THE REGULATION OF LET-7 MICRORNA BIOGENESIS IN EARLY EMBRYOGENESIS

Martin Aaron Newman

A dissertation submitted to the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Cell and Developmental Biology

Chapel Hill

2010

Approved by:

Scott Hammond, Ph.D.

Vytas Bankaitis, Ph.D.

William Marzluff, Ph.D.

Pat Brennwald, Ph.D.

Da-Zhi Wang, Ph.D.

Brian Strahl, Ph.D.

ABSTRACT

MARTIN AARON NEWMAN: THE REGULATION OF LET-7 MICRORNA BIOGENESIS IN EMBRYONIC DEVELOPMENT

(Under the direction of Scott Hammond)

microRNAs (miRNAs) are small non-protein-coding RNAs that silence gene expression post-transcriptionally and have critical functions in development, tissue homeostasis and in the pathogenesis of many diseases. They carry out gene silencing by targeting complementary sequences in messenger RNAs (mRNAs) in the context of a ribonucleoprotein particle named RISC (for RNA induced silencing complex). The mature miRNA contained within RISC is a ~21 nucleotide single stranded RNA that must be processed in two RNase III-catalyzed reactions. In animals, our precise understanding of miRNA funtion *in vivo* has been hindered by the fact that animal miRNAs do not basepair with perfect complementarity to their miRNA targets. However, much insight into miRNA function has been gained by the study of their spatio-temporal expression. Of particular interest is that the majority of known mouse miRNAs are not detectable during early embryogenesis but are robustly induced in later in gestation, during the time of tissue differentiation. Intriguingly, there is a broad down regulation of miRNAs in poorly differentiated tumors; thus it is a critical long-term goal to understand whether the global absence of miRNAs in cancer reflects the reversion to an embryonic-like cellular state. Studies described herein show that the abundance of many miRNAs, including the

let-7 family of tumor suppressor miRNAs, are regulated post-transcriptionally both in murine development and in cancer. Furthermore, RNase III-mediated processing of let-7 is blocked by the embryonic stem (ES) cell-specific RNA binding protein Lin28. While the exact mechanism of the Lin28-mediated processing block is not known, a popular model predicts that Lin28 binds to the short hairpin precursor let-7 (pre-let-7) and recruits the terminal uridyl transferase, TUT4, leading to oligo-uridylation and turnover of pre-let-7. Directly testing this hypothesis is technically challenging, in part, because short-hairpin miRNA precursors (pre-miRNAs) are low in abundance. We developed a novel and powerful technique to detect pre-miRNAs utilizing high-throughput DNA sequencing. Unexpectedly, we observed that pre-let-7 is uridylated in differentiated cells, which lack Lin28; furthermore, many pre-miRNAs are uridylated in vivo, likely through a Lin-28-independent pathway.

ACKNOWLEDGEMENTS

I am indebted to the many people who helped to shape who I am as a scientist and more importantly, as a human being.

First, I am grateful to have had Scott Hammond as my doctoral advisor; he is a brillinat and creative mind, a great motivator and he is always very patient at times in which many people wouldn't be. I also feel very fortunate to have crossed paths with Mike Thomson; his antics, enthusiam for science and endless knowledge of molecular biology techniques were a great inspiration to me.

I would also like to thank my labmates, past and present: Keith Woods, Summer Goodson, Vidya Mani, Handan Kaygun, Andy Baraniak, Claudia Dollins, Carrie Shaw and Robert Sons. They made the lab a fun and entertaining place to be.

I am most thankful for my Mom and Dad, Bob and Hattie; their love, support and wisdom have been instrumental in everything leading up to this point.

TABLE OF CONTENTS

LIST OF TABLES

Table

LIST OF FIGURES

Figure

CHAPTER 1:
INTRODUCTION

In 1993, two research groups discovered that the gene *lin-4*, in the roundworm *Caenorhabditis elegans* (*C. elegans*), encoded a small RNA that could bind to complementary sites in the *lin-14* messenger RNA (mRNA), thereby regulating Lin-14 protein abundance during embryonic development (R. C. Lee, Feinbaum, & Ambros, 1993; Wightman, Ha, & Ruvkun, 1993). The idea that a small non-protein-coding RNA (ncRNA) could post-transcriptionally regulate gene expression was unprecedented and so, for several years, the *lin-4* ncRNA seemed like an odd outlier in molecular biology. RNA-mediated gene silencing gained wide attention in 1998 when Andrew Fire, Craig Mello and colleagues described the phenomenon of RNA interference (or RNAi): the injection of double-stranded RNA (dsRNA) into *C. elegans* could silence the expression of a gene whose nucleotide sequence matched that of the dsRNA (Fire et al., 1998). The discoveries of RNAi and the *lin-4* RNA, at first, seemed unrelated. However, in 1999, another landmark study showed that small RNAs were the molecular determinant of post-transcriptional gene silencing (PTGS) in plants. It soon became clear that RNAi did not just occur with RNA from an external source, but that the cellular machinery responsible for RNAi employed a broad spectrum of genomically-encoded, small ncRNAs; the first class of these small RNAs to be characterized was microRNA (miRNA); lin-4 is considered to be the "founding" miRNA.

