level of transcriptional control (Esquela-Kersher 2005, Ow 2008, Martinez 2008). This data is partly consistent with the work presented here, as we also detect a substantial level of transcriptional up-regulation of *lin-4* primary transcripts when the mature form appears. However, we also find evidence for extended enhancer elements that lie upstream of those incorporated into the reporter utilized in Esquela-Kersher 2005, and these elements drive expression ubiquitously during worm development. The promoter used for the GFP reporter of Ow et al. 2008 and Martinez et al. 2008 extended 2.4kb upstream
of the microRNA, and in these cases the observed GFP expression was limited to late L1 stage and later. There are two possible reasons for the incongruence between their findings and ours: 1. additional regulatory elements located even further upstream of our p1-a construct may control expression of \textit{lin-4}; or 2. the GFP expression reported in these studies largely ignored the weaker ubiquitous expression observed in this work. Interestingly, in their supplemental table 2, describing their \textit{lin-4} promoter-GFP fusion construct, Martinez et al. note that “Weak expression [was] detected ubiquitously”, leaving open to interpretation whether they refer to spatial or temporal expression. However, it is reasonable to assume that different investigators hold different thresholds at which GFP expression is deemed significant, and that these differences (and potentially also differences in optical detection apparatuses) may play a part in the differing results reported here.

Ow et al. 2008 describe the first transcriptional regulators of \textit{lin-4} biogenesis, \textit{flh-1} and \textit{flh-2}, which function semi-redundantly in the embryo to prevent premature expression of \textit{lin-4} microRNA in embryos. The motifs bound by \textit{flh-1} and \textit{flh-2} are present in both p1-a and p1-b, and consistent with transcriptional repression mediated by these elements, the GFP expressed from these constructs is excluded from embryos. However, \textit{flh-1} and \textit{flh-2} do not appear to regulate \textit{lin-4} expression after hatching, because no post-embryonic defects in regulation of the \textit{lin-4} target gene \textit{lin-14}, and no
heterochronic defects, were observed upon mutation of \textit{flh-1} and \textit{flh-2} (Ow et al. 2008). These data suggest that some other mechanism takes over the repression of \textit{lin-4} biogenesis in the early stages after hatching. Our work demonstrates an additional layer of post-transcriptional control which would prevent, until late L1, the processing of premature primary transcripts that may accumulate postembryonically in \textit{flh-1} and \textit{flh-2} mutants, thereby restricting mature microRNA expression to its proper temporal domain. This post-transcriptional regulation thus adds robustness to the expression of \textit{lin-4} microRNA such that very early post-embryonic timepoints, the first 10-12 hours of post-embryonic worm development, are protected from mature \textit{lin-4} microRNA regardless of transcriptional activity. It is tempting to speculate that the post-transcriptional control mechanism relaxes evolutionary constraints upon transcriptional control, and that the ubiquitous expression observed is “transcriptional noise” which evolved once the post-transcriptional control mechanism was in place.

A recent study showed that the RRM-domain protein \textit{hnRNP A1} is important for the Drosha-mediated processing of \textit{miR-18a}, but not for processing of other microRNAs expressed from the same primary transcript (Guil 2007). \textit{hnRNP A1} has been well characterized as a mediator of alternative splicing and mRNA export (Piñol-Roma 1992, Mayeda 1992, Cáceres 1994, Dreyfuss 2002) so it was surprising to find that it also plays a very specific role in the biogenesis of one particular microRNA. In an
analogous way, human and yeast homologs of R05H10.2 are implicated in ribosomal and spliceosomal biogenesis pathways, and the growth inhibition observed in *C. elegans* when R05H10.2 is inactivated is consistent with deficits in multiple gene expression pathways (see discussion below). The importance of multifunctional RRM motif proteins in specifically regulating the microRNA pathway appears to be an emerging motif of microRNA post-transcriptional regulation.

In addition, *hnRNP A1* and R05H10.2 are structural homologs due to their shared RRM motifs, so we wondered whether the mechanism of regulation of *lin-4, mir-2*, and *mir-47* by R05H10.2 might be functionally equivalent to the role played by *hnRNP A1*. Mechanistic analysis in Michlewski et al. 2008 revealed that *hnRNP A1* binding to miR-18a alters its structure such that it becomes accessible to the microprocessor. The binding of *hnRNP A1* to *miR-18a* primary transcript occurs at the terminal loop, but structural changes propagate to the Drosha cleavage site near the base of the hairpin (Guil 2007, Michlewski 2008). Michlewski et al. further noted that evolutionary conservation of terminal loop sequences is a signal of regulation by RNA-binding proteins. However, none of the microRNAs demonstrated to be regulated by R05H10.2 in this work have loop sequences that are conserved across phylogeny (data not shown). In addition, the best human match to R05H10.2 by protein-BLAST is not *hnRNP A1*, but rather is RBM28, a nucleolar protein involved in ribosome and spliceosome biogenesis (see
below). While these findings do not rule out an \textit{hnRNP A1}-like mechanism for R05H10.2, it suggests that R05H10.2 might be functioning in a novel way to mediate microRNA biogenesis.

The mechanism by which R05H10.2 modulates microRNA biogenesis is currently an open question, although the data suggests that regulation occurs at post-transcriptional steps for \textit{lin-4}, while \textit{mir-2} and \textit{mir-47} are regulated at earlier steps, most likely transcription or primary transcript stability. RRM domains most frequently mediate interactions with single-stranded RNA (and can be sequence-specific or nonspecific), although recent work has uncovered RRM domains that bind to proteins and even DNA (Maris et al. 2005, Trowitzsch 2008, Liu 2009). Intriguingly, the R05H10.2 protein product has a highly acidic domain located between RRM-1 and RRM-2, suggesting that this factor could act as a transcription factor upon DNA binding. Future work is also necessary to determine the precise step(s) regulated in \textit{lin-4} biogenesis, since the data presented in this work only demonstrates a decrease in the mature isoform with developmental consequences. One possibility is that Drosha or Dicer processing steps are targeted by R05H10.2. However, as seen in Fig. 7B, \textit{lin-4} primary transcripts remain relatively constant upon inactivation of R05H10.2; likewise no accumulation of precursor forms was observed on Northern blots (data not shown). These data suggest that if Drosha or Dicer processing steps are regulated by R05H10.2, the stability of primary or precursor transcripts must
also be regulated in a compensatory direction. Alternatively, RO5H10.2 may primarily affect the stability of the mature microRNA without altering the processing efficiency at any step.

The expression of RO5H10.2 appears to be ubiquitous during early C. elegans development. We performed RT-PCR of developmentally-staged samples from Fig. 5B and saw strong expression of RO5H10.2 mRNA at every stage (results not shown). In agreement with this result, an RO5H10.2 promoter-GFP fusion protein has been reported to be expressed in most tissues during development but is excluded from neurons (Hunt-Newbury 2007). This finding is intriguing because the lin-4 promoter-GFP fusions created for the current study show ubiquitous expression in nerve tissue, with temporally regulated GFP expression in seam, vulva, and pharynx cells (Fig. 4B). This suggests a model in which RO5H10.2 facilitates lin-4 microRNA processing or stability in all but neuronal tissues. However, this model is difficult to reconcile with the data for mir-2 which is a primarily neuronal microRNA, (Martinez 2008) and which also responds to RO5H10.2 depletion but oppositely to lin-4 (Fig. 7A). While non-cell autonomous effects are one potential explanation for some of these findings, the fact that the effects on two neuronal microRNAs are directly opposite to each other is an indicator of the complexity of the interplay between these microRNAs and RO5H10.2.

Inactivation of RO5H10.2 was observed to cause increased small RNA species on ethidium bromide stained polyacrylamide gels, which may be
tRNAs but were not specifically identified (data not shown). It is possible that the up-regulation of small RNA species (including \textit{mir-2} and \textit{mir-47}) upon depletion of R05H10.2 is indicative of a role in specific RNA degradation pathways. Intriguingly, however, the effect of R05H10.2 depletion on mature \textit{lin-4} is counter to the trend seen for other small RNAs, which accentuates the difference in function that R05H10.2 exhibits toward \textit{lin-4} relative to \textit{mir-2} and \textit{mir-47}. It is likely that mechanistic understanding of the role of R05H10.2 in microRNA biogenesis will require characterization of its interactors—does it directly bind to primary or precursor transcripts of microRNAs, or is it indirectly controlling them by regulating transcription, perhaps directly as a DNA-binding transcription factor? Or perhaps it functions by mediating important protein interactions which regulate microRNA biogenesis or stability. These questions and more will require careful analysis of the RNA, DNA, and protein interactors of R05H10.2.

The role of R05H10.2 in regulating organismal growth mediated by \textit{lin-14} is both significant and novel. Previous work has characterized \textit{lin-14} as a heterochronic regulator, controlling the timing of early developmental events within the somatic tissues including hypodermis, muscle, neurons, intestine, and vulva (Ambros 1984, Ambros 1989, Ruvkun 1989, Slack 1997). However, heterochronic defects are classically defined relative to germline development, which is largely unaltered in these mutants (Ambros 1984). In this work, we present evidence that \textit{lin-14} is also a regulator of germline development.
Retarded heterochronic defects seen in *lin-4* mutants may be thought of as failure of somatic tissues, but not germline tissue, to advance to later developmental stages and thus, to grow; the growth defect induced by R05H10.2 is a repression of growth in both soma and germline. However, we find that *lin-14* mutation can suppress the R05H10.2-RNAi-induced growth defect for the whole animal, an indication of a hitherto unknown function for this protein in regulating germline growth. While *lin-28* also suppresses the growth defect induced by R05H10.2 RNAi, we propose that this effect may be mediated through *lin-14* which exhibits a tight coupling in levels to *lin-28*. Several papers have documented a correlation between *lin-14* and *lin-28* protein levels (Arasu 1991, Moss 1997, Seggerson 2002), and we suspect that the suppression of the growth defect in *lin-28(n719)* mutant animals under R05H10.2 RNAi conditions may be a consequence of reduced *lin-14* protein levels in these animals. However, we cannot rule out a direct role for *lin-28* as well as *lin-14* in repressing germline and somatic growth.

To integrate these findings, we propose the model shown in Fig. 9: R05H10.2 functions as a master control protein for two repressors, one functioning through *lin-4* in regulating developmental timing, and the other regulating germline development and overall organismal growth. Both repressive branches operate through *lin-14* and *lin-28*, which serve to promote earlier fates and also repress later fates, of both soma and germline. This explains how *lin-4* mutants have relatively normal growth of germline tissues
relative to somatic ones—the LIN-14/LIN-28 activity that would repress germline growth in \textit{lin-4} mutants is down-regulated by a \textit{lin-4}-independent mechanism so that germline growth can progress, yet LIN-14/LIN-28 activity controlling the soma remains high, effectively blocking those tissues from reaching later developmental fates. This results in a disjointed organism in which soma fails to advance to later stages but the germline is free to grow. It is only when both \textit{lin-14/lin-28} repressive pathways are inactivated by R05H10.2 RNAi that the growth of the whole animal is impaired because both soma and germline-controlling pathways retain high levels of LIN-14/LIN-28. The net result is a re-coordinating of tissues of the organism which is now holistically blocked from advancing to later fates; therefore, organismal growth stalls. This analysis suggests that \textit{lin-14} serves to promote earlier life stages and prevent later ones in both soma and germline; only when \textit{lin-14} protein is reduced in pathways controlling both tissue types can later life stages be reached, allowing the organism to grow as an integrated whole.

The closest human and yeast homologs of R05H10.2, RBM28 and Nop4p, respectively, are nucleolar proteins with roles in ribosomal RNA processing and spliceosomal RNA biogenesis (Sun 1994, Sun 1997, Damianov 2006, Nousbek 2008). While we have not analyzed potential roles of R05H10.2 in mediating ribosomal RNA biogenesis or spliceosomal RNA biogenesis, we acknowledge that these may contribute to the growth defect observed when R05H10.2 is inactivated by RNAi. However, the ability of \textit{lin-14}
inactivation to substantially suppress the observed growth defect points to a specific and previously unknown role of the heterochronic pathway in mediating organismal growth. The work presented in this paper reveals a larger context for the heterochronic pathway, situating it within a network of genes controlled by R05H10.2, which coordinate the many tissues of the worm to orchestrate overall development of the whole organism.

Acknowledgement

Without the generous financial and intellectual support of Dr. Amy E. Pasquinelli the work presented in this chapter would not have been possible. She is the senior co-author, with the dissertation author as the primary author, for this material currently in review for publication in Development, 2009.
Materials and Methods

Northern blotting:

Total RNA was separated by electrophoresis in 11% denaturing polyacrylamide gels and transferred to nylon membranes (Zeta-Probe GT, Biorad, Hercules, CA). Starfire labeled oligo probe (IDT) was hybridized to the blot in 5X SSC, 7% SDS, 0.02 M sodium phosphate, 1X Denhardt's solution for approximately 12 h at 50°C and then washed twice in 3X SSC, 5% SDS, 0.025 M sodium phosphate, 10X Denhardt's solution at 50°C, before a final two washes in 0.4X SSC, 1.4% SDS. Blots were exposed to a Typhoon Phosphorimager screen overnight and/or to BioMax MR X-ray film overnight. Probes utilized were as follows:

A1916 (lin-4 starfire): 5'-TCACACTTGAGGTCTCAGGGA/STARFIRE-3'
A1356 (miR-2 starfire): 5'-GCACATCAAAGCTGGCTGTGATA/STARFIRE-3'
A1684 (miR-47 starfire): 5'-TGAAGAGAGCGCCTCCATGACA/STARFIRE-3'
A479 (5.8s rRNA F): 5'-CTAGCTTCAGCGATGGATCGGTTGC-3'
A480 (5.8s rRNA R): 5'-GAACCAGACGTACCAAACTGGAGGCCC-3'

5.8s rRNA PCR product was labeled with the Prime-It kit (Stratagene) as per manufacturer's instructions.

