TEMPORAL PATTERNS OF SHORT NON-CODING RNA MODIFICATIONS AND EXPRESSION

by

AMMAR S. NAQVI

A dissertation submitted to the

Graduate School – Camden

Rutgers, The State University of New Jersey

In partial fulfillment of the requirements

For the degree of

Doctor of Philosophy

Graduate Program in Computational and Integrative Biology

Written under the direction of

Dr. Andrey Grigoriev

And approved by

_____

Andrey Grigoriev

_____

Jongmin Nam

_____

Nancy Bonini

Camden, New Jersey

May 2015

ABSTRACT OF THE DISSERTATION

Temporal Patterns of non-coding RNA Modifications and Expression

by AMMAR S. NAQVI

Dissertation Director:
Dr. Andrey Grigoriev

We investigated the function and properties of small RNAs, particularly microRNAs and tRNA-derived fragments (tRFs) with age. We report the characterization of a novel 3'-to-5' exonuclease, Nibbler (Nbr), that generates differing isoforms of miRNAs in Drosophila. We developed a robust approach to help identify and characterize 3' heterogeneity in microRNAs controlled by Nbr, which assisted in identifying age-associated traits, including neurodegeneration and lifespan. Subsequently, given the fact Nbr interacts with Ago1 and not Ago2, we observed an accumulation of certain isoforms, which lead us to ask if there were particular patterns and trends that were Ago-specific. Interestingly, we report a novel age-associated change of select isoforms with age that is Ago2 specific. RNA deep-sequencing analysis coupled with experimental evidence reflected an increased loading of miRNA isoforms into Ago2 with age. Essentially, the loss of methylated miRNAs led to accelerated brain degeneration and shortened lifespan. Intriguingly, we also observed and identified Ago-loaded tRFs, which appear to have properties similar to those of miRNAs. We found this class of small RNAs to also display age-associated changes. For the first time, we found that differentially loaded Drosophila tRFs mapping to both nuclear and mitochondrial tRNA genes associating with all 20 amino acids. These tRFs show a number of similarities with miRNAs, including seed

sequences, suggesting a similar role and function. Moreover, we further characterized and predicted targets for these tRFs and show a significant enrichment in development and neuronal function, suggesting a role in brain-related processes with age. In sum, we discovered a novel component of the canonical microRNA biogenesis pathway, responsible for the generation of multiple isoforms. We also connected specific age-associated patterns and trends of select microRNA isoforms, which were found to impact proper brain development and lifespan. Moreover, we identified differentially loaded tRFs and elucidated their structures, loading, and expression patterns, which corresponded closely with microRNAs. Finally, we were able to identify tRF seed regions that potentially play a role in brain activity or brain changes with age.

ACKNOWLEDGEMENTS

I am extremely thankful for my advisor, Dr. Andrey Grigoriev. His openness and ideas allowed me to pursue the study of small non-coding RNAs in a very meaningful way. Through his mentorship, I was able to contribute to the field. Under his supervision, I have learned a great amount that has given me the confidence and skill set to continue as a scientist. I also would like to thank Dr. Nancy Bonini for her support and leadership that helped me succeed as a graduate student. Her lab's experimental expertise and contributions were invaluable. I would also like to thank Dr. Jongmin Nam for his ideas and feedback throughout my tenure as a graduate student. The combined support of my committee members has really helped me grow as a scientist and I am forever indebted for their help.

I would also like to thank the Grigoriev Lab members for their added insight and assistance and members of the Bonini Lab, Masashi Abe and Virzhiniya Feltzin, particularly for their experimental help. Finally, I want to thank my family and friends for their unconditional support that made my research possible.

Table of Contents

List of Tables

List of Figures

CHAPTER 1 INTRODUCTION

microRNA Biogenesis

microRNAs (miR) are the best understood class of small non-coding regulatory RNAs that work via translational repression or mRNA cleavage. They are approximately 20-24 nucleotides (nt) in length that regulate very broad-based biological processes and are essential for proper development (Flynt and Lai 2008, Lai 2003, Ambros 2004). The canonical biogenesis pathway, though is still being updated with the discoveries of new components and modifications, has been well characterized and has been evolutionarily conserved and maintained in a variety of organisms. First, a primary transcript (pri-miR), which consists of local inverted repeats, that produces a stem loop-like structure is recognized and cleaved by RNase III in a complex with its partner protein, Drosha/Pasha, to yield the pre-miR hairpin (Denli et al. 2004, Han et al. 2004, Han et al. 2006). This generated hairpin is then exported from the nucleus to the cytoplasm by the Exportin-5 protein to undergo further processing (Kim 2004, Lund et al. 2004). It is cleaved in the cytoplasm by RNase III enzyme and its partner protein, Dicer-1 (Dcr-1)/Loquacious-PB (Loqs-PB) to generate a duplex with a 2’ nt overhangs on each side. The duplex consists of a mature and a star strand or a 5' and 3' arm. The mature strand, which may be on either the 5’ or 3' arm is predominately loaded to an Argonaute (Ago) protein, the core component of the RISC complex (Bernstein et al. 2001, Chendrimada et al. 2005, Forstemann et al. 2005, Miyoshi et al. 2010). Upon loading, facilitated by a chaperone complex (Hsp90/Hsp70), the mature strand is bound and retained in Ago through the binding of the 5' end to MID and the 3' end to PAZ domains of the Ago protein (Miyoshi

et al. 2010, Iwasaki et al. 2010, Ma, Ye, and Patel 2004, Parker, Roe, and Barford 2005). Eventually the Ago protein carries the miR to its target transcripts for regulation, where the miR then hybridizes with the mRNA based on sequence complementarity. Though the Ago protein the central component of the RISC complex, it is assisted by other proteins, including GW182, CCR4-NOT and PAN2-PAN3 complexes (Fabian and Sonenberg 2012).

Notably, there have been some exceptions reported that deviate from this traditional pathway (Yang and Lai 2010). For example, the star strand is usually degraded, but at times it may also be loaded (Okamura, Liu, and Lai 2009). Another example is that Dicer does not cleave miR-451, but instead Ago does. This non-Dicer cleavage may be a result of its unusual structure. The miR-451 precursor is structurally unique in that the mature miR-451 sequence harbors a very small (~17nt) stem and a very large loop region. And for proper cleavage, Dicer requires a stem region of >19 nts, which may explain how this miR escapes Dicer cleavage (Yang and Lai 2010, Yang et al. 2010). Similar studies have also showed evidence for single stranded sequences originating outside of the 5'/3' duplex, being loaded onto the Ago proteins, instead of the canonical strand from the duplex portion of the pre-miRs (Okamura, Liu, and Lai 2009, Okamura et al. 2011). Intronic derived miRs called mirtrons, also exist that are cleaved by the splicosome instead of Drosha (Westholm and Lai 2011, Flynt et al. 2010, Ruby, Jan, and Bartel 2007, Okamura et al. 2007). Together, these studies suggest that miR biogenesis is highly complex and advanced. This also leaves the possibility of processing other classes of

small regulatory RNAs, such as tRNA-derived small RNAs (tRFs) to be cleaved in a similar fashion.

MicroRNA modifications, heterogeneity, and diversity

Emerging evidence has also surfaced suggesting that the biogenesis of these miRs is under precise control, which includes specific sequential cleavage, 3'-trimming and similar modifications, such as editing. Editing is a phenomenon that is often observed in RNA and has been shown to play important roles in development, tissue specificity and RNA structure. RNA editing is a molecular process in which the information content in an RNA molecule is modified through a chemical change in the base makeup. RNA editing events generally include nucleoside modifications, cytosine (C) to uracil (U) and adenosine (A) to inosine (I) deaminations, as well as un-templated nucleotide additions, deletions and insertions. These have been observed in tRNA, rRNA, mRNA and more recently in microRNA, where it has been shown that editing may be involved in target selection, degradation and stability, which greatly influence the expression and regulation of the genome (Luciano et al. 2004, Kim et al. 1994, Seton-Rogers 2012, Garcia-Lopez, Hourcade Jde, and Del Mazo 2013, Mehler and Mattick 2007, Li and Church 2013).

Adenosine (A) to inosine (I) mediated RNA editing is an important nucleotide modification that generates RNA and protein diversity in higher eukaryotes, selectively altering coding and non-coding sequences in nuclear transcripts (Maas, Rich, and Nishikura 2003). It is also the best-characterized type of editing that occurs in metazoans. The enzymes responsible for A-to-I editing, the adenosine deaminases acting on RNA

(ADARs) or dADARs in Drosophila, are ubiquitously expressed in mammals and specifically recognize partially double-stranded (ds) or fold back RNA structures where they then modify individual adenosines (Bass 2002, Hurst et al. 1995). It has been shown that edited sites are usually clustered together depending on the length of the ds sequences of the RNA molecule (Morse, Aruscavage, and Bass 2002). Hence, any RNA that forms a fold-back structure may be a template for editing, which makes miR precursor structures ideal for editing, implicating the importance of understanding potential is editing events in miRs. However, in miRs they are site-specific due to the nature of the pre-miR duplex structure, which includes ds and single stranded (ss) bulge-like regions (Luciano et al. 2004).

The editing of miR precursors by ADAR has had major implications for miR analysis, biogenesis and function. It has been noted that editing has influenced strand selection by essentially de-stabilizing the 5' end of the pre-miRs. Sites that lie within the seed portion of the miR, which is the site where hybridization between miR and mRNA occurs, effects target selection, influencing the cells in a very different manner. Hence, editing events may not only change the actual target, but it may also influence its targeting efficiency. There have also been reports of miRs that exhibit stronger editing potential, when compared to other miR loci, suggesting certain miRs may be more dynamic in their function by possessing increased likelihoods for editing. Essentially, editing adds another layer of genome regulation by regulating miRs (Seton-Rogers 2012, Garcia-Lopez, Hourcade Jde, and Del Mazo 2013, Mehler and Mattick 2007, Li and Church 2013, Bass 1997, Polson, Bass, and Casey 1996).

With the advent next generation sequencing (NGS) the discovery of differentially expressed isoforms or isomiRs have been characterized across varying species. This diversification of miR species may be a result of alternative cleavage sites, non-templated additions, 3' deletions, or RNA editing. Accumulating evidence suggests that such heterogeneity varies depending on the specific cell types and conditions (Berezikov et al. 2011, Ruby et al. 2007, Burroughs et al. 2010, Westholm et al. 2012, Wyman et al. 2011).

Alternative cleavage at the 5' end of a miR stem-loop has been shown to be biologically relevant, by virtue of changing loading efficiency and gene targeting (Azuma-Mukai et al. 2008, Fukunaga et al. 2012, Lee and Doudna 2012, Seitz, Ghildiyal, and Zamore 2008). In the context of NGS, this may ascertained by aligning reads in a particular library to canonical pre-miRs and assessing deviations from known start positions of cleavage. Particularly, this type of phenomenon leads to distinct target specificity and potential differences in guide strand selection, since this may occur in and around the seed region of the miR.

On the contrary, the generation of miR isoforms with 3' heterogeneity and its impact on the biology of the cells may be a bit less straightforward. It has been reported that long primary miR transcripts undergo subsequent sequence cleavage to release the embedded miRs. One notable example that was found to display isoMirs that contain different 3' ends with identical start positions was miR-34 (Liu et al. 2011, Liu et al. 2012). This was

an important finding, because mir-34 is conserved across species and is directly involved with aging and neurodegeneration. Shorter isoforms accumulated with age, suggesting a novel biogenesis mechanism involving 3' end processing. It was discovered that the 3'-5' exoribonuclease CG9247/nibbler was necessary for the generation of the smaller isoforms in miR-34. Nibbler (Nbr) interacts with Ago1 and processes miR-34 within the RISC complex, which indicates that both Ago1 and Nbr are required for miR 3' trimming. With the help of deep sequencing analysis and bioinformatics, a subset of miRs were identified to also be modulated and controlled by Nbr. These studies confirm that this new component and process involving Nbr to the miR biogenesis is a key step for processing and maturation. The consequence of 3' trimming may impact the stability by altering miR turnover rate or miR loading or strand selection, as it modifies the degree of duplex pairing (Han et al. 2011, Liu et al. 2011). A more drastic effect, may be related to mRNA targeting. In Drosophila and animal miRs, a "seed" region, which is usually position 2 to 8 of the miR or mature strand, is what primarily binds to the mRNA target, but there have been cases of the nts in the 3' region of the miR playing an important compensatory role. In other words, if there are mismatches between the miR-mRNA interactions within the seed, the 3' non-seed portion may rectify the mRNA regulation by assisting in binding (Brodersen and Voinnet 2009, Brennecke et al. 2005, Elefant, Altuvia, and Margalit 2011). Hence, 3' trimming or those miRs affected by that may have major biological impact.

In addition to trimming, another common 3' end modification that also contributes to 3' heterogeneity is by non-templated nucleotide additions. Again, with the help of NGS,

studies have identified 3' end nucleotide additions across species, including C. elegans, Drosophila, mouse and human (Burroughs et al. 2010, Ruby et al. 2007, Westholm et al. 2012, Berezikov et al. 2011). The most common additions are As or Us. Seven nucleotidyl transferases are implicated in 3' end adenylation and/or uridylation of miRs in humans. These enzymes add non-templated nucleotides to the 3' end to RNAs. These additions also have critical implications in miR processing, including degradation or stabilization. For example, PAPD4/GLD-2, adenylates the 3' end of miR-122 in mammals (human and mouse) and this leads to a stabilization effect on the miR (Katoh et al. 2009). On the other hand, added As negatively regulate or effect the efficiency of miRs, including miR-27a and miR-26a. Uridylation by Zcchc11 (PAPD3/TUT4), an uridyltransferase that add Us to miRs, usually has a negative effect by de-stabilizing miRs (Burroughs et al. 2010). Taken together, these additions either increase or decrease miR efficacy by affecting its stability and turn over rate.

tRNA-derived Small RNAs

Another class of small RNAs, which are in the process of being elucidated and amongst the poorest characterized are tRNA-derived small RNAs. Traditionally, transfer RNAs (tRNAs) have been seen as key players, acting as adaptor molecules in protein translation, but relatively recently there have been multiple attempts to understand them as regulatory molecules (Garcia-Silva et al. 2012). Using NGS technology, population of these small RNAs has become much more easily detectable and identifiable, lending to deeper analysis of them.

Since these fragments are derived from tRNAs, a very conserved type of RNA, it suggests a primitive RNA silencing pathway. They were first discovered in bacterial cells under specific stress conditions, serving as a protective response (Levitz et al. 1990). Later studies discovered these molecules to be present in protozoa, zebrafish, mouse and human (Li et al. 2008, Levitz et al. 1990, Wei et al. 2012, Gong et al. 2013, Li, Ender, et al. 2012, Lee et al. 2009, Yeung et al. 2009, Cole et al. 2009, Haussecker et al. 2010). Since then, there have been multiple attempts to understand these regulatory molecules.

There are two main species of tRFs that are categorized based on length and biogenesis, including tRNA halves and tRNA-derived fragments (tRFs). tRNA halves we discovered first and are usually generated from the mature tRNA, which is cleaved in the anticodon loop portion of the tRNA. Their length may range anywhere from 28 to 40 nts (Thompson and Parker 2009b, Tuck and Tollervey 2011). These halves are generated and cleaved usually after enduring stress, including starvation, temperature stress, hypoxia and oxidative stress (Fu et al. 2009). Cleavage is carried out by the RNase A enzyme angiogenin or RNase T1 family member Rny1 in mammalian and yeast genomes, respectively. These cleaved tRNA halves are functional by repressing translation, promoting stress granule assembly or directly interfering with the siRNA pathways (Emara et al. 2010). In contrast to tRNA halves, tRNA-derived fragments (tRFs) are shorter (~16-24nt) and can be classified into three types based on the tRNA region from which they are generated: 5' tRF, 3'CCA and 3'U tRF. The latter two types originate from the 3’ end of the tRNA, while the former is derived from the 5’ end. The 3'CCA type is generated from the 3' end of the mature tRNA and includes the post-transcriptional CCA

addition. The 3'U type is derived from the precursor tRNA and contains multiple Us in the 3’ end (Lee et al. 2009). This class, in comparison to tRNA halves, has been more heavily studied recently, especially with recent technologies allowing us to detect small RNAs that may be apparently low in quantity or less abundant.

There have been various attempts to determine the biogenesis and function of these different types of tRFs, but currently these questions are still open to investigation. Several studies suggest that the biogenesis of tRFs is similar to that of miRNAs and siRNAs. In one case, levels of a mature 5’ tRF were shown to decrease upon Dicer knockdown in HeLa and HEK293 cells. Moreover, the same study demonstrated in vitro generation of the tRF with recombinant Dicer (Cole et al. 2009). Another study described several 3’ CCA tRFs in HEK-293 cells, which were Dicer-dependent and were 5’ phosphorylated (Babiarz et al. 2008). These modifications are characteristic for Dicer products and contrast with 3’ tRNA halves, which carry a 5’ hydroxyl (Haussecker et al. 2010). In addition, a Dicer-dependent 3’ U tRF was identified in mouse ES cells and it was hypothesized to be generated from an alternative hairpin secondary structure of pre-tRNA Ile (Babiarz et al. 2008). However, two alternative studies, which rely on large-scale computational analysis of sequencing libraries, report that the majority of 5’ and 3’ CCA tRFs in mammalian cells and tissues are Dicer-independent(Kumar et al. 2009, Li et al. 2008, Li, Ender, et al. 2012). In the same study, Angiogenin, which cleaves tRNAs to generate tRNA halves, is also proposed to play a role in the biogenesis by cleaving tRFs (Li, Ender, et al. 2012). As a result, multiple pathways for tRF processing have been proposed, which may be species/context-specific (Gebetsberger et al. 2012).

tRNA-derived fragments have been hypothesized to function like or impact miRNAs, by either regulating mRNAs (similarly to miRNAs) or regulating miRNA loading and disrupting miRNA processing (Cole et al. 2009, Li, Ender, et al. 2012, Miyoshi, Miyoshi, and Siomi 2010). tRFs have also been shown to bind to Argonaute-RISC complexes, further indicating that they participate in RISC-mediated gene silencing. For example, the 3' U tRF and a 3' CCA tRF have been shown to be associated with Ago proteins in humans, suggesting a miR-like function (Haussecker et al. 2010, Kumar et al. 2014, Yeung et al. 2009).

Previous studies have demonstrated regulatory function of these tRFs by postulating that they bind and repress mRNAs in a fashion similar to miRs and at times even compete with miRs. It is unclear if they act like plant miRs that are fully complementary to their targets, or like animal miRNAs that have a specific pairing via a "seed' region found on the 5' end of the molecule. There have been conflicting models of such seed regions, lending to their complexity. One of them has suggested a traditional miRNA-like silencing based on complementarity of the 5' seed sequence of a tRF to a short sub-sequence within a 3' UTR of a transcript; another has shown that the last 8-10 nucleotides (nts) on the 3' end of the tRF in the 5' portion of the full tRNA are responsible for mRNA repression. Regarding their functionality, there have been evidence that their targets and expression have been connected to metabolism, stress, and differentiation, suggesting their significance as regulatory molecules for proper cellular growth and maintenance (Fischer et al. 2011, Haiser et al. 2008, Peng et al. 2012, Wei et al. 2012).

Aging, neurodegeneration and small RNAs

In the context of aging small RNAs, including miRs and tRFs, are differentially expressed (Pincus and Slack 2010, 2008). The impact of these small RNAs on the aging process is still being unraveled, but there have been studies confirming the differential expression of certain miRs, suggesting a biological impact (Kato et al. 2011, Ibanez-Ventoso et al. 2006, de Lencastre et al. 2010). Some miRs are down regulated, while others are up regulated with age. For example, in C. elegans miR-246, miR-71, miR-34, miR-253, miR-238 and miR-239 increased in total levels, while let-7 showed an overall decrease with age. Deletion of miR-71, miR-238, or miR-246 shows a significantly shorter lifespan, while the deletion of miR-239 shows increased lifespan. By contrast, the upregulation of miR-71 and miR-246 extend lifespan, while the upregulation of miR-239 shortens lifespan (de Lencastre et al. 2010). In essence, a miR may regulate multiple pathways, and a pathway may be regulated by multiple miRs in either a positive or negative manner (Chen et al. 2010).

One important example is miR-34, that is upregulated in both C. elegans and Drosophila with age (de Lencastre et al. 2010, Ibanez-Ventoso et al. 2006, Kato et al. 2011, Karp et al. 2011). However, although mir-34 expression changes in C. elegans' mutants either did not affect or extend life span (de Lencastre et al. 2010), while in Drosophila either caused a shorter lifespan, or an extension of life span depending on if its down-regulated or up-regulated, respectively (Liu et al. 2011). This hints at a species-specific effect of mir-34, illustrating small RNA complexity.

In addition to subsequent profiling studies that sought to reveal the role of miRs in aging, disruption of global small RNA biogenesis also suggests a potential importance of miRNAs and other small RNAs to impact lifespan. For example, in C. elegans, a knockdown of an argonaute protein led to shorter lifespan (Kato et al. 2011). Additionally, a knockout of DGCR8/pash-1 also revealed a shortened lifespan in C. elegans (Lehrbach et al. 2012). Furthermore, in Drosophila, the loss-of-function of Loqs, which is required for the cleavage of pre-miRNAs, lead to a shorter lifespan and brain degeneration (Liu et al. 2012). These studies supported a general importance of miRNAs to modulate age-associated events.

More specifically known gene targets of miRs have been actors in regulating lifespan in organisms and modulating aging pathways (Chen et al. 2010). Several pathways have been identified, including insulin/insulin-like growth factor 1 (IGF1) signaling (IIS), DNA damage checkpoint, and mitochondrial function regulate aging in animals (Kenyon 2010, Smith-Vikos et al. 2014, Smith-Vikos and Slack 2012). With that, the best-characterized age-associated pathway that has shown to intersect best with miRs is the IIS pathway. Reduced IIS signaling extends lifespan in C. elegans, Drosophila, mouse, rats, and humans (Kenyon 2010, Smith-Vikos et al. 2014, Smith-Vikos and Slack 2012). In this particular pathway the lin-14, IGF-1R, IR8-1, and IL-1, IGFBP-1 genes are all regulated by lin-4, mir-1, mir-320, mir-206, mir-145, and mir-140 (Boehm and Slack 2005, Elia et al. 2009, La Rocca, Badin, et al. 2009, La Rocca, Shi, et al. 2009, Miyaki et al. 2009, Shan et al. 2009, Wang, Qian, et al. 2009, Yu et al. 2008, Tardif et al. 2009).

The extended lifespan is associated with improved stress resistance and protein homeostasis (O'Neill et al. 2012). A subset of miRs have shown to target genes involved in this pathway in a concordant and parallel manner reflecting an intricate method of regulating age-related pathways (Chen et al. 2010).

Nevertheless, RNA sequencing analyses in C. elegans and Drosophila has also revealed age-associated changes of other classes of small RNAs such as endo-siRNAs, piRNAs, and tRFs. Specifically, in C. elegans tRFs displayed a consistent increase in four different time points (0, 5, 8 and 12 days) by approximately 15%, providing evidence that they also may be associated with age, especially since it has been shown that tRNA cleavage is induced in response to stress in several organisms (Kato et al. 2011).

Nevertheless, a connection between miRs and neurodegenerative conditions has also been established. This class of diseases is a group of late-onset, progressive disorders that lead to cognitive and/or movement disorders. Some of the more common ones include Alzheimer's disease (AD), Parkinson's disease (PD), amyotrophic lateral sclerosis (ALS), and polyglutamine (polyQ) disorders. All of these conditions share the signature feature of the accumulation of key proteins, which can be linked to familial mutations (Ballard et al. 2011, Ballard and Fox 2006, Ballard et al. 1998, Orr and Zoghbi 2007, Ferraiuolo et al. 2011, O'Brien and Wong 2011).

There have been multiple approaches that have been taken to study the influence of miR on these diseases. The most straightforward methods of this sort of study is to disrupt

levels of miR, investigate miR levels, and inspect how disease-associated proteins effect miR processing with age. A functional link between small RNAs and neurodegeneration was discovered in studies of the effect of global disruption of their biogenesis on neuronal development.

Supporting the role of miRs, mir-430 was shown to rescue defects in brain morphology, indicating the importance of this specific miR. Bantam was also identified to specifically modulated Ataxin-3 and Tau toxicity (Schaefer et al. 2007, Kim et al. 2007, Choi et al. 2008). More recent studies on miR-34 revealed a protective function in mitigating the toxicity of pathogenic forms of Ataxin3 (Liu et al. 2012).

Mutations introduced to Dicer revealed a role in brain morphology, development, and differentiation (Giraldez et al. 2005, Kim et al. 2007; Schaefer et al. 2007; Choi et al. 2008; Damiani et al. 2008; Davis et al. 2008; McLoughlin et al. 2012). In mammalian cells, loss of Dicer was linked to brain degradation, including myelin and axon integrity (Shin et al. 2009, Tao et al. 2011, Bremer et al. 2010, Pereira et al. 2010) . In Drosophila, a knockdown of Dicer-1 is associated with dopaminergic neural loss and climbing defects. Additionally, loss of Dicer-1 also enhances the toxicity of neurodegenerative disease proteins, including Ataxin-3 and Tau (Gehrke et al. 2010, Bilen et al. 2006). Other components in the small RNA pathway, including DGCR8, R2D2, and Loqs also lead to neuronal dysfunction in the mouse (Fenelon et al. 2011, Stark et al. 2008, Schofield et al. 2011, Marques et al. 2010). These studies, which target disruption of components of the miR biogenesis pathway, strongly suggest that miR activity impacts

long-term brain integrity. Notably, complimentary studies with microRNAs is also a very critical step, as these components may have broad impact on a plethora of other important functions in the cell, not related to small RNA processing (Marques et al. 2010).

Ascertaining miR targets that are associated with these conditions have also of interest. In general, discovering specific miRNAs that target the 3' UTR of key disease genes, then assessing the expression pattern and level of those miRNAs, can uncover the extent to which they may impact the level of the disease protein and thus pathogenesis. Mutating specific miRs, such as mir-29, miR-107, miR-124, and miR-195 causes disruptions in pathways that cause AD. ALS is characterized by the degeneration of motor neurons in the brain and spinal cord. In ALS, miRs, such as mir-206 and mir-8 affect ALS (Choi et al. 2008, Hebert et al. 2008, Fang et al. 2012, Zhu et al. 2012, Wang et al. 2008, Wang, Liu, et al. 2009). Other small RNAs, such as tRFs have not been directly implicated in such diseases, but the fact they exhibit age associated trends and share properties with miRs, suggests a similar role. Their predicted targets also overlap with targets of mir-29, miR-107, miR-124, and miR-195, which are miRs involved with conditions related to bi-polar disorder and schizophrenia (Moreau et al. 2011, Perkins et al. 2007).

NGS usually refers to the recently developed technology, which enables for high-through-put systematic sequencing based on the Sanger sequencing method. As mentioned earlier with the introduction of NGS and other technological advances, the study of small RNAs have been of great relevance. There have been striking revelations of their biogenesis pathways, which have lead to the identification of novel components

involved in processing and modifying small RNAs. Additionally, the characterization of structural trends, loading patterns, and differential expression in the context of differing conditions has also been unraveled. Nevertheless, even the discovery of new classes of small RNAs, including 21-U-RNAs and exceptions such as mir-454 and mirtrons has also ensued. Combining NGS and computational analyses with traditional experimental techniques, such as immunopreciptation, has allowed us to confirm and validate our predictions and key findings.

In chapter 2, in order to provide an intuitive environment to the understanding of these patterns, including read distribution and secondary structure, we have developed a pipeline to help visualize NGS data in varying conditions by allowing us to overlay modifications, expression, and loading with secondary structure. Chapter 3 details the discovery of a novel 3'-to-5' exonuclease, Nibber (Nbr), that generates diversy population of miRs by 3' end processing in Drosophila. Originally, the Bonini Lab reported an impact of Nbr on mir-34, which is an important miR associated with neurodegeneration and brain function. However with advanced computational and bioinformatics' analysis a subset of miRs were identified as being directly effected by Nbr by observing in-depth comparisons between wild type isoforms and other isoforms. Furthermore, given that miR-34 showed a change in isoform pattern with age via Northerns, we queried whether other Nbr-dependent miRs show changes in their isoform pattern with age. This led us to test other Nbr-dependent miRs for isoform pattern changes with age, detailed in Chapter 4. This not only revealed distinct isoform pattern changes of different miRs with age, but also led use to the novel findings that 2'-O-methylated Ago2 loaded miRs increased with

age. This led us to do an in-depth analysis of Ago1 vs Ago2 miRs. We found that miRs and specific isoforms are differentially loaded; more specifically these miRs showed increase Ago2-loading, while a decrease loading in Ago1 suggesting a mechanism to control differential partitioning of small RNAs into different Ago complexes with age. Mutations in Hen1 and Ago2 increased brain degradation and shortened life span, hinting at the biological impact of these miRs. Nonetheless, when we looked at total small RNA populations we noticed there was also a significant increase of Ago2-loaded small RNAs that were derived from tRNA molecules, suggesting that this class of small non-coding RNAs may also be modulated with age. The significant populations of these small RNAs helped us further characterize properties of these tRNA-derived small RNAs (tRFs) and how they are related to the aging process in the context of expression and loading patterns. In particular, we found tRFs were expressed and also loaded similar to that of miRs. We also found these tRFs to be associated with all known amino acids. Additionally, for the first time, we observed small RNAs cleaved from mitochondrial tRFs that were loaded onto these Ago proteins. Finally, we further characterized their targeting, including tRF-mRNA binding and mRNAs targeted. These predicted targets were enriched for genes controlling neurological or brain processes similar to those of miRs granting us confidence in their biological relevance. Taken together, these studies reveal new insights into a mechanistic and biological understanding the small RNA loading, processing, and expression that are modulated with the aging process and their impact on the brain.

CHAPTER 2 VISUALIZATION OF NUCLEOTIDE SUBSTITUTIONS IN THE (MICRO)TRANSCRIPTOME

This work has been published as follows:

Visualization of nucleotide substitutions in the (micro)transcriptome

Ammar Naqvi, Tiange Cui, Andrey Grigoriev*

Biology Dept., Center for Computational and Integrative Biology, Rutgers University, 315 Penn St, Camden NJ 08055, USA

*Correspondence to: agrigoriev@camden.rutgers.edu

Note

The format of the figure and table numbers, and references have been modified from that published to conform to the format of the dissertation.

Contribution

The project was based on the Genome Navigator code developed in the Grigoriev Lab. I suggested a new application for Genome Navigator and was involved with devising it by adding important layers to the tool, which included editing/snps, secondary structure in a more intuitive manner. This design was implemented by Tiange Cui by modifying/adding code and introducing publically available libraries from the Sequence Reading Archive (SRA) database. The resulting tool produces a comprehensive picture, which is necessary

for large-scale studies and detailed analyses. I performed analyses of the editing patterns and candidates. Together with Dr. Grigoriev, I drafted the manuscript and edited it according to the reviewers' comments.

Abstract

Background

RNA-related applications of the next-generation sequencing (NGS) technologies require context-specific interpretations: e.g., sequence mismatches may indicate sites of RNA editing, or uneven read coverage often points to mature form of microRNA. Existing visualization tools traditionally show RNA molecules in two dimensions, with their base pairing and the resulting secondary structure. However, it is not straightforward to combine a linear NGS data display with the 2-D RNA depictions.

Results

We present a novel approach for interactive representation of nucleotide substitutions and modifications in the transcribed genome. With the focus on RNA secondary structure in the context of NGS data, it provides intuitive visualization of genomic environment, sequence reads, nucleotide polymorphisms and editing events integrated with the structural and functional elements of both coding and non-coding RNA molecules. Using our approach we present and discuss examples and general trends of polymorphisms and editing in the context of the secondary structure of microRNAs. As expected, most of the substitutions comprised A to G and C to T events, consistent with typical RNA editing patterns. However, we did not observe prevalence of editing in double-stranded regions

of the microRNA stem-loop. We describe novel prominent editing event candidates, observed across several small RNA libraries of Drosophila melanogaster.

Conclusions

In contrast to the existing general tools for NGS data visualization, the power of our approach is not only in the display of read alignments and their counts, but the integration of RNA secondary structure, sequencing depth, and rates/patterns of editing or other modifications. It provides a comprehensive picture, important for large-scale studies and detailed analyses, helping to gain insight into the intricate relationships between different events in RNA biogenesis.

Background

RNA molecules are traditionally shown in either one dimension in FASTA format or two dimensions, with the purpose of showing their base pairing and the resulting secondary structures, important for their stability and function. A number of tools for displaying RNA molecules in 2-D has been created, such as RNAViz, VARNA, RnallViewer, jViz.Rna and 4SALE , to name a few. Some of them provide not only a visual display but also tools for analysis and comparison of RNA molecules (De Rijk, Wuyts, and De Wachter 2003, De Rijk and De Wachter 1997, Seibel et al. 2006, Wan, Lin, and Xu 2006, Wiese, Glen, and Vasudevan 2005).

The advent of the next-generation sequencing (NGS) technologies has changed the research approaches in molecular biology and NGS is quickly becoming a standard. The

technologies are based on determining sequences of short fragments of DNA or RNA ("reads"), assembling them de novo into contigs or aligning them to a reference genome sequence and finding meaningful deviation of the sequence itself or the read coverage from expected models. In RNA-related applications, these interpretations are context-specific: for instance, sequence mismatches are often interpreted as potential sites of RNA editing, or uneven read coverage is taken as an indicator of mature form of microRNA. However, it is not straightforward to combine a linear graphical display of NGS data with the 2-D RNA depictions.

Some of the tools mentioned above provide a linear view of the RNA, but none allow for clear connection with genomic features. We have developed a novel representation of RNA secondary structure that is integrated with the display of reads generated by NGS. It is similar to a linear Feynman diagram but implemented in an interactive Java applet and provides intuitive visualization of genomic environment, sequence reads, nucleotide polymorphisms and editing sites together with the structural and functional elements of the encoded RNA molecules. It is not intended to replace the 2-D depictions but, rather, to usefully supplement them with providing visual links to genomic features and experimental data, which can be overwhelming in NGS projects.

MicroRNAs

MicroRNAs are endogenous short non-coding RNAs that are involved with regulation of messenger RNAs through either RNA degradation or translational repression. The widespread functionality of these molecules has been implicated in many different areas

including development and various disease states and conditions (Calin and Croce 2006). As a result, the studies of these molecules have become essential in understanding the plasticity of the transcriptome in relation to gene expression. Emerging evidence has also surfaced suggesting that the biogenesis of these microRNAs is under precise control, which includes specific sequential cleavage, 3'-trimming and similar events (Kim 2005, Liu et al. 2011).

It is important to note that the microRNA secondary structure plays a crucial role for proper processing of these molecules. For example, Drosha, an exonuclease, recognizes a transcript hairpin structure, which it then cleaves and generates a precursor microRNA. Additionally, Dicer, another protein involved with cleavage, must recognize the loop region and a specific nucleotide duplex in the microRNA stem-loop after nuclear transportation. The stem-loop length varies from microRNA to microRNA, but it includes a duplex that contains a mature microRNA and a star strand, as well as a loop region, which is removed and then subsequently degraded (Kim 2005, Czech et al. 2008). Furthermore, the strands consisting of the duplex are sometimes referred to as 5P and 3P arms or species of the microRNA stem-loop referring to their relative positions (1-based). In addition to the duplex and the loop, the stem-loop also includes small bulges across the RNA fragment. These secondary structure elements have been hypothesized to be, together with the 5' most nucleotide, the determinants of recognition-assisted partitioning and loading of the two microRNA strands into Ago1 or Ago2 complexes (Ghildiyal et al. 2010), although our report of microRNA partitioning/loading being age-dependent suggests further complexity (Abe et al. 2014).

RNA editing

Editing is a phenomenon that is often observed in RNA and has been shown to play important roles in development, tissue specificity and RNA structure (Kim et al. 1994). RNA editing is a molecular process in which the information content in an RNA molecule is modified through a chemical change in the base makeup. RNA editing events generally include nucleoside modifications, cytosine (C) to uracil (U) and adenosine (A) to inosine (I) deaminations, as well as un-templated nucleotide additions, deletions and insertions. These have been observed in tRNA, rRNA, mRNA and more recently in microRNA, where it has been shown that editing may be involved in target selection, degradation and stability, which greatly influence the expression and regulation of the genome (Luciano et al. 2004, Seton-Rogers 2012, Garcia-Lopez, Hourcade Jde, and Del Mazo 2013, Teng, Burant, and Davidson 1993).

The RNA editing is performed by the enzyme called Adar (adenosine deaminase acting on RNA), responsible for editing by site-specific deamination of adenosines. It specifically converts A to I. This type of editing is mostly active in the brain, but also has been implicated elsewhere, including various tissues and developmental stages (Mehler and Mattick 2007, Li and Church 2013). It is also worth noting that the position of an editing event may result in truncated products, splice variants, and structural changes. All of these results may change the functionality of a particular gene and contribute to genome's plasticity making it more dynamic than originally thought. Adar has been thus far understood to function and recognize double-stranded RNA substrates (Kim et al.

1994, Teng, Burant, and Davidson 1993). Another, less prevalent, type of editing is a C to U conversion executed by ApoB/APOBEC proteins in mammals (Teng, Burant, and Davidson 1993). No ApoB/APOBEC homolog has been reported in Drosophila, although methylated C is known to spontaneously deaminate to U. In this study, we use a novel visualization approach to overlap RNA secondary structure of known Drosophila melanogaster microRNAs with potential editing events or other single nucleotide changes.

We have developed an approach to integrate the RNA secondary structure display with the results of NGS projects, using linear Feynman diagram as a model. This graphical tool is implemented in an interactive Java applet and provides intuitive visualization of genomic environment, sequence reads, nucleotide polymorphisms and editing sites together with the structural and functional elements of the encoded RNA molecules (Figure 2.1). Here we illustrate this visualization approach with a few examples, highlighting connections between these different sequence and structural features for microRNAs in Drosophila melanogaster.

Results and discussion

General features of visualization

As is common in the analysis of the microRNA libraries, our display shows the number of reads with the given start/end coordinates. Such numbers are often used for the determination of the borders of the 5P and 3P forms of microRNA. A group of identical sequence reads is thus summarized as a single bar with the number of reads. Further,

color segments in the top line of the bar are used to highlight sequence mismatches in this group compared to the reference sequence, thus making it easier to observe specific editing patterns in a compact display. Alternatively, if necessary, individual sequences can be "spelled out" (with the similar highlighting of mismatches), as shown in Figure 2.1(abundant read with a mismatch, shown under the reference sequence).

Figure 2.1 is an illustration of the general features of the display using mir-1012. Line 1 shows the mismatch frequencies or the percentage of that particular mismatch event in all reads for the specified libraries. Line 2 displays the sequencing depth of the reads mapped to the microRNA stem-loop, while line 3 shows the genomic environment (exon-intron structure for a mirtron in this case) and line 4 - the reference sequence.
Library reads are shown in line 5 (some are removed from view for compactness as described below). The list of available libraries is shown here in the menu in the center of the display and the selected library is highlighted in pink in the main display. Libraries can be de-selected to remove from view for less cluttered display. In different libraries, identical reads are summarized in a single bar with the number of counts and, when applicable, the colors on top of the bar reflect the mismatches, as compared to the reference genome.

One can also inspect actual read sequences and the secondary structure, shown in line 6 and 7, respectively. This makes the RNA secondary structure visualization intuitive and as a result we can clearly see the pre-microRNA hairpin. In this case, the secondary structure consists of single stranded and double stranded regions referring to the 5P, loop

region, and 3P fragments of the stem-loop. Finally, the bottom line shows the single-stranded sections, while the double-stranded sections are above. The double stranded portion of the structure is clickable and also contains pertinent base-pairing information. In other words, if a user clicks on a position in this line, its corresponding base-pair in a stem will be connected by a magenta line in the display (as shown). Finally, the green boxes show bulge lengths in the RNA strand opposite to the bulge.

In this particular case (Figure 2.1), we observe that for mir-1012 the sequencing depth (in red) is significant in the 5P and 3P arms, indicating that both the mature and star strands are detected. We can easily extrapolate that mir-1012 is a mirtron and is Drosha independent, since it is located in an intronic region in the genome. More so, this visualization allows us to infer that single nucleotide changes mostly occur on the 3P arm in a bulge region. In sum, the display allows us to connect editing levels, events and locality, RNA structure, genomic regions, and NGS read counts and depth in a cohesive and meaningful manner.

Editing

The structural complexities of microRNAs, including 5'/3' trimming, editing, and base pairing and folding, can be conveniently visualized together using our approach. These features have been shown to be very important for proper modulation and for accurate processing (Kim 2005, Liu et al. 2011, Berezikov et al. 2011).

For illustration, we chose bantam (Figure 2.2), a microRNA that is ubiquitously

expressed and conserved with widespread associations in development and disease (Brennecke et al. 2003). We identified editing events using NGS data from two different libraries, which are specific to ovaries and Ago2 loading (Czech et al. 2008), respectively. Our examples show substitutions present in both double and single stranded regions. Since Adar, implicated in most of RNA editing events, act preferentially on double-stranded regions (Kim et al. 1994), we conclude that a non Adar-mediated editing mechanism may also be involved in producing these substitutions. Our observations of nucleotide changes other than A to G also support this idea (e.g., in bantam and other microRNAs, data not shown). On a more global level, we were also able to identify microRNA editing candidates by analyzing several publicly available libraries (Table 2.1).

For this purpose, we selected all cases where editing level (ratio of edited to total reads) exceeded a threshold of 10% with >20 reads per library. This filtering was done to avoid mismatches that result from sequencing errors. We found some 30 microRNAs that pass the filter (Table 2.2), with mir-986 and mir-971 being amongst the microRNAs with the highest editing level. We also observed that most of these microRNAs possess the same edits in multiple libraries furthering our confidence in these events. We also found that a C to U and A to G conversion events showed the highest level of editing (>60%) and read numbers, while other events were mostly supported by only a few reads and likely represented experimental noise. These results indicate that the canonical type and method of editing is amongst the most prevalent type and supports previous studies (Kim et al. 1994).

mir-986, the strongest candidate for editing, displayed 98-99% editing level in position 29 of the stem-loop. The functionality of mir-986 is still being elucidated, but the editing is significant due to the fact the site is 1-2 nucleotides away from the seed sequence and the event is conserved through multiple libraries. This allows us to speculate that the editing may play a major role in what mRNAs are targeted, specifically in the case of mir-986. The second top editing candidate microRNAs, mir-971 has already been reported as such in an earlier publication (Berezikov et al. 2011), although we did not observe editing events in their other candidates fulfilling our criteria across several libraries.

Editing in Single- and Double-Stranded Regions

We also analyzed the frequency of these editing events in double- and single-stranded regions. We quantified all unique strong editing incidences and positions to help us understand if there was preferential editing based on RNA secondary structure of microRNA. In other words, we wanted to see if there was a significant difference in editing between bulges or double-stranded regions (the duplex) as would be expected from the properties of Adar, if it discriminated against bulges.

For single- and double-stranded regions we obtained the observed (O_{ss} and O_{ds}, respectively) and expected (E_{ss} and E_{ds}, respectively) rates of editing from publicly available libraries (Table 1) as described in Materials and Methods. The ratios of observed to expected events of $O_{ss}/E_{ss} = 1.04$ and $O_{ds}/E_{ds} = 1.00$ for single- and double-

stranded editing sites, respectively, indicated that editing is essentially non-discriminatory when it comes to microRNA stem regions. We observed no editing events in the loop regions, even in those microRNAs that contained significant number of reads (>100) partially (alternative cleavage site) or entirely in loop regions, such as mir-34 (data not shown). A reason for this observation may be due to the fact that the loop regions are contiguously single stranded, which is not a canonical structure of an editing target. In contrast, single-stranded bulge regions in microRNA stems show a very slight prevalence for editing events, suggesting that Adar may tolerate bulges, if it is responsible for the editing events in microRNA.

Differential 5P/3P abundance

As is common in the analysis of microRNA libraries, our display shows the number of counts per read with given start/end coordinates. High read numbers are often used to determine the borders of the 5P and 3P forms of microRNA. In the case of bantam, our display may reflect differential biogenesis as in some libraries abundant reads do not match the canonical end sites (Fig. 2.2, compare the end sites with annotated 5P and 3P forms, shown as grey shades at the bottom). This can be viewed for any set of microRNAs by inspecting start/end positions of the read and comparing it to the grey shades on the bottom of our display (annotated 5P and 3P forms).