To date, a diverse collection of small ncRNA pathways have been discovered beyond the miRNA pathway; in all known cases, small RNAs guide sequence-specific changes in gene expression at either the transcriptional or post-transcriptional level. In plants, small RNA pathways act, in part, as a primitive immune system, processing viral RNAs into short-

interfering RNAs (siRNAs); plants also contain their own complement of genes that encode miRNAs. Small RNA pathways in animals prevent the mobilization of transposon elements (piRNAs), modulate chromatin modifications (siRNAs) and carry out post-transcriptional gene silencing (endo-siRNAs and miRNAs, **Figure 1.1**; for a thorough review see, (Ghildiyal & Zamore, 2009)).

While endogenous small RNA pathways are instrumental in diverse biological processes, they share two common characteristics: 1) each small RNA pathway utilizes a class of small single-stranded RNA (ssRNA) that must be processed from larger dsRNA precursor molecules (an exception is piRNAs, whose precursor's are thought to be ssRNA); 2) each class of small ssRNA associates with a member of the PIWI/Argonaute (Ago) pathway of proteins. PIWI/Ago proteins contain 3 unique functional domains (**Figure 1.1A**). The PAZ domain is located near the N-terminus and accommodates the 3'end of the small RNA. The MID domain binds to the small RNA's phosphorylated 5' end. The PIWI domain possesses an RNase H-like activity that mediates cleavage of target RNAs to which the small RNA is bound (Farazi, Juranek, & Tuschl, 2008).

For reasons that are still not entirely clear, the endogenous RNAi pathways in certain organisms are more numerous than in others. For instance, the C. elegans genome encodes an exceptionally large number of PIWI/Ago genes (Simard and Hutvagner NRMCB, 2008); they even possess a clade of secondary-siRNA-interacting Agos (or "Sagos") that exist only in worms. These Sagos facilitate the amplification of the initial RNAi response. Small RNA effector complexes are named differently depending on the class of small RNA and the PIWI/Ago protein that are complexed together ((Farazi et al., 2008; Ghildiyal & Zamore, 2009) **Figure 1.1B**).

The basic mechanism of miRNA-mediated gene silencing:

miRNAs are 18-25 nucleotide (nt) ssRNAs that associate with Ago proteins to form an RNA induced silencing complex (or RISC). The miRNA mediates RISC-binding to