RT-PCR:

Total RNA from synchronized, staged populations was extracted using TRlzol reagent (GIBCO-BRL). All RNAs were subjected to two sequential deoxyribonuclease digestions using RQ1 DNAse (Promega). cDNA synthesis was performed out by random-primed RT (Invitrogen Superscript II, performed
according to manufacturer's instructions). The resulting RT products were
diluted 5-fold and used as templates for PCR with varying primer sets. All PCR
reactions were carried out with Promega GoTaq as specified in manufacturer's
instructions but supplemented with MgCL to 1.5mM final concentration. All
reactions were carried out at 58°C annealing 30 seconds, followed by
extension at 72°C for 60sec per kb. Reaction cycles were adjusted by primer
set to prevent saturation. Cycle numbers were as follows: lin-4 (A2015+A65),
30 cycles; F59G1.4 intron #9 (A463+A464), 33 cycles; miR-2 (A1549+A1550),
28 cycles; miR-47 (A1551+A1552), 28 cycles; drsh-1 (A145+A146), 30
cycles; pash-1 (A1148+A1149), 30 cycles; actin (A810+A811), 20 cycles; 18S
rRNA (A839+A840), 5 cycles, lin-14 (A425+A426), 25 cycles, lin-28
(A424+A427), 30 cycles. Primer sequences utilized for PCR are as follows:

A65: 5'-CGTTTGACCCTTTTCCCCGAAATACC-3'
A2015: 5'-TTTAGGCTTCCCGGCTTGC-3'
A463: 5'-GAGCGGAAGCTTAAGGGGATGATG-3'
A464: 5'-GTATAGCGGTATGAAATGAG-3'
A001549: 5'-CCTCTGCACTGTTGTTGCTAC-3'
A001550: 5'-TTATATCTACAGTCCGAATATTG-3'
A1551: 5'-GCTTCCTGGCCTGGATCTAC-3'
A1552: 5'-ACACGCGGAACACCTAGTGTTGAAG-3'
A145: 5'-CGCTTTTCATCTTCCGAAAAGC-3'
A146: 5'-TGAAAGGAAACATCTGTTGAG-3'
A1148: 5'-AGACGTCGATGATGAAGAAGAGGATG-3'
A1149: 5'-AGGGATGATGATGAAATCTGC-3'
A810: 5'-GTGTTTCCCATTGGTTGGAGAAGAC-3'
A811: 5'-GCACTTTGGTGAAAGCTTGGAGG-3'
A839: 5'-GCCGACCGCTTGGAGAGCTTGGAC-3'
A840: 5'-GGTCAAGACTAGGGCGCGTAAATGAAC-3'
A1099: 5'-TTATTGGAACCTAGCTCCCAAGGGG-3'
A1111: 5'-CCCTGGCTTGGCAACCTAGG-3'
A892: 5'-ATGGCTTGAGCCCTGATTCTGCTGAAATATG-3'
A974: 5'-TGACTGCTGATCCTGACGACTGATAATTTTGATAG-3'
A904: 5'-GAACCGTTGCACTGCAGTCCAGCCTG-3'
A908: 5'-TATTGGAAACAAAGAGACCTTAGGTTTGGAGGATGGTAG-3'
RACE methods:

Standard GeneRACER (Invitrogen) protocols were utilized for the identification of transcription start sites and polyadenylation site. For detection of Drosha cleavage products, a modified 5'RACE protocol (GeneRACER, Invitrogen) was adapted by the elimination of the initial CIP and TAP treatments. Instead, 4h and 16h RNA samples were double RQ1-DNAse treated and added to the lyophilized RNA oligo. Ligation and all following steps were as per manufacturer's instructions. Reverse transcription was performed with A65; initial PCR reactions were carried out with the 5' oligo from GeneRACER kit plus A1111; products of this reaction were column purified on a Quiagen PCR cleanup kit and used as template for a nested PCR using the 5' nested oligo from the GeneRACER kit plus A1099. Products of this reaction were used as substrates for TOPO cloning and sequencing. Cleavage products were analyzed only if they contained the 5' RNA oligo sequence ligated to lin-4 sequence.

lin-4 promoter fusion constructs:

The p1-a::GFP plasmid (pJRB1) was created by PCR amplification of ~1kb of sequence upstream of lin-4 with A321+A323, and the p1-b plasmid
(pJRB3) was created by PCR amplification of ~400nt with A321+A322. These inserts, and Fire plasmid pPD95.75 were digested with Age-I and ligated by standard subcloning techniques. Plasmids were linearized with Spe-I and injected at 50ng/ul per plasmid, along with the Pha-1 rescuing plasmid pBX (Granato 1994) at 50ng/ul into pha-1(e2123) mutant animals, and transgenic animals were identified by rescue of pharynx development in F1 animals at 25°C.

**lin-4 sensor and transgenic generation:**

pJRB15 was constructed from Fire Vector pPD136.15 (L4809), which contains the let-858 promoter driving GFP (containing 3x SV40 NLS) which is followed by the let-858 3'UTR. This vector was digested with Nhe-I and Apa-I to remove the let-858 3'UTR, which was then replaced by the lin-28 3'UTR that was amplified from genomic DNA with oligos A892 and A974. The –LCE construct (pJRB16) was generated by PCR stitching together two fragments overlapping in the region of the lin-4 site but containing an exact deletion of lin-4 complementary sequences. This was done by PCR off two fragments, (A904+892) and fragment (A908+A974) amplified with outer primers A892+A974. The –LCE lin-28 3'UTR was digested with Nhe-I and Apa-I and ligated into pPD136.15 as described above. For injection, pJRB15 and pJRB16 were digested with Pvu-I to linearize them, and were coinjected at 25ng/ul along with the pha-1 rescuing plasmid pBX (Granato 1994) at 75ng/ul.
This mixture was injected into the gonads of *pha-1(e2123)* mutant animals, and transgenic animals were identified by survival at the non-permissive temperature 25°C.

**RNA interference:**

RNAi was carried out as described in Kamath et al. 2003, with the only difference being that plate IPTG was increased from 1mM to 5mM. The first generation was grown from the L4 stage on lawns of bacteria expressing double-stranded RNA targeting the corresponding gene, then egg-prepped by hypochlorite solution treatment (1% NaClO (~140mM), 500mM KOH), hatched overnight in bacteria-free M9 medium, and then seeded back onto RNAi food for the indicated time before harvesting for analysis.

**Alae scoring:**

Staged populations of wild-type and mutant worms were grown from the L4 stage on bacteria lawns as described above, then egg-prepped, hatched overnight in M9 media, and plated back on the same bacteria lawn. Animals were washed to slides and scored for alae formation and gonad growth.
Figure 4. Analysis of lin-4 transcription. A. Transcription start sites for lin-4 within intron 9 of host gene F59G1.4 were mapped with 5' RACE at 4, 8, 16, and 40 hrs of development. Green box is rescue fragment from Lee et al., 1993. Position of e912 null mutation is indicated, but extent of the deletion is unknown (indicated by dotted line). White boxes indicate regions cloned into GFP expression constructs shown in B and C. B. Graphical summary of expression of GFP from promoter fusion constructs, where green indicates consistently detectable GFP. Light blue shading over times indicate presence of corresponding mature lin-4 microRNA. C. Images of transgenic worms bearing plin4-GFP constructs. p-1a, p-1b as in A. Single cells expressing GFP in early stages of p-1b transgenics appear to be non-specific as they were variable and fainter than the specific tissue expression observed at 18hr and later timepoints.
**Figure 5.** Evidence that *lin-4* biogenesis is regulated post-transcriptionally. A. RT-PCR detection of *lin-4* primary transcripts at both 4h and 16h of wild-type but not *lin-4(e912)* 16hr samples. *mir-2* and *mir-47* included as controls which are ubiquitously expressed and present in the *lin-4(e912)* background, and all experiments were performed with 4 independent populations (N2) or 3 independent populations (*lin-4(e912)*) of staged animals. F59G1.4 intron #9 is detected with oligos residing upstream of *lin-4* primary transcripts and is deleted in *lin-4(e912)* mutant animals. B. Detection of mature and primary microRNAs during a two-hour time-course of the first 18hrs of worm development by RT-PCR and Northern blot. All analysis was performed with the same RNA samples. Primary *lin-4* transcripts are detectable at all time points but undergo an apparent transcriptional upregulation at 10hr of development, corresponding to the detection of mature microRNA at 12h in the PAGE Northern. Notice that F59G1.4 intron #9 mirrors the gradual decrease observed in *lin-4* primary transcripts between 0-8hrs, and thus may contribute to this early signal. The lack of mature *lin-4* before 10hrs might be explained by a lack of Drosha-Pasha activity at these time-points, but the presence of both mature and primary transcripts for *mir-2* and *mir-47* in these samples suggests instead that a specific regulatory mechanism controls *lin-4* processing. C. RNAi of Drosha and Pasha at 16hrs of development induces upregulation of primary transcripts for *lin-4, mir-2*, and *mir-47* but does not cause an increase in levels of intron #9 from F59G1.4. D. Direct mapping of Drosha cleavage by a modified 5’RACE protocol. RNA adaptors were ligated directly to the 3’ cleavage products generated by Drosha activity, followed by cloning and sequencing. At 16hrs, the majority of cleavage products map to the predicted Drosha cleavage site, but at 4hrs no cleavage products corresponding to this site were observed.
B. Figure 6. A candidate-based RNAi screen for factors involved in lin-4 biogenesis. A. Design of a GFP sensor for lin-4. The promoter of let-858 drives expression in all tissues at all times, and was joined to GFP appended with the lin-28 3'UTR. It is known that lin-4 down-regulates lin-28 through a single site in its 3'UTR (Moss 1997), the lin-4 complementary element (LCE). As a control, we created an identical sensor with the LCE deleted (ΔLCE). The screen was carried out in two parts: First, +LCE sensor-bearing strain was grown on RNAi clones targeting candidates (selected as described in text), and testing for persistence of GFP at L3 stage relative to control vector RNAi. Second, candidates passing the first level were tested for effects on the –LCE sensor-bearing animals to ensure that the effects observed on GFP levels were dependent on the intact lin-4 binding site. B. Six candidates passed both levels of the screen and are shown on the table. Northern blot analysis of mature lin-4 was performed to identify candidates affecting lin-4 biogenesis, and only R05H10.2 inactivation consistently depleted mature lin-4.

<table>
<thead>
<tr>
<th>RNAi clone</th>
<th>lin-4 fold change relative to control</th>
<th>comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pasha</td>
<td>-1.64</td>
<td>microRNA processing factor, Drosha interactor</td>
</tr>
<tr>
<td>R05H10.2</td>
<td>-3.85</td>
<td>3x RRM motif protein</td>
</tr>
<tr>
<td>M88.5</td>
<td>-1.20</td>
<td>KH and Q/N (Prion-like) domains</td>
</tr>
<tr>
<td>zim-3</td>
<td>-1.15</td>
<td>Zinc-finger in Meiosis (zim)—important for homolog synapsis</td>
</tr>
<tr>
<td>F53H4.6</td>
<td>-1.79</td>
<td>DNA/RNA helicase</td>
</tr>
<tr>
<td>vbh-1</td>
<td>-1.05</td>
<td>DEAD/DEAH helicase, important for germline development</td>
</tr>
<tr>
<td>Y46G5A.13</td>
<td>-1.16</td>
<td>3x RRM motif protein</td>
</tr>
</tbody>
</table>
Figure 7. Characterization of the role of R05H10.2 in microRNA biogenesis. A. Northern blot of mature microRNAs in worms on control (vector) RNAi and R05H10.2 RNAi bacteria. Because of a severe growth defect observed under R05H10.2 RNAi conditions, equivalent developmental stages were compared as determined by gonad development. Mature *lin-4*, *mir-2*, and *mir-47* are differently affected by inactivation of R05H10.2. B. Primary transcript analysis by RT-PCR demonstrates that while *mir-2* and *mir-47* are increased two-fold or more, *lin-4* primary transcript remains relatively unchanged, as well as F59G1.4 intron 9. However, *lin-4* target genes *lin-14* and *lin-28* are both upregulated at least two-fold.
Figure 8. Interaction of R05H10.2 with the heterochronic pathway. A. representative images of animals reared for 2 generations on R05H10.2 RNAi bacteria at 25°C and photographed at 54hrs. B. Population structure from experiment described in A. Vector control RNAi conditions allow all animals to reach adulthood, while on R05H10.2 N2 worms are mostly L3, an effect that is strongly suppressed in lin-14 in and lin-28 mutant backgrounds. All stages were scored based on gonad development.
Figure 9. A model for the role of *lin-14* and *lin-28* in modulating germline and somatic growth. *lin-14*, and possibly *lin-28*, functions as a repressor of growth (or functions to repress advancement of tissues to later developmental fates). Heterochronic mutants like *lin-4* only affect the somatic growth arm of the integrated pathway, but down-regulation of *lin-14/lin-28* controlling organismal growth still occurs via a yet-undefined repressor. This allows germline growth to occur while somatic tissues become stalled in L1 reiterations, failing to advance to maturity. Since R05H10.2 controls both arms of the pathway, it induces whole-organism growth to stall. Removal of *lin-14* or *lin-28* de-represses growth in the whole organism as seen in Fig. 5.
**Table 1A.** Vulva is absent or defective in adult N2 on R05H10.2 (RNAi). Vulval development was analyzed only in animals that were adult-stage by gonad development: 48 hours for Vector (RNAi) and 94 hours for R05H10.2 (RNAi).

<table>
<thead>
<tr>
<th>RNAi</th>
<th>strain</th>
<th>vulvaless</th>
<th>protruding vulva</th>
<th>normal vulva</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vector Ad</td>
<td>N2</td>
<td>(n=20)</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>R05H10.2 Ad</td>
<td>N2</td>
<td>(n=30)</td>
<td>33%</td>
<td>60%</td>
</tr>
</tbody>
</table>

White = Vector (control) RNAi  
Grey = R05H10.2 RNAi.

