Understanding microRNA biogenesis and function through annotation of primary miRNA transcripts and characterization of functional interactions between microRNAs and RNA-binding proteins.

by

Dustin Haskell

B.S., Central Michigan University, 2014
M.S., Central Michigan University, 2016

AN ABSTRACT OF A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Division of Biology
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2021

Abstract

The gene expression programs that establish and maintain cellular and organism homeostasis require precise, potent, and multifaceted forms of regulation. Post-transcriptional mechanisms of regulation rely on the combinatorial action of two major classes of effectors: RNA-binding proteins (RBPs) and microRNAs (miRNAs). miRNAs are small noncoding RNAs that interact with many developmental and cellular pathways by repressing gene targets and are therefore critical to the execution of gene expression programs. miRNA dysfunction can lead to widespread disruption of gene regulatory networks, contributing to the occurrence and progression of developmental disorders and pathologies such as cancer. Most miRNAs are generated through a complex biogenesis that includes RNA Pol II-dependent transcription, successive enzymatic processing by endonucleases DRSH-1 and DCR-1 and loading into Argonaute proteins to form the miRNA induced silencing complex (miRISC). Guided by the loaded miRNA, miRISC binds the 3'UTR of a target mRNA and actively downregulates its expression through translation repression or mRNA degradation.

Mature miRNAs are produced through a series of enzymatic processing steps. Initial processing of primary miRNA gene transcripts (pri-miRNAs), performed by endonuclease DRSH-1, often occurs co-transcriptionally or shortly thereafter. Hence, pri-miRNA transcripts are largely absent from traditional RNA sequencing data sets, and thus difficult to characterize. The lack of primary miRNA annotations has hindered efforts to understand the mechanisms that modulate miRNA gene expression and complicated our ability to study the regulation of pri-miRNA processing. To fill this gap, we used an auxin-induced degron system to conditionally deplete DRSH-1 and greatly reduce processing of pri-miRNAs, leading to their accumulation. Subsequent RNAseq experiments identified pri-miRNAs and allowed for their annotation,

revealing previously unappreciated, complex genomic features of the miRNA loci and providing an essential resource for future studies of miRNA regulation. In addition, we identified >300 novel transcripts, uncovering existence of previously uncharacterized RNAs that may depend on DRSH-1 for processing, thus expanding the known *C. elegans* transcriptome.

Once miRNAs are processed to their mature form, they exert their repressive functions by targeting miRISC to the 3' UTRs of mRNA transcripts through partial base-pair complementarity. RBPs represent an important class of molecules that contribute to post-transcriptional regulation of gene expression, however, the extent of functional RBP coordination with miRNAs is largely unexplored. Similarly, a comprehensive understanding of how RBPs coordinate with miRNAs to regulate gene expression is lacking. To address the potential functional interaction between miRNAs and RBPs, I performed a targeted RNAi screen of 27 K-homology (KH) domain RBPs to identify factors that genetically interact with five miRNA sensitized mutant backgrounds. I identified multiple KH domain RBPs that functionally interact with all or some of miRNAs families tested, expanding our understanding of the crosstalk between two classes of post-transcriptional gene regulators.

Overall, this work has expanded our understanding of miRNA gene structure and the characteristics of primary miRNA transcripts, ultimately providing a valuable tool for future study of pri-miRNA transcription and processing. Furthermore, this work established a functional relationship between several RNA-binding proteins and developmental miRNA pathways, thus identifying candidates for future studies of functional RBP-miRNA interactions.

Understanding microRNA biogenesis and function through annotation of primary miRNA transcripts and characterization of functional interactions between microRNAs and RNA-binding proteins.

by

Dustin Haskell

B.S., Central Michigan University, 2014
M.S., Central Michigan University, 2016

A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Division of Biology
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2021

Approved by:

Anna Zinovyeva
Major Professor

Copyright

© Dustin Haskell 2021.

Abstract

The gene expression programs that establish and maintain cellular and organism homeostasis require precise, potent, and multifaceted forms of regulation. Post-transcriptional mechanisms of regulation rely on the combinatorial action of two major classes of effectors: RNA-binding proteins (RBPs) and microRNAs (miRNAs). miRNAs are small noncoding RNAs that interact with many developmental and cellular pathways by repressing gene targets and are therefore critical to the execution of gene expression programs. miRNA dysfunction can lead to widespread disruption of gene regulatory networks, contributing to the occurrence and progression of developmental disorders and pathologies such as cancer. Most miRNAs are generated through a complex biogenesis that includes RNA Pol II-dependent transcription, successive enzymatic processing by endonucleases DRSH-1 and DCR-1 and loading into Argonaute (AGO) proteins to form the miRNA induced silencing complex (miRISC). Guided by the loaded miRNA, miRISC binds the 3'UTR of a target mRNA and actively downregulates its expression through translation repression or mRNA degradation.

Mature miRNAs are produced through a series of enzymatic processing steps. Initial processing of primary miRNA gene transcripts (pri-miRNAs), performed by endonuclease DRSH-1, often occurs co-transcriptionally or shortly thereafter. Hence, pri-miRNA transcripts are largely absent from traditional RNA sequencing data sets, and thus difficult to characterize. The lack of primary miRNA annotations has hindered efforts to understand the mechanisms that modulate miRNA gene expression and complicated our ability to study the regulation of pri-miRNA processing. To fill this gap, we used an auxin-induced degron system to conditionally deplete DRSH-1 and greatly reduce processing of pri-miRNAs, leading to their accumulation. Subsequent RNAseq experiments identified pri-miRNAs and allowed for their annotation,

revealing previously unappreciated, complex genomic features of the miRNA loci and providing an essential resource for future studies of miRNA regulation. In addition, we identified >300 novel transcripts, uncovering existence of previously uncharacterized RNAs that may depend on DRSH-1 for processing, thus expanding the known *C. elegans* transcriptome.

Once miRNAs are processed to their mature form, they exert their repressive functions by targeting miRISC to the 3' UTRs of mRNA transcripts through partial base-pair complementarity. RBPs represent an important class of molecules that contribute to post-transcriptional regulation of gene expression, however, the extent of functional RBP coordination with miRNAs is largely unexplored. Similarly, a comprehensive understanding of how RBPs coordinate with miRNAs to regulate gene expression is lacking. To address the potential functional interaction between miRNAs and RBPs, I performed a targeted RNAi screen of 27 K-homology (KH) domain RBPs to identify factors that genetically interact with five miRNA sensitized mutant backgrounds. I identified multiple KH domain RBPs that functionally interact with all or some of miRNAs families tested, expanding our understanding of the crosstalk between two classes of post-transcriptional gene regulators.

Overall, this work has expanded our understanding of miRNA gene structure and the characteristics of primary miRNA transcripts, ultimately providing a valuable tool for future study of pri-miRNA transcription and processing. Furthermore, this work established a functional relationship between several RNA-binding proteins and developmental miRNA pathways, thus identifying candidates for future studies of functional RBP-miRNA interactions.

Table of Contents

List of Figures

List of Tables

Acknowledgements

I would like to thank my Advisor, Anna Zinovyeva and all of the member of my advisory committee for their guidance, expertise, and helpful chats. A special thanks to the members of the Zinovyeva Lab, especially Shilpa Hebbar who always listened to my ideas and crazy schemes, sometimes against her better judgement. Lastly, I would like to thank my loved ones for their unconditional support through all the triumphs and challenges of graduate school.

Chapter 1 - RNA binding proteins and microRNAs functionally interact to exert combinatorial effects on gene expression

Review manuscript in preparation.

Authors: Haskell, D. & Zinovyeva, A, 2021

Regulation of Gene Expression

Overview

Precise and multifaceted regulation of gene expression programs is critical to determination and maintenance of cellular and organism homeostasis. Gene expression begins with transcription of genomic encoded DNA sequences that gives rise to messenger RNA (mRNA), which in turn is translated to make proteins. The faithful execution of these encoded gene expression programs is critical for proper development and homeostasis. Disruption of these programs at any level can lead to the failure of essential cellular functions and ultimately endanger organism health. Therefore, gene expression needs to be precisely coordinated, with regulation occurring at all levels starting with chromatin organization, transcriptional and post-transcriptional regulation, and extending to translation and post-translational regulation.

Gene transcription is tightly regulated by both *trans*-acting factors (such as transcription factors) and *cis*-acting regulatory elements (reviewed in Lee and Young 2013). The combination of these factors first ensures precise patterning of active/repressed transcription and then effectively modulates the rate and efficiency of active transcriptional units (reviewed in Casamassimi and Ciccodicola 2019). This ensures appropriate spatio-temporal production of mRNA transcripts. Once produced, the mRNA transcripts themselves are subject to regulation through modulation of mRNA stability, transport/accessibility of transcripts, and regulation of translation initiation. At the level of translation, regulatory mechanisms affect rates of translation initiation and elongation, thus titrating the levels of protein produced (reviewed in Hershey *et al.* 2012). In transcriptional regulation, loss or mutation of *trans*-acting factors and *cis* regulatory elements have been associated with various pathologies and disorders. For example, mutations or loss of transcription factor activity or regulatory elements result in widespread disruption of gene

expression patterns and are commonly observed in cancers (reviewed in Gebauer *et al.* 2021). Mutations in pancreatic specific transcription factors and their binding sites can disrupt pancreas-specific gene expression profiles, leading to diabetes (Muarano *et al*. 2012). Similarly, loss of translational control generally has widespread consequences and often contributes to the occurrence or progression of cancers (reviewed in Xu and Ruggero 2020). In addition, translational dysfunction has been implicated in autism spectrum disorders (Hooshmandi *et al.* 2020), neurodegenerative disorders (reviewed in Bosco 2018) and aging (Skariah and Todd 2021).

Post-transcriptional regulation of gene expression

While much of the regulation of gene expression occurs at the transcriptional and translational levels, a powerful suite of regulatory mechanisms exists in the transition between these two processes. Post-transcriptional regulation generally targets mRNA transcripts directly, by modulating mRNA processing, localization, or stability. The result is sequestration or downregulation of target transcripts leading to a complete or partial reduction of gene expression (reviewed in Corbett 2018). Post-transcriptional regulation is affected by two classes of regulators: RNA binding proteins (RBPs) and small non-coding RNAs called microRNAs (miRNAs). The combinatorial action of these two regulatory classes allows for extensive regulation of mRNAs. Through specific pairing of RBPs and associated small RNAs, cells can provide highly specialized regulation of mRNAs contributing to the enormous complexity of gene expression patterns (reviewed in Corbett 2018).

In humans, more than 1500 RNA binding proteins are thought to contribute to gene expression regulation (Gerstberger *et al.* 2014). Diverse groups of RBPs regulate mRNA

processing, stability, transport, and function (Schoenberg and Maquat 2012; Xing and Bassell 2013; Buxbaum *et al.* 2015). Some RBPs act on large subsets of mRNAs or even the whole transcriptome (Hogan *et al.* 2008; Nostrand *et al.* 2020). Indeed, it is estimated that each human RBP has tens of thousands of binding sites in the 3'UTRs of mRNAs (Nostrand *et al.* 2016), thus enabling >33 million possible post-transcriptional interactions between human RBPs and mRNAs (Kim *et al.* 2021). Despite the enormous number of possible interactions, it is likely that many RBP-mRNA interactions are more specialized, with RBP expression in a tissue or cell specific manner limiting the number of transcripts targeted (Díaz-Muñoz and Turner 2018). Ultimately, RBP-target mRNA interactions are dictated by the RNA binding sites. Recent large-scale characterization of RBP binding sites suggest that binding sequences are biased towards low compositional complexity (Dominguez *et al.* 2018), consistent with earlier observations that polypyrimidine tracts like A/U or C/U-rich regions, G-rich regions (Wang *et al.* 2012), or polyuridine tracts (Cieniková *et al.* 2014) are all common binding sequences. Increasing in complexity, a subset of RBPs appear to preferentially bind simple bipartite motifs (Dominguez *et al.* 2018) or even defined consensus sequences (Choi *et al.* 2017; Kuang and Wang 2020). In addition, local RNA structure (such as hairpin loops, stems, and stem bulges) also impact RBP binding and are the major binding determinant for proteins like RBM22, ZNF326, and many KH-domain containing proteins (Dominguez *et al.* 2018). Interestingly, it has been observed that most RBPs only bind a subset of available motifs in expressed transcripts (Taliaferro *et al.* 2016). This preferred binding can be explained in part by RBP localization or expression which leads to differential access to transcripts and thus alters RBP-mRNA specificity (Damianov *et al.* 2016).

Through these interactions, some RBPs promote gene expression (Lin *et al.* 2016; Zhang *et al.* 2018; Kim *et al.* 2021) or repress it (Léveillé *et al.* 2011; Degrauwe *et al.* 2016; Kim *et al.* 2021). The functional effect of RBPs on gene expression can vary depending on gene target or molecular context, a phenomenon highlighted by RBPs HuR and PTB, which can either enhance or suppress the downregulation of their targets based on the specific cellular context (Kim *et al.* 2009; Xue *et al.* 2013; Ahuja *et al.* 2016). Underlying the importance of precisely coordinated action of RBPs, several diseases and pathologies result from their mutation and loss. For example, loss of the RBP FUS leads to widespread dysregulation of mRNA decay pathways in diseases like amyotrophic lateral sclerosis (ALS) (Kamelgarn *et al.* 2018). Additionally, loss of RNA binding protein FMRP function disrupts neurodevelopment leading to Fragile X syndrome (Richter and Zhao 2021) and has also been associated with autism spectrum disorders (Hooshmandi *et al.* 2020) and plaque forming neurodegenerative diseases such as Alzheimer's (Ghosh *et al.* 2020).

The second class of effectors in post-transcriptional regulation of gene expression are miRNAs. These molecules are a class of small non-coding RNAs that function as potent repressors of gene expression (reviewed in Bartel 2018), although there is some evidence that miRNAs may act as activators in specific contexts (reviewed in Ramchandran and Chaluvally-Raghavan 2017). miRNAs target mRNA transcripts through partial sequence complementarity between the miRNA and a miRNA binding site, usually located in 3'UTRs of mRNAs. This interaction guides the miRNA induced silencing complex (miRISC) to the target mRNAs and results in their translational inhibition and/or degradation. The number of human miRNA candidates continues to fluctuate, with recent estimates of 2300-2600 miRNAs (Alles *et al.* 2019; Plotnikova *et al.* 2019). Each miRNA is predicted to target multiple transcripts, although the

exact number can range between one and >300 transcripts (Plotnikova *et al.* 2019). Wide-ranging evidence points to the fact that miRNAs and RBPs coordinate to regulate gene expression on target mRNAs (Kim *et al.* 202; reviewed in Bartel 2018; O'Brien *et al.* 2018), resulting in a myriad of combinatorial interactions.

To facilitate the study of these two effectors of post-transcriptional gene regulation, many investigations have utilized the model organism *Caenorhabditis elegans.* In comparison to the much larger number of RBPs/miRNAs in humans, the *C. elegan*s genome encodes ~250 RBPs (Tamburino *et al.* 2013), with 253 miRNAs currently annotated by miRbase (Kozomara *et al.* 2018). The smaller number of possible nematode RBP-miRNA interactions makes *C. elegans* a more manageable system to study the functional effects of these interactions. In addition, assays with quantifiable phenotypical outputs of both RBP and miRNA-mediated gene regulatory function allow for accurate assessments of that activity. Lastly, utilization of both genetic and biochemical methods allows for the dissection of functional and mechanistic interactions between RBPs and miRNAs, facilitating the characterization of their combinatorial effects on gene expression.

This review will discuss the mechanisms of miRNA biogenesis and function as well as the known molecular mechanisms of functional interactions between miRNAs and RNA binding proteins, two potent effectors of post-transcriptional gene expression regulation.

miRNA biogenesis

Overview

miRNAs are a broad class of small non-coding RNAs of 18-24nt in length that have extensive roles in the post-transcriptional regulation of gene expression (reviewed in Bartel

2018). miRNAs are produced through a complex biogenesis pathway that occurs in multiple cellular compartments and requires the successive enzymatic cleavages of two intermediate species by the endonucleases Drosha and Dicer (reviewed in Ha and Kim 2014; O'Brien *et al.* 2018, Figure 1.1). The biogenesis of most miRNAs begins with the transcription of a miRNA gene, transcribed from the genome by RNA polymerase II (Lee *et al.* 2004). A large portion of primary miRNA transcripts (pri-miRNA) are capped and poly-adenylated, providing further evidence of RNA Pol II transcription (Lee *et al.* 2004). However, comprehensive characterization of pri-miRNAs has been difficult due to the short half-life of the primary transcripts (Lee *et al.* 2004; Conrad *et al.* 2014). After transcription, the pri-miRNAs are cleaved by the Microprocessor complex (composed of Drosha and DGCR8) to liberate the ~70 nt long precursor miRNA (pre-miRNA) (Kim *et al.* 2016; Figure 1.1). The pre-miRNA is then exported out of the nucleus, utilizing the pre-miRNA-specific exportin complex. Once in the nucleus, the pre-miRNA is further cleaved to liberate the dsRNA miRNA duplex (Rybak-Wolf *et al.* 2014, Figure 1.1). This duplex is made up the guide strand (that will go on to target a mRNA transcript) and a passenger strand that is later degraded. The duplex is then loaded into an Argonaute protein (AGO), at which point the duplex is unwound, and the passenger strand ejected (Kawamata *et al.* 2009; Yoda *et al.* 2010). The miRNA-loaded AGO forms the minimal miRNA-induced silencing complex (miRISC). This complex then goes on to target a mRNA transcript, guided by the semi-complementarity between the guide miRNA strand and the miRNA target site in the 3'UTR of the mRNA (reviewed in Bartel 2018). The effect of miRISC binding is the active downregulation of the target transcript, either through mRNA decay or translational inhibition (Figure 1.1).

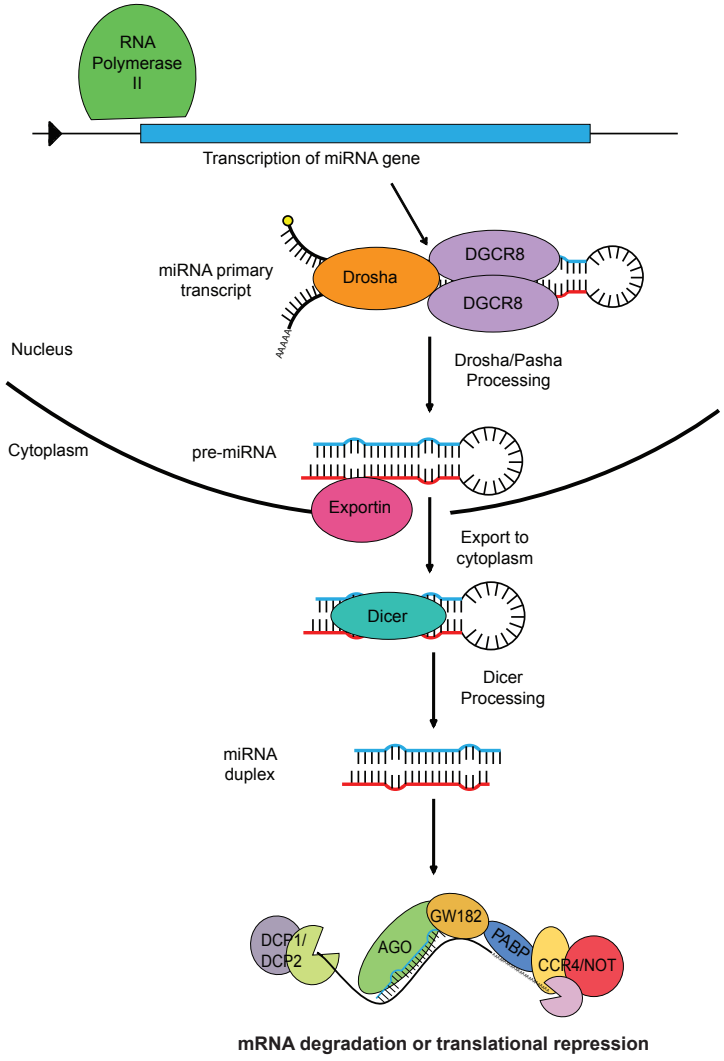


Figure 1.1. Canonical miRNA biogenesis pathway.

miRNAs are transcribed from a miRNA gene to produce a primary transcript (pri-miRNA). The pri-miRNA is processed by Drosha to produce the precursor miRNA (pre-miRNA) which is exported from the nucleus. Once in the cytoplasm, the pre-miRNA is cleaved by Dicer to produce the miRNA duplex containing both the guide and passenger miRNA strand. The guide strand associates with Argonaute (AGO) and accessory proteins to target mRNA for degradation or translational repression.

Drosha Processing

pri-miRNAs undergo enzymatic processing by Drosha (DRSH-1 in *C. elegans*), an RNAse III nuclease responsible specifically for the processing of primary miRNA transcripts (Lee *et al.* 2003), aided by DGCR8 (PSHA-1 in *C. elegans*) as part of the Microprocessor complex (Figure 1.1). Drosha has several functional domains: the dsRNA binding domain (RBD), two RNAase III domains, and a central domain (CED) (Nguyen *et al.* 2015; Li and Patel 2016). The RBD is thought to increase recognition and efficiency of pri-miRNA binding, thus facilitating enzymatic processing in the catalytic core formed by the two RNase III domains (Han *et al.* 2004, 2009). The CED is necessary for processing of pri-miRNAs as it contains both a platform motif and a PAZ-like domain that increase recognition of the pri-miRNA (Kwon *et al.* 2016). Alone, Drosha is capable of endonuclease activity, however efficient enzymatic processing of pri-miRNAs requires the formation of the complete Microprocessor complex, comprised of Drosha and accessory protein DGCR8/Pasha in a 2:1 ratio (Denli *et al.* 2004; Han *et al.* 2004). DGCR8 contains an RNA-binding heme domain (Rhed) necessary for homodimerization (Quick-Cleveland *et al.* 2014), two double stranded RNA-binding domains (dsRBDs) that facilitate Microprocessor binding to the pri-miRNAs (Yeom *et al.* 2006), and a C-terminal region required for Drosha binding (Nguyen *et al.* 2015; Li and Patel 2016). Once formed, the Microprocessor rapidly processes pri-miRNA transcripts as evidenced by the short half-life of most pri-miRNAs (Conrad *et al.* 2014). The speed at which processing occurs suggests that Drosha-mediated processing may be occurring co-transcriptionally, especially in instances where pri-miRNAs are intronic (Kim and Kim 2007; Morlando *et al.* 2008). In further support of the co-transcriptional model, a subset of pri-miRNAs has been observed to lack

poly(A) tails in addition to the 5' cap, suggesting rapid Microprocessor activity after the start of transcription (Ballarino *et al.* 2009).

Early analysis of pri-miRNA structure suggested that several structural landmarks present in all primary miRNA transcripts are responsible for basic recognition by the Microprocessor, specifically the imperfect stem structure flanked by an apical loop and two basal stretches of single stranded RNA (ssRNA) (Han *et al.* 2006). In fact, the perfectly base-paired basal stem region and flanking ssRNA regions are key for recognition of pri-miRNAs (Zeng and Cullen 2005, Figure 1.2A). A UG motif at the base of the hairpin and a UGU/GUG motif in the apical loop ensure efficient recognition by the Microprocessor in humans (Auyeung *et al.* 2013), and a mismatched GHG motif near the base of the hairpin may increase processing efficiency (Fang and Bartel 2015). Interestingly, while these motifs appear critical for efficient recognition and processing of human pri-miRNAs, they appear to be largely dispensable in *C. elegans*, as large-scale sequence/structure analysis showed they are largely absent from nematode pri-miRNAs (Auyeung *et al.* 2013). Taken together, the characteristic motifs and structure allow Microprocessor to properly orient itself on the pri-miRNA, measure appropriately, and cleave ~ 11nt from the basal junction and ~22nt from the apical junction, thus liberating pre-miRNA stemloop (reviewed in Bartel 2018, Figure 1.2A). Despite general agreement in cleavage sites there is some evidence that Microprocessor can make some non-canonical cuts leading to alternative forms of pre-miRNAs, or isomiRs (Nguyen *et al.* 2015, reviewed in Bartel 2018).

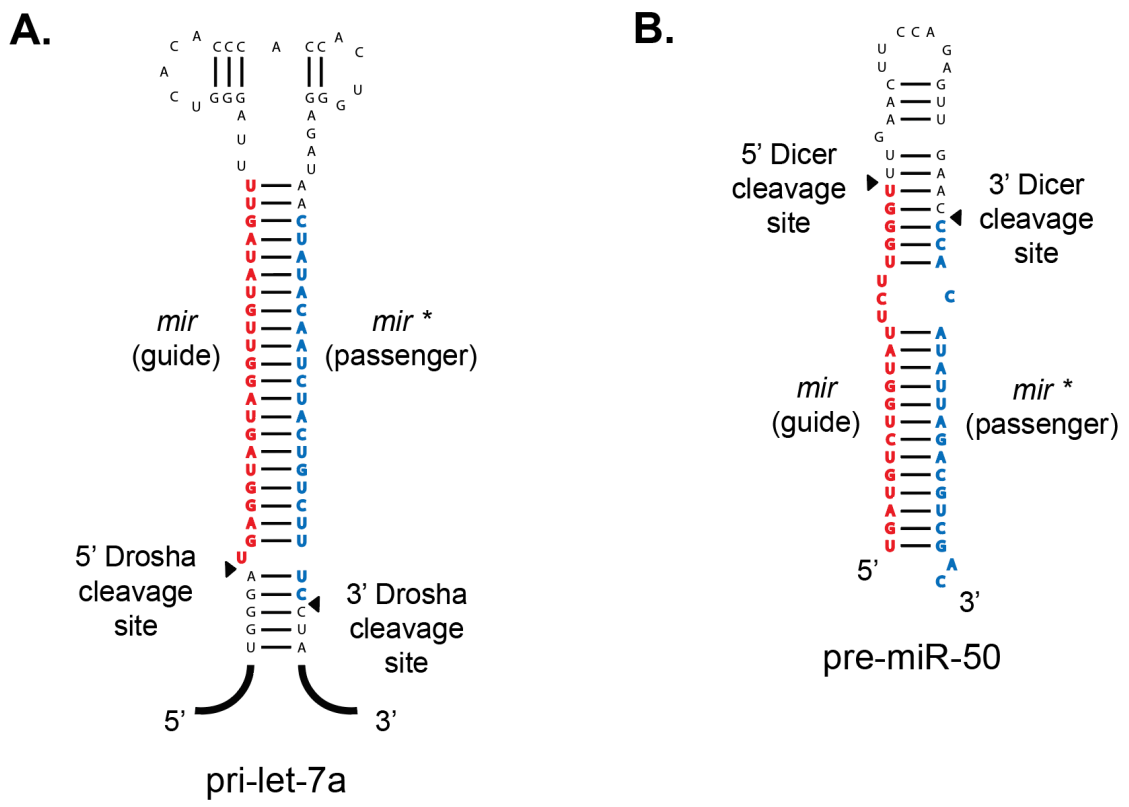


Figure 1.2 Examples of secondary structure and features of primary miRNA transcript (pri-miRNAs) and miRNA precursor (pre-miRNA).

(A) pri-miRNAs are characterized by a long stem region (containing miRNA duplex) with imperfect base pairing and bulges. Depending on the length and sequences, 5' and 3' tails of the pri-miRNA may fold into complex secondary structures. Flanking the stem is the basal region of ~11nt and a apical loop of >10 nt. Drosha cleaves the 5' strand of the stem ~22nt from the apical loop. The 3' strand is cleaved ~11nt from the base of the basal loop resulting in the 3' 2nt overhang characteristic of precursor miRNAs (pre-miRNAs). pri-let-7a structure adapted from Michlewsi and Cacerres 2010. **(B)** The cleavage of pre-miRNAs is completed by Dicer, which cleaves the 5' strand at the base of the loop and the 3' strand 2nt down. This asymmetrical cleavage leaves a miRNA duplex with an offset 2nt overhang, containing both the guide and passenger miRNA strands.

Export from nucleus

Consistent with the co-transcriptional model of Drosha processing, the pre-miRNA product of Drosha is initially localized to the nucleus (reviewed in O'Brien *et al.* 2018) The precursor miRNA is then actively exported into the cytoplasm (Melo *et al.* 2010; Kim *et al.* 2016). To actively cross the nuclear membrane, the pre-miRNAs utilize the Exportin-5 protein (Lund *et al.* 2004). The Exportin-precursor complex is selective as Exportin-5 shows poor affinity for other unrelated species of small RNAs (Lund *et al.* 2004). Consistent with active transport, the presence of Ran-GTP greatly increases efficiency of transport, although a subset of pre-miRNAs interact with Exportin 5 independently of Ran-GTP (Wang *et al.* 2020). Knockdown or mutation of Exportin-5 results in retention of pre-miRNAs in the nucleus and general depletion of downstream products, consistent with the model that it is the primary mechanism of precursor export (Lund *et al.* 2004; Melo *et al.* 2010). However, recent work showed that a subset of human pre-miRNAs utilize Exportin 1, using their 5' trimethylguanosine (TMG) modified cap instead of the typical 5' monophosphate of most precursors (Kim *et al.* 2016). Interestingly, there is increasing evidence that Exportin-5 can function outside its canonical role by interacting with both pri-miRNAs and the Microprocessor complex and stimulating processing of primary transcripts (Wang *et al.* 2020).

Dicer Processing

Once exported from the nucleus, the pre-miRNA associates with the protein Dicer (DCR-1 in *C. elegans*). As a member of the endonuclease RNAase III family, it targets miRNA precursors by binding the characteristic stemloop structure and cleaving out the ~22nt dsRNA miRNA duplex (Song and Rossi 2017). Specifically, Dicer binds both the 3' and 5' end of the

~70 nt long stemloop structure, cleaving off the terminal loop to liberate the miRNA duplex containing the characteristic 2nt 3' overhang (Youngman and Claycomb 2014; Ciechanowska *et al.* 2021; Figure 2B). In determining the cut site, Dicer utilizes a measuring mechanism from both the 3' overhang and 5' end of the precursor, and cuts at a predetermined distance from those points, producing a product that is ~22 nt length (Rybak-Wolf *et al.* 2014; Figure 1.2B). Canonically, precise cutting by Dicer is critical to the production of templated miRNAs, as a single nucleotide shift can cause changes to miRNA seed sequences and the generation of isomiRs, leading to changes in target gene expression (Gu *et al.* 2012; Starega-Roslan *et al.* 2015). However, several recent studies have shown a surprising amount of variability in the Dicer cut sites, suggesting non-canonical Dicer activity is indeed one of several points of isomiR production (Gu *et al.* 2012; Starega-Roslan *et al.* 2015). In addition to its roles in miRNA biogenesis, Dicer isoforms such as the sDCR-1 (short DCR-1) and t-DCR-1 (truncated DCR-1) observed in *C. elegans* can take on other non-canonical and antagonistic roles (Sawh and Duchaine 2013; Ge *et al.* 2014). Lastly, differential expression patterns or active relocation of Dicer may direct activity towards a non-enzymatic function, where Dicer binding of dsRNA essentially sequesters away both enzyme and RNA substrate from other active processes (Song and Rossi 2017; Tong *et al.* 2020).

miRISC and miRNA-mediated target repression

To repress gene expression, miRNAs require the action of effector proteins (reviewed in Ha and Kim 2014). Once the dsRNA miRNA duplex is fully processed by Dicer, it associates with an Argonaute protein (Yoda *et al.* 2010; Figure 1.1). In humans, all four Argonautes (AGO1- AGO4) associate with miRNAs (reviewed in Meister 2013), while *C. elegans* has three

Argonautes (ALG-1, ALG-2, and ALG-5), with ALG-1 and ALG-2 being the two AGO proteins that primarily load miRNAs (Tops *et al.* 2006; Vasquez-Rifo *et al.* 2012; Brown *et al.* 2017). The loading of Argonaute with the miRNA duplex is aided by the action of chaperone proteins such as HSP90 in humans that serve to keep Argonaute in an "open" conformation, thus facilitating binding of the duplex (Iwasaki *et al.* 2010). Subsequent unwinding of the duplex takes place passively, though the process is facilitated by the characteristic structure of the RNA duplex itself. Overall, the thermodynamically unstable ends of the duplex, as well as mismatches present throughout its length favor dissociation of the guide and passenger strand (Kawamata *et al.* 2009). The 5' end of miRNA guide is then tightly bound by the 5' binding pocket of the PIWI domain (Elkayam *et al.* 2012; Schirle and MacRae 2012). Likewise, the 3' end of the guide is anchored in a defined binding pocket within the PAZ domain (Elkayam *et al.* 2012; Schirle and MacRae 2012). Determination of the guide strand relies in part on the relative thermodynamic stability of the duplex ends, where the less stable end generally becomes the 5’ end of the guide (Hu *et al.* 2009). A 5' uracil (U) is also favored for Argonaute loading, with the majority of mature miRNAs beginning with a 5’ U (Frank *et al.* 2010). Together, these two rules partially explain how Argonaute proteins “select” the mature miRNA to be loaded, and therefore dictate the downstream targets. However, the 5’ nucleotide preference and duplex end thermodynamics do not explain strand choice for all miRNAs, suggesting additional rules may apply to miRNA strand selection (Medley *et al.* 2021). Once a strand selection has been made, the passenger strand is ejected from Argonaute in an ATP-independent manner (Kawamata *et al.* 2009; Yoda *et al.* 2010). The association of the guide miRNA and Argonaute, along with the effector protein GW182 (AIN-1/AIN-2 in *C. elegans*) makes up miRISC (Kawamata and Tomari 2010). miRISC is then directed by the guide miRNA to bind a partially complementary sequence in the 3'UTR of

its target mRNA (reviewed in Bartel 2018). The recognition of the appropriate target is determined by the seed sequence of the guide (nucleotides 2-8), which most of the time will have full complementarity to the target UTR (reviewed in Bartel 2009). Target specificity and strength of binding can be further titrated by differential levels of base pairing between the 3' nucleotides of the miRNA and the UTR (Grimson *et al.* 2007). Once on the target mRNA, GW182 recruits proteins involved in de-adenylation and decapping of mRNA targets (Behm-Ansmant *et al.* 2006), triggering translational repression of the target transcript and mRNA degradation (Ipsaro and Joshua-Tor 2015).

RNA-binding proteins impact miRNA-mediated regulation of gene expression

Overview

RNA binding proteins (RBPs) form another large and diverse class of post-transcriptional gene regulators (Baltz *et al.* 2012; Castello *et al.* 2012). RBPs bind diverse sets of RNAs, including mRNAs, small RNAs (including miRNAs), long noncoding RNAs (lncRNAs), transfer RNAs (tRNAs), and dsRNA species (reviewed in Hentze *et al.* 2018). To facilitate binding of a variety of RNA molecules with diverse sequence and secondary structures, RBPs have evolved a number of RNA-binding domains that allow for diverse forms of protein-RNA interaction. Canonical RNA binding domains (RBDs) include RRM (RNA recognition motifs), CCHC Zinc Finger domains, cold shock domain (CSD), K-homology domains (KH), DEAD helicase, and dsRBM (dsRNA binding motif) domains (reviewed in Hentze *et al.* 2018). Interestingly, non-canonical RBPs that lack canonical RBDs and instead interact with RNAs via nonconventional RBDs have been identified (Castello *et al.* 2012). These non-canonical RBPs include metabolic enzymes, heat shock proteins, and proteins with non-canonical RNA binding domains, including

ERM, PDZ, and tubulin binding domains (Castello *et al*. 2016, reviewed in Moore *et al.* 2018). It is thought that most RBPs have some level of specificity for their target RNAs, directed through distinct RNA binding domains and impacted by post-translational modification (Jankowsky and Harris 2015). Furthermore, the presence of accessory domains within RBPs often play an essential role in their mechanism of action (Jankowsky and Harris 2015). Accessory domains can include anything from catalytic domains to domains that scaffold protein-protein interactions. Regardless of their architecture and mechanism of actions, RBP activity is essential for gene regulation and RBPs extensively interface with miRNAs for ultimate combinatorial gene expression outcomes (reviewed in Bartel 2018; Figure 1.3). The functional interactions between RBPs and miRNAs are diverse, and primarily occur at three levels of miRNA biogenesis and activity; pri-miRNA regulation, pre-miRNA regulation, and mature miRNA regulation.

The regulatory mechanisms that act on pri-miRNAs are quite diverse. RBPs have been implicated in transcriptional control of *drsh-1* expression, thus effecting all downstream production of miRNAs (Kim *et al*. 2014). Spliceosome components interact extensively with pri-miRNAs, not only directing the splicing of pri-miRNA sequences and the production of mirtrons (Westholm and Lai 2011; Mondol *et al.* 2015), but also modifying the 3'UTRs of the target mRNAs to alter miRNA-mRNA interactions (Blazie *et al.* 2017; Figure 1.3A). RBPs at this level can also interact with pri-miRNAs to regulate their processing by Drosha/DGCR8, either by RNA editing (Kawahara and Mieda-Sato 2012), methylation (Alarcón *et al.* 2015), or by modulating pri-miRNA-Microprocessor interactions (Michlewski *et al.* 2008; Kawahara and Mieda-Sato 2012; Jiang *et al.* 2017). At the level of pre-miRNA, the mechanisms of regulation are very similar to those of pri-miRNAs (Figure 1.3B). Various RBPs help to regulate precursor processing by modulating the binding/effects of Dicer (Wu *et al.* 2010; Bicker *et al.* 2013).

Furthermore, RNA editing and methylation are also carried out on precursor molecules, altering their processing rates (Heo *et al.* 2009, 2012; Ji and Chen 2012; Pandolfini *et al.* 2019; Figure 1.3B).

Lastly, a great deal of RBP-miRNA interactions occurs at the level of the mature miRNA (Figure 1.3C). RBPs can determine the subcellular localization of mature miRNAs, and directly modulate their stabilization or enzymatic degradation (Chatterjee and Großhans 2009; Chatterjee *et al.* 2011; Chen *et al.* 2012). Several RBPs also modulate mature miRNA stability and functional interactions by altering their sequence composition through RNA tailing (Katoh *et al.* 2009; Yang *et al.* 2020) or by addition of methyl groups (Cheray *et al.* 2020; Liang *et al.* 2020). In addition to action of individual RBPs, this level of regulation also relies heavily on the combinatorial action of many RBPs, generally as part of a protein complex (Figure 1.3C). The major effector complex of the miRNA pathway, miRISC, interacts with RBPs that either promote translational repression of target mRNAs, or directly degrade mRNAs through their enzymatic domains (Jafarnejad *et al.* 2018; Duchaine and Fabian 2019; Mayya *et al.* 2021). In addition to miRISC, other RBP complexes can target the 3'UTR of target transcripts and modulate the binding of miRISC to miRNA target sites and thus alter miRNA targeting ability (Kim *et al*. 2021; Figure 1.3C).

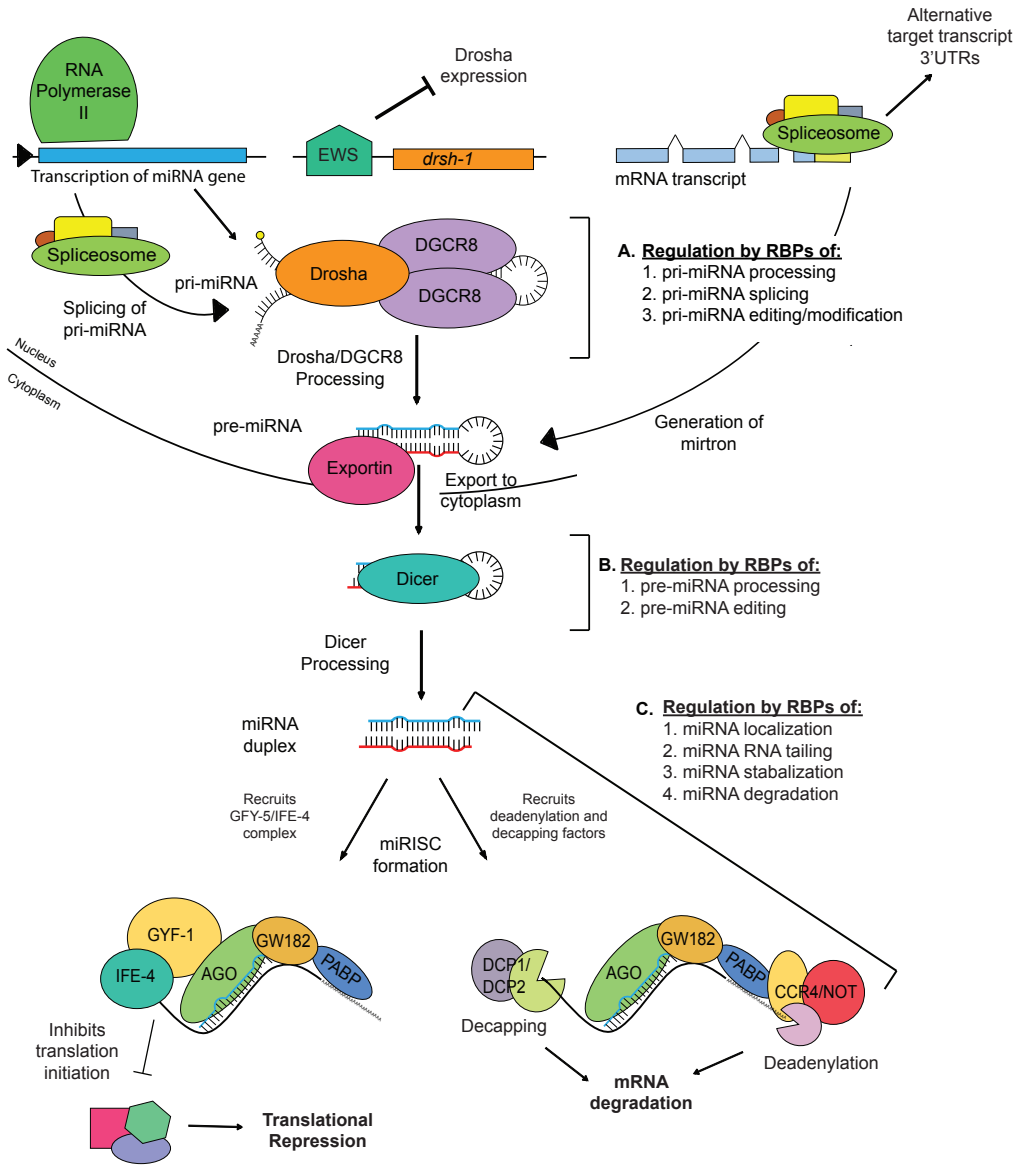


Figure 1.3 Coordination of RBPs with miRNA pathway to modulate gene expression. RBPs interact with the miRNA pathway at every level to modulate miRNA-mediated regulation of gene expression. **(A)** RBPs modulate expression of Drosha, splicing of pri-miRNAs and mirtrons, and splicing of target 3'UTRs. pri-miRNAs-Drosha interactions are modulated by RBPs to influence Drosha processing. **(B)** RBPs modulate interaction between pre-miRNA and Dicer to influence their processing. **(C)** RBPs modulate miRNA stability and decay directly. They also effect stability through RNA tailing of mature miRNA sequences.

RBPs modulate activity and expression of miRNA biogenesis machinery

Given the complexity of the miRNA biogenesis pathway, there are multiple points where miRNA production or activity can be regulated by RNA binding proteins. The enzymatic processing steps carried out by Drosha and Dicer represent critical points of regulation and thus are the targets of multiple regulatory mechanisms. Overall, the careful modulation of expression and activity of the biogenesis machinery imparts both a level of specificity and potency to the interactions between the enzymes and their respective substrates (Figure 1.4).

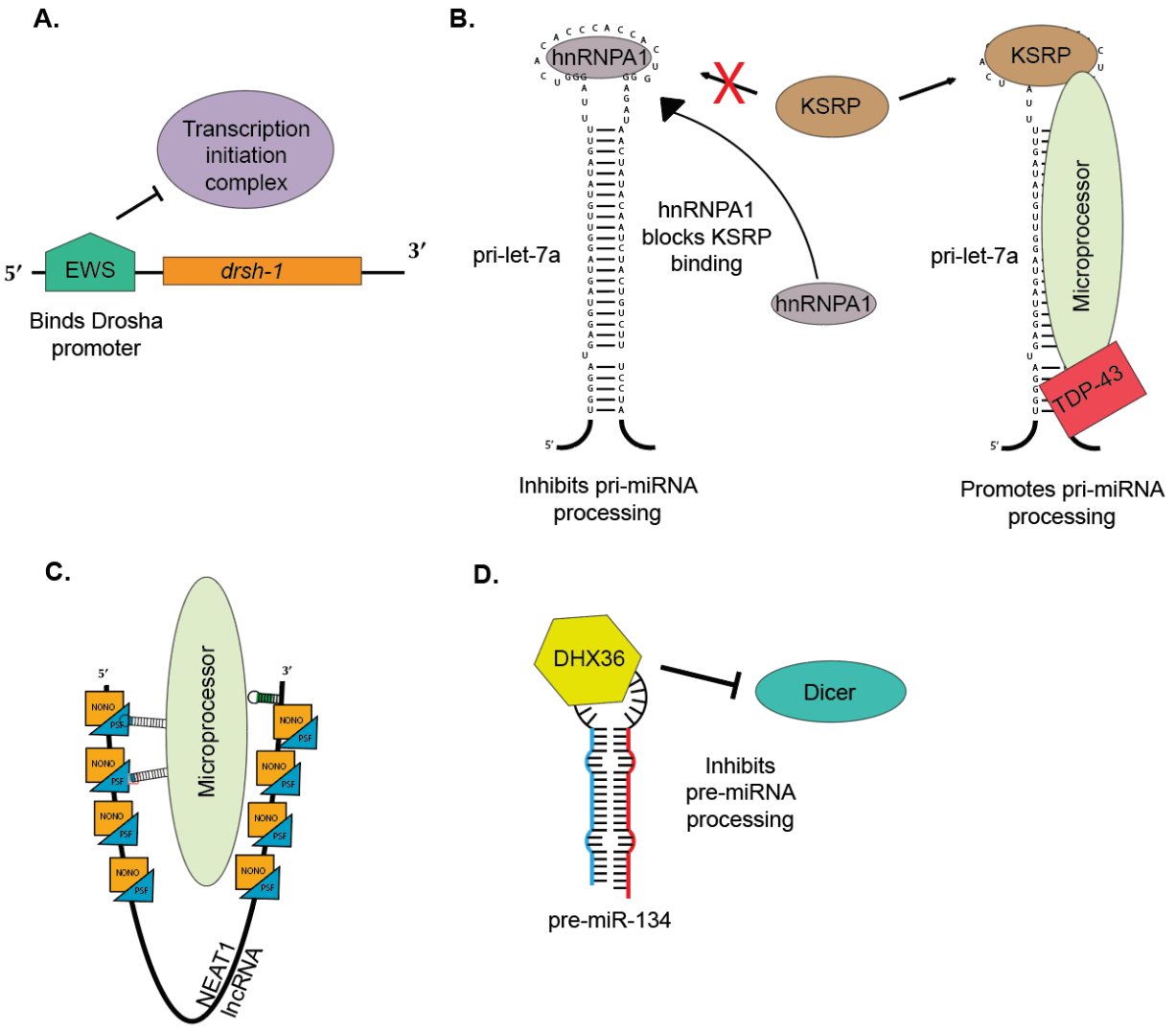


Figure 1.4. RBPs modulate the activity and expression of miRNA biogenesis machinery. **(A)** EWS competitively binds the Drosha promoter, inhibiting the binding of the transcription initiation complex, thus downregulating Drosha mRNA and protein. (**B**) hnRNPA1 negatively regulates pri-miRNA processing by competitively binding the stem-loop of pri-miRNAs to antagonize the binding of KSRP. The conformation of pri-*let-7* is shown pre and post hnRNPA1 binding to highlight the opening of the stem loop structure. When hnRNPA1 is absent, KSRP binds and promotes the processing of pri-miRNAs. Additional factors like TDP-43 can bind Microprocessor to further promote its processing of pri-miRNAs. Model

adapted from Michlewski and Caceres 2011. **(C**) NEAT1 lncRNA cooperatively binds pri-miRNAs with the NONO/PSF complex to recruit the Microprocessor and promote pri-miRNA processing. Model adapted from Jiang *et al*. 2017. **(D)** DHX36 competitively binds the stem-loop of pre-miRNAs to block Dicer binding, thus inhibiting processing of pre-miRNAs.

Like most enzymes, the miRNA biogenesis machinery is subject to transcriptional and translational regulation, whereas cells modulate the expression of the enzymes themselves to modulate the reactions they catalyze. For example, Drosha expression is transcriptionally regulated by the multifaceted regulatory protein EWS, which modulates the transcription of many genes (Kovar 2011). Recent work has shown that EWS specifically regulates the abundance of Drosha, but not other biogenesis machinery like DGCR8 or Dicer (Kim *et al.* 2014; Figure 1.4A). Promoter assays have shown that EWS potently represses the transcription of Drosha in a dosage dependent manner, presumably by occupying the promoter and preventing transcription initiation. Interestingly, EWS also appears to downregulate Drosha at the protein level. Together, the EWS-directed downregulation of both transcript and protein impairs the active processing of pri-miRNAs by the Microprocessor complex (Kim *et al*. 2014; Figure 1.4A), although seemingly conflicting data also suggests that EWS may actually promote Drosha co-transcriptional processing of some pri-miRNAs (Ouyang *et al.* 2017).

The expression of Dicer is also a major point of regulation in the miRNA biogenesis pathway. Dicer expression is partially controlled by AUF1, a protein that modulates gene expression by associating with 3'UTRs of various mRNAs to promote/repress their translation or modulate their decay (Pullmann *et al.* 2007). When AUF1 was shown to interact with Dicer mRNA, it was hypothesized it may represent a far-upstream regulator of miRNA-mediated regulation of gene expression (Mazan-Mamczarz *et al.* 2009). Indeed, AUF-1 shows a strong interaction with Dicer mRNA, suggesting it may regulate the expression of the transcript (Abdelmohsen *et al.* 2012). Knockdown and stability assays support the hypothesis that AUF1 represses Dicer expression by destabilizing Dicer mRNA and targeting it for decay. Furthermore, the repression of Dicer by AUF1 has been shown to have functional consequences, mainly the

general reduction of mature miRNAs (Abdelmohsen *et al*. 2012). Taken together, this data suggests that AUF1 indirectly modulates the Dicer processing step of miRNA biogenesis by regulating the levels of Dicer present in cells.

A number of RBPs directly interact with Drosha and Dicer to modulate their enzymatic activity. The antagonistic roles of hnRNPA1 and KSRP highlight the modulation of Drosha activity by RBPs (Figure 1.4B). hnRNPA1 is a nuclear-cytoplasmic shuttling protein that regulates alternative splicing of mammalian pre-mRNAs by constituently driving splicing by altering pre-mRNA secondary structure (Chen and Manley 2009). A similar role for hnRNPA1 has been hypothesized in the miRNA pathway, whereas the protein is thought to recognize and alter the secondary structure of small RNAs (Michlewski *et al*. 2008). Interestingly, in comparison to it purely promotional role in alternative splicing, hnRNPA1 appears to have roles in both promoting and repressing pri-miRNA processing by the Microprocessor complex (Figure 1.4B). In a promoting role, hnRNPA1 actively binds pri-miR-18 and enhances its processing by relaxing the stem of the hairpin, allowing better access for Drosha and facilitating enhanced binding (Michlewsi *et al*. 2008) (Figure 1.4B). In support of this model, a significant portion of human pri-miRNA have a well-conserved terminal loop (~14%) with the general hnRNPA1 binding consensus sequence of UAGGGA/U, suggesting that hnRNAPA1 may be involved in processing of a subset of pri-miRNAs. In contrast to this role, hnRNPA1 may also serve to oppose primary miRNA processing of certain pri-miRNAs (including *let-7*) by antagonizing the binding of pro-processing factors such as KSRP (Michlewski and Cáceres 2010; Figure 1.4B). In a seemingly conserved mechanism, hnRNPA1 can recognize pri-let-7a through a binding site in the terminal loop (perfect consensus) and inhibit its Drosha dependent processing in a dosage dependent manner (Michlewski & Caceres, 2010).

In addition to proteins that effect the secondary structure of pri-miRNAs, RBPs can also modulate the strength of the interaction between Drosha and pri-miRNA. The protein TDP-43 is another a well-known nuclear-cytoplasmic shuttle that effects RNA processing in both compartments (Ayala *et al*. 2008). TDP-43 has been observed to interact with the miRNA pathway in context of ALS, suggesting it may modulate the biogenesis of miRNAs. Indeed, TDP-43 has been shown to interact with the Microprocessor (Fukuda *et al.* 2007; Kawahara and Mieda-Sato 2012; Figure 1.4B) as well as Exportin 5, suggesting a role in pri-miRNA processing in the nucleus (Kawahara & Mieda-Sato 2012). Binding assays show that the inclusion of TDP-43 within the Microprocessor complex significantly increases the affinity of Drosha for a highly specific subset of human pri-miRNAs (Kawahara & Mieda-Sato 2012; Figure 1.4B). Lack of TDP-43 not only decreases Drosha binding of these pri-miRNAs but impairs their processing, leading to reduced levels of pre-miRNAs and mature miRNAs (Kawahara & Mieda-Sato 2012).