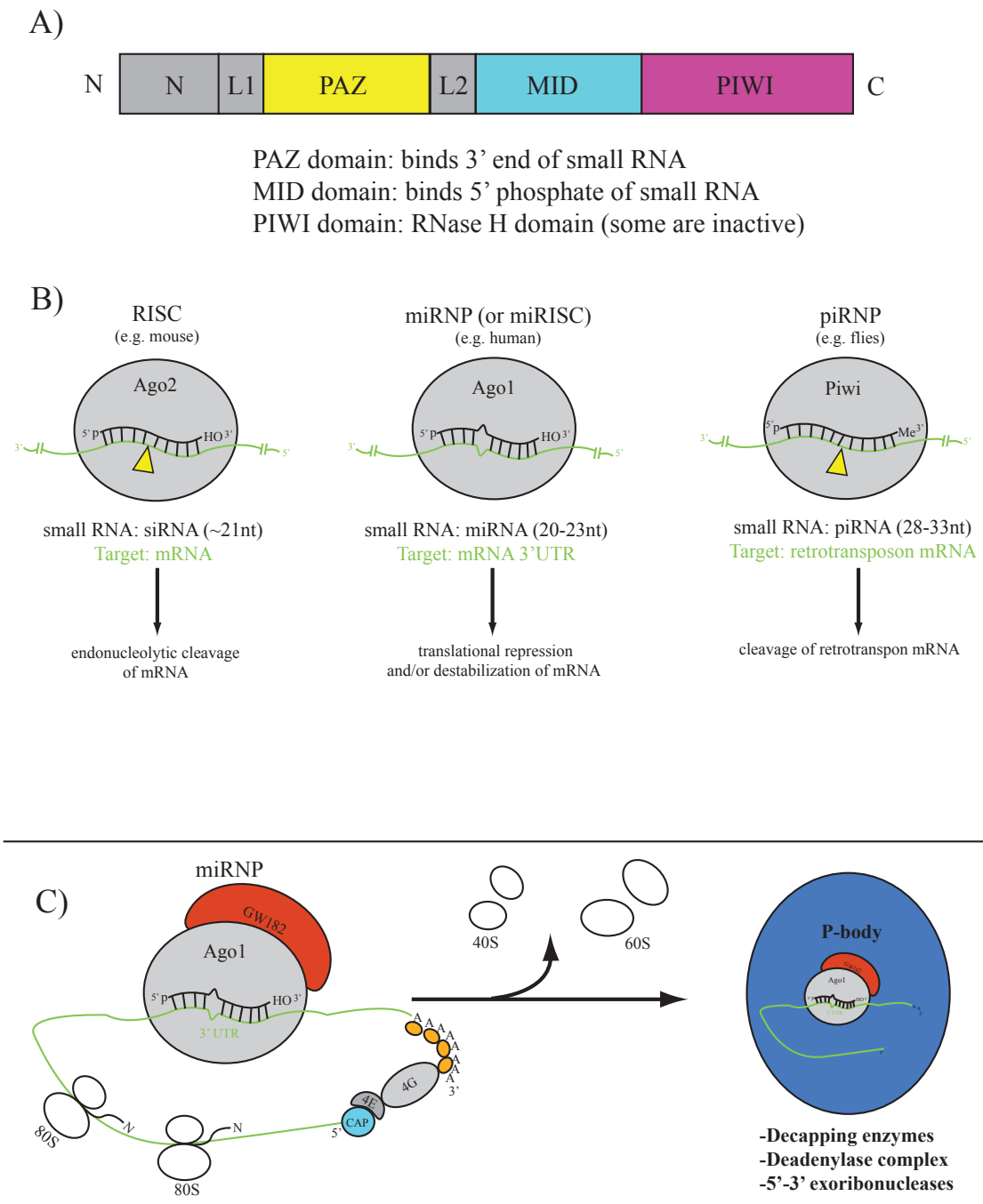


Figure 1.1: Small RNA effector complexes and the mechanism of miRNA-mediated silencing. A) Schematic of the domain organization of a representative PIWI/Ago family protein, Piwi from the prokaryote Aquifex aeolicus (cartoon adapted from Farazi et al., Development 2008). B) Simplified diagrams of diverse small RNA effector complexes. C) Simplified diagram of the mechanism of miRNA mediated gene silencing. The initial repression of mRNA translation occurs in the cytosol, after which the miRISC bound to targeted mRNA localizes to large cytosolic aggregates called P-bodies. Ribosomes are depicted as white ovals. Poly-A binding protein 1 (PABP1) is depicted as the small yellow ovals.

cognate sequences in mRNAs to beget post-transcriptional gene silencing. In mammals, there are 4 Ago proteins, Ago1-4; Ago2 is unique because it is the only Ago protein with the amino acid residues within its PIWI domain necessary for RNase-H-like catalysis (Farazi et al., 2008). In the case of exogenously-introduced siRNAs (e.g. transfected into cultured mammalian cells by a researcher), the extensive base-pairing of an siRNA RISC (bound to Ago2) with its target mRNA allows cleavage of the mRNA at a position corresponding to the region between the 10th and 11th nucleotides of the siRNA (Martinez et al., 2002). In contrast, miRNA-containing RISCs (or miRISCs) bind with less extensive complementarity to the target mRNA, usually within its 3' untranslated region (3' UTR). This results in translational repression and/or mRNA destabilization through a complex series of events.

While the exact mechanism of miRNA-mediated silencing is, to date, intensely debated, it is sufficient here to describe a model where the binding of miRISC to a target mRNA ultimately leads to its translational repression, de-adenylation of the mRNA and, often times, decreased mRNA abundance. It is believed that translational repression occurs in the cytosol whereupon miRISC-targeted mRNAs are sequestered to mRNA processing bodies in the cytosol called "P-bodies." P-bodies are thought to exclude the general translation machinery from miRNA-targeted mRNAs as well as to mediate mRNA degradation. mRNA decay in P-bodies requires Ago proteins, the Ago-associated protein GW182, the decapping enzyme DCP2, additional decapping activating factors and the CAF1-CCR4-NOT deadenylase complex (**Figure 1.1C**, (Eulalio et al., 2008)).

microRNA biogenesis:

The small single-stranded miRNAs that associate with Ago proteins are the biologically active triggers of the miRNA pathway. However, these "mature" miRNAs are initially generated as longer dsRNA precursors that must be processed in a series of RNase-mediated events. The miRNA biogenesis pathway is particularly well-characterized; mammalian miRNA biogenesis is depicted in **Figure 1.2** and is explained in detail below (for

detailed review see (V. N. Kim, Han, & Siomi, 2009). In the simplest sense, miRNAs are processed from larger precursor RNAs by two RNaseIII-catalyzed cleavage events followed by unwinding of the resulting duplex and incorporation of the single-stranded mature miRNA into RISC.

miRNA genes are transcribed by RNA polymerase (pol) II or III giving rise to a primary microRNA transcript (pri-microRNA). Pol II-generated pri-microRNAs (which are better-characterized) exist within introns and exons of non-coding transcripts as well as within introns of protein-coding transcripts; accordingly they have been shown experimentally to be capped and poly-adenylated (Cai et al., 2004). Recent bioinformatic analyses have revealed that a majority of microRNAs are intronic; furthermore, a significant portion (estimated up to 42%) of microRNAs exist in polycistrons containing two or more microRNA precusors (Altuvia et al., 2005).