**Table 1B.** R05H10.2 (RNAi) induces alae phenotypes in N2 adults. Animals were the same as in Table 1A.

<table>
<thead>
<tr>
<th>RNAi</th>
<th>strain</th>
<th>no alae</th>
<th>patchy alae</th>
<th>full alae</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vector Ad</td>
<td>N2</td>
<td>(n=20)</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>R05H10.2 Ad</td>
<td>N2</td>
<td>(n=30)</td>
<td>10%</td>
<td>60%</td>
</tr>
</tbody>
</table>

White = Vector (control) RNAi  
Grey = R05H10.2 RNAi.
Table 2: Suppression of the precocious alae of *lin-14* and *lin-28* mutants by R05H10.2 (RNAi). Only L4 stage animals (assayed by gonad development) were scored.

<table>
<thead>
<tr>
<th>Strain scored at L4 stage</th>
<th>No alae</th>
<th>Patchy alae</th>
<th>Full alae</th>
<th>Hours of development at time of scoring</th>
</tr>
</thead>
<tbody>
<tr>
<td>N2 (n=17)</td>
<td>100%</td>
<td>0%</td>
<td>0%</td>
<td>30 hr</td>
</tr>
<tr>
<td>N2 (n=11)</td>
<td>100%</td>
<td>0%</td>
<td>0%</td>
<td>77 hr</td>
</tr>
<tr>
<td><em>lin-14(n179)</em> (n=13)</td>
<td>15%</td>
<td>54%</td>
<td>31%</td>
<td>40 hr</td>
</tr>
<tr>
<td><em>lin-14(n179)</em> (n=13)</td>
<td>86%</td>
<td>7%</td>
<td>7%</td>
<td>54 hr</td>
</tr>
<tr>
<td><em>lin-14(n179);lin-4(e912)</em> (n=16)</td>
<td>100%</td>
<td>0%</td>
<td>0%</td>
<td>40 hr</td>
</tr>
<tr>
<td><em>lin-14(n179);lin-4(e912)</em> (n=10)</td>
<td>90%</td>
<td>10%</td>
<td>0%</td>
<td>77 hr</td>
</tr>
<tr>
<td><em>lin-4(e912)</em> (n=15)</td>
<td>100%</td>
<td>0%</td>
<td>0%</td>
<td>40 hr</td>
</tr>
<tr>
<td><em>lin-4(e912)</em> (n=8)</td>
<td>100%</td>
<td>0%</td>
<td>0%</td>
<td>77 hr</td>
</tr>
<tr>
<td><em>lin-28(n719)</em> (n=18)</td>
<td>28%</td>
<td>22%</td>
<td>50%</td>
<td>40 hr</td>
</tr>
<tr>
<td><em>lin-28(n719)</em> (n=12)</td>
<td>100%</td>
<td>0%</td>
<td>0%</td>
<td>54 hr</td>
</tr>
</tbody>
</table>

White = Vector (control) RNAi
Grey = R05H10.2 RNAi.
Chapter 2: Analysis of the mechanism of target regulation by \textit{lin-4}
Introduction

In 2005 the prevailing model for microRNA regulation of target mRNAs held that the fate of targeted mRNAs was determined strictly by the degree of base-pairing to the regulating microRNA. In this model, targets with a high degree of complementarity would undergo mRNA degradation through an RNAi-like mechanism of endonucleolytic cleavage, but targets with mismatched interactions with microRNAs would be subjected to translational repression, leaving target mRNA unaffected (Bartel 2004, Doench 2003, Zeng 2003, Hutvágner 2002). Evidence supporting this paradigm for endogenous microRNAs and targets came primarily from lin-4-lin-14 and lin-4-lin-28 interactions in C. elegans (Wightman 1993, Olsen 1999, Seggerson 2002). These studies utilized RNAse protection assays at different developmental times (before and after lin-4 microRNA was expressed) to analyze lin-14 and lin-28 mRNA levels, and used polysome profiling to measure any changes in translation of these targets. Neither the total mRNA measured nor the polysome profiles showed any significant alterations upon lin-4 induction, implicating a model in which microRNAs mediate translational repression of steps after the initiation of translation.

It came as a surprise, then, when Dr. Pasquinelli performed preliminary analysis of the let-7 target gene lin-41 in wild-type and let-7 mutant worms and found a striking increase in lin-41 in the microRNA mutant animals. Based on this observation the Pasquinelli lab put together a collaborative effort to
analyze microRNA target regulation in *C. elegans*. Because of my experience studying the biogenesis of the *lin-4* microRNA, I chose to analyze its regulation of target genes *lin-14* and *lin-28*. Consistent with the initial observations with *lin-41* and *let-7*, I observed that both *lin-14* and *lin-28* mRNA levels were strongly reduced by *lin-4*-mediated repression and that this repression was both dependent on the microRNA and (in the case of *lin-14*) the presence of *lin-4* complementary elements in the 3'UTR of the target gene (see Fig. 11).

These observations, along with similar analysis of the *let-7* mediated regulation of *lin-41*, were published in the journal *Cell* in 2005 (Bagga 2005), and this paper was among the first to report degradation of target mRNAs by microRNAs. Many papers have since characterized complex and—at times—contradictory proposed mechanisms by which microRNAs can mediate both translational repression and target degradation (see dissertation introduction for detailed description). Given our observations that microRNAs cause decreases in target mRNAs without apparently causing endonucleolytic cleavage, we postulated a role for microRNA-mediated deadenylation and mRNA decay. This model has been confirmed by subsequent studies (Wu 2006, Wakiyama 2007, Eulalio 2009a), but other work has demonstrated that translational repression is also a bona fide mechanism of target regulation by microRNAs (Ding 2009, Maroney 2006 Peterson 2006 Nottrott et al. 2006), and that some targets seem to be primarily regulated by translation repression while others are primarily regulated by degradation (Eulalio 2009a). Other
workers report that both translational repression and degradation occur in the same targets (Ding 2009, Baek 2008, Selbach 2008).

The over-proliferation of models for microRNA mediated target regulation (for review, see Filipowicz 2008) are an indication either that many different mechanisms are genuinely utilized by microRNAs to regulate their targets, or that experimental artifacts are endemic to many of these studies, and thus the real signal is buried under a large amount of experimental noise. Marilyn Kozak has weighed in on this topic, in a review paper titled “Faulty old ideas about translational regulation paved the way for current confusion about how microRNAs function” (Kozak 2008) and she makes two points: 1. different transfection techniques, transgenic constructs, cell lines, and cell-free systems tend to support different models of target regulation by microRNAs, and many of these techniques have known methodological problems, and 2. the question we must continually ask is, how does the experimental system being utilized compare to endogenous microRNAs regulating endogenous targets? (Kozak 2008). In her view, the artifacts caused by transfection of cultured cells, transgenic overexpression of target constructs, known problems with IRES’s (Internal Ribosome Entry Sites), and problems with cell-free reconstitution of translational repression (notably RNAse activity) are sufficient to question whether microRNAs drive translational control at all. (She does, however, point out that target degradation is well-established by physiologically relevant studies such as ours). From Kozak’s perspective, the confusion surrounding
the mechanism of translational control (reviewed in the introduction to this
dissertation) is a symptom of a field in crisis rather than a healthy foment of
ideas, and progress will only be made by analyzing genuine microRNAs
regulating their endogenous targets in systems that are as physiologically
relevant as possible. In this, our work has a distinct advantage, and our data
was important in providing physiologically relevant insights into the
mechanisms by which target genes are down-regulated by microRNAs.

**Results: The lin-4 miRNA causes decreased levels of its mRNA targets**

The observation that *let-7* mediated repression of *lin-41* induced
mRNA degradation led us to investigate *lin-4* and how it mediates regulation of
its two well-characterized targets in *C. elegans*, *lin-14* and *lin-28*. These genes
contain sites in their 3'UTRs that are complementary to *lin-4* and these sites
mediate negative regulation by this miRNA (Ha et al., 1996; Moss et al., 1997;
Wightman et al., 1993) To test whether regulation occurs at the level of mRNA
degradation, I first collected synchronized worm populations at different
developmental times, and this RNA was analyzed by two different Northern
blotting techniques: agarose Northern to evaluate the levels of full-length
target mRNAs over developmental time, and PAGE Northern to evaluate *lin-4*
microRNA expression in the same samples. Strikingly, the timing of
accumulation of mature *lin-4* correlated precisely with a strong decrease in *lin-
and lin-28 mRNA levels (Fig. 10), in contradiction to earlier reports (Moss et al., 1997; Olsen and Ambros, 1999; Seggerson et al., 2002; Wightman et al., 1993).

To test whether the decrease in lin-14 lin-28 mRNA levels was due to lin-4 activity, I staged wild-type and lin-4(e912) mutant animals as starved L1 (st. L1), 4h L1 (prior to mature lin-4 induction) and 18h L2 (lin-4 present in wild-type). While both lin-14 and lin-28 were strongly decreased in wild-type animals, lin-4(e912) mutants displayed strongly abrogated down-regulation. lin-28 displayed a slight decrease even in the mutant background which may be evidence of other regulators, possibly microRNAs, that that target this mRNA. Interestingly, the observed fold decrease in lin-14 mRNA levels is consistent with the four-fold reduction in LIN-14 protein levels mediated by lin-4 microRNA (Wightman et al., 1993).

In order to test whether the down-regulation of lin-14 mRNA is dependent on the lin-4 complementary sites (LCEs) in the lin-14 3'UTR, we evaluated the behavior of the lin-14 mRNA in a mutant that deletes all LCE sites. The lin-14(n355gf) allele contains a breakpoint mutation in the lin-14 3'UTR that removes all predicted lin-4 complementary sites and is associated with failure to down-regulate LIN-14 protein expression at the L2 stage (Wightman et al., 1991; Wightman et al., 1993). While some down-regulation still occurs in this mutant background (approximately 1.5-fold), it was much reduced relative to wild-type, which exhibited over 5-fold down-regulation in
lin-14 mRNA levels. Taken together these data demonstrate that down-regulation of lin-14 and lin-28 mRNA levels is mediated by the lin-4 microRNA and also, in the case of lin-14, by the specific binding sites in the 3’UTR.

Discussion

lin-4 mediates degradation of target mRNAs

An early model for gene regulation by miRNAs held that partial base-pairing between a miRNA and its target mRNA induces translational repression without destabilization of the mRNA. This model was based primarily on data indicating that the relative level of lin-14 mRNA remained constant while the protein levels declined in response to lin-4 recognition of the 3’UTR complementary sites (Wightman 1993). Furthermore, the association of lin-14 mRNA with polyribosomes appeared unchanged upon lin-4 mediated inhibition of protein expression (Olsen 1999). The generality of this type of gene regulation was supported by reports that the lin-28 mRNA levels and polyribosome profile were not altered when lin-4 blocked LIN-28 protein accumulation (Moss 1997; Seggerson 2002). However, the data shown in all of these reports were generated by RNase protection experiments, which do not demonstrate the existence of full-length, intact mRNA species. The work presented in this report, and other data from analysis of let-7 and lin-41, presented together in Bagga et al., is evidence that
mRNA degradation does occur *in vivo* in the context of the *C. elegans let-7* and *lin-4* miRNAs and their targets, which contain recognition sequences of imperfect complementarity. Our Northern blot analysis of endogenous, full-length microRNA targets in *C. elegans* demonstrated that they are subject to reduction at the mRNA level. These results argue against the simple model that mRNA fate is entirely dependent upon the degree of base-pairing with a miRNA, although they leave open the possibility that translational repression precedes degradation. However, the very tight correlation between *lin-4* microRNA expression and *lin-14* degradation suggests that such translational repression must be extremely short-lived as the target mRNA is destroyed nearly as soon as the mature microRNA becomes detectable on Northern blot. In a 2-hour timecourse, mature microRNA is detectable at 12hr and both *lin-14* and *lin-28* mRNAs are undetectable by 14hr (Chendrimanda 2007, Holtz 2009). These data are broadly consistent with the findings reported in Baek et al. 2009 who examined global protein and mRNA effects upon transfecting microRNAs into HeLa cells. The authors noted that proteins most strongly affected by a given microRNA tended to also exhibit a significant amount of degradation of the corresponding mRNA. A similar result was observed in Selbach et al. 2008, who reported a significant amount of target degradation as a result of transfecting microRNAs into human cells.

The levels of *lin-14* mRNA degradation mediated by *lin-4* agree with the amount of protein down-regulation that is *lin-4* dependent. While LIN-14
protein levels decrease ~10-fold from the L1 to the L2 stages of development (Olsen and Ambros, 1999; Wightman et al., 1993), this reduction is only partially dependent on *lin-4* miRNA (Wightman et al., 1993). In fact, Wightman et al. demonstrated that the difference in LIN-14 protein down-regulation between wildtype and *lin-4(e912)* mutants was four-fold (Wightman et al., 1993), which is comparable to the fold decrease in *lin-14* mRNA levels reported here (Fig. 11).

We cannot fully explain the discrepancy between the results shown in this chapter and the previous studies of *lin-4* mediated regulation (Wightman 1993, Olsen 1999, Seggerson 2002). For our Northern experiments, we used probes that covered regions analyzed in the RNase protection experiments. However, we did not detect stable degradation products corresponding to these genes that could account for the previous conclusion that the levels of *lin-14* and *lin-28* mRNAs are not significantly altered (Olsen and Ambros, 1999; Seggerson et al., 2002; Wightman et al., 1993). However, Olsen & Ambros report a 2.3-fold (+/- 0.96) reduction in *lin-14* mRNA levels between the L1 and L2 stages of development (Olsen and Ambros, 1999), suggesting that some degree of microRNA-mediated target degradation was observed even in these early studies, but was not discussed.