In comparison to the relatively straightforward examples of RBP modulation previously discussed, recent work has identified several examples of more complex RBP-directed mechanisms of regulation, involving not only the interaction between the biogenesis machinery and miRNA intermediates, but also additional RNA substrates. One example is the interplay between lncRNAs and RBPs to modulate the activity of Drosha in the processing of pri-miRNAs in human cells (Jiang *et al.* 2017; Figure 1.4C). ncRNAs like MALT1 and NEAT1 have been well studied in the context of their localization to nuclear subdomains (Hutchinson *et al.* 2007; Wilusz *et al.* 2008), areas that have been long recognized for the storage of repressed mRNAs and factors involved in RNA processing and metabolism (Fox *et al.* 2002). In addition to these constitutive components, several other diverse RBPs, like NONO and PSF are also sequestered to these RNP complexes suggesting a more complex role involving interaction between lncRNA

and RBP. Indeed, the RBPs NONO and PSF physically interact 2/3 of pri-miRNAs expressed in HeLa cells and are necessary for their efficient processing (Jiang *et al*. 2017), an interaction relies on the lncRNA NEAT1 (Figure 1.4C). Depletion of either NONO/PSF or NEAT1 lncRNA is sufficient to deplete mature miRNAs, suggesting that both function to promote pri-miRNA processing. CLIP-seq data of NONO/PSF shows direct binding of many pri-miRNAS as well as the NEAT1 lncRNA, with enriched binding on both the 5' and 3' ends, with co-IP confirming interactions between NONO-PSF, Microprocessor, and NEAT1 lncRNA (Jiang *et al.* 2017; Figure 1.4C). In further support a role in pri-miRNA processing, NEAT1 contains a number of hairpin-like structures that appear to increase efficiency of Microprocessor binding. Overall, a recent model proposes that NONO/PSF bind the 5' and 3' ends of the NEAT1 lncRNA forming a bird-nest structure, with the NEAT1 lncRNA providing a scaffold to simultaneously recruit and bind the Microprocessor (Jiang *et al.* 2017; Figure 1.4C). Together NONO/PSF and NEAT lncRNA function together to stimulate efficient processing of pri-miRNAs by Microprocessor, thus promoting their biogenesis.

Lastly, the DEAH-box helicase DHX36 has been shown to modulate the interactions between Dicer and pre-miRNAs (Höck *et al.* 2007; Bicker *et al.* 2013). DHX36 was identified by pre-miRNA pulldowns as a factor that regulates precursor processing of pre-miR-134 in the dendrites of neurons (Bicker *et al*. 2013). *In vitro* cleavage assays demonstrated that pre-miR-134 bound by DHX36 has strongly inhibited Dicer-dependent processing, while the processing of other pre-miRNAs was not affected, suggesting a specificity for pre-miR-134. Furthermore, DHX36 directs the pre-miR134 localization to the dendrite of neurons. The impact of DHX36-mediated inhibition of miR-134 was confirmed by the significant de-repression of a dendritically-localized miR-134 reporter upon knockdown of DXH36 (Bicker *et al.* 2013; Figure

1.4D). It is proposed that DXH36 binds and sequesters pre-miR-134 to the dendrites providing a localized pool of precursor that can be rapidly mobilized, although the target and mechanism of localized miR-134 expression remain unknown (Bicker *et al.* 2013; Figure 1.4D).

Impact of splicing-related factors in miRNA biogenesis and activity

Though it is widely understood that the majority of miRNAs are produced through a canonical biogenesis pathway, there is increasing evidence that splicing factors can directly influence miRNA biogenesis and activity. Direct splicing of pri-miRNAs has been observed to produce alternative pri-miRNA isoforms that then undergo normal processing (Mondol *et al*. 2015). Considering the size of the splicing-related RBP repertoire it is perhaps not surprising that these factors appear to interact quite extensively with miRNAs to direct their biogenesis as well as modulate their downstream activity, specifically by altering target 3'UTRs (Blazie *et al.* 2017; Figure 1.5).

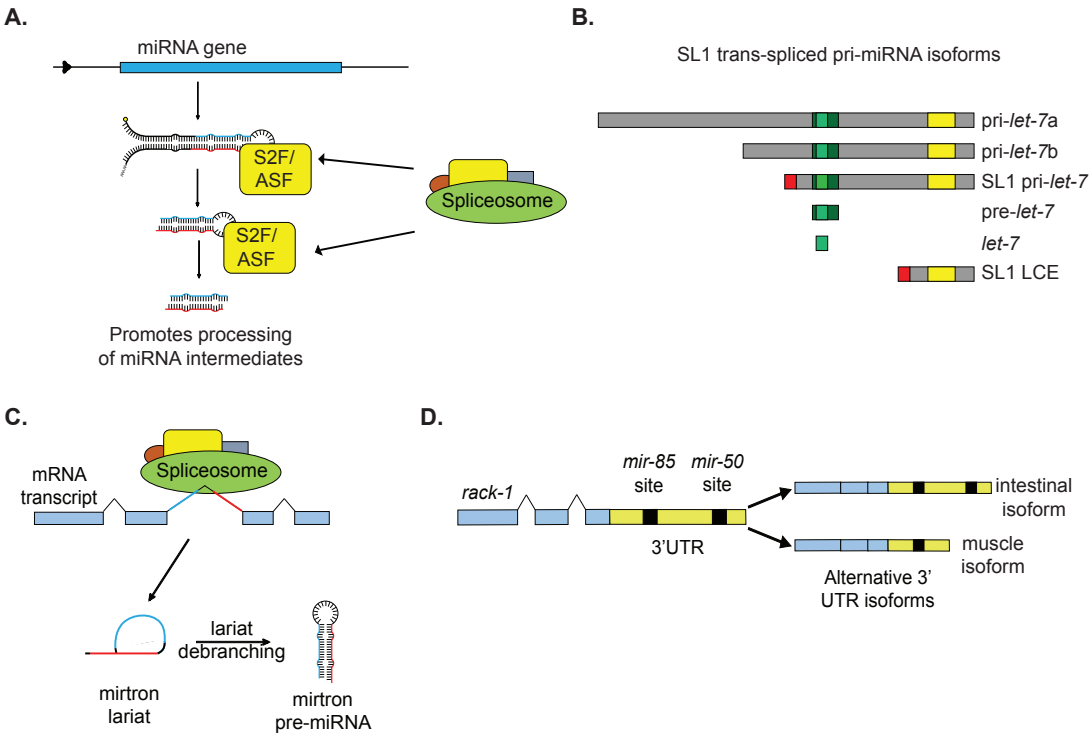


Figure 1.5. Impact of splicing-related factors on miRNA biogenesis.

(A) Subunits of the canonical splicing machinery like S2F/ASF can take on non-canonical roles by binding pri-miRNAs and pre-miRNAs to promote their processing by Drosha and Dicer, respectively. **(B)** pri-let-7 undergoes splicing to produce several splice variants. The canonical pri-let-7a and pri-let-7b represent the two full length (non-spliced) variants, while the SL1 pri-let-7 represents an abundant variant produced by SL1-mediated trans splicing. A much smaller and less abundant splice variant (SL1 LCE) is produced by extensive splicing of the pri-miRNA leaving only the LCE region of the transcript. The SL1 LCE is believed to antagonize both let-7 biogenesis and activity, thus forming a regulatory loop with the canonical and larger spiced let-7 pri-miRNA variants. Model adapted from Nelson and Ambros 2019. **(C)** mirtrons are miRNA precursors that are generated by the splicing of mRNA introns. mirtrons resemble pre-miRNAs in structure and therefore bypass Drosha processing and proceed directly to processing by Dicer. Model adapted from Westholm and Lai 2011.

(D) Alternative polyadenylation of the 3'UTRs of miRNA-targeted transcripts allows tissue specific regulation of gene expression. The 3'UTR of *rack-1* mRNA contains both miR-50 and miR-85 targets sites when expressed in intestinal tissue. However, due to alternative polyadenylation, the miR-85 site is removed from the *rack-1* isoform in muscle tissue thus liberating the transcript of miR-85-mediated regulation in that specific tissue in *C. elegans*. Model adapted from Blazie *et al*. 2017.

S2F/ASF directly binds pri-miRNA stem loop and promotes it processing by Drosha, likely by opening up the secondary structure to allow for more efficient Drosha cleavage (Figure 1.5A). Interestingly, S2F/ASF also appear to promote the processing of select precursor miRNAs such as pre-miR-29b-1, indicating a possible post-Drosha role in miRNA biogenesis (Wu *et al.* 2010)."Canonical" splicing of pri-miRNAs has been hypothesized for some time and highlighted by the discovery of the splicing of *C. elegans* pri-let-7 (Mondol *et al.* 2015). In *C. elegans*, the pri-let-7 undergoes several differential SL1 *trans*-splicing (Graber *et al.* 2007; Figure 1.5B). The intact pri-let-7 is comprised of a relatively long transcript with the precursor embedded near the middle, and a second region containing 7 complementary *let-7* family binding sites (LCE) at the 3' end of the transcript (Zisoulis *et al.* 2012). The larger of the two splice isoforms includes the full 3' end of transcript but removes the sequence directly upstream of the precursor. The second splice isoform (SL1-LCE) encompasses only the LCE and adjacent sequences (Nelson and Ambros 2019). Interestingly, this splice isoform seems to be post-transcriptionally regulated by an RNA binding protein, LIN-28, independently of the other isoform. The *let-7* family sites embedded within SL1-LCE suggests that it may function as a *let-7* family sponge, rescuing LIN-28 from normal *let-7*-mediated downregulation and thus allowing it to negatively regulate *let-7* biogenesis (Nelson & Ambros 2019). Although LIN-28 is clearly involved in the SL1-LCE feedback loop, it is unclear how LIN-28 is promoting SL1-LCE generation.

The second mechanism of miRNA splicing is carried out through an intron splicing mechanism, where miRNA intermediates called mirtrons, are processed directly from intron lariats of mRNAs and resemble precursor miRNAs in both length and secondary structure (Figure 1.5C). By mimicking precursor miRNA, mirtrons bypass Drosha processing entirely and proceeds directly to nuclear export and Dicer processing (reviewed in Westholm and Lai 2011).

Like all intronic splice products, mirtrons contain the canonical splice donor and acceptor sequences at the 5' and 3' termini, respectively, allowing for recognition and processing by the canonical spliceosome. Furthermore, spliced mirtrons take on the typical lariat structure where the 3' branchpoint is ligated to the 5' of the introns (reviewed Westholm and Lai 2011; Figure 1.5C). Mirtron lariats are debranched by a lariat debranching enzyme before taking on their pre-miRNA secondary structure (Okamura *et al.* 2007; Ruby *et al.* 2007; Figure 1.5C). Interestingly, some mirtrons show tailing at the 5 or 3' end, where an unstructured tail extends out from the pre-miRNA hairpin (Babiarz *et al.* 2011). The 3' tails are degraded after debranching by the RNA exosome in Drosophila (Flynt *et al.* 2010). In comparison, the mechanism of degradation is less well understood in *C. elegans*, though it could presumably be carried out the by the XRN1/2 complex responsible for degradation of small RNAs (reviewed in Westholm and Lai 2011).

Splicing can also have an impact on the downstream activity of miRNAs by altering miRNA target sites. In many organisms, alternative polyadenylation enables the expression of multiple isoforms of 3'UTR for a single gene by utilizing different polyadenylation signals within the genomic sequence (Mangone *et al.* 2010; Gupta *et al.* 2014). The differences in 3'UTR length between different isoforms can lead to differential miRNA-mediated regulation of a particular isoform by the loss of gain of miRNA target sites (Blazie *et al.* 2017; Figure 1.5D). In *C. elegans*, alternative polyadenylation is utilized by two well know disease orthologs *rack-1* and *tct-1* to generate 3'UTR isoforms in muscle body wall that lack miR-50 sites, and thus escape regulation by this otherwise ubiquitously acting miRNA (Blazie *et al*. 2017; Figure 1.5D). Interestingly, no specific polyadenylation factors have been directly implicated in the modulation of miRNA "dodgers" like *rack-1* and *tct-1*. Therefore, it is possible that the mechanism to bypass miRNA-mediated regulation may be directed by canonical polyadenylation factors that have

been co-opted for a miRNA-specific function. It is also possible that additional or unique factors are required for polyadenylation events that allow mRNA targets to escape miRNA regulation.

RBPs modulate stability of miRNAs and their intermediates

In addition to affecting the biogenesis and processing of miRNAs intermediates, RPBs also play a part in maintaining the stability and functionality of miRNAs over time (Figure 1.6). Several RBPs are responsible for directly stabilizing mature miRNAs, yet others can also influence the stability of their upstream intermediates. In addition, significant effort has been made to understand how RBPs are able to modify RNAs to modulate their processing and activity. Through RNA editing and nucleotide tailing, RBPs can either impede or promote interactions with biogenesis factors and other downstream effector proteins, ultimately impacting miRNA stability and/or turnover (Figure 1.6).

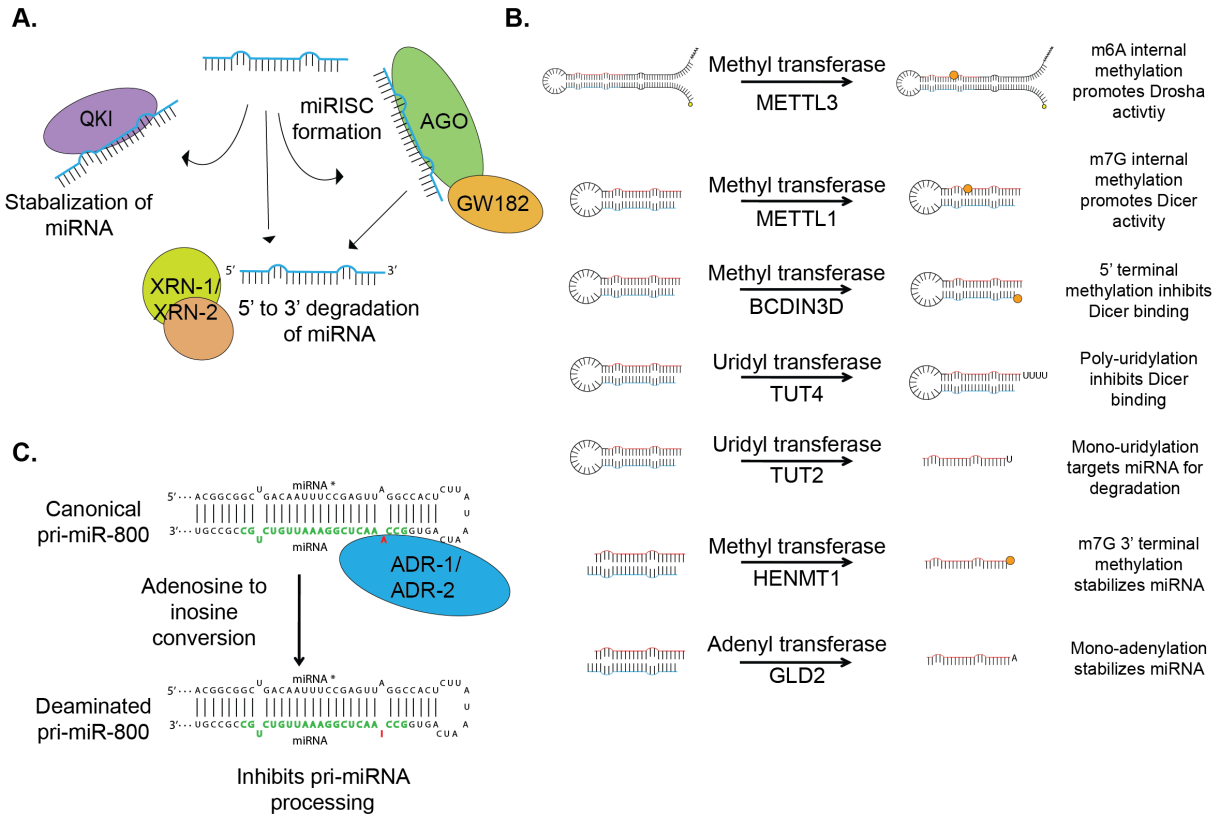


Figure 1.6. RBPs modulate stability of mature miRNAs and intermediates.

(A) The stability of mature miRNAs is heavily regulated in the cell. RBPs like Quaking (QKI) directly bind mature miRNAs and increase their stability. In comparison, miRNAs that are targeted for degradation are bound by exonucleases like XRN-1/XRN-2 which degrade the miRNA in a 5' to 3' orientation. **(B)** miRNA and their intermediates undergo chemical modifications like uridylation, methylation, and adenylation which can affect their stability and impact processing by Drosha and Dicer. Uridylation occurs on both pre-miRNA (polyuridylation) and mature miRNAs (monouridylaton) and inhibits Dicer processing and targets miRNAs for degradation, respectively. Adenylation of mature miRNA is accomplished by GLD-2, stabilizing the miRNA. Methylation of miRNAs and intermediates is the most varied of the chemical modifications effecting all levels of biogenesis. **(C)** Adenosine to

inosine conversion in pri-miRNAs inhibits their processing by Drosha and can change miRNA targeting by altering seed sequences. Model adapted from Warf *et al.* 2012.

One well studied example of an RBP stabilizing miRNAs is the protein Quaking (QKI), which has been broadly implicated in a variety of pathological states including neural pathologies and cancer (reviewed in Chenard and Richard 2008). QKI has been shown to bind and stabilize the mature miR-20 miRNA without effecting the levels of primary or precursor miRNA (Chen *et al.* 2012). QKI bound-miR-20a has a significantly longer half-life and thus maintains the miRNA directed inhibition of miR-20 target TGFbeta-R2, an oncogene that promotes glioblastomas (Chen *et al.* 2012; Figure 1.6A). In contrast to stabilization of miRNAs, cells also utilize exonuclease-based degradation pathways to balance mature miRNA populations (reviewed in Zhang et al., 2012). In *C. elegans*, two exonucleases XRN-1/XRN-2 perform the exonucleolytic degradation of mature miRNAs as well as other species of small RNAs in a 5' to 3' fashion (Chatterjee and Großhans 2009; Chatterjee *et al.* 2011; Figure 1.6A). In *C. elegans* there is compelling evidence that release of miRNAs bound by Argonaute is in part determined by the presence of an appropriate target; whereas the presence of a miRNA stabilizes the Argonaute::miRNA complex (Sanei and Chen 2015). This "protection" even goes so far as to prevent the degradation of mir* strands that would otherwise not be part of an active miRISC complex.

In addition to RBPs responsible for stabilizing and degrading miRNAs, several RBPs have been implicated in the modification of miRNAs and their precursors. In general, these RBPs come in three flavors, methylases, adenylases, and uridylases, adding methyl groups and untemplated adenine and uridine residues, respectively (Figure 1.6B). Observation of the 3' ends of miRNAs noted increased heterogeneity compared to 5' ends, often because of 1-2 nt untemplated additions. The majority of these non-templated additions are terminal uridine or adenine residues. In animals, the primary mechanism of non-templated uridylation occurs

through a TUTase mediated pathway (Martin and Keller 2007; Figure 1.6B). In humans, the uridylyl transferase TUT4 is recruited by LIN28 to miRNA precursors, where it adds an oligo-U tail to the 3' end of the pre-miRNA. The addition of the oligo-U is sufficient to block processing by Dicer and facilitates the decay of the pre-miRNA (Heo *et al.* 2008, 2009; Figure 1.6B). In addition to the uridylation of pre-miRNAs, TUT4 and TUT7 have been shown to oligo-uridylate both 5p and 3p mature miRNAs at similar levels, suggesting tailing can occur post-Dicer processing (Yang *et al.* 2020; Figure 1.6B). Lastly, single uridylyl tailing of pre-miRNAs including pre-let-7 is directed by TUT-2, showing strong preference for Type II miRNAs (1 nt 3' overhang) and appears independent of LIN-28 (Heo *et al.* 2012; Figure 1.6B). Interestingly, TUT4 and TUT7, which had previously been implicated in the poly-uridylation of pre-miRNAs were also shown to mono-uridylate pre-let-7 in the absence of LIN-28. Taken together, the uridylation of miRNAs represents a complicated mechanism by which to regulate miRNA levels. Similar to uridylation, the terminal adenylation of miRNAs can also modulate their stability. In animals, canonical poly(A) polymerases act to regulate mRNA stability through the addition of poly-adenine tracts. In humans, poly(A) polymerase GLD-2 is well known for regulating the translation of mRNAs by the addition of poly(A) tracts, greatly increasing their processing by translational machinery (Nakanishi *et al.* 2006). GLD-2 was found to adenylate the 3' end of mature miR-122 in human liver cells and thus increase their stability (Katoh *et al.* 2009; Figure 1.6B). Further work implicated GLD-2 in the widespread adenylation of many miRNAs, a process that may affect miRNA activity in addition to stability by modulating miRNA targeting effectiveness (D'Ambrogio *et al.* 2012).

Like the untemplated adenylation and uridylation of miRNAs, methylation is also used to modulate the stability of miRNAs and their intermediates (Figure 1.6B). The 3' terminal

methylation of miRNAs is a strongly conserved mechanism in many species including plants and animals (Ji and Chen 2012). In support of a conserved mechanism, the small RNA methyltransferase responsible, HEN1, also appears to be broadly conserved from plant to animals, although its small RNA targets can be different depending on species. HEN1 homologs have been implicated in the 3' terminal methylation of small RNA species like miRNAs, siRNAs, piRNAs, snRNA, and tRNA. In plant and Drosophila, HEN1 and its homolog, respectively, has been implicated in the 2' O-methylation of miRNAs (Yu *et al.* 2005; Abe *et al.* 2014; Figure 1.6B). It has been suggested that one of the functional consequences of miRNA methylation is protection of miRNAs from RNA tailing leading to inhibition of the uridylation-dependent degradation pathway (Jones *et al.* 2009; Ibrahim *et al.* 2010). In support of this model the human homolog of HEN1, HENMT1, was recently shown to 2' O-methylate human miRNAs and promote their stability and function by inhibiting degradation by exonuclease PNPT1 (Laing *et al.* 2020). Conversely, 2'-O 3' methylation has also been reported to impact the binding of miRNAs by Argonaute (Abe *et al.* 2014) and inhibit miRNA/target binding (Cheray *et al.* 2020), potentially representing a mechanism to dampen miRNA-mediated repression of certain targets. In addition to the 3' terminal methylation, miRNA biogenesis can be affected by several other methylation events (Figure 1.6B). For example, human METTL1 has been shown to modulate the m7G methylation of internal guanosines in pre-let-7 to promote its processing by Dicer (Pandolfini *et al.* 2019). A second methyltransferase, BCDIN3D 2' O methylates the 5' end of pre-miR-145, inhibiting binding and processing by Dicer (Xhemalce *et al.* 2012). Lastly, human METTL3 directs the m6A methylation of pri-miRNAs, promoting their processing by Drosha (Alarcon *et al.* 2015). Overall, methylation appears to be an important mechanism to regulate

the biogenesis and activity of miRNAs and is quickly becoming an important biomarker of various cancers (Konno *et al.* 2019; Figure 1.6B).

Internal editing of miRNAs is another regulatory mechanism allowing for fine-tuned modulation of miRNA activity. RNA editing is accomplished by a family of proteins called ADARs that serve to bind dsRNA and deaminate adenosine into inosine (A to I conversion) (Zinshteyn and Nishikura 2009; Figure 1.6C). The conversion of adenosine to inosine can alter base pairing specificity in RNA structures, as inosine preferentially pairs with cytidine. These edits can occur on a number of RNA substrates. Best studied is the effect of deamination in pre-mRNAs where deamination has been shown to increase variability in RNA and proteins (Schaub and Keller 2002). It was hypothesized that as pri-miRNAs and pre-miRNAs both have extensive dsRNA structures that pri-miRNAs may also be targeted by ADAR-dependent deamination. Indeed, A-to-I editing has been observed in mammalian pri-miRNAs and pre-miRNAs and been shown to negatively regulate their processing by Drosha and Dicer, respectively (Kawahara *et al*. 2008; Figure 1.6C). Similarly, editing of human mature miRNAs leads to differential selection of target genes, resulting in altered regulatory networks (Blow et al., 2006, Kawahara et al., 2007). Consistent with a conserved role, two *C. elegans* ADAR proteins, ADR-1/ADR-2, appear to inhibit pri-miRNA processing leading to a general decrease in mature miRNA (Savva *et al.* 2012). In contrast, the editing of mature miRNAs appears exceptionally rare in *C. elegans* with only one miRNA known to be edited (Warf *et al.* 2012).

RBPs modulate the interaction between miRNAs and their targets

While mechanisms that regulate miRNA biogenesis and/or processing are no doubt important, they do not necessarily represent the most critical ways the miRNA pathway is

modulated. Potentially, the most important set of RBPs that regulate miRNA activity do so by influencing the association of miRNAs-miRISC with their target and modulating the repressive effects of the complex (Figure 1.7). RBPs are not only critical components of miRISC itself, but also are the primary effectors of translational repression and mRNA decay. By directing interactions between miRISC and the target transcript, these RBPs are effectively the terminal effectors of the miRNA pathway.

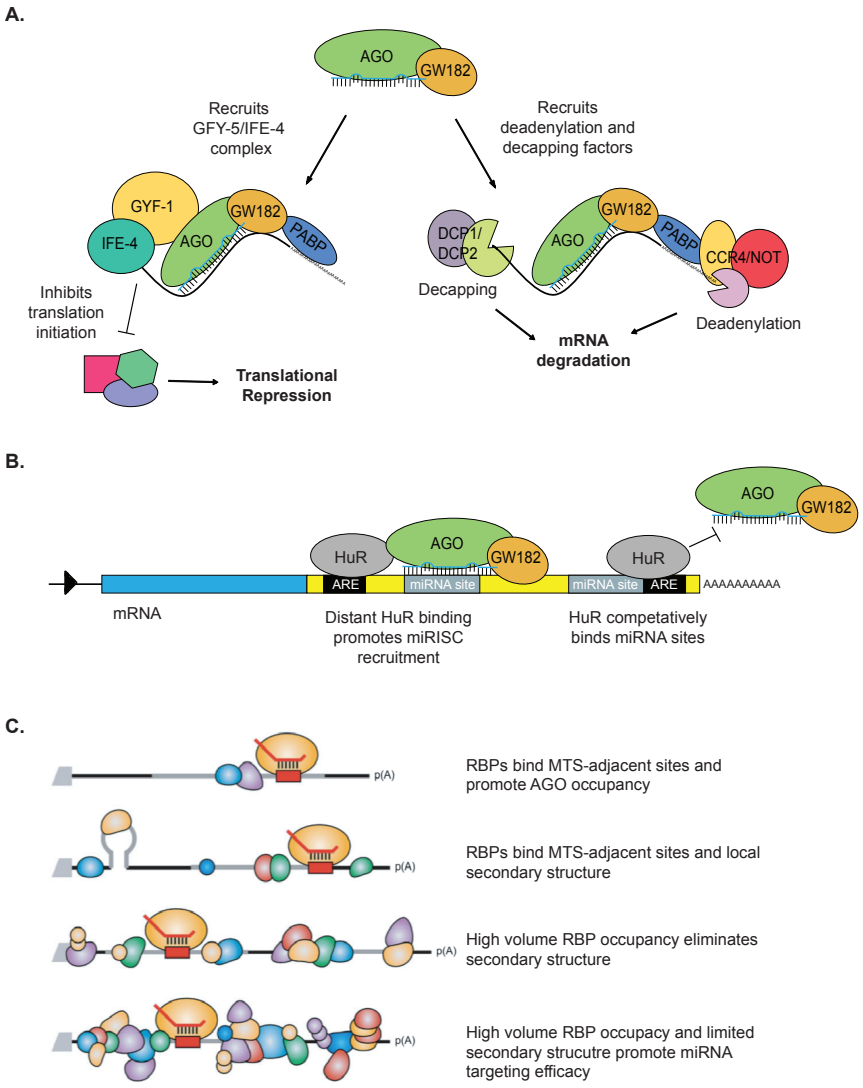


Figure 1.7. RBPs coordinate with miRNAs to modulate their interaction with mRNA targets.

(A) miRISC binding of target mRNA leads to the recruitment of accessory factors that determine the fate of the target transcript. Recruitment of GFY-1/IFE-4 complex leads to binding of the 5' cap region inhibiting the binding of factors responsible for translational initiation. Alternatively, miRISC can recruit the deadenylation and decapping complexes to target the transcript for degradation. **(B)** miRISC-target interactions can be mediated by factors that bind AREs in the 3' UTRs of target transcripts. HuR binding of AREs that lie distant from

miRNA binding sites can promote the recruitment and binding of miRISC to the target. Alternatively, HuR can bind AREs that lie near miRNA binding sites and therefore competitively inhibit miRISC binding. **(C)** RBPs bind sites adjacent to MTS (miRNA target sites) and increase Argonaute occupancy on MTS. Increased RBP binding is highly correlated with a decrease in local secondary structure of target 3'UTR, thus increasing miRISC accessibility to the MTS and increasing the overall efficacy of miRNA targeting. Panel adapted from Kim *et al.* 2021.

miRNAs repress gene expression through translation inhibition and/or inducing target mRNA degradation. 4EHP, a 5' cap binding protein, which blocks translational initiation by eIF4F complex, appears to be directed in part by its ability to interact with miRISC/target complexes, preventing the recruitment of accessory factors involved in decapping and de-adenylation (Jafarnejad *et al.* 2018). Recent work in *C. elegans* have identified new interactors of miRISC and characterized GYF-1 and its binding partner IFE-1, the nematode ortholog of 4EHP (Mayya *et al.* 2021; Figure 1.7A). Like its human ortholog, the GYF-1/IFE-4 complex is a potent inhibitor of translation, presumably by mediating miRISC's interaction with a subset of miRNAs (Figure 1.7A). Furthermore, deadenylation was not induced by the GYF/IFE-4 dependent miRISC complex suggesting that the mRNA decay was not activated, and that GYF-1/IFE-4 are working as pure translational repressors independent of the mRNA decay pathway (Mayya *et al.* 2021; Figure 1.7A).

In contrast to the strictly translational repression model of GYF-1/IFE-4, a larger group of RBPs has been extensively studied for their role in the active repression and degradation of mRNA targets. Poly(A) binding proteins (PABPs), deadenylases, and decapping enzymes, all work cooperatively to elicit target degradation (Figure 1.7A). Poly(A) binding proteins (PABPs) are cytoplasmic proteins responsible for binding 3'UTRs of cytoplasmic proteins to facilitate regulation of translation and mRNA decay (Goss and Kleiman 2013). PABPs primary role in mRNA decay is to bridge the interaction between mRNAs and deadenylase complexes (CCCR/NOT) to facilitate the deadenylation of the transcript and targeting it for decay (Wu *et al.* 2006; Chekulaeva *et al.* 2011). Interestingly, although PABPs and their interactions with deadenylases are well conserved, the mechanism by which they function and their absolute requirement for mRNA decay may be different across species. In *C. elegans*, PAB-1 and PAB-2

actively participate in miRNA-mediated decay of target transcripts (Flamand *et al.* 2016). PAB-2 was identified from miRNA pulldowns along with the core miRISC machinery including ALG-1, ALG-2, AIN-1, and AIN-2; interactions which were confirmed by co-IP of PAB-1 and PAB-2 (Flamand *et al*. 2016). Consistent with a role in miRISC-mediated function, depletion or aberration of PAB-1/PAB-2 impairs the function of multiple developmentally required miRNAs (Flamand *et al.* 2016). In comparison to humans, where PABPs are essential for both miRNA-mediated deadenylation and silencing, *C. elegans* PAB-1/PAB-2 appear to only be essential for silencing. Interestingly, PAB-1/PAB-2 appear to promote miRNA mediated silencing both in the presence and absence of a poly(A) tail suggesting a poly(A) independent role (Flamand *et al.* 2016). Furthermore, the structural characteristics of mRNAs including UTR length, distance between miRNA targets sites and poly(A) appear to modulate the function of PAB-1/PAB-2 providing a distinction between otherwise conflicting requirements in *C. elegans* (Flamand *et al*. 2016; Figure 1.7A).

Consistent with the poly-A tail being a critical target of miRISC, in most species, absence or loss of the poly(A) tail mimics the effects of successful miRISC binding and action, notably leading to rapid mRNA decay (Pérez-Ortín *et al.* 2013). Poly(A) tails are actively targeted by complexes composed of deadenylase proteins (Nicholson and Pasquinelli 2018; Figure 1.7A). In general, there are two distinct deadenylases complexes observed in animals, the PAN-2/PAN-3 and the CCR4/NOT complexes (Fabian *et al.* 2011). In *C. elegans*, the GW182 homolog and miRISC effector protein AIN-1 directly binds PAN-3, bridging miRISC with PAN-2, the enzymatically active subunit of PAN-2/PAN-3 complex (Kuzuoğlu-Öztürk *et al.* 2012; Figure 1.7A). Furthermore, AIN-1 also directly binds NOT-1 and NOT-2 (*C. elegans* homologs of CCR4/NOT subunits). Overall, despite a great deal of sequence divergence between GW182

proteins, the interaction between miRISC and PAN-2/PAN-3 and the CCR4/NOT deadenylase complexes is well conserved across species and required for miRNA mediated repression of target transcripts. In addition to poly(A) tails, eukaryotic mRNAs are universally capped with a m7G cap their 5' end (Moteki and Price 2002); the lack of a 5' cap more often than not signals the transcript for decay (reviewed in Jones *et al.* 2012). The deadenylases CCR4/NOT were partially implicated in the recruitment of the decapping enzymes to miRISC, thus linking deadenylation and decapping with translational silencing of miRNA targeted mRNAs (Braun *et al.* 2011; Figure 1.7A). In mammals, the DCAP1/DCAP2 complex has the task of actively decapping mRNAs to regulate their translation (Charenton and Graille 2018), likely through inhibition of translation initiation (Meijer *et al*. 2013). The presence of decapping enzymes within cytoplasmic foci (especially P-bodies) and in close association with XRN ribonucleases also suggests a tight link between translational repression and mRNA decay in cytoplasm (Eulalio *et al.* 2007).

RBPs are also capable of modulating the binding of miRISC and mRNA, without directly participating in the degradation of the transcript. For example, HuR prevents mRNA decay by antagonizing miRNA-mediated repression by binding mRNA ARE's in the 3'UTR at sequences that may lie overlapping or adjacent to miRNA target sites (Kundu *et al.* 2012). In these cases, HuR binding competes for the miRISC binding, thus preventing miRNA-mediated regulation of that particular transcript (Figure 1.7B). This competition is highlighted by a number of target mRNA/miRNA examples such as *top2a*/miR-548c (Srikantan *et al.* 2011) and *cox2*/miR-16 (Young *et al.* 2012). HuR can also direct the recruitment of miRISC to its bound mRNA to repress *c-myc* (Kim *et al.* 2009). In another example miR-19 RISC binding to target mRNAs is facilitated by HuR (Glorian *et al.* 2011). In both cases, it has been hypothesized that HuR binds

distantly enough from the miRNA binding site that miRISC is able to bind, aided by HuR-dependent conformation changes to the 3’ UTR (Kim *et al*. 2009; Figure 1.7B).

A recent large-scale study bioinformatically and experimentally characterized RBP-UTR interactions for 150 human RBPs (Kim *et al.* 2021). While the size of the RBP interactome had been previously hypothesized, this study confirmed that individual RBPs can target thousands of transcripts and the interactome as a whole may exceed thirty million possible interactions (Kim *et al*. 2021). Not surprisingly, a large number of the observed interactions were between RBPs and the 3'UTRs of miRNA targeted transcripts, with RBPs binding sites adjacent to miRNA target sites (Figure 1.7C). Consistent with this observation, a positive correlation was observed for the number of RBPs occupying a target 3’ UTR and miRNA targeting efficacy (the higher number of adjacent RBPs the higher the targeting efficacy) (Figure 1.7C). Furthermore, the distance between RBP binding sites and miRNA target directly correlated with the efficacy of miRNA targeting of their respective transcripts (Kim *et al*. 2021). Interestingly, direct interactions between RBPs and Argonaute proteins were not required to observe the previous correlations (Kim *et al*. 2021). However, it was noted that RBP occupancy of regions adjacent to miRNA target sites not only opened up the secondary structure of the 3' UTR but also promoted Argonaute residency. Therefore, the authors proposed a model where RPB binding to 3' UTR regions adjacent to miRNA target sites reduce the secondary structure, thereby increasing accessibility for miRISC and facilitating miRNA-mediated repression (Kim *et al*. 2021; Figure 1.7C).

Methods to identify RBP-miRNA interactions

The above study was made possible by existing and newly generated datasets identifying RBP target sites. Continuing efforts to identify and validate RBP and miRNA targets will be important for understanding how RBPs and miRNAs interact on target mRNAs. Protocols that employ various biochemical methods of purification have allowed for the development of a suite of high throughput and omic-scale experiments to identify RBPs and their interacting RNAs. In general, the protocols and techniques that identify RBP-RNA interactions fall into two categories: precipitation of RBPs and identification of their targets and sites and precipitation of RNA substrates and identification of associated RBPs. Below is an overview of both protein-based and RNA-based methods (Table 1.1).

Table 1.1. Methods to identify interactions between RBPs and miRNA.

Method	Concept	Advantages	Disadvantages	Examples
RIP	**Immunoprecipitate specific RBP and sequence interacting RNAs**			
RIP	Identify RNAs that interact with RBP of interest	Identifies interacting RBPs, doesn't require crosslinking, fast and easy ** DO-RIP modified protocol can provide high resolution RBP binding sites	Detection of off target RNAs, protein specific antibody use	Lerner and Steiz, 1979; Lu et al., 2014; **Nicholson et al., 2017
CLIP	**UV Crosslink RBP and RNA, immunoprecipitate RBP and sequence captured RNA**			
HITS-CLIP	Next gen sequencing of interacting RNAs	Crosslinking only captures tightly bound RNAs reducing off target RNAs	Crosslinking is inefficient, IP can miss rare RBP-RNA interactions	Licatalosi, 2008; Gillen et al., 2016
PAR-CLIP	Incorporates nucleoside analogs (4-SU) to crosslink at UV 365 nm	Analogs crosslinking increases RNA capture, high resolution binding sites	Potentially miss analog depleted sequences	Hafner et al., 2010; Danan et al., 2016
iCLIP	Circularization of RT product	Single nucleotide resolution of binding sites	Crosslinking is inefficient, IP can miss rare RBP-RNA interactions	Konig et al., 2010; Haberman et al., 2017
eCLIP	Use size matched input control, differential inline radom-mer incorporated into adaptor	Can distinguish between unique RNAs or PCR duplicates when looking at two identical reads	Crosslinking is inefficient, IP can miss rare RBP-RNA interactions	Van Nostrand et al., 2016; Van Nostrand et al., 2020
CLASH	RNA-RNA ligation to capture dsRNA hybrids	Identify regions of RNA-RNA interaction with high resolution	Crosslinking is inefficient, RNA-RNA ligations are inefficient	Kudla et la., 2012, Helwak et al., 2013
hiCLIP	RNA-RNA ligation to capture dsRNA hybrid, 3' ligation adaptor	Higher confidence identification of dsRNA regions	Crosslinking is inefficient, IP can miss rare RBP-RNA interactions	Sugimoto et al., 2015
Oligo pulldowns	**Pulldown on synthetic oligonucleotides to identify interacting RBPs by MS**			
Antisense oligos	Antisense probe binds complementary target, pulldown RBPs and RNAs	High throughput methods to capture interacting proteome, no metabolic labelling	Crosslinking is inefficient, off target binding	Castello et la., 2012
Sense oligos (mimics)	Sense probes mimic endogenous RNAs, pulldown RBPs and RNAs	High throughput methods to capture interacting proteome, can also identify RNA binding partners (or targets)	Crosslinking is inefficient, off target binding	Subramaniam et la., 2015
RNA Tagging	**RNA is tagged with aptamer tag, aptamer pulldown captures RNAs and interacting RBPs, RBPs identified by MS. RNA is fluorescently labeled and hybridized to protein microarray, interacting proteins identified by fluorescence.**			
RNA aptamers (chemical ligand)	RNA is modified with chemical aptamer which that is recognized by ligand-bead complex	High throughput methods to capture interacting proteome, high efficiency capture of RBPs	Chemical modification of RNAs, low yield of protein	Zheng et al., 2016; Leppek & Stoecklin, 2013
RNA aptamers + protein	RNA is with aptamer, binding proteins expressed in same tissue	Can study interacting RBPS of any type of RNA including structured RNAs, minimal off target binding	Low yield, in vitro transcription required, modification of RNA structure	Tomasso et al., 2016; Slobodin & Gerst, 2010
Protein microarray	Labeled RNAs are hybridized to protein microarray	High throughput identification of RBP interactors	Recombinant proteins may impair relevant RNA-RBP interactions	Kretz et al., 2013; Siprashvili et al., 2016

Many methods identify RBP-RNA by using antibodies to immunoprecipitate the protein of interest and identify its RNA targets using next generation sequencing (Table 1.1). In its simplest form, RIP-seq is used to immunoprecipitate a protein and the associated RNAs, which are captured after washing (Leppek and Stoecklin 2014). Advances in sequencing technology has allowed for the high throughput identification of interacting RNAs (Cook *et al.* 2015). A modification of RIP-seq, called DO-RIP-seq utilizes partial RNA digestion to identify RBP-binding sites with moderate resolution (Nicholson *et al.* 2017). Although RIP-seq allows for high throughput survey of all the interacting RNAs, off target binding of unrelated RNAs is a concern.

To get around off-target binding, other methods were developed to increase the stringency of the RBP-RNA interaction (Table 1.1). This set of methods became known as CLIP (crosslinking and immunoprecipitation), which involves the crosslinking of RBP and RNA using UV or chemicals prior to immunoprecipitations using a protein-specific antibody (Hafner *et al.* 2021). Stringent washes after immunoprecipitation are meant to remove off target RNAs, leaving only those that are covalently bound to the RBP. The most basic form of CLIP methods, HITs-CLIP, adds high-throughput sequencing of cDNA allowing for a genome wide view of RBP-RNA interaction (Licatalosi *et al.* 2008; Lin and Miles 2019). However, this method has a distinct disadvantage of low yield of interacting RNAs, due to the low efficiency of UV crosslinking. The functional consequence of low yield is the possibility that some rare RBP-RNA interactions are missed (Hafner *et al*. 2021).

To address the major issues of older CLIP methods, iCLIP and eCLIP (enhanced CLIP) have increased stringency (less off target RNAs) as well as the higher degree of resolution (single nucleotide) (Van Nostrand *et al*. 2016; Lin & Miles 2019; Table 1.1). Other adapted CLIP methods, like PAR-CLIP (photoactivatable -ribonucleoside-enhanced crosslinking and

immunoprecipitation) utilize alternative crosslinking methods to increase the yield and sensitivity of RBP-RNA interactions to a single nucleotide resolution (Hafner *et al.* 2010; Lin & Miles 2019). This method has been used to identify miRNA and mRNA targets bound to Argonaute (Hafner *et al*. 2010). Additional methods have been developed to explore the RBPs that bind to dsRNA species (Table 1.1). CLASH (crosslinking, ligation, and sequencing of hybrids) (Helwak and Tollervey 2014) and hiCLIP (Sugimoto *et al.* 2015) utilize modified crosslinking that allows for the formation of chimeric RNA duplexes between interacting RNAs, thus making bioinformatic identification of RNA-RNA interactions more precise (Sugimoto *et al.* 2015).

RNA centric methods to identify RNA-RBP interactions have greatly advanced in recent years (reviewed in Ramanathan *et al.* 2019). The basis of many of the RNA-centered precipitations involve affinity purification of specific RNAs followed by mass spectrometry methods (AP-MS) to identify the associated proteins (Tomasso *et al.* 2016; reviewed in Ramanathan *et al.* 2019; Table 1.1). Many AP-MS methods use RNA that has been modified with an affinity tag (chemical tag or RNA aptamers) to achieve high affinity purification of the RNA of interest along with the interacting RBPs. Streptomycin-binding (Bachler *et al.* 1999), tobramycin-binding (Hartmuth *et al.* 2002), and streptavidin-binding aptamers (Leppek and Stoecklin 2014) can bind resin beads, allowing for affinity purification with varying levels of efficiency. In addition, several AP-MS methods utilize exogenous RNA aptamer tags in combination with aptamer-binding proteins, isolated from bacteria and bacteriophages (Table 1.1). The ARiBo method utilizes the λboxB tag at the 3' end of the *in vitro* transcribed RNA of choice, linking the RNA to an affinity bead via the λN-GST protein (Tomasso *et al.* 2016). Similarly, the MS2 stemloop can be embedded within an *in vitro* transcribed RNA of choice thus

linking the RNA to an affinity bead via the MS2-binding protein (Slobodin and Gerst 2010). Interestingly, since these two methods rely on the expression of proteins to bind the RNA tag, fusion proteins like MS2-GFP can be used to not only capture RNAs (bead-based capture) but also to visualize bound RNP complexes *in vivo* (Slobodin and Gerst 2010). Overall, chemical and RNA aptamer tags have the advantage of allowing for high efficiency purification and relatively clean analysis of interacting RBPs using normalization.

In addition to tagging RNAs, RBP-RNA interactions have also been interrogated using synthetic oligonucleotides (Table 1.1). The use of oligos falls into two general categories, antisense and sense, which depends on the strand of RNA used in the pulldown. One of the more straightforward methods utilizes antisense oligonucleotides to bind the mRNA or precursor miRNA of interest and purify using oligo(dT) beads (Castello *et al*. 2012). The use of antisense oligos can readily be applied to study the RBP-RNA interaction of a given RNA, including small RNAs. For example, Fabian *et al.* 2009 used a modified 2'-O-methyl oligo mimicking a let-7 site to identify factors that associate with the *let-7*-loaded miRISC. In contrast, sense oligo pulldowns aim to identify RBPs that directly interact with an RNA of interest, by mimicking its sequence and structure. This method is especially useful to identify RBPs that bind small RNAs like miRNAs (Subramanian *et al.* 2014). The major advantage of oligonucleotide methods is the highly stability of RBP-RNA-oligo complexes which allows for stringent biochemical conditions, ultimately minimizing contaminants (Castello *et al.* 2012). In addition, levels of the oligos can be titrated to study dosage-dependent interactions or identify RPB interactors of RNAs with low abundance.

Lastly, RNA-RBP interactions can be detected using RNA labeling and protein microarrays (Table 1.1). In this method, the RNA of choice is labeled with Cy5 dye and then

hybridized to protein microarray containing 9,400 human proteins (Human Proto Array) (Kretz *et al.* 2013). Proteins that interact with the RNA of interest can be detected by fluorescent signal. Although this method represents a quick method to identify RBP interactors, the recombinant proteins lack post-translational modifications, protein-protein interactions, or cofactors that can influence RNA binding, thus potentially missing biologically relevant RNA-RBP interactions.

Chapter 2 - KH domain containing RNA-binding proteins coordinate with microRNAs to regulate *Caenorhabditis elegans* development.

This chapter was published as:

Haskell, D & Zinovyeva A, 2021. KH domain containing RNA-binding proteins coordinate with microRNAs to regulate *Caenorhabditis elegans* development. G3 Genes, Genomes, Genetics, Volume 11, Issue 2

Abstract

microRNAs (miRNAs) and RNA binding proteins (RBPs) regulate gene expression at the post-transcriptional level, but the extent to which these key regulators of gene expression coordinate their activities and the precise mechanisms of this coordination are not well understood. RNA binding proteins often have recognizable RNA binding domains that correlate with specific protein function. Recently, several RBPs containing K Homology (KH) RNA binding domains were shown to work with miRNAs to regulate gene expression, raising the possibility that KH domains may be important for coordinating with miRNA pathways in gene expression regulation. To ascertain whether additional KH domain proteins functionally interact with miRNAs during *Caenorhabditis elegans* development, we knocked down twenty-four genes encoding KH-domain proteins in several miRNA sensitized genetic backgrounds. Here, we report that a majority of the KH domain-containing genes genetically interact with multiple miRNAs and Argonaute *alg-1*. Interestingly, two KH domain genes, predicted splicing factors *sfa-1* and *asd-2*, genetically interacted with all of the miRNA mutants tested, while other KH domain genes exhibited genetic interactions only with specific miRNAs. Our domain architecture and phylogenetic relationship analyses of the *C. elegans* KH domain-containing proteins revealed potential groups that may share both structure and function. Collectively, we show that many *C. elegans* KH domain RBPs functionally interact with miRNAs, suggesting direct or indirect coordination between these two classes of post-transcriptional gene expression regulators.

Introduction

Most developmental and cellular processes rely on precise choreography of gene regulatory networks that incorporate a wide range of cellular and environmental inputs. Evolution of multiple regulatory pathways provided cells with multifaceted and combinatorial methods of regulating gene expression allowing for robustness, flexibility and rapid remodeling of expression patterns. One of the essential layers of gene regulation occurs at the post-transcriptional level and is effected by two classes of molecules: small non-coding RNAs called microRNAs (miRNAs) and RNA binding proteins (RBPs). The human genome is predicted to encode at least 2000 miRNAs (Alles *et al.* 2019) and approximately 1500 RBPs (Gerstberger *et al.* 2014). In comparison, *C. elegans* genome is predicted to encode more than 180 miRNAs (Ambros and Ruvkun 2018), and at least 850 RBPs (Tamburino *et al.* 2013) making it a more tractable model to study the genetic interactions between miRNAs and RBPs.

Most mature miRNAs are generated via a canonical multi-step biogenesis pathway that starts with transcription of primary miRNA (pri-miRNA) transcripts (reviewed in Gebert and MacRae 2019). Pri-miRNAs are then processed by consecutive enzymatic activities of Drosha and Dicer endonucleases to generate a double stranded RNA duplex, which is ultimately loaded into an Argonaute protein. A single miRNA strand is retained by an Argonaute and the mature miRNA silencing complex (miRISC) is formed when the miRNA-loaded Argonaute associates with a GW182 effector on the target messenger RNA (mRNA) (reviewed in Gebert and MacRae 2019). The miRISC identifies target mRNAs through partial sequence complementarity, ultimately resulting in translation repression and/or mRNA degradation (reviewed in O'Brien *et al.* 2018; Gebert and MacRae 2019).

RNA-binding proteins regulate diverse aspects of mRNA lifecycle, including splicing, transport, and stability (Dassi 2017). Diversity in protein architecture and auxiliary domains, as well as a high degree of modularity allow RBPs to impart specific and potent effects on the gene expression of their targets (Janga 2012). For example, the PUF family of proteins in *C. elegans* inhibit translation of their mRNA targets through sequence specific binding of the 3'UTR in order to promote deadenylation or by physically blocking cap recognition by translation initiation factors (reviewed in Wang and Voronina 2020). Other proteins like OMA-1 appear to play a more nuanced role by concomitantly binding the 3'UTRs of mRNAs along with translational repressors like LIN-41 in order to mediate the selective repression-to-activation transition for a subset of mRNAs essential for oogenesis (Tsukamoto *et al.* 2017). Here, RBPs and miRNAs are thought to cooperate extensively, and de-regulation of their activity can precipitate widespread disruption of gene regulatory networks resulting in a variety of cell pathologies and disease states (Tüfekci *et al.* 2013; O'Brien *et al.* 2018).

To effect post-transcriptional regulation of gene expression, RBP and miRNA activity can intersect on multiple levels. On a most basic level, miRNA biogenesis is performed and aided by RBPs (reviewed in Gebert and MacRae 2019). RBPs may directly associate with miRNA-target complexes to modulate the downstream effects on target gene expression (Hammell *et al*., 2009; Schwamborn *et al.* 2009; Wu *et al.* 2017). Coordination between RBPs and miRNAs can also be indirect, with individual factors affecting the target mRNA in distinct ways, ultimately resulting in a unique combinatorial gene regulatory outcome.

Among RBPs identified as modulators of miRNA activity are three proteins that share a conserved RNA-binding K-homology (KH) domain (Akay *et al.* 2013; Zabinsky *et al.* 2017; Li *et al.* 2019). KH domain was first described in human hnRNP K (Siomi *et al.* 1993, reviewed in

Geuens *et al.* 2016) and is present alone or in tandem in a large group of RBPs associated with transcription or translation regulation (Nicastro *et al.* 2015; Dominguez *et al.* 2018). The type I KH domain, found in eukaryotes, is approximately 70 amino acids and is characterized by three anti-parallel beta sheets abutted by three alpha helices; it includes the GXXG loop, which is thought to be responsible for nucleic acid binding (Grishin 2001; Valverde *et al.* 2008). We recently showed that HRPK-1, a KH domain-containing protein, physically and functionally interacts with miRNA complexes to modulate gene expression during *C. elegans* development (Li *et al.* 2019). Similarly, the KH domain protein VGLN-1 genetically interacts with a diverse set of miRNAs involved in early embryonic and larval development (Zabinsky *et al.* 2017). VGLN-1 binds mRNAs rich with miRNA binding sites in their 3'UTR (Zabinsky *et al.* 2017) and may serve as a platform, bridging interactions between multiple miRNAs, mRNAs, and proteins to regulate gene expression (Zabinsky *et al.* 2017). GLD-1, an RNA-binding protein and a well-characterized translational repressor that regulates germline development (Marin and Evans 2003), has been shown to genetically interact with multiple miRNAs (Akay *et al.* 2013). GLD-1 contains a single KH domain, functionally interacts with miRNA modulators, *nhl-1* and *vig-1,* and physically interacts with ALG-1, CGH-1, and PAB-1, proteins that are key for miRNA gene regulatory activity (Akay *et al.* 2013). Collectively, these findings suggest that RBPs that harbor KH domain(s) may be functionally important for miRNA-dependent gene regulation.