Pri-microRNA processing by the Drosha/Dgcr8 complex:

The first major step forward in understanding the biochemical basis of microRNA processing was a study by V. Narry Kim and colleagues. They demonstrated that pri-miRNA transcripts generated *in vitro* could be processed to a distinct ~70nt product in nuclear extracts, approximately the size of a short hairpin precursor miRNA (pre-miRNA); this was, the first demonstration of an *in vitro* pri-miRNA processing assay. They next verified that this RNA product was a bona fide pre-miRNA by subsequently adding cytoplasmic extract, resulting in appearance of mature miRNA (Y. Lee et al., 2002); as the RNase III enzyme Dicer was already known to process pre-miRNAs in to mature miRNAs (Hutvagner et al., 2001; Knight & Bass, 2001), this study illustrated that miRNA maturation requires two unique RNA processing events. To identify the enzyme responsible for pri-miRNA processing, it was first determined that the 5' and 3' ends of the pre-miRNA generated in vitro were characteristic of RNase III cleavage: the pre-miRNA contained a 2nt, 3' overhang (Y. Lee et al., 2003). As pri-miRNA processing occurs in the nucleus and only one human

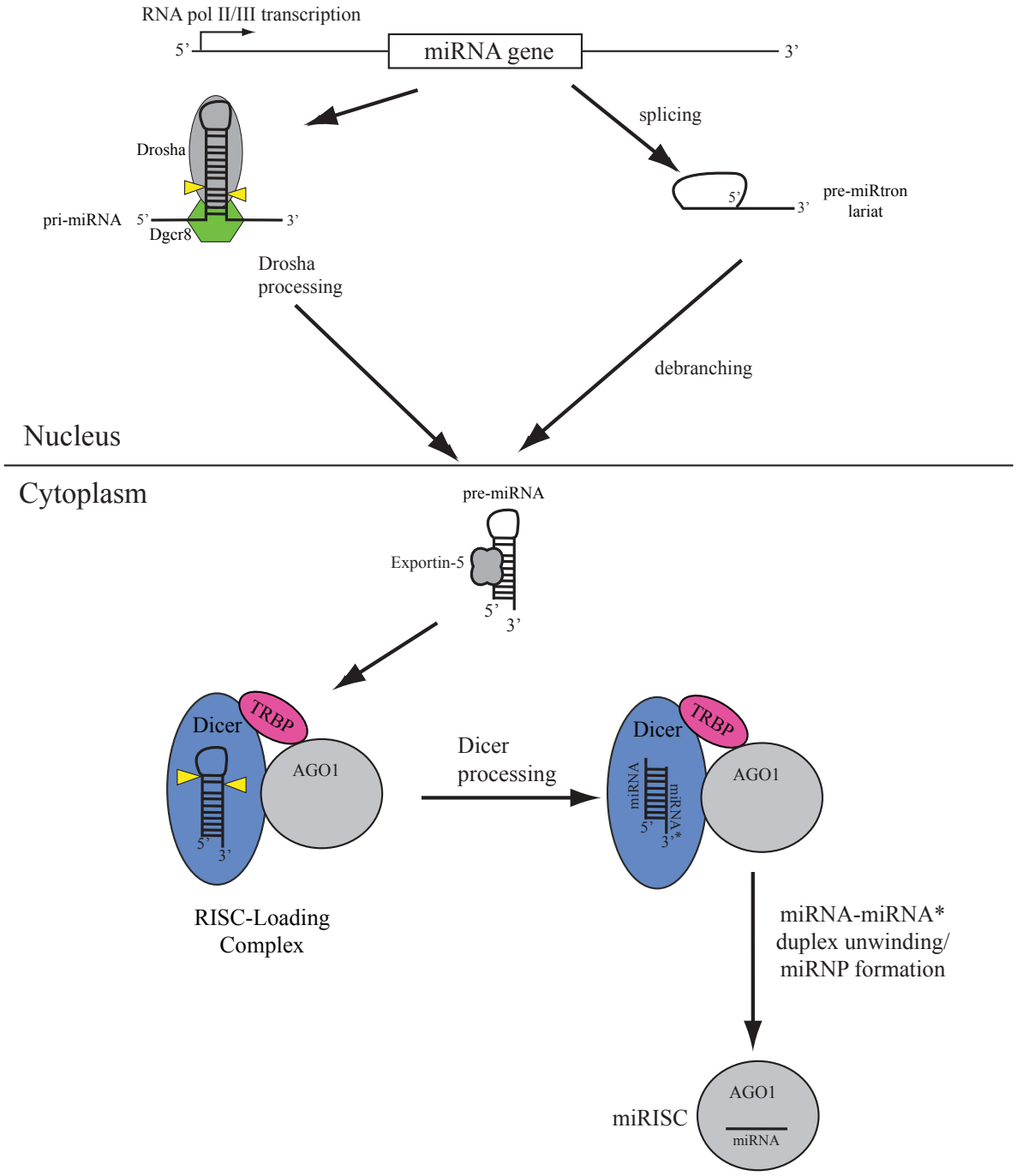


Figure 1.2: The miRNA biogenesis pathway. miRNA genes can be intergenic or situated with protein-coding genes and, in both cases, can be intronic or exonic. Transcription by either RNA pol II or pol III (depending on the miRNA gene) yields a pri-miRNA that is processed by the Microprocessor complex (containing Drosha and Dgcr8) into a pre-miRNA; in the case of mirtrons, pre-miRNAs generation circumnavigates the Drosha step and is created by splicing and debranching of a small intron. The pre-miRNA is then exported to the cytosol where it is further processed by Dicer into a mature miRNA duplex, from which (usually) one strand will be loaded in to an Ago protein to beget a miRISC.

RNase III protein, Drosha, occupies the nuclear compartment, it became clear that Drosha was enzyme responsible for pri-miRNA cleavage (Y. Lee et al., 2003).

Pri-miRNAs form a characteristic stemloop structure that is critical for processing by Drosha; the stemloop consists of an approximately 30 basepair (bp) stemloop region, an unstructured terminal loop ranging in size from approximately 10-30nts and flanked on either side by unstructured ssRNA. The Drosha enzyme alone is not sufficient for pri-microRNA processing *in vitro* but requires the dsRNA-binding domain (dsRBD) protein DeGeorge critical region 8 (Dgcr8, called Pasha in invertebrates;(Denli et al., 2004; Gregory et al., 2004). An elegant *in vitro* study demonstrated that Dgcr8 recognizes the dsRNA-ssRNA junction of pri-microRNAs; in UV-crosslinking studies, Dgcr8 had detectable pri-miRNA binding activity whereas Drosha did not. This led