**Importance of microRNA regulatory mechanism for experimental identification of targets**
Recent work has demonstrated that the computationally predicted potential for microRNA-target interactions is vast, often covering as much as 1/3 to 2/3 of a given organism's genome (Friedman 2009, Lewis 2005). Whether or not these interactions are physiologically relevant, however, requires experimental testing and validation. Knowledge of the mechanism of microRNA-mediated target repression has important ramifications for discovery of direct microRNA targets on a genome-wide scale as a way to validate computational predictions. The ability to detect microRNA-mediated target mRNA degradation by microarray analysis or deep-sequencing is a powerful tool in the arsenal of the microRNA researcher. However, the techniques required to analyze translational repression are generally more technically challenging (2-dimensional gel electrophoresis or isotope labeling of peptides for mass-spectroscopy (Baek 2008, Selbach 2008)). It is therefore of great practical importance to know what sorts of experiments ought to be planned, and that depends on understanding how microRNAs function to regulate their target genes.

The comparison of target mRNAs in wild-type versus microRNA mutant animals will reveal both direct and indirect targets (ie, downstream targets of transcriptional regulators controlled by the microRNA). However, the list of mRNAs upregulated in animals lacking the microRNA can be further culled by analysis for base-pairing capacity to the microRNA. Finally, direct interactions can be tested through analysis of transgenic animals bearing reporter genes.
with the 3’UTR under study; mutation of the cognate microRNA binding site, and observing a loss of regulation, constitutes strong evidence of a direct interaction.

**microRNAs can induce deadenylation of target mRNAs in multiple systems**

The mechanism by which *lin-14* and *lin-28* mRNAs are destabilized are distinct from the mechanism of RNAi-mediated endonucleolytic cleavage. RNA interference utilizes short, sequence-specific guide RNAs (small interfering RNAs, siRNAs) to guide an endonuclease which severs the sugar-phosphate backbone of the target mRNA between nucleotides 10 and 11 of the siRNA (Elbashier 2001), and the cleavage event leaves a 5’ phosphorylated product, which can be detected by ligation and cloning techniques (5’ RACE). We performed 5’ RACE to test whether cleavage products mapping to *lin-4*-complementary elements could be detected for *lin-14* or *lin-28*, but the predicted pattern of cleavage products was not observed (data not shown). Because of this, we conclude that general mRNA degradation pathways are most likely responsible for *lin-4*--mediated target down-regulation, and in Bagga et al. 2005 we put forth a model in which microRNA regulation of a target mRNA induces deadenylation of the target mRNA, followed by decapping and exonucleolytic shortening from the 5’ end and/or the 3’ deadenylated end. Our model is consistent with a report demonstrating that transfection of an exogenous microRNA into tissue culture cells caused global
down-regulation of mRNAs bearing partial complementarity to the transfected microRNA (Lim 2005).

Follow-up work by a variety of groups has confirmed that microRNA-based regulation can induce deadenylation of targets (Wu 2006, Wakiyama 2007, Eulalio 2009a). Extensive work has confirmed that the mechanism of deadenylation involves the interaction of argonaute and GW182 proteins. GW182 proteins (see Eulalio 2009b for review) are glycine/tryptophan-rich and associate with sites of decapping and deadenylation known as P-bodies (Parker 2007). Thus, GW182 proteins serve to recruit microRNA-regulated mRNAs to regions of mRNA decay; however, evidence also suggests that at least part of the function of deadenylation induced by microRNAs is to repress translation (Wakiyama 2007, Eulalio 2008, Takimoto 2009, Beilharz 2009). Thus, deadenylation may serve as a mechanism of translation repression as well as degradation of microRNA-targeted mRNAs.

In sum, translational repression mediated by microRNAs has been reported to occur at nearly every possible step of translation: initiation, elongation, termination, and even post-translationally (ie, proteolytic degradation of nascent peptides) (for review, see Wu 2008). This abundance of mechanisms is conveniently divided into two categories: those relying on GW182 proteins (and, consequently, deadenylation) and those which do not. An intriguing model which bypasses the deadenylation induced by GW182 proteins was recently proposed in which a novel cap-binding motif in
Argonaute proteins specifically recognize the m7G cap and directly competes for cap-binding complex eIF4E (Kiriakidou 2007), thereby directly blocking translation initiation. However, one recent study suggested that the residues originally thought to mediate Argonaute-cap interactions are instead involved in mediating the interactions between Argonaute proteins, GW182 proteins, and microRNAs (Eulalio 2008). Whatever the mechanism for translational control, we predict that deadenylation is a cause of target mRNA destabilization consistent with our observations reported in Bagga 2005.

**Conclusion**

The work performed and published in Bagga 2005 was a collaborative effort that helped shift the paradigm of the microRNA field. The original model, which held that mRNA levels are unaffected by miRNA targeting, led to conclusions that can now be re-examined. For example, in Stark et al. 2005, mutually exclusive expression of microRNAs and target mRNAs was interpreted as an example of evolutionary selection of target gene promoters to mutually exclusive domains of expression relative to their microRNA regulators. However, a more likely interpretation of these data is that the microRNAs actively down-regulate their target genes at the level of mRNA stability, thereby enforcing mutually exclusive domains of expression (Stark 2005). By working with endogenous microRNAs and target genes in C. elegans, we contributed evidence of mRNA degradation to the model of
microRNA-based target regulation, which has very practical implications for target identification and future experiments.

Acknowledgements

The data presented in this chapter was combined with contributions from Shveta Bagga, Shaun Hunter, Katlin Massirer, Janette Holtz, Rachel Eachus, and Amy E. Pasquinelli, and was published in the following report:


John Bracht was the primary researcher and author for the figures and text presented in this chapter of the dissertation.
Materials and Methods

Detection of small RNA species was carried out by polyacrylamide gel electrophoresis (PAGE) Northern methods. Specifically, 15ug total RNA was separated by electrophoresis in 11% denaturing polyacrylamide gels and transferred to nylon membranes (Zeta-Probe GT, Biorad, Hercules, CA). The 5' kinase labeled oligo probe for *lin-4* (p250: ATAGTACACTCACACTTGAGGTCTCAGGG) was hybridized to the blot in 5× SSC, 7% SDS, 0.02 m sodium phosphate, 1× Denhardt's solution for approximately 12 h at 50°C and then washed in 3× SSC, 5% SDS, 0.025 m sodium phosphate, 10× Denhardt's solution at 50°C.

Analyses of larger RNAs were performed by separating 8 μg of total RNA in 1% agarose gels under denaturing conditions. RNA was transferred and UV cross-linked to nylon membranes (Zeta-Probe GT, Biorad), which were then hybridized to PCR-generated DNA probe (*lin-14*: A445+A455, *lin-28*: A424+A427, *eft-2*: A116+A117) radiolabeled using Prime-It II Random Primer Labeling Kit (Stratagene). Hybridization and wash conditions were performed as previously described for PAGE Northernns, except that procedures were carried out at 58°C.

A445: 5' - AACAGTTGAAACAGCTCCACCACCCTC-3'  
A455: 5' - GGATAAGATGGGTGAAGACTGATG-3'  
A424: 5' - GAGCTTTCTGATAGTTTTTCAG-3'  
A427: 5' - GCGTTGCACACAGAATAGCGGAACATCACAG-3'  
A116: 5' - GGTAACCTTCACGGTGATG-3'  
A117: 5' - TCCGAGCTTCTCAACAAGAGC-3'
Quantification of RNA bands was done by PhosphorImager ImageQuant analyses. The *eft-2* mRNA control was used for normalization, and graphs represent the averages and standard deviations for two or more experiments from independent samples.
Figure 10. Developmentally staged RNA samples analyzed for lin-4, lin-14, and lin-28 RNA expression. The wild-type lin-14 mRNA typically does not appear as a discrete band, probably because of the three mRNA isoforms for this gene (Wightman et al., 1991).
Figure 11. *lin-4* miRNA developmentally regulates target mRNA expression (A) RNA from starved L1 (St. L1)- and fed 4hr L1- and L2-stage wt and *lin-4(e912)* worms was analyzed for *lin-14*, *lin-28*, and control *eft-2* mRNAs by Northern analyses. (B) RNA from wt or *lin-14(n355gf)* mutant worms was probed for *lin-14* mRNA by Northern analyses. The lesion in the *lin-14(n355gf)* mutant results in ~2.2 kb bands instead of the ~3.5 kb bands produced by the wt allele (Wightman et al., 1991). (C) PhosphorImager analyses of Northern blots were used to normalize *lin-14* or *lin-28*mRNA levels to the *eft-2* control and the relative change in target mRNA levels between the fed L1 to L2 stages are indicated in the histograms, which represent the averages and standard deviations from two or more independent RNA samples for each.
Conclusions and future directions
Overview

In the work documented in this dissertation, I have contributed in two areas to the understanding of the microRNA \textit{lin-4}: the discovery that \textit{lin-4} biogenesis is regulated at the post-transcriptional level, that an important regulator of its biogenesis is R05H10.2, and that the mechanism of target regulation by \textit{lin-4} involves degradation of target mRNAs. Both are significant advances of the field, and open up multiple new directions for research. The significance of each finding for our understanding of the microRNA field, and a few ideas for follow-up work, are presented below.

\textit{lin-4} biogenesis (chapter 1)

Integration of heterochronic and organismal growth pathways

The discovery of a RRM-motif protein that mediates post-transcriptional regulation of the \textit{lin-4} microRNA fundamentally re-orient the conceptualization of the role of the heterochronic pathway in \textit{C. elegans} development. \textit{lin-14} was initially characterized as a heterochronic gene that controls the developmental timing of somatic tissues such as hypodermis, seam, and vulva, relative to germ-line development, which remained relatively unaltered. However, the work presented in chapter 2 gives a larger context in which to understand this model, by integrating somatic and germline developmental roles for the heterochronic pathway. While \textit{lin-4} appears to be primarily a somatic regulator, its target gene, \textit{lin-14} has an important role as a
germline regulator where an unknown pathway inactivates it to promote whole-organism growth. The nature of the germline-specific lin-14 regulatory pathway remains unknown, but it is under the control of R05H10.2, and may be another microRNA. In this integrated model, two pathways—one controlling somatic development and one germline development—down-regulate lin-14 function, and this regulation is essential for whole-organism growth. The role of lin-14 in germline development was previously masked by the unknown germline-controlling lin-14 regulator, but inactivation of R05H10.2 disables both repressive pathways and unmarks this important function of lin-14. The nature of this germline-specific growth controlling pathway is currently unknown but an important focus of future research. The data currently in hand do not distinguish between signaling pathways which are actually embedded in germline tissue versus somatic pathways which regulate germline growth from other tissues such as the somatic gonad. Other somatic tissues could be involved in regulating germline growth; perhaps neuronal signalling pathways play an important role, since LIN-14 protein was found in the neurons (Ruvkun 1989). Indeed, in that work it was reported that neuronal LIN-14 protein was downregulated with slightly different dynamics than in other tissues, suggesting a separate neuronal regulatory system. Since lin-4 microRNA appears to be expressed in neurons (based on several promoter-GFP fusions from different groups) (Esquela-Kerscher 2005, Ow 2008, Martinez 2008), any hypothetical neuronal, lin-4-independent regulation of lin-14 must somehow
prevent co-expressed *lin-4* from down-regulating *lin-14* in this tissue. Clearly much future work must be done to elucidate the mechanisms of germline regulation by *lin-14*.

Strategies for uncovering the germline-*lin-14* regulatory pathway

On the hypothesis that the germline *lin-14* regulatory pathway is another microRNA, the most straightforward test is to carefully analyze microRNA mutants for germline growth defects. However, since most microRNA mutants have been reported to have no phenotypes (Miska 2007), it is likely that two or more microRNAs mediate this pathway. A simple search for germline-expressed microRNAs with potential to base-pair to *lin-14* failed to identify any candidates for follow-up (data not shown). However, the germline-regulating microRNA need not be expressed in the germline, so a more extensive analysis of microRNAs with capacity to basepair to the *lin-14* 3'UTR may identify a family of microRNAs which could mediate this function. Then, combining mutants of multiple microRNAs from that family may uncover germline phenotypes. Since most microRNA mutants are available from the Caenorhabditis Genetics Center, CGC, testing them for germline effects should be straightforward.

Another strategy is to comprehensively analyze which microRNAs are strongly affected by R05H10.2 RNAi, since potentially one or more of them mediate germline growth. Then, *lin-14* base-pairing capacity would narrow the
list to the most promising candidates to test from the CGC. If a germline-growth-regulating microRNA candidate is upregulated by R05H10.2 (as opposed to the effect on \textit{lin-4} which is downregulated) then it probably represses a \textit{lin-14} repressor. Therefore, inactivation of that microRNA should repress \textit{lin-14} in the germline-growth pathway, allowing the germline to proliferate on R05H10.2 RNAi. The same principle holds true of potential protein-coding regulators of \textit{lin-14} in germline growth; we predict that a \textit{lin-4}-like phenotype (germline growth but somatic stalling) would be evident only on R05H10.2 RNAi. Preliminary evidence suggests that R05H10.2 RNAi bacteria are potent enough to effect \textit{lin-4} even when diluted. The initial RNAi experiments had a mixture of RNAi bacteria due to poor copying of the library, and yet disruption of \textit{lin-4} biogenesis was observed. Therefore, a potentially valuable genetic screen could employ double RNAi of R05H10.2 and other candidate genes to identify suppressors of the growth arrest phenotype.

Before performing this screen, suitable conditions for diluting the R05H10.2 RNAi bacteria need to be identified.

The reduced growth observed on R05H10.2 is accompanied by some morphological changes. The worms look smaller even at adult-stage than vector-fed worms, and they seem thinner in larval stages, which is reminiscent of dauer phenotypes. The dauer stage is an alternative third larval stage which is particularly long-lived and resistant to stresses (Fielenbach 2008).
While animals grown under R05H10.2 RNAi conditions do not fully enter dauer arrest, there may be some cross-talk between dauer pathways and R05H10.2. Furthermore, some dauer-regulating genes are implicated in microRNA biogenesis (Bethke 2009, Hu 2007). For these reasons, investigation of the role of R05H10.2 in mediating microRNA biogenesis and dauer pathways may be fruitful. One approach would be to test dauer-pathway mutants for their ability to suppress the growth and developmental phenotypes of R05H10.2 RNAi.