To determine the extent of functional coordination between the KH domain-containing proteins and miRNAs, we knocked down 24 additional predicted *C. elegans* KH domain genes in sensitized miRNA genetic backgrounds. Strikingly, knock down of nineteen KH domain genes resulted in a modulation of a phenotype associated with a partial loss of miRNA activity. We

found that several genes, including the predicted splicing factors *sfa-1* and *asd-2* genetically interacted with multiple miRNAs families, suggesting that splicing events may influence miRNA gene regulatory activity. Other genes, such as *Y69A2AR.32,* showed miRNA family specificity. Knockdown of most KH domain genes resulted in enhancement of miRNA reduction-of-function phenotypes, suggesting a normally positive functional interaction between KH domain RBPs and miRNAs. However, knockdown of several genes resulted in mild to strong suppression of defects observed in an Argonaute *alg-1* antimorphic mutant, suggesting that some of these factors normally act antagonistically to miRNAs. Overall, this work provides a comprehensive examination of the genetic interactions between miRNAs and KH domain RBPs in *C. elegans,* presents a phylogenetic and a domain analysis of *C. elegans* KH domain-containing proteins, and suggests that these RBPs may directly or indirectly coordinate with miRNA pathways to regulate gene expression.

Results

Multiple KH domain genes genetically interact with lsy-6 miRNA in ASEL neuronal cell fate specification.

The *lsy-6* miRNA controls cell fate specification of the ASEL/ASER sensory neuron pair. *lsy-6* normally represses expression of *cog-1* in the ASEL neuron, ultimately resulting in an ASEL specific gene expression pattern (Johnston and Hobert 2003; Figure 2.1A). Loss of *lsy-6* activity results in an inappropriate cell fate switch of the ASEL neuron to the ASER cell fate (Johnston and Hobert 2003). The *lsy-6(ot150)* reduction-of-function mutation causes a low penetrance phenotype, with ~15% of *lsy-6(ot150)* animals displaying an ASEL cell fate defective

phenotype. This cell fate defect can be observed by the loss of the *Plim-6::gfp* expression within the ASEL neuron (Figure 2.1A, B).

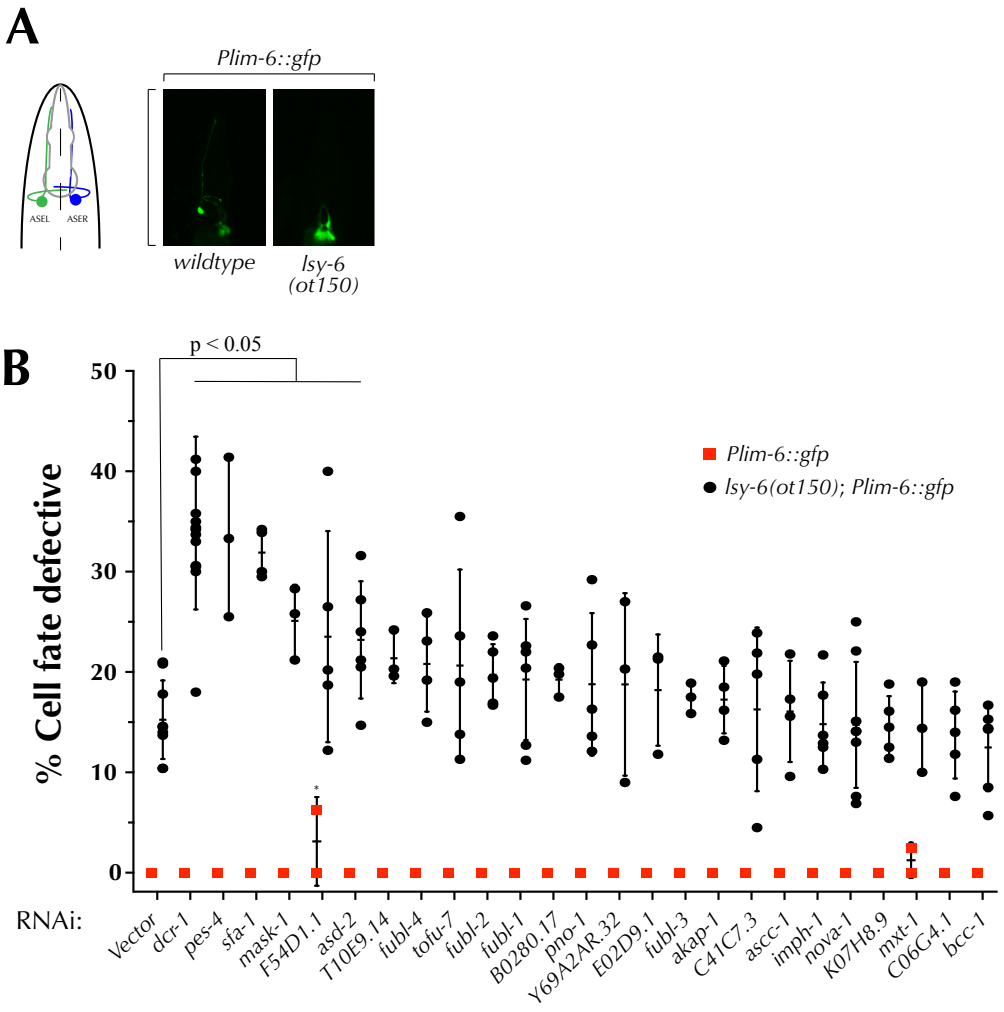


Figure 2.1. Knockdown of several KH domain genes enhances the cell defective phenotype of *lsy-6(ot150)* mutants.

(**A**) *lsy-6* miRNA directs the ASEL cell fate specification, with ASEL cell fate marked by the *lim-6::gfp* reporter. *lsy-6(ot150)* mutations results in partially penetrant loss of *lim-6::gfp* expression in ASEL cells. (**B**) RNAi-mediated knockdown of five KH domain genes significantly enhances the cell fate defective phenotype of *lsy-6(ot150)* animals. ANOVA test was used to determine statistical significance.

To identify whether KH-domain genes play a role in *lsy-6*-dependent neuronal cell fate specification, we knocked down twenty-four KH domain genes in the *lsy-6(ot150)* mutant background (Figure 2.1B) and assayed the penetrance of the ASEL cell fate defect. Knockdown of five of the KH domain genes, *pes-4, sfa-1, mask-1, F54D1.1,* and *asd-2* significantly enhanced the *lsy-6(ot150)* cell fate defective phenotype (Figure 2.1B, Table 2.1). RNAi-mediated knockdown of the KH domain genes did not result in a phenotype in the absence of the *lsy-6(ot150)* allele, with the exception of *F54D1.1* and *mxt-1*, whose depletion caused an occasional loss of *Plim-6::gfp* expression in ASEL (Figure 2.1B). Furthermore, RNAi of *T10E9.14*, *fubl-4, tofu*-7 and *pno-1* resulted in variable and/or mild but not statistically significant enhancement of the *lsy-6(ot150)* phenotype (Figure 2.1B).

KH domain genes coordinate with lsy-6 to regulate the expression of cog-1.

Next, we wanted to determine whether the genes that genetically interacted with *lsy-6(ot150)* were also able to regulate a *lsy-6* target, *cog-1* (Johnston and Hobert 2003). While *lsy-6* expression is normally restricted to neuronal tissues, its endogenous target *cog-1* is more broadly expressed (Palmer *et al.* 2002). Expression of *lsy-6* from the *cog-1* promoter represses the *cog-1::gfp* reporter in the uterine and vulval cells (Johnston and Hobert 2003; Figure 2.2A). Therefore, we can utilize the *lsy-6*-mediated repression of *cog-1* to assay the effects of knocking down potential modulators of *lsy-6* activity. Indeed, RNAi of five genes, *sfa-1, tofu-7, pes-4, asd-2,* and *T10E9.14* resulted in a significant de-repression of *cog-1::gfp* expression in uterine cells (Figure 2.2B, Table 2.1), suggesting that these genes may coordinate with *lsy-6* in repressing *cog-1*. Although not statistically significant the knockdown of *tofu-7, asd-2, and F54DS1.1* mildly repressed *cog-1* expression in the uterine cells in the absence of *Pcog-1::lsy-6*

(Figure 2.2B), suggesting these genes may regulate *cog-1* independently of *lsy-6.* In fact, *tofu-7, asd-2,* and *F54D1.*1 may have a more complex functional relationship, perhaps regulating *cog-1* through multiple genetic pathways, including one that involves the *lsy-6* miRNA. Here, *tofu-7, asd-2,* and *F54D1.1* may act to promote *cog-1::gfp* expression in the absence of *lsy-6*, while the addition of *lsy-6* changes the functional relationship from positive to repressive or may de-regulate target gene expression in either direction (Figure 2.2B).

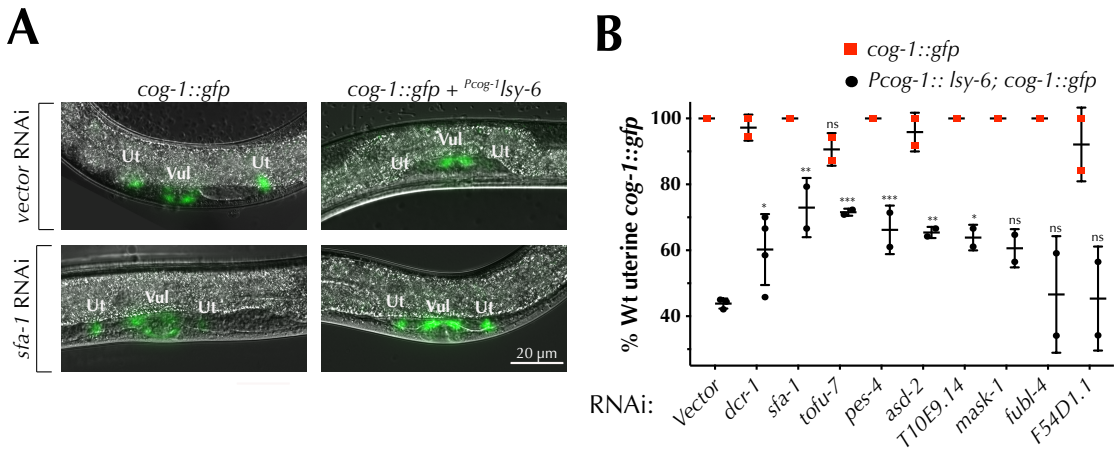


Figure 2.2. Several KH domain genes coordinate with *lsy-6* to regulate *cog-1::gfp* expression in uterine cells.

(A) RNAi of several KH domain genes, including *sfa-1* alleviates the *lsy-6*-mediated repression of *cog-1*::*gfp* in uterine cells (**B**). ANOVA test was used to determine statistical significance. ∗ p≤0.05, ∗∗p≤0.01, ∗∗∗p≤0.001.

KH domains proteins coordinate with let-7 family of miRNAs.

To determine whether KH domain-containing proteins might coordinate with additional miRNAs beyond *lsy-6*, we looked for a genetic interaction between the KH domain genes and the *let-7*-family of miRNAs. The *let-7* miRNA family, as part of a complex genetic network, regulates division patterns and terminal cell differentiation of seam cells during *C. elegans* larval development (Reinhart *et al.* 2000; Slack *et al.* 2000; Abbott *et al.* 2005). Three members of the *let-7* family, *mir-48, mir-241,* and *mir-84* act redundantly to control seam cell divisions by inhibiting the proliferative divisions of the L2 stage and promoting the self-renewing seam cell divisions of the L3 stage (Abbott *et al.* 2005). Loss of *mir-48, mir-241,* and *mir-84* leads to a highly penetrant reiteration of the proliferative L2 seam cell division leading to increased seam cell number, delayed alae formation, and delayed expression of the adult specific reporter, *col-19::gfp* (Abbott *et al.* 2005). Deletion of *mir-48* and *mir-241,* which leaves *mir-84* intact, results in milder heterochronic phenotypes including increased seam cell number and delayed alae formation (Abbott *et al.* 2005).

Table 2.1. KH domain genes functionally interact with miRNA sensitized mutants.

Gene or allele	*lsy-6(ot150)*	*lsy-6*	*mir-48 mir-241(nDf51)*		*let-7(n2853)*	*alg-1(ma202)*
RNAi \ Allele	Cell fate[a] defective	Uterine[b] *cog-1::gfp*	Abnormal[c] *col-19::gfp*	Seam cell[d] number	Bursting[e]	Wildtype[f] *col-19::gfp*
Empty vector	15.2 ± 3.9	48.8 ± 1.5	11.2 ± 7.9	13.4 ± 1.4	13.7 ± 6.9	0
dcr-1	34.9 ± 8.6	60.2 ± 10.8	79.4 ± 15.7	14.6 ± 1.9	33.4 ± 4.8	n.d.
fubl-3	17.4 ± 1.5	n.d.[g]	**50.4 ± 19.2**[h]	**14.2 ± 1.7**	**31.7 ± 8.7**	0
fubl-4	20.8 ± 4.7	46.6 ± 17.7	**49.7 ± 7.5**	14.1 ± 2.1	14.5 ± 10.6	0
fubl-1	19.3 ± 6.0	n.d.	**37.2 ± 29.5**	13.3 ± 1.4	24.2 ± 6.2	2.7 ± 3.8
fubl-2	19.7 ± 3.1	n.d.	22.3 ± 19.7	13.2 ± 1.5	12.2 ± 2.5	2.3 ± 3.2
imph-1	14.8 ± 4.2	n.d.	6.2 ± 5.4	13.5 ± 1.5	9.5 ± 1.8	7.3 ± 7.2
pes-4	**33.4 ± 8.0**	**66.2 ± 7.4**	**34.0 ± 12.2**	12.7 ± 1.1	10.6 ± 2.1	0
T10E9.14	21.4 ± 2.5	**63.9 ± 3.9**	**51.9 ± 18.9**	**14.1 ± 1.9**	15.9 ± 8.3	0
nova-1	14.7 ± 6.3	n.d.	6.7 ± 9.4	14.0 ± 1.6	11.5 ± 4.8	3.0 ± 4.2
mxt-1	14.5 ± 4.5	n.d.	26.6 ± 14.7	**14.2 ± 1.9**	**28.4 ± 11.9**	0
ascc-1	16.1 ± 5.1	n.d.	18.1 ± 13.5	13.2 ± 1.3	15.2 ± 7.1	4.2 ± 5.9
akap-1	17.3 ± 3.4	n.d.	**35.8 ± 2.1**	**14.1 ± 1.5**	13.2 ± 9.4	7.2 ± 10.1
tofu-7	20.6 ± 9.6	**71.6 ± 1.1**	**37.2 ± 18.2**	**14.2 ± 2.1**	17.5 ± 6.5	0
C06G4.1	13.7 ± 4.3	n.d.	13.9 ± 4.0	13.4 ± 1.6	10.9 ± 6.9	2.3 ± 3.2
E02D9.1	18.2 ± 5.5	n.d.	3.9 ± 6.8	13.2 ± 1.2	7.2 ± 5.4	**58.0 ± 19.0**
sfa-1	**32.0 ± 2.5**	**73.0 ± 9.0**	**43.9 ± 27.3**	**14.4 ± 1.8**	**25.3 ± 19.2**	7.2 ± 6.2
asd-2	**23.2 ± 5.8**	**65.4 ± 1.7**	**51.8 ± 23.9**	13.9 ± 1.9	**30.6 ± 19.8**	0
B0280.17	19.3 ± 1.5	n.d.	29.9 ± 19.8	13.5 ± 1.5	**31.5 ± 9.2**	0
F54D1.1	**23.5 ± 10.5**	58.3 ± 25.1	24.9 ± 7.7	13.9 ± 1.4	15.7 ± 7.0	0
K07H8.9	14.7 ± 2.9	n.d.	10.8 ± 14.3	13.7 ± 1.4	9.7 ± 5.4	3.0 ± 4.1
Y69A2AR.32	18.8 ± 9.1	n.d.	**40.8 ± 4.6**	13.5 ± 1.4	12.1 ± 3.0	6.0 ± 1.6
bcc-1	12.5 ± 4.4	n.d.	20.1 ± 12.5	**14.0 ± 1.9**	**31.0 ± 12.5**	0
C41G7.3	16.3 ± 8.2	n.d.	**45.5 ± 15.4**	**15.0 ± 2.3**	13.3 ± 9.2	1.4 ± 2.5
pno-1	18.8 ± 7.1	n.d.	**40.2 ± 12.0**	13.3 ± 1.4	14.0 ± 6.8	0
mask-1	**25.1 ± 3.6**	60.6 ± 5.8	**50.7 ± 18.6**	**14.7 ± 2.3**	12.9 ± 3.0	0

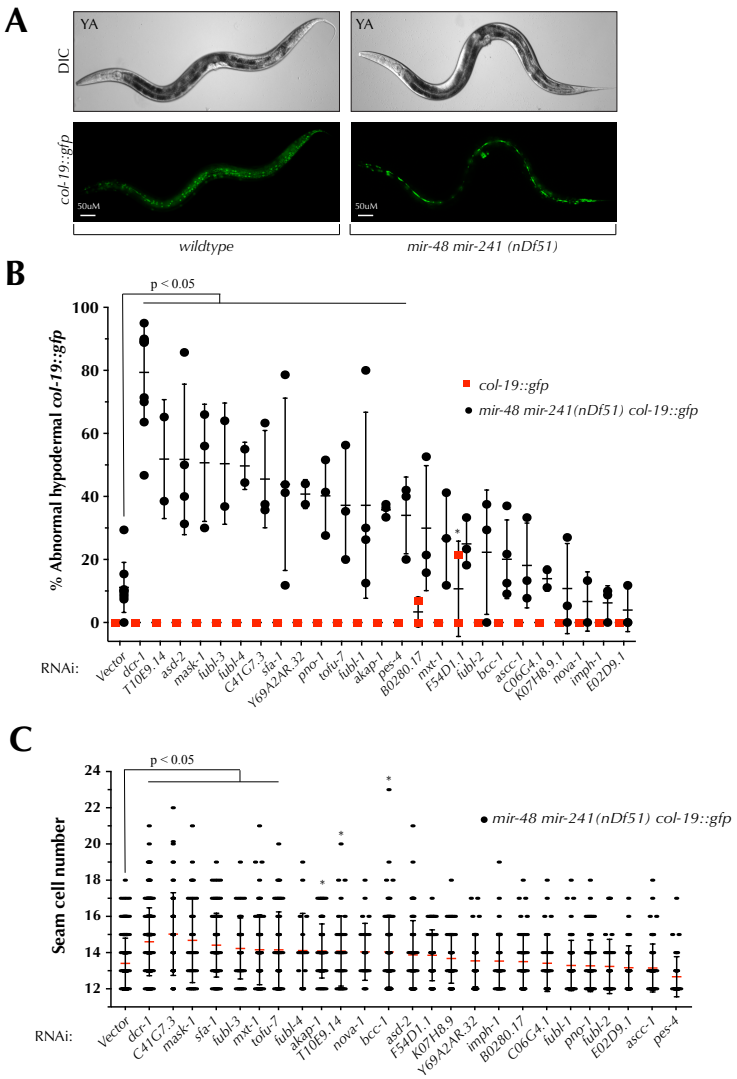


Figure 2.3. RNAi of multiple KH domain genes enhances the *mir-48 mir-241* heterochronic phenotype.

(**A**) Loss of *mir-48 mir-241* results in delayed hypodermal expression of the adult specific marker *col-19::gfp*. (**B**) When compared to vector RNAi, knockdown of 13 KH domain genes by RNAi enhances the delayed hypodermal *col-19::gfp* expression of *mir-48 mir-241(nDf51)* animals. Dots represent experimental replicates. (**C**) RNAi of some KH domain genes increases the seam cell numbers of *mir-48 mir-241(nDf51)* young adults when compared to vector RNAi. ANOVA test was used to determine statistical significance. ∗ $p \leq 0.05$

We performed RNAi for the twenty-four KH domain genes in the *mir-48 mir-241(nDf51)* mutant background and assayed both *col-19::gfp* expression and seam cell number in young adult animals (Figure 2.3A, B). *mir-48 mir-241(nDf51)* young adults fail to properly undergo the adult-specific developmental program, thereby showing a delay in *col-19::gfp* expression consistent with a delay in normal developmental timing. RNAi of thirteen KH domain genes significantly enhanced the abnormal *col-19::gfp* expression phenotype observed in *mir-48 mir-241(nDf51)* animals (Figure 2.3B, Table 2.1). RNAi of the twenty-four genes did not produce a phenotype in the absence of the *mir-48 mir-241* deletion (Figure 2.3B), with the exception of *B0280.17* RNAi, which exhibited a very mild defect in hypodermal *col-19::gfp* expression. In addition, *F54D1.1* RNAi produced a mildly penetrant abnormal *col-19::gfp* expression, but did not enhance the *mir-48 mir-241* phenotype to a statistically significant level (Figure 2.3B). RNAi of nine KH domain genes produced a significant increase in the average number of seam cells in the *mir-48 mir-241(nDf51)* mutants compared to the empty vector control (Figure 2.3C, Table 2.1). Overall, depletion of seven genes both enhanced the delayed hypodermal *col-19::gfp* expression and increased the seam cell number of *mir-48 mir-241(nDf51)* mutants (Figure 2.3B, C, Table 2.1). Together these data suggest that a subset of KH domain genes may coordinate, directly or indirectly, with the *let-7 family* miRNAs to regulate their target gene expression.

To further explore this level of coordination, we examined the role of KH domain genes in the regulation of *hbl-1,* a transcription factor and a known target of the *let-7* family of miRNAs (Abbott *et al*., 2005, Abrahante *et al*., 2003, Lin *et al*., 2003). Expression of *hbl-1* is normally temporally restricted to embryo-L2 animals and upon exit from the L2/L3 molt the expression of *hbl-1* is greatly reduced (Abbott *et al*., 2005). To understand how KH domain

genes may be effecting *hbl-1* expression, either through the *let-7* family of miRNAs or independently, we performed RNAi of the top ten genes identified in the *mir-48 mir-241* assays as well as *F54D1.1* and assessed *hbl-1::gfp* expression. Normally, hypodermal *hbl-1::gfp* becomes downregulated as the animals molt from L2 to L3 and is largely absent in L3 animals (Figure 2.4A). Since reduction of miRNA activity results in inappropriate hypodermal *hbl-1* expression at the L3 stage, we sought to determine the percentage of worms displaying abnormal hypodermal *hbl-1::gfp* expression in early/mid L3 animals (Figure 2.4B). RNAi of *asd-2, C41G7.3,* and *sfa-1* produced a significant change in the abnormal expression of *hbl-1::gfp* in L3 animals, although most genes tested increased the abnormal expression in some RNAi replicates (Figure 2.4B). These data suggest that KH domain genes may play a role in the regulation of *hbl-1,* perhaps through the *let-7* family of miRNAs or through another indirect mechanism.

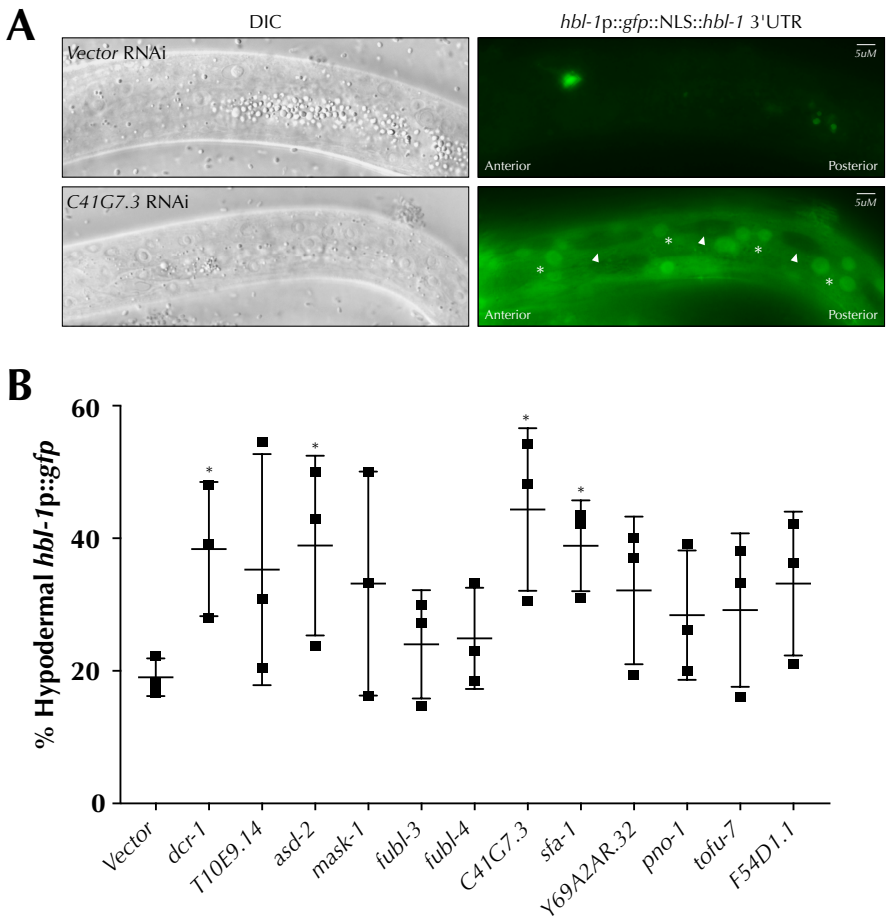


Figure 2.4. RNAi of some KH domain genes affect with *hbl-1::gfp* expression.
(**A**) *hbl-1::gfp::hbl-1* 3′ UTR expression in vector RNAi control and *C41G7.3* RNAi L3 larvae. (**B**) RNAi of *C41G7.3* and other KH domain genes significantly enhances the number of animals displaying hypodermal expression of *hbl-1::gfp::hbl-1* 3' UTR in L3 animals. Hypodermal cells are labeled with white asterisks. Seam cell nuclei are labeled with white arrowheads. Non-hypodermal neuron is indicated with an open arrowhead. ANOVA test was used to determine statistical significance. ∗p≤0.05.

To assess the functional relevance of KH domain genes to miRNA activity later in development, we asked whether reducing KH domain gene function impacts activity of *let-7* itself. *let-7* governs the terminal seam cell differentiation during the transition from L4 to adulthood (Reinhart *et al.* 2000). Compromising *let-7* miRNA activity produces a heterochronic phenotype, which, among other defects, includes vulval rupture during the L4-adult molt (Reinhart *et al.* 2000). *let-7(n2853)*, a temperature sensitive reduction-of-function mutation, causes a mildly penetrant vulval rupture phenotype at 15ºC (Figure 2.5A; Reinhart *et al.* 2000). RNAi of six KH domain genes led to significant enhancement of the vuval bursting phenotype (Figure 2.5B) suggesting these genes may coordinate with *let-7* miRNA in a way that normally promotes its activity.

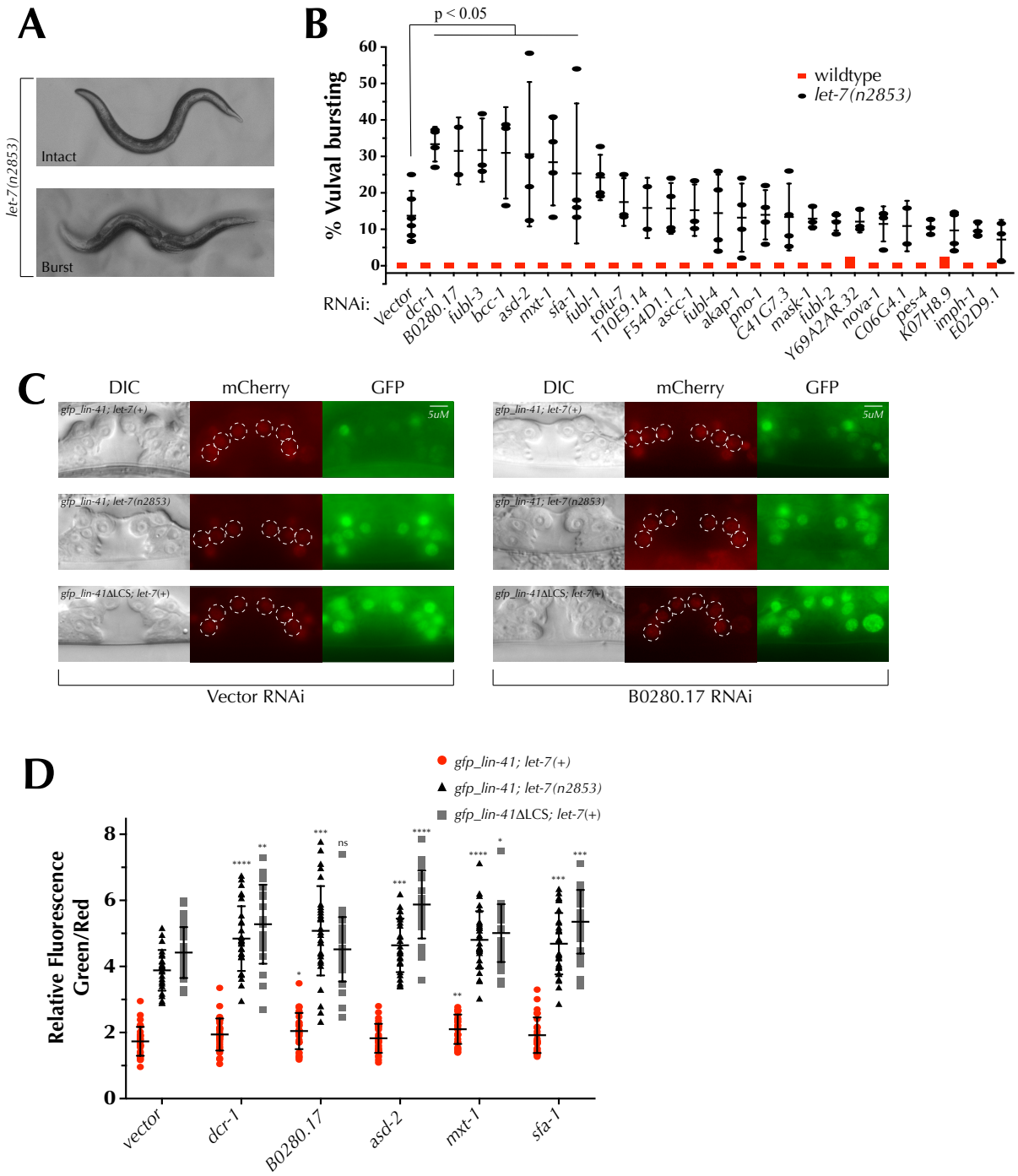


Figure 2.5. Several KH domain genes interact genetically with the *let-7* miRNA and its target *lin-41*.

(**A, B**) *let-5(n2853)* worms display a partially penetrant vulval bursting phenotype at permissive temperature (15ºC). (**B**) RNAi knockdown of six KH domain genes significantly enhances the vulval bursting phenotype of *let-7(n2853)* worms. (**C**) Expression of a reporter system previously designed to assess miRNA activity on miRNA target *lin-41* 3' UTR (Ecsedi et al., 2015) in vector control and B0280.17 RNAi. Three strains express *Pdpy-30::GFP::lin-41* 3'UTR and *Pdpy-30::mCherry* control in vulval cells: wild type (*gfp_lin-41; let-7(+)*), *let-7(n2853)* (*gfp_lin-41; let-4(n2853)*), and wild type *let-7* with *lin-41ΔLCS* reporter lacking the two functional *let-7* complementary sites within the *lin-41* 3′UTR (*gfp_lin-41ΔLCS; let-7(+)*). When *let-7* activity is compromised or *let-7* sites are removed from *lin-41* 3'UTR, *GFP* expression is de-repressed (quantified in **D**). (**D**) RNAi of 4 KH domain genes alleviates the repression on *GFP::lin-413' UTR* expression when *let-7* activity is compromised. Images shown in (C) were adjusted post-quantification to allow reader to more easily visualize vulval cells. ANOVA test was used to determine statistical significance. ∗p≤0.05, ∗∗p≤0.01, ∗∗∗p≤0.001, ∗∗∗∗p<0.0001.

To further explore the genetic interaction between KH domain genes and *let-7* miRNA, we asked whether KH domain genes can regulate expression of a *let-7* target, *lin-41* (Vella *et al*., 2004). We performed RNAi of four of the genes identified in the *let-7(n2853)* assay, utilizing three reporter strains: *lin-41::gfp* alone, *lin-41::gfp* with *let-7(n2853)*, and *lin-41::gfp* delta LCS (two major *let-7* binding sites removed from the *lin-41* 3'UTR) (Ecsedi *et al*., 2015). These strains express *lin-41::gfp* and an mCherry control reporter in the vulval cells, which provides a convenient tool to quantify *lin-41* expression (Figure 2.5C, Ecsedi *et al., 2015*). RNAi of each *B0280.17* and *mxt-1* de-repressed *lin-41::gfp* when *let-7* activity (Figure 2.5C, D). RNAi of the *B0280.17*, *asd-2*, *mxt-1*, and *sfa-1* significantly de-repressed *lin-41::gfp* when *let-7* function was compromised *(lin-41::gfp*; *let-7(n2853))* (Figure 2.5C, D). However, *asd-2*, *mxt-1*, and *sfa-1* also significantly de-repressed *lin-41::gfp* in the delta LCS animals (Figure 2.5D) indicating that these effects on *lin-41::gfp* are not directed through the two *let-7* sites deleted in this reporter strain, at least not exclusively. Interestingly, RNAi of *dcr-1* also increased the *lin-41::gfp* fluorescence level both in the *lin-41::gfp; let-7(n2853)* and *lin-41::gfp delta LCS* backgrounds (Figure 2.5D), suggesting that other miRNA target sites might be engaged in regulation of *lin-41::gfp*. These data support the possibility that *B0280.17* may be coordinating with *let-7* miRNA to regulate *lin-41::gfp*, while other genes may regulate *lin-41::gfp* independently of *let-7* under these conditions.

Knockdown of KH domain genes suppresses compromised miRISC activity.

ALG-1 is one of two *C. elegans* Argonautes (ALG-1 and ALG-2) that primarily associate with miRNAs and are central for miRNA biogenesis and activity (Grishok *et al.* 2005). Mutations abolishing ALG-1 activity result in moderate developmental defects, while abolishing

both *alg-1* and *alg-2* activity results in early lethality (Grishok *et al.* 2005; Vasquez-Rifo *et al.* 2012). In addition, antimorphic mutations in *alg-1,* such as *alg-1(ma202),* result in more pronounced defects in miRNA activity, likely due to sequestration of miRNA pathway components away from ALG-2 (Zinovyeva *et al.* 2014). Specifically, *alg-1(ma202)* animals display severe heterochronic defects (Zinovyeva *et al.* 2014), with 100% of *alg-1(ma202)* young adult animals failing to appropriately express adult cell marker *col-19::gfp* in hyp7 hypodermal cells (Zinovyeva *et al.* 2014; Figure 2.6A,B). Expression of *col-19::gfp* in young adult seam cells of *alg-1(ma202)* is variable, ranging from a complete lack of expression in the seam cells to full seam cell expression (Zinovyeva *et al*. 2014).

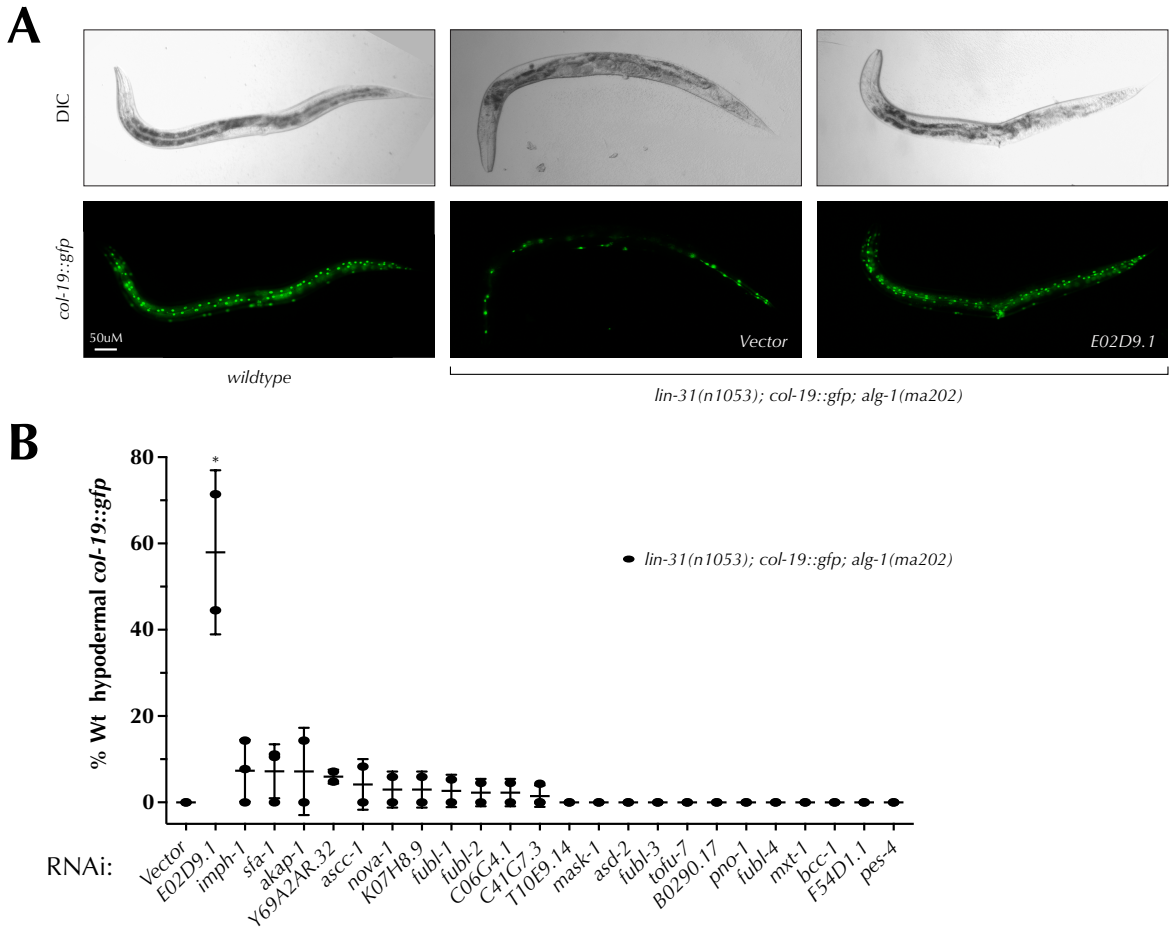


Figure 2.6. Two KH domain containing genes interact genetically with ALG-1(*ma202*). (**A**) *alg-1(ma202)* young adults lack hypodermal *col-19::gfp* expression and display variable *col-19::gfp* expression in seam cells. The *alg-1(ma202)* mutation is present in a *lin-31(n1053)* background to suppress bursting via non-heterochronic mechanisms. RNAi of *E02D9.1* restores *col-19::gfp* expression in young adults (**A, B**). (**B**) RNAi of several genes suppresses the delayed *col-19::gfp* expression phenotype of *alg-1(ma202)* mutants. ANOVA test was used to determine statistical significance. ∗p≤0.05.

To determine whether any of the KH domain genes might normally have a negative genetic relationship with miRNA pathway components, we performed RNAi knockdown of KH domain genes in the *alg-1(ma202)* background and assessed *col-19::gfp* expression in hypodermal cells of young adult animals. We used this background to screen for genes that may normally negatively interact with the miRNA pathways and therefore suppress the *alg-1(ma202)* phenotype when knocked down. Interestingly, RNAi of *E02D9.1* significantly suppressed the abnormal hypodermal *col-19::gfp* expression in *alg-1(ma202)* young adults, with ~60% of *alg-1(ma202)* animals exhibiting wild type hypodermal *col-19::gfp* expression in young adults (Figure 2.6A, B, Table 2.1). Although not statistically significant, possibly due to the variation in RNAi efficiency, RNAi of other genes (most notably *imph-1, sfa-1, akap-1,* and *Y69A2AR.32)* restored wild type *col-19::gfp* expression in *alg-1(ma202)* young adults, something that is never observed in *alg-1(ma202)* mutants alone (Figure 2.6B). As *alg-1(ma202)* suppressors, these KH domain genes may act in a manner that opposes normal miRNA activity, with their depletion perhaps resulting in decreased miRNA target gene expression.

KH domain containing RBPs play a role in early development

To determine whether KH domain genes have a general effect on *C. elegans* development, we assayed the brood size and embryonic lethality of animals with reduced KH domain gene function. Knockdown of seven genes (*fubl-4, pes-4, akap-1, E02D9.1, sfa-1, Y69A2AR.32,* and *bcc-1*) resulted in significant reduction in brood size (Figure 2.7A, Table 2.2), Depletion of *sfa-1* and *pes-4* had significant effects on both brood size and embryonic lethality suggesting that these genes play fundamental roles in *C. elegans* development (Figure 2.7A, B, Table 2.2). Several additional genes disrupted early development, albeit to a degree that was not

statistically significantly by our analysis (Figure 2.7, Table 2.2). These observations are consistent with previously reported roles for *akap-1, E02D9.1, sfa-1, asd-2, K07H8.9,* and *bcc-1* in early *C. elegans* development (Kamath *et al.* 2003; Sönnichsen *et al.* 2005; Ohno *et al.* 2008; Ma and Horvitz 2009; Kapelle and Reinke 2011) and highlight additional genes as important for *C. elegans* fecundity and embryonic development.

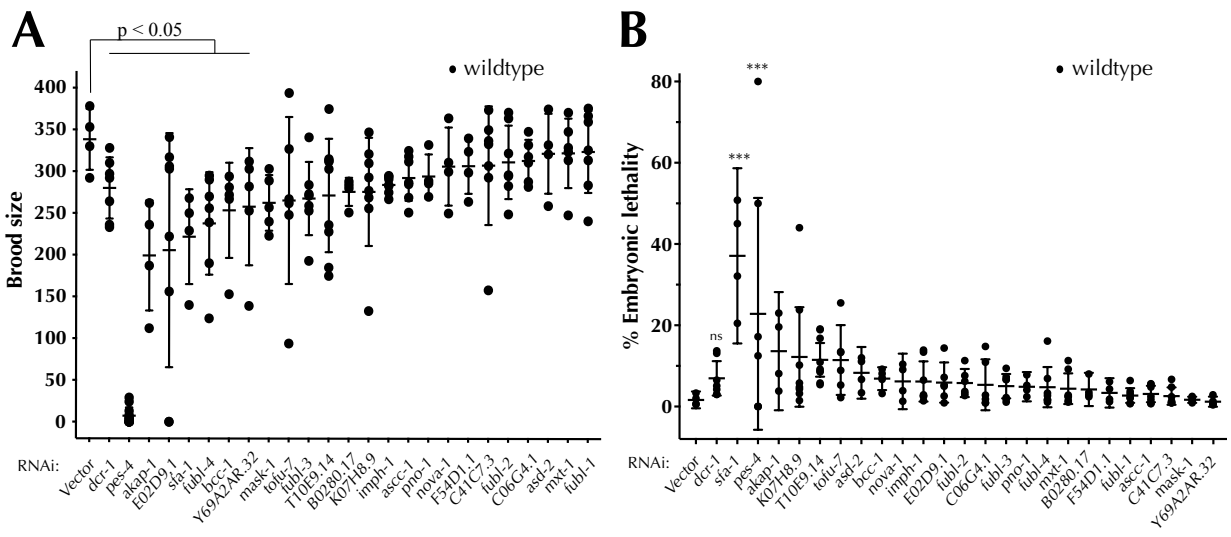


Figure 2.7. KH domain containing RNA-binding proteins may have essential roles in development.

(**A**) RNAi of seven KH domain genes resulted in significant reductions in brood size. (**B**) After knockdown *sfa-1* and *pes-4* significantly enhance embryonic lethality. ANOVA test was used to determine statistical significance. ***$p \leq 0.001$.

Some C. elegans KH domain proteins are evolutionary related and have diverse domain architecture.

Protein domains are discrete functional and structural segments of a protein. The loss, gain, or structural modification of domains can drive evolution, allowing proteins to lose or acquire new functions over evolutionary time. As domains evolve from ancestral forms, proteins containing the same types of domains may be evolutionary related. To understand the evolutionary relationship between the KH domain-containing proteins and to potentially inform our functional analysis, we performed an alignment of *C. elegans* KH domain protein sequences using the MEGAx alignment program (Kumar *et al.* 2018) and generated a phylogenetic tree (Figure 2.8). Interestingly, proteins that appear to coordinate with miRNAs are found in almost every clade in our phylogenetic analysis (Figure 2.8).

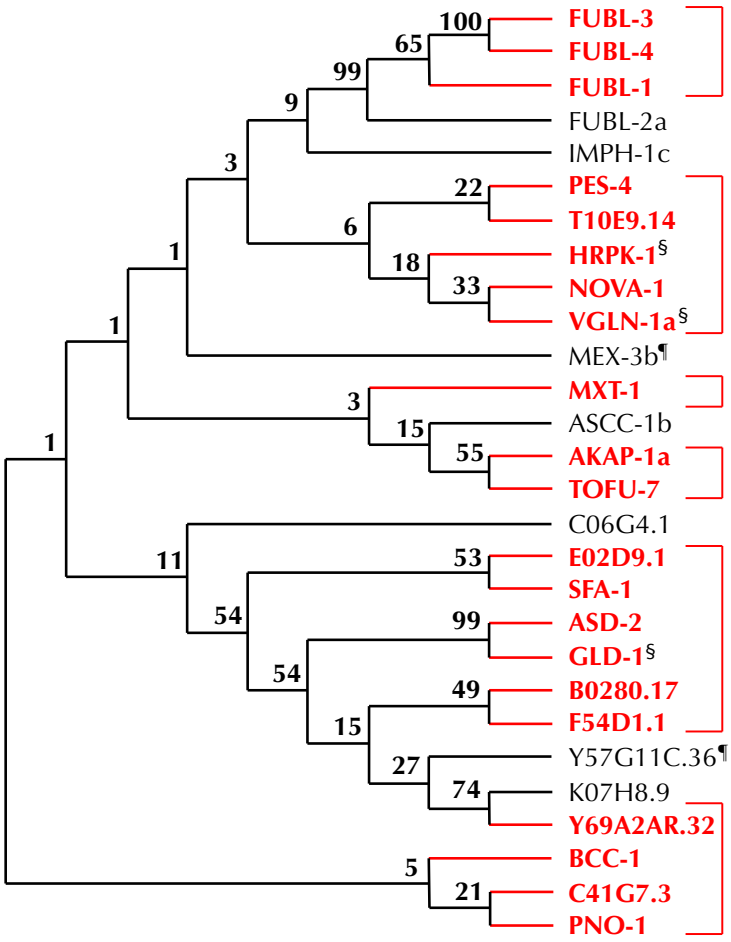


Figure 2.8. Phylogenetic analysis of KH domain containing RNA-binding proteins. Multiple sequence alignment of 28 KH domain containing RNA-binding proteins was performed and the proteins grouped in clades based on sequence similarity; branches are labeled with confidence value. Clades containing proteins that genetically interact with one or more miRNA sensitized background are bracketed and highlighted in red. A ¶ indicates that functional assays were not performed for a particular gene. A § denotes genes identified as interacting with miRNAs in other publications.

KH domains are thought to mediate numerous interactions, including those between proteins (Valverde *et al.* 2008) and proteins and nucleic acids (Grishin 2001; Valverde *et al.* 2008). Due to the KH domain's ability to bind RNA, the *C. elegans* KH domain-containing proteins represent a subset of RBPs, but combinatorial domain arrangements can result in extensive functional diversity among them. To determine the diversity of domain structures of KH domain containing proteins, we analyzed their domain architecture using the Simple Modular Architecture Research Tool (SMART) (Letunic and Bork 2017), which identifies known domain sequences. In addition, we utilized the PLACC web-based tool to identify prion-like domains, or unstructured regions (Lancaster *et al.* 2014). Such low complexity regions are thought to have affinity for RNA (Kato *et al.* 2012) and can play a role in phase-phase separation that is important for forming and reforming of ribonucleoprotein (RNP) bodies (Shin and Brangwynne 2017). We found that KH domain-containing proteins harbor a diverse set of domains, with prion-like domains present in 17/29 of KH domain proteins (Figure 2.9). Unsurprisingly, many proteins of the same clade shared additional domains (Figure 2.8 Figure 2.9). These analyses may in the future may help inform the mechanisms by which these proteins coordinate with miRNA-mediated regulation of gene expression.

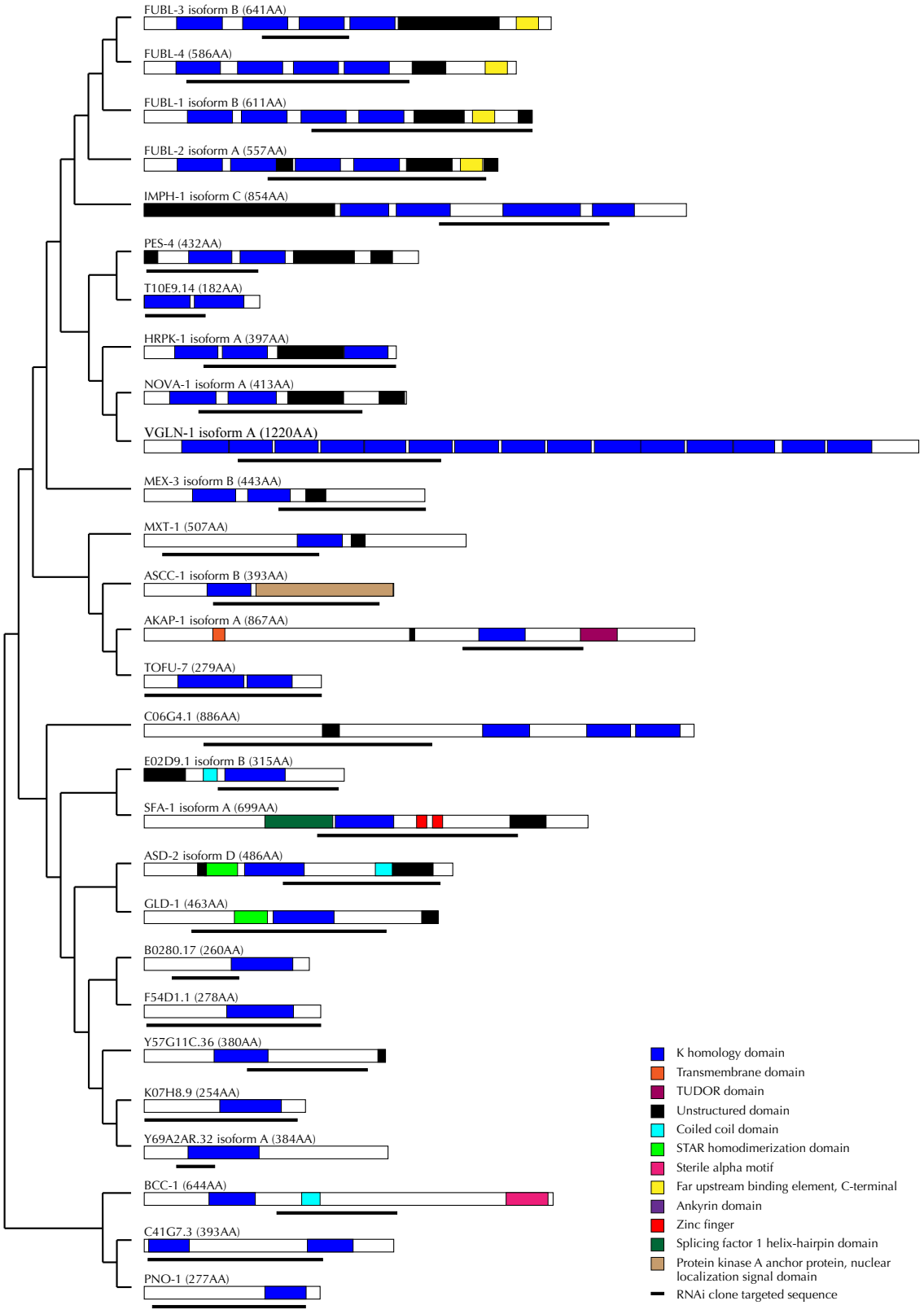

FUBL-3 isoform B (641AA)
FUBL-4 (586AA)
FUBL-1 isoform B (611AA)
FUBL-2 isoform A (557AA)
IMPH-1 isoform C (854AA)
PES-4 (432AA)
T10E9.14 (182AA)
HRPK-1 isoform A (397AA)
NOVA-1 isoform A (413AA)
VGLN-1 isoform A (1220AA)
MEX-3 isoform B (443AA)
MXT-1 (507AA)
ASCC-1 isoform B (393AA)
AKAP-1 isoform A (867AA)
TOFU-7 (279AA)
C06G4.1 (886AA)
E02D9.1 isoform B (315AA)
SFA-1 isoform A (699AA)
ASD-2 isoform D (486AA)
GLD-1 (463AA)
B0280.17 (260AA)
F54D1.1 (278AA)
Y57G11C.36 (380AA)
K07H8.9 (254AA)
Y69A2AR.32 isoform A (384AA)
BCC-1 (644AA)
C41G7.3 (393AA)
PNO-1 (277AA)
K homology domain
Transmembrane domain
TUDOR domain
Unstructured domain
Coiled coil domain
STAR homodimerization domain
Sterile alpha motif
Far upstream binding element, C-terminal
Ankyrin domain
Zinc finger
Splicing factor 1 helix-hairpin domain
Protein kinase A anchor protein, nuclear localization signal domain
RNAi clone targeted sequence

Figure 2.9. Domain architecture analysis of KH domain containing RNA-binding proteins.