In bantam, we can observe a high number of read counts with the end positions on the 5P arm that differ from known mirBase annotations (Griffiths-Jones 2010, 2006). We see that numerous (>100 read count) unedited reads end at position 73, while the mature

microRNA is annotated as ending at position 74, thus indicating possible trimming of one nucleotide. We also observe many other highly expressed (>100 read counts) forms with different terminal sites. Additionally, the two libraries (GSM280082 and GSM280087) seem dramatically different in the assayed expression of the 5P arm (Fig. 2). While in this example we simply illustrate this difference of expression on a qualitative level, one can utilize appropriate normalization (based on RPKM, reads per kilobase per million, on read counts of endogenous siRNA, etc.) for quantitative measures to limit experimental bias or biological fluctuation.

The overall pattern of the NGS analysis of bantam suggests that Dicer cleavage can occur at multiple sites resulting in different isoforms. Notably, both of the most abundant isoforms on the 5P arm display the same substitution (editing event) in base 34. We also observe differential end positions on the 3P arm, which may hint at a 3' trimming event. In addition to our observations and conclusions above, what makes this example more interesting is that we see in different libraries reads that correspond to only mature or both the mature and star strand. We can also see how the same microRNA is modified differently in specific libraries, which may be an effect of the cellular environment. This is a further indication that the star strand may be functional and that functionality as well as editing may be context-dependent.

Conclusions

In this paper, we described an approach to display microRNA features and visualize them in the context of the genome environment, RNA secondary structure, nucleotide

substitution and NGS. A number of visualization tools has been developed for NGS (of which the Integrative Genome Viewer (Robinson et al. 2011) is perhaps the best known). What makes our approach distinct is not only the display of read alignments and their counts, but the integration of RNA secondary structure, sequencing depth, and rates/patterns of editing or other modifications. In essence it produces a more comprehensive picture, which is necessary for large-scale studies and detailed analyses. Using our approach we identified and inspected the microRNA editing candidates detected in several small RNA sequencing libraries. We observed a prevalence of canonical C to U and A to G editing patterns but found no editing bias between bulges and double-stranded regions.

Our results indicate that the modulation and modifications of microRNAs may be context-dependent or specific to experimental conditions; linking together editing, differential expression, modification, and secondary structure, helping us gain insight into the intricate relationships between these events. Finally, we can easily exploit our tool to depict other transcripts (both coding and non-coding, with and without introns) and genome/sequence features relevant for them, which makes our visualization approach both versatile and robust. The Java applet is available from the authors upon request.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

TC wrote the visualization program and scripts for data preparation, AN performed the data analysis and prepared the data for visualization, AG conceived and directed the study, wrote the visualization program. AN and AG wrote the manuscript.

Acknowledgements

This work was in part supported by the NSF grant DBI-1126052 to AG.

Declarations

The publication costs for this article were funded by Rutgers University.

This article has been published as part of BMC Genomics Volume 15 Supplement 4, 2014: SNP-SIG 2013: Identification and annotation of genetic variants in the context of structure, function, and disease. The full contents of the supplement are available online at http://www.biomedcentral.com/bmcgenomics/supplements/15/S4.

Figures and Tables

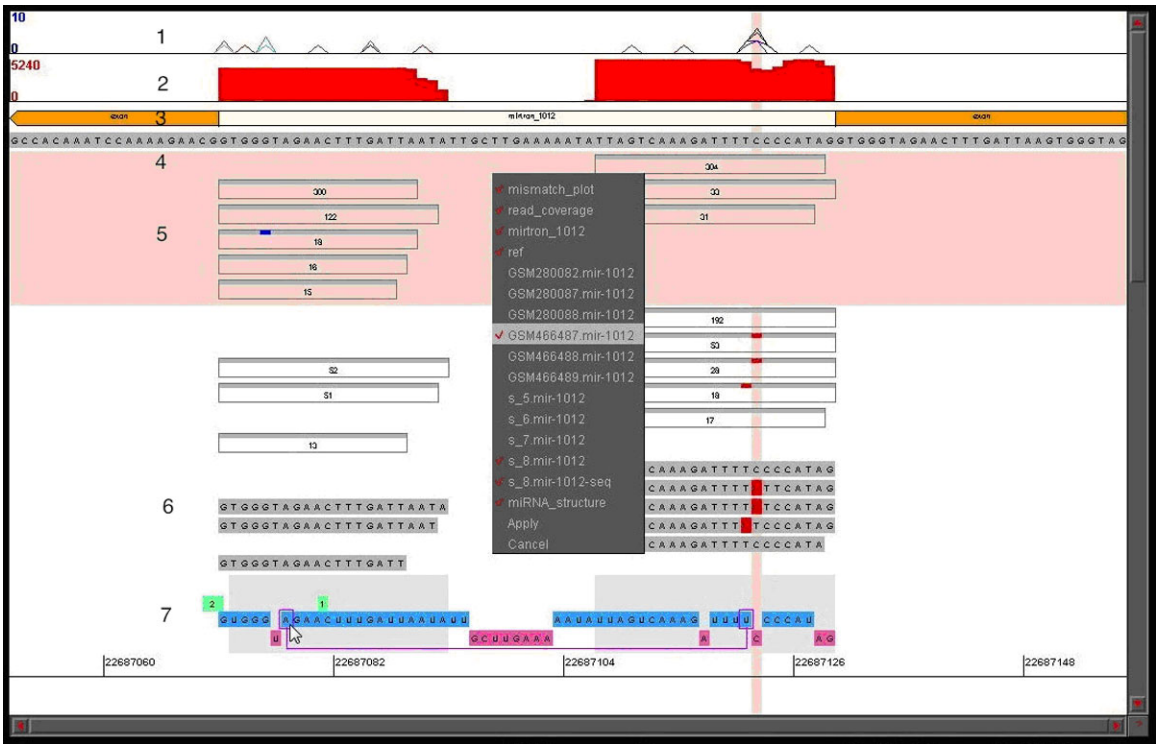


Figure 2.1 General Display Features. Description of menu, buttons, scale

Line 1: Mismatch frequency plot. 2: Match frequency histogram. 3: Structure of theregion (mirtron between two exons). 4: Reference sequence. 5: Reads & Libraries (library list in the menu in the center, selected library highlighted in pink). Identical reads are summarized as a bar with the number of reads. Colors on top of the bar show mismatches (compared to the reference genome sequence). 6: Read sequences are shown (with the similar color of mismatches). 7: Secondary structure. The bottom line shows single- stranded regions. The double-stranded regions are above, with base pairing encoded (if a user clicks on a position in this line, its corresponding base-pair in a stem will be connected by a magenta line in the display, as shown). Green boxes show bulge lengths in the RNA strand opposite to the bulge.

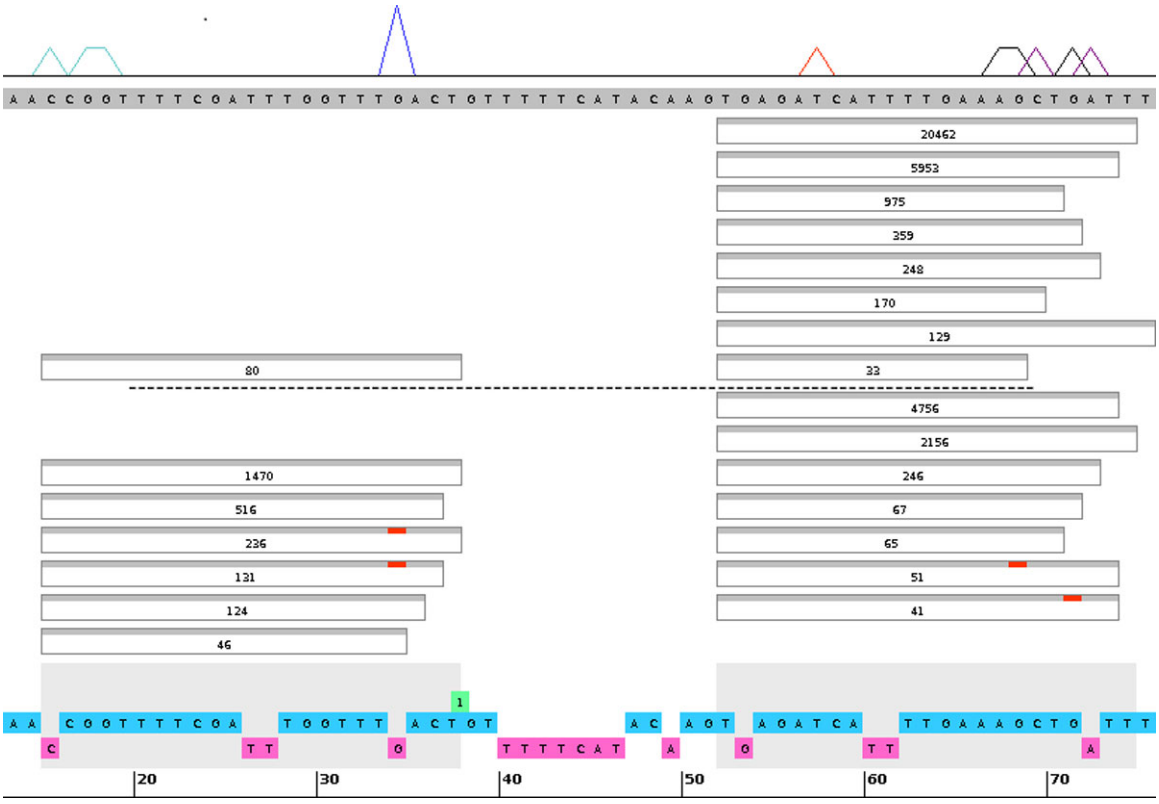


Figure 2.2: The bantam stem-loop edited regions and NGS reads from libraries GSM280082 and GSM280087. Dotted line separates the two libraries, with only groups containing >30 identical reads are shown. Strong editing is occurring in position 34, 71, and 68 in the stem-loop regions on both the 5P and 3P arms. These sites are found in both single stranded (34, bulge) and double stranded regions. We can also see that some very highly expressed perfect reads end at position 73, while the mature microRNA is annotated as ending at position 74, thus indicating possible trimming by one nucleotide. We also observe many other highly expressed forms with different terminal sites. Additionally, the two libraries seem dramatically different in the assayed expression of the 5P arm.

Table 2.1. Small RNA Libraries. All libraries can be downloaded from NCBI's GEO

database with the GSM identifiers below. Full description and references are also reported there.

Library	Description
GSM280082	Ovaries
GSM280087	Ago1-IP
GSM280088	Ago2-IP
GSM466487	Total RNA
GSM466488	Ago1-IP
GSM466489	2-O Methylated
GSM811191	Nbr WT
GSM811192	Nbr -/-
GSM807162	DL-1 cells Nbr WT
GSM807167	DL-1 cells Nbr -/-

Table 2.2. microRNA Candidates for Editing. All microRNAs with >20 reads and an editing level of at least 10% (ratio of edited to total reads) in Table 1 libraries.

microRNA	position	reads (mismatch/total)	Level (%)	type	library
dme-mir-312	63	3/29	11	G>T	GSM466487
dme-mir-11	67	3192/32796	10	G>T	GSM280087
dme-mir-312	65	3/29	11	T>A	GSM466487
dme-mir-1001	22	152/1491	11	A>G	GSM466488
dme-mir-970	82	381/3852	11	A>T	GSM466487
dme-mir-1010	68	11/101	11	G>T	GSM280087
dme-mir-304	32	12/105	12	G>T	GSM280087
dme-mir-3	54	8/78	11	A>G	GSM466488
dme-mir-313	68	3/28	12	A>G	GSM466488

dme-mir-970	82	880/7625	12	A>T	GSM466488
dme-mir-987	35	203/1751	12	G>T	GSM280087
dme-mir-1016	23	26/214	13	T>G	GSM807167
dme-mir-2489	79	6/48	13	G>T	GSM280087
dme-mir-31b	28	15/116	13	A>T	GSM466488
dme-mir-2489	79	4/29	14	G>T	GSM466487
dme-mir-1016	23	15/95	16	T>G	GSM807162
dme-mir-317	80	2888/17496	17	G>T	GSM807167
dme-mir-308	24	1582/8795	18	G>T	GSM280088
dme-mir-4975	48	4/23	18	T>G	GSM466488
dme-mir-317	80	1946/10538	20	G>T	GSM811191
dme-mir-986	29	494/2235	23	C>T	GSM466487
dme-mir-986	29	1924/8739	23	C>T	GSM466488
dme-mir-986	29	1498/6012	25	C>T	GSM807162
dme-mir-31b	28	14/55	27	A>T	GSM466487
dme-mir-988	83	158/527	30	C>T	GSM280087
dme-mir-986	29	42/80	53	C>T	GSM280088
dme-mir-971	75	387/599	65	A>G	GSM466487
dme-mir-971	75	1675/2293	74	A>G	GSM466488
dme-mir-986	29	13057/13492	98	C>T	GSM807162
dme-mir-986	29	9276/9490	99	C>T	GSM807167
dme-mir-986	29	504/514	99	C>T	GSM280087

CHAPTER 3 THE EXORIBONUCLEASE NIBBLER CONTROLS 3' END PROCESSING OF MICRORNAS IN DROSOPHILA

This work has been published as follows: Liu N[1], Abe M1, Sabin LR[2], Hendriks G-J[2], Naqvi A[3], Yu Z[3], Cherry S*, Bonini NM* (2011)

The exoribonuclease Nibbler controls 3' end processing of microRNAs in Drosophila. Curr Biol 21:1888-93. 2011 Nov 3. 1These authors contributed equally to this work. 2These authors contributed equally to this work. 3These authors contributed equally to this work *Corresponding authors

Note

The format of the figure and table numbers, and references have been modified from that published to conform to the format of the dissertation.

Contribution

My contribution to this study consists of the identification Nbr-dependent and independent miRs (Table 3.1, 3.3, 3.4, Figure 3.4F, and Figure 3.7). Originally it was found that mir-34 was impacted by Nbr, but with my deep sequencing computational analyses we were able to identify a subset of miRs that were also directly effected by

Nbr, indicating that this is a general mechanism that is responsible for 3' heterogeneity on multiple miRs. This was accomplished by first developing a pipeline to analyze the small RNA-seq NGS data and to discover trends and patterns of specific isoforms in wild type and Nbr mutant libraries. The computational predictions of these new Nbr-dependent miRs were then validated using RNA Northern blots.

Summary

MicroRNAs (miRNAs) are endogenous noncoding small RNAs with important roles in many biological pathways; their generation and activity are under precise regulation (Ambros 2004, Bartel 2004, O'Connell et al. 2010). Emerging evidence suggests that miRNA pathways are precisely modulated with controls at the level of transcription (Bommer et al. 2007, Chang et al. 2007, Hammond et al. 2001, Johnson, Lin, and Slack 2003, Newman and Hammond 2010, Raver-Shapira et al. 2007), processing (Hagan, Piskounova, and Gregory 2009, Hammond et al. 2001, Heo et al. 2009, Newman and Hammond 2010) and stability (Chatterjee and Grosshans 2009, Ramachandran and Chen 2008) with miRNA deregulation linked with diseases (Chang and Mendell 2007, Chang et al. 2007) and neurodegenerative disorders (Bilen et al. 2006). In the Drosophila miRNA biogenesis pathway, long primary miRNA transcripts undergo sequential cleavage (Bernstein et al. 2001, Denli et al. 2004, Lee et al. 2004) to release the embedded miRNAs. Mature miRNAs are then loaded into Argonaute1 (Ago1) within the RNA-induced silencing complex (RISC) (Hammond et al. 2001, Okamura et al. 2004). Intriguingly, we found that Drosophila miR-34 displays multiple isoforms that differ at the 3′ end, suggesting a novel biogenesis mechanism involving 3′ end processing. To

define the cellular factors responsible, we performed an RNA interference (RNAi) screen and identified a putative 3'→5' exoribonuclease CG9247/nibbler essential for the generation of the smaller isoforms of miR-34. Nibbler (Nbr) interacts with Ago1 and processes miR-34 within RISC. Deep sequencing analysis revealed a larger set of multi-isoform miRNAs that are controlled by nibbler. These findings suggest that Nbr-mediated 3′ end processing represents a critical step in miRNA maturation that impacts miRNA diversity.

Results and Discussion

Although miRNAs are typically annotated and observed as a single species, we found that miR-34 showed a pattern of three major isoforms of 24, 22, and 21 nucleotides (nts) in northern blots from adult Drosophila (Figure 3.1A). Deep sequencing analysis (Zhou et al. 2009) also showed that miR-34 is present in multiple forms that all bear the same 5′ terminus but differ at their 3′ ends, presenting a nested series (Figure 3.1B). To assess the relationship among these, we designed a pulse-chase experiment to follow miR-34 biogenesis. Heat-shock driven primary miR-34 was tightly induced for 30 min and then monitored over time in adult flies. The longest isoform, isoform a (24 nt), was predominant initially, whereas the accumulation of the shorter isoforms was delayed, but then increased over time (Figure 3.1C). Moreover, as the 21 nt isoform accumulated, the 24 nt form was lost in a seemingly reciprocal manner, suggesting that the 24-mer may be converted into the 21-mer.

To define the mechanism, we treated cells with double-stranded RNA (dsRNA) targeting specific genes within the small RNA biogenesis pathways and assessed the miR- 34 pattern by northern blot. Imprecise cleavage of the precursor transcript could result in the production of the multiple forms. However, reduction of either Drosha or Dcr-1, or their binding partners Pasha and Loquacious, or Dicer-2 (Dcr-2), responsible for small interfering RNA (siRNA) generation, did not alter the pattern (Figure 3.2A). Therefore, we reasoned that the smaller isoforms may instead be generated by an exonuclease that sequentially processes the longest isoform into the nested series observed. To test this hypothesis, we performed an RNAi screen against the predicted 3'→5' exonucleases 41 in Drosophila, including components of the RNA exosome (see Table 2.1). This identified one gene, CG9247 (which we named nibbler/nbr), with a striking effect: depletion of nbr led to a dramatic accumulation of the miR-34 large isoform with a concomitant loss of the shorter isoforms (Figure 3.2B; Figure 3.5A,B). In contrast, loss of nbr did not appear to alter the sizes or levels of miRNAs that normally show a single isoform by northern blot, such as miR-14 and miR-277 (Figure 3.2C). We also examined whether nbr knockdown had an effect on endogenous siRNAs but saw no impact on esi-2.1 (Figure 3.2C). These data suggested that the novel putative exoribonuclease Nbr is required to generate the shorter isoforms of the multi-isoform miRNA miR-34 but is not required for general small RNA biogenesis.

The Nbr exoribonuclease domain shows closest sequence homology to human EXD3, falling within the E. coli RNase D protein family; this includes the Werner exoribonuclease and C. elegans Mut-7 involved in transposon silencing (Figure 3.5D;

(Ketting et al. 1999)). Nbr, however, showed no predicted RNA binding domain, suggesting that it may function with a partner with RNA binding capacity, to bring Nbr activity to RNA substrates. To define these, we then performed a second RNAi screen for genes known to bind RNA or associate with small RNA silencing pathways, including the two somatic RISC-associated Argonautes (Table 3.1). Strikingly, loss of Ago1 phenocopied nbr depletion: accumulation of the 24 nt isoform occurred, with reduction of the shorter isoforms (Figure 3.3A). Controls indicated that knockdown of Ago1 had no effect on nbr expression, and nbr knockdown had no effect on Ago1 expression. These data suggested that Ago1 is also required for trimming and that Nbr and Ago1 may act in a complex. Coimmunoprecipitation (coIP) studies indicated that hemagglutinin (HA)-tagged Nbr associates with Flag-tagged Ago1, but not with a control protein (Flag-Ran) (Figure 3.3B). RNase treatment indicated that the association was not RNA dependent (Figure 3.6). Proteomic studies have identified both Ago1 and Nbr as small RNA associated proteins (Gerbasi et al. 2010), underscoring the specificity of the interaction. Because Nbr associates with Ago1, we hypothesized that miR-34 3′ end processing may occur in the context of RISC. Indeed, immunoprecipitation of Ago1 revealed that all miR- 34 isoforms were bound (Figure 3.3C). Furthermore, when Nbr was depleted, the longest miR-34 isoform remained bound to Ago1 (Figure 3.3C). Altogether, these data suggest that the 24 nt miR-34 isoform is first generated by Dcr-1 then loaded into RISC. Next, Nbr, in association with Ago1, processes the long 24 nt isoform into shorter isoforms that remain loaded in RISC.

To assess the in vivo role of Nbr, we analyzed the expression and function of nbr in flies. Northern blots revealed that nbr is expressed during development and in the adult, with peaks during the late larval, early pupal stage and in adults (data not shown). Analysis of nbr messenger RNA (mRNA) levels in animals with a transposon insertion in the coding region (nbrf02257) showed that homozygous mutants (nbr−/−) lacked nbr expression (Figure 3.4A,B). nbr−/− flies were semilethal and sterile, indicating that nbr function is critical. Given the homology to Mut-7, we examined levels of transposons but found no evidence linking nbr to transposon silencing (data not shown). Assessment of miR-34 expression in nbr−/− flies phenocopied cells treated with dsRNA: the shorter isoforms were abolished, whereas the 24 nt form accumulated (Figure 3.4C). As in cells, there was no striking effect on single-form miRNAs like miR-277 (Figure 3.4D). Furthermore, miR- 34* levels and isoform distribution appeared unaffected (Figure 3.5C). These data indicated that nbr modulates the isoform abundance of miR-34 in the animal in vivo.

To assess the broader impact of Nbr function, we screened 65 miRNAs by northern blot of RNA isolated from cultured cells and flies. We identified nine additional miRNAs with multiple isoforms: mir-2, miR-3, miR-12, miR-79, miR-263a/b, miR-274, miR- 279, miR-281-1/2, and miR-305. The expression patterns of five of these were altered in nbr−/− mutants, exhibiting accumulation of the longest isoforms with concomitant loss of the shorter isoforms (Figure 3.4E; miR-2 family was not studied further as a result of cross hybridization between members). Analysis of small RNA profiling data from cells (Zhou et al. 2009) confirmed that two of these (miR-263a and miR-305) had significant levels of multiple forms that differed at the 3' end (Table 3.1); miR-3, miR-12, miR-281, and miR-274 levels were too low for analysis. Three multiple-isoform miRNAs (miR- 79,

miR-274, and miR-279) did not show an altered pattern in nbr−/− flies (Figure 3.4E). The deep sequencing data set revealed that miR-279 displays a series of isoforms that do differ at the 3' end; because miR-279 processing is nbr-independent, nbr may be one member of a larger set of genes or mechanisms responsible for 3' end diversity. miR-79 isoforms differed at the 5' end, suggesting that mechanisms also exist for 5' end diversity of miRNAs.

We further investigated the extent to which trimming is involved in miRNA processing by deep sequencing the small RNAs from flies, comparing nbr mutants to controls. There was no major impact on the size distribution of small RNAs as a whole or miRNAs in particular (Figure 3.7A,B). To more carefully assess isoforms, we mapped reads to the miRNA stem-loop sequences and analyzed for length. For each miRNA, we calculated a ratio of the most frequent length in wild-type to the sum of all other lengths and compared this ratio between nbr and control. The distribution of the length ratios highlighted a cohort of miRNAs with extreme differences between nbr−/− and control. At the two ends of the plot were miRNAs where the most common length isoform of the miRNA was present at a much higher or much lower level in nbr−/−than in wild-type, reflecting an altered pattern of isoform distribution or relative abundance for these miRNAs in the absence of nbr. These included miRNAs we had defined as trimmed and modulated by nbr (miR-34, miR-263a, miR-263b), along with additional candidates (Figure 3.4F, red boxes). Northern blots were performed on the top and bottom eight miRNAs that we had not tested; we confirmed seven new nbr-dependent miRNAs (miR-7, miR-10, miR-11, miR-31b, miR-100, miR-190, miR-317; (Figure 3.7; Table 3.3 and

Table 3.3). Northern blotting revealed some miRNAs that were trimmed were not detected as so by deep sequencing and the reverse: for any given miRNA, the extent of trimming had to be greater than ~10% in isoform level to detect a consistent change by northern blot, whereas deep sequencing analysis suggested that not all isoforms were cloned with equal efficiency.

Trimming exerts a profound and diverse impact on miRNA sequence profiles: nbr promotes the diversity of some miRNAs (miR-34, miR-7, miR-317) and alters the relative abundance of the most prominent isoform of others (miR-190 and miR-10; (Figure 3.7C and Figure 3.7B; Table 3.3). To identify potential Nbr-dependent miRNA targets, we performed transcriptional profiling of Nbr-deficient cells. This would allow identification of target mRNAs whose stability was altered by miRNA trimming, but not targets primarily controlled by translational repression (Bagga et al. 2005). This identified 12 genes whose levels were affected by nbr depletion by >1.5-fold (Figure 3.4G; Table 3.4); of these, one was reduced (nbr) and the others were upregulated. Assessing the levels of eight of these by real-time PCR confirmed increased expression of 6/8 mRNAs (75%) in nbr-depleted cells (Table 3.4). Next, we assessed expression of nine of these genes in nbr−/− flies, compared to wild-type andloquacious mutant flies. loqsf00791 mutants are viable and show deficiency in miRNA maturation and function, thus allowing assessment of miRNA function in adults (Jiang et al. 2005).We reasoned that genes regulated by miRNAs that are impacted by nbr-processing would also show dependence on loqs. We validated 5/9 genes (55%) as upregulated in both nbr−/− and loqsf00791 (two additional genes were upregulated, although did not reach statistical

significance in nbr−/−) (Figure 3.4H; Table 3.4). Sequence analysis of these mRNA targets revealed that 4/7 genes (57%) have potential sites for the miRNAs that showed nbr-dependent processing (Table 3.4). It is unclear, however, whether existing algorithms for miRNA targeting efficiently predict binding sites for miRNAs with 3' end diversity; targets for trimmed miRNAs may use noncanonical recognition motifs that are more dependent on 3' end pairing than seed complementarity.

These data provide evidence for a novel step in miRNA biogenesis: miRNA 3' end terminal trimming mediated by the 3'→5' exoribonuclease Nbr. Notably, small RNA deep sequencing has unveiled a rich pattern of miRNA sequence isoforms, although miRNAs have routinely been annotated as a single mature form. Our findings suggest that miRNA processing by Nbr alters the repertoire of at least a subset of miRNAs in cells and whole animals, contributing to the diversity of the small RNA profile and potentially impacting posttranscriptional gene regulation in Drosophila. Mechanistically, our data indicate that, upon nbr knockdown, miR-34 is still associated with RISC; thus, trimming is not a prerequisite to miR-34 loading and likely occurs after loading.

The impact and biological consequences of trimming may be complex. Nbr may impact strand selection within RISC because strand selection is influenced by the extent of 3' overhang and degree of pairing for any miRNA-miRNA* duplex (Khvorova, Reynolds, and Jayasena 2003, Schwarz et al. 2003). Nbr may impact miRNA stability, because previous studies have demonstrated that tailing and trimming of mature Drosophila miRNAs influence their turnover (Ameres et al. 2010). Trimming may also impact

mRNA silencing by favoring alterative miRNA sites within mRNA targets. Although canonical miRNA-target specificity is thought to be driven largely by complementarity within the seed, noncanonical interactions can depend more heavily on 3' compensatory sites (Bartel 2009a, Brennecke et al. 2005). Therefore, differences in the length of the 3' end of miRNAs may influence both target selection and silencing efficiency of targets that require extensive 3' end pairing. Future analysis of trimmed miRNAs and their range of targets will reveal rules governing miRNA-mRNA pairing specificity that may be impacted by 3' end heterogeneity. Given that some mammalian miRNAs also display multiple isoforms (Cheloufi et al. 2010, Cifuentes et al. 2010), miRNA 3' end processing may be conserved.

Our studies focused on the role of nbr in miRNA pathway function; whether nbr plays a role in additional small RNA pathways remains an open question, although we did not observe effects on transposons, suggesting that it does not globally impact endogenous small RNA pathways. The modification of mature miRNAs and their precursors is an emerging facet of miRNA-mediated gene regulation (Berezikov et al. 2011). Nbr may represent a central player in a larger spectrum of factors that shape miRNA repertoire and miRNA function, through the generation of multi-isoform miRNAs.

Figures and Tables

Table 3.1 miRNA Reads from Deep Sequencing Data

		Read count	Percentage
miR-263a			
AAUGGCACUGGAAGAAUUCACGGG	24	552	56.3%
AAUGGCACUGGAAGAAUUCACGG	23	170	17.3%
AAUGGCACUGGAAGAAUUCACG	22	97	9.9%
AAUGGCACUGGAAGAAUUCAC	21	112	11.4%
AAUGGCACUGGAAGAAUUCA	20	36	3.7%
AAUGGCACUGGAAGAAUUC	19	14	1.4%
		99.6% of reads	
miR-279			
UGACUAGAUCCACACUCAUUAAU	23	5	0.1%
UGACUAGAUCCACACUCAUUAA	22	785	18.8%
UGACUAGAUCCACACUCAUUA	21	757	18.2%
UGACUAGAUCCACACUCAUU	20	1858	44.6%
UGACUAGAUCCACACUCAU	19	758	18.2%
UGACUAGAUCCACACUCA	18	6	0.1%
		99.1% of reads	
miR-305			
AUUGUACUUCAUCAGGUGCUCUGGU	25	8	0.4%
AUUGUACUUCAUCAGGUGCUCUGG	24	225	10.9%
AUUGUACUUCAUCAGGUGCUCUG	23	63	3.0%
AUUGUACUUCAUCAGGUGCUCU	22	295	14.2%
AUUGUACUUCAUCAGGUGCUC	21	1110	53.6%
AUUGUACUUCAUCAGGUGCU	20	357	17.2%
AUUGUACUUCAUCAGGUGC	19	12	0.6%
AUUGUACUUCAUCAGGUG	18	2	0.1%
		99.6% of reads	
miR-79			
AUAAAGCUAGAUUACCAAAGCAU	23	885	17.7%
UAAAGCUAGAUUACCAAAGCAU	22	4044	81.0%
AAAGCUAGAUUACCAAAGCAU	21	66	1.3%
		98.1% of reads	

Table 3.2 Sequence Counts of miRNAs from Ends of the Ratio Plot in Control (WT) and Nbr Mutants

The nine miRNAs from end of the ratio plot with smallest ratios					
miR-7	len	WT	WT specific % of reads	Nbr	Nbr specific % of

					reads
UGGAAGACUAGUGAUUUUGUUGUUU	25	4	0	7	1
UGGAAGACUAGUGAUUUUGUUGUU	24	321	16	513	49
UGGAAGACUAGUGAUUUUGUUGU	23	432	21	339	32
UGGAAGACUAGUGAUUUUGUUG	22	902	45	151	14
UGGAAGACUAGUGAUUUUGUU	21	255	13	25	2
UGGAAGACUAGUGAUUUUGU	20	94	5	12	1
UGGAAGACUAGUGAUUUUG	19	13	1	3	0
		80.9% of reads	75% of reads		
miR-11	length	WT norm	WT specific % of reads	Nbr norm	Nbr specific % of reads
CAUCACAGUCUGAGUUCUUGCU	22	406	19	255	52
CAUCACAGUCUGAGUUCUUGC	21	1526	72	190	38
CAUCACAGUCUGAGUUCUUG	20	153	7	41	8
CAUCACAGUCUGAGUUCUU	19	15	1	3	1
CAUCACAGUCUGAGUUCU	18	9	0	3	1
		97% of reads	95% of reads		
miR-317	length	WT norm	WT specific % of reads	Nbr norm	Nbr specific % of reads
UGAACACAGCUGGUGGUAUCCAGU	24	64	2	123	17
UGAACACAGCUGGUGGUAUCCAG	23	104	3	161	22
UGAACACAGCUGGUGGUAUCCA	22	162	5	120	17
UGAACACAGCUGGUGGUAUCC	21	1564	49	179	25
UGAACACAGCUGGUGGUAUC	20	989	31	93	13
UGAACACAGCUGGUGGUAU	19	153	5	18	2
UGAACACAGCUGGUGGUA	18	103	3	20	3
UGAACACAGCUGGUGGU	17	33	1	5	1
		97% of reads	92% of reads		
miR-283	length	WT norm	WT specific % of reads	Nbr norm	Nbr specific % of reads
AAAUAUCAGCUGGUAAUUCUGGG	23	171	23	170	39
AAAUAUCAGCUGGUAAUUCUGG	22	83	11	74	17
AAAUAUCAGCUGGUAAUUCUG	21	466	63	182	42
AAAUAUCAGCUGGUAAUUCU	20	12	2	5	1
AAAUAUCAGCUGGUAAUUC	19	4	0	0	0
		96% of reads	97% of reads		
miR-210	lengt	WT norm	WT	Nbr	Nbr

	h		specific % of reads	norm	specific % of reads
CUUGUGCGUGUGACAGCGGCUAUU	24	15	5	18	11
CUUGUGCGUGUGACAGCGGCUAU	23	86	29	86	52
CUUGUGCGUGUGACAGCGGCUA	22	26	9	20	12
CUUGUGCGUGUGACAGCGGCU	21	121	40	29	17
CUUGUGCGUGUGACAGCGGC	20	47	16	8	5
CUUGUGCGUGUGACAGCGG	19	5	2	4	2
		56% of reads	56% of reads		
UUGUGCGUGUGACAGCGGCUAUU	23	17	7	12	9
UUGUGCGUGUGACAGCGGCUAU	22	99	42	67	53
UUGUGCGUGUGACAGCGGCUA	21	30	13	17	13
UUGUGCGUGUGACAGCGGCU	20	72	31	21	16
UUGUGCGUGUGACAGCGGC	19	15	6	9	7
UUGUGCGUGUGACAGCGG	18	2	1	0	1
		44% of reads	43% of reads		
miR-307	length	WT norm	WT specific % of reads	Nbr norm	Nbr specific % of reads
UCACAACCUCCUUGAGUGAGCGA	23	2	3	6	9
UCACAACCUCCUUGAGUGAGCG	22	6	11	14	21
UCACAACCUCCUUGAGUGAGC	21	47	81	43	65
UCACAACCUCCUUGAGUGAG	20	2	4	2	4
UCACAACCUCCUUGAGUGA	19	0	1	0	1
		90% of reads	92% of reads		
miR-263b	length	WT norm	WT specific % of reads	Nbr norm	Nbr specific % of reads
CUUGGCACUGGGAGAAUUCACAGU	24	3	0	4	0
CUUGGCACUGGGAGAAUUCACAG	23	233	8	151	15
CUUGGCACUGGGAGAAUUCACA	22	492	17	337	33
CUUGGCACUGGGAGAAUUCAC	21	1557	54	402	39
CUUGGCACUGGGAGAAUUCA	20	487	17	106	10
CUUGGCACUGGGAGAAUUC	19	93	3	22	2
CUUGGCACUGGGAGAAUU	18	6	0	1	0
CUUGGCACUGGGAGAAU	17	2	0	0	0
		99% of reads	98% of reads		
miR-315	length	WT norm	WT specific % of reads	Nbr norm	Nbr specific

					% of reads
UUUUGAUUGUUGCUCAGAAAGCC	23	71	27	75	40
UUUUGAUUGUUGCUCAGAAAGC	22	187	70	104	56
UUUUGAUUGUUGCUCAGAAAG	21	6	2	4	2
UUUUGAUUGUUGCUCAGAAA	20	3	1	1	1
		99% of reads	98% of reads		
The nine miRNAs from end of ratio plot with largest ratios					
miR-190	length	WT norm	WT specific % of reads	Nbr norm	Nbr specific % of reads
AGAUAUGUUUGAUAUUCUUGGUUGUU	26	0	0	0	1
AGAUAUGUUUGAUAUUCUUGGUUGU	25	9	5	13	9
AGAUAUGUUUGAUAUUCUUGGUUG	24	49	26	78	55
AGAUAUGUUUGAUAUUCUUGGUU	23	42	23	29	20
AGAUAUGUUUGAUAUUCUUGGU	22	31	17	12	8
AGAUAUGUUUGAUAUUCUUGG	21	38	20	7	5
AGAUAUGUUUGAUAUUCUUG	20	14	8	3	2
AGAUAUGUUUGAUAUUCUU	19	2	1	0	
		97% of reads	99% of reads		
miR-34	length	WT norm	WT specific % of reads	Nbr norm	Nbr specific % of reads
UGGCAGUGUGGUUAGCUGGUUGUGUA	26	0	0	2	0
UGGCAGUGUGGUUAGCUGGUUGUGU	25	13	1	102	8
UGGCAGUGUGGUUAGCUGGUUGUG	24	807	52	1004	76
UGGCAGUGUGGUUAGCUGGUUGU	23	79	5	111	8
UGGCAGUGUGGUUAGCUGGUUG	22	157	10	60	5
UGGCAGUGUGGUUAGCUGGUU	21	304	20	11	1
UGGCAGUGUGGUUAGCUGGU	20	93	6	14	1
UGGCAGUGUGGUUAGCUGG	19	82	5	11	1
UGGCAGUGUGGUUAGCUG	18	12	1	3	0
UGGCAGUGUGGUUAGCU	17	1	0	0	0
UGGCAGUGUGGUUAGC	16	3	0	2	0
		91% of reads	93% of reads		
miR-10	length	WT norm	WT specific % of reads	Nbr norm	Nbr specific % of reads
CAAAUUCGGUUCUAGAGAGGUUUG	24	0	0	0	0
CAAAUUCGGUUCUAGAGAGGUUU	23	873	46	479	64

CAAAUUCGGUUCUAGAGAGGUU	22	714	38	190	26
CAAAUUCGGUUCUAGAGAGGU	21	181	10	44	6
CAAAUUCGGUUCUAGAGAGG	20	87	5	19	3
CAAAUUCGGUUCUAGAGAG	19	7	0	3	0
CAAAUUCGGUUCUAGAGA	18	8	0	4	1
CAAAUUCGGUUCUAGAG	17	1	0	0	0
CAAAUUCGGUUCUAGA	16	8	0	4	0
		99% of reads	99% of reads		
miR-100	length	WT norm	WT specific % of reads	Nbr norm	Nbr specific % of reads
AACCCGUAAAUCCGAACUUGUG	22	110	86	133	93
AACCCGUAAAUCCGAACUUGU	21	14	11	7	5
AACCCGUAAAUCCGAACUUG	20	3	3	2	2
		97% of reads	97% of reads		
miR-1010	length	WT norm		Nbr norm	Nbr specific % of reads
UUUCACCUAUCGUUCCAUUUGCAG	24	38	76	22	86
UUUCACCUAUCGUUCCAUUUGCA	23	2	4	2	7
UUUCACCUAUCGUUCCAUUUGC	22	9	18	1	5
UUUCACCUAUCGUUCCAUUUG	21	0	1	0	2
		88% of reads	87% of reads		
miR-263a	length	WT norm	WT specific % of reads	Nbr norm	Nbr specific % of reads
AAUGGCACUGGAAGAAUUCACGGGG	25	1	0	4	0
AAUGGCACUGGAAGAAUUCACGGG	24	20494	55	18394	66
AAUGGCACUGGAAGAAUUCACGG	23	10116	27	8037	29
AAUGGCACUGGAAGAAUUCACG	22	3898	11	1225	4
AAUGGCACUGGAAGAAUUCAC	21	2044	6	237	1
AAUGGCACUGGAAGAAUUCA	20	340	1	35	0
AAUGGCACUGGAAGAAUUC	19	129	0	14	0
AAUGGCACUGGAAGAAUU	18	27	0	5	0
AAUGGCACUGGAAGAAU	17	9	0	3	0
AAUGGCACUGGAAGAA	16	4	0	1	0
		100% of	100% of		

		reads	reads		
miR-989	length	WT norm	WT specific % of reads	Nbr norm	Nbr specific % of reads
UGUGAUGUGACGUAGUGGAACAU	23	27	0	19	0
UGUGAUGUGACGUAGUGGAACA	22	4256	32	3196	22
UGUGAUGUGACGUAGUGGAAC	21	8600	64	10278	72
UGUGAUGUGACGUAGUGGAA	20	294	2	406	3
UGUGAUGUGACGUAGUGGA	19	263	2	408	3
UGUGAUGUGACGUAGUGG	18	8	0	14	0
UGUGAUGUGACGUAGUG	17	9	0	6	0
UGUGAUGUGACGUAGU	16	1	0	2	0
		99% of reads	99% of reads		
miR-31b	length	WT norm	WT specific % of reads	Nbr norm	Nbr specific % of reads
UGGCAAGAUGUCGGAAUAGCUGA	23	555	35	347	30
UGGCAAGAUGUCGGAAUAGCUG	22	784	49	674	58
UGGCAAGAUGUCGGAAUAGCU	21	182	11	99	8
UGGCAAGAUGUCGGAAUAGC	20	62	4	38	3
UGGCAAGAUGUCGGAAUAG	19	13	1	7	1
		97% of reads	98% of reads		

Table 3.3. Summary of Results on miRNAs from the Two Ends of the Length Ratio Plot, Related to Figure 3.4.

miRNA	Ratio (nbr-/-/control)	Northern result
The eight miRNAs from end of plot with smallest ratios		
miR-7	0.2148	confirmed by Northern
miR-11	0.2497	confirmed by Northern
miR-317	0.3410	confirmed by Northern
miR-283	0.4284	not detectable by Northern
miR-210	0.4744	5 and 3' end processed; cannot distinguish forms by Northern
miR-307	0.4955	false positive
miR-263b	0.5505	original multiform/altered by nbr
miR-315	0.5585	not detectable by Northern
The eight miRNAs from end of the plot with largest ratios		
miR-190	3.4083	confirmed by Northern
miR-34	2.8973	original multiform/altered in nbr
miR-10	2.0810	confirmed by Northern
miR-100	1.9830	confirmed by Northern
miR-1010	1.9011	false positive
miR-263a	1.5505	original multiform/altered in nbr
miR-989	1.4267	false positive
miR-31b	1.4059	confirmed by Northern

Table 3.4. Genes Identified by Microarray that Changed in Cells upon Nbr Knockdown, related to Figure 3.4. List of all of the genes with a change of 1.5-fold or greater in either direction with a 20% false discovery rate or less (see Figure 3.4G).