**Mechanism of regulation by R05H10.2**

Another important direction for future work is the mechanism of R05H10.2 regulation of lin-4 processing: how does R05H10.2 control lin-4 biogenesis? One simple hypothesis is that the R05H10.2 protein interacts directly with lin-4 primary transcript, perhaps recruiting the microprocessor machinery or in some way licensing the transcript for processing. The most convincing way to demonstrate such an interaction would be to isolate the R05H10.2 protein from worm lysates along with associated primary transcript(s). Generating a peptide antibody would be relatively fast and likely effective for immunoprecipitations, although creating an antibody to recombinant protein might also work. The potential drawbacks of antibodies targeting recombinant protein are 1. more work up-front, and 2. greater chance of non-specific interaction with other RRM proteins, of which there are over
100 in *C. elegans*. Once a suitable antibody is in hand, a good way to proceed would be to use UV light to crosslink RNA species (this captures even transient interactions, but only direct ones since it doesn’t crosslink proteins to each other), then lyse the worms and perform immunoprecipitation. Washes, followed by proteinase-K treatment to liberate the associated RNA species, would complete the protocol. I would first perform random-primed reverse transcription on the isolated RNAs, followed by PCR for the *lin-4* primary transcript. The *lin-4(e912)* mutant can be used as a negative control, and an anti-GFP antibody can serve as a binding control. Because human RBM28 is known to interact with all spliceosomal RNAs, they may serve as positive controls. Since R05H10.2 RNAi depletion affects multiple microRNA primary transcripts and mature isoforms (see chapter 2) it is likely that the pool of interacting RNAs will include other microRNA targets. I would perform deep-sequencing to identify these targets, which may include other RNAs besides microRNAs (PAGE gels showed many small RNA species, currently uncharacterized, that were up-regulated upon R05H10.2 inactivation by RNAi). I would utilize the RNA-Seq techniques described in Wang 2009. Briefly, fragmentation of isolated RNAs can be performed by hydrolysis or sonication of cDNAs followed by ligation of adaptors and deep sequencing (Wang 2009).

Alternatively, R05H10.2 interacting RNAs may be identified by a full RNA-CLIP (RNA cross-linking and immunoprecipitation), in which the protein of interest (R05H10.2) is cross-linked to its associated RNAs by UV light,
immunoprecipitated from biochemical lysates, and then treated with RNAs to
digest single-stranded RNA, leaving protected regions that are crosslinked to
the protein. Purification of the protein by SDS-PAGE and ligation of adaptors
allow deep sequencing of interacting regions of targeted RNAs (Ule 2005).
This method has the advantage of identifying binding sites of target RNAs but
the disadvantage of not distinguishing RNA isoforms such as primary and
precursor microRNAs (mature microRNAs are small enough to potentially be
completely protected by protein from digestion by RNAs). Since all lin-4
isoforms are potential R05H10.2 interactors, identification of R05H10.2-CLIP
sequences corresponding to microRNAs would need to be followed with an
isoform analysis of some sort—potentially through RT-PCR as described
above.

The biochemical purification of R05H10.2 from biochemical extracts
may also immunopurify microRNA processing enzymes, and in vitro tests for
this activity can be performed with purified R05H10.2. However, since neither
Drosha nor Dicer has been immunopurified with R05H10.2 or RBM28,
(Gregory 2004, Chendrimada 2005), and none of the yeast-two-hybrid
interactors with R05H10.2 in C. elegans is associated with Drosha or Dicer,
we consider it unlikely that a direct interaction would be detected.

R05H10.2 EMSA studies
Another direction for analysis of R05H10.2 is to purify recombinant R05H10.2 from *E. coli*. This would open the possibility of performing EMSA assays to test for direct binding of the protein to various microRNA primary transcripts which can be *in-vitro* transcribed. Given that RNA-binding motifs such as RRM might nonspecifically interact with many different RNAs, some other method for confirming R05H10.2-RNA interactions *in vivo* is needed—such as the CLIP procedure described above. Once a genuine target RNA is identified, deletion studies can be performed to identify important motifs recognized by R05H10.2.

**R05H10.2 and DNA**

Since RRM motifs can mediate protein-DNA interactions (Liu 2009), it will be important to test for DNA as well as RNA interactions for R05H10.2. One potential function of this protein is to couple transcription and biogenesis of microRNA primary transcripts, since some RRMs from the same protein might interact with RNA while others might interact with DNA. Once a good antibody is generated, standard ChIP techniques can be utilized to identify potentially interacting regions of the genome. Specifically, regions corresponding to *lin-4*, *mir-2*, and *mir-47* are logical places to start but eventually a ChIP-on-ChIP strategy to identify all sites in the *C. elegans* genome may be illuminating and might highlight specific motifs that R05H10.2 interacts with.
R05H10.2 posttranslational modifications

An important question to address is whether R05H10.2 is post-translationally modified, perhaps by phosphorylation. Since the protein appears to be ubiquitously expressed (Hunt-Newbury 2007 and my RT-PCR analysis of mRNA), the question arises: why is it important for \textit{lin-4} biogenesis at later times, but at earlier timepoints \textit{lin-4} biogenesis is blocked (prior to 12hr, see Fig. 5). Evidence suggests that both \textit{lin-4} primary transcripts and R05H10.2 are present at early timepoints, but no mature microRNA is made. If R05H10.2 is part of the post-transcriptional timer for \textit{lin-4} biogenesis, there may be some signaling pathway or additional cofactors which mediate the activity of R05H10.2 toward \textit{lin-4}. Recent reports demonstrate that hnRNP proteins (which are in the RRM family of proteins along with R05H10.2) can be phosphorylated (Ting 2009) and that phosphorylation can occur in response to signaling pathways (Berglund 2009), making R05H10.2 an intriguing candidate for phosphorylation.

Once immunoprecipitation protocols are worked out, characterizing the phosphorylation state of the protein can be done by running the purified protein on a gel and performing a Western blot with anti-phospho-serine, threonine, or tyrosine antibodies. We predict a difference in phosphorylation state for R05H10.2 between 4h and 16h. In addition, other modifications might be modulating R05H10.2 activity and various other strategies might be
employed to characterize them. Mass-spectroscopic analysis might be necessary to identify potential ubiquitination, sumoylation, and other modifications.

**R05H10.2 and Proteins**

It has been shown that some RRM motifs interact with proteins, not RNA (Maris et al. 2005). Therefore it will be important to characterize interacting protein co-factors in addition to interacting nucleic acid species. Assuming a good antibody to R05H10.2 can be generated, this analysis is theoretically straightforward through co-immunoprecipitation and mass-spectroscopic analysis. However, a list of 64 interactors defined by yeast two hybrid techniques is available on wormbase (www.wormbase.org) and may prove a fruitful place to begin investigations. In scanning through the wormbase information for these interactors, I noticed that many of them have similar RNAi phenotypes to R05H10.2, and some are known nucleolar proteins, consistent with the idea that they may form a functional complex whose function is disrupted by inactivation of any member. It should be fruitful to test, by Northern blot, which of these R05H10.2 interacting proteins are important for biogenesis of *lin-4* and other microRNAs such as *mir-2* and *mir-47*.

**R05H10.2 and other microRNAs**
It is very important to know the size of the set of microRNAs affected by R05H10.2 inactivation, as well as which steps in their biogenesis are affected. Various sorts of microRNA microarrays exist which can assay most or all known microRNAs in *C. elegans*. Are there any microRNAs unaffected by R05H10.2 RNAi? Are there common elements to R05H10.2 interacting microRNAs (or their primary transcripts) which might explain why they are targets of R05H10.2? Recent reports have found several motifs upstream of many *C. elegans* microRNAs (Ohler 2004 Heikkinen 2008), and it will be important to identify correlations between these motifs and R05H10.2 interacting microRNAs.

R05H10.2 and Splicing

Since the human homolog of R05H10.2 interacts with spliceosomal RNAs, there may be splicing defects evident upon depletion of R05H10.2 by RNAi. Alternatively, since many microRNAs are encoded within introns of host genes (Rodriguez 2004), R05H10.2 may associate with components of the spliceosome in order to mediate the processing of all or a subset of these microRNAs. Interestingly, *lin-4* resides within intron 9 of the *C. elegans* gene F59G1.4 (see chapter 1, Fig. 4), but my work has shown that *lin-4* host gene transcripts are preferentially avoided by Drosha, and instead that independently arising transcripts are the productive substrates for Drosha processing and *lin-4* biogenesis (chapter 1, Fig. 5). However, *mir-2* is also
intron-encoded in a host-gene, *ppfr-4*, a protein phosphatase regulatory subunit important in embryonic mitosis. While no information on coordinated expression between *mir-2* and *ppfr-4* is available, sequences immediately upstream of the microRNA were observed to drive GFP expression specifically in somatic, head, and tail neurons (Martinez 2008). In many ways *mir-2* is similar to *lin-4*—both are expressed in neurons (although *lin-4* is not exclusively neuronal) and both are annotated as within introns of host genes. However, they are regulated in opposite directions by R05H10.2, so as-yet undefined differences between the microRNAs may account for the observed different roles played by R05H10.2 in their biogenesis. Regardless of the determinants of R05H10.2 activity toward intronic *mir-2* and *lin-4*, the *mir-47* microRNA is independently transcribed but also regulated by R05H10.2, which suggests that the role of this factor extends beyond intronic microRNAs at least in some cases.

The connection between R05H10.2 and splicing machinery is provocative, especially given recent evidence that splicing and microRNA biogenesis are in some ways coupled. The established model has looked to the snoRNA field for insights into small RNA biogenesis from introns. snoRNAs are short guides of ribosomal and spliceosomal RNA nucleotide modifications, either methylation or pseudouridylation (Meier 2005). In vertebrates, all snoRNAs arise from intronic regions of protein-coding genes, and their biogenesis generally involves post-splicing processing by
exonucleases acting on the debranched lariat structures (Filipowicz 2002).

Since microRNAs are also small noncoding RNAs, it has been assumed that their biogenesis might follow a very similar mechanism. However, recent work has begun to raise questions about these assumptions. The work of Kim 2007 showed that Drosha processing substrates were partially-spliced introns retaining exonic sequences, suggesting that microRNA processing and splicing proceed in parallel. In addition, they found that splicing of mRNAs bearing microRNAs in their introns was unaffected by Drosha knockdown, suggesting that Drosha processing does not antagonize the splicing process.

Kataoka 2009 performed native gel analysis of in vitro splicing reactions with HeLa cell nuclear extract, adding hnRNAs bearing microRNAs in their introns. They demonstrated that microRNA cropping occurs before fully spliced products accumulate. Constructs in which splicing is blocked by mutation of splice sites had no effect on Drosha processing, again supporting the independence of the two processes. Kataoka et al. also demonstrated that Drosha interacts with splicing intermediates, and may serve as a tether to keep the 5’ and 3’ exons together for trans-splicing after removal of the microRNA hairpin loop.

Data from the Proudfoot lab (Morlando 2008) demonstrates that both Drosha and exonucleases localize to introns containing microRNAs. They propose that cotranscriptional Drosha processing recruits exonucleases like the Exosome and XRN2 to rapidly process the exposed intronic ends when
the hairpin is removed. However, they propose also that exonic sequences are protected and splicing proceeds normally, potentially relying on a tethering of upstream and downstream exons (Morlando 2008). A study from the Steitz lab (Pawlicki 2008) reported that alterations which increase residence time of primary transcripts at sites of transcription increase Drosha processing efficiency, and that polyadenylation and release of primary transcripts inhibits Drosha processing. They observe that flanking exons increase Drosha processing, which they attribute to an increased association with the site of transcription and thus greater Drosha processing efficiency (Pawlicki 2008).

An intriguing link between microRNA biogenesis and splicing machinery was published recently by the Rosenfeld lab (Trabucchi 2009); in this work the alternative-splicing protein KSRP (Min 1997) was found to be a member of microprocessor and to facilitate both Drosha and Dicer processing of a subset of microRNAs. They report that KSRP interacts with terminal loop regions of the microRNAs it regulates, and promotes processing at both Drosha and Dicer steps. Intriguingly, they find that KSRP interacts specifically with let-7, and that it acts epigenetically to lin-28 in regulating let-7 biogenesis. Inactivation of lin-28 de-represses let-7 processing, but inactivation of both LIN-28 and KSRP prevents let-7 biogenesis. These findings provide an intimate connection between microRNA biogenesis pathways and splicing pathways and may signal the importance of splicing and microRNA biogenesis. While KSRP is an RNA-binding protein, it has 4 KH domains and
is therefore structurally distinct from R05H10.2, but the general principle of an RNA-binding protein which regulates multiple steps in microRNA biogenesis is instructive and appears to be an emerging motif of microRNA biogenesis regulation.

**lin-4 Target Regulation (chapter 2)**

**Identification of novel lin-4 targets**

The phenotypes that characterize *lin-4* deletion mutants are complex and not fully understood. Given this complexity, we expect that there are more components of the *lin-4* regulome than just *lin-14* and *lin-28*. Work is underway to identify novel targets through crosslinking and immunoprecipitation of RNA associated with the argonaute protein *alg-1*, a technique known as ALG-1-CLIP-seq. Briefly, this technique involves cross-linking of *alg-1* to its associated RNAs by UV light, followed by specific immunoprecipitation of the protein from biochemical lysates. The cross-linked *alg-1*-RNA complexes are then treated with RNase to digest single-stranded RNA, leaving protected regions that are crosslinked to the protein. Purification of the protein by SDS-PAGE and ligation of adaptors allow deep sequencing of interacting regions of targeted RNAs (Ule 2005). While this technique has been utilized by workers in the Pasquinelli lab to identify all microRNA targets in wild-type worms, identification of targets lost in a *lin-4* null mutant will help reveal *bona fide* targets for the microRNA. This work is currently underway and is expected to
reveal other direct targets that may play significant roles in worm and human biology, as is the case for the known lin-4 target lin-28 (see discussion below of the significance of identifying new targets of lin-4). Validation of novel target candidates will require generation of transgenic reporters with candidate 3'UTRs in wild-type versus mutant backgrounds. To verify the direct interaction of a microRNA with a particular target site, removal of the potential lin-4 binding elements and concomitant loss of microRNA repression in wild-type animals needs to be demonstrated. The ability to generate single-copy insertions via mos-transposon mediated homologous recombination (Frøkjaer-Jensen 2008), known as mos-single copy insertion (mos-SCI) is most promising for eliminating the copy-number variables that have plagued transgenic experiments in C. elegans, making these experiments much more sensitive and powerful. Currently the mos-SCI system is being optimized in the Pasquinelli lab and results to date are highly encouraging.