Protein domains prediction analysis was performed on the longest predicted isoform using SMART (Letunic and Bork 2017). Region of each protein targeted by RNAi are highlighted below the predicted protein structure. Proteins are grouped in the clades identified via our phylogenetic analysis.

Materials and Methods

Worm strains

Worm culture and maintenance was performed as previously described (Brenner, 1974). Bristol N2 was used as the wildtype strain. Strains used in this study are OH3646 *lsy-6(ot150)*; *otIs114 [Plim-6-gfp + rol-6(su1006)],* OH812 *otIs114 [Plim-p-gfp + rol-6(su1006)],* PS3662 *syIs63 [cog-1::gfp + unc-119(+)],* OH7310 *otIs193 [Pcog-1::lsy-6 + rol-6(su1006)]; sy1s63,* VT1296 *mir-48 mir-241(nDf51) col-19::gfp (maIs105)*, MT7626 *let-7(n2853*), VT2223 (*lin-31(n1053*); *col-19::gfp(maIs105*); *alg-1(ma202*). All strains were grown at 20° C with the exception of MT7626 *let-7(n2853*) which was grown and maintained at 15° C to prevent excess bursting.

RNA interference

RNAi constructs (pL4440) were obtained from the Ahringer RNAi library (Kamath *et al.* 2000; Source Biosciences) except for *bcc-1* and *E02D9.1* which were obtained from the Vidal RNAi library (Rual *et al.* 2004; Source Biosciences). In addition, 3 RNAi clones were constructed by genomic amplification of the endogenous loci and cloning of the fragment into the L4440 vector. The *fubl-3* clone was generated by using forward 5'-GCCCACTAGTGGACTAACTGCAACGTTCAA-3' and reverse 5'-GTGGGTACCATTTGCCGCCTCAGAATTG-3'. The *Y6A2AR.32* clone was generated using forward 5'-GCTCAGATCTTGCCACGTTTCATGCGAAAC-3' and reverse 5'-GTAGGTACCGGAAGCTCTTCCTCTCACAA-3'. The *B0280.17* clone was generated using forward 5'-GGCCAGATCTCTTCTAGTTCGTGAAATCAA-3' and reverse 5'-ATAGGTACCGCAGTCTCGGGAGGAAAG-3'. The amplified genomic fragments containing

restriction sites were digested using SpeI and Kpn1 (*fubl-3*) or BglII and KpnI (*Y69A2AR.32 and B0280.17*) and were ligated with the digested L4440 vector using NEB (M2200) Quick Ligation protocol. Ligated plasmid was then transformed into *E. coli* HT115 bacteria. Sequence insertion into the L4440 plasmid was confirmed via Sanger sequencing (using M13 forward sequencing primer). Although the Ahringer clone targeting *mex-3* was obtained, RNAi of *mex-3* in *lsy-6(ot150)* and *mir-48 mir-241(nDf51)* resulted in highly penetrant embryonic lethality preventing scoring of the F1 progeny of the RNAi treated animals.

RNAi experiments were done by feeding and performed at 20°C unless otherwise stated and as described below. All RNAi experiments of individual genes were performed in parallel with empty vector RNAi (negative control). *dcr-1* RNAi was used as a positive control as loss or reduction of *dcr-1* eliminates/impairs miRNA biogenesis. RNAi plates were prepared and seeded using standard methods (Kamath *et al.* 2000). Scoring requiring fluorescence was done on a Leica DM6B fluorescent compound microscope. Imaging of fluorescence-based phenotypes was done using the Leica DM6B mounted camera and processed using Leica software. Photoplates were assembled using Adobe Illustrator. Scoring of vulval bursting, brood size, and embryonic lethality were done a standard Leica dissecting microscope. The number of animals scored per replicate as well as the percentage of animals displaying the abnormal phenotype in each replicate are reported in Supplemental Tables 2.1 and 2.2.

Despite the overall relatedness of protein architecture among member of phylogenetic clades (Figure 2.8), BLAST (NCBI) searches for RNAi-targeted regions suggest there is sufficient variation in nucleotide sequence for individual RNAi clones to specifically target the gene of choice. The rare exceptions may be the *fubl* genes, and the *asd-2*/*gld-1* pair which show

a very low level of overlap in targeted sequence, allowing for the possibility that some cross-gene RNAi targeting may occur.

ASEL cell fate differentiation

Plim-6::gfp (otIs114) and *lsy-6(ot150)*; *Plim-6::gfp (otIs114)* worms were placed on RNAi as embryos and F1 progeny were scored as L4 or young adults to increase the ease of detecting fluorescent signal in ASEL neurons. Each group of genes was scored alongside the negative control (empty L4440 vector) and our positive control (*dcr-1* RNAi). Worms were scored as cell fate defective when *lim-6::gfp* was undetectable in the ASEL neuron soma.

Uterine cog-1::gfp

cog-1::gfp (syIs63) and *cog-1::gfp (syIs63); otIs193[Pcog-1::lsy-6*; *rol-6(su1006)]* worms were placed on RNAi as embryos and F1 progeny were scored at mid-late L4s in order to ensure a strong GFP signal in both vulval and uterine cells. Each group of genes was scored alongside the negative control (empty L4440 vector) and our positive control (*dcr-1* RNAi). Worms were considered to have abnormal uterine *cog-1::gfp* if either the anterior or posterior or both uterine cells were lacking GFP. *cog-1* expression was scored as normal when GFP expression was observed in both uterine cells and in vulval cells.

Hypodermal col-19::gfp expression and seam cell number

col-19::gfp (maIs105) and *mir-48 mir-241(nDf51) col-19::gfp (maIs105)* animals were placed on RNAi as young L4s and their F1 progeny were scored as young adults. Each group of genes was scored alongside the negative control (empty L4440 vector) and our positive control

(*dcr-1* RNAi). Worms were scored first for the presence of *col-19::gfp* in the hypodermal cells. Normal expression was defined as all hypodermal cells expressing *col-19::gfp* while abnormal expression was defined as GFP signal absent in many or all of hypodermal cells. Worms were also scored for the number of seam cells present between the pharynx and rectal cells; seam cells were identified using the *col-19::gfp* transgene. *lin-31(n1053)*, *col-19::gfp* (*maIs105)*; *alg-1(ma202)* worms were scored in an identical manner when assaying hypodermal *col-19::gfp* expression.

hbl-1::gfp expression

hbl-1p::gfp::NLS::*hbl-1* 3'UTR (ctIS39) animals were synchronized by bleaching and plated on RNAi plates with the RNAi bacteria supplemented with fluorescent beads for accurate staging (Nika *et al*., 2016). Worms were grown until the majority began the L2 molt, at which point worms were screened for the presence of fluorescent beads within the gut. Worms that lacked beads and therefore had entered the L2 molt were picked to a new plate seeded with the equivalent RNAi bacteria without beads. The molting worms were then screened every 30 minutes for resumption of pumping indicating they had exited the molt into L3. The worms were then scored at 40x magnification for *hbl-1::gfp* expression in hypodermal cells. Representative images were taken at 63X magnification.

Vulval bursting

let-7(n2853) worms were grown and maintained at 15°C. Embryos were synchronized by hypochloride/NaOH solution and embryos plated directly onto RNAi plates as previously described (Parry *et al.* 2007). The embryos were hatched and grown at 15°C until young adults.

Worms were scored for vulval bursting ~ 6 hours after the L4 molt to ensure all animals had reached adulthood. Each group of genes was scored alongside the negative control (empty L4440 vector) and our positive control (*dcr-1* RNAi).

lin-41 reporter assay

[*Pdpy-30*::GFP(PEST)-H2B::*lin-41* 3'UTR (xeSi78); P*dpy-30*::mCherry::H2B::artificial 3'UTR (xeSi36)] (HW1113), **[**P*dpy-30*::GFP(PEST)-H2B::*lin-41* 3'UTR (xeSi78); P*dpy-30*::mCherry::H2B::artificial 3'UTR (xeSi36), *let-7(n2853)*] (HW1114), **[**P*dpy-30*::GFP(PEST)-H2B::*lin-41* delta LCS 3'UTR (xeSi87); P*dpy-30*::mCherry::H2B::artificial 3'UTR (xeSi36)] (HW1159) (Ecsedi et al., 2015) animals were synchronized by bleaching. Embryos were plated onto seeded RNAi plates. Worms were grown until L4 at which point they were imaged at 63x magnification in red and green channels to capture GFP and mCherry expression in the vulval cells. Leica image analysis software was used to determine the fluorescence in each region of interest (ROI) surrounding each of six vulval cells in both red and green channels. Identical exposure and microscope settings were used for all imaging to allow quantification and comparison of signals. To quantify the changes in *lin-41* expression we divided the relative signal intensity of the green channel by the signal intensity in the red channels in each of the vulval cells. The average signal intensity for the vulval tissue was determined by averaging the signal ratios across the six cells scored for each worm. The representative images were adjusted for brightness and contrast post-quantification to allow the reader to more easily observe the fluorescence in cells of interest.

Brood size and embryonic lethality

Wildtype (N2) worms were placed on RNAi as L4s and allowed to lay embryos. When the F1 progeny reached the L4 stage, individual hermaphrodites were moved to their own RNAi plates and allowed to lay embryos for 24 hours. After 24 hours, each animal was moved to a fresh RNAi plate each day for three additional days. Live larvae were counted on each plate (by picking) 24 and 48 hours after the parent has been moved to ensure all larvae were counted. Dead embryos on each plate were counted 48 hours after removal of the parent. The total number of live larvae and dead embryos for each hermaphrodite was tallied and together encompass brood size. Embryonic lethality was calculated as (# dead embryos/total brood size) x 100%. Larval arrest was rarely seen, but when it did occur these worms were counted as "live larvae" because they had successfully hatched and developed beyond the embryonic stage.

Phylogenetic Analysis

Full proteins sequences of the longest isoforms for each protein were collected from Wormbase and entered to the Mega X program. A MUSCLE protein alignment was carried out to provide input for further phylogenetic analysis. In order to construct the tree, we selected the Maximum Likelihood method and bootstrapped the tree-building (1000 iterations) to increase the stringency of the method. A simple LG model was selected for the substitution model, utilizing a Nearest-Neighbor-Interchange (NNI) method. The phylogenic tree shown represents 27 of the 28 KH domain proteins in the *C. elegans* genome: *mask-1* was excluded due to extensive length and sequence/domain variability from the rest of the protein family.

Protein domain and architecture

To generate the protein domain graphics, we first determined the longest isoform of each individual protein. The amino acid sequence of the proteins were obtained from Wormbase.org and entered into Simple Modular Architecture Research Tool (SMART) (Letunic and Bork 2017) under the Genomic options. Domain start and end points were noted and used to generate the proteins graphics in Adobe Illustrator.

To generate the coverage of each RNAi clone used in this study, primer pairs were obtained from the Ahringer library database, aligned to the appropriate transcript. Each RNAi target was then translated in the appropriate frame and aligned to the complete protein sequence. Predicted NLS sites were determined using cNLS Mapper using a threshold of 5.0 (Kosugi *et al.* 2009). Only high confidence (score > 8.0) NLS regions were included in the domain graphics.

Statistical analysis

All statistics were done using GraphPad Prism software. Statistical significance was determined using a one-way ANOVA test. To make the desired comparisons and avoid the loss of statistical power inherent to multiple comparisons, we used planned comparisons to compare each individual gene RNAi with vector control RNAi. Bonferroni correction was applied as a post hoc analysis.

Data Availability

Strains and plasmids are available upon request. All data necessary for confirming the finding of this article are present within the article and the associated figures, and tables.

Discussion

KH-domain containing RBPs functionally interact with multiple miRNA families.

To determine whether *C. elegans* KH domain-containing RBPs may function with miRNAs to regulate gene expression, we asked whether RNAi knockdown of KH domain genes could modify the phenotypes observed in reduction-of-function miRNA or family mutants. Surprisingly, nineteen of the twenty-four tested genes genetically interacted with at least one miRNA mutant background, suggesting widespread functional interaction between KH RBPs and miRNAs. Interestingly, the KH domain genes fell into two groups: those that modified phenotypes of all miRNA sensitized backgrounds tested and those that genetically interacted with specific miRNA reduction-of-function mutants (Table 2.1). *sfa-1 and asd-2* functionally interacted with multiple miRNA families (Table 2.1), suggesting that these two genes have broad roles in regulation of gene expression. The human ortholog of *sfa-1,* SF1 (Splicing Factor 1), participates in the spliceosome assembly by binding 3' branch sites of pre-mRNAs while its partner, U2AF, cooperatively binds the 5' branch site (Rino *et al.* 2008). Likewise, the ortholog of *asd-2*, *quaking*, has established roles in RNA processing, including alternative splicing and generation of select miRNAs and circular RNAs (Darbelli and Richard 2016). In *C. elegans,* both *sfa-1* and *asd-2* are predicted to play a role in splicing, with *asd-2* modulating the alternative splicing of *unc-60* and other transcripts (Kuroyanagi 2013) and *sfa-1* regulating the pre-mRNA splicing of multiple genes (Heintz *et al.* 2017). Depletion of either *sfa-1* or *asd-2* was sufficient to induce embryonic lethality and reduce brood sizes (Table 2.2; Ma and Horvitz 2009; Chu *et al.* 2014), consistent with their essential roles as potential global regulators of splicing. Unbiased reverse genetic screens have previously identified splicing machinery members as important for miRNA-mediated gene regulations (Parry *et al.* 2007). Similarly, factors involved in mRNA processing, including splicing, were found to modulate RNAi efficacy (Kim *et al.*

2005). While splicing and small RNA (including miRNA) pathways intersect, the exact mechanisms by which this occurs remain largely unknown. Given *sfa-1* and *asd-2* potential roles in splicing, it is perhaps not surprising that these factors show broad genetic interaction with miRNAs across all of our assays.

Table 2.2. Knockdown of KH domain gene affects embryonic lethality and brood size.

Assay / RNAi	Embryonic [a] lethality	Brood size [b]
Empty vector	1.6 ± 1.3 (4)	338.3 ± 36.5
dcr-1	6.9 ± 4.6 (7)	280.0 ± 36.6
fubl-3	5.0 ± 3.3(7)	267.7 ± 43.9
fubl-4	4.8 ± 5.4 (7)	**237.7 ± 61.4**[c]
fubl-1	2.1 ± 2.0 (7)	324.0 ± 49.1
fubl-2	5.8 ± 3.3 (6)	311.4 ± 44.1
imph-1	6.2 ± 5.4 (7)	284.1 ± 10.3
pes-4	**22.8 ± 30.9 (7)**	**7.2 ± 9.9**
T10E9.14	11.5 ± 5.4 (9)	271.3 ± 67.9
nova-1	6.2 ± 4.3 (4)	306.3 ± 46.8
mxt-1	4.4 ± 4.1 (7)	322.3 ± 41.6
ascc-1	3.1 ± 2.2 (7)	292.4 ± 27.2
akap-1	13.6 ± 9.1 (4)	**199.3 ± 66.0**
tofu-7	11.5 ± 8.2 (6)	265.3 ± 100.0
C06G4.1	5.4 ± 6.0 (6)	313.2 ± 25.3
E02D9.1	5.9 ± 6.0 (6)	**205.6 ± 140.2**
sfa-1	**37.1 ±13.6 (4)**	**221.8 ± 56.8**
asd-2	8.3 ± 4.0 (4)	321.8 ± 47.9
B0280.17	4.2 ± 2.6 (4)	275.8 ± 16.8
F54D1.1	3.4 ± 2.7 (4)	306.8 ± 33.1
K07H8.9	12.2 ± 14.7 (8)	275.8 ± 64.8
Y69A2AR.32	1.2 ± 1.0 (5)	**257.8 ± 70.2**
bcc-1	6.9 ± 2.3 (5)	**253.4 ± 57.1**
C41G7.3	2.6 ± 2.3 (7)	307.4 ± 71.1
pno-1	4.9 ± 2.3 (4)	294.3 ± 26.5
mask-1	1.7 ± 0.6 (5)	262.4 ± 33.3

To better understand the biological context in which miRNAs and KH domain proteins may interact, we compiled spatial KH domain gene expression patterns using existing promoterome (Hunt-Newbury et al., 2007), tissue-specific transcriptome (Kaletsky et al., 2018), and tissue-specific proteome (Reinke et al., 2018) datasets (Supplemental Table 2.3). Most KH domain genes are broadly expressed with both transcripts and proteins detected in multiple tissues (Supplemental Table 2.3). For the most part, genes whose RNAi produced a phenotype in a particular miRNA background seemed to be expressed in the relevant tissues. For example, Y69A2AR.32 expression in the hypodermis correlated with its knockdown effects on *col-19::gfp* expression in *mir-48 mir-241* mutant animals (Figure 2.3). Similarly, most of the KH domain genes whose knockdown resulted in *lsy-6(ot150)* phenotype enhancement are neuronally expressed, although it remains unknown whether these KH domain genes are specifically expressed in ASEL/R neurons. Future work is needed to characterize precise tissue and cellular expression to fully understand the spatial and temporal overlap among the molecules in question.

In contrast to the splicing-related factors, the majority of the KH domain-containing RBPs genetically interact with specific miRNAs (Table 2.1). RNAi knockdown of *pes-4* and *mask-1* enhances phenotypes of both *lsy-6(ot150)* (Figure 2.1 and Table 2.1) and *mir-48 mir-241(nDf51)* (Figure 2.3 and Table 2.1) mutants, suggesting a somewhat general role in gene regulation that spans multiple tissues. By contrast, *akap-1, C41G7.3, pno-1*, *fubl-1*, *fubl-4*, and *Y69A2AR.32* genetically interacted with *mir-48 mir-241(nDf51)* (Figure 2.3 and Table 2.1), but not *let-7(n2853)* (Figure 2.5 and Table 2.1), suggesting a narrower role for these KH domain genes in target gene regulation. Such functional separation can be achieved through differences in temporal expression or perhaps through distinct specificities of RBPs to target RNAs. In comparison, RNAi of *bcc-1*, *fubl-3* and *mxt-1* genetically interacted with both *mir-48 mir-*

241(nDf51) and *let-7(n2853),* but not *lsy-6(ot150)* (Table 2.1). The *let-7-family* shared interactions suggest that these RBPs may have more general roles in developmental timing or may regulate broader sets of target genes. Interestingly, *fubl-1 (C12D8.1)* was previously identified as a functional interactor of RNAi (Kim *et al.* 2005), suggesting that this gene's activity may impact gene regulation carried out by multiple small RNA pathways. *tofu-7* was previously identified in a screen for regulators of piRNA biogenesis and function (Goh *et al*., 2014). In addition, *fubl-1*, *fubl-3*, *fubl-4*, *imph-1*, and *nova-1* show significant phylogenetic clustering with RNAi related genes when integrating existing immunoprecipitation and *Drosophila* miRNA and siRNA datasets into cluster analysis (Tabach *et al*., 2013). Taken together, these observations suggest that some of the KH domain genes may coordinate with several small RNA pathways.

KH domain protein relatedness.

Protein domains are conserved, structured portions of a protein that can fold and function independently. As distinct functional units of a protein, they can dictate, or add to, the overall cellular and molecular role of the protein. Evolution of protein structure and function is in part driven by addition or removal of domains through genetic recombination of domain-encoding gene sequences. To better understand the evolutionary and functional relatedness of the KH domain-containing proteins in *C. elegans* we performed a phylogenetic analysis (Figure 2.8). Our analysis highlights the overall diversity of these proteins, revealing low levels of similarity between many of the clades, consistent with the observation that in many cases, the proteins sequence similarity is limited to the KH domain(s). However, in contrast to the overall diversity of the proteins, we do see high degrees of relatedness in several of the clades, most notably those

containing the FUBL proteins and the grouping consisting of GLD-1 and ASD-2 (Figure 2.8). This is not surprising given the similarity in domains and overall protein architecture (Figure 2.9). The phylogeny highlights several clades that genetically interact with miRNAs (Figure 2.8), perhaps reflecting the functional relatedness relevant to regulation of gene expression.

Potential models for KH domain RBP and miRNA coordination.

How might KH domain RBPs functionally interact with miRNA pathways to regulate gene expression? Given the evolutionary and domain architecture diversity, the KH RBPs may coordinate with miRNAs, directly or indirectly, via distinct mechanisms. KH RBPs may directly affect aspects of miRNA biogenesis and function or they may indirectly intersect with miRNA pathways by affecting target mRNA processing, transport, stability, and degradation, independent of miRNA activity.

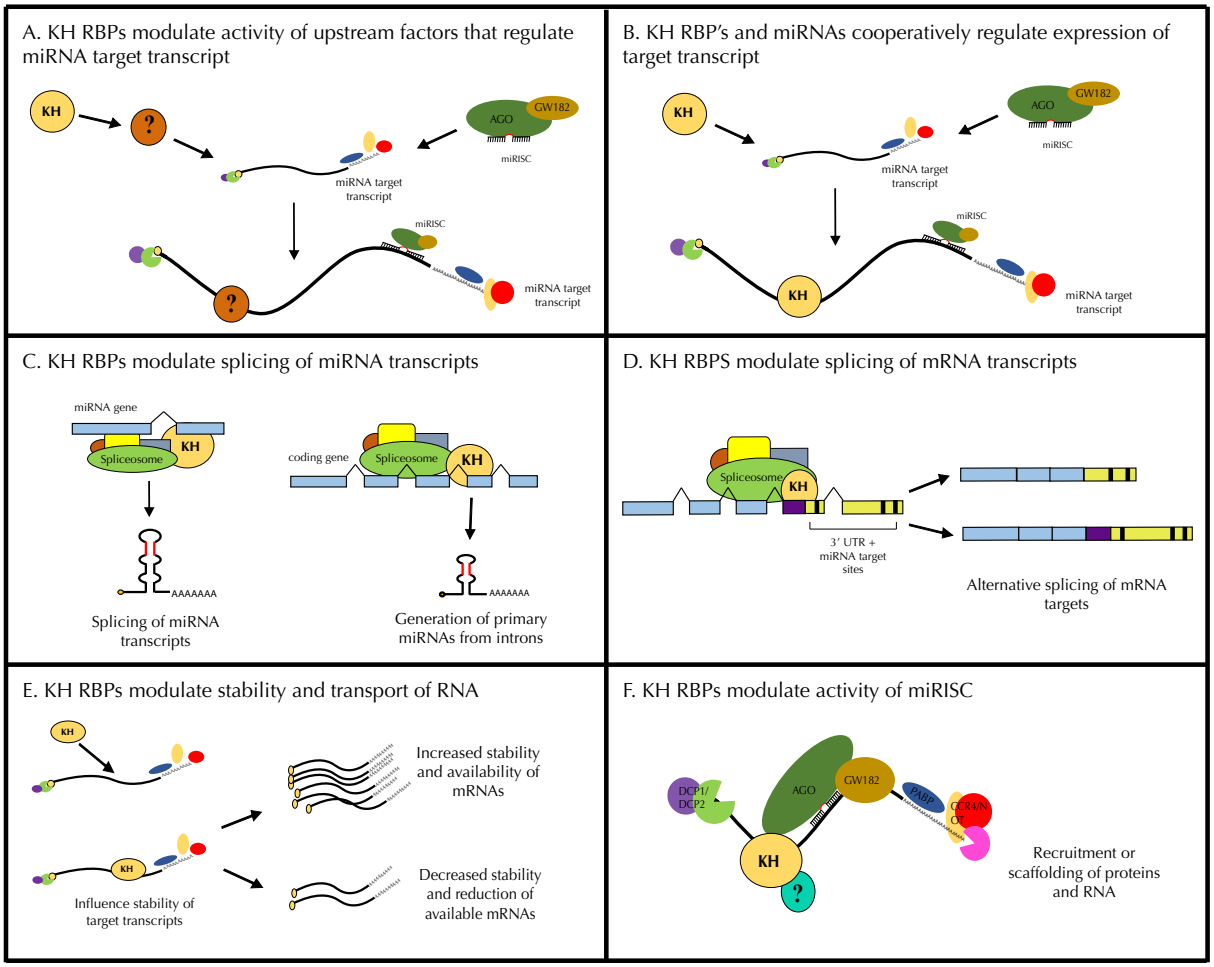


Figure 2.10. Models for potential interactions between miRNAs and KH domain-containing RNA binding proteins.

(**A**) KH RBPs may indirectly modulate miRNA target gene expression through one or more intermediate effector(s). (**B**) KH RBPs may regulate miRNA target gene expression by acting directly on the miRNA target transcripts as elaborated on in **C-F**. (**C**) KH RBPs may modulate the splicing of primary miRNA transcripts. (**D**) RBPs may modulate the splicing of miRNA target transcripts and alter the availability of miRNA target sites in their 3' UTRs. (**E**) RBPs may modulate the activity of miRISC by bridging known RNA and protein components or by the recruitment of additional regulatory factors. (**F**) RBPs may modulate the stability of miRNA target transcripts. Loss of KH domain proteins could increase the pool of target

mRNAs, enhancing the miRNA reduction-of-function phenotypes, or decrease the pool of target mRNAs, resulting in the suppression of miRNA mutant phenotypes.

Overall, KH RBPs may exert their gene regulatory effects on miRNA targets indirectly, through multiple effectors (Figure 2.10A). Alternatively, KH RBPs could more directly regulate the life cycle of miRNA targets by interfacing directly with the miRNAs themselves or with the target mRNAs (Figure 2.10B). There are multiple mechanisms through which KH RBPs could contribute to miRNA target gene regulation. Proteins involved in splicing, such as SFA-1 and ASD-2, may be involved in splicing events that lead to the production of miRNA transcripts either from their independent gene loci or as part of host mRNA processing (Figure 2.10C). In this scenario, loss of a splicing factor's function may reduce the amount of primary miRNA transcript produced, enhancing the reduction of function phenotypes observed in our sensitized backgrounds (Figure 2.10C). In addition, splicing factors may indirectly intersect with miRNA pathways by either increasing or decreasing the availability of a gene target (Figure 2.10D). Alternative splicing of 3' UTRs that eliminate miRNA target sites has been recently observed (Han *et al.* 2018). Under this model, KH domain gene depletion could result in alternatively spliced mRNA isoforms that are no longer able to escape miRNA-mediated regulation (Figure 2.10D), enhancing the phenotypes observed in our reduction of function miRNA mutants.

In another possible scenario, KH domain-containing factors may affect mRNA stability, localization, or transport and thus alter the pool of available miRNA targets (Figure 2.10E). Increased stability of target mRNAs perhaps through sequestration could reduce miRNA efficacy (Figure 2.10E). In contrast, reduced stability of miRNA target mRNAs could result in suppression of miRNA-related phenotypes observed in our assays. Interestingly, *Drosophila* orthologs of MXT-1 (MEXTLI) and B0280.17 (HOW) can enhance the stability of mRNAs (Nabel-Rosen *et al.* 2002; Hernández *et al.* 2013). The B0280.17 ortholog (HOW) shows isoform dependent enhancement or suppression of mRNA stability in order to modulate mRNA

translation (Hernández *et al.* 2013). Likewise, the human orthologs of the FUBL proteins can positively or negatively modulate (depending on the protein) translation of their mRNA targets by binding the 3' UTRs and influencing their stability (Zhang and Chen 2013). These observations lend further support to this model and suggest that the genetic interactions between these RBPs could be complex and context dependent.

Lastly, it is possible that some KH domain-containing RBPs may directly interact with protein components of the miRNA pathway to modulate target gene expression. Several proteins contain additional domains that are predicted to have RNA-binding activity (SAM, zinc finger, splicing factor helix hairpin) (Figure 2.9) and could mediate interactions among proteins and RNA. Other functional domains such as prion-like or low complexity domains were present in approximately 50% of the RBPs tested. These domains have been implicated in driving liquid phase separation and formation of protein aggregates and RNPs (Putnam *et al.* 2019). We also see several examples of domains critical for protein-protein interactions, notably the TUDOR domain present in AKAP-1, the STAR homodimerization domains present ASD-2 and GLD-1, and the ankyrin repeats in VGLN-1. Some KH domain-containing RBPs may alter the activity of miRISC by bridging essential protein components or by recruiting additional regulatory factors (Figure 2.10F). This model is supported by the observation that eight of the twenty-nine KH domain-containing RBPs were previously found to physically interact with miRISC components or DCR-1 (Table 2.3). MASK-1, FUBL-1, -2, and -3 co-precipitate with AIN-1 (Wu *et al.* 2017), while HRPK-1 and IMPH-1 co-precipitate with DCR-1 (Duchaine *et al.* 2006) and ALG-1 (Zinovyeva *et al.* 2015). GLD-1 was found to co-precipitate with ALG-1 (Akay *et al.* 2013; Zinovyeva *et al.* 2015) and AIN-2 (Zhang *et al.* 2007). These proteins may act as scaffolds for the formation of RNP complexes, bridging RNA components (mRNA or miRNA) miRNA

biogenesis factors or miRISC (Figure 2.10F). Overall, KH RBPs could act directly on miRNA targets (Figure 2.10B) via the suggested mechanisms (Figure 2.10A-D) or could indirectly coordinate with miRNAs in regulating gene expression through one or more intermediates (Figure 2.10A).

Table 2.3. Several KH domain containing RBP's physically interact with miRISC components.

Gene	miRISC or biogenesis component
mask-1	AIN-1[a]
fubl-1	AIN-1[a]
fubl-2	AIN-1[a]
fubl-3	AIN-1[a]
hrpk-1	ALG-1[b], DCR-1[c]
imph-1	ALG-1[b], DCR-1[c]
gld-1	ALG-1[b], AIN-2[d]
vlgn-1	AIN-2[e]

Overall, our screen demonstrated that many of the KH domain-containing RBPs in *C. elegans* functionally interact with miRNA-mediated regulation of gene expression. Further work is essential to characterize the mechanisms through which individual KH domain proteins may affect gene expression and how they might functionally intersect with miRNA pathways. This study highlights a number of candidates for future genetic, molecular, and biochemical characterization and demonstrates the extent to which miRNAs and KH domain RBPs may directly or indirectly coordinate to ultimately regulate gene expression.

Acknowledgements

We thank the Zinovyeva lab for helpful technical discussions and assistance. We are grateful to Xantha Karp for critical reading of this manuscript and sharing reagents. We thank Erik Lundquist and Helge Grosshans for sharing reagents and strains. Some of the strains used in this study were provided by the CGC, which is funded by NIH Office of Research Infrastructure Programs (P40 OD010440). We thank Wormbase for providing the various necessary resources.

Supplemental Table 2.1. Percent abnormal phenotype and numbers of animals scored per replicate in *lsy-6* and *mir-48 mir-241* assays.

Gene	*lsy-6*				*mir-48 mir-241(nDf51)*		
Assay / RNAi	***plim-6::gfp*** **% cell fate defective (n)**	***lsy-6(ot150)*; *plim-6::gfp*** **% cell fate defective (n)**	***pcog-1::gfp*** **% uterine *cog-1::gfp* (n)**	***pcog-1::gfp; lsy-6(ot150)*** **% uterine *cog-1::gfp* (n)**	***col-19::gfp*** **% abnormal *col-19::gfp* (n)**	***mir-48 mir-241(nDf51)*** **% abnormal *col-19::gfp* (n)**	***hbl-1::gfp*** **% abnormal expression (n)**
Empty vector	0(50), 0(40)	10.4(96), 17.8(29), 21(108), 20.8(48), 10.49), 14(114), 14.6(48), 113.7(73), 14.5()	100(19), 100(22)	45(22), 44.4(27), 42.1(19)	0(3), 0(14)	8.3(12), 0(10), 10(30), 9.1(11), 7.4(27), 10(10), 10.5(19), 15.4(13), 29.4(17)	18.2(33), 22.2(18), 16.7(18)
dcr-1	0(38), 0(41)	30.5(105), 56.6(53), 35(113), 33.7(104), 30.6(88), 33(104), 35.8(67), 40(35), 30(50), 41.2(34), 34.2(76), 18(53), 34.4(64)	100(13), 94.4(36)	66.6(21), 70(20), 58.5(41), 45.8(24)	0(10), 0(15)	63.6(11), 71.4(14), 88.9(9), 89.7(29), 89.3(28), 70(30), 88.9(18), 95(12), 46.7(15), 90(20)	27.9(43), 48.1(27), 39.1(23)
fubl-3	0(43), 0(34)	18.9(53), 17.5(63), 15.87(76)	n.d. [a]	n.d.	0(16), 0(13)	64(25), 36.8(19)	14.7(34), 30(20), 27.3(22)
fubl-4	0(50), 0(31)	25.9(54), 19.2(83), 15(100), 23.1(78)	100(13), 100(33)	34.1(41), 59.1(22)	0(15), 0(15)	44.4(18), 55(22)	18.4(38), 33.3(21), 23(13)
fubl-1	0(62), 0(41)	22(100), 26.6(53), 12.7(71), 20.4(49), 11.2(169), 22.6(128)	n.d.	n.d.	0(15), 0(12)	30(11), 80(30), 12.5(24), 26.3(19)	n.d.
fubl-2	0(47), 0(46)	23.6(89), 22(150), 19.4(62), 16.9(124), 16.7(60)	n.d.	n.d.	0(15), 0(15)	29.4(17), 37.5(8), 0(10)	n.d.
imph-1	0(49), 0(46)	13.7(87), 21.7(83), 10.3(68), 17.7(147), 12.5(64), 12.9(85)	n.d.	n.d.	0(9), 0(15)	8.7(23), 10(10), 0(12)	n.d.
pes-4	0(28), 0(75)	33.3(27), 41.4(29), 25.5(55)	100(20), 100(16)	n.d.	0(12), 0(11)	20(10), 40(15), 43(14)	n.d.
T10E9.14	0(42), 0(45)	24.2(66), 20.3(59), 19.6(51)	100(25), 100(25)	66.6(18), 61.1(18)	0(13), 0(15)	65.2(23), 38.5(13)	20.5(44), 30.8(13), 54.5(22)
nova-1	0(73), 0(33)	25(28), 22.1(136), 7.6(170), 14.1(78), 13(122), 6.9(116)	n.d.	n.d.	0(12), 0(15)	13.3(15), 0(10)	n.d.
mxt-1	0(43), 2.5(40)	19(84), 14.4(125), 10(40)	n.d.	n.d.	0(14), 0(16)	41.2(17), 26.7(30), 11.8(17)	n.d.
ascc-1	0(61), 0(42)	21.8(55), 9.6(116), 15.6(96), 17.3(23)	n.d.	n.d.	0(10), 0(15)	13.3(30), 33.3(18), 7.7(13)	n.d.
akap-1	0(51), 0(37)	16.2(105), 18.5(54), 21.1(38), 13.2(205)	n.d.	n.d.	0(16), 0(8)	36.4(11), 33.4(30), 37.5(16)	n.d.
tofu-7	0(32), 0(42)	23.6(123), 35.5(45), 13.8(130), 19(84), 11.3(150)	94.1(17), 87.1(31)	70.8(24), 72.8(26)	0(13), 0(15)	56.3(16), 35.3(17), 20(20)	16.1(31), 33.3(18), 38.1(21)
C06G4.1	0(52), 0(40)	7.6(91), 16.2(68), 11.8(85), 19(58), 14(93)	n.d.	n.d.	0(10), 0(15)	11.1(27), 16.7(6)	n.d.
E02D9.1	0(40), 0(44)	21.3(47), 21.5(79), 11.8(34)	n.d.	n.d.	0(11), 0(8)	0(10), 0(16), 11.8(17)	n.d.
sfa-1	0(51), 0(36)	30(10), 33.9(56), 29.5(78), 34.2(38)	100(21), 100(25)	66.6(12), 79.3(29)	0(18), 0(15)	41.2(17), 43.8(16), 11.8(17), 78.6(28)	31(42), 42.1(19), 43.5(23)
asd-2	0(45), 0(25)	24(54), 21.2(52), 27.2(22), 31.6(57), 14.7(102), 20.5(88)	100(15), 91.7(24)	66.6(21), 64.2(28)	0(13), 0(15)	50(10), 40(20), 85.7(7), 31.3(16)	23.8(42), 50(24), 42.9(28)
B0280.17	0(47), 0(28)	19.8(86), 20.4(54), 17.5(80)	n.d.	n.d.	0(16), 6.7(15)	52.6(19), 21.4(14), 15.8(19)	n.d.
F54D1.1	6.3(32), 0(100)	26.5(83), 18.7(107), 40(55), 12.2(245), 20.2(169)	84.2(19), 100(12)	34.2(38), 56.5(23)	0(12), 21.4(14)	23.3(30), 33.3(12), 18.2(11)	21.1(38), 42.1(19), 36.3(22)
K07H8.9	0(63), 0(32)	18.8(48), 14.5(110) ,11.4(70), 12.5(56), 16.1(56)	n.d.	n.d.	0(10), 0(14)	0(13), 27(30), 5.3(19)	n.d
Y69A2AR.32	0(49), 0(41)	27(52), 20.3(64), 9(47)	n.d.	n.d.	0(14), 0(15)	44(25), 37.5(24)	19.4(36), 37(27), 40(20)
bcc-1	0(55), 0(51)	14.3(202), 8.5(82), 14.4(111), 5.6(54), 15.3(118), 16.7(78)	n.d.	n.d.	0(14), 0(13)	37(27), 9.1(11), 21.7(23), 12.5(24)	n.d.
C41G7.3	0(49), 0(33)	21.9(64), 4.5(67), 11.3(71), 23.9(46), 19.8(116)	n.d.	n.d.	0(13), 0(13)	37.5(24), 63.3(30), 35.7(14)	54.2(24), 30.6(36) 48.2(29)
pno-1	0(34), 0(40)	13.6(110), 29.2(89), 12.1(149), 16.3(153), 22.7(132)	n.d.	n.d.	0(15), 0(15)	41.4(29), 51.6(27), 27.6(29)	20(40), 25(36), 39.1(23), 26.1(23)
mask-1	0(60), 0(39)	21.2(85), 28.3(46), 25.8(62)	100(24), 100(27)	64.7(17), 56.5(23)	0(15), 0(15)	56(12), 30(20), 66(25)	16.2(31), 50(15), 33.3(24)

a n.d. - not determined

Supplemental Table 2.2. Percent abnormal phenotype and numbers of animals scored per replicate in *let-7* and *alg-1* mutant assays.

Gene	*let-7*					*alg-1*
Assay / RNAi	N2 % Bursting (n)	*let-7(n2853)* % Bursting (n)	*lin-41::gfp* Relative Fluorescence GFP/mCherry (n)	*lin-41::gfp;let-7(n2853)* Relative Fluorescence GFP/mCherry (n)	*lin-41::gfp* ΔLCS Relative Fluorescence GFP/mCherry (n)	*alg-1(ma202)* % wildtype *col-19::gfp* (n)
Empty vector	0(60), 0(27)	6.7(60), 8.3(60), 13(69), 18.3(60), 11(55), 25(60)	1.73 ± 0.44 (30)	3.88 ± 0.62 (25)	4.42 ± 0.77 (28)	0(14), 0(19)
dcr-1	0(60), 0(30)	27(59), 36.7(60), 37.3(51), 32.4(111)	1.94 ± 0.48 (30)	4.84 ± 0.98 (30)	5.28 ± 1.2 (21)	0(14), 0(12)
fubl-3	0(60), 0(38)	27.6(47), 41.7(48), 25.9(166)	n.d. [a]	n.d.	n.d.	0(17), 0(14)
fubl-4	0(60), 0(30)	20.8(48), 4(50), 25.9(27), 7.1(84)	n.d.	n.d.	n.d.	0(14), 0(18)
fubl-1	0(60), 0(32)	32.7(), 25(0, 20(), 19.1(241)	n.d.	n.d.	n.d.	0(9), 5.3(19)
fubl-2	0(60), 0(48)	14(50), 14(50), 12(50), 8.7(104)	n.d.	n.d.	n.d.	0(8), 4.5(22)
imph-1	0(60), 0(47)	12(50), 8.3(4), 9.5(42), 8.2(183)	n.d.	n.d.	n.d.	0(20), 7.7(13), 14.3(7)
pes-4	0(60), 0(36)	8.6(46), 10.4(58), 12.7(245)	n.d.	n.d.	n.d.	0(8), 0(12)
T10E9.14	0(60), 0(49)	21.7(60), 10(50)	n.d.	n.d.	n.d.	0(23), 0(12)
nova-1	0(60), 0(30)	14.3(49), 14.2(49), 4.3(47), 13.1(160)	n.d.	n.d.	n.d.	0(16), 5.9(17)
mxt-1	0(60), 0(47)	40.8(49), 34(50), 25.5(43), 13.3(181)	2.1 ± 0.44 (29)	4.98 ± 1.1 (30)	5.01 ± 0.87 (27)	0(12), 0(18)
ascc-1	0(60), 0(45)	12.2(49), 23.3(60), 10.2(167)	n.d.	n.d.	n.d.	8.3(12), 0(6)
akap-1	0(60), 0(30)	16.7(48), 10(50), 24(50), 2.1(97)	n.d.	n.d.	n.d.	14.3(21), 0(21)
tofu-7	0(60), 0(31)	25(48), 14(50), 13.4(112)	n.d.	n.d.	n.d.	0(18), 0(16)
C06G4.1	0(60), 0(33)	6.0(50), 15.8(57)	n.d.	n.d.	n.d.	0(8), 4.5(22)
E02D9.1	0(60), 0(21)	11.6(60), 8.8(80), 1.2(81)	n.d.	n.d.	n.d.	71.4(7), 45.5(18)
sfa-1	0(60), 0(33)	13.3(60), 16(50), 54(50), 18(150)	1.92 ± 0.52 (30)	4.69 ± 0.93 (30)	5.35 ± 0.96 (30)	11.1(19), 0(12), 10.5(18)
asd-2	0(60), 0(35)	21.7(60), 58.3(48), 38(50), 12.4(162)	1.83 ± 0.44 (30)	14.64 ± 0.81 (27	5.88 ± 1.03 (29)	0(19), 0(23)
B0280.17	0(60), 0(32)	25(48), 38(50)	2.05 ± 0.55 (30)	5.08 ± 1.35 (30)	4.72 ± 1.47 (30)	0(6), 0(24)
F54D1.1	0(60), 0(58)	10.4(48), 24(50), 19(58), 9.5(168)	n.d.	n.d.	n.d.	0(3), 0(26)
K07H8.9	1.8(54), 0(60)	4.1(49), 14.6(48), 14(50), 6.1(98)	n.d.	n.d.	n.d.	0(16), 5.9(17)
Y69A2AR.32	0(43), 1.6(63)	15.5(45), 10(50), 10.8(111)	n.d.	n.d.	n.d.	7.1(14), 4.8(21)
bcc-1	0(60), 0(25)	37.7(49), 38.7(49), 16.5(176)	n.d.	n.d.	n.d.	0(14), 0(12)
C41G7.3	0(60), 0(38)	8.2(49), 14(50), 26(50), 5.3(113)	n.d.	n.d.	n.d.	0(15), 4.3(23), 0(20)
pno-1	0(60), 0(37)	16(50), 12(50), 22(50), 5.9(136)	n.d.	n.d.	n.d.	0(8), 0(17)
mask-1	0(60), 0(36)	16.3(49), 12(50), 10.5(153)	n.d.	n.d.	n.d.	0(5), 0(20)

a n.d. - not determined

Supplemental Table 2.3. Spatial expression patterns of KH domain RBPs.

Gene	Tissue Transcriptome [a]	Tissue Proteome [b]	Tissue promoterome [c]
fubl-3	Hypodermis, body wall muscle, intestine	Hypodermis, intestine, muscle, neuron	n.d.
fubl-4	n.d. [d]	Hypodermis	n.d.
fubl-1	Hypodermis, body wall muscle, intestine	Hypodermis, intestine, muscle, neuron	pharynx, vulva, spermatheca, muscle, hypodermis, seam cells, neurons
fubl-2	Hypodermis, body wall muscle, intestine	Hypodermis, intestine, muscle, neuron	n.d.
imph-1	Hypodermis, body wall muscle, intestine	Hypodermis, intestine, muscle, neuron	pharynx, intestine, distal tip cell, spermatheca, hypodermis, neurons
pes-4	Intestine	Hypodermis, intestine, muscle, neuron	n.d.
T10E9.14	n.d.	n.d.	n.d.
hrpk-1	Hypodermis, body wall muscle, pharynx, intestine	Hypodermis, intestine, muscle, neuron	n.d.
nova-1	Hypodermis, body wall muscle, pharynx, intestine	Hypodermis, intestine, muscle, neuron	n.d.
vgln-1	Hypodermis, body wall muscle, pharynx, intestine	Hypodermis, intestine, muscle, neuron	n.d.
mex-3	n.d.	Hypodermis, intestine, muscle, neuron	n.d.
mxt-1	Body wall muscle	Hypodermis, intestine, muscle, neuron	n.d.
ascc-1	Hypodermis, intestine	Hypodermis, intestine, muscle, neuron	n.d.
akap-1	Hypodermis, body wall muscle, intestine	Hypodermis, intestine, muscle, neuron	n.d.
tofu-7	n.d.	Intestine	n.d.
C06G4.1	n.d.	Hypodermis, muscle	n.d.
E02D9.1	Hypodermis, body wall muscle, pharynx, intestine	Hypodermis, intestine, muscle, neuron	pharynx, intestine, muscle, vulva, hypodermis, neurons, ventral nerve cord
sfa-1	Hypodermis, intestine	Hypodermis, intestine, muscle, neuron	n.d.
asd-2	Hypodermis, body wall muscle, pharynx	Hypodermis, intestine, muscle, neuron	n.d.
gld-1	n.d.	Hypodermis, intestine, muscle, neuron	n.d.
B0280.17	n.d.	n.d.	n.d.
F54D1.1	n.d.	Intestine	n.d.
Y57G11C.36	n.d.	Hypodermis, intestine, muscle, neuron	n.d.
K07H8.9	n.d.	Hypodermis, intestine	n.d.
Y69A2AR.32	n.d.	Hypodermis, intestine	n.d.
bcc-1	Hypodermis	Hypodermis	n.d.
C41G7.3	n.d.	Intestine	n.d.
pno-1	Hypodermis	Hypodermis, intestine, muscle, neuron	n.d.
mask-1	Hypodermis, body wall muscle, pharynx, intestine	Hypodermis, intestine, muscle, neuron	pharynx, intestine, rectum, vulva, hypodermis, seam cells, neurons

a Kaletsky et al., 2018. Transcriptome analysis of adult Caenorhabditis elegans cells reveals tissue-specific gene and isoform expression

b Reinke et al., 2018. In vivo mapping of tissue- and subcellular-specific proteomes in Caenorhabditis elegans

c Hunt-Newbury et al., 2007. High-Throughput In Vivo Analysis of Gene Expression in Caenorhabditis elegans

d No data available for gene within referenced dataset

Chapter 3 - Annotation of primary microRNA transcripts using conditional depletion of Drosha

Manuscript in preparation. Bioinformatic analyses were completed by Ganesh Panzade.

Authors: Haskell, D*., Panzade, G*., Zinovyeva, A, 2021.

* co-first authorship

Abstract

microRNA (miRNAs) are small non-coding RNAs critical to the regulation of gene expression in eukaryotes. Misregulation of miRNAs themselves can result in cascading effects ultimately resulting in developmental defects or disease. However, characterizations of miRNA gene transcription and processing have been challenging due to the lack of primary miRNA transcript (pri-miRNA) annotations. To facilitate studies of miRNA regulatory mechanisms in *C. elegans* and to circumvent the problem of rapid pri-miRNA processing that has stymied efforts to accurately annotate pri-miRNAs, we utilized the auxin-induced degradation system to conditionally deplete Drosha (DRSH-1) both in somatic and germline tissues. Conditional Drosha (DRSH-1) knockdown leads to reduced pri-miRNA processing and subsequent pri-miRNA accumulation. Using high-throughput sequencing of DRSH-1 knockdown animals' RNA and subsequent bioinformatic analysis, we provide annotations for 74 previously unannotated *C. elegans* pri-miRNAs. Our analysis uncovered novel features of pri-miRNA transcripts, including high degree of overlap with both non-coding and protein coding transcripts, longer than expected average transcript length, and potential pri-miRNA isoforms. Additionally, we identified previously unannotated, Drosha (DRSH-1) dependent transcripts, which expand annotations of non-coding RNAs across the *C. elegans* genome. Overall, our findings provide annotations of *C. elegans* pri-miRNA sequences and associated features and will facilitate further studies into regulation of miRNA gene expression and miRNA biogenesis.

Introduction

microRNAs (miRNAs) are small non-coding RNAs (~22nt) that provide an essential and robust mechanism to regulate gene expression in eukaryotes. miRNAs negatively regulate gene

expression of their target transcripts by repressing the translation of target mRNAs and inducing mRNA decay (reviewed in Huntzinger and Izaurralde 2011). Across species, the diversity of mature miRNAs is staggering, with differences in repertoire size, expression patterns, and divergent targets (Guerra-Assunção and Enright 2012). However, regardless of species differences, the importance of carefully choregraphed miRNA function cannot be overstated. miRNAs are predicted to target nearly 2/3rd of human genes (reviewed in Bartel 2004). miRNA dysfunction contributes to pathologies like cancer (reviewed in Vannini *et al.* 2018), cardiovascular disease (reviewed in Peters *et al.* 2020), metabolic disorders in humans (reviewed in Krol *et al.* 2010), metabolic and developmental dysfunction in *Drosophila* (Teleman and Cohen 2006), and widespread developmental defects in *C. elegans* (Resnick *et al.* 2010). Therefore, much effort has been made to characterize factors and pathways that regulate miRNA biogenesis. However, in depth molecular characterizations of miRNA gene expression control have been significantly hampered by a lack of miRNA gene structure annotations.

Canonical biogenesis of miRNAs begins with the transcription of a primary miRNA transcript (pri-miRNA) by RNA polymerase II. Following transcription, the Microprocessor complex (containing Drosha (DRSH-1) and DGCR8/Pasha) directs the enzymatic cleavage of the pri-miRNA hairpin liberating the ~ 70 nt miRNA precursor (pre-miRNA) (Figure 3.1A). Once processed, the pre-miRNAs are exported from the nucleus to the cytoplasm via Exportin 5 (Muqbil *et al.* 2013). In the cytoplasm Dicer (DCR-1) subsequently cleaves the stem loop from the precursor, freeing the ~22 nt miRNA duplex (reviewed in Ha and Kim 2014). Argonaute protein retains a single miRNA strand (the mature miRNA) from the duplex, ejecting the other strand, referred to as passenger strand or miRNA*. Argonaute associates with accessory proteins to form the fully functional miRNA-Induced Silencing Complex (miRISC) and goes on to target

mRNA transcripts for translation repression or mRNA degradation (reviewed in Bartel 2018; Figure 3.1A).

While many gene transcripts are usually easily annotated with help of high-throughput RNA sequencing (RNAseq), primary miRNA transcript annotations have lagged behind due to lack of pri-miRNA read representation in standard RNAseq datasets. This is primarily due to the speed at which pri-miRNAs are processed, in part because the Microprocessor complex is efficiently recruited to pri-miRNAs, co-transcriptionally or post-transcriptionally (Feng *et al.* 2011; Conrad *et al.* 2014; Louloupi *et al.* 2017). This rapid recruitment encourages expedient processing, in some cases before the transcript has even been fully transcribed, capped and polyadenylated (Ballarino *et al.* 2009). Human pri-miRNAs were recently annotated, having first shown that primary miRNA transcript annotations cannot be obtained using existing RNAseq datasets (Chang *et al.* 2015). This study utilized a dominant-negative human Drosha to disrupt pri-miRNA processing, allowing for computational annotation and validation (by RACE) of the accumulated pri-miRNAs. In *C. elegans,* CAPseq approaches provided presumptive transcription start sites for many, although not all *C. elegans* miRNAs (Gu *et al.* 2012) However, we still lack full annotations of C. *elegans* primary miRNA transcripts and basic features such as length of primary transcripts remain unknown. In addition, a large proportion of *C. elegans* miRNA loci are located within introns of protein coding genes (Martinez *et al.* 2008) and the genetic loci harboring miRNA precursors can be complex, further demonstrating the need for careful primary miRNA transcript annotations. In fact, only two *C. elegans* pri-miRNAs have been previously annotated: *let-7* and *mir-1899,* in part due to transcript reads being present in standard RNAseq datasets and in part due to in depth investigations of *let-7* locus (Mondol *et al.* 2015; Wormbase). To characterize pri-miRNA transcripts in *C. elegans,* we generated a system to conditionally

knockdown Drosha (DRSH-1) using the auxin-induced degron system (Zhang *et al.* 2015). This system allowed us to deplete Drosha (DRSH-1) and prevent efficient processing of pri-miRNAs, leading to their accumulation and detection through standard high-throughput sequencing of the transcriptome. Here, we provide annotations for a total of 74 primary miRNA transcripts. Surprisingly, we identified additional genomic loci that accumulated reads upon DRSH-1 degradation, suggesting the existence of additional non-coding RNAs that that were previously overlooked, perhaps due to rapid Drosha (DRSH-1)-mediated processing. Overall, this work will facilitate future studies of regulation of miRNA expression and miRNA processing.