to the hypothesis that Dgcr8 mediates recognition of pri-microRNAs by Drosha, which then cleaves ~11bp above the ssRNA-dsRNA junction, liberating the pre-miRNA (Han et al., 2006). Several studies recently reported that pri-microRNA processing may occur co-transcriptionally (Y. K. Kim & Kim, 2007; Morlando et al., 2008; Pawlicki & Steitz, 2008). Drosha cleavage does not reduce splicing efficiency of an intron *in vivo* (Y. K. Kim & Kim, 2007). Furthermore, pri-miRNAs are tightly associated with chromatin (Morlando et al., 2008; Pawlicki & Steitz, 2008); also, in fluorescence microscopy experiments pri-miRNAs and Drosha localize to site of Pol II transcription (Pawlicki & Steitz, 2008). Taken together, this data suggests that transcription and Drosha cleavage are tightly coupled processes.

Fractionation of Drosha and Dgcr8 by size-exclusion chromatography suggested the proteins are associated in small (~250 kilo-Dalton [kDa]) and large (>2 mega-Dalton [MDa]) molecular weight complexes. Both of these complexes are referred to as the "Microprocessor" but they differ in protein composition. The small complex is composed solely of Drosha and Dgcr8 whereas the large complex contains many additional, abundant RNA-binding proteins. These additional "auxillary factors" have diverse roles in cellular RNA metabolism (e.g. splicing, mRNA stability, etc…Gregory et al., 2004) although, to date,

most do not have an assigned function in the context of the Microprocessor.

Dicer processing and RISC loading:

The pre-miRNA molecule which results from Drosha cleavage is a short hairpin RNA with a 5' phosphate group, a 3' hydroxyl group and a 3' 2nt overhang. The terminus of the pre-miRNA is recognized by exportin-5 and the pre-miRNA is exported to the cytosol in a Ran-GTP-dependent manner (Bohnsack et al., 2004; Lund et al., 2004; Yi et al., 2003). Once in the cytosol, the pre-miRNA encounters the cytosolic RNaseIII Dicer, which clips off the precursor loop yielding a microRNA duplex. Several biochemical studies showed that Dicer exists in a complex with the dsRBD protein TRBP and an Ago protein; this allows "Dicing" of the precursor, unwinding of the microRNA duplex and loading of the mature microRNA strand into RISC to occur in a concerted series of events ((Chendrimada et al., 2005; Gregory et al., 2005; Maniataki & Mourelatos, 2005); **Figure 1.2**).

Aside from the conventional biogenesis pathway, certain pre-miRNAs arise from splicing of a small intron rather than from pri-miRNA cleavage by Drosha (Ghildiyal & Zamore, 2009), **Figure 1.2**). These "miRtrons" have been detected in invertabrates as well as vertebrates, perhaps suggesting that the differences between the miRNA and miRtron processing pathways may confer important functional differences between their resulting mature miRNA products.