**Significance of lin-4 target identification**

The lin-4 target gene lin-28 is highly conserved across bilaterian phylogeny (Moss 2003). Furthermore, its down-regulation during development and regulation by microRNAs is also conserved (Moss 2003, Wu 2005). Interestingly, neuronal differentiation was found in Wu and Belasco 2005 to require the down-regulation of lin-28 by the human lin-4 microRNA homologs, mir-125a and miR-125b. Consistent with this data, C. elegans lin-4 expression
has been reported in neurons based on promoter-GFP constructs (Esquela-Kerscher 2005, Ow 2008, Martinez 2008, and my work reported in chapter 2), emphasizing the phylogenetic conservation of relative tissue and time-specific expression of lin-4 family members and conservation of at least one known target. Perhaps most intriguingly, lin-28 was recently characterized as one of four genes that when overexpressed, can reprogram human somatic cells into stem cells (Yu 2007). Multiple papers recently identified lin-28 as a post-transcriptional regulator of let-7 biogenesis, although the proposed mechanisms differ (Heo 2008, Rybak 2008, Newman 2008, Piskounova 2008, Viswanathan 2008, see introduction to this dissertation for a discussion). These findings highlight the importance of the lin-4 genetic network: if one target is important for maintenance of stem cells, differentiation of neuronal cells, and regulation of microRNA biogenesis, what might other components of the network be doing? Undoubtedly exciting new discoveries await the full characterization of the lin-4 network in C. elegans.

Conclusion

Taken together, the studies presented in this dissertation represent advances in the understanding the founding microRNA, lin-4. The role of post-transcriptional regulation and biogenesis of lin-4 by R05H10.2 is novel and reveals a previously unknown role for the heterochronic gene lin-14 in mediating whole-organism growth. In addition, my studies on the mechanism
of *lin-4* target regulation played a role in re-shaping the paradigm generally accepted by the microRNA field, and opened new vistas for the discovery of new *lin-4* targets. While these advances are important to increase our understanding of microRNA biogenesis and function, they also beckon with the potential for future discoveries that shape our thinking about stem cells, cancer, and other human diseases.
Appendix 1: Isolation and Characterization of a Novel Alg-1 Mutant
During my rotation in the Pasquinelli lab I demonstrated that the hypomorph hypomorphic let-7(mg279) allele, with no morphological phenotype under ordinary conditions, exhibits enhanced bursting on dcr-1 RNAi food compared to wild-type (data not shown). Somewhat surprisingly, wild-type worms exhibit little or no bursting on dcr-1 RNAi food, likely due to the requirement of dcr-1 for functional RNAi. These preliminary data suggested that dcr-1 (RNAi) is a sensitized background in which to detect mild effects on microRNA biogenesis. This might be important, for instance, if a factor in the biogenesis pathway was essential to viability so strong loss of function would not be tolerated. Because a bursting phenotype can occur due to many developmental defects, the ability to enhance the bursting phenotype under dcr-1 (RNAi) conditions was also a way to ensure specificity for factors in the microRNA pathway.

I decided to screen for factors involved in let-7 biogenesis and function by performing EMS-based mutagenesis on N2 worms and then plating them on dcr-1 RNAi food, scoring the F2 generation (120,000 genomes) for homozygous mutants exhibiting enhanced bursting. Secondary analysis was performed to ensure that on vector control food the worms displayed a relatively wild-type phenotype. In the summer of 2003, I identified two such mutants (both were backcrossed 3X): Dicer RNAi Enhanced -1 and -2 (dre-1 and dre-2). dre-1 and dre-2 exhibit 80-90% bursting on dicer RNAi food but are apparently wild-type (no bursting observed) under control (vector) RNAi
conditions (for dre-2, see Fig. 12). Mapping proved quite challenging due to the fact that the standard Hawaiian strain (CB4856) is partially RNAi defective (mainly in the germline but I was performing 2-generation RNAi so this proved crippling) and tracking the dre-2 mutation required functional RNAi. Even with the reduced RNAi functionality in the CB4856 strain, I succeeded in mapping dre-2 but never could recover bursting worms from crosses between dre-1 and CB4856 worms. We have since received a mapping strain in which the RNAi defect is corrected, so this may now be a viable project.

Mapping of dre-2 indicated a 170kb region of Chromosome X (Fig. 13), which included about 30 genes. However, one gene in the interval was a known microRNA pathway gene: Argonaute-like-1 (alg-1) (Fig. 13). Upon sequencing the coding region of alg-1 in dre-2 mutant worms I found a C-T transition in the PIWI motif, which changes conserved alanine 756 into a valine, and we named this allele ap45 (Fig. 14). While this is a very conservative mutation, sequence alignments of various PIWI motifs demonstrate that this residue (and the surrounding PSIAAV motif) is highly conserved in eukaryotic argonaute proteins, and it falls very close to residues known to be important for argonaute protein function (Fig. 15).

Molecular analysis of the ap45 allele of alg-1 demonstrated that the mutant protein is stable (Fig. 16) and that normal levels of mature let-7, lin-4, mir-48, and mir-241 are generated (Fig. 16). However, regulation of let-7 targets lin-41 and daf-12 were compromised in animals grown under dicer
RNAi conditions (Fig 16), suggesting that ala756 is important for target regulation but not biogenesis of microRNAs. In crystal structures of archaeal argonaute proteins, the analogous residue maps to a beta-strand very close to the active site but on the opposite edge of a beta-sheet, near a potential protein-protein interaction surface. An intriguing possibility is that the dre-2 mutation alters the DICER binding site on alg-1. Previous work has demonstrated that PIWI domains of argonaute proteins interact with DICER, but the site and mechanism of interaction has never been elucidated. I generated a DICER antibody which worked for detecting endogenous protein (expressed at normal levels in ap45, data not shown), and performed several immunoprecipitations of ALG-1 protein from wild-type and ap45 mutant animals but never was able to see dicer interacting with ALG-1 in either background (data not shown). It is possible that the ALG-1-DICER-1 interaction is too transient to detect in wild-type animals and that cross-linking techniques might help. It is also possible that the ALG-1 antibody interferes with the ALG-1-DICER-1 interaction.

What ultimately convinced me to stop work on this project was that the molecular rescue of ap45 proved to be problematic, with transgenic animals displaying variable growth arrest (even at 1 or 0.1 ng/ul of injected DNA), lack of rescue of alg-1(ap45) mutant animals (see Fig. 17) or a stronger allele, alg-1(gk214) (not shown). It remains a possibility that the mutation responsible for the phenotype in dre-2 mutants is not ap45, since they were not rescued by
wild-type alg-1 expression constructs. But the failure of the same constructs to rescue a strong \textit{alg-1} mutant suggests instead that \textit{alg-1} is a sensitive locus, and that even mild over-expression of the protein is not well tolerated. Thus, \textit{ap45} remains to be functionally rescued, and its role in mediating target regulation remains to be elucidated.

The \textit{Alg-1} rescue fragment was generated by Platinum Taq amplification of wild-type genomic DNA with oligos A347+A348, followed by TOPO cloning and sequencing. The rescuing fragment is 7kb and had to be sub-cloned to remove PCR-induced mutations (most sequenced clones had 4+ mutations). To insert the \textit{ap45} mutation, I PCR amplified the relevant region from \textit{ap45} genomic DNA, followed by standard cloning using fortuitously available restriction sites.

\begin{align*}
pJRB13 &= \text{\textit{alg-1} rescue construct, wild-type} \\
pJRB14 &= \text{\textit{alg-1} rescue construct, \textit{ap45} mutation} \\
\end{align*}

linearize for injection with Sph I.

to genotype endogenous \textit{ap45}:
PCR amplify off single worm lysate or genomic DNA prep using A817+A818
(150nt product, 58° anneal, 35 cycles, 10sec extension)
digest with Bbv-I (1ul per tube, the enzyme is relatively inefficient, may take
significant time to digest to completion)
wt cuts to 100 & 50nt fragments
ap45 fails to cut
A817 CGTACGCCCAAGAATCTTCAACG
A818 GTTGTTGGACACGGACGTTGCTGCATATCTG has mismatch in 2
Bbv-I site to increase the size difference between cut and uncut
A347 atgttgactagtTTCTACCGCTACTCTTCCCATC (spe-I site)
A348 atgttgaccggtCGTTTCGGACTTCAGTTGTTTC (age-I site)
Figure 12. Quantitation of the bursting phenotype observed in *alg-1(ap45)* mutant animals after 2-generation RNAi against vector or *dcr-1*. 
Figure 13. Genetic mapping of ap45 mutant (originally called *dre-2*). Vertical axis shows map units in centimorgans (c.m.) Genetic mapping identified the 170kb region bracketed; *alg-1* resides toward the middle of this region as indicated.

Figure 14. Altered DNA and protein sequence in ap45 allele and verification of the snp by genotyping PCR. Notice that Bbv-1 frequently fails to cut to completion but the ap45 mutation abolishes digestion completely.
**Figure 15.** The ap45 amino acid change (A756V) occurs in a highly conserved residue in the PIWI domain of ALG-1.
Figure 16. Molecular analysis of the ap45 allele. A. Western blots demonstrating that the ap45 amino acid change does not destabilize the ALG-1 protein under either vector or dcr-1 RNAi conditions. B. Analysis of microRNA biogenesis in wild-type versus alg-1(ap45) mutant animals showing that under vector and dcr-1 RNAi conditions, let-7, mir-48, mir-241, and lin-4 are expressed at normal levels. C. Analysis of target regulation in alg-1(ap45) mutant animals. Wild-type versus alg-1(ap45) animals were staged on vector and dcr-1 RNAi food and regulation of let-7 targets daf-12 and lin-41 were analyzed. Notice that both lin-41 and daf-12 are de-repressed in alg-1(ap45) animals on dcr-1 RNAi food.
Figure 17. Attempts to rescue \textit{alg-1(ap45)} with a wild-type \textit{alg-1} rescue construct and an \textit{alg-1(ap45)} rescue construct. All animals were grown for two generations on \textit{dcr-1} (RNAi) food, and these transgenic animals displayed non-rescued bursting phenotypes.
Appendix 2: Characterization and Genotyping of the \textit{lin-4(e192)} Rearrangement
The classical null allele *lin-4(e912)* was identified as a lineage mutant generated by p32 mutagenesis by a worker in the Brenner lab (Horvitz and Sulston 1980) and was defined in Lee 1993 as a large deletion. However, the exact breakpoints were left unmapped, and the full extent of the deletion is still unknown. In the course of performing the work described here, I mapped one end of the deletion, and discovered that downstream regions were actually duplicated and inserted, in reverse orientation, into the place where *lin-4* sequence originally was.

![Diagram showing duplication and deletion in the *lin-4(e912)* mutant.](image)

**Figure 18.** Duplication and deletion in the *lin-4(e912)* mutant. The inserted sequence appears to duplicate at least through the *vps-35* gene approximately 7kb from *lin-4* but likely goes even further. See text for breakpoint sequence.

While I was able to sequence approximately 1kb into this duplicated region, I was not able to go further because this sequence exists in two copies in e912 mutants and I would get muddled sequencing data with oligos sitting
in this region. However, an amplified genomic fragment that I TOPO cloned had sequences from \textit{vps-35} at the far end (see discussion below) suggesting that the duplicated region extends over 5kb. The deletion, and duplication/rearrangement data allowed me to design a very reliable genotyping protocol for detecting N2-only and e912-only alleles.

In order to PCR amplify e912 only, use A1666 and A1284. In wild-type animals these oligos both point in the same direction (to the left in Fig. 18) but e912 duplicates and inverts the sequence with A1284 in it, so now it can amplify with A1666. Furthermore, to detect N2 only, you need only use oligos that sit within the region deleted by e912; A463+A464 fit this criteria and quite robustly amplify only wild-type. I have used these oligo sets for many experiments in which various constructs are crossed into an e912 mutant background. (Note: I often get a smaller band, \textasciitilde 400nt, with A1284 and A1666 which is nonspecific and not diagnostic of the e912 rearrangement).

The breakpoint in e912 sits 84nt downstream of the 3' end of mature \textit{lin-4}. The sequence immediately following the breakpoint is as follows:

\[
\text{[deleted region, including } \textit{lin-4}]TTTCACTGGGAGCTAG\ldots\ldots
\]

(Into [deleted region, including \textit{lin-4}] is the downstream sequence which is duplicated and inverted.)