Results

Auxin induced degron system allows for robust knockdown of DRSH-1::AID.

Primary miRNA transcripts were fully annotated using a degradation system to conditionally deplete Drosha (DRSH-1) protein, circumventing the sterility and lethality associated with *drsh-1* loss (Denli *et al.* 2004). To this purpose, we used CRISPR/Cas9 genome editing to endogenously tag *drsh-1* with an Auxin induced degron (AID) tag to degrade DRSH-1 protein and thereby block the rapid pri-miRNA processing that would normally preclude identification of these transcripts in standard RNAseq libraries (Figure 3.1B). The auxin-inducible degradation system takes advantage of the *Arabidopsis thaliana* gene transport inhibitor response 1 (*AtTIR1*), which in the presence of exogenous auxin (Indole-3-acetic acid; IAA), targets AID-tagged proteins for degradation by native proteosomes. This system was adapted to *C. elegans* by ubiquitously expressing TIR-1 (here in the somatic tissues only), which then acts as the Auxin (IAA) sensor responsible for triggering proteasomal degradation of the target protein (Zhang *et al.* 2015). Here, we tagged the DRSH-1 protein with the **A**uxin **I**nducible

Degron (AID) which consists of degron:linker::TEV::3xFLAG residues, allowing DRSH-1 to be targeted by Auxin-activated TIR-1. Subjecting *drsh-1::AID::linker::TEV::3xFLAG (zen80); Peft-3::tir-1::mRuby:: unc-54 3'UTR* larvae to 1 hour, 2 hour and 4 hour long treatments with Auxin (IAA) induced rapid and robust degradation of DRSH-1 (Figure 3.1C). Quantification of DRSH-1 knockdown using 3X FLAG tag showed an average of 78%, 42%, and 89% knockdown when compared to the control (no Auxin) and normalized against tubulin in the 1-hour, 2-hour, and 4-hour timepoints, respectively (Figure 3.1D). The 89% knockdown upon 4-hour auxin exposure suggests that the auxin induced degron system is an effective method to conditionally deplete DRSH-1.

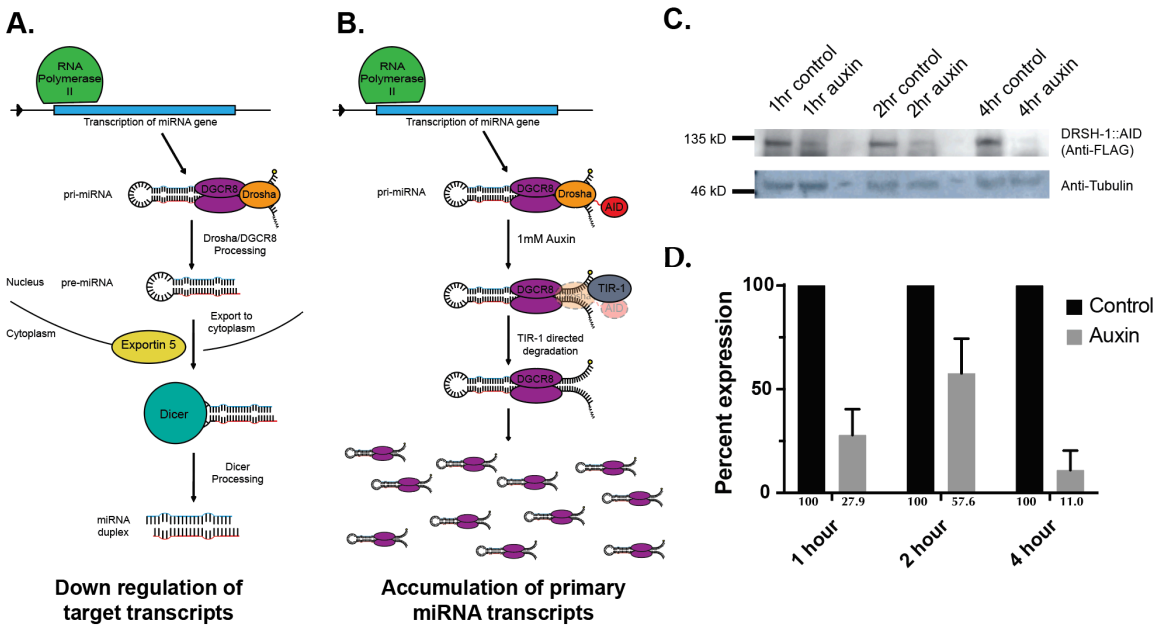


Figure 3.1. Auxin induced degron (AID) leads to strong reduction of DRSH-1, allowing for accumulation of primary miRNA transcripts (pri-miRNAs).

(A) miRNA intermediates are enzymatically processed by Drosha (DRSH-1) and Dicer (DCR-1) giving rise to mature miRNAs. **(B)** The addition of 1mM auxin (IAA) in a *eft-3::tir-1* background induces the degradation of DRSH-1::AID, decreasing levels of pri-miRNA processing and the subsequent accumulation of pri-miRNAs. **(C)** A representative western blot showing depletion of DRSH-1::AID at various timepoints. A 4-hour 1 mM auxin treatment induces the greatest DRSH-1 depletion and was used for subsequent RNAseq. **(D)** Quantification of DRSH-1 depletion showed a ~90% knockdown of DRSH-1 after 4 hours of auxin exposure.

Degron-dependent depletion of Drosha results in accumulated primary miRNA transcripts.

Given the amount of DRSH-1 knockdown, we reasoned that the 4-hour auxin exposure would be sufficient to result in the accumulation of pri-miRNAs and subjected the RNA extracted from these animals to high-throughput RNAseq using the poly(A) selection method for library preparation. After generating 96,128,590 paired-end reads from three replicated DRSH-1 depletion libraries and 62,694,721 paired-end reads from two replicated control libraries, we set out to generate a genome-wide map of *C. elegans* primary miRNA transcripts. Overall, reads overlapping precursor miRNA were obtained for 74 (30%) of miRNA loci. To generate primary miRNA annotations, we evaluated transcriptome assemblers *StringTie* (Kovaka *et al.* 2019)) and TRINITY (Grabherr *et al.* 2011). Manual inspection of assembled transcripts using *JBrowse* (Buels *et al.* 2016) determined that TRINITY provided assemblies that most accurately corresponded to the loci read coverage. RNAseq data and the associated new transcript assemblies are available as GFF3/BAM/FASTA files and can be visualized in standard genome browsers such as *JBrowse* (Buels *et al* 2016) or IGV (Robinson *et al.* 2011).

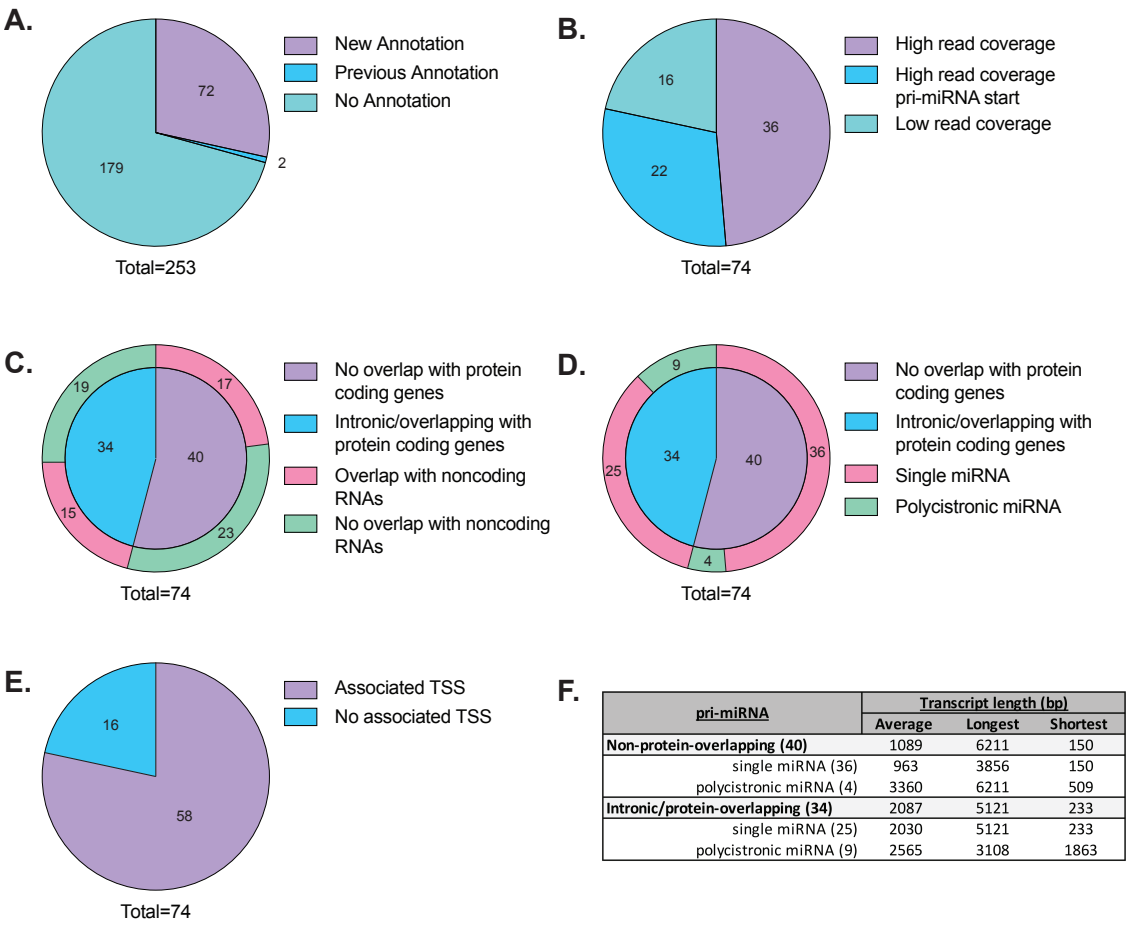
Characterization of primary miRNA annotations

A total of 74 pri-miRNA transcript assemblies were generated, adding 72 miRNAs to the previously known annotations of *let-7* and *mir-18*99 (Figure 3.2A). Of 74 transcripts, 36 had high read coverage with clearly defined transcript start and end (Figure 3.2B). Twenty-two had high read coverage, a clearly defined transcript start, and a transcript end that overlapped with other genomic features, most frequently protein-coding genes, which prevented high-confidence transcript end annotations (Figure 3.2B). Sixteen transcripts had lower read coverage, which was

nonetheless sufficient for transcript annotation (Figure 3.2B). However, we cannot exclude the possibility that these annotations represent partial pri-miRNA transcripts.

Of the 74 pri-miRNA transcripts we provided, 40 do not overlap with protein coding genes, although amongst those ~43% (17) do overlap with a least one noncoding RNAs (Figure 3.2C). The other 34 transcripts represent intronic/protein-overlapping pri-miRNAs, where the primary transcript overlaps with a protein coding region in some way, whether it be a UTR, exon, intron, or some combination therein (Supplemental Table 3.1). Among those 34 intronic/overlapping pri-miRNAs ~44% also overlap with noncoding RNAs (Figure 3.2C, Supplemental Tables 3.2 and 3.3). Overall, the majority of these overlapping noncoding elements were ncRNAS, however lincRNAs, tRNAs, snoRNAs, piRNAs, and snRNAs are also represented (Supplemental Tables 3.2 and 3.3). In addition, we found that 10% of non-protein-overlapping pri-miRNAs and 26.5% of intronic/protein-overlapping pri-miRNAs are contained withing polycistronic clusters (Figure 3.2D). Furthermore, 58 (78%) of our newly annotated pri-miRNA loci had corresponding Transcription Start Site (TSS) as identified by CAPseq (Gu *et al.* 2012; Figure 3.2E), while 16 (22%) did not have a previously identified TSS (Figure 3.2E). Lastly, we characterized transcript length for both intragenic and intergenic pri-miRNAs and found that on average intronic/protein-overlapping pri-miRNAs are longer (Figure 3.2F). We found that many pri-miRNA loci produce surprisingly long transcripts, with the average length of non-protein-overlapping transcripts of 1089 bp and average length of intronic/protein-overlapping transcripts of 2087bp (Figure 3.2F). Interestingly, several miRNA loci produced transcripts longer than 5000bp, with the longest intronic/protein-overlapping transcript (*pri-mir-71*) encodes a single miRNA rather than a polycistronic cluster (Supplemental Table 3.4).

Polycistronic pri-miRNAs were on average longer than single miRNA transcripts (Figure 3.2F, Supplemental Table 3.4).



F.

pri-miRNA	Transcript length (bp)		
	Average	Longest	Shortest
Non-protein-overlapping (40)	1089	6211	150
single miRNA (36)	963	3856	150
polycistronic miRNA (4)	3360	6211	509
Intronic/protein-overlapping (34)	2087	5121	233
single miRNA (25)	2030	5121	233
polycistronic miRNA (9)	2565	3108	1863

Figure 3.2. Characteristics of pri-miRNA loci.

(A) Of 253 miRNAs (miRbase), we provide 74 new pri-miRNA annotations, adding to the 2 previously annotated loci. **(B)** "High read coverage" annotations had good read coverage and a well-defined start and end. " High coverage pri-miRNA start" annotations had good read coverage and a well-defined start; however the end of the transcript overlapped with other genomic features. "Low read coverage" pri-miRNA annotations had low read coverage or overlapped with genomic features that complicated annotations. **(C)** Proportions of *C. elegans* newly annotated pri-miRNAs overlapping with protein-coding or non-coding RNA genomic elements. **(D)** Proportions of single and polycistronic miRNA loci within pri-miRNA that do not overlap with protein coding genes vs those that are

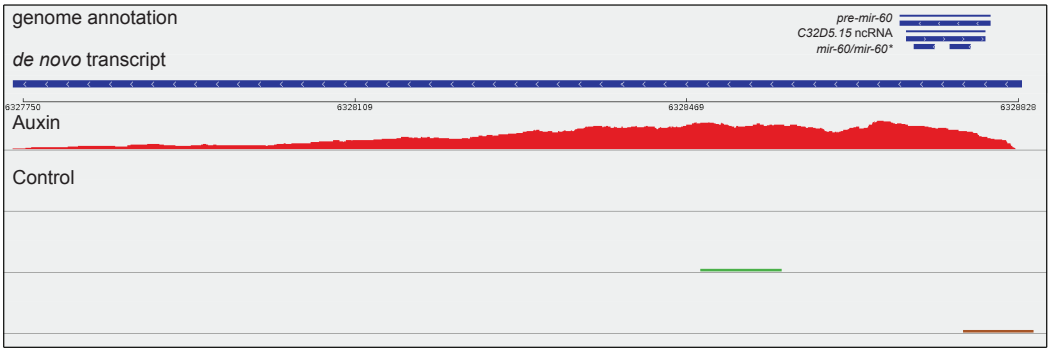
intronic/overlap with protein coding genes. **(E)** Proportion of our pri-miRNA annotations that had or lacked a previously identified transcription start site (TSS) (Wu et al., 2015). **(F)** Length of newly annotated pri-miRNA transcripts length that do not overlap with protein coding genes (40) or those that are intronic/overlap with protein coding genes (34).

miRNA gene structures

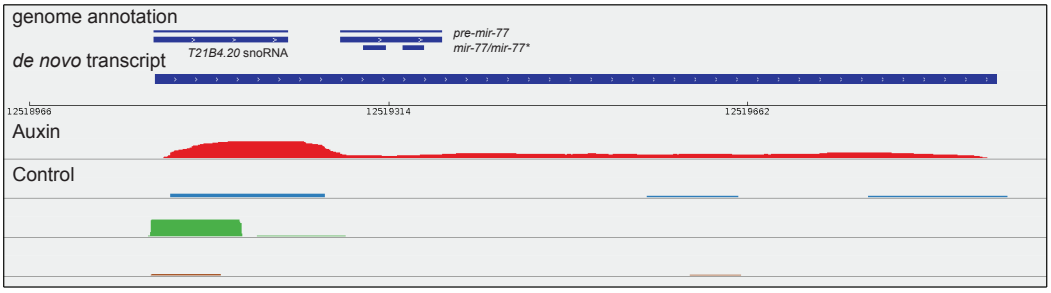
Of the 74 pri-miRNAs we annotated, 40 primary transcripts are non-overlapping with protein-coding loci (Supplemental Table 3.3). *pri-mir-60* and *pri-mir-77* are two examples of high coverage miRNAs that do not overlap with protein-coding genes (Figure 3.3). *mir-60* locus generates 1091 bp long primary transcript and *mir-77* generating an 817 bp long primary transcript (Figure 3.3, Supplemental Table 3.3), with reads accumulating in auxin-treated samples. Both primary transcripts overlap with non-coding RNAs; *C32D5.15* ncRNA in case of *mir-60* (Figure 3.3A) and *T21B4.20* snoRNA in case of *mir-77* (Figure 3.3B). *Pri-mir-77* and *T21B4.20* may be transcribed as a single or separate transcripts, although both appear to be processed by DRSH-1, as DRSH-1-depleted samples accumulate reads overlapping both loci (Figure 3.3B, Supplemental Table 3.3).

The remaining 34 pri-miRNAs we annotated are intronic/overlapping with protein coding genes (Supplemental Table 3.3). Many of the miRNA loci within this category are located within gene intronic regions, however a large percentage (76%) of pri-miRNAs in this category also overlapped with exons (Supplemental Table 3.1). For example, both *mir-67* (Figure 3C) and *mir-90* (Figure 3.3D) loci are intronic to their host genes yet appear to be transcribed independently of their host genes. In each case, transcription start sites correspond to the beginning of the read assemblies (Figure 3.3C, D). However, in each case, the end of the pri-miRNA transcript overlaps with an exon of its host gene (Figure 3.3 C, D), making it difficult to determine the absolute end point of each pri-miRNA transcript without further experimental validation.

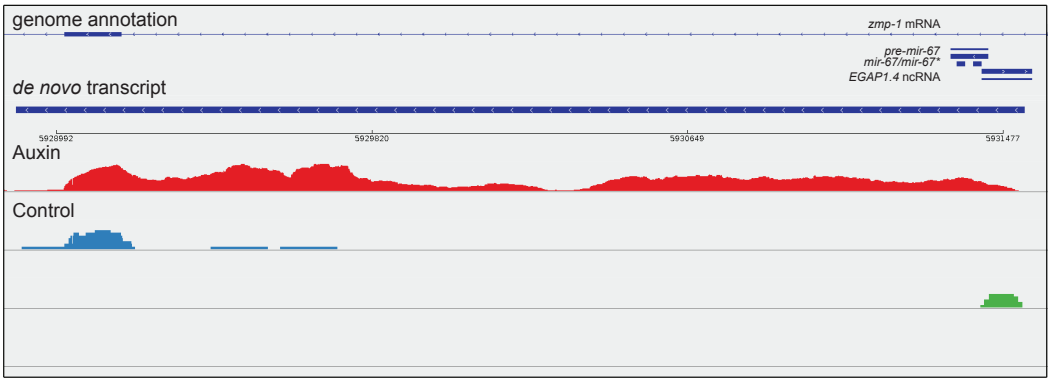
A. ***pri-mir-60***



B. ***pri-mir-77***



C. ***pri-mir-67***



D. ***pri-mir-90***

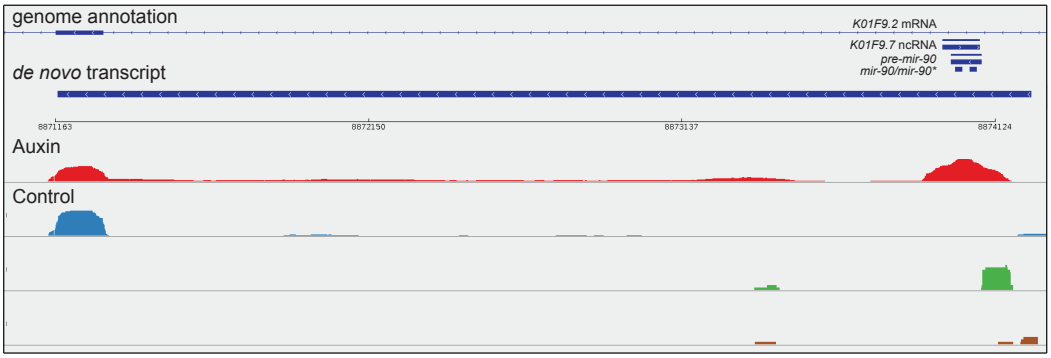


Figure 3.3. Examples of pri-miRNAs encoding single miRNA.

(A) pri-*mir-60* reads do not overlap with any protein coding genes and are accumulated in DRSH-1 depleted samples compared to control. **(B)** pri-*mir-77* and the adjacent snoRNA might be produced from a single non-protein-overlapping transcript, as DRSH-1-processed reads covering both loci accumulate in auxin-treated animals. **(C)** pri-*mir-67* is an intronic pri-miRNAs that overlaps with a *zmp-1* exon and adjacent *EGAP1.4* ncRNA. **(D)** pri-*mir-90* is an intronic pri-miRNA that appears to be transcribed independently of the *K01F9.2* mRNA. IGV-based view of normalized (TPM) read coverage in DRSH-1 depleted samples is shown in red, while control read coverage is shown in light blue. Transcription start sites (TSS), as previously determined by Gu *et al*. 2012, are included for reference and shown in green (average reads across all developmental stages) and brown (reads in young adult animals). Longest transcript assembly is shown (*de novo* transcript).

Putative polycistronic loci are transcribed as single units.

From our data we were able to generate transcript assemblies for five polycistronic clusters (two non-protein-overlapping, three intronic/protein-overlapping) encompassing a total of 13 miRNAs. *mir-229* and *mir-240* clusters represent two examples of the intronic/protein-overlapping polycistronic pri-miRNAs (Figure 3.4). *pri-mir-229* transcription begins in an intron of host gene *gcn-1,* supported by the presence of a TSS, generates a 2725 bp transcript, and end overlapping with a *gcn-1* exon (Figure 3.4A). Interestingly, *pri-mir-240* cluster is transcribed as an 1863 bp transcript that overlaps the 5'UTR and an exon of the adjacent *ppyt-2.2* protein coding gene (Figure 3.4B), although we cannot discount the possibility that a longer transcript exists. Unfortunately, due to significant overlap with *ppyt-2.2,* we cannot confidently annotate a true end, therefore our annotation represents the *de novo* transcript that most accurately reflects the end of the read accumulation (Figure 3.4B).

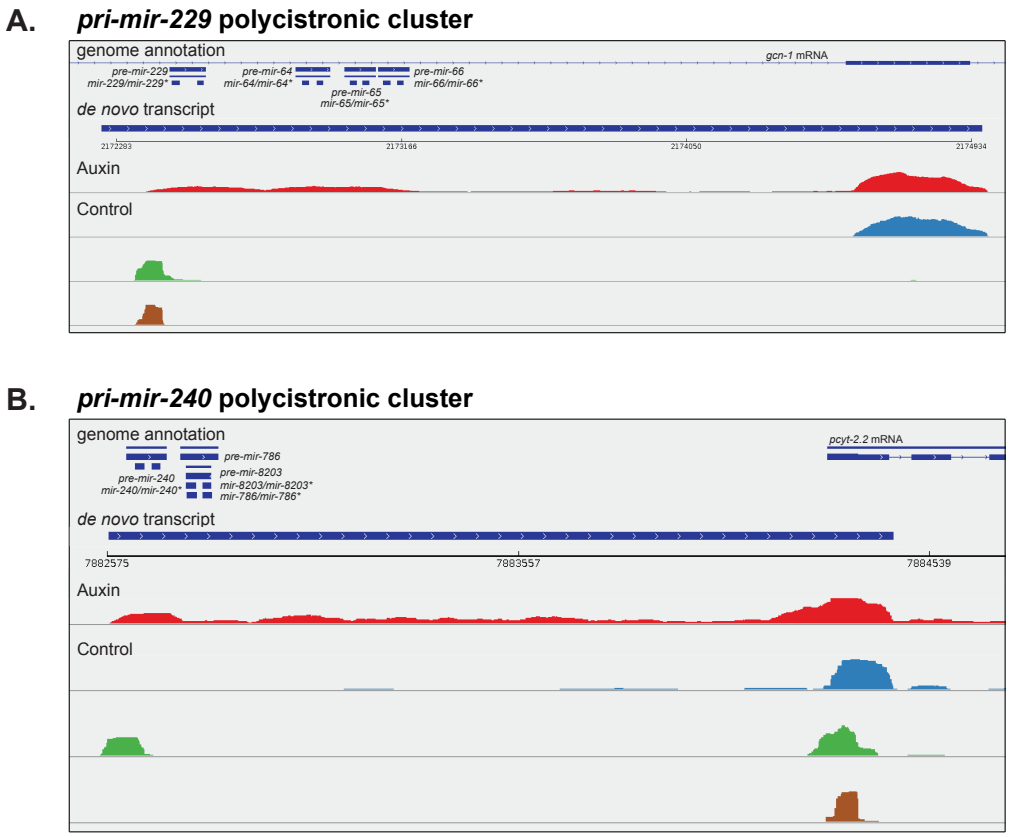


Figure 3.4. Examples of polycistronic pri-miRNA assemblies.

(A) pri-*mir-229* cluster transcript overlaps with *gcn-1*. **(B)** pri-*mir-240* cluster transcript overlaps with the 5' UTR and 1st exon of *pcyt-2.2*. IGV-based view of normalized (TPM) read coverage in DRSH-1 depleted samples is shown in red, while control read coverage is shown in light blue. Transcription start sites (TSS), as previously determined by Gu *et al*. 2012, are included for reference and shown in green (average reads across all developmental stages) and brown (reads in young adult animals).

Drosha depletion results in read accumulation at previously unannotated genomic loci.

Surprisingly, depletion of Drosha also led to accumulation of reads in previously unannotated regions of the genome (Supplemental Table 3.5). More than 300 transcripts were assembled using reads accumulated in DRSH-1 depleted samples, covering loci entirely devoid of reads in control (Supplemental Table 3.5). Figure 3.5 shows examples of previously unannotated transcripts (Figure 3.5). Two novel transcript assemblies were generated for loci adjacent to *pqn-42* (Figure 3.5A). Furthermore, two new transcript assemblies likewise appear to be fully independent of the adjacent *C30G12.11* ncRNA (Figure 3.5B). Importantly, both transcripts have nearby TSSs, providing further support for existence of these transcripts (Figure 3.5B). It is important to note that the generated libraries do not allow us to determine the strandedness of the newly assembled transcripts, which will need to be determined with future experiments using either stranded RNAseq or RACE experiments. In addition, the presence of some reads in the control samples may be due to leakiness of the *drsh-1::AID; tir-1* construct, as *tir-1* has been shown to be activated in the absence if IAA/auxin, likely due to naturally occurring bacterial indole (Hills-Muckey *et al.* 2021). In addition, novel transcripts overlapping with previously annotated ncRNA loci were obtained. The *pri-mir-1* adjacent locus shows Drosha-depleted dependent accumulation and a new transcript assembly overlapping the previously-annotated *T09B4.16* ncRNA (Figure 3.5C).

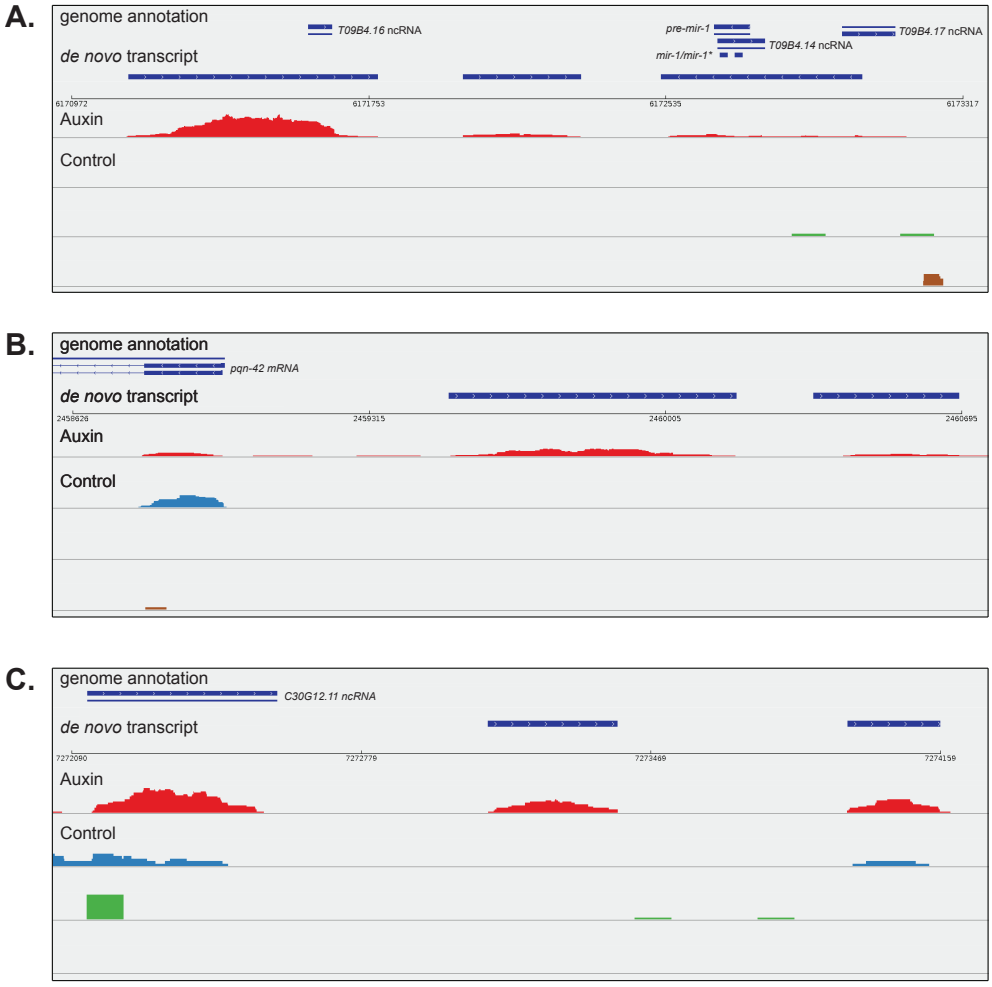


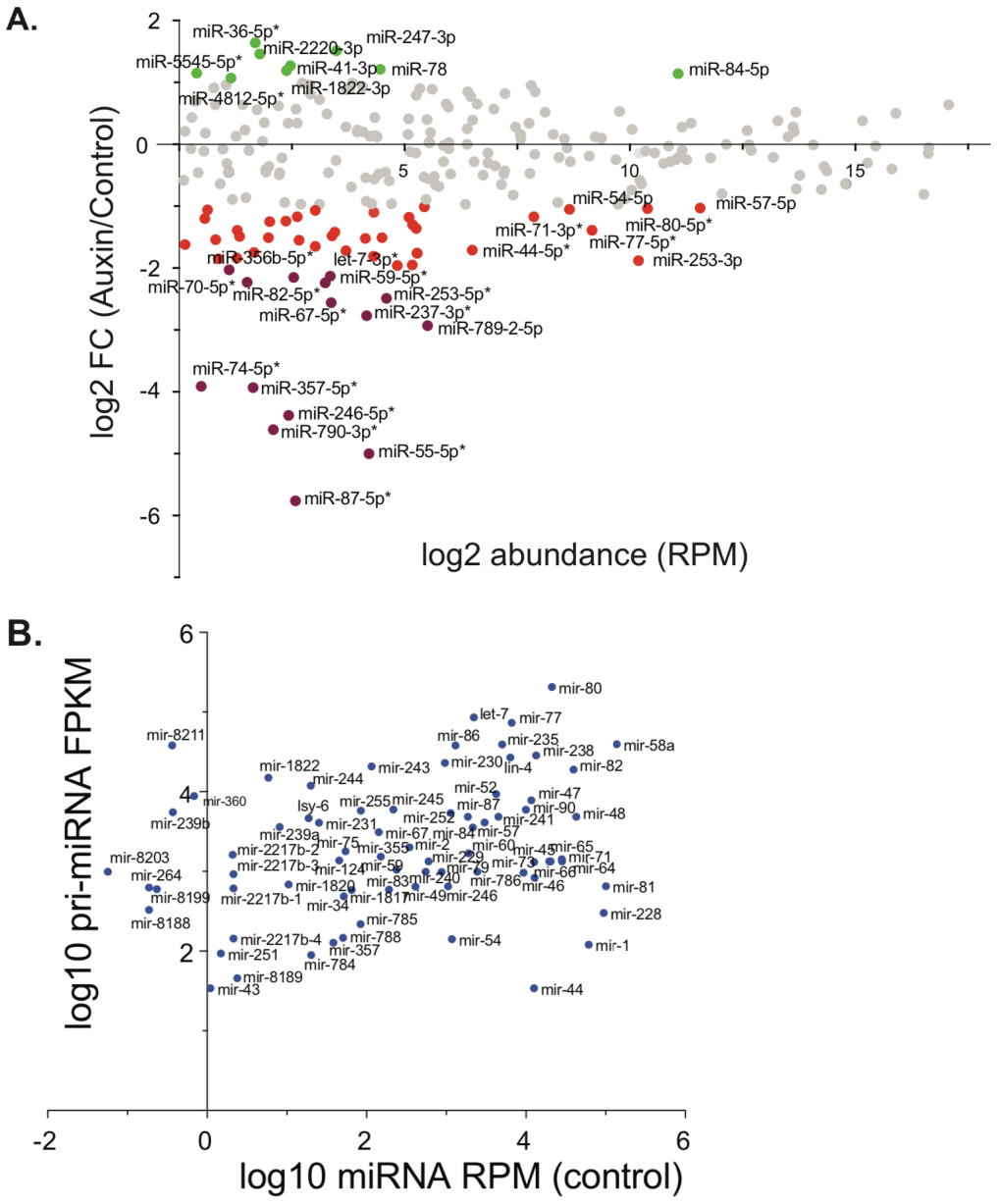
Figure 3.5. Annotation of uncharacterized Drosha-dependent genomic elements. In addition to pri-miRNAs, the depletion of Drosha results in a number of genomic regions with accumulated reads. **(A)** The *mir-1* locus contains a two additional Drosha dependent transcripts, one of which is previously unannotated and the other corresponding to the *T09B4.16* ncRNA. **(B)** Two previously unannotated DRSH-1 dependent transcripts adjacent to the coding gene *pqn-42* at the coordinates II:2459502-2460172 and II:2460352-2460690. **(C)** Two previously unannotated DRSH-1 dependent transcripts adjacent to *C30G12.11* ncRNA at the coordinates II:7273083-7273391 and II:7273939-7274159. IGV-based view of normalized (TPM) read coverage in DRSH-1 depleted samples is shown in red, while control read

coverage is shown in light blue. Transcription start sites (TSS), as previously determined by Gu et al., 2012, are included for reference and shown in green (average reads across all developmental stages) and brown (reads in young adult animals). Longest transcript assembly is shown (*de novo* transcript).

Primary miRNA read accumulation upon DRSH-1 depletion correlates poorly with mature miRNA abundances.

Mature miRNAs have been shown be very stable, with an average half-life of mature miRNAs across species that varies from hours to days (Marzi *et al.* 2016; Kingston and Bartel 2019; Reichholf *et al.* 2019). In contrast, miRNA* strands are significantly less stable, with an average half-life measuring in hours (Marzi *et al*. 2016). To explore the effects of short-term DRSH-1 depletion on mature miRNA stability, we performed small RNAseq on auxin-treated and control samples. We found that majority of mature miRNAs had no change in abundance (Figure 3.6A). Nine miRNAs (6 miRNA, 3 miRNA*) had ~ 2-fold upregulation, a surprising result considering we would expect loss of Drosha processing to ultimately deplete the mature miRNA repertoire. This relatively small upregulation could reflect natural variation among samples, rather than dependency on DRSH-1 depletion, or may be an indirect effect of DRSH-1 depletion. Fifteen miRNAs (14 miRNA*s, 1 miRNA) showed a greater than 4-fold reduction (Figure 3.6A). Thirty-six miRNAs were reduced between 2-4-fold, with 13 miRNAs and 23 miRNA* strands affected (Figure 3.6A). Given the modest changes in miRNA levels, affecting primarily miRNA* strands, these observations are consistent with the relative stabilities of mature miRNAs and miRNA* strands and demonstrate that miRNA*s may be more sensitive to disruptions in pri-miRNA processing as they are degraded more rapidly. Interestingly, short-term depletion of DRSH-1 resulted in very few changes in the overall transcriptome (Supplemental Figure 3.1). This is consistent with modest effects on mature miRNA abundances, as much of the genome-wide changes might be expected to occur as a result of mature miRNA reductions. Transcripts affected by the short-term drsh-1 depletion could represent normally Drosha-processed, previously annotated ncRNAs.

Next, we wished to examine the relationship between the abundance of accumulated of the pri-miRNA transcripts (FPKM, as a rough approximation of transcription rate) and the abundance of mature miRNAs typically present at the equivalent stage, as assessed by small RNAseq in control samples (Figure 3.6B). We found that some of the miRNAs showed pri-miRNA accumulation, despite the low levels of mature miRNAs (Figure 3.6B). These miRNAs may therefore be regulated at the level of pri-miRNA or pre-miRNA processing, or, perhaps, mature miRNA stability, ultimately resulting in lower mature miRNA accumulation despite the production of the primary transcripts. In contrast, we also saw a small proportion with high mature miRNAs abundance despite low pri-miRNA accumulation (Figure 3.6B). This relationship could reflect the stability of mature miRNAs, combined with continuous transcription during earlier developmental stages, which could ultimately result in accumulation of mature miRNAs. While miRNA biogenesis occurs with a 1:1 ratio of pri-miRNA to mature miRNA, we cannot assume a 1:1 pri-miRNA FPKM to mature miRNA RPM ratio using our data. However, given the wide variability in pri-miRNA to mature miRNA ratios across the 74 miRNAs for which primary transcript assemblies were obtained (Figure 3.6C), we can conclude that a combination of variable transcription rates, post-transcriptional pri- and pre-miRNA processing, and variable stability of mature miRNA may all ultimately dictate mature miRNA abundance (Figure 3.6).


A.
log2 FC (Auxin/Control)
log2 abundance (RPM)
miR-36-5p*
miR-247-3p
miR-2220-3p
miR-5545-5p*
miR-41-3p
miR-78
miR-84-5p
miR-1822-3p
miR-4812-5p*
miR-54-5p
miR-57-5p
miR-71-3p*
miR-80-5p*
miR-77-5p*
miR-44-5p*
miR-253-3p
miR-356b-5p*
let-7-3p*
miR-59-5p*
miR-70-5p*
miR-82-5p*
miR-253-5p*
miR-67-5p*
miR-237-3p*
miR-789-2-5p
miR-74-5p*
miR-357-5p*
miR-246-5p*
miR-790-3p*
miR-55-5p*
miR-87-5p*
B.
log10 pri-miRNA FPKM
log10 miRNA RPM (control)
mir-80
let-7
mir-77
mir-8211
mir-86
mir-235
mir-58a
mir-238
mir-1822
mir-243
mir-230
lin-4
mir-82
mir-244
mir-360
mir-52
mir-47
lsy-6
mir-255
mir-245
mir-87
mir-90
mir-48
mir-239b
mir-252
mir-241
mir-231
mir-239a
mir-67
mir-84
mir-57
mir-75
mir-2217b-2
mir-355
mir-2
mir-60
mir-45
mir-65
mir-8203
mir-229
mir-71
mir-2217b-3
mir-124
mir-59
mir-79
mir-73
mir-64
mir-264
mir-1820
mir-83
mir-240
mir-786
mir-66
mir-46
mir-81
mir-8199
mir-2217b-1
mir-1817
mir-49
mir-246
mir-34
mir-228
mir-8188
mir-785
mir-2217b-4
mir-788
mir-54
mir-1
mir-251
mir-357
mir-784
mir-8189
mir-43
mir-44

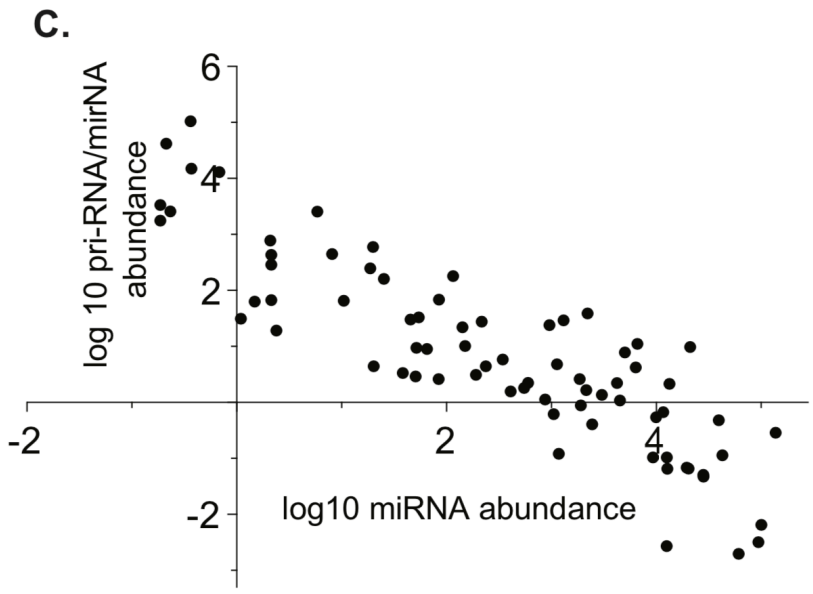


Figure 3.6. DRSH-1 depletion induces limited change to mature miRNA abundances and reveals poor correlation between pri-miRNA accumulation and mature miRNA abundance.

(A) MA Plot showing log2 fold change (FC) in mature miRNAs between Auxin-treated and control samples (Y-axis) and mature miRNA (RPM) abundance on the X-axis. Green dots represent miRNAs upregulated > 2fold. Red dots represent miRNAs 2-4 fold downregulated, while burgundy dots represent those > 4 fold downregulated. Dots are labeled with corresponding miRNA or miRNA*. **(B)** Plot showing relationship between pri-miRNA accumulation (FPKM) (Y-axis) and mature miRNA abundance (RPM) (X-axis). **(C)** MA plot showing ratio between pri-miRNA (FPKM)/ mature miRNA (RPM) abundances (Y-axis) and mature miRNA abundance (RPM) (X-axis).

Materials and Methods

Generation of DRSH-1::AID::LINKER::TEV::3xFLAG

We utilized the minimal Auxin-induced degron (AID) sequence previously described (Zhang et al., 2015), adding ~150 bp overhangs extending upstream and downstream of *drsh-1* C-terminal insertion site to generate a C-terminal gene block donor (IDT) (Table 3.1). The dsDNA donor was prepared using primers with 5' SP9 modifications as previously described (Ghanta & Mello, 2020). The injection mix was prepared as follows: Cas9 (IDT): 0.5uL, tracRNA (IDT): 5uL (0.4 ug/uL stock), *dpy-10* crRNA (IDT): 1.4 uL (0.4 ug/uL stock), *drsh-1* C-terminal crRNA (IDT): 1.8 uL (0.4 ug/uL stock), annealing buffer (IDT): 2 uL, H20: 7.1 uL, dsDNA donor (IDT): 2.2 uL (50 ng/uL). The co-injection marker *dpy-10* was used to screen for successful edits as previously described (Arribere et al., 2014). For complete list of donor, crRNA, and primer sequences, see Table 3.1. *drsh-1::AID::linker::TEV::3xFLAG (zen80*) (UY224) animals were then crossed into the P*eft-3*::*tir-1::mRuby::unc-54* UTR (CA1200) background, which expresses TIR-1 exclusively in somatic cells, generating *drsh-1::AID::linker::TEV::*3xFLAG *(zen80);* P*eft-3*::*tir-1::mRuby::unc-54* UTR (UY251).

Table 3.1. Donor, crRNA, primer sequences

DRSH-1 *C*-terminal donor (AID::linker::TEV::3xFLAG)	AAATAATAGTGTTAGTTTTATTTTTAGCAAAATATAT TTCTTACACCTTTTCAAGTGGTTTCAGAACATGCGCC GTCGTCTTGAACAAGATACCAGCGACGGATCCGGAG GTGGCGGGCCTAAAGATCCAGCCAAACCTCCGGCCA AGGCACAAGTTGTGGGATGGCCACCGGTGAGATCAT ACCGGAAGAACGTGATGGTTTCCTGCCAAAAATCAA GCGGTGGCCCGGAGGCGGCGGCGTTCGTGAAGGAGA ATCTGTACTTTCAATCCGGAAAGGACTACAAAGACCA TGACGGTGATTATAAAGATCATGATATCGATTACAAG GATGACGATGACAAGGGATCCTAATTACGGGGTTAT AATTATACTATGTCTGTTTGAATGTGATTCGGTTCAAT TTATGAATATCATATCTTTATTTTAAGTATGTT
drsh-1 C-terminal crRNA	GATACCAGCGACTAATTACGGGG
drsh-1 C-terminal SP9 forward *drsh-1* C- terminal SP9 reverse	5' GCAAAATATATTTCTTACACCTTTTC 3' 5' GAATATCATATCTTTATTTTAAGTATG 3'
drsh-1 screening forward *drsh-1* screening reverse degron internal forward	5' CATCCGCTGCTGAGTCGAAC 3' 5' CGACTGATCCAAAGGACATGG 3' 5' GTGAGATCATACCGGAAGAAC 3'

Auxin exposure experiments

Animals were grown on NGM media at 20ºC using standard protocols (Brenner 1974). For auxin treatments, NGM media was prepared as previously described and supplemented with Auxin (IAA, Indole-3-acetic acid) in 95% ethanol to a final concentration of 1 mM. Control plates were supplemented with the equivalent amount of ethanol with no auxin. *drsh-1*::AID::*linker::TEV::3xFLAG* (*zen80*); P*eft-3*::*tir-1::mRuby::unc-54* worms were synchronized using bleaching as previously described (Stiernagle 2006; Wormbook) and allowed to grow to early L4 stage on regular NGM, at which point they were washed off normal NGM, washed 3x M9 and 1x with water before being plated on 1mM auxin (IAA) or control plates. At each timepoint (1 hr, 2 hr, or 4 hr), worms were washed off of the plates, and washed three times with M9 and once with water, flash frozen as packed pellets, and stored at -80C. Three replicates were collected for auxin (IAA)-treated animals, and 2 replicates for control animals (normal NGM media).

RNA Preparation and sequencing

RNA preparations were completed as previously described (Li et al. 2019). Total RNA was poly(A) selected and libraries were prepared using NEBNext Ultra II RNA Library Prep Kit for Illumina by Genewiz. Libraries were sequenced using an Illumina Hiseq 2500 instrument (Genewiz Next Generation Sequencing Core). For small RNA libraries preparation, small RNAs were first size selected by gel purification as previously described (Gu et al., 2012). The size selected RNA was used to construct small RNA libraries using the NEXTflex Small RNA Library Prep kit v3 (Bioo Scientific) according to manufacturer's instructions and sequenced on the Illumina NextSeq500 instrument at the Kansas State University Integrated Genomics Facility.

Western Blot and Quantification

Protein preparation was completed as previously described (Li and Zinovyeva 2020). Protein concentrations were determined using Biorad Protein quantification kit (Cat #5000119). Equivalent amounts of protein for each timepoint pair were loaded. The membrane was then cut and probed for DRSH-1::AID::LINKER::TEV::3xFLAG, utilizing mouse M2 anti-FLAG (Sigma F3615) 1:500 and anti-mouse 1:5000. Bands were detected using HRP-based chemiluminescence (LI-COR WesternSure Lot #VH311910) and imaged using Azure Biosystems c600. Quantification of signal was completed using ImageJ (Schneider et al., 2012).

Quality check and filtering RNA-seq samples

RNA-seq 150 bp paired end (PE) reads were generated for control (2 replicates) and auxin-treated (3 replicates) samples. Raw reads were assessed for quality using *fastqc* (https://bioinformatics.babraham.ac.uk/projects/fastqc/). Reads filtering and adapter trimming were done using the *trimmomatic* tool (Bolger *et al.* 2014).

Genomic alignment and coverage calculations

The quality-filtered reads were mapped to *C. elegans* reference genome (WS279) using *bowtie2* (Langmead and Salzberg 2012). The mapped reads were sorted using *samtools* v1.13 (Li *et al.* 2009). *StringTie* v2.1.4 (Kovaka *et al.* 2019) was used for reference-based assembly. Genome-wide per-base coverage RPKM normalization was performed for mapped reads using *bamCoverage* utility (Ramírez *et al.* 2016) .

de novo transcript assembly

Genome-guided *de novo* transcriptome assembly was constructed using the TRINITY pipeline (Grabherr *et al*. 2011; Haas *et al*. 2013). After assembly, *de novo* transcripts were annotated to reference genome using PASA pipeline (Haas *et al.* 2008). Quantification of transcript expression was calculated by first generating the index of pri-miRNA sequences and mapping filtered reads to this index using *Salmon* aligner (Patro *et al.* 2017). TPM values were converted to FPKM utilizing the *neuMatidx* package in R (He and Yu 2018). To better visualize the data, the bam files from the 3 auxin-treated replicates were merged into a single bam file using the *samtools* merge command. Likewise, the 2 control replicates were merged into a single bam file.

Processing of small RNA-seq reads

Quality of small RNAseq reads was assessed using *fastqc FastQX* v0.11.8. (https://www.bioinformatics.babraham.ac.uk/projects/fastqc). Read trimming and adapter clipping were performed using *cutadapt* tool (Marcel 2011). Trimmed reads ranging from 17-29 nt were used for further analysis. Reads were mapped to the *C. elegans* reference genome (WS279) using *bowtie v1.2.2* (Langmead *et al.* 2009). Mapped reads were input into the quantifier utility from *miRDeep2* pipeline (Friedländer *et al.* 2012) to generate expression of miRNAs (RPM).

Data Visualization

JBrowse 1.16.3 was utilized for data visualization (Buels *et al.* 2016). Visualized tracks included *C. elegans* reference genome annotation (WS279). Precursor miRNA data was extracted from miRbase v22 (Kozomara *et al.* 2018). To visualize the read coverage from auxin and control treatments, we displayed the merged files showing normalized read coverage. To correlate our pri-miRNA annotations with existing TSS data, we downloaded CAP-Seq reads from GEO (GSE40053) (Gu *et al* 2012),and mapped them to the reference genome. We visualized both GSM984429 (ya0217) and GSM984430 (avr0217) (Gu *et al*. 2012), representing CAPseq reads from young adults and average reads across development, respectively. Our *de novo* transcript assembly was also visualized in *Jbrowse*. To generate figures, the same tracks were loaded into IGV (Robinson *et al*. 2011), and sashimi plots were generated to visualize the pri-miRNA loci and surrounding genomic features.

Discussion

The study of miRNA function and regulation is critical to our understanding of gene regulatory networks and their relationship to normal development and manifestation of disease states or pathologies. To this end, research efforts have focused on elucidating miRNA biogenesis, regulation of miRNA activity, and their downstream effects. However, the transcriptional and post-transcriptional study of miRNAs has been significantly impeded by the lack of comprehensive annotations of primary miRNA transcripts, including miRNA loci structures. In addition, we (Chapter 2) and others (see Chapter 1 for review) have been interested in understanding the molecular crosstalk between two classes of post-transcriptional regulators: RNA binding proteins (RBPs) and miRNAs. RBPs have been shown to affect primary miRNA processing (see Chapter 1), however, probing effects of RBP activity on primary miRNA processing has remained challenging due to lack of pri-miRNA annotations.

To facilitate the study of miRNA gene transcription, regulation, and pri-miRNA processing, we generated an Auxin-inducible degron system to conditionally knockdown expression of DRSH-1. 4-hour exposure to Auxin (IAA) was sufficient to significantly knockdown DRSH-1 expression, leading to accumulation of primary miRNA transcripts which were then subjected to RNAseq analysis. Pri-miRNA transcripts were generated, leading to the annotation of 74 pri-miRNAs, 72 or which have not been previously described (Figure 3.1A, Supplemental Table 3.4). We characterized the 74 pri-miRNAs by genomic location, overlap with protein-coding and non-coding genomic elements (Figure 3.2C, Supplemental Table 3.3), and their relationship to previously identified transcription start sites (Gu *et al*. 2012; Figure 3.2E). For loci where predicted transcription start sites were annotated, we were gratified to see that many primary miRNA assemblies had corresponding TSS, supporting the accuracy of our annotations.