The biological function of miRNAs:

The discovery of miRNAs was built upon pioneering studies in C. elegans, where two "founding" miRNAs, *lin-4* and *let-7* were characterized as critical non-protein-coding genes in early and late embryonic development, respectively(R. C. Lee et al., 1993; Reinhart et al., 2000; Wightman et al., 1993). Worm embryos mutant for the *lin-4* gene displayed precociously differentiated larval tissue whereas let-7 mutants exhibited a failure of certain cell lineages to differentiate, resulting in a vulval bursting phenotype. miRNAs are now

recognized as important components of gene expression circuitry and have crucial roles in normal developmental as well as in the pathogenesis of many diseases including cancer. Thus, the overall function of miRNAs *in vivo* is currently a subject of intense study.

An effective approach to understanding miRNA biological function has been to identify natural mutations or to design targeted knockouts of miRNA genomic loci. For example, a key study revealed that the copy-number of a genomic locus harboring the miR-17-92 polycistronic cluster is amplified in many cancers (L. He et al., 2005), making miR-17-92 the first known miRNA oncogene (also the original group of miRNAs referred to an "oncomiR"). Conversely, a locus that contains the miR-15a and miR-16-1 genes is frequently deleted in B-cell chronic lymphocytic leukemia (Calin et al., 2002), revealing the function of these miRNAs as tumor suppressors. In an example of targeted deletion, transgenic mice lacking miR-1-2 showed severe defects in the development of cardiac precursor cells despite the fact that one miR-1 locus was unmodified (Zhao et al., 2007). This study not only attributed a role for miR-1 in cardiogenesis but also revealed that the proper dosage of a miRNA can be critical for proper development. Depsite being an effective tool for understanding miRNA function *in vivo*, genetic manipulation of miRNAs may not feasible for all miRNAs; for instance, total disruption of let-7 miRNA function would require targeted knockout of 12 genomic loci. In a related approach, total miRNA expression in embryonic stem (ES) cells was ablated by the targeted deletion of the Dgcr8 gene (Wang et al., 2007); these ES cells were viable but were unable to differentiate implicating a role for microRNAs in shutting down stem cell self-renewal programs.

An ideal approach in understanding miRNA function would be to identify the comprehensive set mRNAs targets for a given miRNA. This has been a monumental challenge in the field because, in animals, miRNAs bind with imperfect complimentarity to mRNAs making bioinformatic analysis of mRNA targets very difficult (reviewed in (Bartel, 2009)). One consistent finding has been the clear requirement for a miRNA "seed match," in which a stretch of nucleotides in the mRNA perfectly corresponds to nucleotides 2-8 of

the miRNA (this seed match can vary between 7-8 nucleotides of complementarity; for an extensive review (Bartel, 2009)). To date, rules for the contribution of the remainder of the miRNA in basepairing to its target mRNA have not been experimentally established (Bartel, 2009). Clearly it will be important to employ novel methods in the identification of miRNA targets as bioinformatic approaches suggest that more than half of the human protein-coding genome may be targeted by miRNAs (Bartel, 2009). Indeed, one potentially powerful alternative technique in the miRNA target identification is the use of dual-label mass spectrometry to directly analyze cellular protein compositioin after depletion of a selected miRNA (Baek et al., 2008; Selbach et al., 2008).

Understanding miRNA function is even further complicated by the recent idea (and observation) that the direct outcome of a miRNA binding to certain targets might simply be the sequestration (or "sponging up") of miRNAs without affecting the expression levels of the bound target (Seitz, 2009). This hypothesis is supported by a recent study where pseudogene transcripts of PTEN and KRAS compete with their protein-coding mRNA counterparts for miRNA binding, thereby controlling PTEN and K-Ras protein levels (Poliseno et al., 2010).

Intensive efforts in analyzing miRNA experession patterns have significantly contributed to our understanding of the biological function of miRNAs. Initial studies to detect mature miRNAs utilized northern blot analysis in various tissues and at different developmental stages (R. C. Lee et al., 1993; Reinhart et al., 2000; Sempere et al., 2004). Subsequently, miRNA microarray platforms were developed which greatly expedited expression analysis of the rapidly growing catalog of known miRNAs in various organisms (Thomson et al., 2004). More recently, high-throughput or "deep-sequencing" technology has given miRNA researchers an immensely powerful tool in the detection of both abundant and rare miRNAs as well as in the identification of novel miRNA genes (Hafner et al., 2008).