The full sequence of the duplicated, inverted region which has been inserted, as much as I know, is as follows (this sequence was obtained from a
TOPO-cloned e912 genomic fragment which was amplified from e912 genomic DNA in a single-primer A306 PCR reaction):

[breakpoint]AAATTGCAAACAAATTTCAGCAGTCCTGAAGACTACACCGCA
GTGTTTGAAGTCTCCAAAAAGTTGCTCGAAACTCCAGATACAAATCATCGAG
ATCTCTCATTAAATCGAAATCTTCAATTCTTACTTTCTTTGTTCATTTCTCGTTT
GAATAAAACTGAATAAAAGTTTTATACAAAATGGTCGTTTGATGCGAGTCT
TCTCTATCTATTTGTGTGATGTCCTGGAGTTTTACTGTTATTTGAAAAACGATAA
CTAAATCGAGTTTTATAAGTAGCTTGTCTCCTGCTGTTGAAAAATTCTAGAG
ATTCTTTTAGTTGGATCTTTTCAATGCTGCCAGATCAAGATGACTGAG
TACCTTTATTCAAAATTTTTATGATCTTCTGTCTTGTGCTCAGTTTTC
TATTAATTTTGTATTGTGTGTGCGACTTCTCGAGAGGAGGGCCATTCC
CATGCTAAATTGGAAGACCTACTCGGAATGCAAGATGACTCAG
TCTTTTCAAATTCACTGTGGTTACATATAATTTTCTAGATCAATTCTTGTATAG
AGTTCATTTCTTATTTGTGATGAGTTATACACTTTTCTTTTCTCTGCATTTCC
TCAGAATATGATACAAATGACAAACTTTTAATGACTGTCGAGTTGCGACAC
TTATAGATTGATCATGAAATCTCGACATATTGAATTATTGGAACACTAAC
GAATCAATGGAAATCTGATTTCTAGTGATGCAATTCTGCTGATCT
ACGATGCAAATCTTCTTTATGCTTTACCAAAAAGTTTTTAGATTTGAAAT
ACTCGTGTTTGTGTGAACTCCAGTATTTTTGAGTGTCTGTCAAG
TTGACAAATTTCTATTCTACCTCAAACGGAAACAAATGACGCTTGAATGAGT
GGAGTACTGTTCAGTACCTTAAAGCTGCTGTAACACAGTTTCTGTTC
AA.....

(signal faded out)

The other end of the genomic fragment (which has a site where A306 mis-primed) is in vps-35, a gene approximately 7kb downstream from the lin-4 breakpoint (making for a total of 5.4kb of duplicated sequence, which starts ~2kb downstream of the breakpoint). I do not know whether the intervening regions are intact in the e912 insert, but the size is approximately right if they were simply inserted without further rearrangement. Because this sequence starts with an A306 mispriming site, it does not represent the full size of the inserted sequence. I include it here simply for completeness:
[mispriming site for A306, pointing back toward the breakpoint site] -
GTTGGTCGAGCTAAACAAAACTTCTGAATATTCGAATAGATGAATACAA
AACGTGCTCAGACTTTCCCCAACTTGGCGAATACCATCAAAGTTGATCCT
TTGATTATCAGAGAGACTGAAATTGCTTCTATATGATTCAAAATATG
CTCGAGAAGAAGAAGCTTTCTCGTAATCAAGATGACGTGACCTAGCGTT
CTCACTTTATTCTCTTTACTCAAGAGTACGGAGAAACATCAGAAGTATTG
ACTGAGACGAGAGAATTGCCGATGCAAAATTTGATAGCTCTAGACGTTC
TACATTATACAGAGCGGATGATGGATGATTCCGAAATTTTTGGTAAAAATC
GTTTGACTACTATAGCAAAATATTATTTTTCTTCTACCTTCTCAACAGTG
CCGTTAAACTCTGGTTGAAGGTCGCGCCATCGTCTCCGATACACGCTT
CCTCCTATAATTTTTGAACTCTCTATCTGCTTGGTCTGATTTTCAGACATG
AAAAGATGAGAGATGACAGTGGGATGCAAGATTAAAGAGTTCTGGTG
TGCAATGGGAAACCATCGGAGCAGCTTTGATCGACAGCTGAAATTCACAG
CTTCCAATGAAGTGTATCTTTAATGTAAGTATCCAGAAAACCTGATAATTTT
TAACAAATGTGTTTTAGGGTGCAATCAGCACCGATCGTGTTCCTCTTCAAG
GAT AATCACACAG TAGTCTATGA ATTTGTATCG(signal faded out)

For genotyping e912: (note, need to run 2 reactions for each worm tested,
since you can only test for the presence or absence of one allele at a time; if
you detect the presence of e912, the worms may still be heterozygous and a
separate test for N2 is necessary to confirm homozygosity).
A1284+A1666, 1kb
58º, 1min extend, 35cycles

for genotyping N2:
A463+A464, 900nt
58º, 1min extend, 35cycles

A1284 gatcgctcagATTTTCGACAGAGATTCGATCAAAC (apa-I site)
A1666 GCACCGCTCAGCTCGTTGCGATG
A463  GAGCGAAGCTTAAGGGGATGATGTGTC
A464  GTATACCGGTCATGAAAAGTGACACTTGC
Appendix 3: Plasmids and Strains Generated
<table>
<thead>
<tr>
<th>Construct</th>
<th>Description</th>
<th>Construction &amp; Strains</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>pJRBl</td>
<td>1kb <em>lin-4</em> promoter fused to gfp. Start site ~200nt from <em>lin-4</em>: tcctccagtcct</td>
<td>A321+A323 digested Age-I, cloned into Age-I site of pPD95.75. apEx 20,25,26,27,28,39,40 PQ9,10,11,12,13,A321+A323 digested Age-I, cloned into Age-I site of pPD95.75. apEx 20,25,26,27,28,39,40 PQ9,10,11,12,13</td>
<td>GFP in neurons (constitutive), pharynx, seam cells, vulval cells, distal tip cells, hypodermis</td>
</tr>
<tr>
<td>pJRBl</td>
<td>same as pJRBl but promoter is inserted in reverse as a control</td>
<td>no lines</td>
<td></td>
</tr>
<tr>
<td>pJRBl</td>
<td><em>lin-4</em> rescue fragment promoter-GFP. Start site ~200nt from <em>lin-4</em>: tcctccagtcct</td>
<td>A321+A322, digest Age-I, insert into pPD95.75 apEx21,22,23,24</td>
<td>GFP in pharynx, seam cells, vulval cells, distal tip cells, hypodermis</td>
</tr>
<tr>
<td>pJRBl</td>
<td>1.5kb start site promoter-GFP. attgccctcatatc</td>
<td>insert: A439+A440, digest Age-I, Hind-III cloned into pPD95.75 apEx41, 53</td>
<td>GFP non-specifically in posterior intestine</td>
</tr>
<tr>
<td>pJRBl</td>
<td>2kb <em>lin-4</em> promoter-GFP. gggaggttttaactt</td>
<td>A463+A464, digest Age-I, Hind-III apex 42,43,54,55 PQ5,6</td>
<td>GFP in intestine throughout development</td>
</tr>
<tr>
<td>pJRBl (clone 456)</td>
<td><em>lin-4</em> rescue fragment (minus unc54 3'UTR)</td>
<td>A370+A372, digest Apa-I, Hind-III clone into pFS1030 backbone with pcol-10-lacZ-lin41utr, unc-54utr removed. apEx29,35,36,37,38</td>
<td>rescues e912 mutant and apEx35, 36 express <em>lin-4</em> early (4h).</td>
</tr>
<tr>
<td>pJRBl (clone 454)</td>
<td><em>lin-4</em> rescue fragment with <em>lin-4</em> deleted (minus unc54)</td>
<td>same as pJRBl, but fortuitous mutation deleted 50nt including <em>lin-4</em> hairpin.</td>
<td>Does not cause changes in timing or levels of endogenous <em>lin-4</em></td>
</tr>
<tr>
<td>Construct</td>
<td>Description</td>
<td>Construction &amp; Strains</td>
<td>Comments</td>
</tr>
<tr>
<td>--------------</td>
<td>--------------------------------------</td>
<td>----------------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>pJRB8</td>
<td>lin-4 rescue fragment (plus unc-54 3'UTR)</td>
<td>A370+A371, digest with Nco-I, Hind-III, insert into pFS1030 backbone with pcol-10-lacZ-lin41utr removed. No surviving lines.</td>
<td>Harder to get lines, had to go lower in concentration than pJRB6</td>
</tr>
<tr>
<td>pJRB9 (clone 700)</td>
<td>lin-4 rescue fragment, extended to 2nd exon (plus unc-54 3'UTR)</td>
<td>A370+A373, digest with Nco-I, Hind-III, insert into pFS1030 backbone with pcol-10-lacZ-lin41utr removed. No surviving lines.</td>
<td></td>
</tr>
<tr>
<td>pJRB10 (clone 677)</td>
<td>myo2-GFP construct used in the lab (was clone #677)</td>
<td>Insert amplified by PCR from pYC2.1 from Schaefer lab. pYC2.1 was pmyo2-chamelion, so subcloned the promoter into pPD95.75. Oligos were A368+369, Age-1, Xba-1 (not Hind-III, see p. 32 lab notebook 6). Many, many strains!</td>
<td>Bright and consistent pharyngeal expression. Generally use 50ng/ul for good expression.</td>
</tr>
<tr>
<td>pJRB11 (clone 1022)</td>
<td>lin-4 sensor but DO NOT USE</td>
<td>Inserted lin-28 UTR into pSEH33, but removed lin-28 stop codon in the process. apEx130,131,132,133,134</td>
<td>DO NOT USE. Fixed clone=pJRB15.</td>
</tr>
<tr>
<td>pJRB12 (clone 1078)</td>
<td>lin-4 sensor, deleted LCE, but DO NOT USE</td>
<td>Inserted lin-28 UTR (LCE) into pSEH33, but removed lin-28 stop codon in the process. apEx135,136,137,138,139</td>
<td>DO NOT USE. Fixed clone=pJRB16.</td>
</tr>
<tr>
<td>pJRB13 (clone 1067)</td>
<td>wt alg-1 rescue fragment in TOPO vector</td>
<td>TOPO-cloned wt alg-1 rescue fragment (7kb), constructed with A347+A348. (linearize Sph-I) apEx140 (1ng/ul) apEx141 (0.01ng/ul) apex142 (0.01ng/ul)</td>
<td>failed to rescue alg-1(ap45) or alg-1(gk214) (no surviving lines in gk214 background). No apparent protein expression by Western.</td>
</tr>
<tr>
<td>Construct</td>
<td>Description</td>
<td>Construction &amp; Strains</td>
<td>Comments</td>
</tr>
<tr>
<td>---------------</td>
<td>------------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>pJRB14</td>
<td>alg-1 rescue fragment (7kb), with ap45 mutation, subcloned from pJRB13.</td>
<td>apEx154 (0.1 ng/ul) apEx155 (0.1 ng/ul) apEx156 (0.1 ng/ul)</td>
<td>failed to rescue alg-1 (ap45) or alg-1 (gk214) (line in gk214 background is not available). Some protein expression evident on Western.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>apex157 (1 ng/ul) apEx158 (1 ng) apEx159 (0.1 ng)</td>
<td></td>
</tr>
<tr>
<td>(clone 1104)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pJRB16</td>
<td>Fixed lin-4 sensor control (-LCE)</td>
<td>Amplified (-LCE) UTR sequence from pJRB12. Cloned into pPD136.15 with Nhe-I and Apa-I. *Note: inserted a unique Avr-II site where the LCE initially was. This allows any sequence to be inserted and tested for regulation. apEx190, 191, 192, 193 (all 25 ng/ul)</td>
<td>GFP well expressed until adult stage (probably downregulated at that time by let-7, which has an intact binding site in this lin-28 UTR sequence). Would likely work well as a let-7 sensor. Also can insert any sequence to be tested (Avr-II site)</td>
</tr>
<tr>
<td>(clone 1127)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pJRB17</td>
<td>lin-4 downstream intron (F59G1.4 intron #10) inserted into</td>
<td>To test whether the intron #10 of F59G1.4 contains expression regulatory elements. Replaced first intron of GFP with F59G1.4 intron #10.</td>
<td>GFP expression was indistinguishable from control lines with pPD136.15 (containing the</td>
</tr>
<tr>
<td>Construct</td>
<td>Description</td>
<td>Construction &amp; Strains</td>
<td>Comments</td>
</tr>
<tr>
<td>----------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>GFP</td>
<td>Was clone 1153.</td>
<td>apEx177,178,179,180,181,182 (1ng/ul) control (pPD136.15, 1ng/ul): apEx183,184,185</td>
<td>original, nonregulated intron. Concluded that the intron was not regulating lin-4 primary transcript processing.</td>
</tr>
<tr>
<td>pJRB18 (clone 1513)</td>
<td>Is a control for a TAP-Alg-1 fusion created by Shveta Bagga (see See page 103 of Shveta Bagga lab notebook VI). pJRB18 deleted the Alg-1 CDS.</td>
<td>To test whether GFP-Alg-1 was mislocalized because of Alg-1 sequences, deleted Alg-1 CDS from Shveta Bagga's SB2(corrected) and injected into worms. apEx188 (pJRB18, 50ng/ul) compare to: apEx186,187, PQ46 (SB2corrected, 50ng/ul)</td>
<td>pJRB18 is a control for SB2(corrected), a tap-tagged pCol10:GFP::Alg-1 translation fusion construct made by Shveta Bagga. We observed expression in hypodermal cells and exclusion from seam cells during development so in pJRB18 I deleted the alg-1 CDS to test whether the Alg-1 protein sequence was causing the GFP to localize to non seam cells. pJRB18 construct exhibited precisely the same expression pattern as SB2 (corrected), so we concluded that the col-10 promoter is not seam-cell specific in all C. elegans stages.</td>
</tr>
<tr>
<td>pJRB19 (clone)</td>
<td>Alg-1 rescue in bombardment cut out of pJRB13 with Spe-I, Not-I(blunt), cloned</td>
<td>bombardment did not work.</td>
<td></td>
</tr>
<tr>
<td>Construct</td>
<td>Description</td>
<td>Construction &amp; Strains</td>
<td>Comments</td>
</tr>
<tr>
<td>-----------</td>
<td>-------------</td>
<td>------------------------</td>
<td>----------</td>
</tr>
<tr>
<td>1303)</td>
<td>vector</td>
<td>into pPDMM016b digested with XbaI and SacII(blunt). no lines.</td>
<td></td>
</tr>
<tr>
<td>pJRB20</td>
<td>A construct designed to overexpress the <em>lin-4</em> primary transcript without <em>lin-4</em> hairpin. I hoped to see phenotypes as I titrated out factors necessary for processing.</td>
<td>I fused pJRB7 (454, rescue fragment with deleted <em>lin-4</em> hairpin) with pJRB9 (rescue fragment, extended) to capture as much of the primary transcript as possible. I linearized pJRB7 with BglII, and pJRB9 with Kpn-I, then mixed linearized plasmids as templates for PCR amplification with A1127+A1128.</td>
<td>Transgenics looked normal (no phenotypes) but no lines were isolated due to problems with contaminated plates. Did not follow up.</td>
</tr>
<tr>
<td>pJRB22</td>
<td><em>lin-4</em> rescue fragment, deleted from 10nt downstream of 16h start site (tcctcccatgtcc)</td>
<td>Δ10-10 rescue made by stitching fragment (A1153+A372) and fragment (A1154+A370) with A370+A372 as outside primers. Construct is in TOPO vector.</td>
<td>Saw a failure to rescue but had to switch this over to the bombardment vector as well.</td>
</tr>
<tr>
<td>(clone 5354)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>