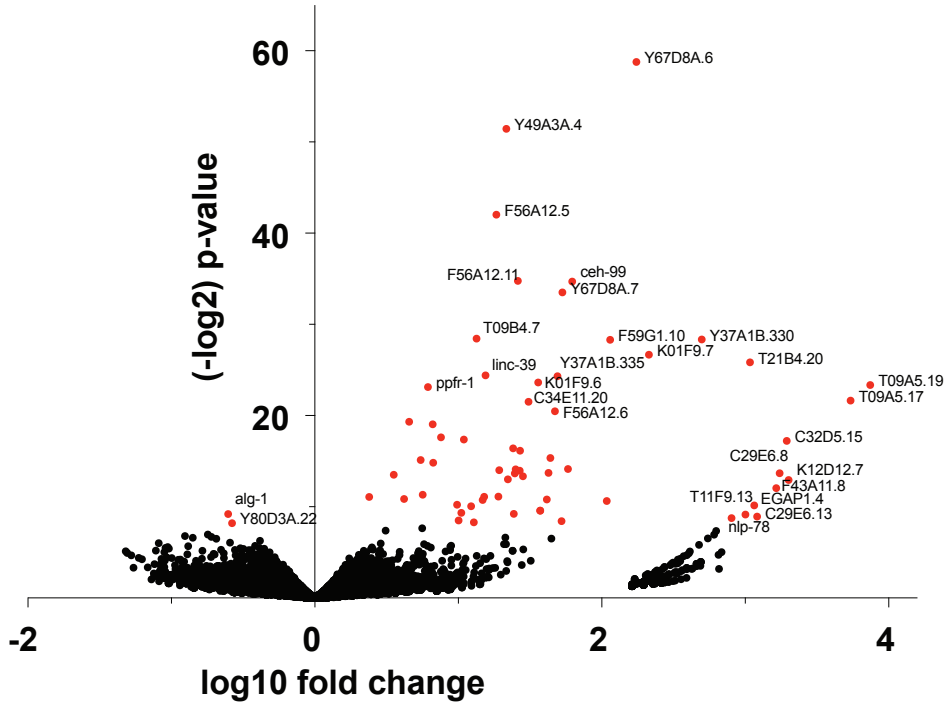
Complexity of miRNA gene loci

This work has highlighted the increasingly recognized complexity of miRNA loci. Consistent we previous observations, we found that majority of miRNA transcripts overlapped significantly with other genomic elements, including non-coding elements, and protein-coding genes, or both (Supplemental Table 3.3). Many of the *C. elegans* miRNA loci overlap with antisense non-coding RNAs (Figures 3.3 A,B,C, Figure 3.5 A,C and Supplemental Table 3.3). Antisense transcripts to miR398 genes have been recently shown to repress processing of the *miR398* pri-miRNA (Li *et al.* 2020). Others have reported RNAs antisense to miRNA transcripts (Faghihi and Wahlestedt 2009; Song *et al.* 2020), however, the biological significance of these, and their effect on miRNA biogenesis have not been fully explored. Future work should be focused on refining annotations of pri-miRNA transcripts and careful annotations of their cognate antisense RNAs, followed by functional studies into the roles antisense RNAs may play in miRNA loci regulation.

In addition to the potential impact of strandedness, we noted patterns of read accumulation that would seem to support the presence of multiple pri-miRNA isoforms for many miRNA loci (Supplemental Table 3.6). In support of this hypothesis, our *de novo* transcript assembly identified alternative isoforms for the majority of pri-miRNAs we annotated (Supplemental Table 3.6). These observations, in combination with the recent work done on regulatory roles of pri-*let-7* isoforms (Nelson & Ambros 2019), would suggest that miRNAs may be expressed from differentially produced pri-miRNA isoforms, as a layer of regulation of miRNA biogenesis.

Lastly, analysis of our *de novo* transcriptome assemblies allowed us to identify greater than 300 transcripts that who appear to be dependent on DRSH-1, presumably for their processing (Figure 3.5, Supplemental Table 3.5). Many of these transcripts are relatively short (between 200-300 nt long, Supplemental Figure 3.5) and *a priori* do not overlap with any other genomic elements. Although we cannot discount the fact that lower read coverage impacted detailed annotation of these novel loci, their identification expands our understanding of *C. elegans* ncRNA transcriptome.

Future efforts will be focused on validating the newly annotated pri-miRNA transcripts. Experimental validation (through, for example, RACE) will be crucial to confirm the presence of alternative pri-miRNA isoforms, as well the as the novel DRSH-dependent transcripts. As our study was performed utilizing somatically expressed TIR-1, DRSH-1 depletions were largely ineffective in the germline tissues, meaning that we failed to capture germline enriched pri-miRNAs. Future work will repeat the conditional knockdown using both somatic and germline expressed TIR-1 (under a truly ubiquitous promoter), allowing for a more comprehensive analysis of primary miRNA transcripts, regardless of tissue. Overall, this work will facilitate future studies into regulation of miRNA transcription and processing, ultimately furthering our understanding of miRNA regulation in development and disease.



Supplemental Figure 3.1. Genome-wide differential expression analysis of mRNAs after depletion of DRSH-1 by auxin treatment.

X-axis shows fold change (auxin/control). Y-axis shows the -log2 of p-values from for each gene. Significant p-values (< 0.05) are highlighted in red.

Supplemental Table 3.1. Coding gene features that interact with de novo pri-miRNA transcripts.

Overlaping feature	Number of pri-miRNAs
Exon only	1
Intron only	8
Intron + exon	13
Exon+ UTR	7
Intron + exon + UTR	5

Supplemental Table 3.2. Total number of non-coding genomic elements across 74 pri-miRNA annotations.

Overlapping non-coding feature	Total number
ncRNA	40
lnc-RNA	5
piRNA	4
snoRNA	4
tRNA	1
snRNA	1

Supplemental Table 3.3. Overlapping features of *de novo* pri-miRNA transcripts.

	miR	Overlapping ncRNA	Overlapping coding	TSS
Intronic or overlaps with protein coding gene	cel-mir-244	T04D1.7 lincRNA (linc-39)	T04D1.2 exon + UTR (+strand)	corresponding TSS
	cel-mir-79	C12C8.6 ncRNA (- strand)	lin-41 exon + UTR	corresponding TSS
	cel-mir-2	F16A11.5 ncRNA (+ strand), F16A11.11a lncRNA (+ strand), F16A11.11b lncRNA (+ strand), F16A11.12 snoRNA (- strand), M04C9.10 ncRNA (- strand)	ppfr-1 intronic only	corresponding TSS
	cel-mir-71	F16A11.6 ncRNA (+ strand), F16A11.8 ncRNA (+ strand), F16A11.13 ncRNA (+ strand), F16A11.10 lncRNA (+ strand), M04C9.10 ncRNA (- strand)	ppfr-1 intron + exons	corresponding TSS
	cel-lin-4	F59G1.10 ncRNA (- strand), F59G1.12 ncRNA (- strand)	F59G1.4 intron + exon (+ strand)	corresponding TSS
	cel-mir-1822	n/a	ZK84.2 intron + exon + UTR (+ strand)	n/a
	cel-mir-57	T09A5.19 ncRNA (+ strand), T09A5.17 ncRNA (- strand)	acr-7 exon + UTR (- strand)	corresponding TSS
	cel-mir-252	W02B12.17 ncRNA (+ strand)	W02B12.13 intron + exon (+ strand)	corresponding TSS
	cel-mir-355	n/a	clec-144 exon + intron (- strand)	corresponding TSS
	cel-mir-229 cel-mir-64 cel-mir-65 cel-mir-66	n/a	gcn-1 intron + exon (+ strand)	corresponding TSS
	cel-mir-67	EGAP1.4 ncRNA (+ strand)	zmp-1 intron + exon (- strand)	corresponding TSS
	cel-mir-90	K01F9.7 ncRNA (+ strand)	K01F9.2 intron + exon	corresponding TSS
	cel-mir-86 cel-mir-8211	n/a	Y56A3A.7 intron + exon + UTR (- strand)	corresponding TSS
	cel-mir-58a	Y67D8A.6 ncRNA (+ strand), Y67D8A.7 ncRNA (+ strand)	Y67D8A.2 intron only	corresponding TSS
	cel-mir-243	n/a	R08C7.12 intron only	corresponding TSS
	cel-mir-124	C29E6.13 ncRNA (+ strand), C29E6.8 ncRNA (- strand), C29E6.11 ncRNA (+ strand)	trpa-1 intron + exons	corresponding TSS
	cel-mir-87	F10C2.11 ncRNA (+ strand)	kup-1 intron + exon (- strand)	corresponding TSS
	cel-mir-81	n/a	T07D1.2 intron only (- strand)	n/a
	cel-mir-82	n/a	T07D1.2 intron only (- strand)	corresponding TSS
	cel-mir-230	n/a	F13D11.3 intron + exon + UTR (- strand)	corresponding TSS
	cel-mir-360	Y23B4A.5 ncRNA (+ strand)	capa-1 intron + exon + UTR (+ strand)	n/a
	cel-mir-240 cel-mir-786 cel-mir-8203	n/a	pcyt-2.2 UTR + exon (+ strand)	corresponding TSS
	cel-mir-54	F09A5.11 ncRNA (- strand)	F09A5.3 intron only (+ strand)	corresponding TSS
	cel-mir-47	n/a	meg-2 UTR + exon (- strand)	corresponding TSS
	cel-let-7	C05G5.8 ncRNA (+ strand)	C05G5.7 exon only (- strand),	corresponding TSS
	cel-mir-84	n/a	nhx-1 intron + exon + UTR (+ strand)	corresponding TSS
	cel-mir-2217	n/a	C16H3.3 intron only (- strand)	n/a
	cel-mir-357	n/a	symk-1 intron only (+ strand)	corresponding TSS
Does not overlap with protein coding gene	cel-mir-235	T09B4.19 ncRNA (+strand), T09B4.t2 tRNA (- strand)	n/a	corresponding TSS
	cel-mir-1	T09B4.14 ncRNA (+ strand), T09B4.17 ncRNA (+ strand)	n/a	corresponding TSS
	cel-mir-245	n/a	n/a	corresponding TSS
	cel-mir-1817	n/a	n/a	n/a
	cel-mir-819	n/a	n/a	n/a
	cel-mir-60	n/a	n/a	corresponding TSS
	cel-mir-8188	n/a	n/a	n/a
	cel-mir-45	ZK930.13 ncRNA (+ strand)	n/a	corresponding TSS
	cel-mir-77	T21B4.20 snoRNA (+ strand)	n/a	corresponding TSS
	cel-mir-2217	n/a	n/a	n/a
	cel-mir-231	R13A5.17 (linc-57) (-strand)	n/a	corresponding TSS
	cel-mir-80	K01F9.6 ncRNA (+ strand)	n/a	corresponding TSS
	cel-mir-238	n/a	n/a	corresponding TSS
	cel-mir-46	ZK525.4 ncRNA (- strand), ZK525.6 ncRNA (+ strand)	n/a	corresponding TSS
	cel-mir-8189	n/a	n/a	corresponding TSS
	cel-mir-228	T12E12.44 ncRNA (+ strand), T12E12.32 piRNA (- strand), T12E12.25 piRNA (- strand), T12E12.37 piRNA (- strand), T12E12.35 piRNA (- strand)	n/a	corresponding TSS
	cel-mir-1820	n/a	n/a	corresponding TSS
	cel-mir-83	C06A6.9 ncRNA (+ strand)	n/a	corresponding TSS
	cel-mir-246	n/a	n/a	corresponding TSS
	cel-mir-59	n/a	n/a	corresponding TSS
	cel-mir-52	Y37A1B.330 ncRNA (+ strand), Y37A1B.335 ncRNA (- strand)	n/a	corresponding TSS
	cel-mir-2217	n/a	n/a	n/a
	cel-mir-2217	n/a	n/a	n/a
	cel-mir-255	F08F3.13 ncRNA (- strand)	n/a	corresponding TSS
	cel-lsy-6	n/a	n/a	n/a
	cel-mir-48 cel-mir-241	F56A12.5 snRNA (- strand), F56A12.6 snoRNA (+ strand), F56A12.8 ncRNA (+ strand)	n/a	corresponding TSS + independent mir-48 TSS
	cel-mir-73	T24D8.11 ncRNA (- strand), T24D8.14 ncRNA (+ strand), T24D8.16 ncRNA (- strand)	n/a	corresponding TSS
	cel-mir-75	n/a	n/a	corresponding TSS
	cel-mir-264	n/a	n/a	n/a
	cel-mir-34	n/a	n/a	n/a
	cel-mir-784	n/a	n/a	n/a
	cel-mir-788	n/a	n/a	corresponding TSS
	cel-mir-49	n/a	n/a	n/a
	cel-mir-251	n/a	n/a	n/a
	cel-mir-239b	n/a	n/a	corresponding TSS
	cel-mir-239a	C34E11.20 snoRNA (+ strand)	n/a	corresponding TSS
	cel-mir-785	n/a	n/a	corresponding TSS
	cel-mir-43 cel-mir-44	ZK930.12 ncRNA (- strand)	n/a	corresponding TSS

Supplemental Table 3.4. Characteristics of *de novo* pri-miRNA transcripts.

		***de novo* transcript coordinates**		**Transcript Characteristics**	
	miR			Transcript length	Orientation
Intronic or overlaps with protein coding gene	cel-mir-244	4682313	4684646	2333	(-) strand
	cel-mir-79	9332799	9337426	4627	(+) strand
	cel-mir-2	9371954	9375961	4007	(-) strand
	cel-mir-71	9377844	9382965	5121	(-) strand
	cel-lin-4	5900636	5902493	1857	(+) strand
	cel-mir-1822	6015445	6015926	481	(-) strand
	cel-mir-57	7849618	7852365	2747	(-) strand
	cel-mir-252	11446594	11446968	374	(-) strand
	cel-mir-355	11832506	11833829	1323	(-) strand
	cel-mir-229	2172242	2174967	2725	(+) strand
	cel-mir-64				
	cel-mir-65				
	cel-mir-66				
	cel-mir-67	5928887	5931534	2647	(-) strand
	cel-mir-90	8871165	8874234	3069	(-) strand
	cel-mir-86	11933990	11937098	3108	(-) strand
	cel-mir-8211				
	cel-mir-58a	3232622	3234774	2152	(+) strand
	cel-mir-243	4450719	4451352	633	(+) strand
	cel-mir-124	11871340	11873780	2440	(+) strand
	cel-mir-87	12037052	12039017	1965	(-) strand
	cel-mir-81	2430607	2431221	614	(+) strand
	cel-mir-82	2435223	2435456	233	(-) strand
	cel-mir-230	5803780	5805027	1247	(+) strand
	cel-mir-360	5919100	5919835	735	(+) strand
	cel-mir-240	7882579	7884442	1863	(+) strand
	cel-mir-786				(+) strand
	cel-mir-8203				(-) strand
	cel-mir-54	13144929	13145245	316	(-) strand
	cel-mir-47	13920780	13923096	2316	(+) strand
	cel-let-7	14743589	14745311	1722	(-) strand
	cel-mir-84	16020487	16023981	3494	(-) strand
	cel-mir-2217	17605528	17609347	3819	(+) strand
	cel-mir-357	8580353	8580828	475	(-) strand
Does not overlap with protein coding gene	cel-mir-235	6161953	6163534	1581	(-) strand
	cel-mir-1	6172523	6173053	530	(-) strand
	cel-mir-245	7887336	7887486	150	(+) strand
	cel-mir-1817	10868839	10869625	786	(-) strand
	cel-mir-819	2136652	2136923	271	(+) strand
	cel-mir-60	6327740	6328831	1091	(-) strand
	cel-mir-8188	11840104	11840767	663	(+) strand
	cel-mir-45	11880601	11881150	549	(-) strand
	cel-mir-77	12519089	12519906	817	(+) strand
	cel-mir-2217	13618472	13622328	3856	(-) strand
	cel-mir-231	7543475	7546436	2961	(-) strand
	cel-mir-80	8863917	8865652	1735	(-) strand
	cel-mir-238	8867025	8868452	1427	(-) strand
	cel-mir-46	13660218	13662466	2248	(+) strand
	cel-mir-8189	2078879	2079130	251	(+) strand
	cel-mir-228	5561902	5563327	1425	(+) strand
	cel-mir-1820	7767231	7767576	345	(+) strand
	cel-mir-83	7841425	7842227	802	(+) strand
	cel-mir-246	10940099	10940280	181	(+) strand
	cel-mir-59	11308616	11309326	710	(-) strand
	cel-mir-52	14033873	14034553	680	(+) strand
	cel-mir-2217	14451423	14455223	3800	(-) strand
	cel-mir-2217	15798890	15799167	277	(-) strand
	cel-mir-255	5440502	5440968	466	(+) strand
	cel-lsy-6	10647182	10647527	345	(+) strand
	cel-mir-48	14361129	14367340	6211	(-) strand
	cel-mir-241				
	cel-mir-73	2368707	2369014	307	(+) strand
	cel-mir-75	2372260	2373222	962	(+) strand
	cel-mir-264	2818993	2819581	588	(+) strand
	cel-mir-34	2969467	2969993	526	(-) strand
	cel-mir-784	8074873	8075040	167	(-) strand
	cel-mir-788	8484604	8485409	805	(-) strand
	cel-mir-49	9989152	9989436	284	(+) strand
	cel-mir-251	11004702	11004984	282	(+) strand
	cel-mir-239b	11790629	11791445	816	(-) strand
	cel-mir-239a	11792126	11793527	1401	(+) strand
	cel-mir-785	3004449	3005027	578	(+) strand
	cel-mir-43	11889865	11890374	509	(+) strand
	cel-mir-44				(+) strand

Supplemental Table 3.5. Novel DRSH-1 dependent transcripts

Transcript Location	Length	TPM (avg-3 reps)	FPKM (avg-3 reps)
transcript::IV:7839079-7839428(-)	349	360.49	1052.10
transcript::II:2459502-2460172(+)	670	331.18	947.82
transcript::I:6337095-6337300(+)	205	304.63	890.51
transcript::X:480377-480598(+)	221	254.32	742.73
transcript::X:3469140-3469403(+)	263	233.07	673.41
transcript::X:10410437-10410790(-)	353	206.31	611.92
transcript::II:7273082-7273391(+)	309	201.33	592.47
transcript::IV:8200343-8200544(-)	201	198.13	575.38
transcript::IV:11027204-11027436(-)	232	195.05	559.97
transcript::II:2460351-2460690(+)	339	185.92	540.80
transcript::II:9901973-9902191(+)	218	173.65	504.35
transcript::II:5345856-5346067(+)	211	160.28	476.84
transcript::II:12692838-12693043(+)	205	159.35	464.23
transcript::I:69794-70137(+)	343	158.82	460.41
transcript::II:306407-306627(+)	220	148.48	432.58
transcript::IV:10783149-10783391(-)	242	146.89	441.48
transcript::X:7681668-7681919(+)	251	142.58	408.07
transcript::III:9277793-9278045(-)	252	142.02	406.45
transcript::X:8494117-8494391(+)	274	139.03	399.37
transcript::X:11490664-11490870(+)	206	138.98	397.77
transcript::I:7321800-7322025(-)	225	134.93	391.84
transcript::X:2710106-2710319(-)	213	133.29	387.14
transcript::V:10003239-10003584(-)	345	132.14	391.60
transcript::V:13531059-13531277(-)	218	131.33	375.87
transcript::I:6172002-6172312(+)	310	130.32	375.58
transcript::II:7032859-7033178(-)	319	130.26	382.06
transcript::X:2968970-2969195(-)	225	127.25	364.18
transcript::X:1405114-1405324(-)	210	126.96	381.56
transcript::V:4210992-4211193(+)	201	126.91	371.55
transcript::I:529213-529650(-)	437	126.52	370.87
transcript::I:3977779-3977991(-)	212	125.76	377.96
transcript::I:10516476-10516698(+)	222	125.29	367.94
transcript::II:10445605-10445891(-)	286	125.13	358.13
transcript::V:9504620-9504829(+)	209	124.81	361.05
transcript::IV:2065642-2065945(+)	303	123.82	358.58
transcript::I:7888922-7889213(+)	291	122.98	351.98
transcript::IV:10611426-10611663(-)	237	119.79	347.94
transcript::V:3984067-3984370(-)	303	117.32	339.75
transcript::V:13301201-13301508(-)	307	116.57	333.63
transcript::V:14983974-14984380(-)	406	116.55	338.51
transcript::II:9776003-9776392(-)	389	116.33	333.96
transcript::X:13250043-13250325(+)	282	115.84	337.47
transcript::X:9754198-9754467(-)	269	115.01	330.64
transcript::I:11591114-11591378(-)	264	115.00	333.96
transcript::X:9261174-9261438(-)	264	115.00	333.96
transcript::II:2872830-2873063(+)	233	114.42	343.90
transcript::X:6581573-6581989(-)	416	114.32	329.12
transcript::X:4801486-4801737(-)	251	114.07	326.46
transcript::V:5114163-5114435(-)	272	113.74	326.99
transcript::X:13996021-13996241(-)	220	113.32	335.93
transcript::III:7375112-7375380(-)	268	113.28	328.97
transcript::V:1006137-1006348(-)	211	112.70	324.44
transcript::X:13420981-13421223(+)	242	112.55	334.32
transcript::X:8207525-8207740(-)	215	110.60	318.40
transcript::X:13004884-13005094(-)	210	110.49	322.28
transcript::III:7463285-7463603(+)	318	109.97	321.28
transcript::IV:3898180-3898440(-)	260	109.20	317.16
transcript::X:7286655-7286852(+)	197	109.00	311.96
transcript::IV:13005702-13005986(+)	284	108.93	313.18
transcript::III:4822737-4822956(-)	219	108.58	312.59
transcript::II:14467361-14467629(-)	268	106.83	305.75
transcript::X:9257522-9257724(+)	202	106.30	304.23
transcript::X:6011708-6011932(+)	224	106.16	305.61
transcript::V:1074802-1075141(+)	339	105.57	302.14
transcript::V:13397720-13397949(-)	229	103.84	298.94
transcript::X:6121548-6121802(+)	254	102.70	297.08
transcript::IV:8544437-8544709(-)	272	102.26	300.31
transcript::V:8155610-8155821(-)	211	101.77	291.26
transcript::X:8484603-8484831(-)	228	101.76	296.84
transcript::IV:2214514-2214771(-)	257	101.50	293.61
transcript::V:1851515-1851817(+)	302	100.53	291.94
transcript::II:9452334-9452549(-)	215	99.87	285.84
transcript::III:13671922-13672137(-)	215	99.87	285.84
transcript::X:15233445-15233662(+)	217	98.95	283.20
transcript::V:16823058-16823276(-)	218	98.50	281.90
transcript::V:9274091-9274357(+)	266	98.06	283.68
transcript::V:11341095-11341405(+)	310	97.94	284.40
transcript::III:7817284-7817505(-)	221	97.16	278.08
transcript::IV:14265755-14265976(-)	221	97.16	278.08
transcript::V:11691858-11692080(-)	222	96.73	276.83
transcript::III:9668728-9668938(-)	210	96.57	279.54
transcript::V:19038504-19038825(-)	321	96.38	277.08
transcript::III:1786789-1787049(-)	260	95.89	284.25
transcript::IV:13126801-13127044(-)	243	95.48	278.52
transcript::V:12667587-12667814(+)	227	94.59	270.73
transcript::V:4895364-4895669(-)	305	93.87	268.66

Transcript Location	Length	TPM (avg-3 reps)	FPKM (avg-3 reps)
transcript::III:4451526-4451751(+)	225	71.31	209.75
transcript::IV:9367376-9367826(-)	450	71.31	209.75
transcript::I:5699713-5699914(+)	201	71.22	203.83
transcript::IV:6493376-6493577(-)	201	71.22	203.83
transcript::X:3256449-3256650(-)	201	71.22	203.83
transcript::IV:10374467-10374734(+)	267	70.89	205.89
transcript::II:3077461-3077688(+)	227	70.68	207.91
transcript::II:7564770-7564973(-)	203	70.52	201.82
transcript::IV:3195279-3195482(+)	203	70.52	201.82
transcript::IV:6493165-6493368(+)	203	70.52	201.82
transcript::V:1627226-1627454(-)	228	70.37	206.99
transcript::V:6864873-6865077(-)	204	70.17	200.83
transcript::X:6609723-6609927(+)	204	70.17	200.83
transcript::X:690907-691112(-)	205	69.83	199.85
transcript::IV:9717350-9717589(+)	239	69.55	200.72
transcript::I:9081601-9081807(+)	206	69.49	198.88
transcript::V:9175556-9175762(-)	206	69.49	198.88
transcript::X:16962735-16962991(+)	256	69.43	208.67
transcript::X:9320213-9320469(-)	256	69.43	208.67
transcript::V:8599835-8600067(-)	232	69.16	203.42
transcript::IV:11382292-11382499(+)	207	69.16	197.92
transcript::I:7327662-7327871(+)	209	68.49	196.03
transcript::V:7009392-7009601(+)	209	68.49	196.03
transcript::V:15347500-15347768(+)	268	68.47	202.22
transcript::X:16226100-16226335(+)	235	68.27	200.83
transcript::II:6311394-6311604(+)	210	68.17	195.10
transcript::IV:11733363-11733573(+)	210	68.17	195.10
transcript::V:17749936-17750146(-)	210	68.17	195.10
transcript::II:4659244-4659593(-)	349	68.13	196.15
transcript::X:16131986-16132222(+)	236	67.99	199.98
transcript::III:5823228-5823439(-)	211	67.85	194.17
transcript::V:635333-635578(-)	245	67.84	195.80
transcript::V:14940715-14940994(-)	279	67.84	197.04
transcript::III:5368252-5368532(-)	280	67.60	196.34
transcript::II:4935185-4935397(-)	212	67.53	193.26
transcript::IV:12832292-12832564(+)	272	67.47	199.25
transcript::V:18452578-18452859(-)	281	67.36	195.64
transcript::II:9038406-9038653(+)	247	67.29	194.22
transcript::III:1601422-1601635(+)	213	67.21	192.35
transcript::V:17233840-17234105(-)	265	67.07	201.58
transcript::IV:16182313-16182561(-)	248	67.02	193.44
transcript::II:8087516-8087897(+)	381	66.95	196.01
transcript::II:8193815-8194029(-)	214	66.89	191.45
transcript::X:11027793-11028007(-)	214	66.89	191.45
transcript::I:8392802-8393051(-)	249	66.75	192.66
transcript::V:553568-553843(-)	275	66.73	197.08
transcript::III:9272722-9272964(-)	242	66.30	195.02
transcript::X:6794952-6795229(-)	277	66.25	195.65
transcript::I:3846144-3846361(+)	217	65.97	188.80
transcript::X:10435158-10435375(-)	217	65.97	188.80
transcript::X:13929723-13930049(+)	326	65.87	188.51
transcript::I:9805203-9805530(-)	327	65.67	187.94
transcript::III:13670934-13671918(-)	984	65.47	187.36
transcript::V:11024720-11024939(-)	219	65.37	187.08
transcript::X:3295846-3296119(+)	273	65.11	195.67
transcript::X:10762244-10762464(+)	220	65.07	186.23
transcript::X:6262643-6262863(+)	220	65.07	186.23
transcript::II:7420642-7420889(-)	247	64.96	191.07
transcript::X:8887501-8887722(-)	221	64.78	185.38
transcript::II:11838456-11838678(+)	222	64.48	184.55
transcript::IV:1646351-1646573(+)	222	64.48	184.55
transcript::X:2089370-2089592(-)	222	64.48	184.55
transcript::I:10012514-10012737(+)	223	64.19	183.72
transcript::III:6279222-6279445(+)	223	64.19	183.72
transcript::I:12812010-12812270(-)	260	63.93	184.51
transcript::V:14153620-14153871(-)	251	63.92	188.03
transcript::IV:601914-602251(-)	337	63.72	182.36
transcript::I:5713843-5714104(+)	261	63.68	183.80
transcript::X:7916370-7916631(+)	261	63.68	183.80
transcript::IV:17424993-17425218(-)	225	63.62	182.09
transcript::V:2828078-2828303(-)	225	63.62	182.09
transcript::V:597179-597404(+)	225	63.62	182.09
transcript::X:14786602-14786882(+)	280	63.48	190.78
transcript::III:1997559-1997785(+)	226	63.34	181.28
transcript::IV:9929167-9929620(-)	453	63.20	180.88
transcript::X:9956150-9956377(+)	227	63.06	180.49
transcript::IV:11812153-11812381(-)	228	62.79	179.69
transcript::I:787514-787743(+)	229	62.51	178.91
transcript::X:16977142-16977372(-)	230	62.24	178.13
transcript::IV:9287738-9288033(+)	295	62.21	183.72
transcript::IV:7300720-7300951(+)	231	61.97	177.36
transcript::IV:7452302-7452534(+)	232	61.70	176.60
transcript::V:5041246-5041478(-)	232	61.70	176.60
transcript::I:419623-419893(+)	270	61.56	177.67
transcript::I:5543497-5543730(+)	233	61.44	175.84
transcript::IV:6522568-6522802(-)	234	61.18	175.09

transcript::V:2267417-2267646(-)	229	93.77	268.36		transcript::X:325092-325392(-)	300	61.17	180.65
transcript::X:8484142-8484372(-)	230	93.36	267.20		transcript::II:8739977-8740249(-)	272	61.11	176.37
transcript::I:7471015-7471247(-)	232	92.56	264.89		transcript::III:5501429-5501781(-)	352	61.00	174.59
transcript::IV:13980356-13980607(+)	251	92.44	269.64		transcript::II:11785660-11785952(-)	292	60.87	182.94
transcript::I:8588257-8588510(+)	253	91.71	267.51		transcript::V:9077861-9078125(-)	264	60.77	178.77
transcript::V:17899512-17899719(+)	207	91.44	265.57		transcript::IV:2943710-2944064(-)	354	60.66	173.60
transcript::X:2153246-2153454(-)	208	91.00	264.30		transcript::X:8743132-8743370(+)	238	60.15	172.14
transcript::V:5852153-5852547(-)	394	90.83	259.96		transcript::II:10896274-10896513(-)	239	59.90	171.42
transcript::V:976387-976703(-)	316	90.60	259.31		transcript::IV:11389694-11389933(+)	239	59.90	171.42
transcript::IV:5110750-5110989(-)	239	89.85	257.13		transcript::IV:2783660-2783899(+)	239	59.90	171.42
transcript::V:5571409-5571622(+)	213	88.86	258.09		transcript::X:4577111-4577350(+)	239	59.90	171.42
transcript::II:4570041-4570337(-)	296	88.13	254.93		transcript::X:5260322-5260600(+)	278	59.79	172.56
transcript::II:4814812-4815191(-)	379	87.71	253.15		transcript::V:12466238-12466637(-)	399	59.60	171.57
transcript::V:4110055-4110272(-)	217	87.22	253.33		transcript::IV:17452165-17452407(-)	242	59.15	169.30
transcript::I:8987595-8987813(+)	218	86.82	252.17		transcript::III:12934210-12934482(+)	272	58.99	173.51
transcript::II:5957296-5957597(-)	301	86.66	250.69		transcript::III:4482179-4482423(+)	244	58.67	167.91
transcript::III:9010321-9010569(+)	248	86.58	247.80		transcript::IV:6653222-6653467(-)	245	58.43	167.22
transcript::X:5509181-5509641(-)	460	86.57	251.42		transcript::IV:7265713-7265959(-)	246	58.19	166.55
transcript::III:10653359-10653578(+)	219	86.43	251.02		transcript::I:6198348-6198595(-)	247	57.96	165.87
transcript::IV:7718171-7718390(-)	219	86.43	251.02		transcript::II:12855804-12856051(+)	247	57.96	165.87
transcript::X:3324732-3324938(-)	206	86.28	259.31		transcript::IV:12763513-12763760(+)	247	57.96	165.87
transcript::V:6611425-6611639(-)	214	85.75	253.25		transcript::X:3137423-3137670(-)	247	57.96	165.87
transcript::X:14601670-14601864(-)	194	85.68	247.28		transcript::X:1394267-1394515(-)	248	57.72	165.20
transcript::V:5678154-5678472(-)	318	85.65	254.42		transcript::X:13741245-13741494(-)	249	57.49	164.54
transcript::X:12671907-12672128(-)	221	85.64	248.75		transcript::III:4453513-4453958(+)	445	57.32	167.82
transcript::V:10212258-10212483(+)	225	84.12	244.33		transcript::IV:7858856-7859107(+)	251	57.03	163.23
transcript::III:4393148-4393361(-)	213	83.45	250.79		transcript::V:6725971-6726222(+)	251	57.03	163.23
transcript::II:4772160-4772387(+)	227	83.38	242.18		transcript::III:9281270-9281523(-)	253	56.58	161.94
transcript::V:14949526-14949753(+)	227	83.38	242.18		transcript::V:19083373-19083626(-)	253	56.58	161.94
transcript::IV:11388851-11389109(+)	258	83.23	238.20		transcript::X:9937449-9937702(+)	253	56.58	161.94
transcript::II:4508704-4509050(+)	346	82.75	236.82		transcript::III:1427695-1427949(-)	254	56.36	161.30
transcript::X:3186220-3186393(+)	173	82.75	236.82		transcript::IV:4306969-4307256(-)	287	55.90	164.44
transcript::X:4721905-4722106(+)	201	82.69	238.67		transcript::IV:5581148-5581405(+)	257	55.70	159.42
transcript::III:4547872-4548102(+)	230	82.29	239.02		transcript::V:12555323-12555612(+)	289	55.52	163.30
transcript::X:2589759-2590302(-)	543	82.27	237.82		transcript::X:12591658-12591917(-)	259	55.27	158.19
transcript::V:14659423-14659706(+)	283	81.99	239.15		transcript::II:6379634-6379894(-)	260	55.06	157.58
transcript::I:13292199-13292461(+)	262	81.96	234.56		transcript::III:3686231-3686533(-)	302	55.04	158.85
transcript::X:16839135-16839397(+)	262	81.96	234.56		transcript::IV:3865261-3865522(+)	261	54.85	156.97
transcript::X:15154735-15154953(+)	218	81.53	245.04		transcript::I:6400786-6401079(-)	293	54.76	161.07
transcript::V:17574834-17575098(+)	264	81.34	232.79		transcript::X:15223542-15223804(-)	262	54.64	156.37
transcript::V:8287539-8287804(-)	265	81.03	231.91		transcript::X:685971-686322(+)	351	53.92	156.62
transcript::III:7048662-7048928(-)	266	80.73	231.03		transcript::X:8828631-8828929(-)	298	53.84	158.37
transcript::X:17115256-17115455(-)	199	80.63	237.16		transcript::II:7110499-7110766(-)	267	53.62	153.45
transcript::III:9050900-9051128(+)	228	80.49	237.70		transcript::I:5396862-5397130(+)	268	53.42	152.87
transcript::III:10201378-10201585(+)	207	80.30	231.75		transcript::V:1071948-1072216(+)	268	53.42	152.87
transcript::X:7544329-7544536(+)	207	80.30	231.75		transcript::V:4721671-4721939(+)	268	53.42	152.87
transcript::IV:12841884-12842152(+)	268	80.12	229.31		transcript::X:7186428-7186696(-)	268	53.42	152.87
transcript::IV:11166912-11167120(+)	208	79.91	230.63		transcript::IV:16036319-16036588(+)	269	53.22	152.31
transcript::V:3984889-3985158(+)	269	79.83	228.46		transcript::V:5404407-5404676(+)	269	53.22	152.31
transcript::II:5605996-5606266(-)	270	79.53	227.61		transcript::X:13172353-13172622(-)	269	53.22	152.31
transcript::X:2905213-2905506(+)	293	79.19	230.99		transcript::X:5098884-5099153(-)	269	53.22	152.31
transcript::II:13770341-13770551(-)	210	79.15	228.44		transcript::II:121312-121626(-)	314	52.93	152.78
transcript::III:2068983-2069193(+)	210	79.15	228.44		transcript::I:2579990-2580294(+)	304	52.78	155.24
transcript::V:8258484-8258709(+)	225	78.99	237.42		transcript::II:8586128-8586403(+)	275	52.06	148.98
transcript::I:7109635-7109907(+)	272	78.94	225.94		transcript::IV:8851369-8851645(+)	276	51.87	148.44
transcript::II:3528974-3529246(-)	272	78.94	225.94		transcript::IV:8944656-8944932(-)	276	51.87	148.44
transcript::I:3648345-3648730(-)	385	78.86	229.00		transcript::I:10601-10878(-)	277	51.68	147.91
transcript::V:6748762-6748997(+)	235	78.09	230.62		transcript::IV:3474777-3475054(-)	277	51.68	147.91
transcript::X:6120731-6120974(-)	243	77.89	226.23		transcript::IV:8086823-8087102(-)	279	51.31	146.85
transcript::V:6750791-6751068(+)	277	77.52	221.86		transcript::I:8011204-8011520(+)	316	50.77	149.35
transcript::X:5781302-5781510(-)	208	77.14	226.90		transcript::II:1277191-1277473(-)	282	50.76	145.28
transcript::I:10074545-10074783(+)	238	77.10	227.72		transcript::IV:14469557-14469840(-)	283	50.58	144.77
transcript::X:2157959-2158238(+)	279	76.96	220.27		transcript::V:14970871-14971162(-)	291	49.19	140.79
transcript::IV:11048882-11049113(-)	231	76.94	231.25		transcript::V:12548099-12548437(-)	338	49.18	141.93
transcript::X:3298391-3298671(-)	280	76.69	219.48		transcript::X:16619763-16620200(-)	437	49.14	140.63
transcript::II:7317071-7317318(+)	247	76.63	222.57		transcript::II:10567393-10567700(-)	307	46.63	133.45
transcript::V:5693861-5694142(+)	281	76.42	218.70		transcript::X:11132225-11132537(-)	312	45.88	131.31
transcript::I:14212759-14212977(+)	218	76.24	220.05		transcript::V:11541834-11542149(+)	315	45.45	130.06
transcript::III:6198749-6198967(-)	218	76.24	220.05		transcript::V:4722239-4722560(-)	321	44.60	127.63
transcript::IV:8253998-8254209(+)	211	76.04	223.67		transcript::V:9191032-9191353(-)	321	44.60	127.63
transcript::IV:7318363-7318598(+)	235	75.63	227.31		transcript::I:7375345-7375676(+)	331	43.25	123.78
transcript::II:7772058-7772373(-)	315	75.49	217.32		transcript::III:10020633-10020972(+)	339	42.23	120.86
transcript::II:6523156-6523369(+)	213	75.33	221.57		transcript::X:11409806-11410146(-)	340	42.10	120.50
transcript::IV:3826613-3826849(-)	236	75.31	226.35		transcript::X:7087422-7088044(-)	622	41.01	120.07
transcript::I:5713284-5713536(+)	252	75.11	218.15		transcript::II:2622034-2622394(-)	360	39.76	113.81
transcript::IV:5563610-5563897(+)	287	74.82	214.13		transcript::V:5253150-5254132(-)	982	38.55	111.96
transcript::I:4535237-4535453(+)	216	74.28	218.49		transcript::V:4212984-4213233(+)	249	38.01	110.39
transcript::I:5389671-5389864(-)	193	74.17	212.28		transcript::IV:5491549-5491788(+)	239	37.18	111.75
transcript::II:3010031-3010375(-)	344	74.15	217.10		transcript::V:10563711-10563970(+)	259	36.54	106.13
transcript::II:8649293-8649534(-)	241	73.75	221.65		transcript::IV:10030038-10030299(+)	261	36.26	105.31
transcript::X:5983135-5983428(-)	293	73.29	209.75		transcript::II:1162880-1163088(+)	208	34.41	98.49
transcript::I:3070334-3070585(+)	251	73.11	215.92		transcript::X:11643018-11643243(+)	225	31.81	91.04
transcript::IV:7881707-7882004(-)	297	72.30	206.92		transcript::V:12822938-12823396(-)	458	31.26	89.45
transcript::X:15167167-15167520(-)	353	72.26	211.56		transcript::III:13174527-13174757(-)	230	31.12	89.07
transcript::III:401910-402164(-)	254	72.25	213.37		transcript::X:6671964-6672267(-)	303	14.46	41.39
transcript::V:20198962-20199209(+)	247	71.96	216.27		transcript::III:2444253-2444499(+)	246	14.33	41.62
transcript::V:9064146-9064370(-)	224	71.63	210.69					

Supplemental Table 3.6. Alternative *de novo* transcripts for pri-miRNA loci.

	miR	Annotated de novo transcript coordinates		Alternative *de novo* transcripts
Intronic or overlaps with protein coding gene	cel-mir-244	4682313	4684646	4681619..4684646; 4683044..4684646; 4682632..4684646
	cel-mir-79	9332799	9337426	9332799..9337168; 9332799..9336126
	cel-mir-2	9371954	9375961	9372750..9375961
	cel-mir-71	9377844	9382965	9379916..9382965
	cel-lin-4	5900636	5902493	5900636..5904048
	cel-mir-1822	6015445	6015926	n/a
	cel-mir-57	7849618	7852365	7849428..7852365
	cel-mir-252	11446594	11446968	11446710..11446968
	cel-mir-355	11832506	11833829	11832318..11833829; 11833531..11833829
	cel-mir-229 cel-mir-64 cel-mir-65 cel-mir-66	2172242	2174967	2172242..2173209
	cel-mir-67	5928887	5931534	5929037..5931534
	cel-mir-90	8871165	8874234	8873913..8874234
	cel-mir-86 cel-mir-8211	11933990	11937098	n/a
	cel-mir-58a	3232622	3234774	3232622..3233375
	cel-mir-243	4450719	4451352	n/a
	cel-mir-124	11871340	11873780	11871340..11873416; 11871340..11872336; 11871340..11871853
	cel-mir-87	12037052	12039017	12038660..12039017
	cel-mir-81	2430607	2431221	2430607..2431193
	cel-mir-82	2435223	2435456	2434733..2435456;
	cel-mir-230	5803780	5805027	5803780..5804930; 5803780..5804629; 5803780..5804024
	cel-mir-360	5919100	5919835	5919100..5919419; 5919100..5919256
	cel-mir-240 cel-mir-786 cel-mir-8203	7882579	7884442	7882579..7884762; 882579..7884589; 7882579..7882732
	cel-mir-54	13144929	13145245	n/a
	cel-mir-47	13920780	13923096	13921165..13922652
	cel-let-7	14743589	14745311	14743980..14745311
	cel-mir-84	16020487	16023981	n/a
	cel-mir-2217b-1	17605528	17609347	n/a
	cel-mir-357	8580353	8580828	n/a
Does not overlap with protein coding gene	cel-mir-235	6161953	6163534	6162265..6163534
	cel-mir-1	6172523	6173053	n/a
	cel-mir-245	7887336	7887486	n/a
	cel-mir-1817	10868839	10869625	n/a
	cel-mir-8199	2136652	2136923	n/a
	cel-mir-60	6327740	6328831	n/a
	cel-mir-8188	11840104	11840767	n/a
	cel-mir-45	11880601	11881150	n/a
	cel-mir-77	12519089	12519906	12519089..12519400
	cel-mir-2217b-2	13618472	13622328	13618472..13621025
	cel-mir-231	7543475	7546436	7545040..7546436
	cel-mir-80	8863917	8865652	n/a
	cel-mir-238	8867025	8868452	8867314..8868452
	cel-mir-46	13660218	13662466	n/a
	cel-mir-8189	2078879	2079130	n/a
	cel-mir-228	5561902	5563327	n/a
	cel-mir-1820	7767231	7767576	n/a
	cel-mir-83	7841425	7842227	7841425..7841838
	cel-mir-246	10940099	10940280	n/a
	cel-mir-59	11308616	11309326	11309148..11309326
	cel-mir-52	14033873	14034553	n/a
	cel-mir-2217b-3	14451423	14455223	n/a
	cel-mir-2217b-4	15798890	15799167	n/a
	cel-mir-255	5440502	5440968	n/a
	cel-lsy-6	10647182	10647527	10647182..10647430; 10647182..10647348
	cel-mir-48 cel-mir-241	14361129	14367340	14364530..14367340
	cel-mir-73	2368707	2369014	n/a
	cel-mir-75	2372260	2373222	2372260..2372559
	cel-mir-264	2818993	2819581	n/a
	cel-mir-34	2969467	2969993	n/a
	cel-mir-784	8074873	8075040	n/a
	cel-mir-788	8484604	8485409	n/a
	cel-mir-49	9989152	9989436	n/a
	cel-mir-251	11004702	11004984	n/a
	cel-mir-239b	11790629	11791445	n/a
	cel-mir-239a	11792126	11793527	n/a
	cel-mir-785	3004449	3005027	na
	cel-mir-43 cel-mir-44	11889865	11890374	n/a

Chapter 4 - Conclusions

Precisely controlled gene expression is an essential part of cellular and organism development and homeostasis. Disruption of appropriate gene expression can result in a variety of diseases and disorders. Therefore, it is unsurprising that cells employ complex gene regulatory networks to ensure that gene expression is precisely controlled. RNA-binding proteins and miRNAs are known to be important regulators of gene expression, but we are only beginning to characterize the mechanisms through which RBPs and miRNAs coordinate and influence each other in their gene regulatory activity.

Recent research efforts have focused on filling this gap in our knowledge by identifying RNA binding proteins that physically and functionally interact with the miRNA pathways to modulate their gene regulatory effects (Van Nostrand *et al*. 2016; Gebauer *et al.* 2021). These types of studies have ranged from large computational efforts (Kim et al., 2021) to more focused functional studies. In Chapter 1, I review the currently known mechanisms of RPB-miRNA coordination in regulation of gene expression. In Chapter 2, I report on functional interactions between KH domain-containing subclass of RBPs and several *C. elegans* miRNAs and begin to dissect the genetic mechanisms through which they might interact (Appendix B). In Chapter 3, using a conditional knockdown system targeting a key component of miRNA biogenesis pathway, DRSH-1, I annotate a significant proportion of primary miRNA transcripts, thus facilitating future studies in pri-miRNA transcription regulation and the roles of RBPs in this process. I also report a novel set of previously unannotated transcripts that appear to depend on DRSH-1 processing, thereby expanding the *C. elegans* transcriptome. The transcript annotation work has led to several interesting biological questions, including ones about the nature of the previously unknown transcripts, the mechanisms of primary miRNA processing,

and the roles that overlapping loci, especially antisense RNAs, play in regulation of miRNA biogenesis.

Given the number of the RBP and miRNA regulators and the combinatorial possibilities among them that can lead to distinct gene regulatory outcomes, it is important to continue characterizing the genetic and molecular mechanisms by which these two classes of post-transcriptional gene regulators interact. To that end, it is important to recognize that functional interactions among RBPs and miRNAs can be directed through distinct mechanisms that range from physical complex association to indirect targeting of the same developmental pathways.

To better understand the mechanisms of RBP-miRNA interactions, future efforts should be focused on hypotheses-driven experiments that can distinguish between the models proposed in Chapter 2, outlining potential mechanisms of interaction between KH RBPs and miRNAs. The most direct mechanism by which RBPs can be coordinating with miRNAs is by physically interacting with miRISC components to regulate the activity of the complex (Figure 2.10F). To test this prediction, immunoprecipitation of miRISC components like ALG-1 or AIN-1 can be used to test specific RBP co-precipitation, with reciprocal IPs of any RBPs of interest performed to confirm the interactions (Li *et al*. 2019). IPs can also be done in the absence or presence of RNase to test whether potential protein-protein interactions are RNA dependent. Additionally, fluorescent tag co-localization and SPLIT-GFP assays can be used to further confirm physical RBP-miRISC interactions and have the added benefit of identifying sub-cellular co-localization (Foglieni *et al*. 2017). *In vitro* interactions between purified Argonautes and RBPs can test for direct interaction. ALG-1 Argonaute IPs in the absence of the RBP followed by assays aimed at assessing miRISC activity (ex. ALG-1 interaction with miRNAs, mRNA targets, downstream

effectors, etc) can further refine the mechanism by which a direct interactor of miRISC can affect miRISC activity.

In addition to more direct models of RBP-miRNA interactions, RBPs may indirectly coordinate with miRNAs to ultimately regulate gene expression by modulating their targets (Figure 2.10A,B). Here, RBPs regulate expression of target mRNAs independently of miRISC activity. CLIP experiments can be employed to identify precise RNA targets and resolve RBP binding sites (Hafner *et al*. 2021). RNAseq could be performed in the absence of the RBP of interest to understand how the transcriptome is affected by RBP depletion, and how depletion of certain miRNA targets correlates with miRNA targeting and downregulation.

In contrast to RBPs that exclusively effect mRNA expression levels, CLIP studies could also identify small RNA binding partners, as well as differentially spliced RNA variants of small RNA transcripts. Figure 2.10.C-D outlines two models by which RBPs can modulate miRNA biogenesis by coordinating splicing processes to modulate the expression of spliced miRNAs. These models predict that RBPs may coordinate with the spliceosome to contribute to the splicing of primary miRNA transcripts or the generation of mirtrons from intron lariats (Figure 2.10.C), or affect alternative splicing of miRNA targets in order to allow differential targeting (Figure 2.10.D) Testing the first of these two models would require quantification of both spliced primary miRNA transcripts (like pri-let-7) and known mirtrons by RT-qPCR or RNAseq in a KH RBP null or depleted mutant. To fully assess primary miRNA transcripts levels on a genome wide scale, it may also be necessary to conditionally deplete Drosha using the auxin-inducible degron (see Chapter 3) to accumulate pri-miRNAs that would normally be absent in RNA samples. To test the second model, RNAseq could be utilized to uncover changes in mRNA

structures, such as transcripts with differentially spliced 3'UTRs, potentially altering the availability of miRNA target sites.

While not directly outlined by a model, we must also consider the possibility that RBPs are directly interacting with miRNA biogenesis factors like Drosha or Dicer to affect their activity. In these cases, we would predict that the RBP would physically interact with the biogenesis factors, which could be detected by IP of the RBP and probing for Drosha/Dicer co-precipitation. Reciprocal IPs can be similarly performed. Direct RBP-miRNA interactions can be confirmed by 2' *O*-methyl miRNA pulldowns (Fabian *et al*. 2009). Lastly, to confirm a RBP's role in modulating miRNA biogenesis, we could assay the effects of RBPs depletion on the processing of the appropriate miRNA intermediates using RT-qPCR or small RNAseq. Overall, if RBPs are affecting miRNA biogenesis in some way, we would expect to see an overall depletion of the appropriate subset of miRNA intermediates, which would ultimately disrupt downstream miRNA targeting and produce the developmental defects observed in Chapter 2.

Overall, the accurate identification of miRNA targets genes and characterization of physical relationships among RBP and miRNA sites will be an important step forward. These approaches will need to continue on a genome-wide scale as well as through careful mechanistic characterizations of individual interactions. A major focus in the future will not only be teasing apart the mechanisms by which individual RPBs are able to interact with miRNAs and how they modulate their activity, but also how these interactions fit within a larger network of gene regulatory mechanisms. A better understanding of the role that miRNA play in the overall context of gene regulatory mechanisms and how they are controlled will better allow us to identify and treat pathologies stemming from miRNA dysfunction.

References

Abbott, A. L., E. Alvarez-Saavedra, E. A. Miska, N. C. Lau, D. P. Bartel *et al.*, 2005 The let-7 MicroRNA Family Members mir-48, mir-84, and mir-241 Function Together to Regulate Developmental Timing in *Caenorhabditis elegans*. Dev Cell 9: 403–414.

Abrahante, J. E., A. L. Daul, M. Li, M. L. Volk, J. M. Tennessen *et al.*, 2003 The *Caenorhabditis elegans* hunchback-like Gene lin-57/hbl-1 Controls Developmental Time and Is Regulated by MicroRNAs. Dev Cell 4: 625–637.

Abdelmohsen, K., K. Tominaga-Yamanaka, S. Srikantan, J.-H. Yoon, M.-J. Kang *et al.*, 2012 RNA-binding protein AUF1 represses Dicer expression. Nucleic Acids Res 40: 11531–11544.

Abe, M., A. Naqvi, G.-J. Hendriks, V. Feltzin, Y. Zhu *et al.*, 2014 Impact of age-associated increase in 2′-O-methylation of miRNAs on aging and neurodegeneration in Drosophila. Gene Dev 28: 44–57.

Ahuja, D., A. Goyal, and P. S. Ray, 2016 Interplay between RNA-binding protein HuR and microRNA-125b regulates p53 mRNA translation in response to genotoxic stress. Rna Biol 13: 1–14.

Akay, A., A. Craig, N. Lehrbach, M. Larance, E. Pourkarimi *et al.*, 2013 RNA-binding protein GLD-1/quaking genetically interacts with the mir-35 and the let-7 miRNA pathways in *Caenorhabditis elegans*. Open Biol 3: 130151–130151.

Alarcón, C. R., H. Lee, H. Goodarzi, N. Halberg, and S. F. Tavazoie, 2015 N6-methyladenosine marks primary microRNAs for processing. Nature 519: 482–485.

Alles, J., T. Fehlmann, U. Fischer, C. Backes, V. Galata *et al.*, 2019 An estimate of the total number of true human miRNAs. Nucleic Acids Res 47: gkz097.