Probeset ID	Gene Symbol	Nbr-dependent miRNAs predicted to target the mRNA[1]	Fold-change by microarray of DL1 cells (Nbr dsRNA vs. control)	Change in DL1 cells by realtime PCR (Nbr dsRNA vs control)*	Change in adult flies by realtime PCR ($nbr^{-/-}$ mutant animals vs control)*
1634849_at	CG9247/nbr		-8.9181	down	down
1634823_at	mRpS25		2.06252	2.3 ($p<0.001$)	— (not upregulated)
1630515_s_at	Glt		1.95666	(not done)	(not done)
1639500_at	nub	miR-2a,b,c, miR-11, miR-10	1.51219	1.4	5.2 ($p<0.001$; Fig 4G)
1627962_at	Fas3	miR-3, miR305, miR-210	1.74339	1.2	3 ($p<0.001$; Fig 4G)
1637358_at	CG10232	miR-281-1, miR-263a, miR-283	1.87433	— (not upregulated)	2.8 ($p<0.01$; Fig 4G)
1628666_at	CG3328		1.69937	1.2	2.2 ($p<0.05$; Fig 4G)
1629903_at	CG30359		1.54676	(not done)	(not done)
1634569_at	CG34051		1.56768	(not done)	2 ($p>0.05$)
1630710_at	GluRIIA		1.60267	1.3	2.6 ($p<0.001$; Fig4G)
1627637_at	d	miR-263b, miR-10, miR-7	1.53422	1.4	3.2 ($p>0.05$)

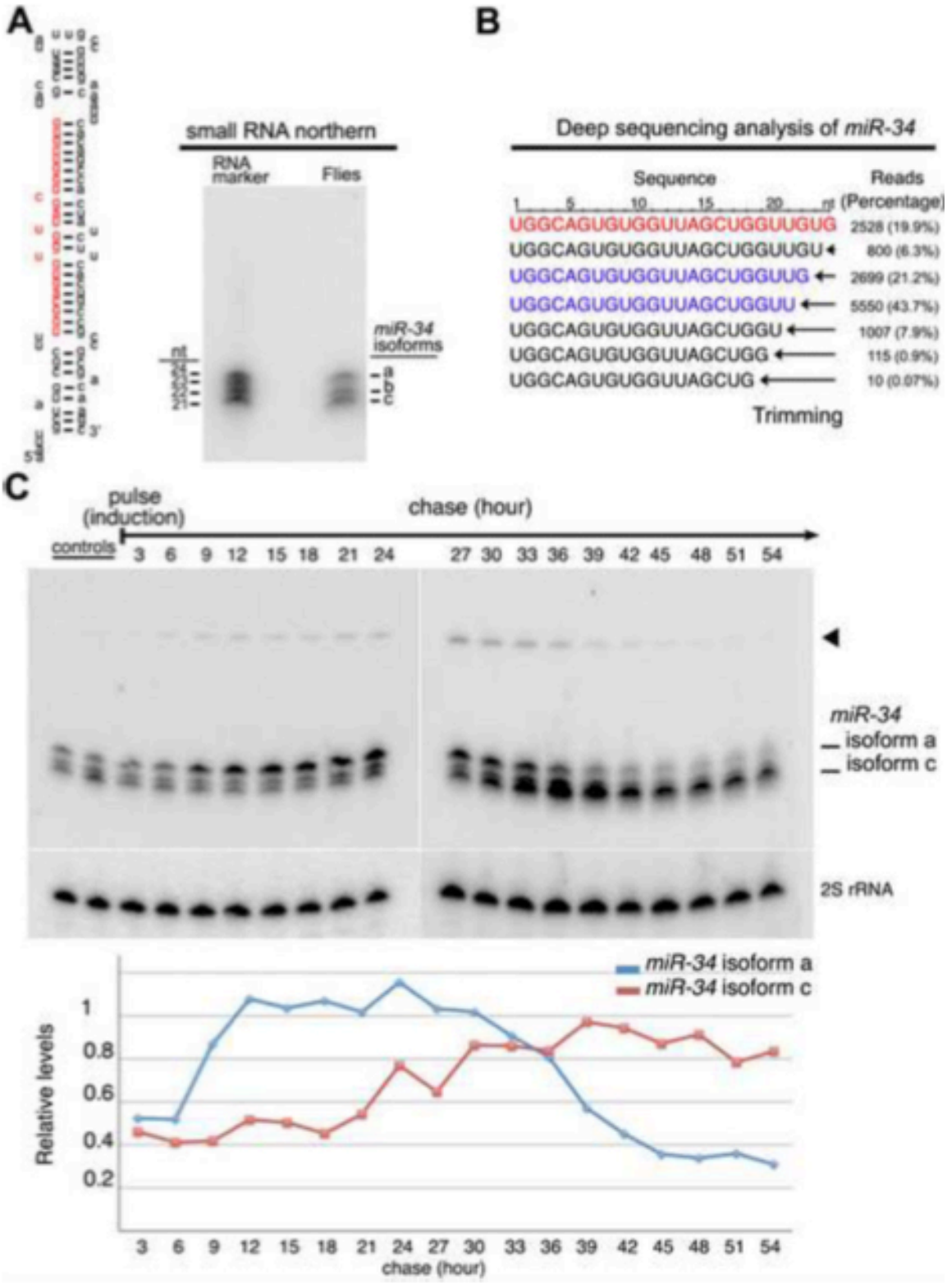

A
small RNA northern
RNA marker
Flies
miR-34 isoforms
B
Deep sequencing analysis of miR-34
Sequence
Reads
(Percentage)
UGGCAGUGUGGUUAGCUGGUUGUG 2528 (19.9%)
UGGCAGUGUGGUUAGCUGGUUGU 800 (6.3%)
UGGCAGUGUGGUUAGCUGGUUG 2699 (21.2%)
UGGCAGUGUGGUUAGCUGGUU 5550 (43.7%)
UGGCAGUGUGGUUAGCUGGU 1007 (7.9%)
UGGCAGUGUGGUUAGCUGG 115 (0.9%)
UGGCAGUGUGGUUAGCUG 10 (0.07%)
Trimming
C
pulse (induction)
chase (hour)
controls
3 6 9 12 15 18 21 24 27 30 33 36 39 42 45 48 51 54
miR-34
isoform a
isoform c
2S rRNA
miR-34 isoform a
miR-34 isoform c
Relative levels
1
0.8
0.6
0.4
0.2
chase (hour)

Figure 3.1. Drosophila miR-34 Shows Multiple Isoforms Whose Generation Appears Dependent on 3' End Trimming. (A) miR-34 has multiple forms in adult flies. Left is the miR-34 precursor, with the mature 24 nucleotide (nt) sequence in red, and right is a northern blot for miR-34. Isoforms of 24, 22, and 21 nt are labeled a, b, and c, respectively. (B) miR-34 isoforms from a deep sequencing Drosophila S2 cell data set (Zhou et al. 2009). In red is the 24 nt isoform a, and in blue are isoforms b and c. These reads are 99.1% of the total miR-34 reads. (C) Northern blot analysis of miR-34 isoform accumulation in vivo. Transient induction of pri-miR-34 by hs-GAL4 in adult flies leads to initial accumulation of isoform a, which is lost over time while the shorter isoforms accumulate. Arrowhead notes pre- miR-34. (D) Quantification of miR-34 isoforms from pulse-chase in (C). Values normalized to 2S ribosomal RNA.

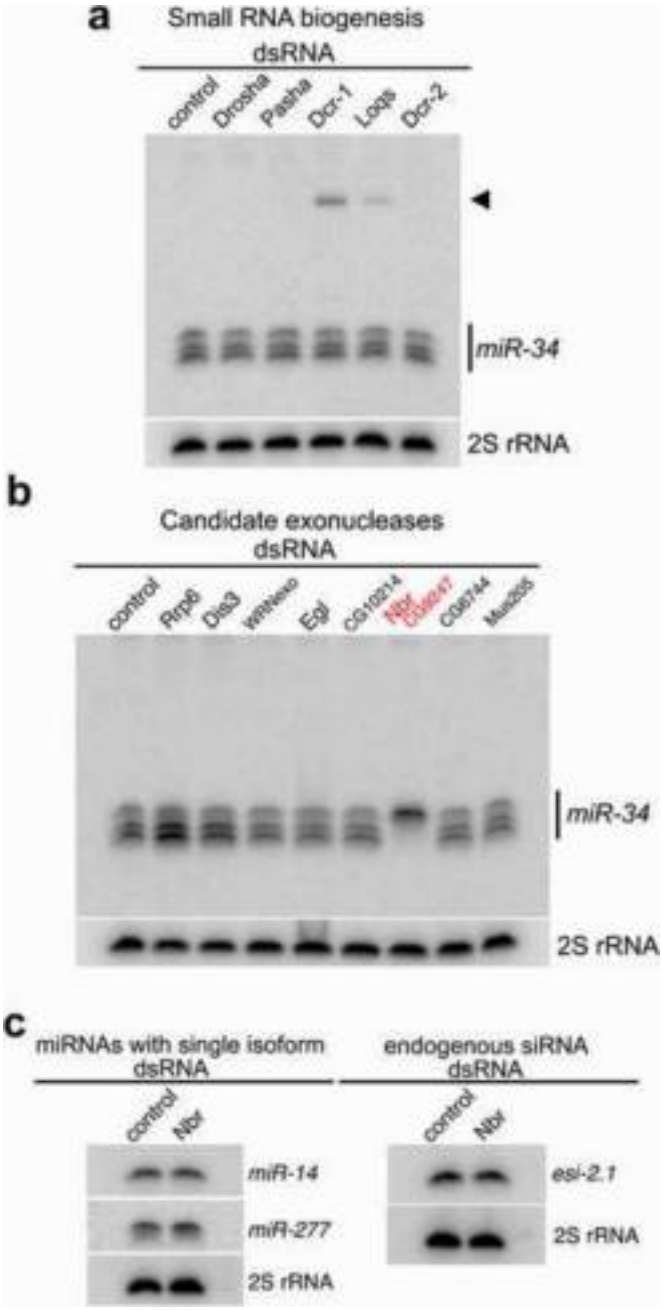


Figure 3.2. nbr Is Required to Generate the Isoforms of miR-34. (A) Depletion of known factors in the small RNA biogenesis pathways has no effect on miR-34. (B) Depletion of candidate exoribonucleases shows that loss of CG9247/Nbr (red) leads to an accumulation of the 24 nt isoform, with dramatic reduction of the shorter isoforms. (C)

Cells depleted of Nbr are not altered in single isoform microRNAs (miRNAs) or endogenous small interfering RNA (siRNA) esi-2.1.

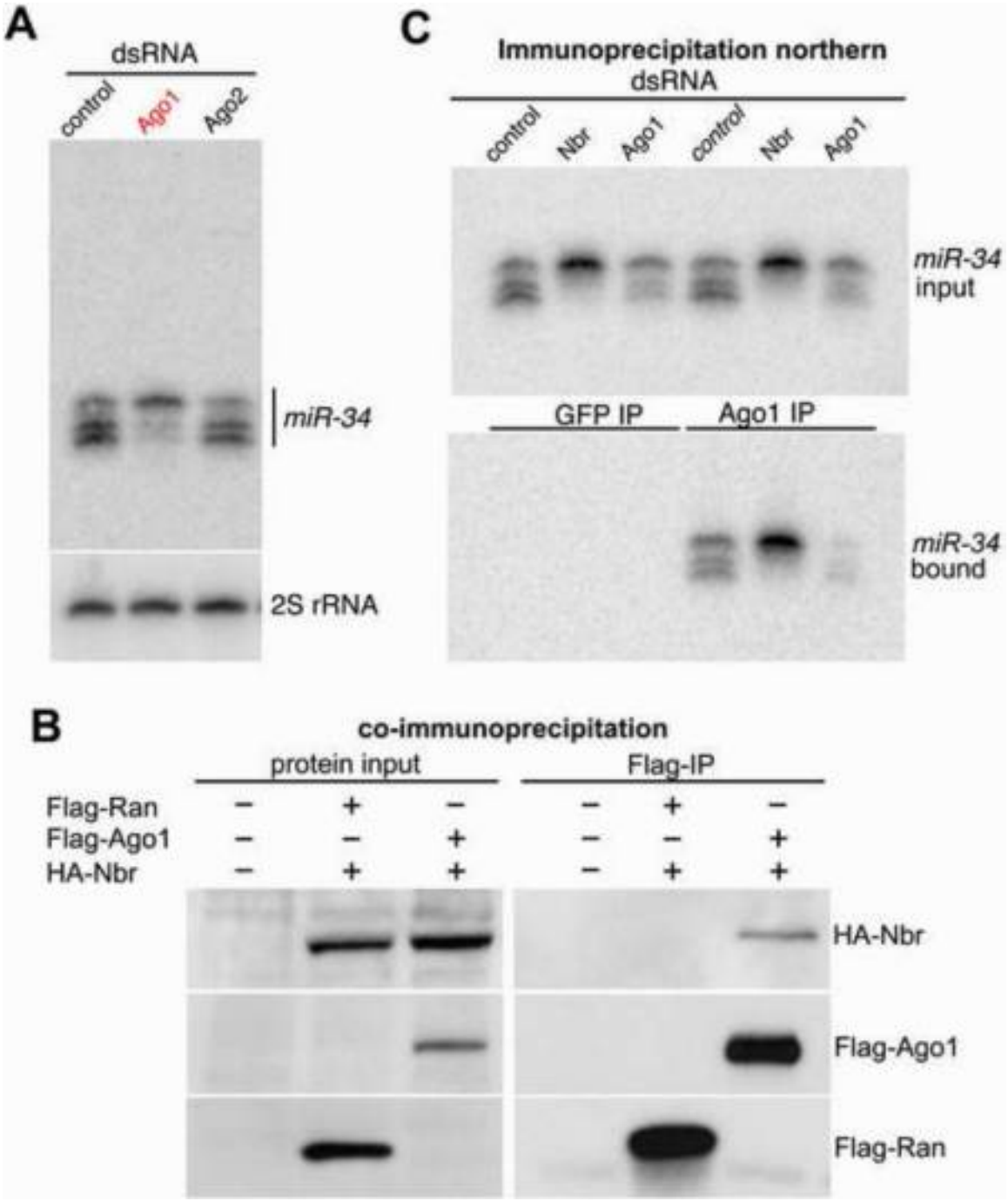


Figure 3.3. Nbr Interacts with Ago1-RNA-Induced Silencing Complex.

(A) Small RNA northern blot analysis of mir-34 isoforms. Depletion of Ago1 phenocopies Nbr knockdown. (B) Ago1 and Nbr interact by coimmunoprecipitation (coIP). Cells were untreated or transfected with hemagglutinin (HA)-Nbr and Flag-Ago1 or Flag-Ran (control). Following immunoprecipitation (IP), interacting proteins were probed by immunoblot. Input is 10% of Flag-IP. (C) All miR-34 isoforms coimmunoprecipitated with Ago1. Cells were treated with double-stranded (dsRNA) to control (LacZ), Nbr, or Ago1, and IPs were performed with anti-GFP (control) or Ago1 antibodies. Input and coimmunoprecipitated RNA were analyzed by northern blotting for miR-34.

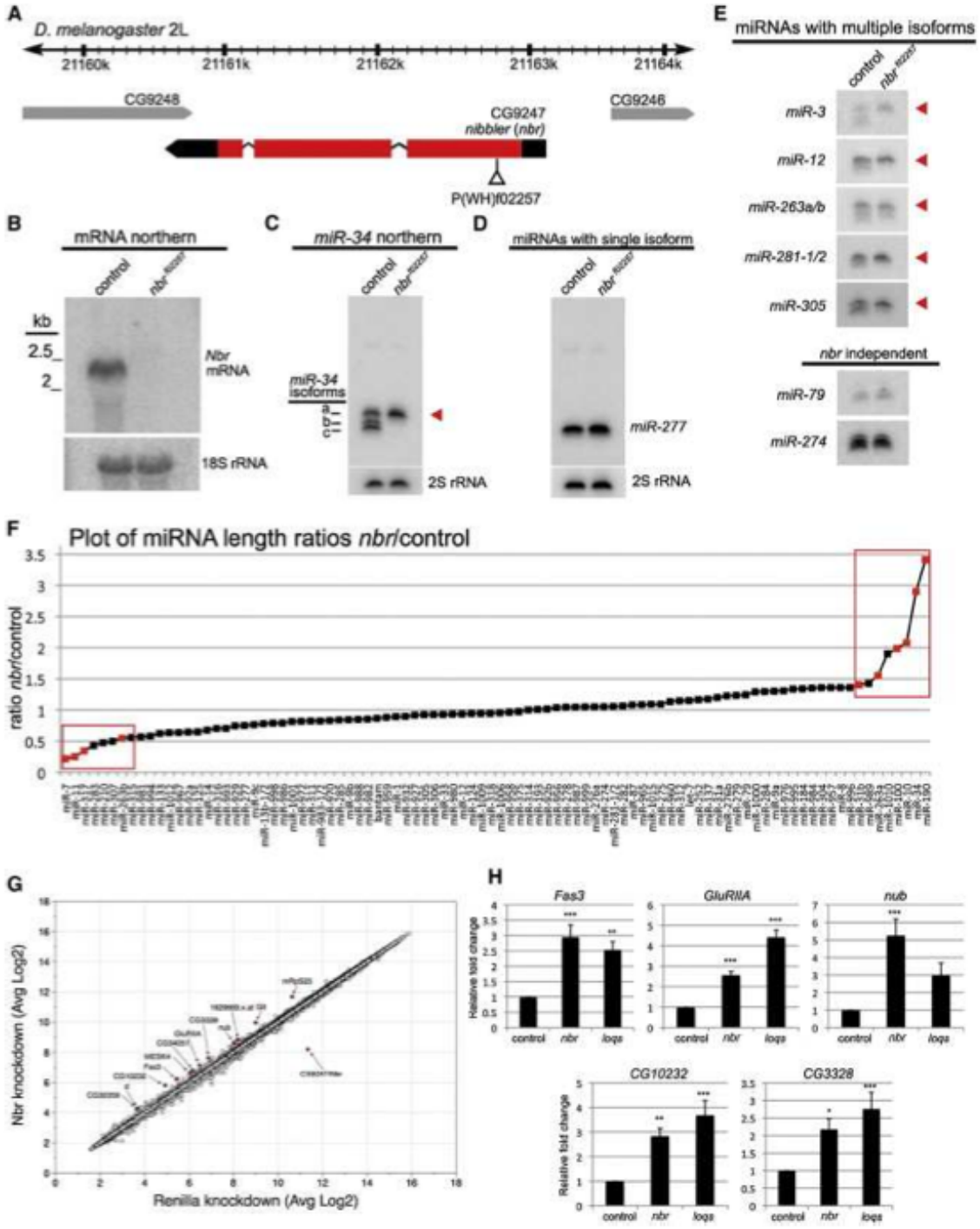


Figure 3.4. nbr Is Required In Vivo to Process Select miRNAs and Silence Target Messenger RNAs. (A) Genomic map of the nbr locus. Coding region is shown in red, with transposon insertion highlighted. (B) Northern blot for nbr. The nbrf02257 mutant shows complete messenger RNA (mRNA) loss. (C) Shorter isoforms of miR-34 are

abolished inthe nbr mutant. Arrowhead notes isoform a. (D) Northern blot of single-isoform miR-277, which is not altered in nbr−/−. (E) Comparison of multiple-isoform miRNAs from control and nbrf02257 flies. Some miRNA isoforms require nbr (red arrowheads), whereas others are nbr independent. (F) The ratio of the most frequent form of the miRNA in wild-type, compared to the sum of all other forms, was generated for nbr and control. The ratios were compared (nbr ratio/control ratio) and plotted. The ratio was excessively high or low when isoform biogenesis is defective. Red boxes highlight miRNAs with extreme ratios that were further analyzed. Red symbols are confirmed Nbr targets (Table 3.3). (G) Scatterplot of microarray data from cells treated with dsRNA against Nbr or Renilla control. Highlighted are all of the genes >1.5 fold changed in either direction. (H) Real- time PCR for mRNAs from nbr and loqs mutant flies (Mean ± standard error of the mean, 4–6 experiments; *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$).

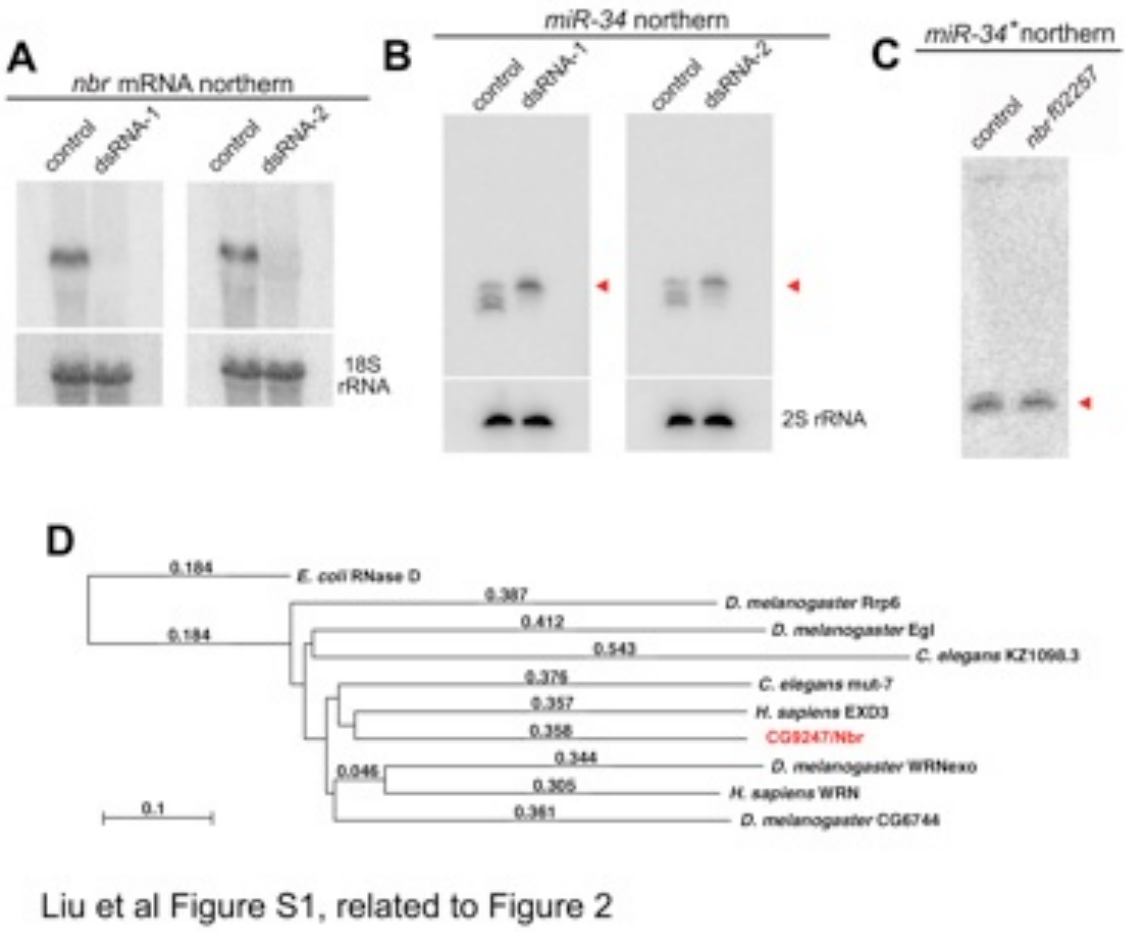


Figure 3.5. Reduction of nbr Affects Biogenesis of miR-34 Shorter Isoforms, Related to Figure 3.2. (A) mRNA Northern shows that treating cells with two independent dsRNAs directed to different regions of the nbr gene depleted nbr mRNA levels. Loading control, 18S rRNA. (B) Upon reduction of nbr, biogenesis of miR-34 shorter isoforms was affected. Loading control, 2S rRNA. (C) Loss of nbr has minimal effects on the level or pattern of miR-34*. Northern analysis of miR-34* in wild type versus nbrf00257 mutant flies. (D) Neighbor joining homology tree showing the relatedness of the 3'→5' exoribonuclease domains of Nbr and other exoribonucleases, rooted to the E. coli RNase D exonuclease domain.

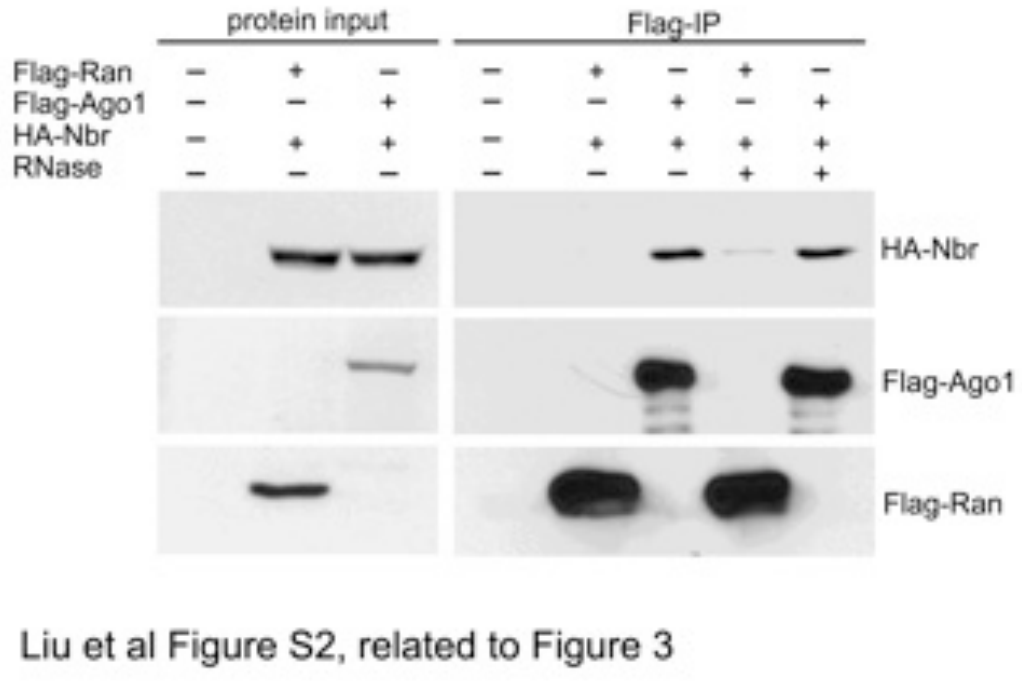
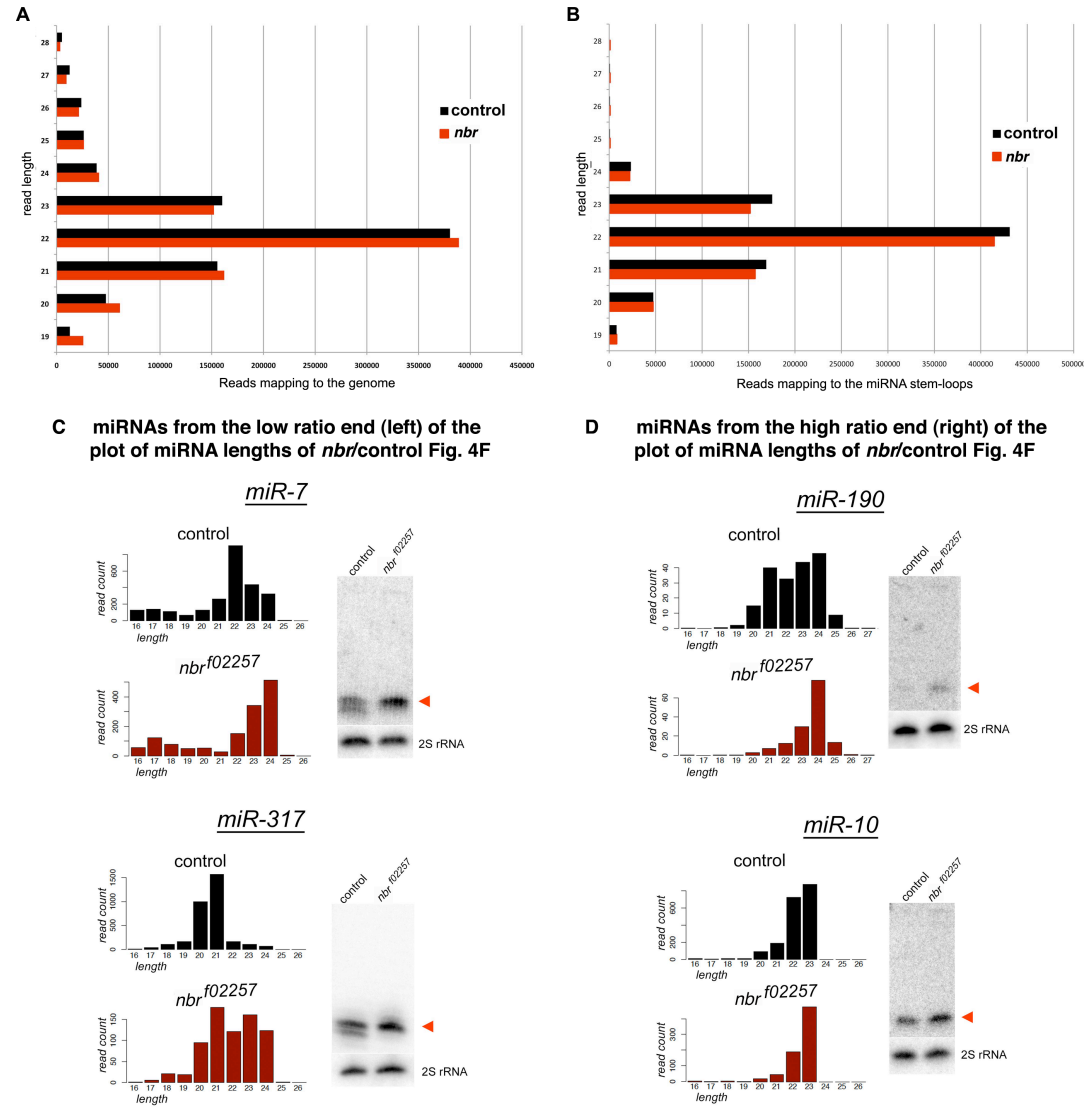


Figure 3.6. The Interaction between Nbr and Ago1 Is Not Dependent on RNA, Related to Figure 3.3. Ago1 and Nbr interact by co-immunoprecipitation in a manner that is RNA-independent. Cells were either left untreated or were transfected with HA-Nbr and either Flag-Ago1 or Flag-Ran (control), in the presence or absence of added RNase. Flag-tagged protein was then immunoprecipitated, and interacting proteins were probed by Western immunoblot. Protein input is 10% of Flag-IP.



Liu et al Figure S3, related to Figure 4

Figure 3.7. New nbr-Dependent Candidate miRNAs. Related to Figure 3.4. (A and B) Distribution of (A) overall reads and (B) reads mapping to the miRNA stemloops in control and nbrf02257 mutants. These show no overall difference in read lengths between control and mutant. (C and D) The left panels are the distribution plots of read lengths for these miRNAs, from which the ratio plot in Figure 3.4F was generated. The right panels are small RNA

Northerns from adult flies of control and nbr-/-mutants. In C are candidate miRNAs from the low ratio end of the plot in Figure 3.4F where the most abundant isoform in wild type is trimmed. The mutant shift in the distribution is observed by Northern blot analysis. In D, are two miRNAs are from the high ratio side of the plot in Figure 3.4F. For these miRNAs, only one major isoform is detectable by Northern in controls, but the level of this isoform becomes more abundant in nbr-/- (a 2.6-fold increase for miR-190, a 1.8-fold increase for miR-10). Deep sequencing analysis confirmed an impact on isoform distribution in the nbr-/- mutant (Table 3.2).

CHAPTER 4 IMPACT OF AGE-ASSOCIATED INCREASE IN 2'-O-METHYLATION OF MIRNAS ON AGING AND NEURODEGENERATION IN DROSOPHILA

This work has been published as follows:

Impact Of Age-Associated Increase In 2'-O-Methylation Of miRNAs On Aging And Neurodegeneration In Drosophila

Masashi Abe[1], Ammar Naqvi[2], Gert-Jan Hendriks[1,3], Virzhiniya Feltzin[1], Yongqing Zhu[1], Andrey Grigoriev[2,4*], and Nancy M. Bonini[1,3*]

[1]Department of Biology, University of Pennsylvania, [3]Howard Hughes Medical Institute, Philadelphia, PA 19104, USA

[2]Department of Biology, Center for Computational and Integrative Biology, Rutgers University, Camden, NJ 08102, USA

[3]Current address: Friedrich Miescher Institute for Biomedical Research, Maulbeerstrasse 66, 4058 Basel, Switzerland

*co-senior authors.

[4]Corresponding authors

E-mail nbonini@sas.upenn.edu

E-mail agrigoriev@camden.rutgers.edu

Keywords: miRNA, 2'-O-methylation, miRNA sorting, aging, neurodegeneration.

Note

The format of the figure and table numbers, and references have been modified from that published to conform to the format of the dissertation.

Contribution

The computational and bioinformatics analyses were performed in the Grigoriev Lab. I developed a pipeline to analyze the small RNA-seq NGS data and to discover trends and patterns, which resulted in identifying a subset of microRNAs that were modulated with the aging process and brain development. Using this pipeline I performed the analyses and prepared several figures/tables as well as parts of the manuscript text. Our results were then validated and confirmed by examining the expression pattern of miRNAs by Northern blot analysis by our collaborators. (Table 4.2 and 4.3, Figure 4.6 to 4.8).

Abstract

microRNAs (miRNAs) are 20~24nt small RNAs that impact a variety of biological processes, from development to age-associated events. To study the role of miRNAs in aging, studies have profiled the levels of miRNAs with time. However, evidence suggests that miRNAs show heterogeneity in length and sequences in different biological contexts. Here, by examining the expression pattern of miRNAs by northern blot analysis, we found that Drosophila miRNAs show distinct isoform pattern changes with age. Surprisingly, an increase of some miRNAs reflects increased 2'-O-methylation of select

isoforms. Small RNA deep-sequencing revealed a global increase of miRNAs loaded into Ago2, but not into Ago1, with age. Our data suggests increased loading of miRNAs into Ago2 but not Ago1 with age, indicating a mechanism for differential loading of miRNAs with age between Ago1 and Ago2. Mutations in Hen1 and Ago2, which lack 2'-O-methylation of miRNAs, result in accelerated neurodegeneration and shorter lifespan, suggesting a potential impact of the age-associated increase of 2'-O-methylation of small RNAs on age-associated processes. Our study highlights that miRNA 2'-O-methylation at the 3'end is modulated by differential partitioning of miRNAs between Ago1 and Ago2 with age, and that this process might impact age-associated processes in Drosophila.

Introduction

miRNAs are 20-24nt small RNAs that regulate gene expression post-transcriptionally through translational repression and/or mRNA cleavage. Starting from the discovery of the first miRNA lin-4 in C. elegans, hundreds of miRNAs have now been identified in various species. miRNAs were classically studied for their role on developmental timing (Bartel 2004, Lee, Feinbaum, and Ambros 1993, Reinhart et al. 2000). By now, functional analyses have identified some miRNAs, such as lin-4 in C. elegans (Boehm and Slack 2005, Kenyon 2010, Smith-Vikos and Slack 2012) and miR-34 in Drosophila (Liu et al. 2012) as miRNAs with critical roles in organismal and brain aging, respectively. Although most studies have focused on changes in the abundance of miRNAs with age (de Lencastre et al. 2010, Ibanez-Ventoso et al. 2006, Kato et al. 2011), evidence suggests that miRNAs also show heterogeneity in length and sequence in different cell types and biological contexts (Burroughs et al. 2010, Li, Liao, et al. 2012,

Llorens et al. 2013, Marti et al. 2010, Westholm et al. 2012). Such heterogeneity is mediated by imprecise Drosha and Dicer cleavages (Azuma-Mukai et al. 2008, Calabrese et al. 2007, Wyman et al. 2011), 3'end adenylation/uridylation (Burroughs et al. 2010, Landgraf et al. 2007, Ruby et al. 2006, Wyman et al. 2011), and RNA editing events (Kawahara et al. 2007; Nishikura 2010; Alon et al. 2012; Ekdahl et al. 2012). Importantly, effects of this heterogeneity on miRNA target silencing are also reported (Azuma-Mukai et al. 2008, Fukunaga et al. 2012, Lee and Doudna 2012, Seitz, Ghildiyal, and Zamore 2008), suggesting a functional impact of regulating miRNA heterogeneity in animals. Our study on the role of miR-34 on aging and age-associated neurodegeneration in Drosophila revealed an intriguing pattern of miR-34 isoforms: although multiple length isoforms of miR-34 are generated, only the short isoform accumulates with age (Liu et al. 2012). Furthermore, the generation of isoforms of miR-34 requires 3'end trimming by a novel 3'-to-5' exonuclease, Nibbler (Nbr) (Han et al. 2011, Liu et al. 2011) highlighting the potential importance of regulation of miRNA length on age-associated processes.

In addition to length and sequence heterogeneity, the 2'-OH of the 3' terminal ribose of miRNAs can be modified by 2'-O-methylation (Zhao et al. 2012). In plants, this modification occurs on nearly all miRNAs and siRNAs (Yu et al. 2005, Zhao, Mo, and Chen 2012, Zhao et al. 2012), and protects the small RNAs from HESO1-mediated uridylation and degradation (Li et al. 2005, Ren, Chen, and Yu 2012, Yu et al. 2005, Zhao et al. 2012). In Drosophila, most miRNAs are loaded into Ago1 and remain unmodified (Hutvagner et al. 2001, Okamura et al. 2004, Vagin et al. 2006). However, a

subset of miRNAs are found to be 2'-O-methylated (Forstemann et al. 2007, Ghildiyal et al. 2008, Horwich et al. 2007, Okamura, Liu, and Lai 2009). Importantly, this modification is also protective in animals: loss of 2'-O-methylation of small RNAs leads to destabilization, and tailing and trimming of the small RNAs (Ameres et al. 2010, Kamminga et al. 2010, Kurth and Mochizuki 2009).

In Drosophila, 2'-O-methylation of small RNAs is associated with loading of the miRNAs into different Ago complexes: while most miRNAs are loaded into Ago1 (miRISC) and remain unmodified, those loaded into Ago2 (siRISC) are 2'-O-methylated (Czech et al. 2009, Ghildiyal et al. 2010, Okamura, Liu, and Lai 2009). The miRNA/miRNA* duplex structure and 5' nucleotide preference are two suggested mechanisms by which miRNAs are directed into the two distinct Ago complexes (Forstemann et al. 2007, Ghildiyal et al. 2010, Okamura, Liu, and Lai 2009). miRNAs loaded into Ago2 are also active on translational silencing of the target genes, like miRNAs loaded into Ago1, but by a potentially different mechanism and with different silencing efficiency (Czech et al. 2009, Iwasaki, Kawamata, and Tomari 2009, Okamura, Liu, and Lai 2009). However, despite the discovery of functional Ago2-loaded and 2'-O-methylated miRNAs in Drosophila, whether such loading of miRNAs to Ago2 is biologically regulated, or biologically important in vivo, is not yet clear.

Here, we pursued study of several Nbr-dependent miRNAs that like miR-34 show multiple isoforms. Examination of changes in their isoform pattern with age unexpectedly revealed a diversity of patterns. The age-associated increase of some

Drosophila miRNAs reflected increased 2'-O-methylation of select isoforms. Associated with this increase, we found increased loading of specific miRNA isoforms into Ago2, but not Ago1, with age. Furthermore, the lack of 2'-O-methylation by Hen1 and Ago2 mutations resulted in reduced lifespan and brain degeneration, suggesting that the increased protection of small RNAs with age may impact age-associated events.

Results

Distinct age-associated isoform patterns of Nbr-trimmed miRNAs

Drosophila miR-34-5p, which is a target of Nbr-dependent 3'end trimming (Han et al. 2011, Liu et al. 2011), shows an increase in the short isoform with age (Liu et al. 2012) (Figure 4. 1A, B). This observation led us to examine the age-associated pattern of other Nbr-trimmed miRNAs. We focused on several miRNAs that could be readily detected by northern analysis. Northern blots of these miRNAs at 3d and 30d in heads revealed intriguing changes in isoform patterns. One Nbr-trimmed miRNA, miR-317-3p, increased with age like miR-34-5p (Figure 4. 1A,B). The increase in both of these miRNAs reflected an increase in their short isoforms with age. By contrast, although both miR-305-5p and miR-263a-5p increased in total amount with age, their isoform pattern was opposite: these miRNAs showed an increase in the longer isoforms with age (Figure 4. 1C,D). miR-11-3p showed a single isoform by northern and increased by 30d (Figure 4. 1E,F). These observations raised the possibility that multiple mechanisms are used to generate distinct miRNA isoform patterns with age.

Age-associated increase in 2'-O-methylation of Nbr-trimmed miRNAs

One mechanism to explain an increase in long isoforms with age could be an increased protection of these specific isoforms with age. Almost all plant miRNAs and siRNAs (Li et al. 2005, Ren, Chen, and Yu 2012, Yu et al. 2005, Zhao, Mo, and Chen 2012, Zhao et al. 2012), and animal siRNAs and piRNAs are 2'-O-methylated (Horwich et al. 2007, Kirino and Mourelatos 2007a, b, c). Recent studies also suggest that a subset of Drosophila miRNAs are 2'-O-methylated (Forstemann et al. 2007; Horwich et al. 2007; Czech et al. 2009; Okamura et al. 2009; Ghildiyal et al. 2010). Importantly, 2'-O-methylation of animal small RNAs is also protective such that loss of 2'-O-methylation leads to their destabilization (Ameres et al. 2010, Kamminga et al. 2010). Therefore, we hypothesized that the increase of the long isoforms of miR-305-5p, miR-263a-5p, and miR-11-3p may result from increased protection at the 3’end of specific isoforms by 2’-O-methylation. To address this, we performed oxidation/β-elimination assays on total RNA from 3d and 30d wild-type head and body. If the 3'end ribose of small RNAs is protected by 2'-O-methylation, as for piRNAs and siRNAs, those small RNAs cannot react with sodium periodate in the oxidation/β-elimination assay and will not decrease in size. However if the 3'end ribose is unmodified, they react, will be β-eliminated, and will shift down in length (Horwich et al. 2007). Therefore a change in the mobility of the small RNAs by northern indicates the presence or absence of a protective modification at the 3'end. We performed this assay and found that miR-34-5p shifted down in size, thus is not protected at the 3’end (Figure 4. 2A, left panel). By contrast, miR-305-5p, miR-263a-5p, and miR-11-3p showed an increase in protection of the long isoforms with age (Figure 4. 2A, arrows, right panel). Quantification of the ratio of protected to unprotected isoforms of miR-305-5p, miR-263a-5p, and miR-11-3p showed an increase in the ratio of

protected to unprotected isoforms with age. If the increase of the protected isoforms simply reflected an increase of the overall level of these miRNAs, the ratio of protected to unprotected isoforms after oxidation/β-elimination would not change. However the ratio increased for these three miRNAs (Figure 4. 2B), indicating a mechanism to increase the protection of the long isoforms of these miRNAs with age.

We next determined whether the protection of specific miRNA isoforms is due to 2'-O-methylation at the 3'end. 2'-O-methylation of small RNAs depends on the methyltransferase Hen1. In addition, 2'-O-methylation of small RNAs requires loading into Ago2: loss of Ago2 and Ago2-loading machinery (Dcr2 and R2D2) leads to loss of modification (Horwich et al. 2007). Therefore, to confirm that the miRNAs were 2'-O-methylated and loaded into Ago2, we tested the effect of loss of Hen1 and Ago2 on the pattern of miR-305-5p, miR-263a-5p, and miR-11-3p with age. $Hen1^{f00810}$ and $Ago2^{BL16608}$ mutant flies were aged to 3d or 30d, RNA isolated and oxidation/β-elimination performed. This revealed that miR-305-5p, miR-263a-5p, and miR-11-3p were no longer protected, but rather depended on Hen1 and Ago2 for protection at both 3d and 30d (Figure 4. 3, arrowheads). This finding indicates that the increase of the long isoforms of these miRNAs is dependent on Hen1, thus is due to 2'-O-methylation at the 3'end. These data also indicate that these miRNA isoforms are loaded into Ago2 for 2'-O-methylation. We note that miR-317-3p is exceptional in that it showed protection of the long isoform from oxidation/β-elimination in both young and old flies (Figure 4. 2A, arrowheads); however, to our surprise, protection persisted even after the loss of Hen1

and Ago2 (Figure 4. 3, arrowheads). This suggests that miR-317-3p is protected in a manner independent of Hen1 and Ago2 function.

An age-associated increase in Ago2-loading of 2'-O-methylated isoforms of miRNAs

In Drosophila, while most miRNAs are loaded into Ago1 and thus are unmodified, a subset of miRNAs have been observed to be loaded into Ago2 (siRISC) and are 2'-O-methylated (Czech et al. 2009, Ghildiyal et al. 2010, Okamura, Liu, and Lai 2009). Therefore, based on our findings, we proceeded to determine whether the protected isoforms of miR-305-5p, miR-263a-5p, and miR-11-3p increased their Ago2 loading with age. To do this, we immunoprecipitated Ago1 and Ago2 from whole flies, then performed northern blots on the precipitated small RNAs, comparing 3d to 30d. Ago1 could be immunoprecipitated with an antibody; for Ago2, we used flies bearing a genomically-tagged FLAG-HA-Ago2 (Czech et al. 2008). This approach revealed a specific loading pattern for each small RNA with age in Ago1 vs Ago2. miR-34-5p, which was not protected at the 3'end (see Figure 4. 2A), was predominantly loaded into Ago1, and not detectably present in Ago2 (Figure 4. 4A). Consistent with the age-associated pattern of miR-34-5p by northern from total RNA (see Figure 4.1A, Liu et al., 2012), the short isoforms of miR-34-5p accumulated in Ago1 with age (Figure 4. 4A). esi-2.1, an abundant endogenous siRNA in flies (Czech et al. 2008, Ghildiyal et al. 2008, Kawamura et al. 2008, Okamura et al. 2008) was selectively loaded into Ago2, and was not detectable in Ago1 (Figure 4. 4A). By contrast, miR-305-5p and miR-263a-5p showed an intriguing loading pattern: while multiple isoforms were present in Ago1 at both 3d and 30d, only the long isoforms (24nt) were present in northerns of Ago2 IP

material (Figure 4. 4A). The single isoform of miR-11-3p was present in Ago1, however, a small portion was loaded into Ago2 (Figure 4. 4A). In addition to the loading of specific length of isoforms into Ago2, there was also a trend of increased levels of the loaded isoforms of miR-305-5p, miR-263a-5p, and miR-11-3p into Ago2 with age. To quantify this, we normalized the level of the Ago2-loaded miRNAs to Ago2-loaded esi-2.1 at the corresponding age; the total level of esi-2.1 did not change with age (Figure 4. 4B), and is loaded into Ago2 and not into Ago1 (Figure 4. 4A). After normalizing, we confirmed a significant increase of Ago2-loaded miR-305-5p, miR-263a-5p, and miR-11-3p with age (Figure 4. 4C,D,E). These findings suggest that the age-associated increase of 2'-O-methylation of miRNAs reflects increased loading of selective isoforms into Ago2 with age.