<table>
<thead>
<tr>
<th>Construct</th>
<th>Description</th>
<th>Construction &amp; Strains</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>pJRB23 (clone 1691)</td>
<td>to 10nt upstream of 4h-1 start site. (tttcggtcact) See p. 46 of book 12. Called this Δ10-10.</td>
<td>No lines.</td>
<td></td>
</tr>
<tr>
<td>pJRB24 (clone 1701)</td>
<td><em>lin</em>-4 rescue Δ16h start site (10nt on either side)</td>
<td>Created by stitching fragment (A1156+A370) and fragment (A1155+A372) with A370+A372 as outside primers.</td>
<td>eventually put into bombardment vector for testing.</td>
</tr>
<tr>
<td>pJRB25 (clone 1705)</td>
<td><em>plin</em>-4::GFP construct, with GFP inserted just after 16h start site (<em>pJRB3 includes sequence past 4h-1)</em>.</td>
<td>Amplified A1213+A1214, cloned into pPD95.75 with Hind-III, Age-I. apEx221,222,223,224,225,226,227, 228</td>
<td>GFP looked very similar to pJRB3. In pharynx, neurons, and seam cells.</td>
</tr>
<tr>
<td>pJRB26 (clone 1752)</td>
<td><em>plin</em>-4::GFP construct, with GFP inserted just after 16h start site but with Burge Motif deleted exactly (8nt). Burge Motif (BM) = CTCCGCCC</td>
<td>Amplified A1214+A1215, cloned into pPD95.75 with Hind-III, Age-I. apEx229,230</td>
<td>GFP looked exactly the same as pJRB24. In pharynx, neurons, and seam cells. It appears the Burge Motif is not important for transcriptional regulation.</td>
</tr>
</tbody>
</table>

* BM = CTCCGCCC
<table>
<thead>
<tr>
<th>Construct</th>
<th>Description</th>
<th>Construction &amp; Strains</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>(clone 1756)</td>
<td>(9nt deleted total). Keeps GFP in frame with an upstream start codon.</td>
<td></td>
<td>transcriptional regulation.</td>
</tr>
<tr>
<td>pJRB27</td>
<td>lin-4 promoter::GFP fusion with GFP in place of mature microRNA.</td>
<td>Amplified A1214+A1247, cloned into pPD95.75 with Hind-III, Age-I. no lines.</td>
<td>never injected.</td>
</tr>
<tr>
<td>pJRB28 (clone 1809)</td>
<td>original lin-4 rescue fragment cloned into bombardment vector pPDMM016b</td>
<td>digested lin-4 rescue fragment out of pJRB6 with Hind-III and Apa-I, then cloned into the same sites of pPDMM016b.</td>
<td>never bombarded</td>
</tr>
<tr>
<td>pJRB29 (clone 1802)</td>
<td>lin-4 rescue fragment extended to endogenous pA signal (3kb rescue fragment), cloned into bombardment vector pPDMM016b.</td>
<td>PCR amplified A1283+A1284 from N2 genomic DNA, digested with Ava-I, Apa-I and cloned into corresponding sites of pPDMM016b. apls245 (integrated, no GFP) apEx246 (array)</td>
<td>Does not have GFP so bombards were carried out with both pJRB28 and pJRB10; integrated lines only had pJRB28. Decided to include the GFP on the same plasmid as the rescue construct in all future experiments, pJRB31 is the bombardment vector with a myo2::GFP expression construct.</td>
</tr>
<tr>
<td>pJRB30 (clone)</td>
<td>Same as pJRB29,</td>
<td>PCR amplified A1112+A1293, digest with</td>
<td>I thought proper lin-4 expression</td>
</tr>
<tr>
<td>Construct</td>
<td>Description</td>
<td>Construction &amp; Strains</td>
<td>Comments</td>
</tr>
<tr>
<td>-----------</td>
<td>-------------</td>
<td>------------------------</td>
<td>----------</td>
</tr>
<tr>
<td>1903)</td>
<td>except rescue extended ~3kb in 5’ direction from <em>lin-4</em>. (making approximately 6kb rescue fragment).</td>
<td>Apa-I, SnaBl, cloned into equivalent sites of pJRB29.</td>
<td>might require even more upstream sequence. This clone was never used because it also lacks a GFP expression cassette.</td>
</tr>
<tr>
<td>pJRB31 (clone 2202)</td>
<td>Is pPDMM016b with pmyo2::GFP cassette, so that rescuing constructs can be followed outside unc-119 background.</td>
<td>Created by digesting pJRB10 with Spe-I, then blunting; and cutting pPDMM016b with Sac-II, and blunting. Both were then digested with Xba-I and ligated together.</td>
<td>Used as foundation for pJRB32, pJRB33, pJRB34, pJRB35, pJRB36.</td>
</tr>
<tr>
<td>pJRB32 (clone 2501)</td>
<td>3kb rescue construct put into pJRB31.</td>
<td>Removed 3kb rescue from pJRB29 by digestion with Apa-I and Ava-I, cloned into equivalent sites of pJRB31. Had to partially digest pJRB31 with Ava-I because of an extra site in the plasmid backbone. aPlS243: PQ91 (in unc-119), PQ97(N2 background), PQ95(e912 background). aPlS244: PQ92 (in unc-119), PQ (in e912 background). aPlS386: PQ269 (e912 background), PQ265 (N2 background).</td>
<td>Bombarde and crossed into <em>lin-4</em>(e912) background. Still saw variable rescue and variable mature <em>lin-4</em> expression. Concluded that mos-SCI system is best way to perform rescue experiments and ran out of time to work on characterizing role of cis-elements in <em>lin-4</em> expression.</td>
</tr>
<tr>
<td>pJRB33 (clone 2430)</td>
<td>6kb rescue in pJRB31</td>
<td>Removed 6kb rescue from pJRB30 by digestion with Apa-I and Ava-I, cloned</td>
<td>no lines made.</td>
</tr>
<tr>
<td>Construct</td>
<td>Description</td>
<td>Construction &amp; Strains</td>
<td>Comments</td>
</tr>
<tr>
<td>-----------</td>
<td>-------------</td>
<td>------------------------</td>
<td>----------</td>
</tr>
<tr>
<td>pJRB34 (clone D3)</td>
<td>3kb rescue ΔBurge Motif (ΔBM) in pJRB31. See notes for pJRB21 for details on this deletion.</td>
<td>Amplified fragment (A1151+A1284) and fragment (A1152+A1283), stitched together with A1284+A1283, then digested with Apa-I and Ava-I. Ligated into partially-digested pJRB31 (see note for pJRB32). in unc-119: PQ98,138,139,140 in e912: PQ182,184,186 in N2: PQ183, 185</td>
<td>Bombarded and crossed into lin-4(e912) background. Still saw variable rescue and variable mature lin-4 expression. Concluded that mos-SCI system is best way to perform rescue experiments and ran out of time to work on characterizing role of cis-elements in lin-4 expression.</td>
</tr>
<tr>
<td>pJRB35 (clone H12)</td>
<td>3kb rescue Δ16ss in pJRB31 See notes for pJRB23 for details on this deletion.</td>
<td>Amplified fragment (A1155+A1284) and fragment (A1156+A1283), stitched together with A1284+A1283, then digested with Apa-I and Ava-I. Ligated into partially-digested pJRB31</td>
<td>Full rescue and lin-4 was expressed normally. Only got one cross into e912 to work.</td>
</tr>
<tr>
<td>Construct</td>
<td>Description</td>
<td>Construction &amp; Strains</td>
<td>Comments</td>
</tr>
<tr>
<td>-----------</td>
<td>-------------</td>
<td>-------------------------</td>
<td>----------</td>
</tr>
<tr>
<td><strong>pJRB36 (clone H4)</strong></td>
<td>3kb rescue Δ10-10 in pJRB31</td>
<td>(see note for pJRB32). in unc-119: PQ179, PQ180 in e912: PQ259 (crossed PQ180 into e912)</td>
<td>Amplified fragment (A1153+A1284) and fragment (A1154+A1283), stitched together with A1284+A1283, then digested with Apa-I and Ava-I. Ligated into partially-digested pJRB31 (see note for pJRB32). In unc-119: PQ141,142,143,144,145,146,147 In e912: PQ262,250,252,254,255,256, (not able to homozygose PQ142 into e912) In N2: PQ263,249,251,253,257</td>
</tr>
<tr>
<td><strong>pJRB37 (clone A1)</strong></td>
<td>*note: lines PQ107-116 did not have GFP expression and apparently did not have the correct transgene integrated. Genotyping them demonstrated no unc-119 rescue fragment, although they were crawlers. Perhaps they were mutationally induced suppressors of unc-119.</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>pJRB37 (clone A1)</strong></td>
<td>lin-4 sensor control (ΔLCE) pJRB16 with</td>
<td>PCR amplified lin-4 hairpin (~70nt) with A1636+A1637, digested</td>
<td>My idea was that this construct would measure</td>
</tr>
<tr>
<td>Construct</td>
<td>Description</td>
<td>Construction &amp; Strains</td>
<td>Comments</td>
</tr>
<tr>
<td>-----------</td>
<td>-------------</td>
<td>------------------------</td>
<td>----------</td>
</tr>
<tr>
<td>pJRB38 (clone B1)</td>
<td><em>lin-4</em> sensor control (ΔLCE) <em>pJRB16</em> with the <em>lin-4</em> precursor hairpin extended upstream and downstream (~300nt total), cloned into LCE site (at unique Avr-II site).</td>
<td>PCR amplified <em>lin-4</em> hairpin with upstream, downstream regions (~300nt) with A1638+A1639, digested with Avr-II, and cloned into Avr-II site of pJRB16. Constructed in the event that <em>lin-4</em> surrounding regions are necessary for proper Drosha processing.</td>
<td>Same as pJRB37 except I included regions surrounding the <em>lin-4</em> hairpin. This construct was to test for regulated processing of <em>lin-4</em> hairpin but I was not able to generate transgenic lines; apparently <em>lin-4</em> was efficiently processed. No lines were generated. I did not re-inject at lower concentration.</td>
</tr>
<tr>
<td></td>
<td>the <em>lin-4</em> precursor hairpin cloned into LCE site (by unique Avr-II site).</td>
<td>with Avr-II, and cloned into Avr-II site of pJRB16. No lines.</td>
<td>Drosha processing of <em>lin-4</em>, since cleavage of the hairpin in the 3’UTR of GFP mRNA ought to lead to its degradation. If processing of <em>lin-4</em> by Drosha is regulated, I would see GFP expression decrease as the worms get older. Initially we injected at 25ng/ul which only produced growth arrested transgenic worms suggesting that <em>lin-4</em> was efficiently processed. No lines were generated. I did not re-inject at lower concentration.</td>
</tr>
<tr>
<td>Construct</td>
<td>Description</td>
<td>Construction &amp; Strains</td>
<td>Comments</td>
</tr>
<tr>
<td>------------</td>
<td>--------------------------------------</td>
<td>----------------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>pJR39</td>
<td>3kb lin-4 rescue in mos-SCI chromosome I integration plasmid from Angela Parrish.</td>
<td>Removed 3kb rescue from pJR29 by digesting with Apa-I, then blunting, then Xho-I digesting and gel purifying the 3kb band. Cloned into mosA plasmid which was digested with Not-I and blunted, followed by digestion with Xho-I to create a sticky end compatible with the rescuing insert.</td>
<td>Intermediate in building rescue vector. Next step was insertion of GFP expression cassette (see pJR40).</td>
</tr>
<tr>
<td>pJR41</td>
<td>Inserted 6x MS2-stem-loop into pJR32</td>
<td>Partially digested pJR32 with Bst XI to get single digested plasmid, then inserted 6x stem-loops into site just downstream of lin-4 hairpin (verified by colony PCR and sequencing). The 6x stem-loop insert was amplified with A1640+A1320 (template was 6x stem-loop plasmid).</td>
<td>Never tested due to time constraints. The goal was to perform MS2::GFP localization of primary transcripts and also biochemical pull-down of primary transcripts and isolate interacting regulatory factors.</td>
</tr>
<tr>
<td>pJR42</td>
<td>Inserted 12x MS2-stem-loop into pJR32</td>
<td>Partially digested pJR32 with Bst XI to get single digested plasmid, then inserted 12x stem-loops into site just downstream</td>
<td>Never tested due to time constraints. The goal was to perform MS2::GFP localization of</td>
</tr>
<tr>
<td>Construct</td>
<td>Description</td>
<td>Construction &amp; Strains</td>
<td>Comments</td>
</tr>
<tr>
<td>-----------</td>
<td>-------------</td>
<td>------------------------</td>
<td>----------</td>
</tr>
<tr>
<td></td>
<td>of <em>lin-4</em> hairpin (verified by colony PCR and sequencing). The 12x stem-loop insert was amplified with A1640+A1320 (template was 12x stem-loop plasmid).</td>
<td>primary transcripts and also biochemical pull-down of primary transcripts to isolate interacting regulatory factors.</td>
<td></td>
</tr>
</tbody>
</table>
References