Ambros, V., and G. Ruvkun, 2018 Recent Molecular Genetic Explorations of *Caenorhabditis elegans* MicroRNAs. Genetics 209: 651–673.

Auyeung, V. C., I. Ulitsky, S. E. McGeary, and D. P. Bartel, 2013 Beyond Secondary Structure: Primary-Sequence Determinants License Pri-miRNA Hairpins for Processing. Cell 152: 844–858.

Babiarz, J. E., R. Hsu, C. Melton, M. Thomas, E. M. Ullian *et al.*, 2011 A role for noncanonical microRNAs in the mammalian brain revealed by phenotypic differences in Dgcr8 versus Dicer1 knockouts and small RNA sequencing. Rna 17: 1489–1501.

Bachler, M., R. Schroder, and U. V. Ahsen, 1999 StreptoTag: A novel method for the isolation of RNA-binding proteins. Rna 5: 1509–1516.

Ballarino, M., F. Pagano, E. Girardi, M. Morlando, D. Cacchiarelli *et al.*, 2009 Coupled RNA Processing and Transcription of Intergenic Primary MicroRNAs ▿ †. Mol Cell Biol 29: 5632–5638.

Baltz, A. G., M. Munschauer, B. Schwanhäusser, A. Vasile, Y. Murakawa *et al.*, 2012 The mRNA-Bound Proteome and Its Global Occupancy Profile on Protein-Coding Transcripts. Mol Cell 46: 674–690.

Bartel, D. P., 2018 Metazoan MicroRNAs. Cell 173: 20–51.

Bartel, D. P., 2004 MicroRNAs Genomics, Biogenesis, Mechanism, and Function. Cell 116: 281–297.

Bartel, D. P., 2009 MicroRNAs: Target Recognition and Regulatory Functions. Cell 136: 215–233.

Behm-Ansmant, I., J. Rehwinkel, T. Doerks, A. Stark, P. Bork *et al.*, 2006 mRNA degradation by miRNAs and GW182 requires both CCR4:NOT deadenylase and DCP1:DCP2 decapping complexes. Gene Dev 20: 1885–1898.

Bicker, S., S. Khudayberdiev, K. Weiß, K. Zocher, S. Baumeister *et al.*, 2013 The DEAH-box helicase DHX36 mediates dendritic localization of the neuronal precursor-microRNA-134. Gene Dev 27: 991–996.

Blazie, S. M., H. C. Geissel, H. Wilky, R. Joshi, J. Newbern *et al.*, 2017 Alternative Polyadenylation Directs Tissue-Specific miRNA Targeting in Caenorhabditis elegans Somatic Tissues. Genetics 206: 757–774.

Bolger, A. M., M. Lohse, and B. Usadel, 2014 Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30: 2114–2120.

Bosco, D. A., 2018 Translation dysregulation in neurodegenerative disorders. Proc National Acad Sci 115: 201818493.

Brown, K. C., J. M. Svendsen, R. M. Tucci, B. E. Montgomery, and T. A. Montgomery, 2017 ALG-5 is a miRNA-associated Argonaute required for proper developmental timing in the Caenorhabditis elegans germline. Nucleic Acids Res 45: gkx536-.

Buels, R., E. Yao, C. M. Diesh, R. D. Hayes, M. Munoz-Torres *et al.*, 2016 JBrowse: a dynamic web platform for genome visualization and analysis. Genome Biol 17: 66.

Buxbaum, A. R., G. Haimovich, and R. H. Singer, 2015 In the right place at the right time: visualizing and understanding mRNA localization. Nat Rev Mol Cell Bio 16: 95–109.

Casamassimi, A., and A. Ciccodicola, 2019 Transcriptional Regulation: Molecules, Involved Mechanisms, and Misregulation. Int J Mol Sci 20: 1281.

Castello, A., B. Fischer, K. Eichelbaum, R. Horos, B. M. Beckmann *et al.*, 2012 Insights into RNA Biology from an Atlas of Mammalian mRNA-Binding Proteins. Cell 149: 1393–1406.

Chang, T.-C., M. Pertea, S. Lee, S. L. Salzberg, and J. T. Mendell, 2015 Genome-wide annotation of microRNA primary transcript structures reveals novel regulatory mechanisms. Genome Res 25: 1401–1409.

Charenton, C., and M. Graille, 2018 mRNA decapping: finding the right structures. Philosophical Transactions Royal Soc B Biological Sci 373: 20180164.

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

Chatterjee, S., and H. Großhans, 2009 Active turnover modulates mature microRNA activity in Caenorhabditis elegans. Nature 461: 546–549.

Chekulaeva, M., H. Mathys, J. T. Zipprich, J. Attig, M. Colic *et al.*, 2011 miRNA repression involves GW182-mediated recruitment of CCR4–NOT through conserved W-containing motifs. Nat Struct Mol Biol 18: 1218–1226.

Chen, M., and J. L. Manley, 2009 Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches. Nat Rev Mol Cell Bio 10: 741–754.

Chen, A.-J., J.-H. Paik, H. Zhang, S. A. Shukla, R. Mortensen *et al.*, 2012 STAR RNA-binding protein Quaking suppresses cancer via stabilization of specific miRNA. Gene Dev 26: 1459–1472.

Cheray, M., A. Etcheverry, C. Jacques, R. Pacaud, G. Bougras-Cartron *et al.*, 2020 Cytosine methylation of mature microRNAs inhibits their functions and is associated with poor prognosis in glioblastoma multiforme. Mol Cancer 19: 36.

Choi, S., C. Park, K. E. Kim, and K. K. Kim, 2017 An in vitro technique to identify the RNA binding-site sequences for RNA-binding proteins. Biotechniques 63: 28–33.

Ciechanowska, K., M. Pokornowska, and A. Kurzyńska-Kokorniak, 2021 Genetic Insight into the Domain Structure and Functions of Dicer-Type Ribonucleases. Int J Mol Sci 22: 616.

Cieniková, Z., F. F. Damberger, J. Hall, F. H.-T. Allain, and C. Maris, 2014 Structural and Mechanistic Insights into Poly(uridine) Tract Recognition by the hnRNP C RNA Recognition Motif. J Am Chem Soc 136: 14536–14544.

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

Cook, K. B., T. R. Hughes, and Q. D. Morris, 2015 High-throughput characterization of protein–RNA interactions. Brief Funct Genomics 14: 74–89.

Chu, J. S.-C., S.-Y. Chua, K. Wong, A. M. Davison, R. Johnsen *et al.*, 2014 High-throughput capturing and characterization of mutations in essential genes of *Caenorhabditis elegans*. Bmc Genomics 15: 361.

D'Ambrogio, A., W. Gu, T. Udagawa, C. C. Mello, and J. D. Richter, 2012 Specific miRNA Stabilization by Gld2-Catalyzed Monoadenylation. Cell Reports 2: 1537–1545.

Damianov, A., Y. Ying, C.-H. Lin, J.-A. Lee, D. Tran *et al.*, 2016 Rbfox Proteins Regulate Splicing as Part of a Large Multiprotein Complex LASR. Cell 165: 606–619.

Darbelli, L., and S. Richard, 2016 Emerging functions of the Quaking RNA-binding proteins and link to human diseases. Wiley Interdiscip Rev Rna 7: 399–412.

Dassi, E., 2017 Handshakes and Fights: The Regulatory Interplay of RNA-Binding Proteins. Frontiers Mol Biosci 4: 67.

Degrauwe, N., T. B. Schlumpf, M. Janiszewska, P. Martin, A. Cauderay *et al.*, 2016 The RNA Binding Protein IMP2 Preserves Glioblastoma Stem Cells by Preventing let-7 Target Gene Silencing. Cell Reports 15: 1634–1647.

Denli, A. M., B. B. J. Tops, R. H. A. Plasterk, R. F. Ketting, and G. J. Hannon, 2004 Processing of primary microRNAs by the Microprocessor complex. Nature 432: 231–235.

Díaz-Muñoz, M. D., and M. Turner, 2018 Uncovering the Role of RNA-Binding Proteins in Gene Expression in the Immune System. Front Immunol 9: 1094.

Dominguez, D., P. Freese, M. S. Alexis, A. Su, M. Hochman *et al.*, 2018 Sequence, Structure, and Context Preferences of Human RNA Binding Proteins. Mol Cell 70: 854-867.e9.

Drake, M., T. Furuta, K. M. Suen, G. Gonzalez, B. Liu *et al.*, 2014 A Requirement for ERK-Dependent Dicer Phosphorylation in Coordinating Oocyte-to-Embryo Transition in C. elegans. Dev Cell 31: 614–628.

Duchaine, T. F., and M. R. Fabian, 2019 Mechanistic Insights into MicroRNA-Mediated Gene Silencing. Csh Perspect Biol 11: a032771.

Duchaine, T. F., J. A. Wohlschlegel, S. Kennedy, Y. Bei, D. Conte *et al.*, 2006 Functional Proteomics Reveals the Biochemical Niche of C. elegans DCR-1 in Multiple Small-RNA-Mediated Pathways. Cell 124: 343–354.

Elkayam, E., C.-D. Kuhn, A. Tocilj, A. D. Haase, E. M. Greene *et al.*, 2012 The Structure of Human Argonaute-2 in Complex with miR-20a. Cell 150: 100–110.

Ecsedi, M., M. Rausch, and H. Großhans, 2015 The let-7 microRNA Directs Vulval Development through a Single Target. Dev Cell 32: 335–344.

Eulalio, A., I. Behm-Ansmant, D. Schweizer, and E. Izaurralde, 2007 P-Body Formation Is a Consequence, Not the Cause, of RNA-Mediated Gene Silencing ▿ †. Mol Cell Biol 27: 3970–3981.

Fabian, M. R., M. K. Cieplak, F. Frank, M. Morita, J. Green *et al.*, 2011 miRNA-mediated deadenylation is orchestrated by GW182 through two conserved motifs that interact with CCR4–NOT. Nat Struct Mol Biol 18: 1211–1217.

Faghihi, M. A., and C. Wahlestedt, 2009 Regulatory roles of natural antisense transcripts. Nat Rev Mol Cell Bio 10: 637–643.

Fang, W., and D. P. Bartel, 2015 The Menu of Features that Define Primary MicroRNAs and Enable De Novo Design of MicroRNA Genes. Mol Cell 60: 131–145.

Feng, Y., X. Zhang, Q. Song, T. Li, and Y. Zeng, 2011 Drosha processing controls the specificity and efficiency of global microRNA expression. Biochimica Et Biophysica Acta Bba - Gene Regul Mech 1809: 700–707.

Flamand, M. N., E. Wu, A. Vashisht, G. Jannot, B. D. Keiper *et al.*, 2016 Poly(A)-binding proteins are required for microRNA-mediated silencing and to promote target deadenylation in C. elegans. Nucleic Acids Res 44: 5924–5935.

Flynt, A. S., J. C. Greimann, W.-J. Chung, C. D. Lima, and E. C. Lai, 2010 MicroRNA Biogenesis via Splicing and Exosome-Mediated Trimming in Drosophila. Mol Cell 38: 900–907.

Fox, A. H., Y. W. Lam, A. K. L. Leung, C. E. Lyon, J. Andersen *et al.*, 2002 Paraspeckles A Novel Nuclear Domain. Curr Biol 12: 13–25.

Frank, F., N. Sonenberg, and B. Nagar, 2010 Structural basis for 5′-nucleotide base-specific recognition of guide RNA by human AGO2. Nature 465: 818–822.

Friedländer, M. R., S. D. Mackowiak, N. Li, W. Chen, and N. Rajewsky, 2012 miRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades. Nucleic Acids Res 40: 37–52.

Fukuda, T., K. Yamagata, S. Fujiyama, T. Matsumoto, I. Koshida *et al.*, 2007 DEAD-box RNA helicase subunits of the Drosha complex are required for processing of rRNA and a subset of microRNAs. Nat Cell Biol 9: 604–611.

Ge, X., X. Zhao, A. Nakagawa, X. Gong, R. R. Skeen-Gaar *et al.*, 2014 A novel mechanism underlies caspase-dependent conversion of the dicer ribonuclease into a deoxyribonuclease during apoptosis. Cell Res 24: 218–232.

Gebauer, F., T. Schwarzl, J. Valcárcel, and M. W. Hentze, 2021 RNA-binding proteins in human genetic disease. Nat Rev Genet 22: 185–198.

Gebert, L. F. R., and I. J. MacRae, 2019 Regulation of microRNA function in animals. Nat Rev Mol Cell Bio 20: 21–37.

Gerstberger, S., M. Hafner, and T. Tuschl, 2014 A census of human RNA-binding proteins. Nat Rev Genet 15: 829–845.

Geuens, T., D. Bouhy, and V. Timmerman, 2016 The hnRNP family: insights into their role in health and disease. Hum Genet 135: 851–867.

Ghosh, A., K. Mizuno, S. S. Tiwari, P. Proitsi, B. G. Perez-Nievas *et al.*, 2020 Alzheimer's disease-related dysregulation of mRNA translation causes key pathological features with ageing. Transl Psychiat 10: 192.

Glorian, V., G. Maillot, S. Polès, J. S. Iacovoni, G. Favre *et al.*, 2011 HuR-dependent loading of miRNA RISC to the mRNA encoding the Ras-related small GTPase RhoB controls its translation during UV-induced apoptosis. Cell Death Differ 18: 1692–1701.

Goh, W.-S. S., J. W. E. Seah, E. J. Harrison, C. Chen, C. M. Hammell *et al.*, 2014 A genome-wide RNAi screen identifies factors required for distinct stages of *C. elegans* piRNA biogenesis. Gene Dev 28: 797–807.

Goss, D. J., and F. E. Kleiman, 2013 Poly(A) binding proteins: are they all created equal? Wiley Interdiscip Rev Rna 4: 167–179.

Graber, J. H., J. Salisbury, L. N. Hutchins, and T. Blumenthal, 2007 C. elegans sequences that control trans-splicing and operon pre-mRNA processing. Rna 13: 1409–1426.

Grabherr, M. G., B. J. Haas, M. Yassour, J. Z. Levin, D. A. Thompson *et al.*, 2011 Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol 29: 644–652.

Grimson, A., K. K.-H. Farh, W. K. Johnston, P. Garrett-Engele, L. P. Lim *et al.*, 2007 MicroRNA Targeting Specificity in Mammals: Determinants beyond Seed Pairing. Mol Cell 27: 91–105.

Grishin, N. V., 2001 KH domain: one motif, two folds. Nucleic Acids Res 29: 638–643.

Grishok, A., J. L. Sinskey, and P. A. Sharp, 2005 Transcriptional silencing of a transgene by RNAi in the soma of *C. elegans*. Gene Dev 19: 683–696.

Gu, S., L. Jin, Y. Zhang, Y. Huang, F. Zhang *et al.*, 2012a The Loop Position of shRNAs and Pre-miRNAs Is Critical for the Accuracy of Dicer Processing In Vivo. Cell 151: 900–911.

Gu, W., H.-C. Lee, D. Chaves, E. M. Youngman, G. J. Pazour *et al.*, 2012b CapSeq and CIP-TAP Identify Pol II Start Sites and Reveal Capped Small RNAs as C. elegans piRNA Precursors. Cell 151: 1488–1500.

Guerra-Assunção, J. A., and A. J. Enright, 2012 Large-scale analysis of microRNA evolution. Bmc Genomics 13: 218.

Gupta, I., S. Clauder-Münster, B. Klaus, A. I. Järvelin, R. S. Aiyar *et al.*, 2014 Alternative polyadenylation diversifies post-transcriptional regulation by selective RNA–protein interactions. Mol Syst Biol 10: 719.

Ha, M., and V. N. Kim, 2014a Regulation of microRNA biogenesis. Nat Rev Mol Cell Bio 15: nrm3838.

Ha, M., and V. N. Kim, 2014b Regulation of microRNA biogenesis. Nat Rev Mol Cell Bio 15: 509–524.

Haas, B. J., S. L. Salzberg, W. Zhu, M. Pertea, J. E. Allen *et al.*, 2008 Automated eukaryotic gene structure annotation using EVidenceModeler and the Program to Assemble Spliced Alignments. Genome Biol 9: R7.

Hafner, M., M. Katsantoni, T. Köster, J. Marks, J. Mukherjee *et al.*, 2021 CLIP and complementary methods. Nat Rev Methods Primers 1: 20.

Hafner, M., M. Landthaler, L. Burger, M. Khorshid, J. Hausser *et al.*, 2010 Transcriptome-wide Identification of RNA-Binding Protein and MicroRNA Target Sites by PAR-CLIP. Cell 141: 129–141.

Hammell, C. M., I. Lubin, P. R. Boag, T. K. Blackwell, and V. Ambros, 2009 *nhl-2* Modulates MicroRNA Activity in *Caenorhabditis elegans.* Cell 136: 926–938.

Han, J., Y. Lee, K.-H. Yeom, Y.-K. Kim, H. Jin *et al.*, 2004 The Drosha-DGCR8 complex in primary microRNA processing. Gene Dev 18: 3016–3027.

Han, J., Y. Lee, K.-H. Yeom, J.-W. Nam, I. Heo *et al.*, 2006 Molecular Basis for the Recognition of Primary microRNAs by the Drosha-DGCR8 Complex. Cell 125: 887–901.

Han, J., J. S. Pedersen, S. C. Kwon, C. D. Belair, Y.-K. Kim *et al.*, 2009 Posttranscriptional Crossregulation between Drosha and DGCR8. Cell 136: 75–84.

Han, S., D. Kim, M. Shivakumar, Y.-J. Lee, T. Garg *et al.*, 2018 The effects of alternative splicing on miRNA binding sites in bladder cancer. Plos One 13: e0190708.

Hartmuth, K., H. Urlaub, H.-P. Vornlocher, C. L. Will, M. Gentzel *et al.*, 2002 Protein composition of human prespliceosomes isolated by a tobramycin affinity-selection method. Proc National Acad Sci 99: 16719–16724.

He, Z., and Q. Yu, 2018 Identification and characterization of functional modules reflecting transcriptome transition during human neuron maturation. Bmc Genomics 19: 262.

Heintz, C., T. K. Doktor, A. Lanjuin, C. C. Escoubas, Y. Zhang *et al.*, 2017 Splicing factor 1 modulates dietary restriction and TORC1 pathway longevity in *C. elegans*. Nature 541: 102–106.

Helwak, A., and D. Tollervey, 2014 Mapping the miRNA interactome by cross-linking ligation and sequencing of hybrids (CLASH). Nat Protoc 9: 711–728.

Hentze, M. W., A. Castello, T. Schwarzl, and T. Preiss, 2018 A brave new world of RNA-binding proteins. Nat Rev Mol Cell Bio 19: 327–341.

Heo, I., M. Ha, J. Lim, M.-J. Yoon, J.-E. Park *et al.*, 2012 Mono-Uridylation of Pre-MicroRNA as a Key Step in the Biogenesis of Group II let-7 MicroRNAs. Cell 151: 521–532.

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

Hernández, G., M. Miron, H. Han, N. Liu, J. Magescas *et al.*, 2013 Mextli Is a Novel Eukaryotic Translation Initiation Factor 4E-Binding Protein That Promotes Translation in *Drosophila melanogaster*. Mol Cell Biol 33: 2854–2864.

Hershey, J. W. B., N. Sonenberg, and M. B. Mathews, 2012 Principles of Translational Control: An Overview. Csh Perspect Biol 4: a011528.

Höck, J., L. Weinmann, C. Ender, S. Rüdel, E. Kremmer *et al.*, 2007 Proteomic and functional analysis of Argonaute-containing mRNA–protein complexes in human cells. Embo Rep 8: 1052–1060.

Hogan, D. J., D. P. Riordan, A. P. Gerber, D. Herschlag, and P. O. Brown, 2008 Diverse RNA-Binding Proteins Interact with Functionally Related Sets of RNAs, Suggesting an Extensive Regulatory System. Plos Biol 6: e255.

Hooshmandi, M., C. Wong, and A. Khoutorsky, 2020 Dysregulation of translational control signaling in autism spectrum disorders. Cell Signal 75: 109746.

Hu, H. Y., Z. Yan, Y. Xu, H. Hu, C. Menzel *et al.*, 2009 Sequence features associated with microRNA strand selection in humans and flies. Bmc Genomics 10: 413.

Hunt-Newbury, R., R. Viveiros, R. Johnsen, A. Mah, D. Anastas *et al.*, 2007 High-Throughput In Vivo Analysis of Gene Expression in Caenorhabditis elegans. Plos Biol 5: e237.

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

Hutchinson, J. N., A. W. Ensminger, C. M. Clemson, C. R. Lynch, J. B. Lawrence *et al.*, 2007 A screen for nuclear transcripts identifies two linked noncoding RNAs associated with SC35 splicing domains. Bmc Genomics 8: 39.

Ibrahim, F., L. A. Rymarquis, E.-J. Kim, J. Becker, E. Balassa *et al.*, 2010 Uridylation of mature miRNAs and siRNAs by the MUT68 nucleotidyltransferase promotes their degradation in Chlamydomonas. Proc National Acad Sci 107: 3906–3911.

Ipsaro, J. J., and L. Joshua-Tor, 2015 From guide to target: molecular insights into eukaryotic RNA-interference machinery. Nat Struct Mol Biol 22: 20–28.

Iwasaki, S., M. Kobayashi, M. Yoda, Y. Sakaguchi, S. Katsuma *et al.*, 2010 Hsc70/Hsp90 Chaperone Machinery Mediates ATP-Dependent RISC Loading of Small RNA Duplexes. Mol Cell 39: 292–299.

Jafarnejad, S. M., C. Chapat, E. Matta-Camacho, I. A. Gelbart, G. G. Hesketh *et al.*, 2018 Translational control of ERK signaling through miRNA/4EHP-directed silencing. Elife 7: e35034.

Janga, S. C., 2012 From specific to global analysis of posttranscriptional regulation in eukaryotes: posttranscriptional regulatory networks. Brief Funct Genomics 11: 505–521.

Jankowsky, E., and M. E. Harris, 2015 Specificity and nonspecificity in RNA–protein interactions. Nat Rev Mol Cell Bio 16: 533–544.

Ji, L., and X. Chen, 2012 Regulation of small RNA stability: methylation and beyond. Cell Res 22: 624–636.

Jiang, L., C. Shao, Q.-J. Wu, G. Chen, J. Zhou *et al.*, 2017 NEAT1 scaffolds RNA-binding proteins and the Microprocessor to globally enhance pri-miRNA processing. Nat Struct Mol Biol 24: 816–824.

Jones, M. R., L. J. Quinton, M. T. Blahna, J. R. Neilson, S. Fu *et al.*, 2009 Zcchc11-dependent uridylation of microRNA directs cytokine expression. Nat Cell Biol 11: 1157–1163.

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

Kaletsky, R., V. Yao, A. Williams, A. M. Runnels, A. Tadych *et al.*, 2018 Transcriptome analysis of adult Caenorhabditis elegans cells reveals tissue-specific gene and isoform expression. Plos Genet 14: e1007559.

Kamath, R. S., A. G. Fraser, Y. Dong, G. Poulin, R. Durbin *et al.*, 2003 Systematic functional analysis of the *Caenorhabditis elegans* genome using RNAi. Nature 421: 231–237.

Kamath, R. S., M. Martinez-Campos, P. Zipperlen, A. G. Fraser, and J. Ahringer, 2000 Effectiveness of specific RNA-mediated interference through ingested double-stranded RNA in *Caenorhabditis elegans*. Genome Biol 2: research0002.1.

Kamelgarn, M., J. Chen, L. Kuang, H. Jin, E. J. Kasarskis *et al.*, 2018 ALS mutations of FUS suppress protein translation and disrupt the regulation of nonsense-mediated decay. Proc National Acad Sci 115: 201810413.

Kapelle, W. S., and V. Reinke, 2011 *C. elegans* meg-1 and meg-2 differentially interact with nanos family members to either promote or inhibit germ cell proliferation and survival. Genesis 49: 380–391.

Kato, M., T. W. Han, S. Xie, K. Shi, X. Du *et al.*, 2012 Cell-free Formation of RNA Granules: Low Complexity Sequence Domains Form Dynamic Fibers within Hydrogels. Cell 149: 753–767.

Katoh, T., Y. Sakaguchi, K. Miyauchi, T. Suzuki, S. Kashiwabara *et al.*, 2009 Selective stabilization of mammalian microRNAs by 3′ adenylation mediated by the cytoplasmic poly(A) polymerase GLD-2. Gene Dev 23: 433–438.

Kawahara, Y., and A. Mieda-Sato, 2012 TDP-43 promotes microRNA biogenesis as a component of the Drosha and Dicer complexes. Proc National Acad Sci 109: 3347–3352.

Kawamata, T., H. Seitz, and Y. Tomari, 2009 Structural determinants of miRNAs for RISC loading and slicer-independent unwinding. Nat Struct Mol Biol 16: 953–960.

Kawamata, T., and Y. Tomari, 2010 Making RISC. Trends Biochem Sci 35: 368–376.

Kim, J. K., H. W. Gabel, R. S. Kamath, M. Tewari, A. Pasquinelli *et al.*, 2005 Functional Genomic Analysis of RNA Interference in *C. elegans*. Science 308: 1164–1167.

Kim, K. Y., Y. J. Hwang, M.-K. Jung, J. Choe, Y. Kim *et al.*, 2014 A multifunctional protein EWS regulates the expression of Drosha and microRNAs. Cell Death Differ 21: 136–145.

Kim, Y., and V. N. Kim, 2007 Processing of intronic microRNAs. Embo J 26: 775–783.

Kim, S., S. Kim, H. R. Chang, D. Kim, J. Park *et al.*, 2021 The regulatory impact of RNA-binding proteins on microRNA targeting. Nat Commun 12: 5057.

Kim, Y.-K., B. Kim, and V. N. Kim, 2016a Re-evaluation of the roles of DROSHA, Exportin 5, and DICER in microRNA biogenesis. Proc National Acad Sci 113: E1881–E1889.

Kim, Y.-K., B. Kim, and V. N. Kim, 2016b Re-evaluation of the roles of DROSHA, Exportin 5, and DICER in microRNA biogenesis. Proc National Acad Sci 113: E1881–E1889.

Kim, H. H., Y. Kuwano, S. Srikantan, E. K. Lee, J. L. Martindale *et al.*, 2009 HuR recruits let-7/RISC to repress c-Myc expression. Gene Dev 23: 1743–1748.

Kingston, E. R., and D. P. Bartel, 2019 Global analyses of the dynamics of mammalian microRNA metabolism. Genome Res 29: 1777–1790.

Konno, M., J. Koseki, A. Asai, A. Yamagata, T. Shimamura *et al.*, 2019 Distinct methylation levels of mature microRNAs in gastrointestinal cancers. Nat Commun 10: 3888.

Kosugi, S., M. Hasebe, M. Tomita, and H. Yanagawa, 2009 Systematic identification of cell cycle-dependent yeast nucleocytoplasmic shuttling proteins by prediction of composite motifs. P Natl Acad Sci Usa 106: 10171–6.

Kovaka, S., A. V. Zimin, G. M. Pertea, R. Razaghi, S. L. Salzberg *et al.*, 2019 Transcriptome assembly from long-read RNA-seq alignments with StringTie2. Genome Biol 20: 278.

Kovar, H., 2011 Dr. Jekyll and Mr. Hyde: The Two Faces of the FUS/EWS/TAF15 Protein Family. Sarcoma 2011: 837474.

Kosugi, S., M. Hasebe, M. Tomita, and H. Yanagawa, 2009 Systematic identification of cell cycle-dependent yeast nucleocytoplasmic shuttling proteins by prediction of composite motifs. P Natl Acad Sci Usa 106: 10171–6.

Kozomara, A., M. Birgaoanu, and S. Griffiths-Jones, 2018a miRBase: from microRNA sequences to function. Nucleic Acids Res 47: gky1141-.

Kozomara, A., M. Birgaoanu, and S. Griffiths-Jones, 2018b miRBase: from microRNA sequences to function. Nucleic Acids Res 47: gky1141-.

Kretz, M., Z. Siprashvili, C. Chu, D. E. Webster, A. Zehnder *et al.*, 2013 Control of somatic tissue differentiation by the long non-coding RNA TINCR. Nature 493: 231–235.

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

Kuang, S., and L. Wang, 2020 Identification and analysis of consensus RNA motifs binding to the genome regulator CTCF. Nar Genom Bioinform 2: lqaa031-.

Kumar, S., G. Stecher, M. Li, C. Knyaz, and K. Tamura, 2018 MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. Mol Biol Evol 35: 1547–1549.

Kundu, P., M. R. Fabian, N. Sonenberg, S. N. Bhattacharyya, and W. Filipowicz, 2012 HuR protein attenuates miRNA-mediated repression by promoting miRISC dissociation from the target RNA. Nucleic Acids Res 40: 5088–5100.

Kuroyanagi, H., 2013 Switch-like regulation of tissue-specific alternative pre-mRNA processing patterns revealed by customized fluorescence reporters. Worm 2: e23834.

Kuzuoğlu-Öztürk, D., E. Huntzinger, S. Schmidt, and E. Izaurralde, 2012 The Caenorhabditis elegans GW182 protein AIN-1 interacts with PAB-1 and subunits of the PAN2-PAN3 and CCR4-NOT deadenylase complexes. Nucleic Acids Res 40: 5651–5665.

Kwon, S. C., T. A. Nguyen, Y.-G. Choi, M. H. Jo, S. Hohng *et al.*, 2016 Structure of Human DROSHA. Cell 164: 81–90.

Lancaster, A. K., A. Nutter-Upham, S. Lindquist, and O. D. King, 2014 PLAAC: a web and command-line application to identify proteins with prion-like amino acid composition. Bioinformatics 30: 2501–2502.

Langmead, B., and S. L. Salzberg, 2012 Fast gapped-read alignment with Bowtie 2. Nat Methods 9: 357–359.

Langmead, B., C. Trapnell, M. Pop, and S. L. Salzberg, 2009 Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 10: R25.

Lee, Y., C. Ahn, J. Han, H. Choi, J. Kim *et al.*, 2003 The nuclear RNase III Drosha initiates microRNA processing. Nature 425: 415–419.

Lee, Y., M. Kim, J. Han, K. Yeom, S. Lee *et al.*, 2004 MicroRNA genes are transcribed by RNA polymerase II. Embo J 23: 4051–4060.

Lee, T. I., and R. A. Young, 2013 Transcriptional Regulation and Its Misregulation in Disease. Cell 152: 1237–1251.

Leppek, K., and G. Stoecklin, 2014 An optimized streptavidin-binding RNA aptamer for purification of ribonucleoprotein complexes identifies novel ARE-binding proteins. Nucleic Acids Res 42: e13–e13.

Letunic, I., and P. Bork, 2017 20 years of the SMART protein domain annotation resource. Nucleic Acids Res 46: gkx922-.

Léveillé, N., R. Elkon, V. Davalos, V. Manoharan, D. Hollingworth *et al.*, 2011 Selective inhibition of microRNA accessibility by RBM38 is required for p53 activity. Nat Commun 2: 513.

Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan *et al.*, 2009 The Sequence Alignment/Map format and SAMtools. Bioinformatics 25: 2078–2079.

Li, Y., X. Li, J. Yang, and Y. He, 2020 Natural antisense transcripts of MIR398 genes suppress microR398 processing and attenuate plant thermotolerance. Nat Commun 11: 5351.

Li, S., and D. J. Patel, 2016 Drosha and Dicer: Slicers cut from the same cloth. Cell Res 26: 511–512.

Li, L., I. Veksler-Lublinsky, and A. Zinovyeva, 2019 HRPK-1, a conserved KH-domain protein, modulates microRNA activity during Caenorhabditis elegans development. Plos Genet 15: e1008067.

Li, L., and A. Y. Zinovyeva, 2020 Protein Extract Preparation and Co-immunoprecipitation from Caenorhabditis elegans.

Liang, H., Z. Jiao, W. Rong, S. Qu, Z. Liao *et al.*, 2020 3′-Terminal 2′-O-methylation of lung cancer miR-21-5p enhances its stability and association with Argonaute 2. Nucleic Acids Res 48: gkaa504-.

Licatalosi, D. D., A. Mele, J. J. Fak, J. Ule, M. Kayikci *et al.*, 2008 HITS-CLIP yields genome-wide insights into brain alternative RNA processing. Nature 456: 464–469.

Lin, C., and W. O. Miles, 2019 Beyond CLIP: advances and opportunities to measure RBP–RNA and RNA–RNA interactions. Nucleic Acids Res 47: 5490–5501.

Lin, S.-Y., S. M. Johnson, M. Abraham, M. C. Vella, A. Pasquinelli *et al.*, 2003 The *C. elegans* hunchback Homolog, *hbl-1*, Controls Temporal Patterning and Is a Probable MicroRNA Target. Dev Cell 4: 639–650.

Lin, X., B. Yang, W. Liu, X. Tan, F. Wu *et al.*, 2016 Interplay between PCBP2 and miRNA modulates ARHGDIA expression and function in glioma migration and invasion. Oncotarget 7: 19483–19498.

Louloupi, A., E. Ntini, J. Liz, and U. A. Ørom, 2017 Microprocessor dynamics shows co- and post-transcriptional processing of pri-miRNAs. Rna 23: 892–898.

Lund, E., S. Güttinger, A. Calado, J. E. Dahlberg, and U. Kutay, 2004 Nuclear Export of MicroRNA Precursors. Science 303: 95–98.

Ma, L., and H. R. Horvitz, 2009 Mutations in the *Caenorhabditis elegans* U2AF Large Subunit UAF-1 Alter the Choice of a 3′ Splice Site In Vivo. Plos Genet 5: e1000708.

Mangone, M., A. P. Manoharan, D. Thierry-Mieg, J. Thierry-Mieg, T. Han *et al.*, 2010 The Landscape of C. elegans 3′UTRs. Science 329: 432–435.

Marin, V. A., and T. C. Evans, 2003 Translational repression of a *C. elegans* Notch mRNA by the STAR/KH domain protein GLD-1. Development 130: 2623–2632.

Martin, G., and W. Keller, 2007 RNA-specific ribonucleotidyl transferases. Rna 13: 1834–1849.

Martinez, N. J., M. C. Ow, J. S. Reece-Hoyes, M. I. Barrasa, V. R. Ambros *et al.*, 2008 Genome-scale spatiotemporal analysis of Caenorhabditis elegans microRNA promoter activity. Genome Res 18: 2005–2015.

Marzi, M. J., F. Ghini, B. Cerruti, S. de Pretis, P. Bonetti *et al.*, 2016 Degradation dynamics of microRNAs revealed by a novel pulse-chase approach. Genome Res 26: 554–565.

Mayya, V. K., M. N. Flamand, A. M. Lambert, S. M. Jafarnejad, J. A. Wohlschlegel *et al.*, 2021 microRNA-mediated translation repression through GYF-1 and IFE-4 in C. elegans development. Nucleic Acids Res 49: gkab162-.

Mazan-Mamczarz, K., Y. Kuwano, M. Zhan, E. J. White, J. L. Martindale *et al.*, 2009 Identification of a signature motif in target mRNAs of RNA-binding protein AUF1. Nucleic Acids Res 37: 204–214.

Medley, J. C., G. Panzade, and A. Y. Zinovyeva, 2021 microRNA strand selection: Unwinding the rules. Wiley Interdiscip Rev Rna 12: e1627.

Meister, G., 2013 Argonaute proteins: functional insights and emerging roles. Nat Rev Genet 14: 447–459.

Melo, S. A., C. Moutinho, S. Ropero, G. A. Calin, S. Rossi *et al.*, 2010a A Genetic Defect in Exportin-5 Traps Precursor MicroRNAs in the Nucleus of Cancer Cells. Cancer Cell 18: 303–315.

Melo, S. A., C. Moutinho, S. Ropero, G. A. Calin, S. Rossi *et al.*, 2010b A Genetic Defect in Exportin-5 Traps Precursor MicroRNAs in the Nucleus of Cancer Cells. Cancer Cell 18: 303–315.

Michlewski, G., and J. F. Cáceres, 2010 Antagonistic role of hnRNP A1 and KSRP in the regulation of let-7a biogenesis. Nat Struct Mol Biol 17: 1011–1018.

Michlewski, G., S. Guil, C. A. Semple, and J. F. Cáceres, 2008 Posttranscriptional Regulation of miRNAs Harboring Conserved Terminal Loops. Mol Cell 32: 383–393.

Mondol, V., B. C. Ahn, and A. E. Pasquinelli, 2015 Splicing remodels the let-7 primary microRNA to facilitate Drosha processing in Caenorhabditis elegans. Rna 21: 1396–1403.

Moore, S., A. I. Järvelin, I. Davis, G. L. Bond, and A. Castello, 2018 Expanding horizons: new roles for non-canonical RNA-binding proteins in cancer. Curr Opin Genet Dev 48: 112–120.

Mori, M. A., P. Raghavan, T. Thomou, J. Boucher, S. Robida-Stubbs *et al.*, 2012 Role of MicroRNA Processing in Adipose Tissue in Stress Defense and Longevity. Cell Metab 16: 336–347.

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

Moteki, S., and D. Price, 2002 Functional Coupling of Capping and Transcription of mRNA. Mol Cell 10: 599–609.

Muqbil, I., B. Bao, A. Abou-Samra, R. Mohammad, and A. Azmi, 2013 Nuclear Export Mediated Regulation of MicroRNAs: Potential Target for Drug Intervention. Curr Drug Targets 14: 1094–1100.

Nabel-Rosen, H., G. Volohonsky, A. Reuveny, R. Zaidel-Bar, and T. Volk, 2002 Two Isoforms of the *Drosophila* RNA Binding Protein, How, Act in Opposing Directions to Regulate Tendon Cell Differentiation. Dev Cell 2: 183–193.

Nakanishi, T., H. Kubota, N. Ishibashi, S. Kumagai, H. Watanabe *et al.*, 2006 Possible role of mouse poly(A) polymerase mGLD-2 during oocyte maturation. Dev Biol 289: 115–126.

Nelson, C., and V. Ambros, 2019 Trans-splicing of the C. elegans let-7 primary transcript developmentally regulates let-7 microRNA biogenesis and let-7 family microRNA activity. Development 146: dev172031.

Nguyen, T. A., M. H. Jo, Y.-G. Choi, J. Park, S. C. Kwon *et al.*, 2015 Functional Anatomy of the Human Microprocessor. Cell 161: 1374–1387.

Nicastro, G., I. A. Taylor, and A. Ramos, 2015 KH–RNA interactions: back in the groove. Curr Opin Struc Biol 30: 63–70.

Nicholson, C. O., M. B. Friedersdorf, L. S. Bisogno, and J. D. Keene, 2017 DO-RIP-seq to quantify RNA binding sites transcriptome-wide. Methods 118: 16–23.

Nicholson, A. L., and A. E. Pasquinelli, 2018 Tales of Detailed Poly(A) Tails. Trends Cell Biol 29: 191–200.

Nika, L., T. Gibson, R. Konkus, and X. Karp, 2016 Fluorescent Beads Are a Versatile Tool for Staging *Caenorhabditis elegans* in Different Life Histories. G3 Genes Genomes Genetics 6: 1923–1933.

Nostrand, E. L. V., P. Freese, G. A. Pratt, X. Wang, X. Wei *et al.*, 2020 A large-scale binding and functional map of human RNA-binding proteins. Nature 583: 711–719.

Nostrand, E. L. V., G. A. Pratt, A. A. Shishkin, C. Gelboin-Burkhart, M. Y. Fang *et al.*, 2016 Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP). Nat Methods 13: 508–514.

O'Brien, J., H. Hayder, Y. Zayed, and C. Peng, 2018 Overview of MicroRNA Biogenesis, Mechanisms of Actions, and Circulation. Front Endocrinol 9: 402.

Ohno, G., M. Hagiwara, and H. Kuroyanagi, 2008 STAR family RNA-binding protein ASD-2 regulates developmental switching of mutually exclusive alternative splicing *in vivo*. Gene Dev 22: 360–374.

Okamura, K., J. W. Hagen, H. Duan, D. M. Tyler, and E. C. Lai, 2007 The Mirtron Pathway Generates microRNA-Class Regulatory RNAs in Drosophila. Cell 130: 89–100.

Ouyang, H., K. Zhang, K. Fox-Walsh, Y. Yang, C. Zhang *et al.*, 2017 The RNA binding protein EWS is broadly involved in the regulation of pri-miRNA processing in mammalian cells. Nucleic Acids Res 45: gkx912-.

Palmer, R. E., T. Inoue, D. R. Sherwood, L. I. Jiang, and P. W. Sternberg, 2002 *Caenorhabditis elegans cog-1* Locus Encodes GTX/Nkx6.1 Homeodomain Proteins and Regulates Multiple Aspects of Reproductive System Development. Dev Biol 252: 202–213.

Pandolfini, L., I. Barbieri, A. J. Bannister, A. Hendrick, B. Andrews *et al.*, 2019 METTL1 Promotes let-7 MicroRNA Processing via m7G Methylation. Mol Cell 74: 1278-1290.e9.

Parry, D. H., J. Xu, and G. Ruvkun, 2007 A Whole-Genome RNAi Screen for *C. elegans* miRNA Pathway Genes. Curr Biol 17: 2013–2022.

Patro, R., G. Duggal, M. I. Love, R. A. Irizarry, and C. Kingsford, 2017 Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods 14: 417–419.

Pavelec, D. M., J. Lachowiec, T. F. Duchaine, H. E. Smith, and S. Kennedy, 2009 Requirement for the ERI/DICER Complex in Endogenous RNA Interference and Sperm Development in Caenorhabditis elegans. Genetics 183: 1283–1295.

Pérez-Ortín, J. E., P. Alepuz, S. Chávez, and M. Choder, 2013 Eukaryotic mRNA Decay: Methodologies, Pathways, and Links to Other Stages of Gene Expression. J Mol Biol 425: 3750–3775.

Peters, L. J. F., E. A. L. Biessen, M. Hohl, C. Weber, E. P. C. van der Vorst *et al.*, 2020 Small Things Matter: Relevance of MicroRNAs in Cardiovascular Disease. Front Physiol 11: 793.

Plotnikova, O., A. Baranova, and M. Skoblov, 2019 Comprehensive Analysis of Human microRNA–mRNA Interactome. Frontiers Genetics 10: 933.

Pullmann, R., H. H. Kim, K. Abdelmohsen, A. Lal, J. L. Martindale *et al.*, 2007 Analysis of Turnover and Translation Regulatory RNA-Binding Protein Expression through Binding to Cognate mRNAs ▿ . Mol Cell Biol 27: 6265–6278.

Putnam, A., M. Cassani, J. Smith, and G. Seydoux, 2019 A gel phase promotes condensation of liquid P granules in *Caenorhabditis elegans* embryos. Nat Struct Mol Biol 26: 220–226.

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

Ramanathan, M., D. F. Porter, and P. A. Khavari, 2019 Methods to study RNA–protein interactions. Nat Methods 16: 225–234.

Ramírez, F., D. P. Ryan, B. Grüning, V. Bhardwaj, F. Kilpert *et al.*, 2016 deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res 44: W160–W165.

Reichholf, B., V. A. Herzog, N. Fasching, R. A. Manzenreither, I. Sowemimo *et al.*, 2019 Time-Resolved Small RNA Sequencing Unravels the Molecular Principles of MicroRNA Homeostasis. Mol Cell 75: 756-768.e7.

Reinhart, B. J., F. J. Slack, M. Basson, A. E. Pasquinelli, J. C. Bettinger *et al.*, 2000 The 21-nucleotide *let-7* RNA regulates developmental timing in *Caenorhabditis elegans*. Nature 403: 901.

Reinke, A. W., R. Mak, E. R. Troemel, and E. J. Bennett, 2017 In vivo mapping of tissue- and subcellular-specific proteomes in *Caenorhabditis elegans*. Sci Adv 3: e1602426.

Resnick, T. D., K. A. McCulloch, and A. E. Rougvie, 2010 miRNAs give worms the time of their lives: Small RNAs and temporal control in Caenorhabditis elegans. Dev Dynam 239: 1477–1489.

Richter, J. D., and X. Zhao, 2021 The molecular biology of FMRP: new insights into fragile X syndrome. Nat Rev Neurosci 22: 209–222.

Robinson, J. T., H. Thorvaldsdóttir, W. Winckler, M. Guttman, E. S. Lander *et al.*, 2011 Integrative genomics viewer. Nat Biotechnol 29: 24–26.

Rino, J., J. M. P. Desterro, T. R. Pacheco, T. W. J. Gadella, and M. Carmo-Fonseca, 2008 Splicing Factors SF1 and U2AF Associate in Extraspliceosomal Complexes ▿ †. Mol Cell Biol 28: 3045–3057.

Rual, J.-F., J. Ceron, J. Koreth, T. Hao, A.-S. Nicot *et al.*, 2004 Toward Improving *Caenorhabditis elegans* Phenome Mapping With an ORFeome-Based RNAi Library. Genome Res 14: 2162–2168.

Ruby, J. G., C. H. Jan, and D. P. Bartel, 2007 Intronic microRNA precursors that bypass Drosha processing. Nature 448: 83–86.

Rybak-Wolf, A., M. Jens, Y. Murakawa, M. Herzog, M. Landthaler *et al.*, 2014a A Variety of Dicer Substrates in Human and C. elegans. Cell 159: 1153–1167.

Rybak-Wolf, A., M. Jens, Y. Murakawa, M. Herzog, M. Landthaler *et al.*, 2014b A Variety of Dicer Substrates in Human and C. elegans. Cell 159: 1153–1167.

Sanei, M., and X. Chen, 2015 Mechanisms of microRNA turnover. Curr Opin Plant Biol 27: 199–206.

Savva, Y. A., L. E. Rieder, and R. A. Reenan, 2012 The ADAR protein family. Genome Biol 13: 252.

Sawh, A. N., and T. F. Duchaine, 2013 A Truncated Form of Dicer Tilts the Balance of RNA Interference Pathways. Cell Reports 4: 454–463.

Schaub, M., and W. Keller, 2002 RNA editing by adenosine deaminases generates RNA and protein diversity. Biochimie 84: 791–803.

Schirle, N. T., and I. J. MacRae, 2012 The Crystal Structure of Human Argonaute2. Science 336: 1037–1040.

Schoenberg, D. R., and L. E. Maquat, 2012 Regulation of cytoplasmic mRNA decay. Nat Rev Genet 13: 246–259.

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

Shin, Y., and C. P. Brangwynne, 2017 Liquid phase condensation in cell physiology and disease. Science 357: eaaf4382.

Siomi, H., M. J. Matunis, W. M. Michael, and G. Dreyfuss, 1993 The pre-mRNA binding K protein contains a novel evolutionary conserved motif. Nucleic Acids Res 21: 1193–1198.

Skariah, G., and P. K. Todd, 2021 Translational control in aging and neurodegeneration. Wiley Interdiscip Rev Rna 12: e1628.

Slack, F. J., M. Basson, Z. Liu, V. Ambros, H. R. Horvitz *et al.*, 2000 The *lin-41* RBCC Gene Acts in the *C. elegans* Heterochronic Pathway between the let-7 Regulatory RNA and the LIN-29 Transcription Factor. Mol Cell 5: 659–669.

Slobodin, B., and J. E. Gerst, 2010 A novel mRNA affinity purification technique for the identification of interacting proteins and transcripts in ribonucleoprotein complexes. Rna 16: 2277–2290.

Song, Y., L. Li, W. Yang, Q. Fu, W. Chen *et al.*, 2020 Sense–antisense miRNA pairs constitute an elaborate reciprocal regulatory circuit. Genome Res 30: 661–672.

Song, M.-S., and J. J. Rossi, 2017a Molecular mechanisms of Dicer: endonuclease and enzymatic activity. Biochem J 474: 1603–1618.

Sönnichsen, B., L. B. Koski, A. Walsh, P. Marschall, B. Neumann *et al.*, 2005 Full-genome RNAi profiling of early embryogenesis in *Caenorhabditis elegans*. Nature 434: 462–469.

Srikantan, S., K. Abdelmohsen, E. K. Lee, K. Tominaga, S. S. Subaran *et al.*, 2011 Translational Control of TOP2A Influences Doxorubicin Efficacy. Mol Cell Biol 31: 3790–3801.

Starega-Roslan, J., T. M. Witkos, P. Galka-Marciniak, and W. J. Krzyzosiak, 2015 Sequence Features of Drosha and Dicer Cleavage Sites Affect the Complexity of IsomiRs. Int J Mol Sci 16: 8110–8127.

Subramanian, M., X. L. Li, T. Hara, and A. Lal, 2014 Regulatory Non-Coding RNAs, Methods and Protocols. Methods Mol Biology 1206: 29–37.

Sugimoto, Y., A. Vigilante, E. Darbo, A. Zirra, C. Militti *et al.*, 2015 hiCLIP reveals the in vivo atlas of mRNA secondary structures recognized by Staufen 1. Nature 519: 491–494.

Tabach, Y., A. C. Billi, G. D. Hayes, M. A. Newman, O. Zuk *et al.*, 2013 Identification of small RNA pathway genes using patterns of phylogenetic conservation and divergence. Nature 493: 694–698.

Taliaferro, J. M., N. J. Lambert, P. H. Sudmant, D. Dominguez, J. J. Merkin *et al.*, 2016 RNA Sequence Context Effects Measured In Vitro Predict In Vivo Protein Binding and Regulation. Mol Cell 64: 294–306.

Tamburino, A. M., S. P. Ryder, and A. J. M. Walhout, 2013 A Compendium of Caenorhabditis elegans RNA Binding Proteins Predicts Extensive Regulation at Multiple Levels. G3 Genes Genomes Genetics 3: 297–304.

Teleman, A. A., and S. M. Cohen, 2006 Drosophila lacking microRNA miR-278 are defective in energy homeostasis. Gene Dev 20: 417–422.

Tomasso, G. D., L. M. M. Jenkins, and P. Legault, 2016 ARiBo pull-down for riboproteomic studies based on label-free quantitative mass spectrometry. Rna 22: 1760–1770.

Tong, X., N. Yu, R. Han, and T. Wang, 2020 Function of Dicer with regard to Energy Homeostasis Regulation, Structural Modification, and Cellular Distribution. Int J Endocrinol 2020: 1–7.

Tops, B. B. J., R. H. A. Plasterk, and R. F. Ketting, 2006 The Caenorhabditis elegans Argonautes ALG-1 and ALG-2: Almost Identical yet Different. Cold Spring Harb Sym 71: 189–194.

Tsukamoto, T., M. D. Gearhart, C. A. Spike, G. Huelgas-Morales, M. Mews *et al.*, 2017 LIN-41 and OMA Ribonucleoprotein Complexes Mediate a Translational Repression-to-Activation Switch Controlling Oocyte Meiotic Maturation and the Oocyte-to-Embryo Transition in *Caenorhabditis elegans*. Genetics 206: 2007–2039.

Tüfekci, K. U., M. G. Öner, R. L. J. Meuwissen, and Ş. Genç, 2013 miRNomics: MicroRNA Biology and Computational Analysis. Methods Mol Biology Clifton N J 1107: 33–50.
Valverde, R., L. Edwards, and L. Regan, 2008 Structure and function of KH domains. Febs J 275: 2712–2726.
Vannini, I., F. Fanini, and M. Fabbri, 2018 Emerging roles of microRNAs in cancer. Curr Opin Genet Dev 48: 128–133.
Valverde, R., L. Edwards, and L. Regan, 2008 Structure and function of KH domains. Febs J 275: 2712–2726.

Vasquez-Rifo, A., G. Jannot, J. Armisen, M. Labouesse, S. I. A. Bukhari *et al.*, 2012 Developmental Characterization of the MicroRNA-Specific C. elegans Argonautes alg-1 and alg-2. Plos One 7: e33750.

Vella, M. C., E.-Y. Choi, S.-Y. Lin, K. Reinert, and F. J. Slack, 2004 The *C. elegans* microRNA let-7 binds to imperfect let-7 complementary sites from the lin-41 3′UTR. Gene Dev 18: 132–137.

Wang, J., J. E. Lee, K. Riemondy, Y. Yu, S. M. Marquez *et al.*, 2020 XPO5 promotes primary miRNA processing independently of RanGTP. Nat Commun 11: 1845.

Wang, X., and E. Voronina, 2020 Diverse Roles of PUF Proteins in Germline Stem and Progenitor Cell Development in *C. elegans*. Frontiers Cell Dev Biology 8: 29

Wang, Y., M. Ma, X. Xiao, and Z. Wang, 2012 Intronic splicing enhancers, cognate splicing factors and context-dependent regulation rules. Nat Struct Mol Biol 19: 1044–1052.

Warf, M. B., B. A. Shepherd, W. E. Johnson, and B. L. Bass, 2012 Effects of ADARs on small RNA processing pathways in C. elegans. Genome Res 22: 1488–1498.

Westholm, J. O., and E. C. Lai, 2011 Mirtrons: microRNA biogenesis via splicing. Biochimie 93: 1897–1904.

Wilusz, J. E., S. M. Freier, and D. L. Spector, 2008 3′ End Processing of a Long Nuclear-Retained Noncoding RNA Yields a tRNA-like Cytoplasmic RNA. Cell 135: 919–932.

Wu, E., A. A. Vashisht, C. Chapat, M. N. Flamand, E. Cohen *et al.*, 2017 A continuum of mRNP complexes in embryonic microRNA-mediated silencing. Nucleic Acids Res 45: 2081–2098.