Research in miRNA detection over nearly the past decade has culminated in the classification of miRNA expression patterns during development and in the pathogenesis

of many diseases. For example the pattern of miRNA expression during murine embryonic development generally falls into three categories and is illustrated in **Figure 1.3A**. In embryonic stem (ES) cells and in very early embryonic development, most miRNAs are not detectable save for the "early embryonic" miRNAs (which include the miR-302 and miR-290 families); it is now known that these miRNAs help to maintain the unique cell-cycle of ES-cells which contributes to ES-cell self renewal (Wang et al., 2008). Conversely, the majority of known miRNAs are robustly induced later in development, during tissue formation and throughout the remainder of life. This differentiation-specific cohort includes not only tissue-restricted miRNAs, such as miR-124 in the brain, miR-1 and miR-133 in the heart and skeletal muscle and miR-122 in the liver but also more widely-expressed miRNAs such as the let-7 and miR-30 families. Last, expression of the miR-17-92 polycistron and related miRNAs is continuous throughout development although these miRNAs are more abundant early on.

Expression profiling has also been widely used to create miRNA "signatures" for many diseases, especially in cancer research. This is best exemplified in a seminal study that employed a bead-based fluorescence assay to compare miRNA levels in hundreds of different tumor samples with those found in normal tissue(J. Lu et al., 2005). Amazingly, the miRNA signatures in various cancers faithfully reflected their tissue origin; furthermore, using miRNA profiles to identify the origin of poorly differentiated tumors was more effective than using mRNA expression data. Most intriguingly, this study revealed that there is a widespread down-regulation of mature miRNAs in many types cancer; this global absence of miRNAs is reminiscent of miRNA expression in early embryogenesis **(Figure 1.3B)**. Thus, it is tempting to speculate that one overall function of the miRNA pathway is to maintain the differentiated state of adult tissues.

Initial evidence for the post-transcriptional regulation of miRNA biogenesis:

Many early studies of miRNA expression contributed provided insight into the

potential functions of miRNAs in vivo. However, what was largely underappreciated for some time was the possibility that miRNA abundance is controlled not only at the level of transcription but that post-transcriptional mechanism could be involved. Over the past several years, many studies have made it clear that miRNA abundance is often post-transcriptionally regulated, suggesting that the precise spatio-temporal control of miRNA expression is crucial for normal cell function. One of the earliest indications of post-transcriptional regulation was the subtle observation that pri-let-7a-1 is detectable in both differentiated and undifferentiated ES cells by reverse transcription followed by polymerase chain reaction (RT-PCR, (Suh et al., 2004)); this presented an exciting scenario where mature let-7a did not accumulate possibly due to a block of Dicer or Drosha processing. Another interesting report demonstrated that pre-miR-138-2 was ubiquitously expressed in mouse tissues but the corresponding mature miRNA is restricted to only certain compartments by late embryogenesis; this was first evidence that some miRNAs might be blocked at the point of Dicer cleavage (Obernosterer et al., 2006). Furthermore, an intriguing study of breast cancer samples found dysregulation of ceratin genes encoding miRNA biogenesis machinery, providing a partial explanation for the broad down-regulation of miRNA abundance in cancer(Blenkiron et al., 2007).

miRNA levels are post-transcriptionally regulated not only during miRNA processing but also at the point of mature miRNA turnover. Mendell and colleagues first observed that miR-29 abundance was regulated by a novel mechanism in human cell culture that involved nuclear import of mature miR-29, possibly followed by its degradation (Hwang, Wentzel, & Mendell, 2007). A recent report in C. elegans revealed that mature miRNA turnover by the exonuclease XRN-2 is coupled to RISC loading in vitro and that the mature miRNA is protected from turnover when annealed to a target RNA (Chatterjee & Grosshans, 2009). Two somewhat related discoveries showed that the degradation of a miRNA is dependent on the extent to which it is base paired with its target RNA (Ameres et al., 2010; Cazalla, Yario, & Steitz, 2010). An equally interesting study revealed that certain miRNAs in photoreceptor