One possibility for increased loading into Ago2 could be an increase in the level of Hen1, Ago2 and/or the Ago2-loading machinery (Dcr2 and R2D2) with age (Liu et al. 2003, Marques et al. 2010, Nishida et al. 2013, Okamura et al. 2011, Tomari et al. 2004). To address this possibility, we performed western immunoblots on Ago2 (FLAG-HA-Ago2 flies), R2D2, and Dcr2 with age (no antibody is available for Hen1). Both Ago2 and R2D2 were unchanged with age (Figure 4. 4F,G,H), and Dcr2 was not detectable in adult flies (Figure 4. S1). To further confirm these results, we examined microarray data from aged Drosophila brains (3d, 30d, and 60d) (Liu et al. 2012). Analysis confirmed no significant increase in Hen1, Ago2, Dcr-2, or R2D2 RNA levels with age (Table 4. S1). In addition, the level of Ago1 mRNA and protein was unchanged, suggesting that a decrease in Ago1 levels would not explain the relative increase in Ago2-loading of

miRNAs with age (Table 4. S1; Figure 4. S2). We observed a potential increase in mRNA levels for Hsp70 (Table 4. S1), a component of RISC-loading machinery in a complex with Hsp90 (Iki et al. 2010, Iwasaki et al. 2010, Miyoshi et al. 2010). However, since the Hsp90 level is unchanged, there is likely no increase in the Hsp90/Hsp70 complex with age. Together, these results suggest that the increase of Ago2-loading of miRNAs with age is independent of a change in the level of Ago2 or Ago2-loading machinery.

A global shift of Ago1 vs Ago2-loaded miRNAs with age

To confirm the increase in miRNA loading into Ago2 with age, we immunopurified Ago1 and Ago2 from flies aged 3d and 30d, and performed deep sequencing of the small RNA fraction. We normalized the number of observed reads to the total number of non-miRNA reads in Ago1, and to the esi-2.1 reads in Ago2. This analysis showed that the total read number of miRNAs in Ago1 was largely unchanged, while that of Ago2 increased with age (Figure 4. 5A). In addition, the share of miRNAs among total small RNAs occupying Ago1 showed little change with age, while in Ago2 the miRNA reads increased dramatically with age (Figure 4. 5B, 30.7% at 3d to 69.6% at 30d). During the analysis, we noticed that miR-263a-5p was a singular most abundant miRNA in the 3d and 30d Ago2 libraries. Therefore, to rule out that the increased trend of total miRNA reads and percentage in Ago2 with age simply reflected a change in miR-263a-5p, we also removed miR-263a-5p reads from 3d and 30d Ago2-IP libraries, and re-calculated the total miRNA read number (normalized to esi-2.1) and the percentage of total miRNAs in 3d and 30d Ago2 libraries. This confirmed that, even after removing miR-263a-5p, the

trend persisted: the read number and percentage of the total miRNAs in Ago2 increased with age (Figure 4. S3). These results, together with the increased ratio of protected to unprotected isoforms of miRNAs with age (see Figure 4. 2), support the idea that more miRNA molecules are becoming loaded into Ago2, and not Ago1, with age.

We analyzed in detail the three miRNAs that were the focus of northern analysis: miR-305-5p, miR-263a-5p, and miR-11-3p. Supporting the earlier observations, miR-305-5p, miR-263a-5p, and miR-11-3p showed increased association in Ago2 with age (Figure 4. 5C, bottom panels). By contrast, the amount of these miRNAs in Ago1 decreased with age (Figure 4. 5C, top panels). Together, these data support the idea that the loading pattern of miRNAs between Ago1 and Ago2 changes dramatically with age, such that Ago2-loading of selective miRNA isoforms increases, while Ago1-loading of the miRNA isoforms generated from the same stem-loops decreases.

Detailed analysis of the libraries allowed identification of additional miRNAs with the same trend of an increase of select isoforms in Ago2 with age. We compared relative loading ratios (Ago2 over Ago1, R_{21}, see Methods) of the most abundant isoform for each miRNA in 30d flies vs 3d flies (Figure 4. 6A). This showed the three miRNAs that were the focus of our tests with northerns (miR-305-5p, miR-263a-5p, and miR-11-3p) are among those preferentially loaded into Ago2 at both 3d and 30d (Figure 4. 6A).

Among the 240 Drosophila miRNA genes, 135 were detected with at least one raw read in both Ago1 and Aog2-IP libraries at 3d, while 143 were detected at 30d.

Among these, 67 5p and 3p sequences had at least 1000 raw reads of at least one isoform at 3d or 30d, of which 15 mature and 9 star sequences showed preferential loading of specific isoforms into Ago2 relative to Ago1 at 3d (R_{21}> 1.2), while 16 mature and 9 star sequences showed that trend at 30d (R_{21}> 1.2; Figure 4. 6A, Table 4.s S2 and S3). (5p and 3p sequences were classified as "mature" or "star", based on their relative abundance in the Ago1-IP 3d library and a consistent behavior across all four libraries: those with a higher read number compared to the read number from the other strand of the stemloop were defined as "mature". See Methods for details.).

Next, we determined whether the miRNA isoforms showing preferential loading into Ago2 compared to Ago1 at either age (R_{21}>1.2 at 3d or 30d; Figure 4. 6A) also showed an increase in loading into Ago2 with age. To assess the latter value, we calculated for each miRNA the fold change between 30d and 3d in normalized read numbers of the main isoform loaded into Ago2 (Figure 4. 6B). A change of >1 would indicate miRNAs with an increase in specific miRNA loading into Ago2 with age. We confirmed that miR-305-5p, miR-263a-5p, and miR-11-3p showed increased loading into Ago2 with age (Figure 4. 4A-E; Figure 4. 6B, red bars). Together, these data indicate that 11 mature and 8 star sequences out of 67 (5p and 3p) total sequences, with >1000 raw reads of at least one isoform, showed preferential loading of specific isoforms into Ago2 relative to Ago1, and such loading increased with age. We also analyzed a library from a previous study comparing Ago1-IP miRNAs to 3'end protected miRNA from heads (Ghildiyal et al. 2010). This analysis also indicated a selective enrichment for particular miRNAs (SFigure 4. 4). From these analyses (Fig 6B and SFigure 4. 4), we then

performed northern blots on several highlighted miRNAs that would be sufficiently abundant to detect, to examine whether their protection increases with age: miR-1000, miR-100 and miR-8 (black arrows Figure 4. 6B and green bars SFigure 4. 4). Northern blots confirmed accumulation of long isoforms protected at the 3'end that increased both in level and in protected isoforms with age (Figure 4. 6C).

To gain potential insight into the functional effects of the shift in miRNAs from Ago1 to Ago2 with age, we analyzed gene expression changes in the brain with age comparing 3d to 30d from a previous microarray analysis (Liu et al. 2012). We selected down- and upregulated genes with high stringency ($p<0.001$, Benjamini-Hochberg FDR < 5%) with expression change (down or up) above two-fold. At 30d vs 3d, 719 genes were significantly downregulated and 199 significantly upregulated, giving a global ratio of age-related downregulation of 3.61 for the whole genome. We then considered the potential effects of miRNAs with increased loading into Ago2 by deep sequencing that were confirmed by northerns (miR-1000, miR-100, miR-305, miR-11, miR-263a, miR-263b, miR-8). In a combined list of mRNA targets of these miRNA (predicted by TargetScan), a downregulation ratio is substantially higher ($104/13=8$, $p<0.001$, Benjamini-Hochberg FDR < 5%), thus these target genes are twice as likely to show stronger downregulation with age compared to the whole genome. We further analyzed the splicing array data of Taliaferro et al. (2013) following knockdowns of Ago proteins in S2 cells. While these arrays were designed for quantifying alternative splicing rather than gene expression they still show trends consistent with our observations. Of 159 deregulated genes with Ago1 knockdown, eight predicted targets of the miRNAs shifted to Ago2 showed changed expression (downregulated, except for the cals gene); all were

also downregulated in the ageing microarray at 30d vs 3d. Among the 51 de-regulated genes in Ago2 knockdown, four predicted targets of the miRNAs shifted to Ago2 showed a change of expression but it was upregulation (except for the bbc gene), while they were downregulated in the ageing microarray at 30d vs 3d. Thus, for the group of genes in common, there is a striking reversal of expression following Ago2 knockdown, but not Ago1 knockdown, compared to the general downregulation with age.

Loss of Hen1 and Ago2 are associated with shortened lifespan and neurodegeneration

To address the biological significance of the global shift in the population of Ago2-loaded miRNAs, we examined the impact of global loss of 2'-O-methylation of small RNAs by analyzing Hen1 and Ago2 mutant animals. Ago2 loss is known to cause defects in synaptic structure in third instar larvae (Pepper et al. 2009), indicating the importance of Ago2 on neuronal development. Here we focused on adult-specific age-associated effects. Lifespan analysis indicated that both $Hen1^{f00810}$ and $Ago2^{BL16608}$ mutants showed a shorter lifespan compared to genetic background-matched controls (Figure 4. 7D). Detailed analysis of these animals also showed that both $Hen1^{f00810}$ and Ago2 ($Ago2^{414}$ and $Ago2^{BL16608}$) mutants have increased brain vacuolization, indicative of brain degeneration with age (Figure 4. 7A,B,C). While $Hen1^{f00810}$ mutants showed more brain vacuolization in the optic lamina (Figure 4. 7B), Ago2 mutants had excessive vacuoles in the retina (Figure 4. 7C). Both Hen1 and Ago2 mutants are predicted to impact siRNAs which may contribute to the effects, and the difference in phenotypes between Hen1 and Ago2 mutants might reflect a broader role of Ago2 beyond its effects on siRNAs (Taliaferro et al. 2013). However, defects in both Hen1 and Ago2 mutants

suggest that 2'-O-methylation of small RNAs, one class of which is miRNAs, is critical for age-associated processes in Drosophila.

Discussion

Dynamic regulation of Ago1 vs Ago2-loading of miRNAs with age

Increasing evidence suggests that miRNAs show heterogeneity at the 5'end, 3'end and even in precise sequence in different cell types and biological contexts (Neilsen, Goodall, and Bracken 2012). Despite this, it is less understood whether and how such heterogeneity is regulated biologically or the biological impact. Here, by examining the age-associated pattern of miRNA isoforms, we show that 2'-O-methylation of miRNAs changes in an age-dependent manner in Drosophila. This change is correlated with a shift in partitioning of miRNAs into Ago2 vs Ago1 with age. These data suggest that the partitioning of miRNAs between Ago1 and Ago2 is modulated with age in Drosophila.

The mechanism of miRNA-loading into the different Ago complexes in Drosophila, especially in vivo, remains unclear. Previously, miR-277 was shown to be loaded into both Ago1 and Ago2 because of the lack of an extensive central bulge in the miR-277/miR-277* duplex (Forstemann et al. 2007). However, our analysis of the miRNAs that show higher levels of 2'-O-methylation with age failed to identify a consistent lack of central bulges on miRNA/miRNA* duplexes (Figure 4. S5). Rather, our results indicate that miRNA isoforms with the same sequence and length can be partitioned into both Ago1 and Ago2, and that a change in this distribution is an age-associated phenomenon. This effect, at least in vivo, is likely to be driven by a bulge-

independent mechanism. Bulge-independent partitioning of small RNAs into different Ago complexes is suggested from other work as well: the loop sequence from miR-34 precursor stem-loop is abundant in some cells and loaded into Ago1, despite its single-stranded nature (Okamura et al. 2013).

In vivo significance of differential partitioning of miRNAs between Ago1 and Ago2 with age

Our analyses of Hen1 and Ago2 mutants suggests that 2'-O-methylation of small RNAs affects age-associated traits of brain degeneration and lifespan. In addition to effects on other small RNA classes, it is possible that the loss of 2'-O-methylation of miRNAs leads to destabilization of the miRNAs, thus affecting target silencing and causing accelerated age-associated defects. In this case, Ago2-loaded, but not Ago1-loaded, miRNAs would be selectively affected. Interestingly, both Ago1 and Ago2 are active on silencing miRNA targets. However, Drosophila Ago1 and Ago2 employ different mechanisms to silence target mRNAs, as silencing by Ago1-let-7 RISC entails the removal of the polyA tail, whereas that by Ago2-let-7 RISC entails retention of the polyA tail (Iwasaki, Kawamata, and Tomari 2009). Such mechanistic differences might contribute to the differential silencing efficiency of the two Ago complexes of the reporters: Ago1-let-7 is more efficient than Ago2-let-7 to silence the reporters (Iwasaki, Kawamata, and Tomari 2009). Our analysis of gene expression changes with age, indicates a global downregulation in the brain; correlating this with Ago1 vs. Ago2 knockdown data in cells indicates greater downregulation of predicted mRNA targets of miRNAs shifted to Ago2. While we cannot exclude that such changes result from other

activities of Ago2 (Taliaferro et al. 2013), the observed reversal of expression is also consistent with a model of regulation resulting from increased loading of miRNAs shifted to Ago2 with age. Thus, potentially the shift of these miRNAs into Ago2 may affect downregulation of their targets with age, and disruption of this process in Hen1 and Ago2 mutants may contribute to the lifespan and neurodegenerative phenotypes observed.

An interesting possibility is that, by shifting miRNAs towards Ago2 from Ago1 with age, the organism might be adjusting the efficiency of target gene expression for ongoing or upcoming age-associated stresses. Since Ago2-mediated translational silencing causes retention of the polyA tail (Iwasaki, Kawamata, and Tomari 2009), this might make it possible to respond to age-associated internal or external stresses more rapidly by re-activation of target mRNAs. Another possibility is the effect of loss of 2'-O-methylation on other classes of small RNAs, such as endo-siRNAs or piRNAs (Czech et al. 2008, Ghildiyal et al. 2008, Horwich et al. 2007, Kawamura et al. 2008, Okamura et al. 2008). Loss of Ago2 leads to decreased production of endo-siRNAs, which is correlated with upregulation of transposons (Czech et al. 2008, Ghildiyal et al. 2008, Kawamura et al. 2008, Okamura et al. 2008). Recently, an age-associated increase of transposon expression and shorter lifespan were noted in Ago2 mutants (Li et al. 2013). Although Ago2 mutants show developmental defects (Pepper et al. 2009), both $Hen1^{f00810}$ and Ago2 mutants are maintained in homozygous condition, and that the overall brain morphology and climbing activity of these mutants when young is normal (see Figure 4. 7). Although it is possible that loss of these gene activities during development might sensitize the adult state, taken together these data suggest that the defects observed in

adult Hen1 and Ago2 mutants more likely reflect defects in adult age-associated processes. These findings support the importance of terminal modification of small RNAs on impacting age-associated traits.

Dynamic isomiR regulation with age

It has become evident that the pattern of miRNA isoforms (isomiRs) varies in different biological contexts, such as different stages of development, different tissues, and with disease (Fernandez-Valverde, Taft, and Mattick 2010, Li, Liao, et al. 2012, Llorens et al. 2013, Marti et al. 2010, Neilsen, Goodall, and Bracken 2012).. It is also becoming evident that such isoform distributions are regulated (Neilsen et al. 2012). For example, Nbr defines the 3'end of many Drosophila miRNAs (Han et al. 2011; Liu et al. 2011), and Loqs-PB (Drosophila homologue of TRBP) partnering with Dcr-1 defines the cleavage position of several Drosophila miRNAs, which seems conserved in mammals as well (Fukunaga et al. 2012; Lee and Doudna 2012). Our study reveals another repertoire of biologically heterogeneity of miRNAs: differential 2'-O-methylation of miRNA with age. It is intriguing that the loss of such miRNA heterogeneity leads to clear biological defects (this study and Abe et al., submitted). These results raise the possibility that, in the aging adult, fine-tuning of miRNA heterogeneity might be critical for combating age-associated stresses.

Acknowledgements:

We thank members of the Bonini laboratory for critical reading, Xiuyin Teng for outstanding technical assistance, and Matthew Taliaferro and Dr. Donald Rio for kindly

sharing data. This work was funded, in part, by the Ellison Medical Foundation (to N.M.B), and the NSF (grant DBI-1126052, to A.G.).

Figure and Tables

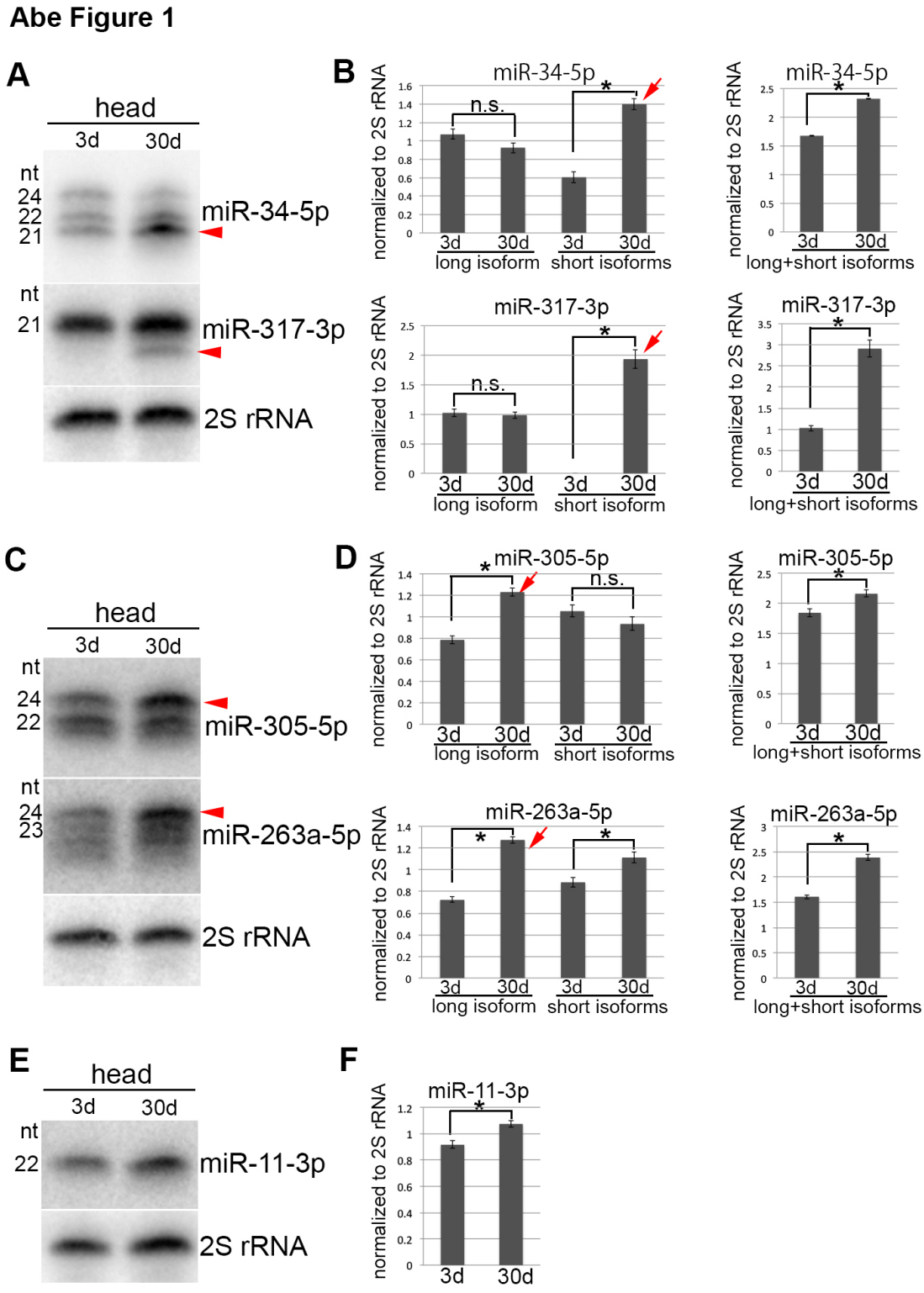


Figure 4.1. Nbr-dependent miRNAs show distinct isoform patterns with age.

Northern blots of different Nbr-dependent miRNAs with age. (A, B) miR-34-5p and miR-317-3p showed accumulation of short isoforms with age (A, (B left panel)). Both miR-34-5p and miR-317-3p also increased in total amount with age (B, right panel). (C, D) miR-305-5p and miR-263a-5p showed accumulation of long isoforms with age (C, D left panel). Both miR-305-5p and miR-263a-5p increased in total amount with age (D, right panel). (E, F) miR-11-3p, a single form miRNA by northern, accumulated with age. (B, D, F) Left panel: Quantification of the different isoforms of each miRNA. Right panel: Quantification of the total amount of each miRNA with age. Red arrowheads and arrows indicate the isoforms that increased with age. Mean ±SD (n=3, *p<0.05 (student's t-test)).

Abe Figure 2

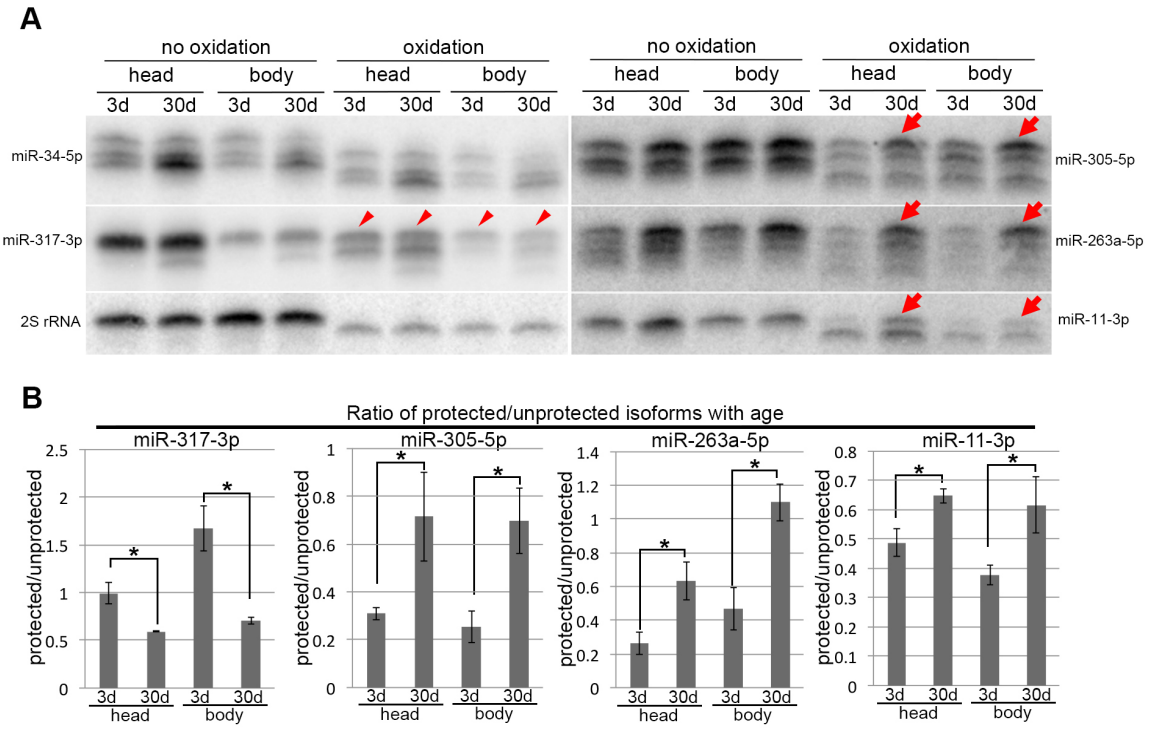


Figure 4.2. Age-associated increase of long isoforms of miR-305, miR-263a/b, and miR-11 is associated with increased protection from oxidation/β-elimination. Northern blots of miR-305, miR-263a/b and miR-11, miR-34 and miR-317, without (left side of each miRNA panel) and with (right side of each miRNA panel) treatment by oxidation/β-elimination, with quantitation. (A) All isoforms of miR-34 were sensitive to oxidation/β-elimination. The longest isoform of miR-317 was protected (arrowheads), but there was a decrease in the ratio of the protected isoform with age. miR-305, miR-263a/b, and miR-11 show accumulation of a protected long isoform with age (arrows). (B) Ratio of oxidation/β-elimination protected isoforms to unprotected isoforms with age. Mean ±SD (n=3, *$p<0.05$ (Student's t-test)).

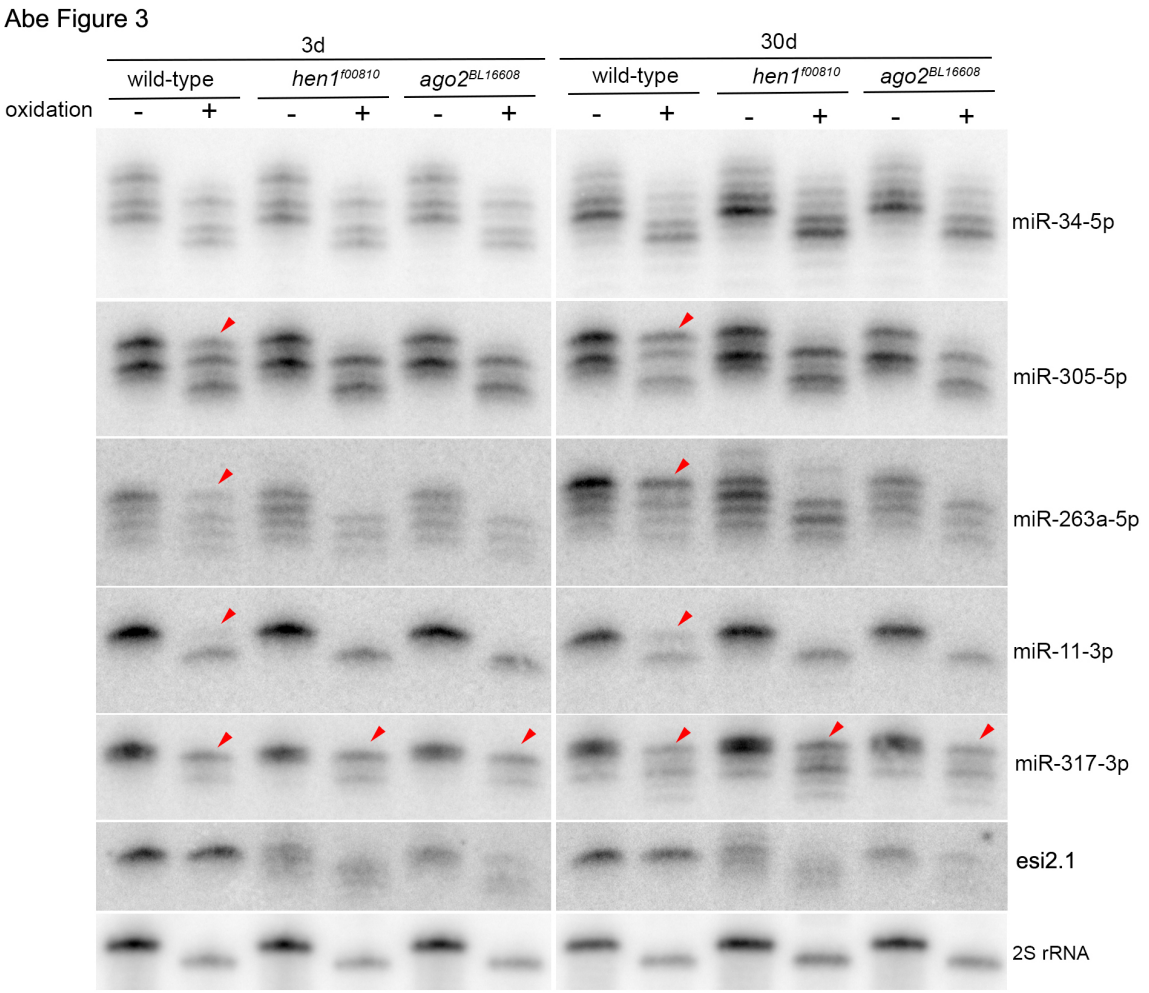


Figure 4.3. The age-associated increase of the long isoforms of miR-305, miR-263a, and miR-11 is eliminated upon Hen1 and Ago2 mutation.

Northern blots of each miRNA with or without oxidation/β-elimination at 3d and 30d in wild-type, Hen1^{f00810}, and Ago2^{BL16608} animals. Red arrowheads indicate the isoforms that are protected after oxidation/β-elimination in wild-type; these forms for miR-305, miR-263a/b, and miR-11 were no longer protected in Hen1^{f00810} and Ago2^{BL16608} mutant animals. The long forms of miR-317 remained protected in Hen1^{f00810} and Ago2^{BL16608} animals, indicating an alternative mechanism by which the long isoform accumulates for this miRNA.

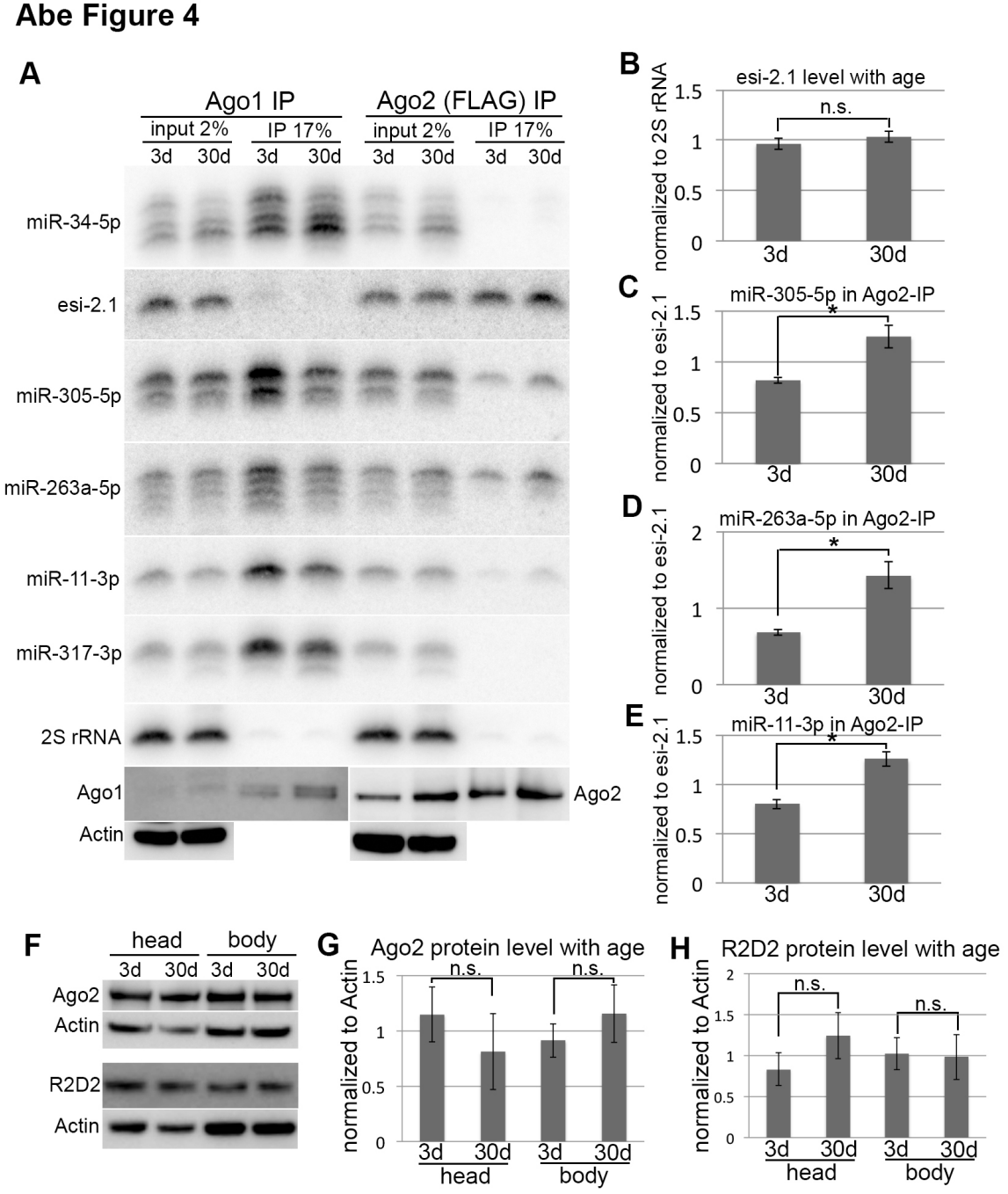


Figure 4.4. Increased Ago2 loading of long miRNA isoforms with age.

Northern blots with quantitation of miRNAs associated with Ago1 vs Ago2 with age. (A) Northern blots for small RNAs on RNA isolated from (left) Ago1-IP and (right)

Ago2-IP (FLAG-HA-Ago2). (B) Quantification of esi-2.1 level with age (mean ±SD (n=4), student's t-test confirmed no significant differences with age). There is no change in esi-2.1 levels with age. (C~E) Quantification of miRNA isoforms loaded into Ago2 with age, normalized to esi-2.1 in Ago2-IP. Mean ± SD (n=3, * $p<0.05$ (Student's t-test)). For all three miRNAs, the Ago2-loaded isoforms increase with age. (F) Western immunoblot for Ago2 and R2D2 with age. (G) Quantification of Ago2 protein level with age. Mean ± SD (n=3, not significant, student's t-test). (H) Quantification of R2D2 protein level with age. Mean ± SD (n=4, not significant by student's t-test).

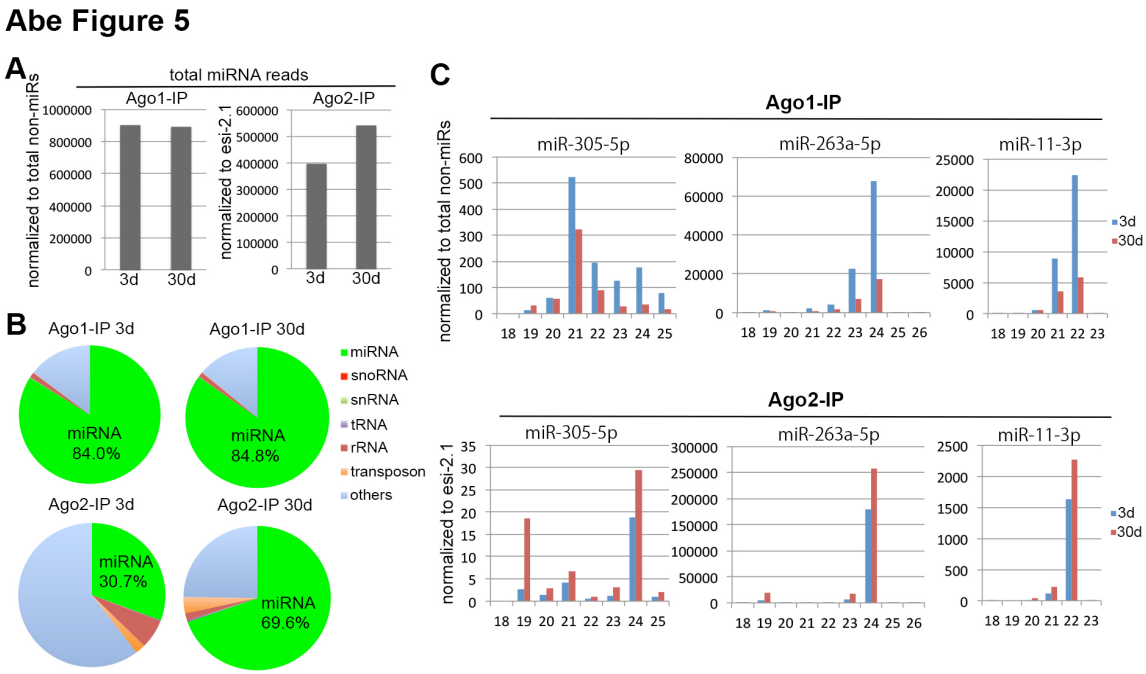


Figure 4.5. Ago1 vs Ago2-IP small RNA deep-sequencing with age.

(A) Normalized read number of total miRNAs in Ago1 or Ago2 with age. (B) Percentage of miRNAs and other small RNA classess in Ago1 and Ago2 with age. (C) miR-305-5p,

miR-263a-5p, and miR-11-3p show a decrease in Ago1, but an increase of specific isoforms in Ago2, with age.

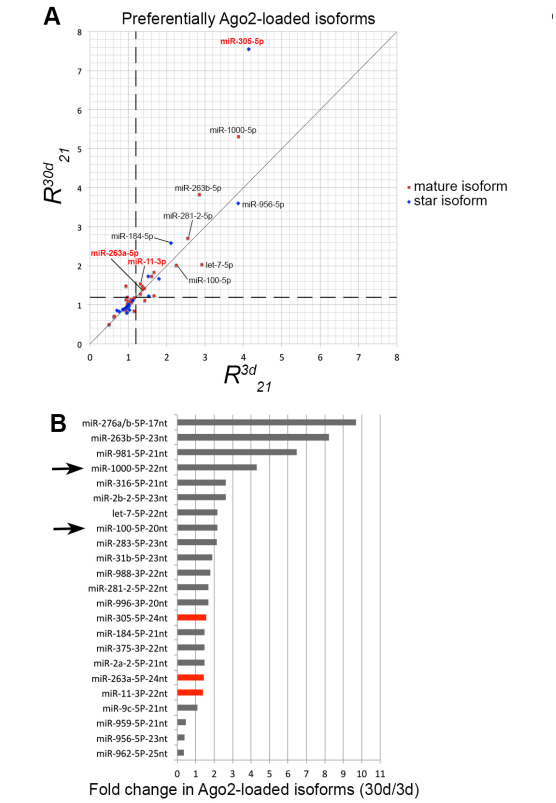

A
Preferentially Ago2-loaded isoforms
miR-305-5p
miR-1000-5p
miR-263b-5p
miR-956-5p
miR-281-2-5p
miR-184-5p
miR-263a-5p
miR-11-3p
let-7-5p
miR-100-5p
mature isoform
star isoform
B
miR-276a/b-5P-17nt
miR-263b-5P-23nt
miR-981-5P-21nt
miR-1000-5P-22nt
miR-316-5P-21nt
miR-2b-2-5P-23nt
let-7-5P-22nt
miR-100-5P-20nt
miR-283-5P-23nt
miR-31b-5P-23nt
miR-988-3P-22nt
miR-281-2-5P-22nt
miR-996-3P-20nt
miR-305-5P-24nt
miR-184-5P-21nt
miR-375-3P-22nt
miR-2a-2-5P-21nt
miR-263a-5P-24nt
miR-11-3P-22nt
miR-9c-5P-21nt
miR-959-5P-21nt
miR-956-5P-23nt
miR-962-5P-25nt
Fold change in Ago2-loaded isoforms (30d/3d)
C
no oxidation
oxidation
head
body
3d
30d
miR-8
miR-100
miR-1000
normalized to 2S rRNA
protected/unprotected

Figure 4.6. Identification of Ago2-loaded miRNA isoforms whose loading increases with age.

(A) Dot plot showing the ratio of preferentially Ago2-loaded isoforms at 3d (X-axis) to 30d (Y-axis). We selected miRNAs with values R_{21}>1.2 as candidates with greater loading into Ago2 at 3d or 30d (dashed lines parallel to the axes; miR-305-5p, miR-263a-5p, and miR-11-3p are seen, full lists in Table S2 and S3). miR-34-18nt, miR-276a/b-5p-17nt, and miR-981-5p-21nt showed an extremely high R_{21} ratio at 3d and/or 30d (> 16, Tables S2 and S3). This suggests much higher accumulation of these miRNA isoforms in Ago2-IP compared to Ago1-IP at 3d and/or 30d. This could happen because of biologically relevant upregulation of these miRNA isoforms, or cloning bias and degradation from the corresponding longer isoforms. These miRNA isoforms were removed from Fig. 4.6, because including these data obscured the distribution of the other data in the plot. miR-34-18nt was undetectable by Ago1 vs Ago2-IP northerns. Therefore most-likely miR-34-18nt is an artifact of library preparation. (B) Fold change of the main isoform of preferentially Ago2-loaded miRNA isoforms with age (30d/3d). Red are miRNAs known from our northern analyses that show an increase in Ago2 association with age. Black arrows here and green bars in SFig. 3 indicate abundant miRNAs that were tested for accumulation in Ago2 with age in C. (C) Northern analysis (top) and quantitation (bottom) of miRNAs highlighted by the analyses in Fig. 6B and SFig. 3.

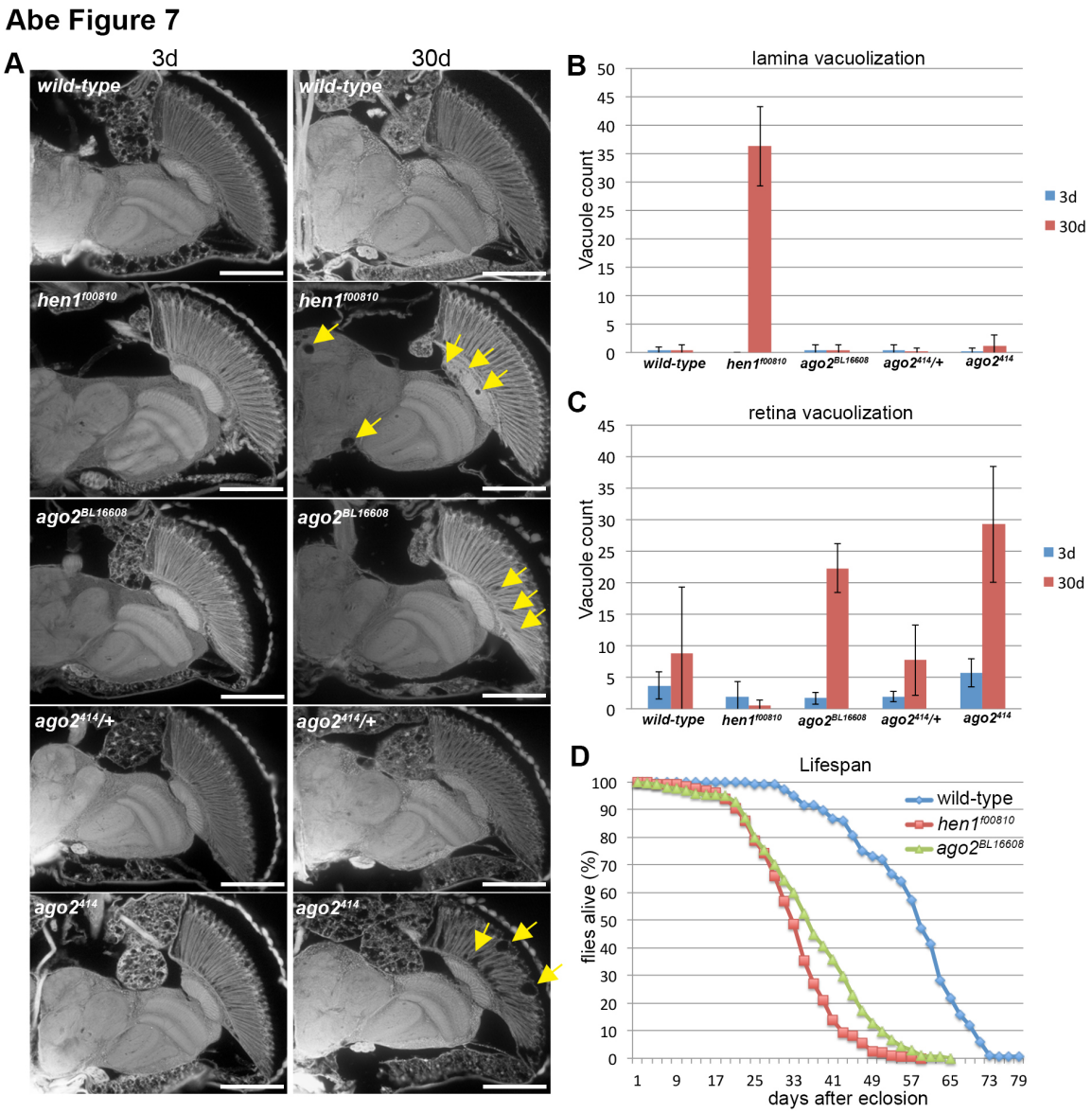


Figure 4.7. Mutations in Hen1 and Ago2 are associated with age-dependent brain degeneration and shorter lifespan.