Wu, L., J. Fan, and J. G. Belasco, 2006 MicroRNAs direct rapid deadenylation of mRNA. P Natl Acad Sci Usa 103: 4034–4039.

Wu, H., S. Sun, K. Tu, Y. Gao, B. Xie *et al.*, 2010 A Splicing-Independent Function of SF2/ASF in MicroRNA Processing. Mol Cell 38: 67–77.

Xhemalce, B., S. C. Robson, and T. Kouzarides, 2012 Human RNA Methyltransferase BCDIN3D Regulates MicroRNA Processing. Cell 151: 278–288.

Xing, L., and G. J. Bassell, 2013 mRNA Localization: An Orchestration of Assembly, Traffic and Synthesis. Traffic 14: 2–14.

Xue, Y., K. Ouyang, J. Huang, Y. Zhou, H. Ouyang *et al.*, 2013 Direct Conversion of Fibroblasts to Neurons by Reprogramming PTB-Regulated MicroRNA Circuits. Cell 152: 82–96.

Yang, A., T.-J. Shao, X. B.-D. Ros, C. Lian, P. Villanueva *et al.*, 2020 AGO-bound mature miRNAs are oligouridylated by TUTs and subsequently degraded by DIS3L2. Nat Commun 11: 2765.

Yeom, K.-H., Y. Lee, J. Han, M. R. Suh, and V. N. Kim, 2006 Characterization of DGCR8/Pasha, the essential cofactor for Drosha in primary miRNA processing. Nucleic Acids Res 34: 4622–4629.

Yoda, M., T. Kawamata, Z. Paroo, X. Ye, S. Iwasaki *et al.*, 2010a ATP-dependent human RISC assembly pathways. Nat Struct Mol Biol 17: 17–23.

Yoda, M., T. Kawamata, Z. Paroo, X. Ye, S. Iwasaki *et al.*, 2010b ATP-dependent human RISC assembly pathways. Nat Struct Mol Biol 17: 17–23.

Young, L. E., A. E. Moore, L. Sokol, N. Meisner-Kober, and D. A. Dixon, 2012 The mRNA Stability Factor HuR Inhibits MicroRNA-16 Targeting of COX-2. Mol Cancer Res 10: 167–180.

Youngman, E. M., and J. M. Claycomb, 2014 From early lessons to new frontiers: the worm as a treasure trove of small RNA biology. Frontiers Genetics 5: 416.

Yu, B., Z. Yang, J. Li, S. Minakhina, M. Yang *et al.*, 2005 Methylation as a Crucial Step in Plant microRNA Biogenesis. Science 307: 932–935.

Zeng, Y., and B. R. Cullen, 2005 Efficient Processing of Primary microRNA Hairpins by Drosha Requires Flanking Nonstructured RNA Sequences*. J Biol Chem 280: 27595–27603.

Zabinsky, R. A., B. M. Weum, M. Cui, and M. Han, 2017 RNA Binding Protein Vigilin Collaborates with miRNAs To Regulate Gene Expression for *Caenorhabditis elegans* Larval Development. G3 Genes Genomes Genetics 7: 2511–2518.

Zhang, J., and Q. M. Chen, 2013 Far upstream element binding protein 1: a commander of transcription, translation and beyond. Oncogene 32: 2907–2916.

Zhang, L., L. Ding, T. H. Cheung, M.-Q. Dong, J. Chen *et al.*, 2007 Systematic Identification of *C. elegans* miRISC Proteins, miRNAs, and mRNA Targets by Their Interactions with GW182 Proteins AIN-1 and AIN-2. Mol Cell 28: 598–613.

Zhang, L., J. D. Ward, Z. Cheng, and A. F. Dernburg, 2015 The auxin-inducible degradation (AID) system enables versatile conditional protein depletion in C. elegans. Development 142: 4374–4384.

Zhang, T., Y.-C. Wu, P. Mullane, Y. J. Ji, H. Liu *et al.*, 2018 FUS Regulates Activity of MicroRNA-Mediated Gene Silencing. Mol Cell 69: 787-801.e8.

Zinovyeva, A. Y., S. Bouasker, M. J. Simard, C. M. Hammell, and V. Ambros, 2014 Mutations in Conserved Residues of the *C. elegans* microRNA Argonaute ALG-1 Identify Separable Functions in ALG-1 miRISC Loading and Target Repression. Plos Genet 10: e1004286.

Zinovyeva, A. Y., I. Veksler-Lublinsky, A. A. Vashisht, J. A. Wohlschlegel, and V. R. Ambros, 2015 *Caenorhabditis elegans* ALG-1 antimorphic mutations uncover functions for Argonaute in microRNA guide strand selection and passenger strand disposal. Proc National Acad Sci 112: E5271–E5280.

Zinshteyn, B., and K. Nishikura, 2009 Adenosine-to-inosine RNA editing. Wiley Interdiscip Rev Syst Biology Medicine 1: 202–209.

Zisoulis, D. G., Z. S. Kai, R. K. Chang, and A. E. Pasquinelli, 2012 Autoregulation of microRNA biogenesis by let-7 and Argonaute. Nature 486: 541–544.

Appendix A - DCR-1 Auxin inducible degron

Introduction

As one of the core biogenesis enzymes, Dicer is critical to the processing and biogenesis of miRNAs (Ryback-Wolf *et al.* 2014, reviewed in Bartel 2018). Specifically, Dicer binds precursor miRNAs (pre-miRNAs) generated from Drosha processing and exported into the cytoplasm. After binding the stemloop structure of pre-miRNAs, Dicer cleaves off the loop leaving the ~ 22 nt dsRNA miRNA duplex with the characteristic 3' 2nt overhang (reviewed in Ciechanowska *et al.* 2021). Consistent with its critical role in miRNA biogenesis, Dicer activity is thought to be carefully modulated to ensure the coordinated production of miRNAs duplexes. Although this is still an area of active interests, several mechanisms by which Dicer activity is modulated have been studied. At the level of translation, RBPs like AUF1 bind *dcr-1* mRNA and repress its translation (Abdelmohsen *et al.* 2012). Furthermore, Dicer protein can also be regulated through post-translational modification such as phosphorylation (Drake *et al.* 2014). Outside of factors that regulate its expression or structure, several RBPs appear to modulate the interaction between Dicer and its substrates. For example, the RBP DHX36 modulates the interaction between Dicer and pre-miR-134 in neurons, thus modulating the levels of pre-miRNA's processing by Dicer (Bicker *et al.* 2013).

Despite a decent understanding of Dicer processing and how Dicer activity can be regulated, we still lack a comprehensive overview of the more intricate details of pre-miRNA dynamics, particularly an in depth understanding of precursor processing rates and stability. This gap in knowledge is not entirely surprising given that perturbations of Dicer generally lead to sterility and reduced lifespan (Pavelic *et al.* 2009; Mori *et al.* 2012), with both phenotypes thought to be the result of the general depletion of the miRNA repertoire due disrupted pre-

miRNA processing (Duchaine *et al.* 2006). The sterility and lethality phenotypes caused by Dicer loss make it challenging to assess pre-miRNA processing rates and stability in *C. elegans*.

To overcome these issues, I designed and utilized an auxin-inducible degron (AID) system in order to conditionally deplete Dicer (DCR-1) and study the effects of its loss on miRNA dynamics. In general, the conditional depletion of Dicer allows us to bypass the detrimental effects of a null mutation, while also allowing for the precise control of timing and dosage of protein knockdown. Overall, the construction of the DCR-1::AID strain provides a powerful tool to better understand the processing rates of pre-miRNAs, as well as the overall stability of these molecules in the absence of Dicer processing. Lastly, the auxin-inducible degron system provides a novel tool to study the effects of DCR-1 loss on other RNA substrates, potentially identifying additional species of RNA that are dependent on Dicer for processing.

Methods

To build the DCR-1 degron I utilized CRISPR/Cas9 to edit the C-terminal tail of the protein at the endogenous loci. The following mix was prepared and injected: Cas9 0.5 uL (3.78ng/ul stock) (IDT), *dpy-10* crRNA 1.3 uL (40ng/ul stock) (IDT), *dcr-1* guide 5 crRNA 1.6 uL (100ng/ul dilution) (IDT), annealing buffer 1 ul (IDT), *dcr-1* C-terminal degron donor 3.2 ul (480 ng/uL) (IDT). The donor template sequence, comprised of the 5' homology arm, linker, degron, TEV, 3xFLAG and 3' homology arm is shown in shown in Table A.1. The mix was incubated at 37 C for 5 minutes and spun down at 10,000 rpm for 2 minutes. Young adult animals were injected and the F1 generation screened for dpy/roller animals, indicating edits had occurred. F1 animals were then isolated and self-propagated, after which F2 animals were screened for successful insertions using the listed screening primers (Table A.1). Two

independent edits, UY210(*zen78*) and UY212(*zen79*) were isolated and homozygozed before being confirmed by Sanger sequencing. Moving forward, *zen79* was then crossed into the p*eft*-1::TIR-1::mRuby::*unc-54* 3'UTR background to generate the complete degron strain *dcr-1*::AID(*zen79*); *eft*-1::TIR-1::mRuby::*unc-54* 3'UTR (UY237), which expresses TIR-1 in somatic tissues only. To conditionally knockdown DCR-1, we exposed *dcr-1*::AID(*zen79*); *eft*-1::TIR-1::mRuby::*unc-54* 3'UTR animals to Auxin in an identical manner to the DRSH-1::AID animals (See chapter 3). The depletion of DCR-1 in auxin exposed (and matched control animals) was quantified using western blot, whereas DCR-1::AID levels were estimated by probing for the 3xFLAG embedded within the AID tag. Based on quantification, 4-hour auxin exposure resulted in the highest level of knockdown and thus was used for RNA preparations and subsequent sequencing. RNA preparation and sequencing was performed as previously described (see Chapter 3).

Table A.1. Donor, crRNA, and primer sequences for creation of DCR-1::AID

DCR-1 C-terminal donor sequence	TTGATGGAAAATGCTCAAAGCTCCATAAATTAAC ACGTTTTGCAGATACCTCCATCAAATTGAGCAGC AAAGAAGACAAAGCCCATCATTAACAACTGTTG GATCCGGAGGTGGCGGGCCTAAAGATCCAGCCA AACCTCCGGCCAAGGCACAAGTTGTGGGATGGCC ACCGGTGAGATCATACCGGAAGAACGTGATGGTT TCCTGCCAAAAATCAAGCGGTGGCCCGGAGGCG GCGGCGTTCGTGAAGGAGAATCTGTACTTTCAAT CCGGAAAGGACTACAAAGACCATGACGGTGATT ATAAAGATCATGATATCGATTACAAGGATGACGA TGACAAGGGATCCTAAATTACATCTTCACTTTCT GTGATATGCTAAGTATTAAGCTATGTGTTTCTAG GATCTATTGATCTGATTTTCCTAATTCTCCAATTT TTACTCGTTT
dcr-1 C-terminal guide crRNA	5'-GCATATCACAGAAAGTGAAGG-3'
forward screening primer	5'-CTGGAATGGGAAGGAACTATC-3'
reverse screening primer	5'-GTGTGCAATTCATAAATGGCG-3'

Results/Conclusions

Quantification of DCR-1::AID knockdown by western blot revealed significant depletion of the protein in 1 hour, 2 hour, and 4 hour exposure to auxin, suggesting that we may have disrupted pre-miRNA processing (Figure A.1). While DCR-1 was not completely depleted, the 4-hour exposure to Auxin induced the greatest levels of knockdown (99.4%) of DCR-1 as assayed by western blot (Figure A.1B, C). We wondered whether blocked DCR-1 processing at the level of the precursor could result in primary miRNA accumulation. However, we did not see accumulation of pri-miRNAs, consistent with rapid pri-miRNA processing to precursor miRNAs.

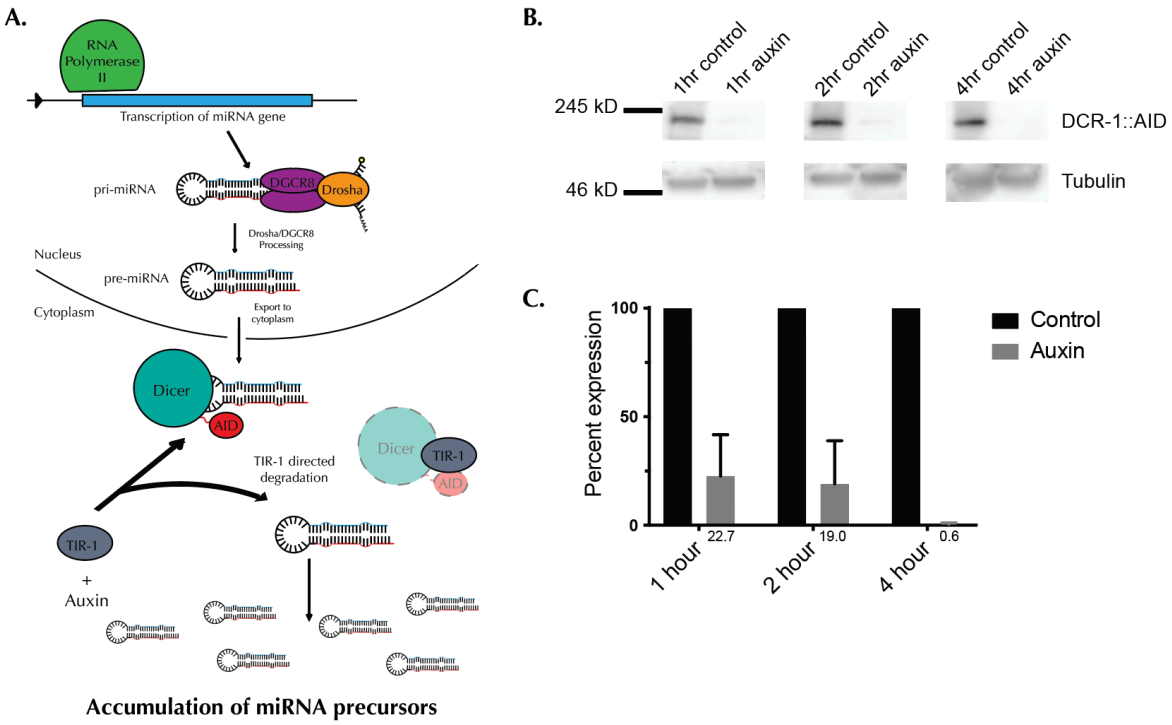


Figure A.1. DCR-1 levels are depleted after exposure to auxin. (A) Schematic of auxin-induced knockdown of DCR-1. **(B)** Representative western blot showing depletion of DCR-1 levels after 1hour, 2hour, and 4hour exposure to auxin. Time matched controls are shown for comparison. **(C)** Average DCR-1 expression for 3 replicates is quantified. Quantifications were done using ImageJ tool.

In all, we created a useful tool for future studies of Dicer and its role in the biogenesis of miRNAs. The AID system allows us to conditionally deplete of DCR-1, facilitating future experiments that will focus on the characterizing processing rates of pre-miRNAs. In addition, the conditional disruption of DCR-1 processing will allow us to uncover the relative stability of individual miRNA precursors, and possible helping to uncover additional factors that influence pre-miRNA stability. Overall, these types of analyses could add to our understanding of Dicer's role not only in miRNA biogenesis but a role in the processing of other species of RNA.

Appendix B - Characterization of KH domain RBP, IMPH-1 and generation of KH RBP::HALO reagents

Introduction

In recent years a great deal of effort has been made to understand how RBPs coordinate with miRNAs to modulate gene expression. A targeted screen of the KH domain family of RBPs found that many of this family of protein functionally interact with miRNAs to modulate gene expression in *C. elegans* development (See Chapter 2). Among the KH domain genes found to functionally interact with miRNA pathways, *imph-1* was chosen for further characterization. This choice was based in part on the fact that IMPH-1 interacts with miRISC components as shown by the IP of ALG-1 (Zinoveyva *et al*. 2015) and IP of DCR-1 (Duchaine *et al*. 2006). RNAi of *imph-1* led to a mild (but not statistically significant) repression in multiple miRNA sensitized background including *mir-48 mir-241(nDf51), let-7(n2853), and alg-1(ma202)* mutants (Haskell and Zinovyeva 2021). Taken together these data suggests that *imph-1* may modulate multiple miRNA families across several different developmental stages, perhaps by modulating miRISC function. To better understand the role of *imph-1* in the miRNA pathway, we set out to genetically characterize this gene and uncover interactions with miRNAs, miRNA intermediates, and other proteins. In addition, we also sought to generate reagents for the interrogation of four other KH domain RBPs, *E02D9.1*, *pes-4*, *sfa-1*, and *mask-1*, which show genetic interactions with multiple miRNA families (Haskell and Zinovyeva 2021).

Methods

To initially characterize the role of *imph-1* in the miRNA pathway I first generated a full gene deletion using CRISPR/Cas9. To accomplish this, I targeted the endogenous *imph-1* loci (which contains four isoforms); using two guide crRNAs to cut immediately upstream and downstream of the start and stop codon of *imph-1* isoform C (longest isoform). The following injection mix was prepared and injected as previously described in Chapter 3 and Appendix A: Cas9 (IDT): 0.5uL, tracRNA (IDT): 5uL (0.4 ug/uL stock), *dpy-10* crRNA (IDT): 1.4 uL (0.4 ug/uL stock), *imph-1* N-terminal crRNA (IDT): 1.8 uL (0.4 ug/uL stock), *imph-1* C-terminal crRNA (IDT): 1.8 uL (0.4 ug/uL stock), annealing buffer (IDT): 2 uL, H20: 7.1 uL. The co-injection marker *dpy-10* was used to screen for successful edits. Screening for edits was done using the *imph-1* forward screening primer (5' CAGGTCTCGAAGAAGA CATGAC-3') and *imph-1* reverse screening primer (5'- GCCCCCAACCAGAATTTGAGA-3') which are placed immediately upstream and downstream of the *imph-1* locus. Four full gene deletions were identified by PCR screening, homozygozed and confirmed by Sanger sequencing. To assess the impact of the loss of *imph-1,* I crossed UY235*(zen89*), representing largest *imph-1* deletion, into several miRNA sensitized backgrounds, including the *mir-48 mir-241 (nDf51)* and *alg-1 (ma202)* backgrounds and scored for *col-19::gfp* expression in hypodermal cells. Two additional strains were produced where *imph-1(zen89)* was crossed into the *mir-35 mir-36 mir-37 mir-38 mir-39 mir-40 mir-41(nDf50)* and *lsy-6(ot150)* miRNA sensitized backgrounds. Although these strains were completed and confirmed, they were not scored for their corresponding phenotype.

While deletion alleles provide information about the effect of complete loss of a gene, I also sought to generate mutates that may help us to better understand *imph-1* on a more mechanistic level. To do this, I targeted the KH domains, specifically the GXXG RNA binding

motifs located in the last two KH domains (KH3 and KH4) of IMPH- 1. To generate the mutants, I utilized CRISPR/Cas9 to first target the KH3 domain, modifying the GXXG motif to GDDG, and thus eliminating its RNA binding ability. Young adult worms were injected with the following mix: Cas9 (IDT): 0.5uL, tracRNA (IDT): 5uL (0.4 ug/uL stock), *dpy-10* crRNA (IDT): 1.4 uL (0.4 ug/uL stock), *imph-1* KH3 guide crRNA (IDT): 1.8 uL (0.4 ug/uL stock), *imph-1* KH3 GDDG donor: 2.2uL (IDT) (500 ng/uL), annealing buffer (IDT): 2 uL, H20: 9.1 uL. The co-injection marker *dpy-10* was used to screen for successful edits. Screening for edits was done using the KH3 forward and KH3 reverse primers, placed immediately upstream and downstream of the edit (Table B.1.). Dpy/roller animals were singled, and the edited region amplified by PCR. To screen for edits I utilized the enzyme *bccI*, which will only digest the edited loci due to the introduction of a *bccI* site. Two alleles*, zen182* and *zen183* were confirmed by Sanger sequencing to have the expected edit. To generate the mutant carrying both KH3 and KH4 GDDG mutations, I injected *zen182* animals with the following mix: 0.5uL, tracRNA (IDT): 5uL (0.4 ug/uL stock), dpy-10 crRNA (IDT): 1.4 uL (0.4 ug/uL stock), *imph-1* KH4 guide crRNA (IDT): 1.8 uL (0.4 ug/uL stock), *imph-1* KH4 GDDG donor: 2.2uL (IDT) (500 ng/uL), annealing buffer (IDT): 2 uL, H20: 9.1 uL. The co-injection marker *dpy-10* was used to screen for successful edits. Screening for edits was done using the KH4 forward.1 and KH4 reverse.1 primers, placed immediately upstream and downstream of the edit (Table B.1). Dpy/roller animals were singled, and the edited region amplified by PCR in the F2 progeny. To screen for edits I utilized the enzyme *hpy811* which will only cut wildtype loci, due to the removal of the *hpy811* cut site from the repair template. One successful edit was identified and Sanger sequencing of both KH3 and KH4 GDDG motifs was completed to confirm the double mutant UY435 (*zen182zen183*).

Table B.1. IMPH-1 KH3 and KH4 crRNA, donor, and primer sequences.

imph-1 KH3 GDDG guide	5'-GGAGCACTCATCGGCGCAAAGGG-3'
imph-1 KH3 GDDG donor	GGGTTCCTGATAGTATGATTGGAGCTCTCATTGGCACGATGGAAAGAATATCAAGATGATAATTCGTGATACTGG
KH 3 forward.1 primer	5'-CCACCAATGATGATGCCAC-3'
KH 3 reverse.1 primer	5'-GGAGGATAATGCAACAATCAACTC-3'
imph-1 KH4 GDDG guide	5'-GGTAGAATAATTGGAAAAGGTGG-3'
imph-1 KH4 GDDG donor	5'-CCGTTCCAACTAGAATCATTGGTAGAATCATCGGAGACGATGGACAGAATGTTCGTGAGCTACAGAGAATTACGGG-3'
KH 4 forward.1 primer	5'-CATCTGATGCTATTGAGGAGAAGC-3'
KH 4 reverse.1 primer	5'-CCGGAGATCATCGCAATAAGTC-3'

To confirm the interaction between IMPH-1 and ALG-1, as well as interrogate the interaction between IMPH-1 and other proteins/RNAs I also generated an IMPH-1::HALO allele. To generate this mutant, I utilized CRISPR/Cas9 in order to insert the HALO tag at the endogenous C-terminus of IMPH-1, which is shared between all 4 isoforms. Young adult animals were injected with the following mix: Cas9 (IDT): 0.5uL, tracRNA (IDT): 5uL (0.4 ug/uL stock), *dpy-10* crRNA (IDT): 1.4 uL (0.4 ug/uL stock), *imph-1* C-terminal guide crRNA (IDT): 1.8 uL (0.4 ug/uL stock), *imph-1* C-terminal HALO donor: 2.2uL (IDT) (500 ng/uL), annealing buffer (IDT): 2 uL, H20: 9.1 uL. The sequence of the HALO donor, guide crRNA, and screening primers are shown in Table B.2, B.3, and B.4, respectively. Dpy/roller were singled from the F1 generation. F2 animals were screening for the insertion using *imph-1* forward and reverse screening primers that sit upstream and downstream of the insertion site. Several possible insertions were isolated. The resulting strains were sequenced, with one allele (*zen105*) being confirmed by sequencing. The full length of *imph-1* was then sequenced in UY272 (*zen105*) to ensure there were no mutations introduced into the locus.

In addition to IMPH-1::HALO, four additional HALO tags were designed to tag EO2D9.1, PES-4, SFA-1, and MASK-1 for purification and subsequent identification of their RNA targets. HALO tags were designed for the N-terminus of E02D9.1 and PES-4, and for the C-terminus of SFA-1 and MASK-1. Donor sequences for the HALOtag are shown in Table B.2, guide crRNA sequences in Table B.3, and primer sequences in Table B.4. In addition to IMPH-1::HALO we also attempted to generate a MASK-1::HALO and SFA-1::HALO allele by CRISPR/Cas9. For MASK-1::HALO young adult animals were injected with the following mix: Cas9 (IDT): 0.5uL, tracRNA (IDT): 5uL (0.4 ug/uL stock), *dpy-10* crRNA (IDT): 1.4 uL (0.4 ug/uL stock), *mask-1* C-terminal guide crRNA (IDT): 1.8 uL (0.4 ug/uL stock), *mask-1* C-terminal HALO donor: 2.2uL (IDT) (500 ng/uL), annealing buffer (IDT): 2 uL, H20: 9.1 uL. For SFA-1::HALO young adult animals were injected with the following mix: Cas9 (IDT): 0.5uL, tracRNA (IDT): 5uL (0.4 ug/uL stock), *dpy-10* crRNA (IDT): 1.4 uL (0.4 ug/uL stock), *sfa-1* C-terminal guide crRNA (IDT): 1.8 uL (0.4 ug/uL stock), *sfa-1* C-terminal HALO donor: 2.2uL (IDT) (500 ng/uL), annealing buffer (IDT): 2 uL, H20: 9.1 uL. Dpy/roller were singled from the F1 generation. F2 animals were screening for the insertion using the appropriate forward and reverse screening primers that sit upstream and downstream of the insertion site. Two alleles were isolated for MASK-1::HALO, *zen268* and *zen269*. Both were confirmed by Sanger sequencing. In addition, the full length of *mask-1* was sequenced in UY269 (*zen104*) to ensure no off-target mutations had occurred. A single allele of SFA-1::HALO was isolated, *zen106*, but has not been confirmed by sequencing. All strains generated for this project are shown in Table B.5.

***T*able B.2. KH domain RBP HALO tag donor sequences.**

IMPH-1 C-terminal HALO donor	TTCAGCCCCATCATCAGGTCAAGAAAAAGATGGTAAGCTACTGAAA GGACTTGACACTTTTTTTTTTACCGAGTGAACAGTTTTCAGGTTCCG CACTCGAAAAAATGGACCAGCTGGGCACAATTGCACCAATTAGCA ATTCGAATCGGGCTTCTCCAAAATCTGTTTCTCCTCCAAAGTCTAAG TCGCCAGGATCCGGAGGTGGCGGGAAGGACTACAAAGACCATGAC GGTGATTATAAAGATCATGATATCGATTACAAGGATGACGATGACA AGGGATCCGAGAATCTGTACTTTCAATCCGGAATGGCTGAAATTGG TACTGGCTTTCCTTTTGATCCCCACTACGTTGAGGTTCTTGGAGAGC GCATGCACTACGTTGATGTTGGACCACGCGATGGAACCCCAGTCCT TTTCCTTCACGGAAACCCAACTTCTTCTTACGTTTGGCGCAACATCA TCCCACATGTTGCTCCAACCCACCGCTGCATCGCTCCAGATCTTATC GGAATGGG:AAAGTCTGATAAGCCAGATCTTGGATACTTCTTCGATG ATCACGTCCGTTTCATGGATGCTTTCATCGAGGCCCTTGGACTTGAG GAGGTTGTCCTTGTTATCCACGATTGGGGATCCGCTCTTGGATTCCA CTGGGCTAAGCGCAACCCAGAGCGCGTT:AAGGGAATCGCTTTCAT GGAATTCATCCGCCCAATCCCAACTTGGGATGAATGGCCAGAGTTC GCTCGCGAGACCTTCCAAGCTTTCCGCACCACCGATGTTGGACGTA AGCTTATCATCGATC:AAAACGTTTTCATCGAAGGAACCCTTCCAAT GGGAGTTGTTCGTCCAC:TTACCGAGGTTGAGATGG:ATCACTACCG CGAGCCATTCCTT:AACCCAGTTGATCGCGAGCCACTTTGGCGCTTC CCAAACGAACTTCCAATCGCTGGAGAGCCAGCTAACATCGTTGCTC TTGTTGAGGAGTACATGGATTGGCTTCACCAATCTCCAGTTCCAAA GCTTCTTTTCTGGGGAACTCCAGGAGTTCTTATCCCACCAGCTGAGG CTGCCCGTCTTGCTAAGTCTCTTCCAAACTGCAAGGCTGTCGATATC GGACCAGGACTTAACCTTCTTCAAGAGGATAACCCAGATCTTATCG GATCTGAGATCGCTCGTTGGCTTTCTACCCTTGAGATCTCCGGAGGT TAGAATATAGTTCTACGGTTTCCAGTTTTTTGTTGAAATTTTCCATA TTATTCTTTTGTTTTTTTTTCTTTTTCAATTTCTTCCCTTTGGTTTTA TGTCCCGCACCATCAAAACAAGCATTTTCCCGTTGTCTTTTCCCCGT CCCGCGTTGAGCAATTTTTTGAGAATTTAAGTTTAGCACATATTTCG CAACAAATTTACATTCATTTCTTTACGAGCATATCCCGAACCAATAC ATTTTCTGAAGAATACCCCCCTCCATCCCCAATGTTATTCCAATTGT TTTTTAGGAATATTTT
E02D9.1 N-terminal HALO donor	GTACCTTTTATAAGAATAAGGTCCTGCCGTGGCTGAGATAGCGCTG GTACACGGAGCCTACCTAAAATGGATACGTAATACTATCGATTCAC AAAAAAACTGTATATGATTAAAAATCTTAATAATTAATGTTTCAAA GTACTCCTTCGTAGCTTGTCTGTTAATTTGTTTGTTTCTGCTGCAATC TTTTCAACTGTCTAACGCCTACTATTCCAGAGCTTCTCTCTTTGATTT CTTGACAACTCACAGTCATATTTATGGCTGAAATTGGTACTGGCTTT CCTTTTGATCCCCACTACGTTGAGGTTCTTGGAGAGCGCATGCACTA CGTTGATGTTGGACCACGCGATGGAACCCCAGTCCTTTTCCTTCACG GAAACCCAACTTCTTCTTACGTTTGGCGCAACATCATCCCACATGTT GCTCCAACCCACCGCTGCATCGCTCCAGATCTTATCGGAATGGGAA AGTCTGATAAGCCAGATCTTGGATACTTCTTCGATGATCACGTCCGT TTCATGGATGCTTTCATCGAGGCCCTTGGACTTGAGGAGGTTGTCCT

	TGTTATCCACGATTGGGGATCCGCTCTTGGATTCCACTGGGCTAAG CGCAACCCAGAGCGCGTTAAGGGAATCGCTTTCATGGAATTCATCC GCCCAATCCCAACTTGGGATGAATGGCCAGAGTTCGCTCGCGAGAC CTTCCAAGCTTTCCGCACCACCGATGTTGGACGTAAGCTTATCATCG ATCAAAACGTTTTCATCGAAGGAACCCTTCCAATGGGAGTTGTTCG TCCACTTACCGAGGTTGAGATGGATCACTACCGCGAGCCATTCCTT: AACCCAGTTGATCGCGAGCCACTTTGGCGCTTCCCAAACGAACTTC CAATCGCTGGAGAGCCAGCTAACATCGTTGCTCTTGTTGAGGAGTA CATGGATTGGCTTCACCAATCTCCAGTTCCAAAGCTTCTTTTCTGGG GAACTCCAGGAGTTCTTATCCCACCAGCTGAGGCTGCCCGTCTTGC TAAGTCTCTTCCAAACTGCAAGGCTGTCGATATCGGACCAGGACTT AACCTTCTTCAAGAGGATAACCCAGATCTTATCGGATCTGAGATCG CTCGTTGGCTTTCTACCCTTGAGATCTCCGGAGGTGAGAATCTGTAC TTTCAATCCGGAAAGGACTACAAAGACCATGACGGTGATTATAAAG ATCATGATATCGATTACAAGGATGACGATGACAAGGGATCCGGATC CGGAGGTGGCGGGGACCCATATCAACAAGGAGGCCGTGGCGGTGG ATTCCCAGCACGTGGAGGTCGTGGCGGAGGTCATGGTGGAGGATAT CCACAAGAAGGTTATGGTGCCGCTGCTGGTGGCTACGGTGGATACG ATCCATACAATCCATATGGAGCCGCTGGTGGATATGGAATGTATCC AGGTCAAGGATACCCACCACAAGAAATGACTTCACCTTTGGATGCC GAGATTCAGGCAGTTTTACGAGAAATTCATTTGGAAGTAACTGGAC TGGAAACCTCGGGTGATCAGTTTAGAAATGCCCGA
PES-4 N-terminal HALO donor	TTTTAGTTGACAATTTTCAATTTAAACTTGGGGTTGTTGAGATAGCA CCGATTAAAAGTTAAGGTTAGCAAGTTAAGAAAAAGACGGCTGGA AGGGGGTGTGAGATGGGCGGTCAGTGTGTCCCATGTCATTTTTCGA TAAGATAAGCAGAGCAGTTTGCCGTTCAACTACACCACATTAACGT CGTTTCTTTCTCTCGCCTTCTCTCCCACTTTATTACCTAAATTATTAC AATTCAAACGCAGTTTTATAGTTTCTGGAGCCTTTTCTGGATATTTT ATATTATCGATTTTTTTTTTGCAGATGGCTGAAATTGGTACTGGCTT TCCTTTTGATCCCCACTACGTTGAGGTTCTTGGAGAGCGCATGCACT ACGTTGATGTTGGACCACGCGATGGAACCCCAGTCCTTTTCCTTCAC GGAAACCCAACTTCTTCTTACGTTTGGCGCAACATCATCCCACATGT TGCTCCAACCCACCGCTGCATCGCTCCAGATCTTATCGGAATGGGA AAGTCTGATAAGCCAGATCTTGGATACTTCTTCGATGATCACGTCC GTTTCATGGATGCTTTCATCGAGGCCCTTGGACTTGAGGAGGTTGTC CTTGTTATCCACGATTGGGGATCCGCTCTTGGATTCCACTGGGCTAA GCGCAACCCAGAGCGCGTTAAGGGAATCGCTTTCATGGAATTCATC CGCCCAATCCCAACTTGGGATGAATGGCCAGAGTTCGCTCGCGAGA CCTTCCAAGCTTTCCGCACCACCGATGTTGGACGTAAGCTTATCATC GATCAAAACGTTTTCATCGAAGGAACCCTTCCAATGGGAGTTGTTC GTCCACTTACCGAGGTTGAGATGGATCACTACCGCGAGCCATTCCT TAACCCAGTTGATCGCGAGCCACTTTGGCGCTTCCCAAACGAACTT CCAATCGCTGGAGAGCCAGCTAACATCGTTGCTCTTGTTGAGGAGT ACATGGATTGGCTTCACCAATCTCCAGTTCCAAAGCTTCTTTTCTGG GGAACTCCAGGAGTTCTTATCCCACCAGCTGAGGCTGCCCGTCTTG CTAAGTCTCTTCCAAACTGCAAGGCTGTCGATATCGGACCAGGACT

	TAACCTTCTTCAAGAGGATAACCCAGATCTTATCGGATCTGAGATC GCTCGTTGGCTTTCTACCCTTGAGATCTCCGGAGGTGAGAATCTGTA CTTTCAATCCGGAAAGGACTACAAAGACCATGACGGTGATTATAAA GATCATGATATCGATTACAAGGATGACGATGACAAGGGATCCGGA TCCGGAGGTGGCGGGGACCCATCCATGTACCCGGGCGCATCTGCAA CATACTTATCATATCCACAACAGAATGCTAGTCAAGACTACCTGGG AGCATACGCAGCACAGACAACTACAACGACAAGGAATACTGCAAA TGGGTAGGGGTTTTTTGGATTTTTTGCATTCAAATTGTAGCCCTTGA AATTCTACACATTTTGTTAGTTGGAAAGTTTTTGATAGCTCTATTCC CCGCTGAGTTATACGGCGTCAAAGATTCAGAGTGAGTGAGTGTGCG GCGAGGGGAGACGCAGGAGGCGATGGCGTCTGCGT
SFA-1 C-terminal HALO donor	CTATGCCGATGCCTGTACCTCCACCAGGCGGCCTAGGCGGTTTCAT GCCTCCACCACCTCCACCACCACCAATGCCTGGAGATTTGTCGTCA CTTTTGGCCGCTGCTCCACCGCCACCGCCTAGTGGATCCGGAGGTG GCGGGAAGGACTACAAAGACCATGACGGTGATTATAAAGATCATG ATATCGATTACAAGGATGACGATGACAAGGGATCCGAGAATCTGT ACTTTCAATCCGGAATGGCTGAAATTGGTACTGGCTTTCCTTTTGAT CCCCACTACGTTGAGGTTCTTGGAGAGCGCATGCACTACGTTGATG TTGGACCACGCGATGGAACCCCAGTCCTTTTCCTTCACGGAAACCC AACTTCTTCTTACGTTTGGCGCAACATCATCCCACATGTTGCTCCAA CCCACCGCTGCATCGCTCCAGATCTTATCGGAATGGGAAAGTCTGA TAAGCCAGATCTTGGATACTTCTTCGATGATCACGTCCGTTTCATGG ATGCTTTCATCGAGGCCCTTGGACTTGAGGAGGTTGTCCTTGTTATC CACGATTGGGGATCCGCTCTTGGATTCCACTGGGCTAAGCGCAACC CAGAGCGCGTTAAGGGAATCGCTTTCATGGAATTCATCCGCCCAAT CCCAACTTGGGATGAATGGCCAGAGTTCGCTCGCGAGACCTTCCAA GCTTTCCGCACCACCGATGTTGGACGTAAGCTTATCATCGATCAAA ACGTTTTCATCGAAGGAACCCTTCCAATGGGAGTTGTTCGTCCACTT ACCGAGGTTGAGATGGATCACTACCGCGAGCCATTCCTTAACCCAG TTGATCGCGAGCCACTTTGGCGCTTCCCAAACGAACTTCCAATCGC TGGAGAGCCAGCTAACATCGTTGCTCTTGTTGAGGAGTACATGGAT TGGCTTCACCAATCTCCAGTTCCAAAGCTTCTTTTCTGGGGAACTCC AGGAGTTCTTATCCCACCAGCTGAGGCTGCCCGTCTTGCTAAGTCTC TTCCAAACTGCAAGGCTGTCGATATCGGACCAGGACTTAACCTTCT TCAAGAGGATAACCCAGATCTTATCGGATCTGAGATCGCTCGTTGG CTTTCTACCCTTGAGATCTCCGGAGGTTAAATTTGATTTTCACAATT TTCCTAATATTCGATACTGTAATTCGTTTTTTTTTCCAGTTTTTAATG TTCTTTTAAAAGTGGAGTAGGAGTGAGGAAAAATCCACGCGATTTT TC
MASK-1 C-terminal HALO donor	GCAATACGGACAGAGTTCACAGCAACAGCCTTACGGTCAGATGCCT CAAGCGTAAGTTCTACGATTTCTAATTAAATATTTGTTTATTTATTA CAGAATGGATTGGAACCGACTTGGACAGCAACAGCAGTCTGCATCT GGCCAACAAAATCATCAGTCATCCTCGTCAAATAAATGGTCTTCCA ACTGG:GGATCCGGAGGTGGCGGGAAGGACTACAAAGACCATGAC GGTGATTATAAAGATCATGATATCGATTACAAGGATGACGATGACA AGGGATCCGAGAATCTGTACTTTCAATCCGGAATGGCTGAAATTGG

	TACTGGCTTTCCTTTTGATCCCCACTACGTTGAGGTTCTTGGAGAGC GCATGCACTACGTTGATGTTGGACCACGCGATGGAACCCCAGTCCT TTTCCTTCACGGAAACCCAACTTCTTCTTACGTTTGGCGCAACATCA TCCCACATGTTGCTCCAACCCACCGCTGCATCGCTCCAGATCTTATC GGAATGGGAAAGTCTGATAAGCCAGATCTTGGATACTTCTTCGATG ATCACGTCCGTTTCATGGATGCTTTCATCGAGGCCCTTGGACTTGAG GAGGTTGTCCTTGTTATCCACGATT:GGGGATCCGCTCTTGGATTCC ACTGGGCTAAGCGCAACCCAGAGCGCGTTAAGGGAATCGCTTTCAT GGAATTCATCCGCCCAATCCCAACTTGGGATGAATGGCCAGAGTTC GCTCGCGAGACCTTCCAAGCTTTCCGCACCACCGATGTTGGACGTA AGCTTATCATCGATCAAAACGTTTTCATCGAAGGAACCCTTCCAAT GGGAGTTGTTCGTCCACTTACCGAGGTTGAGATGGATCACTACCGC GAGCCATTCCTTAACCCAGTTGATCGCGAGCCACTTTGGCGCTTCCC AAACGAACTTCCAATCGCTGGAGAGCCAGCTAACATCGTTGCTCTT GTTGAGGAGTACATGGATTGGCTTCACCAATCTCCAGTTCCAAAGC TTCTTTTCTGGGGAACTCCAGGAGTTCTTATCCCACCAGCTGAGGCT GCCCGTCTTGCTAAGTCTCTTCCAAACTGCAAGGCTGTCGATATCG GACCAGGACTTAACCTTCTTCAAGAGGATAACCCAGATCTTATCGG ATCTGAGATCGCTCGTTGGCTTTCTACCCTTGAGATCTCCGGAGGTT AGATATCGCTCACAAAATCGTCTTCTTGTCAATTAGTCCCCCCTCCC CTCATAATAATTCGGTGTTTACTTGTTTTTTATTCAACCGTGAAACA ATATGTCTTCCCGGTTGAATTGCATCCATTTTTCTTTATTTCGTCTCC TCGTTCCGGGTTTTGTAACACTTTTTCTTCATTTTAAATCTAGCCTTT TTTGGTCCATTTTTTCTTTCAACTCTCAAATGTTCGATCTTTCCCATC CCTAGCAAGTGTTTTCTTGCGAGCTCAGCCTTGTGGAAATTGTACTC TTCTGTCCAAATTA

Table B.3. IMPH-1 HALO tag donor, crRNA, and primer sequences.

IMPH-1 C-terminal guide	TATGGCGATTTGGATTTAGG
E02D9.1 N-terminal guide	TTGACAACTCACAGTCATAA
PES-4 N-terminal guide	TCTGGATATTTTATATTATCG
SFA-1 C-terminal guide	GAAAATCAAATTTAACTAGG
MASK-1 C-terminal guide	TGTGAGCGATATCTACCAGT

Table B.4. KH domain RBP HALO tag SP9 and screening primer sequences.

imph-1 SP9 forward primer	CAGCCCCATCATCAGGTCAAG
imph-1 SP9 reverse primer	CCCTCCATCCCCAATGTTATTCC
imph-1 forward screening primer	CAGGTCTCGAAGAAGACATGAC
imph-1 reverse screening primer	GCCCCCAACCAGAATTTGAGA
E02D9.1 SP9 forward primer	GAATAAGGTCCTGCCGTGGCTGAG
E02D9.1 SP9 reverse primer	GGACTGGAAACCTCGGGTGATCAG
E02D9.1 forward screening primer	GGCTACATTTTTCACAAATTTGGTC
E02D9.1 reverse screening primer	CAGTTAAGGTTTGCAAGAAGATTC

pes-4 SP9 reverse primer	GTTGGAAAGTTTTTGATAGCTC
pes-4 forward screening primer	GTCTCTCTCCTTCCACACACTC
pes-4 reverse screening primer	CTCAACCTCATCGGACCACC
sfa-1 SP9 forward primer	GATGCCTGTACCTCCACCAGG
sfa-1 SP9 reverse primer	GTAGGAGTGAGGAAAAATCCACGCG
sfa-1 forward screening primer	CACTATCAACAATACGCCTATCC
sfa-1 reverse screening primer	CCGATTTTTGTTCTCAAAAATGCTGC
mask-1 SSP9 forward primer	GGACAGAGTTCACAGCAACAGCC
mask-1 SP9 reverse primer	CTTGCGAGCTCAGCCTTGTGG
mask-1 forward screening primer	GCAAATCAATCGCGACTCG
mask-1 reverse screening primer	GTATTTGAGACGGATGTCAAATGG

Table B.5. Strains generated during characterization of *imph-1*.

Strain	Genotype
UY233	*imph-1 (zen87)*
UY234	*imph-1 (zen88)*
UY235	*imph-1 (zen89)*
UY236	*imph-1 (zen90)*
UY268	*mask-1::HALO(zen103)*
UY269	*mask-1::HALO(zen104)*
UY270	*imph-1 (zen89); col-19::gfp (maIs105)*
UY271	*imph-1 (zen89); mir-48 mir-241 (nDf51); col-19::gfp (maIs105)*
UY272	*imph-1::HALO (zen105)*
UY273	*sfa-1::HALO(zen106)*
UY433	*imph-1(zen182)*
UY434	*imph-1(zen183)*
UY435	*imph-1(zen183zen184)*
UY436	*imph-1 (zen89); lsy06(ot150)*
UY437	*imph-1 (zen89) ;nDf50*
UY439	*imph-1 (zen89); alg-1(ma202); lin-31(n1083); col-19::gfp (maIs105)*

Results/Discussion

To facilitate the study of *imph-1* I targeted the endogenous loci, containing 4 *imph-1* isoforms, for mutation (Figure B.1A). Using CRISPR Cas9 we generated 4 full gene deletions, *zen87- zen90* (corresponding to UY233-UY236) (Figure B1B, Table B.5). All 4 deletion alleles (*zen87- zen90*) are hypothesized to be complete nulls, as they delete the vast majority of the

coding sequence. The largest deletion (*zen89*) was used in future assays as it deleted the full coding region of the gene as well as the regions immediately flanking the start and stop codons.

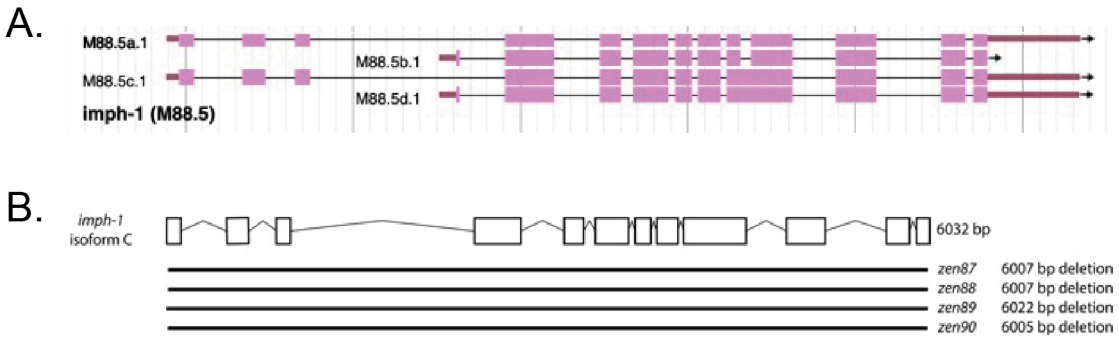


Figure B.1. Structure of *imph-1* locus and deletion alleles. (A) *imph-1* is encoded by 4 isoforms (A-D). The N-terminus of the isoforms are variable, although they share a common C-terminus. (B) *imph-1* alleles *zen87*, *zen88*, *zen89*, *zen90* represent full gene deletions. *imph-1(zen89)* is the most complete deletion and was used for subsequent genetic assays.

In addition to generating full gene deletions, I also wanted to generate several alleles targeting the GXXG RNA binding motifs within the KH domains of IMPH-1 (Figure B.2, Table B.5). To that end, I generated two alleles *zen182* (UY433) and *zen-182* (UY434) that carry GDDG mutations at the RNA binding motif of KH domain 3 of IMPH-1. I also generated an additional strain UY435(*zen183zen184*) that carries GDDG mutations at both RNA binding motif of both KH3 and KH4 domains of IMPH-1 (Table B.5).

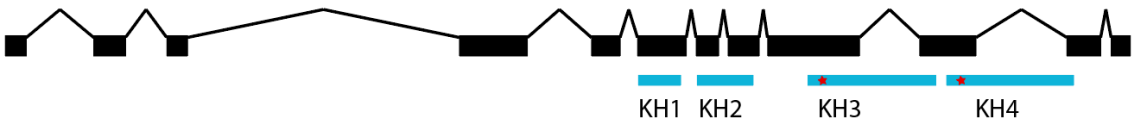


Figure B.2. IMPH-1 KH domain GDDG mutants. Schematic showing location of GDDG motif mutations in KH3 and KH4 domains.

Knockdown of *imph-1* was shown to mildly suppress abnormal hypodermal expression of adult specific marker *col-19::gfp* in both *mir-48 mir-241* and *alg-1* antimorphic background (Haskell and Zinovyeva 2021). The RNAi results were recapitulated in *imph-1(zen89);mir-48 mir-241(nDf51);col-19::gfp (maIs105)* animals which showed ~17% abnormal *col-19::gfp* when compared to *mir-48 mir-241(nDf51)* animals (Figure B.3A). While not statistically significant, this this suppression may still be biologically relevant. Loss of *imph-1* does not impact *col-19::gfp* expression in the absence of the *mir-48 mir-241(nDf51*) mutation. We also saw a recapitulation of RNAi data in *imph-1(zen89); lin-31(n1053)*, *col-19::gfp* (*maIs105)*; *alg-1(ma202)* animals, where ~8% animals show normal hypodermal *col-19::gfp* expression compared to the control *lin-31(n1053)*, *col-19::gfp* (*maIs105)*; *alg-1(ma202)* animals at 0% (Figure B.3B).

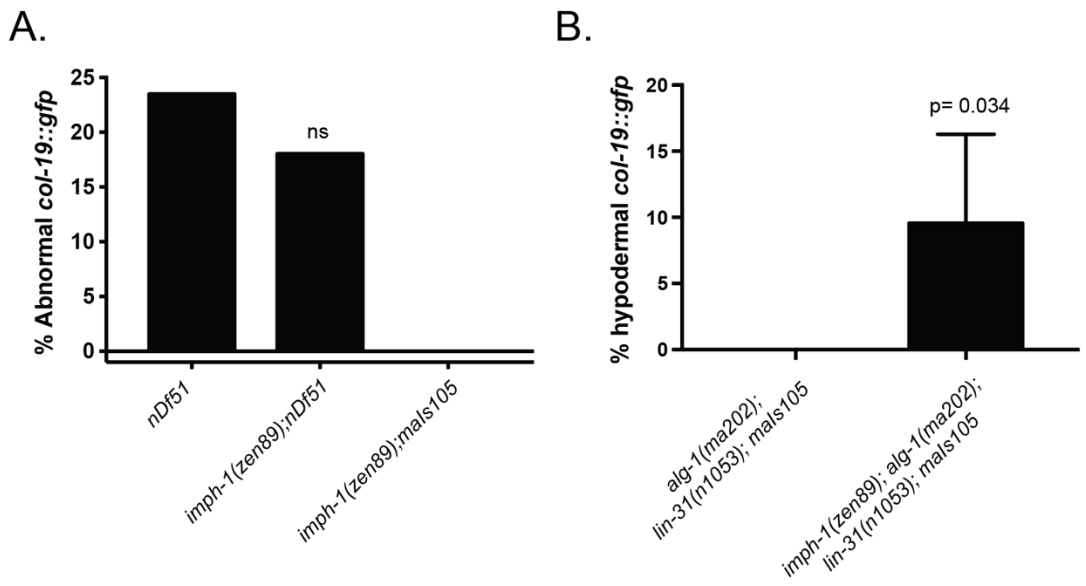


Figure B.3. The loss of imph-1 mildly suppresses *mir-48 mir-241 (nDf51)* and *alg-1(ma202)* phenotypes. (A) *imph-1*(*zen89*);*nDf51* animals show a mild suppression in the percentage of animals showing abnormal *col-19::gfp* when compared to the nDf51 strain alone. (B) A significant percentage of *imph-1(zen89);alg-1(ma202)* animals show hypodermal expression of *col-19::gfp* which never occurs in the *alg-1(ma202)* strain.

In order to confirm the expression of IMPH-1::HALO, I performed a western blot as previously described, and probed for FLAG in order to detect the 3xFLAG embedded within the HALO tag. When compared to wildtype, the IMPH-1::HALO showed a band at approximately 135 kDa which is consistent with the IMPH-1C (longest isoform) + the HALO tag (Figure B.4).

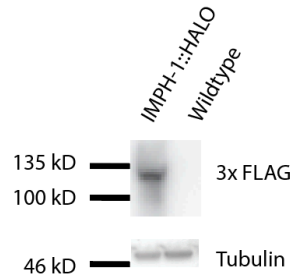


Figure B.4. IMPH-1::HALO expression and size. IMPH-1::HALO is expressed a workable levels and can be detected by anti-FLAG antibody. Size detected on western is consistent with IMPH-1C + HALOtag.

Overall, I have generated reagents to facilitate the future study of IMPH-1. As a KH domain containing RBP, IMPH-1 has been shown to not only physically interact with ALG-1 and DCR-1, but also to functionally interact with multiple miRNA families. Although the genetic suppression observed in both RNAi and genetic assays were subtle, there was an observable modulation of miRNA phenotypes, suggesting that *imph-1* may normally negatively regulate miRNA activity and function. Furthermore, the interaction between IMPH-1 and miRISC components suggests that this RBP may exert its effects on miRNA function indirectly by modulating miRISC activity of function. Future study will be needed to confirm a role, for IMPH-1 in miRNA-mediated regulation of gene expression; work that will facilitated by genetic and biochemical reagents produced in this work.