(A) Paraffin sections of wild-type, Hen1^{f00810} , Ago2^{BL16608}, Ago2^{414}/+, and Ago2^{414} heads at 3d and 30d. Scale bar = 0.1 mm. The abundant vacuoles that are present in 30d animals in the mutants are highlighted in yellow. (B) Quantification of brain vacuoles in lamina. (C) Quantification of brain vacuoles in retina. (D) Lifespan of wild-type, Hen1^{f00810}, and Ago2^{BL16608} mutants. Mutant lines have been backcrossed into the genetic

background of the control wild type (see Methods). The lifespans of the mutants are significantly different from wild-type ($P<0.0001$ for each compared to wild-type, log rank analysis; chi-square value with $Hen1^{f00810}$ 327.7, with $Ago2^{BL16608}$ is 253.3)

CHAPTER 5 AGE-DRIVEN MODULATION OF TRNA-DERIVED FRAGMENTS IN DROSOPHILA AND THEIR POTENTIAL TARGETS

This chapter forms the basis of a manuscript submitted to Biology Direct:

Age-driven modulation of tRNA-derived fragments in Drosophila and their potential targets

Ammar S. Naqvi[1#], Spyros Karaiskos[1#], Karl E. Swanson[1],

Nancy M. Bonini[2], Andrey Grigoriev[1]*

[1] Department of Biology, Center for Computational and Integrative Biology, Rutgers University, Camden, New Jersey 08102, USA

[2]Department of Biology, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA

* Correspondence: Andrey Grigoriev, Department of Biology, Center for Computational and Integrative Biology, Rutgers University, Camden, New Jersey 08102, USA
andrey.grigoriev@rutgers.edu

# These authors provided equal contribution to this work.

Keywords: RISC, Argonaute, aging, smallRNAs, ncRNAs, tRNAs, tRFs.

Contribution

My contribution to this study involves developing the pipeline to analyze the NGS data. It also involves identifying patterns and trends in the context of Ago and age related context that was then used in downstream analysis (Figure 3-6). In particular, I performed sequence alignments, tRF isoform identification and characterization (Figure 1 and Figure 2). Additionally, also did comparisons between different small RNA populations and amongst additional datasets (data not shown). Finally, I also drafted and edited the manuscript.

ABSTRACT

Background: Development of sequencing technologies and supporting computation enable discovery of small RNA molecules that previously escaped detection or were ignored due to low count numbers. While the focus in the analysis of small RNA libraries has been primarily on microRNAs (miRNAs), recent studies have reported findings of fragments of transfer RNAs (tRFs) across a range of organisms.

Results: Here we describe Drosophila melanogaster tRFs, which appear to have a number of structural and functional features similar to those of miRNAs but are less abundant. As is the case with miRNAs, (i) tRFs seem to have distinct isoforms preferentially originating from 5' or 3' end of a precursor molecule (in this case, tRNA), (ii) ends of tRFs appear to contain short "seed" sequences matching conserved regions across 12 Drosophila genomes, preferentially in 3' UTRs; (iii) tRFs display specific isoform

loading into Ago1 and Ago2 and thus likely function in RISC complexes; (iii) levels of loading in Ago1 and Ago2 differ considerably; and (iv) both tRF expression and loading appear to be age-dependent, indicating potential regulatory changes from young to adult organisms.

Conclusions: We found that Drosophila tRF reads mapped to both nuclear and mitochondrial tRNA genes for all 20 amino acids, while previous studies have usually reported fragments from only a few tRNAs. These tRFs show a number of similarities with miRNAs, including seed sequences. Based on homology with conserved Drosophila regions we systematically identified such seed sequences and their possible targets with matches in the 3'UTR regions. Strikingly, the potential target genes of the most abundant tRFs show significant Gene Ontology enrichment in development and neuronal function. The latter suggests that involvement of tRFs in the RNA interfering pathway may play a role in brain activity or brain changes with age.

BACKGROUND

Transfer RNAs (tRNAs) have been traditionally seen as key players in protein translation, but recently there have been multiple attempts to understand them as regulatory molecules (Gong et al. 2013, Li et al. 2008, Wei et al. 2012). There are two main species of tRNA-derived small RNAs that are categorized based on length and biogenesis, including tRNA-derived small RNAs (tsRNAs, ~28-40 nt) and tRNA-derived fragments (tRFs, ~16-24nt) (Thompson and Parker 2009a, Tuck and Tollervey 2011). In this study, we focus specifically on tRFs, represented by three different fragment types

based on cleavage pattern. One type is produced from the tRNA 5' part (ending before the anticodon loop), while the other two types originate from the 3' region, and contain either multiple uracils or a CCA modification at the end (Lee et al. 2009, Li et al. 2008, Sobala and Hutvagner 2007). There have been various attempts to determine the biogenesis pathways and potential cleavage events that make these tRFs distinct from one another (Anderson and Ivanov 2014, Gebetsberger et al. 2012, Gong et al. 2013, Lee et al. 2009, Miyoshi, Miyoshi, and Siomi 2010, Sobala and Hutvagner 2007, Loss-Morais, Waterhouse, and Margis 2014).

Previous studies have demonstrated regulatory function of these tRFs by postulating that they bind and repress mRNAs in a fashion similar to microRNAs (miRNAs) and at times even compete with miRNAs (Fischer et al. 2011, Garcia-Silva et al. 2012, Gebetsberger et al. 2012, Ivanov et al. 2011, Li et al. 2008, Sobala and Hutvagner 2007, Tuck and Tollervey 2011). It is unclear if they act like plant miRNAs that are fully complementary to their targets, or like animal miRNAs that have a specific pairing “seed’ region. Conflicting models of such seed regions have been proposed. One of them has suggested a traditional miRNA-like silencing based on complementarity of the 5' seed sequence of a tRF to a short sub-sequence within a 3' UTR of a transcript (Miyoshi, Miyoshi, and Siomi 2010); another has shown that the last 8-10 nucleotides (nts) on the 3’ end of the tRF in the 5’ portion of the full tRNA are responsible for mRNA repression (Wang et al. 2012).

In the present study, we elucidated tRF/mRNA pairing further by developing a computational approach and a pipeline analogous to miRNA seed-pairing studies

(Grimson et al. 2007, Lewis, Burge, and Bartel 2005, Lewis et al. 2003). Searching for conserved regions among 12 Drosophila species, we predicted tRF seeds and hybridization patterns similar to that of miRNAs. In a striking parallel to the experimental observations, we also found cases of both 3'- and 5'-located potential seeds for different tRF species. Some of the functions of tsRNAs/tRFs have been connected to stress, metabolism, and differentiation suggesting the species may be critical regulatory molecules for proper cellular growth and maintenance (Anderson and Ivanov 2014, Fischer et al. 2011, Gebetsberger et al. 2012, Haiser et al. 2008, Maute et al. 2013, Sobala and Hutvagner 2007, Wang et al. 2012, Wei et al. 2012, Miyoshi, Miyoshi, and Siomi 2010). Expanding this functional catalog in our study, we observed significant enrichment in neuronal function and development among potential targets of the prominent tRF isoforms.

We further analyzed the association with age. Recent studies have highlighted that miRNAs are associated with the aging process, showing differential isoform expression and differential RISC loading of specific miRNAs with age, related to modifications on the 3' end, including untemplated additions, 2'-O-methylation or imprecise Drosha/Dicer cleavages (Abe et al. 2014, Liu et al. 2011). Here, we present a follow-up computational analysis of the same deep-sequencing libraries, this time focusing on tRFs originating from multiple tRNAs. In addition to the in silico prediction of seed regions, we examined changes in individual tRF isoforms with age. This unexpectedly revealed diverse patterns, resembling those of miRNA and suggesting that tRFs may impact age-associated events, while simultaneously being modulated with age. Taken together, these

findings confirm that despite the lower counts in deep-sequencing experiments, tRFs represent not degradation products but potentially important players in Argonaute pathways, increasing our understanding of these regulatory molecules.

RESULTS

Using four different D. melanogaster small RNA libraries, including co-immunopreciptations of Ago1 and Ago2 in flies aged 3 days and 30 days (Abe et al. 2014), we observed striking patterns of age-dependent expression, structure and preferential loading of tRFs into RISC complexes. Following the similarity of tRF features with miRNAs, we predicted potential targets for further experimental validation that would be the ultimate test of the biological functionality of tRFs.

Read Distributions of tRNA Fragments are Similar to miRNAs

The read distributions mapping to known miRNAs usually show an asymmetry favoring the mature arm of a given miRNA stem-loop sequence. This is usually is demonstrated by observing a high relative frequency of the reads aligning to one of the arms (5' or 3'). At times, we also observe reads that originate from the middle or loop section, which is inferred by a very low frequency of reads mapping to the middle section of the RNA molecule. This type of visualization is particularly useful because it may shed light into potential 5' or 3' modifications, which may include alternative cleavage sites, deletions, non-templated additions, and RNA editing events (Chen et al. 2011, Ghildiyal et al. 2009, Kim 2005).

We investigated whether tRF-tRNA alignments displayed similar patterns to miRNAs in the read distributions. First, we found that tRF reads, which were more abundant in the Ago2 libraries, mapped to >100 nuclear and mitochondrial Drosophila tRNA genes covering the whole spectrum of 20 amino acids. This is in contrast to previous studies, which have usually reported fragments from only a few tRNAs (Maute et al. 2013, Peng et al. 2012, Wang et al. 2012, Miyoshi, Miyoshi, and Siomi 2010). We also observed multiple isoforms of the same tRF being expressed. Interestingly, these mappings showed very specific patterns: the reads typically aligned to either the 5' or 3' region of the tRNA molecule, and often had identical start positions or presumed cleavage sites (see below). In other words, the distribution of reads that mapped appeared non-random and precise, strongly suggesting that their source was not degradation but rather a targeted biological process.

All detected Drosophila tRFs and their relative read distributions in visual format can be found on our website [50]; here we illustrate the findings with the two examples of tRFs of different level of abundance, AlaAGC and MetCAT tRFs (Fig. 1). As was typical for most tRFs, the read distributions invite comparisons to a canonical miRNA structure, suggesting that specific cleavage mechanisms may be at work. We observed clearly defined boundaries for 5' and 3' regions. The uneven read distribution allows one to speculate that, in case of AlaAGC (Fig. 1A), the 5' arm is the analog of a miRNA mature and/or functional strand, while the 3' arm is similar to a passenger strand (that would eventually be degraded). The low frequency of reads mapping to the middle region is akin to miRNA loop regions. MetCAT is an example of the opposite case of prevalent

read counts in the 3'-end (Fig 1B). Generally, the majority of reads showed a miRNA-like asymmetric distribution by aligning to either the 5’ or 3’ region of the tRNAs.

Age-associated Global Shift of Ago1 vs Ago2-loaded tRFs

A number of further similarities to miRNAs were suggested by the association of the tRFs with the Argonaute proteins (Ago1 and Ago2) of the two RISC complexes. Previously, we have analyzed Ago1 and Ago2 loading of microRNAs and found age-specific patterns (Abe et al. 2014).

As with miRNAs, we observed that the total levels of Ago-loaded tRFs changed with age. In Ago1 the normalized read counts for 3 days and 30 days stayed relatively constant at ~5,000. In contrast, in Ago2 there was a 4-fold increase (from 5,000 to ~20,0000 normalized total read counts) between 3 to 30 days. Amongst tRFs with counts >100, 8 were downregulated and 4 upregulated in Ago1, while all 40 Ago2-associated tRFs were upregulated with age. These findings raise the possibility of a functional importance in an age-related manner, especially in those tRFs that were upregulated with age in Ago2.

Further investigating this result, we determined whether the differences in loading into Ago2 reflected an increased association of specific isoforms over others. This particular phenomenon is seen in miRNAs (Abe et al. 2014), so it was of interest to assess if there was a similarity in tRF behavior. We first identified two tRFs, GluCTC and AspGTC, that displayed multiple isoforms in both the Ago1-IP and Ago2-IP libraries and that also showed differential loading with age, with the most abundant isoform changing two-fold

or more (Fig. 2). For GluCTC we observed the same isoform, the 25mer, being loaded onto both RISC complexes, but in Ago1 it showed a decrease with age, while in Ago2 it showed an increase with age, hinting at a mechanism that actively partitions these fragments at the loading step in the biogenesis pathway (Fig. 2A-B). In the case of AspGTC, the isoform (29mer) that is most abundant was not detected at all in Ago1, while it was readily loaded into Ago2, which also showed increased loading with age (Fig. 2C-D).

Further, we considered loading ratios of 30 days to 3 days for each tRF. Our findings indicated that loading onto Ago2 increased at 30 days, while Ago1 loading decreased or stayed the same as at 3 days (Fig 3). Not all the tRFs are shown: e.g., Gly-related ones did not have any reads in the Ago1 libraries, and no reads were found in Ago1 for the major Ago2 isoform of AspGTC tRF depicted in Fig. 2. In several cases, distinct fragments from different tRNA genes with the same anticodon were detected, e.g., for GluCTC. When tRF sequences allowed us to distinguish such tRNA genes, we named them tRNAgene-1, tRNAgene-2, an so on (note that a union of all GluCTC isoforms in Fig. 2 corresponds to GluCTC-2 in Fig. 3). When tRF sequences could be assigned to more than one of such tRNA genes, we assumed all of these genes contributed equally to the observed tRF counts.

We then examined tRFs that were both Ago1- and Ago2-loaded in order to ascertain any age preference. We specifically looked at tRFs at the two different time points and compared their ratios in Ago2- and Ago1-associated libraries (Fig. 3). At 3 days, we

observe that the ratios are either below 1 or very close to 1, with the exception of GluTTC. Thus at 3 days, there is either a preference for Ago1 or no preference at all. However, at 30 days the reverse is the case: for most tRFs we detected at least a two-fold increase in Ago2 loading. Hence, tRFs are more likely to be loaded onto Ago2 and not Ago1 in older flies, confirming an age preference amongst loaded tRFs.

We next focused specifically on the tRF species containing CCA at the 3'-end and examined their accumulation with age in both Ago1 and Ago2. Such species showed a two-fold increase in Ago1 libraries (6% to 12%) from 3 days to 30 days, and even higher in Ago2 libraries (6% to 16%), supporting the notion that fragments of mature tRNAs contribute to the global increase of loading with age.

Together, these data support the idea that the loading patterns of tRFs between Ago1 and Ago2 change dramatically with age, such that Ago2-loading of select isoforms increases, while Ago1-loading of tRF isoforms belonging to the same tRNA decreases. These results are similar to findings of age-dependent loading of miRNAs (Abe et al. 2014) and they also indicate that there may be distinct pathways for Ago loading by recognizing and partitioning specific isoforms, which may change as a function of age.

Seed Sequences in Conserved Regions

Our results suggest dynamic tRF loading into Ago1 and Ago2 with age. The mechanism of tRF action still remains unclear, but there are clues to suggest a miRNA-like pathway of execution. For example, the fragments have been detected in the cytoplasmic fraction

of cells (Wang et al. 2012), several studies have shown trans-silencing capabilities of tRFs, and the silencing of a mock mRNA fully complementary to a tRF (Ivanov et al. 2011, Jochl et al. 2008) has been demonstrated. One proposal (Miyoshi, Miyoshi, and Siomi 2010, Maute et al. 2013) suggests a traditional miRNA-like silencing based on complementarity of the 5' seed sequence of a tRF to a short sub-sequence within a 3' UTR of a transcript (Lewis et al. 2003, Lewis, Burge, and Bartel 2005, Grimson et al. 2007). Another study, however, using luciferase assays, suggests a 3' seed sequence, while ruling out a 5' and a mid-tRF seed binding (Wang et al. 2012).

To further explore the notion of mRNA targeting, we developed a computational pipeline to detect a potential location of a seed sequence (analogous to that of animal miRNAs) in the tRFs. In miRNAs, 3'-compensatory sites (Bartel 2009b) and central pairing sites (Shin et al. 2010) have been reported in addition to the most prevalent 5' seeds (Grimson et al. 2007, Lewis, Burge, and Bartel 2005, Lewis et al. 2003, Wang et al. 2012). Following the same approach used to identify the seed sequence in microRNAs (Grimson et al. 2007, Lewis, Burge, and Bartel 2005, Lewis et al. 2003, Wang et al. 2012), we used short sequence windows sliding along the tRF sequence, without any location constraints. Then we aligned these sequence windows to the conserved 5', 3' and coding sequence (CDS) regions of the 12 Drosophila genomes (Clark et al. 2007). Following our hypothesis that tRFs may harbor miRNA-like short seed sequences, 7-nt windows showed good discrimination between conservation levels of 5' and 3' ends of the most tRFs we analyzed (Fig. 4).

The results for the most abundant Ago2-loaded Gly-derived tRFs detected in our studies strongly supported the 3' seed location (Fig. 4A-B). We observed that the tRF GlyGCC 7mer located at position 12 to 18 has the highest frequency of reverse complement occurrences in the conserved regions of Drosophila genomes (regions associated with >14,000 genes in total), making it a candidate seed sequence. A very similar tRF, GlyTCC (attcccggccgaCgcacca), contained a one nt difference to GlyGCC (attcccggccgaTgcacca) and had a candidate seed shifted one nt towards the 5’ end (from position 11 to 17) (Fig 4B).

We found no overlap between the lists of D. melanogaster transcripts with matches to the seeds of GlyGCC compared to GlyTCC. Thus, although a single nucleotide difference in/near the seed region may influence tRF targeting and hybridization, it is remarkable that a very different set of conserved sequence matches/potential targets still corresponds to the same 3' location of the seed sequence. Though many tRFs showed a peak similar to those in Fig 4A, we noted that a few tRFs showed such peaks in the 5’ region, suggesting a 5’ seed targeting. For example, in the tRF mt:SerAGC the 7mer window matches peaked at the 5’ end of the sequence (Fig. 4C), as opposed to a 3’ maximum found in the Gly-related tRFs. Thus we also observed potential seed on both 5' and 3' ends of tRFs, in parallel to what was detected experimentally.

Potential Targeted Regions in mRNA

To find potential targeted regions, we separately analyzed the conserved 5', 3' and CDS parts of genes in 12 Drosophila genomes. Per unit of length the 3’ UTR regions matched

the potential tRF seeds most frequently (Fig. 5) suggesting a prevalence of a 3' UTR targeting mode. This further supports the hypothesis that tRFs may behave similar to miRNAs.

We observed significant enrichment of the 3'UTR for mt:SerAGC seed matches in the D. melanogaster genome ($p < 0.001$), both among random heptamers and among reshuffled nucleotides comprising the seed. Gly tRF seed matches, with less extreme AT-richness, did not show such enrichment. However, we note that shuffling of the seed sequence is not an ideal random model and statistical testing of the tRF seed regions is complicated by the fact that a tRNA sequence is under multiple selective constraints for its structure and function related to translation (and furthermore different from the constraints of a miRNA).

We also scanned for nearly perfect complementary matching between full-length tRFs and 3’ UTRs, which would inform us if some of these tRFs acted like plant miRNA. This analysis, however, yielded no significant results, suggesting that the tRF binding mode may be more consistent with animal miRNAs.

Assuming the latter (animal-like) binding mode, we observed a variety of seed sequence matches in the conserved fly genome regions. As with miRNAs, there were 7mer-m8, 7mer-1a and 8mer-1a match types. These types have been studied and confirmed previously for miRNAs (Grimson et al. 2007) and are as follows. 7mer-m8 is a match of 7 nts (Fig. 6A and Fig. 6B). 7mer-1a and 8mer-1a refer to matches of first 6 (Fig. 6C,

GlyGCC) or 7 (Fig. 6C, mt:SerAGC) nts of the seed, respectively, followed by an extra A (added to the elongated match region in the Fig. 6C). All three illustrated targets (Fig. 6A-C) possess 3 'UTR regions highly conserved amongst all 12 Drosophila genomes analyzed.

Notably, some seeds showed overlap with the seed of either another tRF or a miRNA (Fig. 6A and Fig. 6C). For example, both GlyGCC and mir-277 seeds overlap by 5 nts and this sometimes led to their complementarity against the same target (Fig. 6A). Such overlaps could theoretically lead to competition of tRFs and miRNAs for the same targets, potentially adding another layer of complexity to the regulatory processes.

As demonstrated by our results, there is clear evidence that tRFs interact and are loaded onto Argonaute proteins and may target primarily the 3’ UTR regions of mRNAs, which allows us to hypothesize a potential post-transcriptional regulatory mechanism similar to that of miRNAs. The fact that the candidate seeds aligned predominantly to the 3’ UTRs suggests that one of the mechanisms for suppression may be translational inhibition. Alternatively, some tRFs may employ mRNA cleavage for regulation, since we observed CDS regions that also aligned to our candidate seeds (Grimson et al. 2007, Lewis, Burge, and Bartel 2005, Lewis et al. 2003).

Gene Ontology Analysis of Potential Targets

Given the difference in seed localization, we predicted targets for the divergent cases of the Gly and mt:SerAGC tRFs. Following the link between the Ago-loading change of

miRNA and brain degeneration with age (Abe et al. 2014), we assessed whether targets of these tRF were also associated with a particular biological process. Using the identified seed sequences, we sought targets for the tRFs in D. melanogaster genome based on perfect matches. We then conducted a gene ontology (GO) enrichment analysis using the AmiGO 2 software (Ashburner et al. 2000) to understand the nature of the predicted targets.

Stringent criteria for enrichment revealed several interesting trends. Notably, neuronal and developmental processes were the most dominant among the significantly enriched terms (p-value < 0.001) belonging to the GO category "biological process". In particular, for GlyGCC we observed 52% of enriched GO terms related to development and 15% related to neuronal function, while for mt:SerAGC these numbers were 39% and 12%, respectively (Supp Table 1). In the GO analysis, the most populated process terms (if one counts potential targets, described by these terms) are often generic ones, like "biological process" or "biological regulation". For both of these tRFs, the most populated GO terms after the generic ones were GO:0032502 (developmental process) or GO:0048856 (anatomical structure development). Pertinent to the tRF involvement in the neuronal regulation, synapse- or axon-related GO terms accounted for 20% (in mt:SerAGC) to about half (in GlyGCC) of the significantly enriched terms (p-value < 0.001) in the category "cellular localization" (Supp Table 2). The targets, exemplified in Fig. 6, belong to these GO categories, e.g., Dlar, a targeted gene of mt:SerAGC (Fig. 6B) is a conserved member of the tyrosine phosphatase family with a fundamental role in axon targeting/development and organization of actin filaments (Prakash et al. 2009, Krueger

et al. 1996); see Discussion). For the category "molecular function", terms related to DNA and RNA binding (with variations including regulatory region or nucleotide binding) were frequently enriched for mt:SerAGC and GlyGCC (Supp Table 3).

DISCUSSION

In this bioinformatics report we extensively characterized tRFs found in Ago1 and Ago2 IP libraries from Drosophila to reveal expression and loading patterns in the context of age. We also identified potential targets and a likely mode for targeting.

We identified tRFs in both Ago1 and Ago2 co-immunoprecipitated libraries, indicating miRNA-like functionality of loading of these tRFs into RISC complexes. Alignment to the mature tRNA sequence revealed a high read-depth on one side of the tRNA molecule and size distributions of 16-30 base pairs in length, which suggests a similar structural motif as miRNAs. Although the library was size-selected for these distributions, we observed very precise boundaries of tRFs, strongly suggestive of a biological process rather than random degradation responsible for their generation.

By examining age-associated patterns of tRF expression, distinct isoforms changed in an age-dependent manner in Drosophila. For example, for GluCTC we observed the same isoforms present in both Ago1 and Ago2 libraries, but we saw an increase in individual isoforms in Ago2, and a decrease in Ago1, especially for most abundant or major isoform. Additionally, the major isoform of AspGTC in Ago2, was not present at all in Ago1 (see Fig. 2). These types of change are correlated with a shift in loading of these

fragments into Ago2 vs Ago1 with age. Thus, the partitioning of multiple tRFs between Ago1 and Ago2 appears to be a coordinated process modulated with age in Drosophila (see Fig. 3).

One explanation for this phenomenon is that the cells are adjusting regulatory efficiency for upcoming age-associated stresses. Since Ago2-mediated translational silencing causes retention of the polyA tail (Iwasaki, Kawamata, and Tomari 2009), Ago2-association might make it possible to respond to age-associated internal or external stimuli more rapidly and effectively by allowing for re-activation of target mRNAs.

For miRNAs it is known that there are specific features associated with differential loading, such as methylation of the 3' end. 2'-O-methylation is known to stabilize Ago2-loaded miRNAs and these small RNAs impact brain degeneration and lifespan (Abe et al. 2014, Ghildiyal et al. 2009, Iwasaki, Kawamata, and Tomari 2009, Liu et al. 2011, Okamura, Liu, and Lai 2009, Ren et al. 2014). Since increased miRNA Ago2-loading is associated with 3' protection by methylation (Abe et al. 2014), it is tempting to speculate that Ago2-loaded tRFs may also be protected at the 3' end via a similar mechanism. To investigate this possibility, we scanned for tRFs in an oxidized RNA library (Ghildiyal et al. 2009), highly enriched for 2'-O-methylated small RNAs. Using the threshold of >100 raw counts, we could only detect one tRF in that library. Strikingly, this was the 25bp isoform of GluCTC, which also showed increased Ago2 loading with age in our study (see Fig. 3). Thus 2'-O-methylation may be a protective mechanism for accumulating tRFs, although more experimentation is required for confirmation.

Increased Ago2 loading of miRNAs was observed in aging flies and Ago2 mutants have been shown to develop neurodegenerative phenotypes in the study of miRNA involvement in the aging process (Abe et al. 2014). This may serve as further support of our hypothesis since in these mutants the disrupted stabilizing modification, lack of tRF RISC loading, and subsequent deregulation of the neuronal targets would further contribute to such phenotypes. One possible target of GlyGCC and mt:SerAGC (see Fig. 6C), the gene Atg8a, is intimately linked to aging pathways, e.g., the insulin/IGF-signaling pathway that mediates the lifespan in Drosophila through Smad binding (Bai et al. 2013).

Although the mechanism of silencing is still being unraveled, our results suggest a miRNA-like seed region in tRFs that is key for targeting to potential mRNA targets. While for animal miRNAs 5' seed location is most common, 3'-compensatory sites (Bartel 2009b) and central pairing sites (Shin et al. 2010) have been reported. In our examples, the Gly-associated tRF in Drosophila has a putative 3' seed region, while the mt:SerAGC tRF has a 5' seed. Thus, in parallel to experimental data showing two possible seed locations (Maute et al. 2013, Wang et al. 2012, Miyoshi, Miyoshi, and Siomi 2010), our results demonstrate that regions of conservation can be present at either the 5’ or the 3’ end in different tRFs. We also provide evidence that the 3' UTR or CDS may be where targeting occurs, allowing us to speculate that the mode of action may include translational repression or mRNA cleavage.

Drosophila Ago1 and Ago2 employ different mechanisms to silence target mRNAs and in particular Ago2 mutants show neurodegeneration and a shortened lifespan (Abe et al. 2014). The fact that most tRFs are loaded and/or show a dramatic change in loading with age in Ago2 suggests that these small RNAs may also be involved in such pathways. In this regard, it is notable that despite the difference in seed localization (and no common targets), putative targets of tRFs from both mt:SerAGC and GlyCTC are significantly enriched in developmental and neuronal functions (Supp. Table 1-3). Further, we found that these target lists overlap (with up to 29 targets) with the well-studied miRNAs mir-34, mir-277, mir-190, and mir-10. All of these miRNAs impact brain function, affecting neurodegeneration, bi-polar disorder, and schizophrenia (Liu et al. 2011, Moreau et al. 2011, Perkins et al. 2007), strengthening our hypothesis of tRF influence on the brain and age-related events. An overlap of the tRF seed with that of mir-277 is of importance, as it may relate one of the most abundant tRFs (GlyGCC) to brain deterioration, since mir-277 has been reported to modulate neurodegeneration (Tan et al. 2012).

Amongst the common targets of GlyGCC and mir-277, we observed Dlg (FBgn0001624), coding for the Drosophila discs large tumor suppressor protein (see Fig 6A). The mechanism of regulation of this gene would be of interest since it has been previously associated with neuron development (Zhang et al. 2007, Kumar et al. 2009) and it also shows homology with a human tumor suppressor protein (Azim et al. 1995). Another common target, Toll-7 (FBgn0034476), may also be of significance, since it acts as a neurotrophin receptor and neurotrophism is only starting to be elucidated in insects (McIlroy et al. 2013).

Some of the tRF targets in the significantly enriched GO categories are closely related to the RNA regulatory pathways, e.g., Fmr1 (FBgn0028734, a homolog of the fragile X mental retardation 1 gene in human). This is an RNA-binding protein that interacts with the RISC complex itself and is necessary for proper development (Garber et al. 2006, Jin et al. 2004, Ishizuka, Siomi, and Siomi 2002). Of note, this gene is located in the Drosophila genome in the immediate vicinity (a few hundred basepairs) of mir-34 and mir-277, hinting at a potentially deeper regulatory connection.

CONCLUSIONS

This is the first time such a detailed analysis has been performed on tRFs. We developed a robust pipeline to identify candidate "seed" regions that clearly showed a stronger binding pattern based on specific positions, restricting it to the 5’ or 3’ end and a binding preference for 3’ UTRs. The results reveal tRFs features that in many respects resemble structural and functional properties of miRNAs and strongly suggest that these small RNAs are not simply tRNA degradation products, but are specific, biologically-generated species. The targets predicted with candidate seeds showed enrichment in processes related to neuronal function and development, hinting at the biological significance of these tRF molecules. Thus, the trends observed with tRFs likely represent bona fide targeted processing of tRNAs, and the tRF association with different RISC complexes in the context of age may reflects an important regulatory function.

ACKNOWLEDGMENTS

This work was funded in part by the Ellison Medical Foundation (to N.M.B.), and the National Science Foundation (grant DBI-1126052 to A.G.).

COMPETING INTERESTS

The authors declare that they have no competing interests.

AUTHORS' CONTRIBUTIONS

AN performed sequence alignments, tRF and isoform identification and characterization, drafted the manuscript. SK identified seeds and targets and performed gene ontology analysis. KS with participation of AN, SK and AG created the public repository of the data. NB provided the library sequences, helped to draft the manuscript. AG conceived the study, participated in its design, execution and coordination, and drafted the manuscript. All authors read and approved the final manuscript.

Figures and Tables

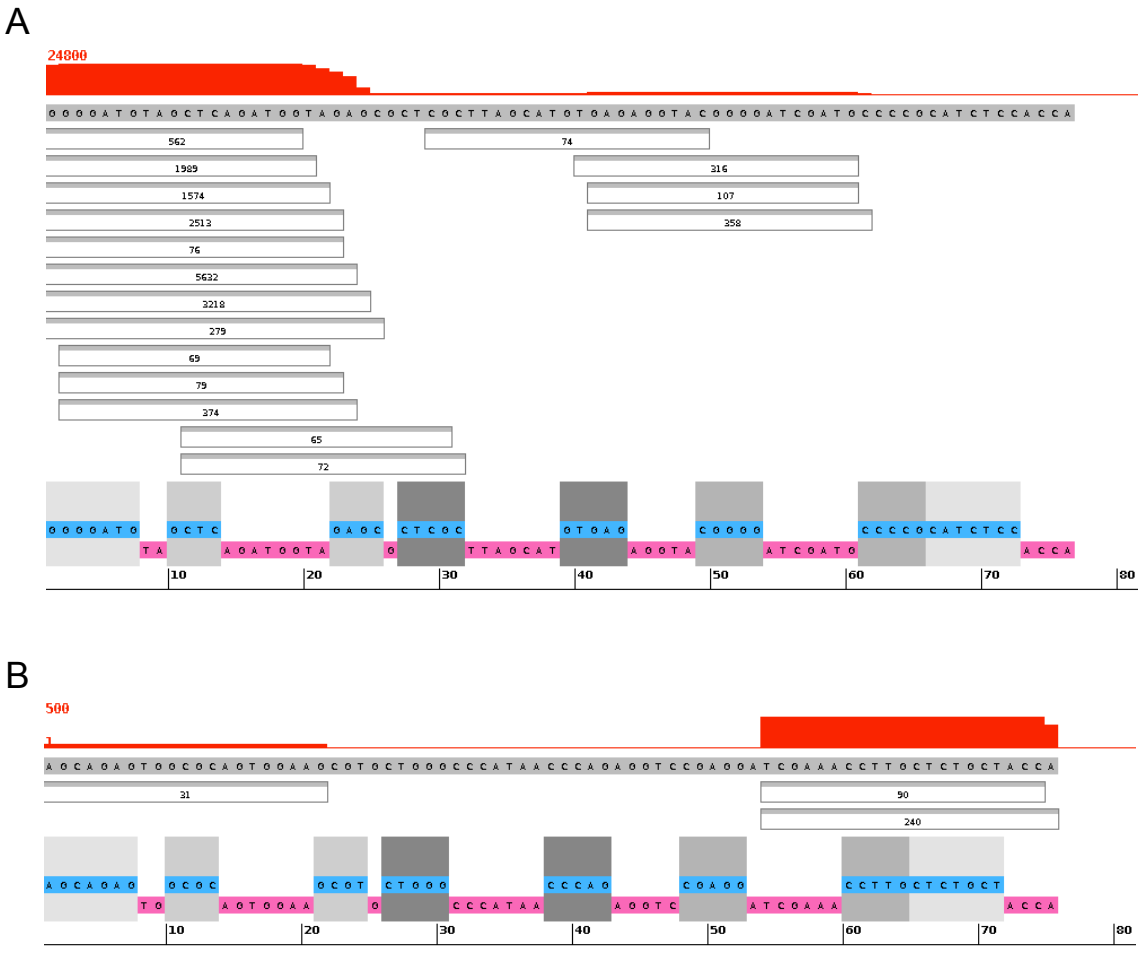


Figure 5.1. Fragment distribution pattern that aligns to tRNA-Ala in the Ago2, 30 days library and MetCAT in Ago1, 30 days.

Screenshot of our RNA display (Naqvi, Cui, and Grigoriev 2014). Sequence at the bottom with the magenta background indicates single-stranded (loop) regions in the tRNA molecule, while the cyan background and matching grey boxes indicate stems. The red on top indicates read depth coverage of specific regions of the tRNA. Reads (boxes in the middle) with counts of at least 1% of the most abundant read are displayed; lower count reads are omitted for compact visualization.

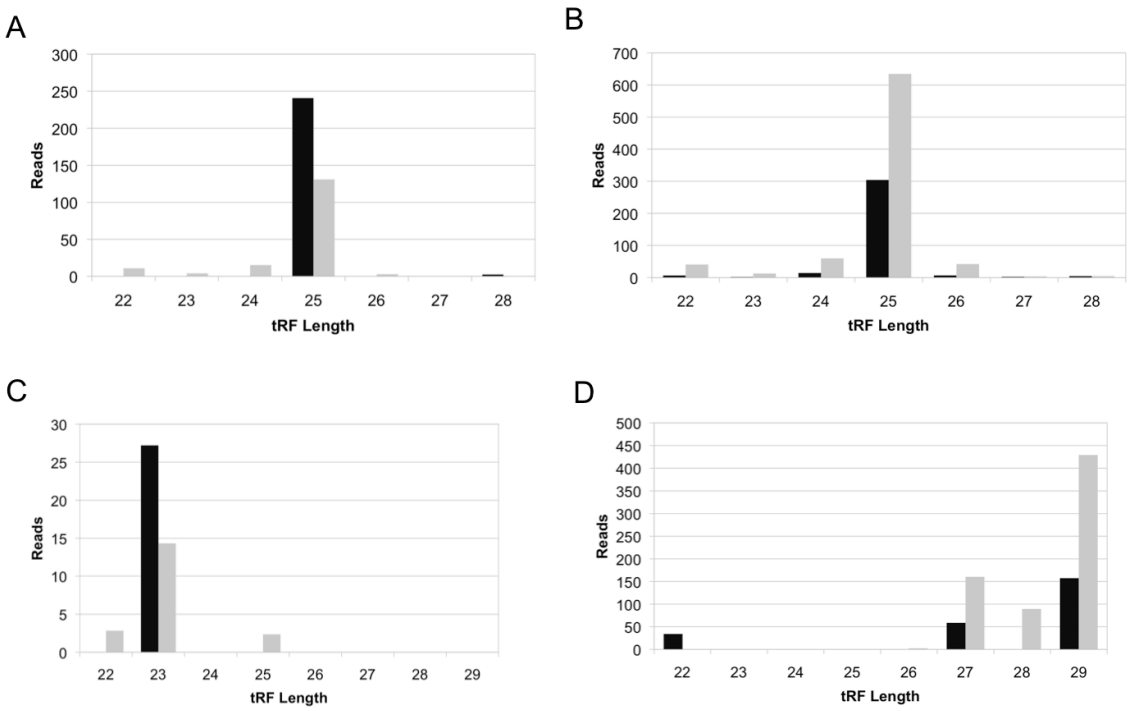


Figure 5.2. Distinct tRF Isoform Changes with Age.

GluCTC in (a) Ago1-IP and (b) Ago2-IP, (c) AspGTC in Ago1-IP and (d) Ago2-IP. Both tRFs show a decrease in Ago1, but an increase of specific isoforms in Ago2 with age. Black bars represent normalized counts at 3 days, while grey bars represent normalized counts at 30 days. The same GluCTC is also presented in Fig. 3 (GluCTC-2).

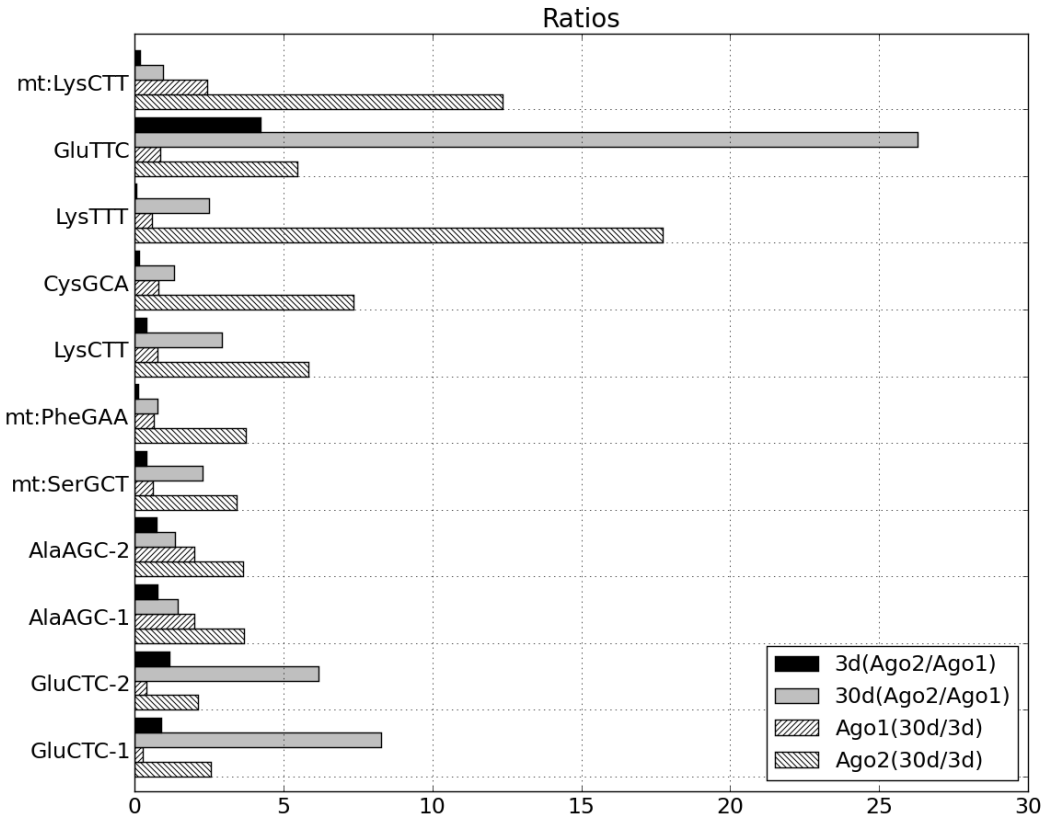


Figure 5.3. Differential and Preferential Loading.

Plot of abundant tRFs that are present in all four libraries. Plots show relative ratios of reads: Ago2 to Ago1 in 3 days (black) and 30 days (gray); 30 days to 3 days in Ago1 (slash) and Ago2 (back slash).

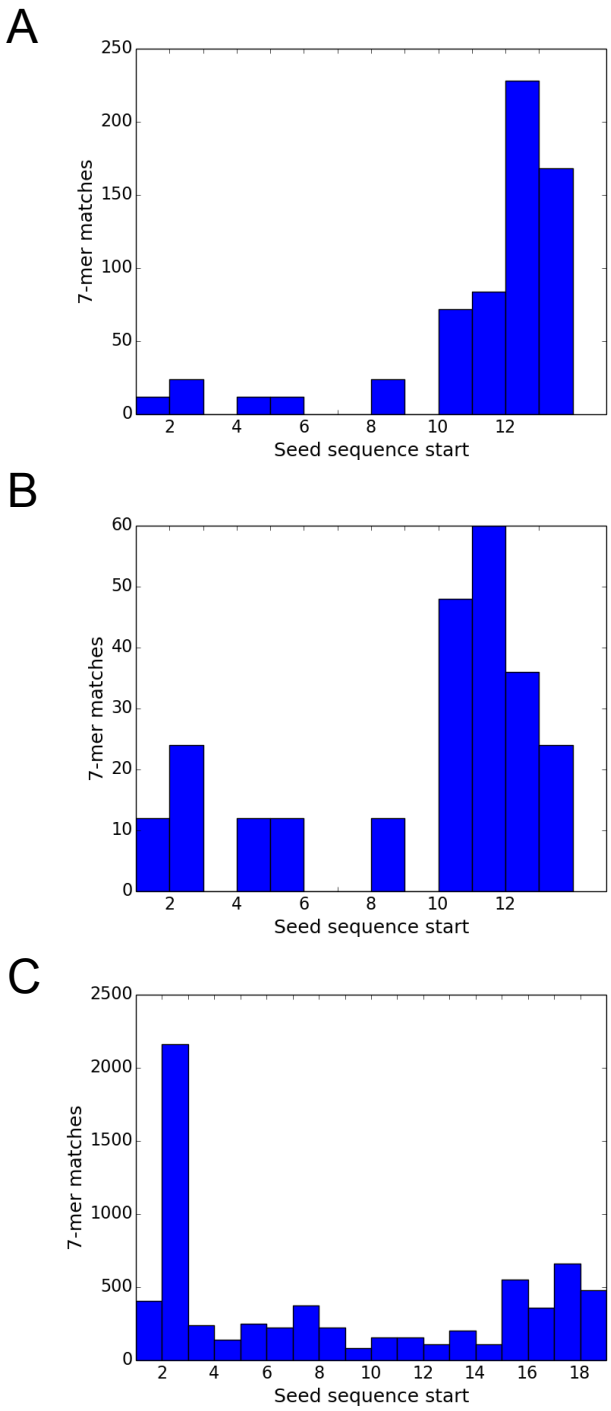


Figure 5.4. Candidate Seed Regions for tRFs.

The numbers of conserved sequence matches in the 3' UTR regions are plotted vs window start positions of 7mer windows of (a) GlyGCC, (b) GlyTCC tRFs and (c) mt:SerAGC tRF in Drosophila.

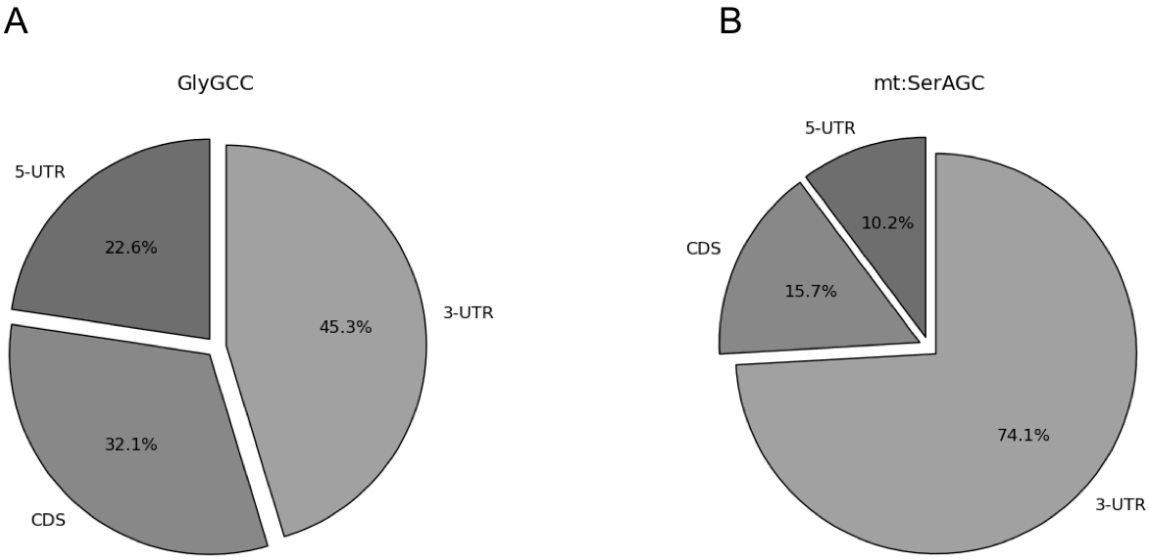


Figure 5.5. Percentage of Seed Alignments by Region.

The percentages by region (5’ UTR, 3’ UTR and CDS) of the matches to the most abundant 7mer shown in Fig. 4A and 4C.

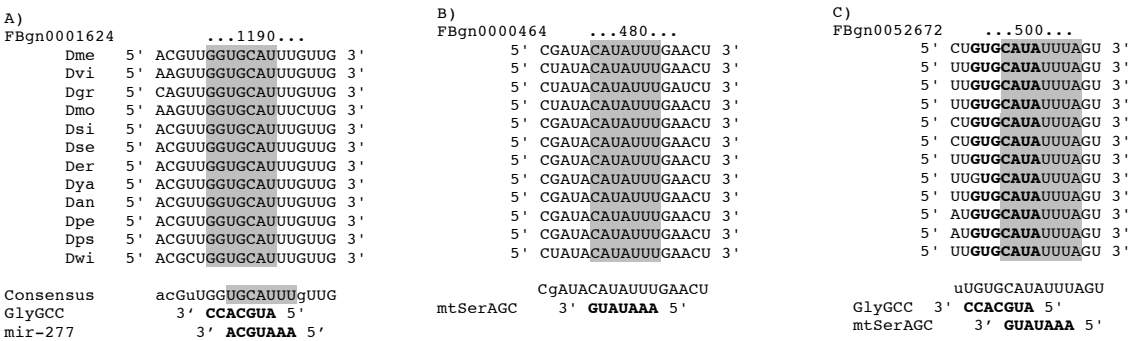


Figure 5.6. Conserved Seed Region Matches.

Grey highlights and bold text indicate seed complementarity to conserved (12 Drosophila genomes) 3' UTR regions. Targeted genes with overlapping coordinates in the genome are shown on top. (a) Both GlyGCC and mir-277 having a 7mer-m8 match, (b)

mt:SerAGC 7mer-m8 match and (c) GlyGCC having a 7mer-1a match and mt:SerAGC having a 8mer-1a match, with additional A for the 1a matches are also highlighted (mt:SerAGC) or bolded (GlyGCC).

List of supplementary tables

CHAPTER 6 MATERIALS AND METHODS

Chapter 2 Materials and methods

Edit calling and filtering

Publicly available small RNA-Seq libraries were downloaded from the NCBI SRA database. Before clipping and mapping, we removed (i) the 3 adaptors from the reads, (ii) all reads shorter than 16 nucleotides and (iii) those with one or more "N" nucleotides present. The remaining reads were then mapped against the stem-loop sequences obtained from mirBase (Griffiths-Jones 2006) using Bowtie ver. 1.0 (-v 1, -a options) (Langmead et al. 2009). Reads that mapped to the microRNA stem-loops with one mismatch were then mapped against the whole Drosophila genome (dm3) using Bowtie. In order to filter out potential noise and artifacts, all mismatched reads that also successfully aligned elsewhere in the genome with at most one mismatch were removed. These steps were performed for each of the ten libraries utilized in this study. We then quantified observed mismatches for every position in all known microRNAs in Drosophila.

To limit the experimental artifacts, we used specific thresholds for read count (>20) and editing rate (>10%), multiple library and sequence mappings (i.e. different sequences from the same library having the same position edited and different sequences from different libraries having the same position edited), and other similar measures. While we cannot guarantee that we removed all sequencing artifacts, they are significantly reduced with such filtering.

Evaluating editing events in structural context

MicroRNA secondary structure annotation was taken from mirBase (Griffiths-Jones 2006, 2010). For our random model, with editing independent of the secondary structure, we determined the expected number of edits occurring in bulge (E_{ss}) and double-stranded regions (E_{ds}) as follows:

$$Eds=S*Nds/(Nds+Nss)$$

$$Ess=S*Nss/(Nds+Nss)$$

where S is the total number of edits observed and N_{ds} and N_{ss} are the total numbers of nucleotides in double- and single-stranded regions, respectively. We then obtained ratios of observed vs expected events for these region types and did not detect any significant difference, with a slight prevalence of the observed single-stranded edits.

Chapter 3 Materials and methods

Fly Stocks and Culture

Flies were grown in standard cornmeal molasses agar medium with dry yeast, at 25°C unless otherwise specified. General stock lines and GAL4 driver lines were obtained from the Drosophila Stock center at Bloomington. nbrf02257 was obtained from the Exelixis collection (Harvard University). Fly transgenics were generated by standard procedures (Genetic Services, Inc).

Constructs

Fly genomic DNA was prepared from whole flies with the Puregene DNA purification kit

(Qiagen). Using genomic DNA as the template, a 286bp of miR-34 genomic sequence was amplified by PCR (primers: 5'-CCG TTA CAC ACG ACTA TTC TCA AT-3'/5'-CCA TCT GAT ACA GGT CCT ACA TTT TCT AAA A-3'), and used to generate a miR-34 pUAST construct. To generate Nbr constructs, PCR amplification was conducted using single stranded cDNA as the template, with primer pairs of HA-Nbr (5'- GAA TTC ATG TAC CCA TAC GAT GTT CCA GAT TAC GCT GCA CGC AAG AGC CAC ATG-3'/5'- GGT ACC TCA CTT AAC ATG GGC ACC CCG). PCR products were then cloned into the pRmHa3 vector.

mRNA Northern and Small RNA Northerns

Total RNA was isolated from cells or flies using Trizol reagent (Invitrogen) according to manufacturer's protocol. For mRNA Northern, 5?g RNA was run on a 1% MOPS/formaldehyde gel, and transferred onto nylon plus (Northernmax, Ambion). The

RNA blots were then hybridized following standard procedures at 68°C, with prehybridization (~ 1 hr), hybridization (~ 12 hr or overnight) with P32 labeled probe, washed and exposed to Phosphoimager (Amersham). RNA probes were used that were made by in vitro transcription of cDNA templates using Maxiscript-T7 in vitro transcription kit (Ambion), supplemented with P32-labled UTP. The cDNA templates were prepared from total RNA of DL1 cells by one-step RT-PCR (SuperScript One-Step RT-PCR with Platinum Taq, Invitrogen, CA), with primers: T7-nbr (5'-GAATTCATGGCACGCAAGAGCCACATG-3'/ 5'-GAT AAT ACG ACT CAC TAT AGG GAG AGG CTT CAG AAT GAG CTC CAG-3') and18S rRNA loading control

(5’-GAT AAT ACG ACT CAC TAT AGG GAG A-3’/ 5'-AGG GAG CCT GAG AAA CGG CTA CCA CAT CTA AGG AAT CTC CCT ATA GTG AGT CGT ATT ATC -3’). For small RNA Northerns, 3-15ug of RNA was fractionated on a 15% Tris-UREA gel (NuPage) with 1XTBE buffer. The transfer was performed with 0.5X TBE buffer. Prior to hybridization, the RNA blots were first prehybridized with Oligohyb (Ambion), and then incubated with radioactive labeled RNA probes for ~12 hr to overnight at 50°C. RNA probes were used, and made by in vitro transcription of oligo templates using Maxiscript-T7 in vitro transcription kit (Ambion), supplemented with P32-labeled UTP. Oligo DNA templates were prepared by annealing two single stranded DNA oligos into duplex (99°C 5min and cool down to room temperature). Oligos used were miR-2b-1 (5’-GAT AAT ACG ACT CAC TAT AGG GAG A-3’/5'-AAA AAA TAT CAC AGC CAG CTT TGA GGA GCT CTC CCT ATA GTG AGT CGT ATT ATC-3’); miR-3 (5’-GAT AAT ACG ACT CAC TAT AGG GAG A-3’/5'-AAA AAA TCA CTG GGC AAA GTG TGT CTC ATC TCC CTA TAG TGA GTC GTA TTA TC-3’); miR-7 (5’-GAT AAT ACG ACT CAC TAT AGG GAG A-3’/5'-AAA AAA ATG GAA GAC TAG TGA TTT TGT TGT TCT CCC TAT AGT GAG TCG TAT TAT C-3’); miR-10 (5’-GAT AAT ACG ACT CAC TAT AGG GAG A-3’/5'-AAA AAA ACC CTG TAG ATC CGA ATT TGT TTC TCC CTA TAG TGA GTC GTA TTA TC-3’); miR-11 (5’-GAT AAT ACG ACT CAC TAT AGG GAG A-3’/5'-AAA AAA ACA TCA CAG TCT GAG TTC TTG CTC TCC CTA TAG TGA GTC GTA TTA TC- 3’); miR-12 (5’-GAT AAT ACG ACT CAC TAT AGG GAG A-3’/5'-AAA AAA TGA GTA TTA CAT CAG GTA CTG GTT CTC CCT ATA GTG AGT CGT ATT ATC); miR-31b (5’-GAT AAT ACG ACT CAC TAT AGG GAG A-3’/5'-AAA AAA TGG CAA GAT GTC GGA ATA GCT GTC TCC

CTA TAG TGA GTC GTA TTA TC-3’); miR-34 (5’-GAT AAT ACG ACT CAC TAT AGG GAG A-3’/5’-AAA AAA TGG CAG TGT GGT TAG CTG GTT GTG TCT CCC TAT AGT GAG TCG TAT TAT C-3’); miR-34* (GAT AAT ACG ACT CAC TAT AGG GAG A-3’/5'- AAA AAA CAG CCA CTA TCT TCA CTG CCG CCT CTC CCT ATA GTG AGT CGT ATT ATC-3'); miR- 100 (5’-GAT AAT ACG ACT CAC TAT AGG GAG A-3’/5’-AAA AAA AAC CCG TAA ATC CGA ACT TGT GTC TCC CTA TAG TGA GTC GTA TTA TC-3’); miR- 190 (5’-GAT AAT ACG ACT CAC TAT AGG GAG A-3’/5’-AAA AAA AGA TAT GTT TGA TAT TCT TGG TTG TCT CCC TAT AGT GAG TCG TAT TAT C-3’); miR-210 (5’-GAT AAT ACG ACT CAC TAT AGG GAG A-3’/5’-AAA AAA ATT GTG CGT GTG ACA GCG GCT ATC TCC CTA TAG TGA GTC GTA TTA TC-3’); miR-263 (5’-GAT AAT ACG ACT CAC TAT AGG GAG A-3’/5'-GTT AAT GGC ACT GGA AGA ATT CAC TCT CCC TAT AGT GAG TCG TAT TAT C-3’); miR-277 (5’-GAT AAT ACG ACT CAC TAT AGG GAG A-3’/5’-TAA ATG CAC TAT CTG GTA CGA CAT AAA TGC ACTATCTGGTACGACA TCT CCC TAT AGT GAG TCG TAT TAT C-3’); miR-274 (5’-GAT AAT ACG ACT CAC TAT AGG GAG A-3’/5'-AAA AAA TTT TGT GAC CGA CAC TAA CGG GTA ATT CTC CCT ATA GTG AGT CGT ATT ATC-3’); miR-281rev (5’-GAT AAT ACG ACT CAC TAT AGG GAG A- 3’/5'-TGT CAT GGA ATT GCT CTC TTT GTT GTC ATG GAA TTG CTC TCT TTG TTC TCC CTA TAG TGA GTC GTA TTA TC-3’); miR-283 (5’-GAT AAT ACG ACT CAC TAT AGG GAG A-3’/5'AAA AAA TAA ATA TCA GCT GGT AAT TCT TCT CCC TAT AGT GAG TCG TAT TAT C-3’); miR-305 (5’-GAT AAT ACG ACT CAC TAT AGG GAG A-3’/5'-AAA AAA ATT GTA CTT CAT CAG GTG CTC TGT CTC CCT ATA GTG AGT

CGT ATT ATC-3’); miR-307 (5’-GAT AAT ACG ACT CAC TAT AGG GAG A-3’/5'-AAA AAA TCA CAA CCT CCT TGA GTG AGT CTC CCT ATA GTG AGT CGT ATT ATC-3’); miR-307rev (5’-GAT AAT ACG ACT CAC TAT AGG GAG A-3’/5'-AAA AAA TCA CAC CCA GGT TGA GTG AGT CTC TCC CTA TAG TGA GTC GTA TTA TC-3’); miR-315 (5’-GAT AAT ACG ACT CAC TAT AGG GAG A-3’/5'-AAA AAA TTT TGA TTG TTG CTC AGA AAG CTC TCC CTA TAG TGA GTC GTA TTA TC-3’); miR-317 (5’-GAT AAT ACG ACT CAC TAT AGG GAG A-3’/5'-AAA AAA ATG AAC ACA GCT GGT GGT ATC CAG TTC TCC CTA TAG TGA GTC GTA TTA TC-3’); miR-986 (5’-GAT AAT ACG ACT CAC TAT AGG GAG A-3’/5'- AAA AAA ATC TCG AAT AGC GTT GTG ACT GAT CTC CCT ATA GTG AGT CGT ATT ATC-3’); miR-1010 (5’-GAT AAT ACG ACT CAC TAT AGG GAG A-3’/5'- AAA AAA TTT CAC CTA TCG TTC CAT TTG CAG TCT CCC TAT AGT GAG TCG TAT TAT C-3’); esi2.1 (5’-GAT AAT ACG ACT CAC TAT AGG GAG A-3’/5'-TTG ACT CCA ACA AGT TCG CTC CTC TCC CTA TAG TGA GTC GTA TTA TC-3’) and 2S rRNA (5’- GAT AAT ACG ACT CAC TAT AGG GAG A-3’/5'-TGC TTG GAC TAC ATA TGG TTG AGG GTT GTA TCT CCC TAT AGT GAG TCG TAT TAT C-3’).

Cell culture, dsRNA Synthesis, and RNAi

Drosophila DL1 cells were grown and maintained in Schneider’s media supplemented with 10% FBS, penicillin/streptomycin and glutamine as described (Cherry and Perrimon 2004). dsRNAs for RNAi were generated as described (Boutros et al. 2004). Briefly, gene-specific primers containing T7 polymerase binding sites were used to amplify ~500

nucleotide regions within genes of interest by PCR. PCR products were used as templates for in vitro transcription using MEGAscript T7 (Ambion), and dsRNA products were purified using RNeasy columns (Qiagen). For RNAi knockdowns, cells were bathed into serum free media containing dsRNA for 45min-1h. Complete media was then added and cells were incubated for three more days.

Protein and RNA Immunoprecipitations

pMT-FLAG-Ran (Sabin et al. 2009) and pMT-FLAG-Ago1 are as described (Zhou et al. 2008). For protein immunoprecipitations, 8x106 cells were seeded into 10 cm plates and transfected the next day with 4 ?g pMT-HA-Nbr and 4 ?g of either pMT-Flag-Ran or pMT-Flag-Ago1 using Effectene (Qiagen). Plasmid expression was induced 24 hours later with 500 ?M CuSO4, and cells were collected 36 hours post-induction. Cells were processed as described (Saito et al. 2009). Briefly, cells were lysed into Buffer A+KOAc:150mM KOAc, 30mM Hepes pH 7.4, 2mM MgOAc, 0.1% NP40, 5mM DTT, PMSF, and a complete protease inhibitor cocktail (Roche). FLAG-tagged proteins were immunoprecipitated overnight at 4°C using anti-FLAG M2 agarose beads (Sigma). Beads were washed six times in Buffer A+KOAc, and bound proteins were separated by SDS-PAGE and immunoblotted with monoclonal anti-FLAG M2 antibody diluted 1:2,500 (Sigma #F3165) and HRP-conjugated anti-HA antibody diluted 1:2,000 (Roche #12013819001) as described (Cherry and Perrimon 2004).

For RNA immunoprecipitation, 1.2x107 DL1 cells were seeded into 10 cm plates in serum-free media with 12 ?g dsRNA. One hour later, complete media was added and

cells were incubated for 5 days. Endogenous Ago1 was immunoprecipitated as described (Czech et al. 2008). Cells were lysed in lysis buffer: 20mM HEPES pH 7.0, 150mM NaCl, 2.5mM MgCl2, 0.3% Triton-X, 30% glycerol, PMSF, and a complete protease inhibitor cocktail (Roche). Pre-cleared lysates were incubated with rabbit polyclonal AGO1 antibody (1:20; Abcam #ab5070) or control rabbit polyclonal GFP antibody (1:20; Invitrogen #A-6455) overnight at 4°C. AGO1 and control antibodies were isolated using protein A/G beads (1:10; Pierce #20421) for 1 hour at 4°C. Beads were then washed 6 times, 10 minutes each in wash buffer: 30mM HEPES pH 7.4, 800mM NaCl, 2mM MgCl2, 0.1% NP-40, PMSF and a complete protease inhibitor cocktail (Roche). 1 mL Trizol (Invitrogen) was added to beads following the final wash. RNA was extracted and analyzed by small RNA northern blotting.

Small RNA Deep Sequence Analysis

To make small RNA sequencing libraries, total RNA was extracted using Trizol Reagent (Invitrogen) from ~3d old nbr mutants and control 5905 flies (1:1 ratio between males and females). 40ug RNA was fractioned in a 15% TBE-Urea gel (Novex, Invitrogen), followed by gel-purification of small RNA ranging between 18 nt and 30 nt. The library was then prepared following Small RNA v1.5 Sample Preparation Guide (Illumina) with some modifications. To perform sequence analysis, adaptor sequences (5’ adapter- 5’-GTT CAGA GTT CTA CAG TCC GAC GAT C-3’; 3’ adapter 5’-ATC TCG TAT GCC GTC TTC TGC TTG AA-3’) were removed from the raw reads in the Illumina fastQ generated files using the FASTQ/A Clipper program in the fastx-toolkit (http://hannonlab.cshl.edu/fastx_toolkit/). Reads less than 16 bp or

more than 30 bp were discarded. Remaining reads were then mapped to the Drosophila genome (Flybase v5.34), and to the microRNA stemloop sequences (BDGP5.0, http://www.mirbase.org/) using Bowtie. Mapped files generated from Bowtie were formatted and analyzed by customized Perl scripts. To enrich for miRNAs affected by nbr function, the length distribution for all reads corresponding to individual miRNAs was analyzed. To be included in the analysis the miRNA had to have more than 70 reads. The ratio of the most frequent length to the sum of all other lengths in wild type was calculated, and compared to the ratio of that same most frequent form divided by the sum of all other lengths in nbr (ratio nbr/control), and ratios plotted (Fig. 4F). miRNAs whose trimming is impacted by nbr presented at either end of the ratio graph. At one end of the graph were miRNAs with exceptionally high ratios of nbr/control and at the other end of the graph were miRNAs with exceptionally low ratios of nbr/control. The ratio equals [(the number in nbr of most common form in wild type/sum all forms in nbr) divided by (the most common form in wild type/sum of all forms in wild type)]. Thus, the ratio is excessively large or excessively low when the most common length in nbr is either much greater or much lower than the percentage of reads of that length isoform for the miRNA in wild type. In addition, another deep sequencing dataset from Drosophila S2 cells GSM430030 was also used (Zhou et al. 2009). Reads were mapped to the miRNA stemloop, delineated by read length and sequence, and analyzed.

Transcriptional Profiling

For microarray analysis, DL1 cells were treated with dsRNAs (Renilla control or nbr). Total RNA was extracted from 2.5 million DL1 cells per replicate with Trizol Reagent

(Invitrogen). Microarray hybridization and reading was performed at the Penn Microarray Core Facility. For mRNA microarrays, total RNA was reverse-transcribed to ss-cDNA, followed by two PCR cycles using the Ovation RNA amplification system V2 (Ovation). Quality control on both RNA and ss-cDNA was performed using 2100 Agilent Bioanalyzer (Quantum Analytics). The cDNA was labeled using the FL-OvationTM cDNA Biotin Module V2 (Ovation), hybridized to GeneChip Drosophila Genome 2.0 Arrays (Affymetrix) and scanned with an Axon Instruments 4000B Scanner using GenePix Pro 6.0 image acquisition software (Molecular Devices).

Five biological replicates of each set of cells, and each genotype of flies were used. Affymetrix .cel (probe intensity) files were exported from GeneChip Operating Software (Affymetrix). The .cel files were imported to ArrayAssist Lite (Agilent) in which GCRMA probeset expression levels and Affymetrix absent/present/marginal flags were calculated. Statistical analysis for those genes passing the flag filter was performed using Partek Genomics Suite v6.6 (Partek). The signal values were log2 transformed and a 2-way ANOVA was performed. DataGraph 2.3.2 was used to generate the scatterplot (http://www.visualdatatools.com/DataGraph/).

Real-time RT-PCR Analysis

Total RNA was prepared from DL-1 cells treated with dsRNAs (Renilla control or Nbr), and flies, control, nbr-/-, loqsf00791. cDNA was synthesized by High-Capacity cDNA Reverse Transcription kit (Applied Biosystems). The realtime-PCR reaction was performed by Power SYBR Green PCR Master Mix (Applied Biosystems) in 7500 Fast

Realtime PCR System (Applied Biosystems). Each target gene was normalized to endogenous control (Rp49), followed by calculation of relative fold change compared to control. 500 Fast System SDS Software (Applied Biosystems)

Luciferase assays in DL1 cells

8?104 DL1 cells were plated and bathed in 30 ?l of serum-free medium with 60 ng of dsRNA in each well of a 96-well plate. The next day, the following amount of plasmids were transfected by Effectene (Qiagen): (A) 1.6ng of pMT-Firefly, 400ng of pMT-miR-34, and 400ng of pMT-Renilla. (B) 1.6ng of pMT-Firefly, 400ng of pMT-miR-34, and 400ng of pMT-Renilla-E74A- 3'UTR. (C) 1.6ng of pMT-Firefly, 400ng of pMT-miR-277, and 400ng of pMT-Renilla-miR-277- 4xbulged targets. (D) 1.6ng of pMT-Firefly and 800ng of pMT-Renilla-mRpS25-3'UTR. (E) 1.6ng of pMT-Firefly and 800ng of pMT-Renilla-Fas3 3'UTR. (F) 1.6ng of pMT-Firefly and 800ng of pMT-Renilla-CG3328-3'UTR. (G) 1.6ng of pMT-Firefly and 800ng of pMT-Renilla-CG30359 3'UTR. (H) 1.6ng of pMT-Firefly and 800ng of pMT-Renilla-GluRIIA 3'UTR. (I) 1.6ng of pMT-Firefly and 800ng of pMT-Renilla-nub 5'UTR.

Two days after transfection, the expression of the reporters and miR-34 was induced by CuSO4. Twenty-four hours after induction, luminescence assays were performed by the Dual-Glo Luciferase Assay System (#E2920, Promega, Fitchburg, WI).

Chapter 4 Materials and methods

Small RNA northern hybridization

Total RNA was extracted from Fly tissues using Trizol Reagent (#15596-018, Life Technologies, Carlsbad, CA), following the manufacturer's protocol. 3~10ug of total RNA was loaded/lane in 15% TBE-urea gel (#EC6885BOX, Life Technologies, Carlsbad, CA), followed by the transferring to nylon membrane (Hybond N+, GE Healthcare, Piscataway, NJ). After UV crosslinking and pre-hybridization (50°C, 1h), the membranes were hybridized with P32-labeled probes overnight at 50°C. DNA oligonucleotides were annealed to obtain the templates for RNA probes. The sequences of the DNA oligonucleotides used to make probes were: miR-34-5p (5'-GAT AAT ACG ACT CAC TAT AGG GAG A-3'/5'-AAA AAA TGG CAG TGT GGT TAG CTG GTT GTG TCT CCC TAT AGT GAG TCG TAT TAT C-3'), miR-263a-5p (5'-GAT AAT ACG ACT CAC TAT AGG GAG A-3'/5'-GTT AAT GGC ACT GGA AGA ATT CAC TCT CCC TAT AGT GAG TCG TAT TAT C-3'), miR-305-5p (5'-GAT AAT ACG ACT CAC TAT AGG GAG A-3'/5'-AAA AAA ATT GTA CTT CAT CAG GTG CTC TGT CTC CCT ATA GTG AGT CGT ATT ATC-3'), miR-317-3p (5'-GAT AAT ACG ACT CAC TAT AGG GAG A-3'/5'- AAA AAA ATG AAC ACA GCT GGT GGT ATC CAG TTC TCC CTA TAG TGA GTC GTA TTA TC-3'), esi2.1 (5'-GAT AAT ACG ACT CAC TAT AGG GAG A-3'/5'-TTG ACT CCA ACA AGT TCG CTC CTC TCC CTA TAG TGA GTC GTA TTA TC-3'), miR-8 (5'-GAT AAT ACG ACT CAC TAT AGG GAG A-3'/5'-AAA AAA TAA TAC TGT CAG GTA AAG ATG TTC TCC CTA TAG TGA GTC GTA TTA TC-3'), miR-100 (5'- AAA AAA AAC CCG TAA ATC CGA ACT TGT GTC TCC CTA TAG TGA GTC GTA TTA TC-3'), miR-1000 (5'-GAT AAT ACG ACT CAC TAT AGG GAG A-3'/5'-AAA AAA ATA TTG TCC TGT CAC AGC AGT TCT CCC TAT AGT GAG TCG TAT TAT C-3') and 2S rRNA (5'- GAT

AAT ACG ACT CAC TAT AGG GAG A-3’/5'-TGC TTG GAC TAC ATA TGG TTG AGG GTT GTA TCT CCC TAT AGT GAG TCG TAT TAT C-3’). P32-labeled probes were synthesized by in vitro transcription using MAXIscript T7 Kit (#AM1312, Life Technologies, Burlingame, CA), supplemented with P32-?-UTP.

Oxidation/B-elimination of RNA

Total RNA was extracted from fly tissues using Trizol Reagent (#15596-018, Life Technologies, Carlsbad, CA), following the manufacturer's protocol. Total RNA was resuspended in 1xborate/borax buffer with or without NaIO4 (final concentration: 25mM) (#311448, Sigma-Aldrich, St. Louis, MO), and incubated at room temperature in dark for 30min. After adding 1/10 volume of 100% glycerol, the samples were incubated further for 10min at room temperature at dark. After purifying the RNA, the RNA samples were resuspended into 1xborate/borax buffer with supplemented with NaOH to the final concentration of 0.1 mM. The samples were incubated at 45°C for 90min, followed by purification and analysis by northern blots. 10~20ug of total RNA was used for each miRNA.

IP-northern

Fly lysate was prepared from 3d and 30d male FLAG-HA-ago2 whole flies (Czech et al. 2008). Approximately 80 flies were used for each IP (Ago1 or Ago2) followed by northerns (miR-34-5p, esi-2.1, miR-305-5p, miR-263a-5p, miR-11-3p, and miR-317-3p). Immunoprecipitation was performed as described (Kirino et al. 2011), except that 800mM of NaCl (final concentration) was used for Ago2-IP. Anti-Ago1 (ab5070,

Abcam) or M2 beads (#A2220, Sigma-Aldrich, St. Louis, MO) was used for Ago1 or Ago2-IP respectively. After extraction of RNA from beads as described in Kirino et al., 2011, the purified RNA was loaded onto 15% TBE-urea gel, followed by northern blots.

Preparation of Ago1 and Ago2-IP small RNA libraries

40 whole male flies (FLAG-HA-ago2, 3d and 30d) were used for Ago1-IP, and 200 whole male flies (3d and 30d) were used for Ago2 (M2)-IP. After immunoprecipitation of Ago1 or Ago2, the purified RNA was P32-labeled as described (Kirino et al. 2011), and the radiolabeled RNA was run on a 15% TBE-urea gel (#EC6885BOX, Life Technologies, Carlsbad, CA). The gel was exposed to phosphorimager screen, and the fractions corresponding to small RNAs were excised from the gel. After purifying the small RNAs from the gel, small RNA libraries were generated using Illumina's TruSeq small RNA sample prep kit (#RS-200-0012, Illumina, Inc. San Diego, CA), following the manufacturer's protocol. The libraries were sequenced on HiSeq2000 platform (Illumina).

Western immunoblots

Fly tissues were resuspended in RIPA buffer, followed by grinding and centrifugation to remove debris. The supernatant was measured by Bradford assay, and 25~50ug of protein was loaded each lane. NuPAGE Novex 4-12% Bis-Tris gel (#NP0321BOX, Life Technologies, Carlsbad, CA) was used to run the samples in 1xNuPAGE MES SDS Running Buffer (#NP0002, Life Technologies, Carlsbad, CA). Proteins were transferred to PVDF membrane, and the membrane was blocked by 5% milk/TBST for 1h at 4°C. The membrane was incubated with primary at 4°C overnight. After washing the

membrane in TBST buffer 3 times (5 min each), the membrane was incubated with secondary antibody at 4°C for 2h. The membrane was washed in TBST 3 times (5 min each), followed by signal development by Pierce ECL plus Western Blotting Substrate (#32132, Thermo Fisher Scientific, Waltham, MA). The image was scanned by Fujifilm LAS-3000 Imager (Fujifilm, Tokyo, Japan). For Ago2 western, FLAG-HA-ago2 male flies were used. For R2D2 and Dcr2 WB, wild-type male flies were used. The primary antibodies used are anti-FLAG-HRP (1:2000, #A8592, Sigma-Aldrich, St. Louis, MO) for Ago2. Anti-Ago1 (1: 500-1:1000) for westerns was a kind gift of the Mourelatos lab. Anti-R2D2 (rabbit polyclonal) and anit-Dcr2 (mouse monoclonal) antibodies were the kind gifts from Siomi lab. Anti-Actin Ab (Abcam Ab8227) was used at 1:2000. The secondary antibodies used are anti-mouse IgG-HRP (1:2000, #7076S, Cell Signaling Technology, Inc. Danvers, MA), anti-rabbit IgG-HRP (1:2000, #sc-2030, Santa Cruz Biotechnology, Inc., Dallas, Texas).

Fly stocks

Flies were grown in standard cornmeal molasses agar medium at 25°C. Hen1f00810 (FlyBase ID: FBst1016506, The Exelixis Collection at the Harvard Medical School) and Ago2414 (Flybase ID: FBst0313641, DGRC Kyoto Stock Center) flies were the kind gifts from Sara Cherry lab. Ago2BL16608 (FlyBase ID: FBst0016608) and FLAG-HA-Ago2 flies (FlyBase ID: FBst0033242) were obtained from the Bloomington Stock Center. Hen1f00810 and Ago2BL16608 lines were backcrossed into a homogenous wild-type background (Bloomington Stock Line 5905 (BL5905), FlyBaseID: FBst0005905, w1118) for five generations.

Lifespan assay

180~200 flies were used for each lifespan replicate. Male flies were collected on the day of eclosion, aged at 25°C, 15 flies per vial, and transferred to new fly food vials every other day while scored for the survival. The assay was repeated in triplicate, and was analyzed by Excel (Microsoft) for survival curves.

Brain paraffin sections

Adult female heads (3d and 30d) were used for paraffin sections as described (Li et al. 2008). Brain vacuoles were counted for lamina or retina through 10 continuous horizontal sections, defining the center section with oesophagus being most prominent. 4 heads per genotype were used for quantification.

Computational analyses

Mapping and Histograms:

Adaptor sequences were removed from the 3' end of the reads in the Illumina fastQ generated files using the fastx-toolkit (http://hannonlab.cshl.edu/fastx_toolkit/). The adapter sequences are as follows:

5' adapter = 5'- GUUCAGAGUUCUACAGUCCGACGAUC- 3'

3' adapter = 5'- TGGAATTCTCGGGTGCCAAGG- 3’

Reads were then collapsed and annotated with the number of times each was sequenced,

and only unique reads were analyzed. Reads were filtered based on bp size, such that reads less than 16 bp or greater than 30 bp were discarded. The remaining reads were then mapped using Bowtie to the Drosophila melagonaster (dm5) genome and microRNA stem-loop sequences obtained from miRBase (release 20). Bowtie parameters were restricted to only output perfectly aligned matches to the stemloop sequence. The reads were aligned and mapped to entire miRNA stem-loops. After aligning, each read was either annotated as a 5P or a 3P. To do this, the stem-loop sequence was split in half, with the first half designated the 5' arm and the second half designated the 3' arm, as the start and end positions of most microRNAs are imprecise.

Change in the read number and percentage of total microRNAs and other classes of small RNAs:

All processed reads were first mapped to transposons annotated by the Flybase (ftp://ftp.flybase.net/genomes/Drosophila_melanogaster/current/fasta/). The rest of the reads were stored and mapped to microRNA stem-loop regions, as well as known ncRNAs (snRNAs, snoRNAs, tRNAs, and rRNAs) and the rest of the genome (defined as "ncRNA" and "genome" by Flybase).

Identification of preferentially Ago2-loaded miRNA isoforms at 3d and 30d:
We identified miRNA isoforms preferentially loaded into Ago2 compared to Ago1 for flies of different age using a measure of "Ago2/Ago1 relative load ratio", or R21. The most abundant isoform from each miRNA stem (5p or 3p) from Ago2-IP 3d library was

selected. For all the isoforms in each library that start with the same 5' nt position, we denoted the share of each isoform silib, such that ∑isilib = 1 in each library, and the share of the most abundant isoform Slib. The age ratio for 3d flies was R3d21 = S3dAgo2/S3dAgo1 (x-axis in Fig. 6A) and for 30d, R30d21 = S30dAgo2/S30dAgo1 (y-axis of Fig. 6A). 5p and 3p miRNA stem sequences were classified as “mature” and “star”, based on their relative (high and low, respectively) deep-sequencing read counts in the Ago1 sample at 3d (strand with S3dAgo1 being the mature strand). In order to compare the ratio of a specific isoform between Ago1 and Ago2, after assigning "mature" or "star" category, we only kept the cases with the mature and star categories consistent among all four libraries. Thus, if there was a discrepancy of the relatively more abundant stem (5p or 3p) between Ago1-3d and any of the other three libraries (Ago2-3d, Ago1-30d, and Ago2-30d), those cases were removed. This was to assure that the same isoform was consistently called "mature" or "star" in all four libraries (Ago1-3d, Ago1-30d, Ago2-3d, Ago2-30d). We also filtered out any miRNAs that were below a 1000 read count threshold. Note that miR-34-18nt, miR-276a/b-5p-17nt, and miR-981-5p-21nt showed extremely high R21 ratio at 3d and/or 30d (> 16, Table S2 and S3). This suggests much higher accumulation of these miRNA isoforms in Ago2-IP, compared to Ago1-IP at 3d and/or 30d. This could happen either due to biologically relevant upregulation of these miRNA isoforms, or to cloning bias or degradation from the corresponding longer isoforms. These miRNA isoforms were removed from Fig. 6, because including the dots corresponding to these miRNA isoforms obscured the distribution of the other dots in the plot (see Fig. 6A). miR-34-18nt was undetectable by Ago1 vs Ago2-IP northerns, therefore most-likely miR-34-18nt is an abnormal by-product of the library preparation.

After identifying preferentially loaded Ago2 isoforms, we calculated a ratio of normalized reads at 30d and 3d for each specific isoform. The ratios (fold change) were plotted as in Fig. 6B. To check for selective enrichment of particular miRNAs from data of Ghildiyal et al. (2010), we first identified the most abundant isoform for each miRNA in the total RNA set (GSM466487). For each of the libraries (GSM466488 and GSM466489) we counted the number of reads M for these isoforms and calculated their relative enrichment as E = N / (N – M), where N is the total number of library reads for a given miRNA. We then plotted the enrichment ratio EGSM466488/EGSM466489 in SFig 3. Note that we used both sequence analysis for discovery of new Ago2-preferentially loaded miRNAs (see Figures 6A,B, SFigure 3), and also northern analysis (see Figure 6C) as each technique (deep sequencing and northern analysis), will have strengths and weaknesses. We also analyzed previously published microarray data GSE25009 from Liu et al (2012) and data from Taliaferro et al (2013).

Chapter 5 Materials and Methods

Mapping and quantifying tRFs

Adaptor sequences were removed from the 3' end of the reads in the Illumina fastQ generated files using the fastx-toolkit (http://hannonlab.cshl.edu/fastx_toolkit/). The adapter sequences are as follows:

5' adapter = 5'- GUUCAGAGUUCUACAGUCCGACGAUC- 3'

3' adapter = 5'- TGGAATTCTCGGGTGCCAAGG- 3’

Reads were then collapsed and annotated with the number of times each was sequenced, so only unique reads were analyzed. The reads were then mapped using Bowtie to the Drosophila melagonaster (dm5) genome and tRNAs obtained from FlyBase. Bowtie parameters were restricted to only output perfectly aligned matches to the tRNA sequence. The reads were aligned and mapped to the entire tRNA sequence with the CCA addition. After aligning, each read was annotated as a 5P, 3P, or middle. To do this, the tRNA sequence was split into three even parts, with the first part designated as the 5' arm, the second part being the middle and the last part designated as the 3' arm. If the read spanned multiple parts, any part that had at least the starting 5 nts was chosen. After mapping reads to their respective tRNAs, each library was independently normalized by the total number of reads mapped to the D. melanogaster genome (v. R6.03).

Differential/Preferential loading with age

We identified differential loading of tRFs with age in Ago1 and Ago2 using a ratio metric. We first identified the most abundant isoform in our 30 day libraries and used the read count numbers of that specific isoform for our ratio calculations. We calculated the ratio of 30 days to 3 days for Ago1 and Ago2 of highly abundant (>=1000) tRFs. We then plotted the ratios to see loading changes that may occur with age.

To observe what was preferentially loaded (Ago1 vs Ago2) with age, we obtained a different ratio. The ratio of this measure was the ratio of reads of a particular tRF of Ago2 to Ago1 at 3 days and at 30 days.

Targeting Analysis and GO Analysis

In order to identify a potential seed sequence in our dataset, we generated 7mer subsequences of the tRF by applying a sliding window by shifting one nt towards the 3' end after each subsequent 7mer generation. We then mapped not allowing for any mismatches each of these subsequences to conserved 5' UTR, CDS, and 3' UTR regions of 12 Drosophila genomes provided by UCSC (Rosenbloom et al. 2014). In order to compare and since the sizes of each region differ, we normalized by dividing the total hits by the mean length of each respective region. Since, the 3' UTR regions had the most tRFs alignments (see results), we then plotted the number of hits for each subsequence to see if there was preferential binding. In order to predict targets, we aligned the candidate seeds of each tRF to the 3’ UTRs, only allowing perfect matches similar to that of TargetScan (Grimson et al. 2007, Lewis, Burge, and Bartel 2005, Lewis et al. 2003). We did not use TargetScan since its "context score" was unlikely to be applicable for our cases of both 3' and 5' seeds. AmiGO (Ashburner et al. 2000) was used to find enriched GO-terms in our target list for each tRF.

CHAPTER 7 DISCUSSION AND FUTURE DIRECTIONS

It has been commonly observed via small RNA sequencing that there is great diversity, especially originating from 3' end in most miR species. Some components, for non-templated additions and RNA editing have been identified, but for the first time in Drosophila, we discovered and further characterized a 3' to 5' exonuclease protein, Nbr that co-immunoprecipitated with Ago1, which is not only responsible for processing the 3' ends in mir-34, but a subset of miRs. With the assistance of advanced computational capabilities, we identify Nbr-dependent miRs. Furthermore, we report evidence of the biological impact of these miRs in brain development and aging in our studies and others. In order to assist us in recognizing these types of modifications, we introduce a tool helping us visualize miR stem-loops, NGS reads with secondary structure and nucleotide modification information. It is designed in a way where we could apply our tool to other RNA molecules and we demonstrate that with the investigation of tRFs in the later chapters.

Select isoforms impacted by Nbr and other components involved

We have determined that Nbr selects a subset of miRs for trimming. We showed there was no a global impact on miRs (Figure 3.7), which suggests that there may be a signal for specific isoforms that recruits Nbr for further processing. For Nbr-dependent miRs, we observed two patterns upon Nbr knockdown, longer isoforms becoming the most abundant, or shorter isoforms becoming less frequent. The selection process may be related to sequence length. It is known that recruitment of other components in the miR maturation process is dependent on sequence length, such as Dicer. Another reason may

be specific motifs present in the miR, as it is known that certain motifs recruit RNA binding proteins (RBPs).

It is very probable that the 3' end of a miR contains a recognition sequence for the recruitment of other proteins. The 3' is long enough to harbor a potential RBP signal, which may be 2-6 nucleotides in length. With NGS data for Nbr wildtype and Nbr mutation libraries, one could computationally test this. First, one could scan each Nbr-dependent isoform and compute isoform size to find any representative lengths and compare that with Nbr-independent isoforms. For finding motifs, one could scan for particularly 2-6 nucleotide patterns to see if there are any statistically significant patterns over represented.

This will give us a list of potential proteins that may be recruited and work with Nbr. One could check to see if they are present in the RBPmap (Paz et al. 2014) or other known databases. After identifying a set of potential RBPs, then one could run an RNAi assay targeting Nbr and this newly identified protein, similar to how Nbr was identified in Chapter 3. Then one would help us identify components that co-immunoprecepitate with Nbr. This could also be helpful in identifying how Ago2-loaded small RNAs, including piRNAs and endo-siRNAs, are trimmed (Feltzin et al. 2015) by allowing us to identify bound proteins to Nbr. In essence, this will reveal additional components that may work with Nbr, and may also reveal potentially another exonuclease protein, since we observed Nbr-independent miRs as well.

3’ processing in other species

There are known homologous proteins to Nbr, so it hints at a conserved process that spans beyond the fly. According to our phylogenetic analysis, Nbr belongs to the RNase D protein family. The closest homolog of Nbr in C. elegans is mut-7 (Figure 3.5D). One could do a similar test on mut-7, by disrupting mut-7 via RNAi and apply our computational approach to identify relative patterns of miR isoforms and potential 3’ processing. Additionally, with publically available datasets, especially with the completion of modEncode, it would be interesting to look at homologous proteins in other species as well or look for potential 3’ processing. This will tell one if this process of 3’ trimming is selected for and is conserved across species, adding an evolution perspective for this particular mechanism. This may also further suggest functional relevance for this process.

Distinct pathways for Ago loading

We initially observed an accumulation of certain isoforms, which prompted us to pursue computational and bioinformatics analyses on specific Ago1 and Ago2 immunoprecipitated libraries at 3 days and 30 days. This allowed us to discover global trends of miR loading that were modulated with age. Ago2 was the more dynamic Ago protein, as global populations of small RNA loading changed more readily, including miRs and tRFs. We then sought out to investigate individual isoforms, and we were able to find that specific isoforms contributed to this effect, adding to our knowledge of differential partitioning between Ago proteins. The results of chapters 2 and 3, lead us to question if there was a difference in loading patterns between Ago1 and Ago2 in the

context of age. In plant miRs the first nucleotide and the sequence length dictate its loading destinations (Voinnet 2009) . Moreover, there are a few proposed reasons related to the star strand and anti-sense strands and loading destination (Czech and Hannon 2011, Okamura, Liu, and Lai 2009, Okamura et al. 2011). However, it is unclear why there would be an accumulation and an Ago2 preference with age. The first question would be to ask if Ago2 or any of its known interacting protein components increase with age. It was shown that was not the case (Fig. 3.4). Hence, there is a process that is independent of such components that contributes to this effect. One possibility is the existence of specific factors that might be in a complex with Ago2 and the Ago2-loading machinery. Identifying proteins interacting with Ago2, and not Ago1, in 3 days vs 30 days flies might reveal the factors that are in a complex with Ago2 in age-associated manner. After identifying potential components, one could then apply a similar analysis performed in Chapter 3 to assess if small RNA populations show changes of loading levels.

Other small RNAs impacting the brain

Though, we provide evidence for methylated Ago2-loaded miRs influence on the brain and lifespan, mutations of Ago2 and Hen1 does not rule out other small RNAs impacting age-associated traits. Thus, it would b helpful to identify other classes of small RNAs that may also play a role. Using computational biology, one can identify other Ago2 loaded small RNAs that are methylated. A starting point may be to further inspect small RNAs populations derived from other molecules, such as rRNAs and tRNAs, that displayed age-associated changes. Once a subset of differentially Ago2-loaded small RNAs is identified, one can check to see if they are present in our oxidated/b-eliminated libraries

to determine if they are methylated. Nonetheless and intriguingly, the long isoform for Ago2-loaded mir-317 remained intact after disrupting Hen1 activity suggesting an alternative protective mechanism (Abe et al. 2014). This alternative mechanism may also be present in protecting tRF isoforms, as we did not see many of them present in our oxidative/beta-eliminated libraries, leaving the possibility of the involvement of non-methylated small RNAs in such processes.

tRF and its miR-like function

Chapters 5 and 6 were an extensive analysis of tRNA fragments in Drosophila that provides strong evidence that they play miR-like roles. First, we show that the distribution of reads along the tRNA molecule mimic the asymmetric read distribution usually observed in miR expression profiles. They are also loaded into the RISC complex and the loading exhibits trends similar to that of miRs. We were able to confirm reports of tRF binding biases through computational analysis. These identified seeds also mapped mostly to 3' UTRs, allowing us to speculate their mode of gene regulation to be similar to miRs. These identified seeds also mapped mostly to 3' UTRs, allowing us to speculate their mode of gene regulation. Nevertheless, given these seeds, their targets, according to our GO analysis, often times focused on brain related processes and development, hinting at a critical role for these small RNAs. Also, the fact these seeds matched to 3' UTRs suggest they are more active in the brain, and since there is extensive 3' UTR expansion in the brain (Miura et al. 2013) it strengthens the hypothesis of them functioning like miRs.

Further characterization of Ago2-loaded tRF targets

We are confident that are tRFs are active, as we did cross species analysis (data not shown) and confirmed our most abundant tRFs in other studies (Haussecker et al. 2010, Thompson et al. 2008), but one could further characterize our findings. To further elucidate the targeting of Ago2-loaded tRFs one could do a transcriptome analysis (RNA-seq) at the two time points, 3 days and 30 days. Then one could quantify gene expression levels of their predicted targets and test to see if there are inversely (down-regulated) related to tRF loading patterns. Importantly, one would need to apply statistics to this step to make sure that the down regulation is significant. Once confirmed, the next step would be to see if these targets are shared with known differentially Ago2-loaded miRs. If there is overlap, the targets' fold change should be much higher than those that share no overlap with miR predictions (tRF exclusive). This will also suggest that tRFs work in conjunction with miRs and help further regulate overlapping gene targets. To compliment and strengthen this test, one could do a similar analysis after disrupting Hen1 (removing methylated miRs) and see if similar patterns persist.

Tissue-specificity of tRFs

We initially observed different population of tRFs originating from the ovaries (data not shown), which leads us to suspect differential expression and loading patterns in the context of the genomic environment. Hence, one can profile tRF expression levels based on the tissue type. This will further characterize tRFs that either play a broader or more specific role. tRFs possessing a broader role may be the ones that are more consistently expressed and loaded onto the Ago proteins across different tissues, since they modulate

more fundamental networks, while the more specific tRFs may be less expressed in general, but enriched in specific tissues. These specific tRFs would be necessary for the cells to regulate and modulate processes and events that are distinct to the tissue, helping maintain tissue identity. This is highly likely and one would expect this, since their miR counterparts are expressed in a very tissue-specific manner, especially during development (Wienholds and Plasterk 2005, Aboobaker et al. 2005).

Development-associated Profiling of tRFs

As a follow up to the previous direction, one could profile tRF expression and loading based on development. We investigated tRFs in adult flies and established age-associated changes of tRFs, hence, another direction to take is to ask if there is differential expression with development. Furthermore, miRs have already been established to be key players in very fundamental roles in development (Wienholds and Plasterk 2005, Aboobaker et al. 2005, Lai, Tam, and Rubin 2005). For example, the Notch signaling pathway is essential for patterning and development, and studies in Drosophila have shown that Notch-targeted genes are regulated by miRNAs via conserved motifs (Lai, Tam, and Rubin 2005). Another example in C. elegans are the let-7 miR family members mir-48, mir-84, and mir-241 in regulating developmental timing pathway (Abbott et al. 2005). Since we showed tRF targets to be enriched in development-related processes, one would expect them to play a critical role. One could use publicly available datasets of different developmental time points and profile tRF expression looking for differential patterns amongst these time points. This would help prioritize tRFs based for specific development processes and help to identify tRFs or specific tRF isoforms that may be

active in early or later developmental time points. It will help one understand their role in cell differentiation.

tRF silencing with identified seeds

For our most abundant tRF species, we observed clear biases using our seed discovery pipeline, hinting at a seed region within the tRF. We also predicted targets, which suggested important functional relevance. To test this experimentally, one can perform similar luciferase assays as before (Haussecker et al. 2010, Wang et al. 2013).

Essentially one would perform a siRNA assay, using a plasmid containing a sensor harboring a reverse complementary sequence of a tRF of interest, such as GlyGCC, in the 3' UTR region of the firefly luciferase gene. If the tRF acts on the target as expected, it would recognize the complementary sequence as a target site, therefore leading to suppressed expression of the gene. To test our seed predictions and further characterize how targeting may occur one can: (i) mutate or shuffle nucleotides in the inserted sequence away from the predicted seed region, (ii) mutate nucleotides within the seed region only and (iii) mutate nucleotides within the seed region and away from the region. After repeating with a few different/random combinations, the relative luciferase activities should reveal properties, including the necessity of the seed region and a possible 3' compensatory role in targeting.

Another direction worth considering is to apply our seed discovery pipeline to other tRFs across different species, and/or to other classes of small RNAs. Studies have shown

certain animal miRs deviate from the "seed region" type binding, with the 3' end playing a compensatory role (Elefant, Altuvia, and Margalit 2011), so one could apply our pipeline to known miRs to computationally catalogue miRs that bind in such a manner expanding our knowledge base for miR:mRNA targeting interactions. Additionally, one could apply our pipeline to other less known small RNAs. Since we observed age-associated changes in loading of rRNA-derived fragments in our libraries, making this class a good candidate in trying to understand how they may interact with their targets. This may elucidate evolutionary insight and also may help one understand the mode of regulation of other small RNAs that may display unusual targeting patterns.

Small RNA Targeting Complexity

Gene complexity will be measured as a function of a developmental time course. Qiang Tu et al. were able to identify expression patterns of genes in sea urchin, suggesting “complex” genes (Tu et al. 2012). Our current analysis in miRs and tRFs uses GO terms, which is a good first step, but sometimes does not contain enough information. This particular future direction will provide an important component by connecting small RNAs to these genes by suggesting the nature of microRNA-targeted transcripts in terms of complexity.

One would first retrieve all known small RNAs (miRs and tRFs) in the sea urchin and their predicted targets. One would then use spbase (Tu, Cameron, and Davidson 2014) to query all known targets, which will give us all relevant information on expression levels and directions (increase/decrease) for the time course. Once all relevant information is

retrieved, one would compute the complexity of each gene. Particularly, this will be computed and assigned using two parameters, including the slope and the relative abundance levels. For example, if the slope changes dramatically and the abundance significantly increases/decreases that gene would be annotated as "complex." With this information one would then cluster the targets into different bins (high, low, and medium complexity). Finally, one would then overlap each target with its associated miR and/or tRF to help characterize these clusters in relation to their small RNA.

In addition in helping one cluster small RNAs based on the complexity of their targets, it will also inform one if there is a distinction between miRs and tRFs, potentially explaining why the cell utilizes both types of small RNAs. This will help one gain greater insight into gene regulation and help refine small RNA targeting and regulation. This would also be important, especially when looking for structural variations, because it will reflect the significance of such modifications. The literature and past studies already illustrate the role of editing in gene regulation; editing is much more frequent in the UTR regions, and may be associated with avoidance of miR-mediated repression (Friedman et al. 2009), but these set of results will provide another layer of pertinent information, by allowing one to rank structural changes based on gene complexity.

tRF evolution and modifications

Comparative genomics has provided us important insights into miR evolution, including common signatures and constraints, which point to functional importance (Lai et al.

2003, Czech and Hannon 2011, Nozawa, Miura, and Nei 2010). One can exploit the availability of the 12 sequenced Drosophila genomes, helping uncover important features.

One can first identify well-conserved tRFs, which will allow them to identify tRFs that may be related to more fundamental processes. A deeper analysis would help assess species-specific tRFs originating from different or even the same tRNA. This type of analysis also may reveal potential cleavage sites or signatures that were not known before, as it has for miRs (Czech and Hannon 2011). Particularly it may show conserved cleavage sites that may be associated with certain structures in the tRNA molecule, including the D, T and anticodon loops.

Coupling this comparative analysis, one could also scan for RNA editing and non-templated modifications in tRFs to see if they undergo similar modifications. This will help one understand the diversity amongst tRFs and help us determine if certain tRFs undergo more modifications than others and which type of modifications are preferred. Certain modifications may also have impact on tRF function by effecting turnover rate, stability or even targeting, so it would be helpful to identify specific trends or patterns of modifications that may be present and to ascertain their consistency across species.

Conclusion

We have been able to characterize and have identified additional steps in the small RNA processing pathway. We were able to connect select isoforms of miR and tRFs to age-associated events, impacting brain and neuronal processes. We also were able to assess

properties, including structural changes, expression and loading patterns of tRFs on a global level in Drosophila, which highlighted novel patterns and Ago-loaded tRFs. Altogether, our studies have added crucial steps to the canonical small RNA processing pathway, expanding our knowledge in small RNA biogenesis and has opened the window for further studies that will help in the discovery and characterization of other classes of small RNAs and their potential function.

BIBLIOGRAPHY

Abbott, A. L., E. Alvarez-Saavedra, E. A. Miska, N. C. Lau, D. P. Bartel, H. R. Horvitz, and V. Ambros. 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 (3):403-14. doi: 10.1016/j.devcel.2005.07.009.

Abe, M., A. Naqvi, G. J. Hendriks, V. Feltzin, Y. Zhu, A. Grigoriev, and N. M. Bonini. 2014. "Impact of age-associated increase in 2'-O-methylation of miRNAs on aging and neurodegeneration in Drosophila." Genes Dev 28 (1):44-57. doi: 28/1/44 [pii]
10.1101/gad.226654.113.

Aboobaker, A. A., P. Tomancak, N. Patel, G. M. Rubin, and E. C. Lai. 2005. "Drosophila microRNAs exhibit diverse spatial expression patterns during embryonic development." Proc Natl Acad Sci U S A 102 (50):18017-22. doi: 10.1073/pnas.0508823102.

Ambros, V. 2004. "The functions of animal microRNAs." Nature 431 (7006):350-5. doi: 10.1038/nature02871.

Ameres, S. L., M. D. Horwich, J. H. Hung, J. Xu, M. Ghildiyal, Z. Weng, and P. D. Zamore. 2010. "Target RNA-directed trimming and tailing of small silencing RNAs." Science 328 (5985):1534-9. doi: 10.1126/science.1187058.

Anderson, P., and P. Ivanov. 2014. "tRNA fragments in human health and disease." FEBS Lett:4297-4304. doi: S0014-5793(14)00665-6 [pii]
10.1016/j.febslet.2014.09.001.

Ashburner, M., C. A. Ball, J. A. Blake, D. Botstein, H. Butler, J. M. Cherry, A. P. Davis, K. Dolinski, S. S. Dwight, J. T. Eppig, M. A. Harris, D. P. Hill, L. Issel-Tarver, A. Kasarskis, S. Lewis, J. C. Matese, J. E. Richardson, M. Ringwald, G. M. Rubin, and G. Sherlock. 2000. "Gene ontology: tool for the unification of biology. The Gene Ontology Consortium." Nat Genet 25 (1):25-9. doi: 10.1038/75556.

Azim, A. C., J. H. Knoll, S. M. Marfatia, D. J. Peel, P. J. Bryant, and A. H. Chishti. 1995. "DLG1: chromosome location of the closest human homologue of the Drosophila discs large tumor suppressor gene." Genomics 30 (3):613-6. doi: S0888-7543(85)71286-4 [pii]
10.1006/geno.1995.1286.

Azuma-Mukai, A., H. Oguri, T. Mituyama, Z. R. Qian, K. Asai, H. Siomi, and M. C. Siomi. 2008. "Characterization of endogenous human Argonautes and their miRNA partners in RNA silencing." Proc Natl Acad Sci U S A 105 (23):7964-9. doi: 10.1073/pnas.0800334105.

Babiarz, J. E., J. G. Ruby, Y. Wang, D. P. Bartel, and R. Blelloch. 2008. "Mouse ES cells express endogenous shRNAs, siRNAs, and other Microprocessor-independent,

Dicer-dependent small RNAs." Genes Dev 22 (20):2773-85. doi: 10.1101/gad.1705308.

Bagga, S., J. Bracht, S. Hunter, K. Massirer, J. Holtz, R. Eachus, and A. E. Pasquinelli. 2005. "Regulation by let-7 and lin-4 miRNAs results in target mRNA degradation." Cell 122 (4):553-63. doi: 10.1016/j.cell.2005.07.031.

Bai, H., P. Kang, A. M. Hernandez, and M. Tatar. 2013. "Activin signaling targeted by insulin/dFOXO regulates aging and muscle proteostasis in Drosophila." PLoS Genet 9 (11):e1003941. doi: 10.1371/journal.pgen.1003941
PGENETICS-D-13-01286 [pii].

Ballard, C., and C. Fox. 2006. "Effectiveness of collaborative care for elderly adults with Alzheimer's disease in primary care." Lancet Neurol 5 (8):644-5. doi: 10.1016/S1474-4422(06)70507-5.

Ballard, C., S. Gauthier, A. Corbett, C. Brayne, D. Aarsland, and E. Jones. 2011. "Alzheimer's disease." Lancet 377 (9770):1019-31. doi: 10.1016/S0140-6736(10)61349-9.

Ballard, C., J. Grace, I. McKeith, and C. Holmes. 1998. "Neuroleptic sensitivity in dementia with Lewy bodies and Alzheimer's disease." Lancet 351 (9108):1032-3.

Bartel, D. P. 2004. "MicroRNAs: genomics, biogenesis, mechanism, and function." Cell 116 (2):281-97.

Bartel, D. P. 2009a. "MicroRNAs: target recognition and regulatory functions." Cell 136 (2):215-33. doi: 10.1016/j.cell.2009.01.002.

Bartel, David P. 2009b. "MicroRNAs: target recognition and regulatory functions." Cell 136 (2):215-233.

Bass, B. L. 1997. "RNA editing and hypermutation by adenosine deamination." Trends Biochem Sci 22 (5):157-62.

Bass, B. L. 2002. "RNA editing by adenosine deaminases that act on RNA." Annu Rev Biochem 71:817-46. doi: 10.1146/annurev.biochem.71.110601.135501.

Berezikov, E., N. Robine, A. Samsonova, J. O. Westholm, A. Naqvi, J. H. Hung, K. Okamura, Q. Dai, D. Bortolamiol-Becet, R. Martin, Y. Zhao, P. D. Zamore, G. J. Hannon, M. A. Marra, Z. Weng, N. Perrimon, and E. C. Lai. 2011. "Deep annotation of Drosophila melanogaster microRNAs yields insights into their processing, modification, and emergence." Genome Res 21 (2):203-15. doi: 10.1101/gr.116657.110.

Bernstein, E., A. A. Caudy, S. M. Hammond, and G. J. Hannon. 2001. "Role for a bidentate ribonuclease in the initiation step of RNA interference." Nature 409 (6818):363-6. doi: 10.1038/35053110.

Bilen, J., N. Liu, B. G. Burnett, R. N. Pittman, and N. M. Bonini. 2006. "MicroRNA pathways modulate polyglutamine-induced neurodegeneration." Mol Cell 24 (1):157-63. doi: 10.1016/j.molcel.2006.07.030.

Boehm, M., and F. Slack. 2005. "A developmental timing microRNA and its target regulate life span in C. elegans." Science 310 (5756):1954-7. doi: 10.1126/science.1115596.

Bommer, G. T., I. Gerin, Y. Feng, A. J. Kaczorowski, R. Kuick, R. E. Love, Y. Zhai, T. J. Giordano, Z. S. Qin, B. B. Moore, O. A. MacDougald, K. R. Cho, and E. R. Fearon. 2007. "p53-mediated activation of miRNA34 candidate tumor-suppressor genes." Curr Biol 17 (15):1298-307. doi: 10.1016/j.cub.2007.06.068.

Boutros, M., A. A. Kiger, S. Armknecht, K. Kerr, M. Hild, B. Koch, S. A. Haas, R. Paro, N. Perrimon, and Consortium Heidelberg Fly Array. 2004. "Genome-wide RNAi analysis of growth and viability in Drosophila cells." Science 303 (5659):832-5. doi: 10.1126/science.1091266.

Bremer, J., T. O'Connor, C. Tiberi, H. Rehrauer, J. Weis, and A. Aguzzi. 2010. "Ablation of Dicer from murine Schwann cells increases their proliferation while blocking myelination." PLoS One 5 (8):e12450. doi: 10.1371/journal.pone.0012450.

Brennecke, J., D. R. Hipfner, A. Stark, R. B. Russell, and S. M. Cohen. 2003. "bantam encodes a developmentally regulated microRNA that controls cell proliferation and regulates the proapoptotic gene hid in Drosophila." Cell 113 (1):25-36. doi: S0092867403002319 [pii].

Brennecke, J., A. Stark, R. B. Russell, and S. M. Cohen. 2005. "Principles of microRNA-target recognition." PLoS Biol 3 (3):e85. doi: 10.1371/journal.pbio.0030085.

Brodersen, P., and O. Voinnet. 2009. "Revisiting the principles of microRNA target recognition and mode of action." Nat Rev Mol Cell Biol 10 (2):141-8. doi: 10.1038/nrm2619.

Burroughs, A. M., Y. Ando, M. J. de Hoon, Y. Tomaru, T. Nishibu, R. Ukekawa, T. Funakoshi, T. Kurokawa, H. Suzuki, Y. Hayashizaki, and C. O. Daub. 2010. "A comprehensive survey of 3' animal miRNA modification events and a possible role for 3' adenylation in modulating miRNA targeting effectiveness." Genome Res 20 (10):1398-410. doi: 10.1101/gr.106054.110.

Calabrese, J. M., A. C. Seila, G. W. Yeo, and P. A. Sharp. 2007. "RNA sequence analysis defines Dicer's role in mouse embryonic stem cells." Proc Natl Acad Sci U S A 104 (46):18097-102. doi: 10.1073/pnas.0709193104.

Calin, G. A., and C. M. Croce. 2006. "MicroRNA signatures in human cancers." Nat Rev Cancer 6 (11):857-66. doi: nrc1997 [pii]
10.1038/nrc1997.

Chang, T. C., and J. T. Mendell. 2007. "microRNAs in vertebrate physiology and human disease." Annu Rev Genomics Hum Genet 8:215-39. doi: 10.1146/annurev.genom.8.080706.092351.

Chang, T. C., E. A. Wentzel, O. A. Kent, K. Ramachandran, M. Mullendore, K. H. Lee, G. Feldmann, M. Yamakuchi, M. Ferlito, C. J. Lowenstein, D. E. Arking, M. A. Beer, A. Maitra, and J. T. Mendell. 2007. "Transactivation of miR-34a by p53 broadly influences gene expression and promotes apoptosis." Mol Cell 26 (5):745-52. doi: 10.1016/j.molcel.2007.05.010.

Chatterjee, S., and H. Grosshans. 2009. "Active turnover modulates mature microRNA activity in Caenorhabditis elegans." Nature 461 (7263):546-9. doi: 10.1038/nature08349.

Cheloufi, S., C. O. Dos Santos, M. M. Chong, and G. J. Hannon. 2010. "A dicer-independent miRNA biogenesis pathway that requires Ago catalysis." Nature 465 (7298):584-9. doi: 10.1038/nature09092.

Chen, C. J., Q. liu, Y. C. Zhang, L. H. Qu, Y. Q. Chen, and D. Gautheret. 2011. "Genome-wide discovery and analysis of microRNAs and other small RNAs from rice embryogenic callus." RNA Biol 8 (3):538-47. doi: 15199 [pii].

Chen, L. H., G. Y. Chiou, Y. W. Chen, H. Y. Li, and S. H. Chiou. 2010. "MicroRNA and aging: a novel modulator in regulating the aging network." Ageing Res Rev 9 Suppl 1:S59-66. doi: 10.1016/j.arr.2010.08.002.

Chendrimada, T. P., R. I. Gregory, E. Kumaraswamy, J. Norman, N. Cooch, K. Nishikura, and R. Shiekhattar. 2005. "TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing." Nature 436 (7051):740-4. doi: 10.1038/nature03868.

Cherry, S., and N. Perrimon. 2004. "Entry is a rate-limiting step for viral infection in a Drosophila melanogaster model of pathogenesis." Nat Immunol 5 (1):81-7. doi: 10.1038/ni1019.

Choi, P. S., L. Zakhary, W. Y. Choi, S. Caron, E. Alvarez-Saavedra, E. A. Miska, M. McManus, B. Harfe, A. J. Giraldez, H. R. Horvitz, A. F. Schier, and C. Dulac. 2008. "Members of the miRNA-200 family regulate olfactory neurogenesis." Neuron 57 (1):41-55. doi: 10.1016/j.neuron.2007.11.018.

Cifuentes, D., H. Xue, D. W. Taylor, H. Patnode, Y. Mishima, S. Cheloufi, E. Ma, S. Mane, G. J. Hannon, N. D. Lawson, S. A. Wolfe, and A. J. Giraldez. 2010. "A novel miRNA processing pathway independent of Dicer requires Argonaute2 catalytic activity." Science 328 (5986):1694-8. doi: 10.1126/science.1190809.

Clark, A. G., M. B. Eisen, D. R. Smith, C. M. Bergman, B. Oliver, T. A. Markow, T. C. Kaufman, M. Kellis, W. Gelbart, V. N. Iyer, D. A. Pollard, T. B. Sackton, A. M. Larracuente, N. D. Singh, J. P. Abad, D. N. Abt, B. Adryan, M. Aguade, H.

Akashi, W. W. Anderson, C. F. Aquadro, D. H. Ardell, R. Arguello, C. G. Artieri, D. A. Barbash, D. Barker, P. Barsanti, P. Batterham, S. Batzoglou, D. Begun, A. Bhutkar, E. Blanco, S. A. Bosak, R. K. Bradley, A. D. Brand, M. R. Brent, A. N. Brooks, R. H. Brown, R. K. Butlin, C. Caggese, B. R. Calvi, A. Bernardo de Carvalho, A. Caspi, S. Castrezana, S. E. Celniker, J. L. Chang, C. Chapple, S. Chatterji, A. Chinwalla, A. Civetta, S. W. Clifton, J. M. Comeron, J. C. Costello, J. A. Coyne, J. Daub, R. G. David, A. L. Delcher, K. Delehaunty, C. B. Do, H. Ebling, K. Edwards, T. Eickbush, J. D. Evans, A. Filipski, S. Findeiss, E. Freyhult, L. Fulton, R. Fulton, A. C. Garcia, A. Gardiner, D. A. Garfield, B. E. Garvin, G. Gibson, D. Gilbert, S. Gnerre, J. Godfrey, R. Good, V. Gotea, B. Gravely, A. J. Greenberg, S. Griffiths-Jones, S. Gross, R. Guigo, E. A. Gustafson, W. Haerty, M. W. Hahn, D. L. Halligan, A. L. Halpern, G. M. Halter, M. V. Han, A. Heger, L. Hillier, A. S. Hinrichs, I. Holmes, R. A. Hoskins, M. J. Hubisz, D. Hultmark, M. A. Huntley, D. B. Jaffe, S. Jagadeeshan, W. R. Jeck, J. Johnson, C. D. Jones, W. C. Jordan, G. H. Karpen, E. Kataoka, P. D. Keightley, P. Kheradpour, E. F. Kirkness, L. B. Koerich, K. Kristiansen, D. Kudrna, R. J. Kulathinal, S. Kumar, R. Kwok, E. Lander, C. H. Langley, R. Lapoint, B. P. Lazzaro, S. J. Lee, L. Levesque, R. Li, C. F. Lin, M. F. Lin, K. Lindblad-Toh, A. Llopart, M. Long, L. Low, E. Lozovsky, J. Lu, M. Luo, C. A. Machado, W. Makalowski, M. Marzo, M. Matsuda, L. Matzkin, B. McAllister, C. S. McBride, B. McKernan, K. McKernan, M. Mendez-Lago, P. Minx, M. U. Mollenhauer, K. Montooth, S. M. Mount, X. Mu, E. Myers, B. Negre, S. Newfeld, R. Nielsen, M. A. Noor, P. O'Grady, L. Pachter, M. Papaceit, M. J. Parisi, M. Parisi, L. Parts, J. S. Pedersen, G. Pesole, A. M. Phillippy, C. P. Ponting, M. Pop, D. Porcelli, J. R. Powell, S. Prohaska, K. Pruitt, M. Puig, H. Quesneville, K. R. Ram, D. Rand, M. D. Rasmussen, L. K. Reed, R. Reenan, A. Reily, K. A. Remington, T. T. Rieger, M. G. Ritchie, C. Robin, Y. H. Rogers, C. Rohde, J. Rozas, M. J. Rubenfield, A. Ruiz, S. Russo, S. L. Salzberg, A. Sanchez-Gracia, D. J. Saranga, H. Sato, S. W. Schaeffer, M. C. Schatz, T. Schlenke, R. Schwartz, C. Segarra, R. S. Singh, L. Sirot, M. Sirota, N. B. Sisneros, C. D. Smith, T. F. Smith, J. Spieth, D. E. Stage, A. Stark, W. Stephan, R. L. Strausberg, S. Strempel, D. Sturgill, G. Sutton, G. G. Sutton, W. Tao, S. Teichmann, Y. N. Tobari, Y. Tomimura, J. M. Tsolas, V. L. Valente, E. Venter, J. C. Venter, S. Vicario, F. G. Vieira, A. J. Vilella, A. Villasante, B. Walenz, J. Wang, M. Wasserman, T. Watts, D. Wilson, R. K. Wilson, R. A. Wing, M. F. Wolfner, A. Wong, G. K. Wong, C. I. Wu, G. Wu, D. Yamamoto, H. P. Yang, S. P. Yang, J. A. Yorke, K. Yoshida, E. Zdobnov, P. Zhang, Y. Zhang, A. V. Zimin, J. Baldwin, A. Abdouelleil, J. Abdulkadir, A. Abebe, B. Abera, J. Abreu, S. C. Acer, L. Aftuck, A. Alexander, P. An, E. Anderson, S. Anderson, H. Arachi, M. Azer, P. Bachantsang, A. Barry, T. Bayul, A. Berlin, D. Bessette, T. Bloom, J. Blye, L. Boguslavskiy, C. Bonnet, B. Boukhgalter, I. Bourzgui, A. Brown, P. Cahill, S. Channer, Y. Cheshatsang, L. Chuda, M. Citroen, A. Collymore, P. Cooke, M. Costello, K. D'Aco, R. Daza, G. De Haan, S. DeGray, C. DeMaso, N. Dhargay, K. Dooley, E. Dooley, M. Doricent, P. Dorje, K. Dorjee, A. Dupes, R. Elong, J. Falk, A. Farina, S. Faro, D. Ferguson, S. Fisher, C. D. Foley, A. Franke, D. Friedrich, L. Gadbois, G. Gearin, C. R. Gearin, G. Giannoukos, T. Goode, J. Graham, E. Grandbois, S. Grewal, K.

Gyaltsen, N. Hafez, B. Hagos, J. Hall, C. Henson, A. Hollinger, T. Honan, M. D. Huard, L. Hughes, B. Hurhula, M. E. Husby, A. Kamat, B. Kanga, S. Kashin, D. Khazanovich, P. Kisner, K. Lance, M. Lara, W. Lee, N. Lennon, F. Letendre, R. LeVine, A. Lipovsky, X. Liu, J. Liu, S. Liu, T. Lokyitsang, Y. Lokyitsang, R. Lubonja, A. Lui, P. MacDonald, V. Magnisalis, K. Maru, C. Matthews, W. McCusker, S. McDonough, T. Mehta, J. Meldrim, L. Meneus, O. Mihai, A. Mihalev, T. Mihova, R. Mittelman, V. Mlenga, A. Montmayeur, L. Mulrain, A. Navidi, J. Naylor, T. Negash, T. Nguyen, N. Nguyen, R. Nicol, C. Norbu, N. Norbu, N. Novod, B. O'Neill, S. Osman, E. Markiewicz, O. L. Oyono, C. Patti, P. Phunkhang, F. Pierre, M. Priest, S. Raghuraman, F. Rege, R. Reyes, C. Rise, P. Rogov, K. Ross, E. Ryan, S. Settipalli, T. Shea, N. Sherpa, L. Shi, D. Shih, T. Sparrow, J. Spaulding, J. Stalker, N. Stange-Thomann, S. Stavropoulos, C. Stone, C. Strader, S. Tesfaye, T. Thomson, Y. Thoulutsang, D. Thoulutsang, K. Topham, I. Topping, T. Tsamla, H. Vassiliev, A. Vo, T. Wangchuk, T. Wangdi, M. Weiand, J. Wilkinson, A. Wilson, S. Yadav, G. Young, Q. Yu, L. Zembek, D. Zhong, A. Zimmer, Z. Zwirko, P. Alvarez, W. Brockman, J. Butler, C. Chin, M. Grabherr, M. Kleber, E. Mauceli, and I. MacCallum. 2007. "Evolution of genes and genomes on the Drosophila phylogeny." Nature 450 (7167):203-18. doi: nature06341 [pii]
10.1038/nature06341.

Cole, C., A. Sobala, C. Lu, S. R. Thatcher, A. Bowman, J. W. Brown, P. J. Green, G. J. Barton, and G. Hutvagner. 2009. "Filtering of deep sequencing data reveals the existence of abundant Dicer-dependent small RNAs derived from tRNAs." RNA 15 (12):2147-60. doi: rna.1738409 [pii]
10.1261/rna.1738409.

Czech, B., and G. J. Hannon. 2011. "Small RNA sorting: matchmaking for Argonautes." Nat Rev Genet 12 (1):19-31. doi: 10.1038/nrg2916.

Czech, B., C. D. Malone, R. Zhou, A. Stark, C. Schlingeheyde, M. Dus, N. Perrimon, M. Kellis, J. A. Wohlschlegel, R. Sachidanandam, G. J. Hannon, and J. Brennecke. 2008. "An endogenous small interfering RNA pathway in Drosophila." Nature 453 (7196):798-802. doi: 10.1038/nature07007.

Czech, B., R. Zhou, Y. Erlich, J. Brennecke, R. Binari, C. Villalta, A. Gordon, N. Perrimon, and G. J. Hannon. 2009. "Hierarchical rules for Argonaute loading in Drosophila." Mol Cell 36 (3):445-56. doi: 10.1016/j.molcel.2009.09.028.

de Lencastre, A., Z. Pincus, K. Zhou, M. Kato, S. S. Lee, and F. J. Slack. 2010. "MicroRNAs both promote and antagonize longevity in C. elegans." Curr Biol 20 (24):2159-68. doi: 10.1016/j.cub.2010.11.015.

De Rijk, P., and R. De Wachter. 1997. "RnaViz, a program for the visualisation of RNA secondary structure." Nucleic Acids Res 25 (22):4679-84. doi: gka734 [pii].

De Rijk, P., J. Wuyts, and R. De Wachter. 2003. "RnaViz 2: an improved representation of RNA secondary structure." Bioinformatics 19 (2):299-300.

Denli, A. M., B. B. Tops, R. H. Plasterk, R. F. Ketting, and G. J. Hannon. 2004. "Processing of primary microRNAs by the Microprocessor complex." Nature 432 (7014):231-5. doi: 10.1038/nature03049.

Elefant, N., Y. Altuvia, and H. Margalit. 2011. "A wide repertoire of miRNA binding sites: prediction and functional implications." Bioinformatics 27 (22):3093-101. doi: 10.1093/bioinformatics/btr534.

Elia, L., R. Contu, M. Quintavalle, F. Varrone, C. Chimenti, M. A. Russo, V. Cimino, L. De Marinis, A. Frustaci, D. Catalucci, and G. Condorelli. 2009. "Reciprocal regulation of microRNA-1 and insulin-like growth factor-1 signal transduction cascade in cardiac and skeletal muscle in physiological and pathological conditions." Circulation 120 (23):2377-85. doi: 10.1161/CIRCULATIONAHA.109.879429.

Emara, M. M., P. Ivanov, T. Hickman, N. Dawra, S. Tisdale, N. Kedersha, G. F. Hu, and P. Anderson. 2010. "Angiogenin-induced tRNA-derived stress-induced RNAs promote stress-induced stress granule assembly." J Biol Chem 285 (14):10959-68. doi: M109.077560 [pii]
10.1074/jbc.M109.077560.

Fabian, M. R., and N. Sonenberg. 2012. "The mechanics of miRNA-mediated gene silencing: a look under the hood of miRISC." Nat Struct Mol Biol 19 (6):586-93. doi: 10.1038/nsmb.2296.

Fang, M., J. Wang, X. Zhang, Y. Geng, Z. Hu, J. A. Rudd, S. Ling, W. Chen, and S. Han. 2012. "The miR-124 regulates the expression of BACE1/beta-secretase correlated with cell death in Alzheimer's disease." Toxicol Lett 209 (1):94-105. doi: 10.1016/j.toxlet.2011.11.032.

Fenelon, K., J. Mukai, B. Xu, P. K. Hsu, L. J. Drew, M. Karayiorgou, G. D. Fischbach, A. B. Macdermott, and J. A. Gogos. 2011. "Deficiency of Dgcr8, a gene disrupted by the 22q11.2 microdeletion, results in altered short-term plasticity in the prefrontal cortex." Proc Natl Acad Sci U S A 108 (11):4447-52. doi: 10.1073/pnas.1101219108.

Fernandez-Valverde, S. L., R. J. Taft, and J. S. Mattick. 2010. "Dynamic isomiR regulation in Drosophila development." RNA 16 (10):1881-8. doi: 10.1261/rna.2379610.

Ferraiuolo, L., J. Kirby, A. J. Grierson, M. Sendtner, and P. J. Shaw. 2011. "Molecular pathways of motor neuron injury in amyotrophic lateral sclerosis." Nat Rev Neurol 7 (11):616-30. doi: 10.1038/nrneurol.2011.152.

Fischer, S., J. Benz, B. Spath, A. Jellen-Ritter, R. Heyer, M. Dorr, L. K. Maier, C. Menzel-Hobeck, M. Lehr, K. Jantzer, J. Babski, J. Soppa, and A. Marchfelder. 2011. "Regulatory RNAs in Haloferax volcanii." Biochem Soc Trans 39 (1):159-62. doi: BST0390159 [pii]
10.1042/BST0390159.

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 (6):900-7. doi: 10.1016/j.molcel.2010.06.014.

Flynt, A. S., and E. C. Lai. 2008. "Biological principles of microRNA-mediated regulation: shared themes amid diversity." Nat Rev Genet 9 (11):831-42. doi: 10.1038/nrg2455.

Forstemann, K., M. D. Horwich, L. Wee, Y. Tomari, and P. D. Zamore. 2007. "Drosophila microRNAs are sorted into functionally distinct argonaute complexes after production by dicer-1." Cell 130 (2):287-97. doi: 10.1016/j.cell.2007.05.056.

Forstemann, K., Y. Tomari, T. Du, V. V. Vagin, A. M. Denli, D. P. Bratu, C. Klattenhoff, W. E. Theurkauf, and P. D. Zamore. 2005. "Normal microRNA maturation and germ-line stem cell maintenance requires Loquacious, a double-stranded RNA-binding domain protein." PLoS Biol 3 (7):e236. doi: 10.1371/journal.pbio.0030236.

Friedman, R. C., K. K. Farh, C. B. Burge, and D. P. Bartel. 2009. "Most mammalian mRNAs are conserved targets of microRNAs." Genome Res 19 (1):92-105. doi: 10.1101/gr.082701.108.

Fu, H., J. Feng, Q. Liu, F. Sun, Y. Tie, J. Zhu, R. Xing, Z. Sun, and X. Zheng. 2009. "Stress induces tRNA cleavage by angiogenin in mammalian cells." FEBS Lett 583 (2):437-42. doi: 10.1016/j.febslet.2008.12.043.

Fukunaga, R., B. W. Han, J. H. Hung, J. Xu, Z. Weng, and P. D. Zamore. 2012. "Dicer partner proteins tune the length of mature miRNAs in flies and mammals." Cell 151 (3):533-46. doi: 10.1016/j.cell.2012.09.027.

Garber, K., K. T. Smith, D. Reines, and S. T. Warren. 2006. "Transcription, translation and fragile X syndrome." Curr Opin Genet Dev 16 (3):270-5. doi: S0959-437X(06)00071-2 [pii]
10.1016/j.gde.2006.04.010.

Garcia-Lopez, J., D. Hourcade Jde, and J. Del Mazo. 2013. "Reprogramming of microRNAs by adenosine-to-inosine editing and the selective elimination of edited microRNA precursors in mouse oocytes and preimplantation embryos." Nucleic Acids Res 41 (10):5483-93. doi: 10.1093/nar/gkt247.

Garcia-Silva, M. R., F. Cabrera-Cabrera, M. C. Guida, and A. Cayota. 2012. "Hints of tRNA-Derived Small RNAs Role in RNA Silencing Mechanisms." Genes (Basel) 3 (4):603-14. doi: genes3040603 [pii]
10.3390/genes3040603.

Gebetsberger, J., M. Zywicki, A. Kunzi, and N. Polacek. 2012. "tRNA-derived fragments target the ribosome and function as regulatory non-coding RNA in Haloferax volcanii." Archaea 2012:260909. doi: 10.1155/2012/260909.

Gehrke, S., Y. Imai, N. Sokol, and B. Lu. 2010. "Pathogenic LRRK2 negatively regulates microRNA-mediated translational repression." Nature 466 (7306):637-41. doi: 10.1038/nature09191.

Gerbasi, V. R., D. E. Golden, S. B. Hurtado, and E. J. Sontheimer. 2010. "Proteomics identification of Drosophila small interfering RNA-associated factors." Mol Cell Proteomics 9 (9):1866-72. doi: 10.1074/mcp.M900614-MCP200.

Ghildiyal, M., H. Seitz, M. D. Horwich, C. Li, T. Du, S. Lee, J. Xu, E. L. Kittler, M. L. Zapp, Z. Weng, and P. D. Zamore. 2008. "Endogenous siRNAs derived from transposons and mRNAs in Drosophila somatic cells." Science 320 (5879):1077-81. doi: 10.1126/science.1157396.

Ghildiyal, M., J. Xu, H. Seitz, Z. Weng, and P. D. Zamore. 2009. "Sorting of Drosophila small silencing RNAs partitions microRNA* strands into the RNA interference pathway." RNA 16 (1):43-56. doi: rna.1972910 [pii]
10.1261/rna.1972910.

Ghildiyal, M., J. Xu, H. Seitz, Z. Weng, and P. D. Zamore. 2010. "Sorting of Drosophila small silencing RNAs partitions microRNA* strands into the RNA interference pathway." RNA 16 (1):43-56. doi: 10.1261/rna.1972910.

Gong, B., Y. S. Lee, I. Lee, T. R. Shelite, N. Kunkeaw, G. Xu, K. Lee, S. H. Jeon, B. H. Johnson, Q. Chang, T. Ha, N. L. Mendell, X. Cheng, D. H. Bouyer, P. J. Boor, T. G. Ksiazek, and D. H. Walker. 2013. "Compartmentalized, functional role of angiogenin during spotted fever group rickettsia-induced endothelial barrier dysfunction: evidence of possible mediation by host tRNA-derived small noncoding RNAs." BMC Infect Dis 13:285. doi: 1471-2334-13-285 [pii]
10.1186/1471-2334-13-285.

Griffiths-Jones, S. 2006. "miRBase: the microRNA sequence database." Methods Mol Biol 342:129-38. doi: 1-59745-123-1:129 [pii]
10.1385/1-59745-123-1:129.

Griffiths-Jones, S. 2010. "miRBase: microRNA sequences and annotation." Curr Protoc Bioinformatics Chapter 12:Unit 12 9 1-10. doi: 10.1002/0471250953.bi1209s29.

Grimson, A., K. K. Farh, W. K. Johnston, P. Garrett-Engele, L. P. Lim, and D. P. Bartel. 2007. "MicroRNA targeting specificity in mammals: determinants beyond seed pairing." Mol Cell 27 (1):91-105. doi: S1097-2765(07)00407-8 [pii]
10.1016/j.molcel.2007.06.017.

Hagan, J. P., E. Piskounova, and R. I. Gregory. 2009. "Lin28 recruits the TUTase Zcchc11 to inhibit let-7 maturation in mouse embryonic stem cells." Nat Struct Mol Biol 16 (10):1021-5. doi: 10.1038/nsmb.1676.

Haiser, H. J., F. V. Karginov, G. J. Hannon, and M. A. Elliot. 2008. "Developmentally regulated cleavage of tRNAs in the bacterium Streptomyces coelicolor." Nucleic Acids Res 36 (3):732-41. doi: gkm1096 [pii]
10.1093/nar/gkm1096.

Hammond, S. M., S. Boettcher, A. A. Caudy, R. Kobayashi, and G. J. Hannon. 2001. "Argonaute2, a link between genetic and biochemical analyses of RNAi." Science 293 (5532):1146-50. doi: 10.1126/science.1064023.

Han, B. W., J. H. Hung, Z. Weng, P. D. Zamore, and S. L. Ameres. 2011. "The 3'-to-5' exoribonuclease Nibbler shapes the 3' ends of microRNAs bound to Drosophila Argonaute1." Curr Biol 21 (22):1878-87. doi: 10.1016/j.cub.2011.09.034.

Han, J., Y. Lee, K. H. Yeom, Y. K. Kim, H. Jin, and V. N. Kim. 2004. "The Drosha-DGCR8 complex in primary microRNA processing." Genes Dev 18 (24):3016-27. doi: 10.1101/gad.1262504.

Han, J., Y. Lee, K. H. Yeom, J. W. Nam, I. Heo, J. K. Rhee, S. Y. Sohn, Y. Cho, B. T. Zhang, and V. N. Kim. 2006. "Molecular basis for the recognition of primary microRNAs by the Drosha-DGCR8 complex." Cell 125 (5):887-901. doi: 10.1016/j.cell.2006.03.043.

Haussecker, D., Y. Huang, A. Lau, P. Parameswaran, A. Z. Fire, and M. A. Kay. 2010. "Human tRNA-derived small RNAs in the global regulation of RNA silencing." RNA 16 (4):673-95. doi: 10.1261/rna.2000810.

Hebert, S. S., K. Horre, L. Nicolai, A. S. Papadopoulou, W. Mandemakers, A. N. Silahtaroglu, S. Kauppinen, A. Delacourte, and B. De Strooper. 2008. "Loss of microRNA cluster miR-29a/b-1 in sporadic Alzheimer's disease correlates with increased BACE1/beta-secretase expression." Proc Natl Acad Sci U S A 105 (17):6415-20. doi: 10.1073/pnas.0710263105.

Heo, I., C. Joo, Y. K. Kim, M. Ha, M. J. Yoon, J. Cho, K. H. Yeom, J. Han, and V. N. Kim. 2009. "TUT4 in concert with Lin28 suppresses microRNA biogenesis through pre-microRNA uridylation." Cell 138 (4):696-708. doi: 10.1016/j.cell.2009.08.002.

Horwich, M. D., C. Li, C. Matranga, V. Vagin, G. Farley, P. Wang, and P. D. Zamore. 2007. "The Drosophila RNA methyltransferase, DmHen1, modifies germline

piRNAs and single-stranded siRNAs in RISC." Curr Biol 17 (14):1265-72. doi: 10.1016/j.cub.2007.06.030.

Hurst, S. R., R. F. Hough, P. J. Aruscavage, and B. L. Bass. 1995. "Deamination of mammalian glutamate receptor RNA by Xenopus dsRNA adenosine deaminase: similarities to in vivo RNA editing." RNA 1 (10):1051-60.

Hutvagner, G., J. McLachlan, A. E. Pasquinelli, E. Balint, T. Tuschl, and P. D. Zamore. 2001. "A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA." Science 293 (5531):834-8. doi: 10.1126/science.1062961.

Ibanez-Ventoso, C., M. Yang, S. Guo, H. Robins, R. W. Padgett, and M. Driscoll. 2006. "Modulated microRNA expression during adult lifespan in Caenorhabditis elegans." Aging Cell 5 (3):235-46. doi: 10.1111/j.1474-9726.2006.00210.x.

Iki, T., M. Yoshikawa, M. Nishikiori, M. C. Jaudal, E. Matsumoto-Yokoyama, I. Mitsuhara, T. Meshi, and M. Ishikawa. 2010. "In vitro assembly of plant RNA-induced silencing complexes facilitated by molecular chaperone HSP90." Mol Cell 39 (2):282-91. doi: 10.1016/j.molcel.2010.05.014.

Ishizuka, A., M. C. Siomi, and H. Siomi. 2002. "A Drosophila fragile X protein interacts with components of RNAi and ribosomal proteins." Genes Dev 16 (19):2497-508. doi: 10.1101/gad.1022002.

Ivanov, P., M. M. Emara, J. Villen, S. P. Gygi, and P. Anderson. 2011. "Angiogenin-induced tRNA fragments inhibit translation initiation." Mol Cell 43 (4):613-23. doi: S1097-2765(11)00524-7 [pii]
10.1016/j.molcel.2011.06.022.

Iwasaki, S., T. Kawamata, and Y. Tomari. 2009. "Drosophila argonaute1 and argonaute2 employ distinct mechanisms for translational repression." Mol Cell 34 (1):58-67. doi: 10.1016/j.molcel.2009.02.010.

Iwasaki, S., M. Kobayashi, M. Yoda, Y. Sakaguchi, S. Katsuma, T. Suzuki, and Y. Tomari. 2010. "Hsc70/Hsp90 chaperone machinery mediates ATP-dependent RISC loading of small RNA duplexes." Mol Cell 39 (2):292-9. doi: 10.1016/j.molcel.2010.05.015.

Jiang, F., X. Ye, X. Liu, L. Fincher, D. McKearin, and Q. Liu. 2005. "Dicer-1 and R3D1-L catalyze microRNA maturation in Drosophila." Genes Dev 19 (14):1674-9. doi: 10.1101/gad.1334005.

Jin, P., D. C. Zarnescu, S. Ceman, M. Nakamoto, J. Mowrey, T. A. Jongens, D. L. Nelson, K. Moses, and S. T. Warren. 2004. "Biochemical and genetic interaction between the fragile X mental retardation protein and the microRNA pathway." Nat Neurosci 7 (2):113-7. doi: 10.1038/nn1174
nn1174 [pii].

Jochl, C., M. Rederstorff, J. Hertel, P. F. Stadler, I. L. Hofacker, M. Schrettl, H. Haas, and A. Huttenhofer. 2008. "Small ncRNA transcriptome analysis from Aspergillus fumigatus suggests a novel mechanism for regulation of protein synthesis." Nucleic Acids Res 36 (8):2677-89. doi: gkn123 [pii] 10.1093/nar/gkn123.

Johnson, S. M., S. Y. Lin, and F. J. Slack. 2003. "The time of appearance of the C. elegans let-7 microRNA is transcriptionally controlled utilizing a temporal regulatory element in its promoter." Dev Biol 259 (2):364-79.

Kamminga, L. M., M. J. Luteijn, M. J. den Broeder, S. Redl, L. J. Kaaij, E. F. Roovers, P. Ladurner, E. Berezikov, and R. F. Ketting. 2010. "Hen1 is required for oocyte development and piRNA stability in zebrafish." EMBO J 29 (21):3688-700. doi: 10.1038/emboj.2010.233.

Karp, X., M. Hammell, M. C. Ow, and V. Ambros. 2011. "Effect of life history on microRNA expression during C. elegans development." RNA 17 (4):639-51. doi: 10.1261/rna.2310111.

Kato, M., X. Chen, S. Inukai, H. Zhao, and F. J. Slack. 2011. "Age-associated changes in expression of small, noncoding RNAs, including microRNAs, in C. elegans." RNA 17 (10):1804-20. doi: 10.1261/rna.2714411.

Katoh, T., Y. Sakaguchi, K. Miyauchi, T. Suzuki, S. Kashiwabara, T. Baba, and T. Suzuki. 2009. "Selective stabilization of mammalian microRNAs by 3' adenylation mediated by the cytoplasmic poly(A) polymerase GLD-2." Genes Dev 23 (4):433-8. doi: 10.1101/gad.1761509.

Kawamura, Y., K. Saito, T. Kin, Y. Ono, K. Asai, T. Sunohara, T. N. Okada, M. C. Siomi, and H. Siomi. 2008. "Drosophila endogenous small RNAs bind to Argonaute 2 in somatic cells." Nature 453 (7196):793-7. doi: 10.1038/nature06938.

Kenyon, C. 2010. "A pathway that links reproductive status to lifespan in Caenorhabditis elegans." Ann N Y Acad Sci 1204:156-62. doi: 10.1111/j.1749-6632.2010.05640.x.

Ketting, R. F., T. H. Haverkamp, H. G. van Luenen, and R. H. Plasterk. 1999. "Mut-7 of C. elegans, required for transposon silencing and RNA interference, is a homolog of Werner syndrome helicase and RNaseD." Cell 99 (2):133-41.

Khvorova, A., A. Reynolds, and S. D. Jayasena. 2003. "Functional siRNAs and miRNAs exhibit strand bias." Cell 115 (2):209-16.

Kim, J., K. Inoue, J. Ishii, W. B. Vanti, S. V. Voronov, E. Murchison, G. Hannon, and A. Abeliovich. 2007. "A MicroRNA feedback circuit in midbrain dopamine neurons." Science 317 (5842):1220-4. doi: 10.1126/science.1140481.

Kim, U., Y. Wang, T. Sanford, Y. Zeng, and K. Nishikura. 1994. "Molecular cloning of cDNA for double-stranded RNA adenosine deaminase, a candidate enzyme for nuclear RNA editing." Proc Natl Acad Sci U S A 91 (24):11457-61.

Kim, V. N. 2004. "MicroRNA precursors in motion: exportin-5 mediates their nuclear export." Trends Cell Biol 14 (4):156-9.

Kim, V. N. 2005. "MicroRNA biogenesis: coordinated cropping and dicing." Nat Rev Mol Cell Biol 6 (5):376-85. doi: nrm1644 [pii]
10.1038/nrm1644.

Kirino, Y., and Z. Mourelatos. 2007a. "2'-O-methyl modification in mouse piRNAs and its methylase." Nucleic Acids Symp Ser (Oxf) (51):417-8. doi: 10.1093/nass/nrm209.

Kirino, Y., and Z. Mourelatos. 2007b. "The mouse homolog of HEN1 is a potential methylase for Piwi-interacting RNAs." RNA 13 (9):1397-401. doi: 10.1261/rna.659307.

Kirino, Y., and Z. Mourelatos. 2007c. "Mouse Piwi-interacting RNAs are 2'-O-methylated at their 3' termini." Nat Struct Mol Biol 14 (4):347-8. doi: 10.1038/nsmb1218.

Krueger, N. X., D. Van Vactor, H. I. Wan, W. M. Gelbart, C. S. Goodman, and H. Saito. 1996. "The transmembrane tyrosine phosphatase DLAR controls motor axon guidance in Drosophila." Cell 84 (4):611-22. doi: S0092-8674(00)81036-3 [pii].

Kumar, P., J. Anaya, S. B. Mudunuri, and A. Dutta. 2014. "Meta-analysis of tRNA derived RNA fragments reveals that they are evolutionarily conserved and associate with AGO proteins to recognize specific RNA targets." BMC Biol 12 (1):78. doi: 10.1186/s12915-014-0078-0.

Kumar, V., R. Fricke, D. Bhar, S. Reddy-Alla, K. S. Krishnan, S. Bogdan, and M. Ramaswami. 2009. "Syndapin promotes formation of a postsynaptic membrane system in Drosophila." Mol Biol Cell 20 (8):2254-64. doi: E08-10-1072 [pii]
10.1091/mbc.E08-10-1072.

Kurth, H. M., and K. Mochizuki. 2009. "2'-O-methylation stabilizes Piwi-associated small RNAs and ensures DNA elimination in Tetrahymena." RNA 15 (4):675-85. doi: 10.1261/rna.1455509.

La Rocca, G., M. Badin, B. Shi, S. Q. Xu, T. Deangelis, L. Sepp-Lorenzinoi, and R. Baserga. 2009. "Mechanism of growth inhibition by MicroRNA 145: the role of the IGF-I receptor signaling pathway." J Cell Physiol 220 (2):485-91. doi: 10.1002/jcp.21796.

La Rocca, G., B. Shi, M. Badin, T. De Angelis, L. Sepp-Lorenzino, and R. Baserga. 2009. "Growth inhibition by microRNAs that target the insulin receptor substrate-1." Cell Cycle 8 (14):2255-9.

Lai, E. C. 2003. "microRNAs: runts of the genome assert themselves." Curr Biol 13 (23):R925-36.

Lai, E. C., B. Tam, and G. M. Rubin. 2005. "Pervasive regulation of Drosophila Notch target genes by GY-box-, Brd-box-, and K-box-class microRNAs." Genes Dev 19 (9):1067-80. doi: 10.1101/gad.1291905.

Lai, E. C., P. Tomancak, R. W. Williams, and G. M. Rubin. 2003. "Computational identification of Drosophila microRNA genes." Genome Biol 4 (7):R42. doi: 10.1186/gb-2003-4-7-r42.

Landgraf, P., M. Rusu, R. Sheridan, A. Sewer, N. Iovino, A. Aravin, S. Pfeffer, A. Rice, A. O. Kamphorst, M. Landthaler, C. Lin, N. D. Socci, L. Hermida, V. Fulci, S. Chiaretti, R. Foa, J. Schliwka, U. Fuchs, A. Novosel, R. U. Muller, B. Schermer, U. Bissels, J. Inman, Q. Phan, M. Chien, D. B. Weir, R. Choksi, G. De Vita, D. Frezzetti, H. I. Trompeter, V. Hornung, G. Teng, G. Hartmann, M. Palkovits, R. Di Lauro, P. Wernet, G. Macino, C. E. Rogler, J. W. Nagle, J. Ju, F. N. Papavasiliou, T. Benzing, P. Lichter, W. Tam, M. J. Brownstein, A. Bosio, A. Borkhardt, J. J. Russo, C. Sander, M. Zavolan, and T. Tuschl. 2007. "A mammalian microRNA expression atlas based on small RNA library sequencing." Cell 129 (7):1401-14. doi: 10.1016/j.cell.2007.04.040.

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 (3):R25. doi: 10.1186/gb-2009-10-3-r25.

Lee, H. Y., and J. A. Doudna. 2012. "TRBP alters human precursor microRNA processing in vitro." RNA 18 (11):2012-9. doi: 10.1261/rna.035501.112.

Lee, R. C., R. L. Feinbaum, and V. Ambros. 1993. "The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14." Cell 75 (5):843-54.

Lee, Y. S., K. Nakahara, J. W. Pham, K. Kim, Z. He, E. J. Sontheimer, and R. W. Carthew. 2004. "Distinct roles for Drosophila Dicer-1 and Dicer-2 in the siRNA/miRNA silencing pathways." Cell 117 (1):69-81.

Lee, Y. S., Y. Shibata, A. Malhotra, and A. Dutta. 2009. "A novel class of small RNAs: tRNA-derived RNA fragments (tRFs)." Genes Dev 23 (22):2639-49. doi: 23/22/2639 [pii]
10.1101/gad.1837609.

Lehrbach, N. J., C. Castro, K. J. Murfitt, C. Abreu-Goodger, J. L. Griffin, and E. A. Miska. 2012. "Post-developmental microRNA expression is required for normal

physiology, and regulates aging in parallel to insulin/IGF-1 signaling in C. elegans." RNA 18 (12):2220-35. doi: 10.1261/rna.035402.112.

Levitz, R., D. Chapman, M. Amitsur, R. Green, L. Snyder, and G. Kaufmann. 1990. "The optional E. coli prr locus encodes a latent form of phage T4-induced anticodon nuclease." EMBO J 9 (5):1383-9.

Lewis, B. P., C. B. Burge, and D. P. Bartel. 2005. "Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets." Cell 120 (1):15-20. doi: S0092867404012607 [pii]
10.1016/j.cell.2004.12.035.

Lewis, B. P., I. H. Shih, M. W. Jones-Rhoades, D. P. Bartel, and C. B. Burge. 2003. "Prediction of mammalian microRNA targets." Cell 115 (7):787-98. doi: S0092867403010183 [pii].

Li, J. B., and G. M. Church. 2013. "Deciphering the functions and regulation of brain-enriched A-to-I RNA editing." Nat Neurosci 16 (11):1518-22. doi: 10.1038/nn.3539.

Li, J., Z. Yang, B. Yu, J. Liu, and X. Chen. 2005. "Methylation protects miRNAs and siRNAs from a 3'-end uridylation activity in Arabidopsis." Curr Biol 15 (16):1501-7. doi: 10.1016/j.cub.2005.07.029.

Li, S. C., Y. L. Liao, M. R. Ho, K. W. Tsai, C. H. Lai, and W. C. Lin. 2012. "miRNA arm selection and isomiR distribution in gastric cancer." BMC Genomics 13 Suppl 1:S13. doi: 10.1186/1471-2164-13-S1-S13.

Li, W., L. Prazak, N. Chatterjee, S. Gruninger, L. Krug, D. Theodorou, and J. Dubnau. 2013. "Activation of transposable elements during aging and neuronal decline in Drosophila." Nat Neurosci 16 (5):529-31. doi: 10.1038/nn.3368.

Li, Y., J. Luo, H. Zhou, J. Y. Liao, L. M. Ma, Y. Q. Chen, and L. H. Qu. 2008. "Stress-induced tRNA-derived RNAs: a novel class of small RNAs in the primitive eukaryote Giardia lamblia." Nucleic Acids Res 36 (19):6048-55. doi: gkn596 [pii]
10.1093/nar/gkn596.

Li, Z., C. Ender, G. Meister, P. S. Moore, Y. Chang, and B. John. 2012. "Extensive terminal and asymmetric processing of small RNAs from rRNAs, snoRNAs, snRNAs, and tRNAs." Nucleic Acids Res 40 (14):6787-99. doi: gks307 [pii]
10.1093/nar/gks307.

Liu, N., M. Abe, L. R. Sabin, G. J. Hendriks, A. S. Naqvi, Z. Yu, S. Cherry, and N. M. Bonini. 2011. "The exoribonuclease Nibbler controls 3' end processing of microRNAs in Drosophila." Curr Biol 21 (22):1888-93. doi: 10.1016/j.cub.2011.10.006.

Liu, N., M. Landreh, K. Cao, M. Abe, G. J. Hendriks, J. R. Kennerdell, Y. Zhu, L. S. Wang, and N. M. Bonini. 2012. "The microRNA miR-34 modulates ageing and neurodegeneration in Drosophila." Nature 482 (7386):519-23. doi: 10.1038/nature10810.

Liu, Q., T. A. Rand, S. Kalidas, F. Du, H. E. Kim, D. P. Smith, and X. Wang. 2003. "R2D2, a bridge between the initiation and effector steps of the Drosophila RNAi pathway." Science 301 (5641):1921-5. doi: 10.1126/science.1088710.

Llorens, F., M. Hummel, L. Pantano, X. Pastor, A. Vivancos, E. Castillo, H. Matllin, A. Ferrer, M. Ingham, M. Noguera, R. Kofler, J. C. Dohm, R. Pluvinet, M. Bayes, H. Himmelbauer, J. A. Del Rio, E. Marti, and L. Sumoy. 2013. "Microarray and deep sequencing cross-platform analysis of the mirRNome and isomiR variation in response to epidermal growth factor." BMC Genomics 14 (1):371. doi: 10.1186/1471-2164-14-371.

Loss-Morais, G., P. M. Waterhouse, and R. Margis. 2014. "Description of plant tRNA-derived RNA fragments (tRFs) associated with argonaute and identification of their putative targets." Biol Direct 8:6. doi: 1745-6150-8-6 [pii]
10.1186/1745-6150-8-6.

Luciano, D. J., H. Mirsky, N. J. Vendetti, and S. Maas. 2004. "RNA editing of a miRNA precursor." RNA 10 (8):1174-7. doi: 10.1261/rna.7350304
10/8/1174 [pii].

Lund, E., S. Guttinger, A. Calado, J. E. Dahlberg, and U. Kutay. 2004. "Nuclear export of microRNA precursors." Science 303 (5654):95-8. doi: 10.1126/science.1090599.

Ma, J. B., K. Ye, and D. J. Patel. 2004. "Structural basis for overhang-specific small interfering RNA recognition by the PAZ domain." Nature 429 (6989):318-22. doi: 10.1038/nature02519.

Maas, S., A. Rich, and K. Nishikura. 2003. "A-to-I RNA editing: recent news and residual mysteries." J Biol Chem 278 (3):1391-4. doi: 10.1074/jbc.R200025200.

Marques, J. T., K. Kim, P. H. Wu, T. M. Alleyne, N. Jafari, and R. W. Carthew. 2010. "Loqs and R2D2 act sequentially in the siRNA pathway in Drosophila." Nat Struct Mol Biol 17 (1):24-30. doi: 10.1038/nsmb.1735.

Marti, E., L. Pantano, M. Banez-Coronel, F. Llorens, E. Minones-Moyano, S. Porta, L. Sumoy, I. Ferrer, and X. Estivill. 2010. "A myriad of miRNA variants in control and Huntington's disease brain regions detected by massively parallel sequencing." Nucleic Acids Res 38 (20):7219-35. doi: 10.1093/nar/gkq575.

Maute, R. L., C. Schneider, P. Sumazin, A. Holmes, A. Califano, K. Basso, and R. Dalla-Favera. 2013. "tRNA-derived microRNA modulates proliferation and the DNA damage response and is down-regulated in B cell lymphoma." Proc Natl Acad Sci U S A 110 (4):1404-9. doi: 1206761110 [pii]

10.1073/pnas.1206761110.

McIlroy, G., I. Foldi, J. Aurikko, J. S. Wentzell, M. A. Lim, J. C. Fenton, N. J. Gay, and A. Hidalgo. 2013. "Toll-6 and Toll-7 function as neurotrophin receptors in the Drosophila melanogaster CNS." Nat Neurosci 16 (9):1248-56. doi: nn.3474 [pii]
10.1038/nn.3474.

Mehler, M. F., and J. S. Mattick. 2007. "Noncoding RNAs and RNA editing in brain development, functional diversification, and neurological disease." Physiol Rev 87 (3):799-823. doi: 10.1152/physrev.00036.2006.

Miura, P., S. Shenker, C. Andreu-Agullo, J. O. Westholm, and E. C. Lai. 2013. "Widespread and extensive lengthening of 3' UTRs in the mammalian brain." Genome Res 23 (5):812-25. doi: 10.1101/gr.146886.112.

Miyaki, S., T. Nakasa, S. Otsuki, S. P. Grogan, R. Higashiyama, A. Inoue, Y. Kato, T. Sato, M. K. Lotz, and H. Asahara. 2009. "MicroRNA-140 is expressed in differentiated human articular chondrocytes and modulates interleukin-1 responses." Arthritis Rheum 60 (9):2723-30. doi: 10.1002/art.24745.

Miyoshi, K., T. Miyoshi, and H. Siomi. 2010. "Many ways to generate microRNA-like small RNAs: non-canonical pathways for microRNA production." Mol Genet Genomics 284 (2):95-103. doi: 10.1007/s00438-010-0556-1.

Miyoshi, T., A. Takeuchi, H. Siomi, and M. C. Siomi. 2010. "A direct role for Hsp90 in pre-RISC formation in Drosophila." Nat Struct Mol Biol 17 (8):1024-6. doi: 10.1038/nsmb.1875.

Moreau, M. P., S. E. Bruse, R. David-Rus, S. Buyske, and L. M. Brzustowicz. 2011. "Altered microRNA expression profiles in postmortem brain samples from individuals with schizophrenia and bipolar disorder." Biol Psychiatry 69 (2):188-93. doi: S0006-3223(10)01008-5 [pii]
10.1016/j.biopsych.2010.09.039.

Morse, D. P., P. J. Aruscavage, and B. L. Bass. 2002. "RNA hairpins in noncoding regions of human brain and Caenorhabditis elegans mRNA are edited by adenosine deaminases that act on RNA." Proc Natl Acad Sci U S A 99 (12):7906-11. doi: 10.1073/pnas.112704299.

Naqvi, A., T. Cui, and A. Grigoriev. 2014. "Visualization of nucleotide substitutions in the (micro)transcriptome." BMC Genomics 15 Suppl 4:S9. doi: 1471-2164-15-S4-S9 [pii]
10.1186/1471-2164-15-S4-S9.

Neilsen, C. T., G. J. Goodall, and C. P. Bracken. 2012. "IsomiRs--the overlooked repertoire in the dynamic microRNAome." Trends Genet 28 (11):544-9. doi: 10.1016/j.tig.2012.07.005.

Newman, M. A., and S. M. Hammond. 2010. "Emerging paradigms of regulated microRNA processing." Genes Dev 24 (11):1086-92. doi: 10.1101/gad.1919710.

Nishida, K. M., K. Miyoshi, A. Ogino, T. Miyoshi, H. Siomi, and M. C. Siomi. 2013. "Roles of R2D2, a cytoplasmic D2 body component, in the endogenous siRNA pathway in Drosophila." Mol Cell 49 (4):680-91. doi: 10.1016/j.molcel.2012.12.024.

Nozawa, M., S. Miura, and M. Nei. 2010. "Origins and evolution of microRNA genes in Drosophila species." Genome Biol Evol 2:180-9. doi: 10.1093/gbe/evq009.

O'Brien, R. J., and P. C. Wong. 2011. "Amyloid precursor protein processing and Alzheimer's disease." Annu Rev Neurosci 34:185-204. doi: 10.1146/annurev-neuro-061010-113613.

O'Connell, R. M., D. S. Rao, A. A. Chaudhuri, and D. Baltimore. 2010. "Physiological and pathological roles for microRNAs in the immune system." Nat Rev Immunol 10 (2):111-22. doi: 10.1038/nri2708.

O'Neill, C., A. P. Kiely, M. F. Coakley, S. Manning, and C. M. Long-Smith. 2012. "Insulin and IGF-1 signalling: longevity, protein homoeostasis and Alzheimer's disease." Biochem Soc Trans 40 (4):721-7. doi: 10.1042/BST20120080.

Okamura, K., W. J. Chung, J. G. Ruby, H. Guo, D. P. Bartel, and E. C. Lai. 2008. "The Drosophila hairpin RNA pathway generates endogenous short interfering RNAs." Nature 453 (7196):803-6. doi: 10.1038/nature07015.

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 (1):89-100. doi: 10.1016/j.cell.2007.06.028.

Okamura, K., A. Ishizuka, H. Siomi, and M. C. Siomi. 2004. "Distinct roles for Argonaute proteins in small RNA-directed RNA cleavage pathways." Genes Dev 18 (14):1655-66. doi: 10.1101/gad.1210204.

Okamura, K., E. Ladewig, L. Zhou, and E. C. Lai. 2013. "Functional small RNAs are generated from select miRNA hairpin loops in flies and mammals." Genes Dev 27 (7):778-92. doi: 10.1101/gad.211698.112.

Okamura, K., N. Liu, and E. C. Lai. 2009. "Distinct mechanisms for microRNA strand selection by Drosophila Argonautes." Mol Cell 36 (3):431-44. doi: 10.1016/j.molcel.2009.09.027.

Okamura, K., N. Robine, Y. Liu, Q. Liu, and E. C. Lai. 2011. "R2D2 organizes small regulatory RNA pathways in Drosophila." Mol Cell Biol 31 (4):884-96. doi: 10.1128/MCB.01141-10.

Orr, H. T., and H. Y. Zoghbi. 2007. "Trinucleotide repeat disorders." Annu Rev Neurosci 30:575-621. doi: 10.1146/annurev.neuro.29.051605.113042.

Parker, J. S., S. M. Roe, and D. Barford. 2005. "Structural insights into mRNA recognition from a PIWI domain-siRNA guide complex." Nature 434 (7033):663-6. doi: 10.1038/nature03462.

Peng, H., J. Shi, Y. Zhang, H. Zhang, S. Liao, W. Li, L. Lei, C. Han, L. Ning, Y. Cao, Q. Zhou, Q. Chen, and E. Duan. 2012. "A novel class of tRNA-derived small RNAs extremely enriched in mature mouse sperm." Cell Res 22 (11):1609-12. doi: cr2012141 [pii]
10.1038/cr.2012.141.

Pepper, A. S., R. W. Beerman, B. Bhogal, and T. A. Jongens. 2009. "Argonaute2 suppresses Drosophila fragile X expression preventing neurogenesis and oogenesis defects." PLoS One 4 (10):e7618. doi: 10.1371/journal.pone.0007618.

Pereira, J. A., R. Baumann, C. Norrmen, C. Somandin, M. Miehe, C. Jacob, T. Luhmann, H. Hall-Bozic, N. Mantei, D. Meijer, and U. Suter. 2010. "Dicer in Schwann cells is required for myelination and axonal integrity." J Neurosci 30 (19):6763-75. doi: 10.1523/JNEUROSCI.0801-10.2010.

Perkins, D. O., C. D. Jeffries, L. F. Jarskog, J. M. Thomson, K. Woods, M. A. Newman, J. S. Parker, J. Jin, and S. M. Hammond. 2007. "microRNA expression in the prefrontal cortex of individuals with schizophrenia and schizoaffective disorder." Genome Biol 8 (2):R27. doi: gb-2007-8-2-r27 [pii]
10.1186/gb-2007-8-2-r27.

Pincus, Z., and F. J. Slack. 2008. "Transcriptional (dys)regulation and aging in Caenorhabditis elegans." Genome Biol 9 (9):233. doi: 10.1186/gb-2008-9-9-233.

Pincus, Z., and F. J. Slack. 2010. "Developmental biomarkers of aging in Caenorhabditis elegans." Dev Dyn 239 (5):1306-14. doi: 10.1002/dvdy.22224.

Polson, A. G., B. L. Bass, and J. L. Casey. 1996. "RNA editing of hepatitis delta virus antigenome by dsRNA-adenosine deaminase." Nature 380 (6573):454-6. doi: 10.1038/380454a0.

Prakash, S., H. M. McLendon, C. I. Dubreuil, A. Ghose, J. Hwa, K. A. Dennehy, K. M. Tomalty, K. L. Clark, D. Van Vactor, and T. R. Clandinin. 2009. "Complex interactions amongst N-cadherin, DLAR, and Liprin-alpha regulate Drosophila photoreceptor axon targeting." Dev Biol 336 (1):10-9. doi: S0012-1606(09)01194-4 [pii]
10.1016/j.ydbio.2009.09.016.

Ramachandran, V., and X. Chen. 2008. "Degradation of microRNAs by a family of exoribonucleases in Arabidopsis." Science 321 (5895):1490-2. doi: 10.1126/science.1163728.

Raver-Shapira, N., E. Marciano, E. Meiri, Y. Spector, N. Rosenfeld, N. Moskovits, Z. Bentwich, and M. Oren. 2007. "Transcriptional activation of miR-34a contributes to p53-mediated apoptosis." Mol Cell 26 (5):731-43. doi: 10.1016/j.molcel.2007.05.017.

Reinhart, B. J., F. J. Slack, M. Basson, A. E. Pasquinelli, J. C. Bettinger, A. E. Rougvie, H. R. Horvitz, and G. Ruvkun. 2000. "The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans." Nature 403 (6772):901-6. doi: 10.1038/35002607.

Ren, G., X. Chen, and B. Yu. 2012. "Uridylation of miRNAs by hen1 suppressor1 in Arabidopsis." Curr Biol 22 (8):695-700. doi: 10.1016/j.cub.2012.02.052.

Ren, G., M. Xie, S. Zhang, C. Vinovskis, X. Chen, and B. Yu. 2014. "Methylation protects microRNAs from an AGO1-associated activity that uridylates 5' RNA fragments generated by AGO1 cleavage." Proc Natl Acad Sci U S A 111 (17):6365-70. doi: 1405083111 [pii]
10.1073/pnas.1405083111.

Robinson, J. T., H. Thorvaldsdottir, W. Winckler, M. Guttman, E. S. Lander, G. Getz, and J. P. Mesirov. 2011. "Integrative genomics viewer." Nat Biotechnol 29 (1):24-6. doi: 10.1038/nbt.1754.

Rosenbloom, K. R., J. Armstrong, G. P. Barber, J. Casper, H. Clawson, M. Diekhans, T. R. Dreszer, P. A. Fujita, L. Guruvadoo, M. Haeussler, R. A. Harte, S. Heitner, G. Hickey, A. S. Hinrichs, R. Hubley, D. Karolchik, K. Learned, B. T. Lee, C. H. Li, K. H. Miga, N. Nguyen, B. Paten, B. J. Raney, A. F. Smit, M. L. Speir, A. S. Zweig, D. Haussler, R. M. Kuhn, and W. J. Kent. 2014. "The UCSC Genome Browser database: 2015 update." Nucleic Acids Res. doi: gku1177 [pii]
10.1093/nar/gku1177.

Ruby, J. G., C. H. Jan, and D. P. Bartel. 2007. "Intronic microRNA precursors that bypass Drosha processing." Nature 448 (7149):83-6. doi: 10.1038/nature05983.

Ruby, J. G., C. Jan, C. Player, M. J. Axtell, W. Lee, C. Nusbaum, H. Ge, and D. P. Bartel. 2006. "Large-scale sequencing reveals 21U-RNAs and additional microRNAs and endogenous siRNAs in C. elegans." Cell 127 (6):1193-207. doi: 10.1016/j.cell.2006.10.040.

Ruby, J. G., A. Stark, W. K. Johnston, M. Kellis, D. P. Bartel, and E. C. Lai. 2007. "Evolution, biogenesis, expression, and target predictions of a substantially expanded set of Drosophila microRNAs." Genome Res 17 (12):1850-64. doi: 10.1101/gr.6597907.

Saito, K., S. Inagaki, T. Mituyama, Y. Kawamura, Y. Ono, E. Sakota, H. Kotani, K. Asai, H. Siomi, and M. C. Siomi. 2009. "A regulatory circuit for piwi by the large Maf gene traffic jam in Drosophila." Nature 461 (7268):1296-9. doi: 10.1038/nature08501.

Schaefer, A., D. O'Carroll, C. L. Tan, D. Hillman, M. Sugimori, R. Llinas, and P. Greengard. 2007. "Cerebellar neurodegeneration in the absence of microRNAs." J Exp Med 204 (7):1553-8. doi: 10.1084/jem.20070823.

Schofield, C. M., R. Hsu, A. J. Barker, C. C. Gertz, R. Blelloch, and E. M. Ullian. 2011. "Monoallelic deletion of the microRNA biogenesis gene Dgcr8 produces deficits in the development of excitatory synaptic transmission in the prefrontal cortex." Neural Dev 6:11. doi: 10.1186/1749-8104-6-11.

Schwarz, D. S., G. Hutvagner, T. Du, Z. Xu, N. Aronin, and P. D. Zamore. 2003. "Asymmetry in the assembly of the RNAi enzyme complex." Cell 115 (2):199-208.

Seibel, P. N., T. Muller, T. Dandekar, J. Schultz, and M. Wolf. 2006. "4SALE--a tool for synchronous RNA sequence and secondary structure alignment and editing." BMC Bioinformatics 7:498. doi: 1471-2105-7-498 [pii]
10.1186/1471-2105-7-498.

Seitz, H., M. Ghildiyal, and P. D. Zamore. 2008. "Argonaute loading improves the 5' precision of both MicroRNAs and their miRNA* strands in flies." Curr Biol 18 (2):147-51. doi: 10.1016/j.cub.2007.12.049.

Seton-Rogers, S. 2012. "MicroRNAs: Editing changes the meaning." Nat Rev Cancer 12 (12):797. doi: 10.1038/nrc3413.

Shan, Z. X., Q. X. Lin, Y. H. Fu, C. Y. Deng, Z. L. Zhou, J. N. Zhu, X. Y. Liu, Y. Y. Zhang, Y. Li, S. G. Lin, and X. Y. Yu. 2009. "Upregulated expression of miR-1/miR-206 in a rat model of myocardial infarction." Biochem Biophys Res Commun 381 (4):597-601. doi: 10.1016/j.bbrc.2009.02.097.

Shin, Chanseok, Jin-Wu Nam, Kyle Kai-How Farh, H Rosaria Chiang, Alena Shkumatava, and David P Bartel. 2010. "Expanding the microRNA targeting code: functional sites with centered pairing." Molecular cell 38 (6):789-802.

Shin, D., J. Y. Shin, M. T. McManus, L. J. Ptacek, and Y. H. Fu. 2009. "Dicer ablation in oligodendrocytes provokes neuronal impairment in mice." Ann Neurol 66 (6):843-57. doi: 10.1002/ana.21927.

Smith-Vikos, T., A. de Lencastre, S. Inukai, M. Shlomchik, B. Holtrup, and F. J. Slack. 2014. "MicroRNAs mediate dietary-restriction-induced longevity through PHA-4/FOXA and SKN-1/Nrf transcription factors." Curr Biol 24 (19):2238-46. doi: 10.1016/j.cub.2014.08.013.

Smith-Vikos, T., and F. J. Slack. 2012. "MicroRNAs and their roles in aging." J Cell Sci 125 (Pt 1):7-17. doi: 10.1242/jcs.099200.

Sobala, A., and G. Hutvagner. 2007. "Transfer RNA-derived fragments: origins, processing, and functions." Wiley Interdiscip Rev RNA 2 (6):853-62. doi: 10.1002/wrna.96.

Stark, K. L., B. Xu, A. Bagchi, W. S. Lai, H. Liu, R. Hsu, X. Wan, P. Pavlidis, A. A. Mills, M. Karayiorgou, and J. A. Gogos. 2008. "Altered brain microRNA biogenesis contributes to phenotypic deficits in a 22q11-deletion mouse model." Nat Genet 40 (6):751-60. doi: 10.1038/ng.138.

Taliaferro, J. M., J. L. Aspden, T. Bradley, D. Marwha, M. Blanchette, and D. C. Rio. 2013. "Two new and distinct roles for Drosophila Argonaute-2 in the nucleus: alternative pre-mRNA splicing and transcriptional repression." Genes Dev 27 (4):378-89. doi: 10.1101/gad.210708.112.

Tan, H., M. Poidevin, H. Li, D. Chen, and P. Jin. 2012. "MicroRNA-277 modulates the neurodegeneration caused by Fragile X premutation rCGG repeats." PLoS Genet 8 (5):e1002681. doi: 10.1371/journal.pgen.1002681
PGENETICS-D-11-02474 [pii].

Tao, J., H. Wu, Q. Lin, W. Wei, X. H. Lu, J. P. Cantle, Y. Ao, R. W. Olsen, X. W. Yang, I. Mody, M. V. Sofroniew, and Y. E. Sun. 2011. "Deletion of astroglial Dicer causes non-cell-autonomous neuronal dysfunction and degeneration." J Neurosci 31 (22):8306-19. doi: 10.1523/JNEUROSCI.0567-11.2011.

Tardif, G., D. Hum, J. P. Pelletier, N. Duval, and J. Martel-Pelletier. 2009. "Regulation of the IGFBP-5 and MMP-13 genes by the microRNAs miR-140 and miR-27a in human osteoarthritic chondrocytes." BMC Musculoskelet Disord 10:148. doi: 10.1186/1471-2474-10-148.

Teng, B., C. F. Burant, and N. O. Davidson. 1993. "Molecular cloning of an apolipoprotein B messenger RNA editing protein." Science 260 (5115):1816-9.

Thompson, D. M., Cheng Lu, Pamela J. Green, and Roy Parker. 2008. "tRNA cleavage is a conserved response to oxidative stress in eukaryotes." RNA 14 (10):2095-2103. doi: 10.1261/rna.1232808.

Thompson, D. M., and R. Parker. 2009a. "The RNase Rny1p cleaves tRNAs and promotes cell death during oxidative stress in Saccharomyces cerevisiae." J Cell Biol 185 (1):43-50. doi: jcb.200811119 [pii]
10.1083/jcb.200811119.

Thompson, D. M., and R. Parker. 2009b. "Stressing out over tRNA cleavage." Cell 138 (2):215-9. doi: 10.1016/j.cell.2009.07.001.

Tomari, Y., C. Matranga, B. Haley, N. Martinez, and P. D. Zamore. 2004. "A protein sensor for siRNA asymmetry." Science 306 (5700):1377-80. doi: 10.1126/science.1102755.

Tu, Q., R. A. Cameron, and E. H. Davidson. 2014. "Quantitative developmental transcriptomes of the sea urchin Strongylocentrotus purpuratus." Dev Biol 385 (2):160-7. doi: 10.1016/j.ydbio.2013.11.019.

Tu, Q., R. A. Cameron, K. C. Worley, R. A. Gibbs, and E. H. Davidson. 2012. "Gene structure in the sea urchin Strongylocentrotus purpuratus based on transcriptome analysis." Genome Res 22 (10):2079-87. doi: 10.1101/gr.139170.112.

Tuck, A. C., and D. Tollervey. 2011. "RNA in pieces." Trends Genet 27 (10):422-32. doi: S0168-9525(11)00089-8 [pii]
10.1016/j.tig.2011.06.001.

Vagin, V. V., A. Sigova, C. Li, H. Seitz, V. Gvozdev, and P. D. Zamore. 2006. "A distinct small RNA pathway silences selfish genetic elements in the germline." Science 313 (5785):320-4. doi: 10.1126/science.1129333.

Voinnet, O. 2009. "Origin, biogenesis, and activity of plant microRNAs." Cell 136 (4):669-87. doi: 10.1016/j.cell.2009.01.046.

Wan, X. F., G. Lin, and D. Xu. 2006. "Rnall: an efficient algorithm for predicting RNA local secondary structural landscape in genomes." J Bioinform Comput Biol 4 (5):1015-31. doi: S0219720006002363 [pii].

Wang, Q., I. Lee, J. Ren, S. S. Ajay, Y. S. Lee, and X. Bao. 2012. "Identification and functional characterization of tRNA-derived RNA fragments (tRFs) in respiratory syncytial virus infection." Mol Ther 21 (2):368-79. doi: mt2012237 [pii]
10.1038/mt.2012.237.

Wang, Q., I. Lee, J. Ren, S. S. Ajay, Y. S. Lee, and X. Bao. 2013. "Identification and functional characterization of tRNA-derived RNA fragments (tRFs) in respiratory syncytial virus infection." Mol Ther 21 (2):368-79. doi: mt2012237 [pii]
10.1038/mt.2012.237.

Wang, W. X., B. W. Rajeev, A. J. Stromberg, N. Ren, G. Tang, Q. Huang, I. Rigoutsos, and P. T. Nelson. 2008. "The expression of microRNA miR-107 decreases early in Alzheimer's disease and may accelerate disease progression through regulation of beta-site amyloid precursor protein-cleaving enzyme 1." J Neurosci 28 (5):1213-23. doi: 10.1523/JNEUROSCI.5065-07.2008.

Wang, X. H., R. Z. Qian, W. Zhang, S. F. Chen, H. M. Jin, and R. M. Hu. 2009. "MicroRNA-320 expression in myocardial microvascular endothelial cells and its relationship with insulin-like growth factor-1 in type 2 diabetic rats." Clin Exp Pharmacol Physiol 36 (2):181-8. doi: 10.1111/j.1440-1681.2008.05057.x.

Wang, X., P. Liu, H. Zhu, Y. Xu, C. Ma, X. Dai, L. Huang, Y. Liu, L. Zhang, and C. Qin. 2009. "miR-34a, a microRNA up-regulated in a double transgenic mouse model of Alzheimer's disease, inhibits bcl2 translation." Brain Res Bull 80 (4-5):268-73. doi: 10.1016/j.brainresbull.2009.08.006.

Wei, C., L. Salichos, C. M. Wittgrove, A. Rokas, and J. G. Patton. 2012. "Transcriptome-wide analysis of small RNA expression in early zebrafish development." RNA 18 (5):915-29. doi: rna.029090.111 [pii]
10.1261/rna.029090.111.

Westholm, J. O., E. Ladewig, K. Okamura, N. Robine, and E. C. Lai. 2012. "Common and distinct patterns of terminal modifications to mirtrons and canonical microRNAs." RNA 18 (2):177-92. doi: 10.1261/rna.030627.111.

Westholm, J. O., and E. C. Lai. 2011. "Mirtrons: microRNA biogenesis via splicing." Biochimie 93 (11):1897-904. doi: 10.1016/j.biochi.2011.06.017.

Wienholds, E., and R. H. Plasterk. 2005. "MicroRNA function in animal development." FEBS Lett 579 (26):5911-22. doi: 10.1016/j.febslet.2005.07.070.

Wiese, K. C., E. Glen, and A. Vasudevan. 2005. "JViz.Rna--a Java tool for RNA secondary structure visualization." IEEE Trans Nanobioscience 4 (3):212-8.

Wyman, S. K., E. C. Knouf, R. K. Parkin, B. R. Fritz, D. W. Lin, L. M. Dennis, M. A. Krouse, P. J. Webster, and M. Tewari. 2011. "Post-transcriptional generation of miRNA variants by multiple nucleotidyl transferases contributes to miRNA transcriptome complexity." Genome Res 21 (9):1450-61. doi: 10.1101/gr.118059.110.

Yang, J. S., and E. C. Lai. 2010. "Dicer-independent, Ago2-mediated microRNA biogenesis in vertebrates." Cell Cycle 9 (22):4455-60.

Yang, J. S., T. Maurin, N. Robine, K. D. Rasmussen, K. L. Jeffrey, R. Chandwani, E. P. Papapetrou, M. Sadelain, D. O'Carroll, and E. C. Lai. 2010. "Conserved vertebrate mir-451 provides a platform for Dicer-independent, Ago2-mediated microRNA biogenesis." Proc Natl Acad Sci U S A 107 (34):15163-8. doi: 10.1073/pnas.1006432107.

Yeung, M. L., Y. Bennasser, K. Watashi, S. Y. Le, L. Houzet, and K. T. Jeang. 2009. "Pyrosequencing of small non-coding RNAs in HIV-1 infected cells: evidence for the processing of a viral-cellular double-stranded RNA hybrid." Nucleic Acids Res 37 (19):6575-86. doi: 10.1093/nar/gkp707.

Yu, B., Z. Yang, J. Li, S. Minakhina, M. Yang, R. W. Padgett, R. Steward, and X. Chen. 2005. "Methylation as a crucial step in plant microRNA biogenesis." Science 307 (5711):932-5. doi: 10.1126/science.1107130.

Yu, X. Y., Y. H. Song, Y. J. Geng, Q. X. Lin, Z. X. Shan, S. G. Lin, and Y. Li. 2008. "Glucose induces apoptosis of cardiomyocytes via microRNA-1 and IGF-1." Biochem Biophys Res Commun 376 (3):548-52. doi: 10.1016/j.bbrc.2008.09.025.

Zhang, Y., H. Guo, H. Kwan, J. W. Wang, J. Kosek, and B. Lu. 2007. "PAR-1 kinase phosphorylates Dlg and regulates its postsynaptic targeting at the Drosophila

neuromuscular junction." Neuron 53 (2):201-15. doi: S0896-6273(06)01021-X [pii]
10.1016/j.neuron.2006.12.016.

Zhao, Y., B. Mo, and X. Chen. 2012. "Mechanisms that impact microRNA stability in plants." RNA Biol 9 (10):1218-23. doi: 10.4161/rna.22034.

Zhao, Y., Y. Yu, J. Zhai, V. Ramachandran, T. T. Dinh, B. C. Meyers, B. Mo, and X. Chen. 2012. "The Arabidopsis nucleotidyl transferase HESO1 uridylates unmethylated small RNAs to trigger their degradation." Curr Biol 22 (8):689-94. doi: 10.1016/j.cub.2012.02.051.

Zhou, R., B. Czech, J. Brennecke, R. Sachidanandam, J. A. Wohlschlegel, N. Perrimon, and G. J. Hannon. 2009. "Processing of Drosophila endo-siRNAs depends on a specific Loquacious isoform." RNA 15 (10):1886-95. doi: 10.1261/rna.1611309.

Zhu, H. C., L. M. Wang, M. Wang, B. Song, S. Tan, J. F. Teng, and D. X. Duan. 2012. "MicroRNA-195 downregulates Alzheimer's disease amyloid-beta production by targeting BACE1." Brain Res Bull 88 (6):596-601. doi: 10.1016/j.brainresbull.2012.05.018.