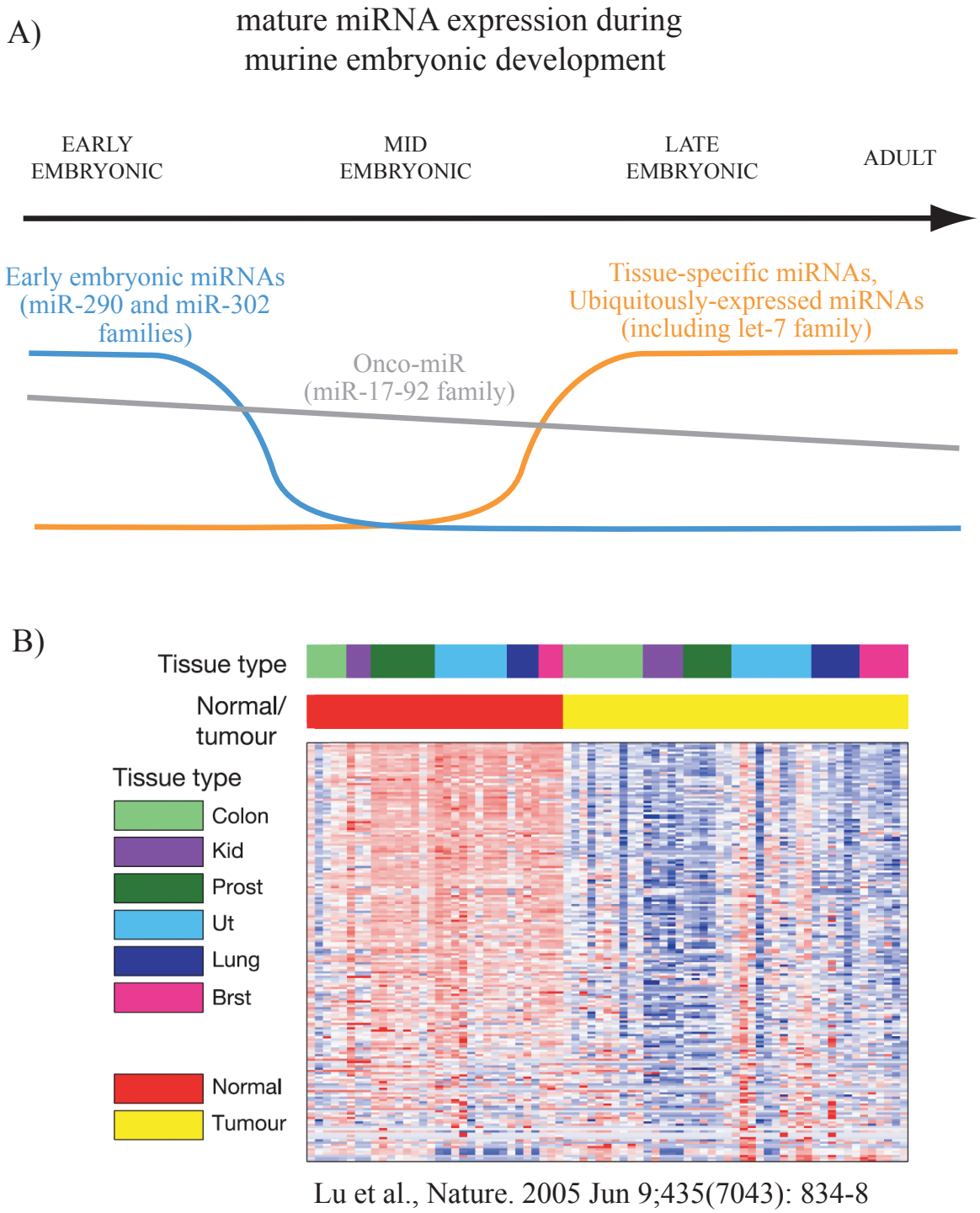


Figure 1.3: miRNA expression patterns. A) A generalized view of miRNA expression patterns during murine embryonic development. B) miRNA expression heat map from a key study (Lu et al., Nature. 2005 Jun 9;435(7043): 834-8) where hundreds of different tumors were profiled for miRNA expression versus corresponding normal tissue. This experiment was the first to indicate that there is a global down-regualtion of miRNA abundance in cancer.

cells of the retina, as well as in certain neurons are rapidly turned over (Krol et al., 2010). Taken together, there is a broad spectrum of biochemical mechanisms by which miRNA expression is post-transcriptionally regulated.

The let-7 family of miRNAs:

let-7 is an ancient and highly conserved miRNA, likely existing in most, if not all bilaterian animals (Christodoulou et al., 2010). While simpler organisms such as flies and worms possess only one let-7 miRNA gene, vertebrate genomes encode an expanded family of closely related let-7 genes. For example, mice and humans have 12 let-7 genes that encode a total of 9 mature miRNAs (**Figure 1.4**). let-7 was the second of the two "founding" miRNAs discovered by the Ruvkun lab in 2000 (Reinhart et al., 2000). While function of this miRNA family isn't exactly understood, it exhibits similar expression patterns in diverse animals and is generally associated with cell differentiation. In *C. elegans let-7* was first found to target critical heterochronic genes such as *lin-41, lin-14 and lin-28* (Reinhart et al., 2000); in mammals, let-7 miRNAs target a variety of oncogenes such as Ras, HMGA2, Myc, Cdc25a and Cdk6 (Bussing, Slack, & Grosshans, 2008). These findings have further revealed the function of let-7 miRNAs as tumor suppressors. Consistent with this notion, mature Let-7 is down regulated in many cancers and is most depleted in poorly differentiated tumors (Shell et al., 2007).

Lin28:

Lin28 is an ES-cell specific RNA binding protein that was originally identified in *C. elegans* as a heterochronic gene (Ambros & Horvitz, 1984); mutations in the *lin-28* gene caused precocious differentiation of many cell lineages implicating *lin-28* as a critical regulator of early development in the worm. It was later discerned that *lin-28* encodes an RNA binding protein with a cold-shock domain (CSD) and 2 CCHC zinc finger motifs,

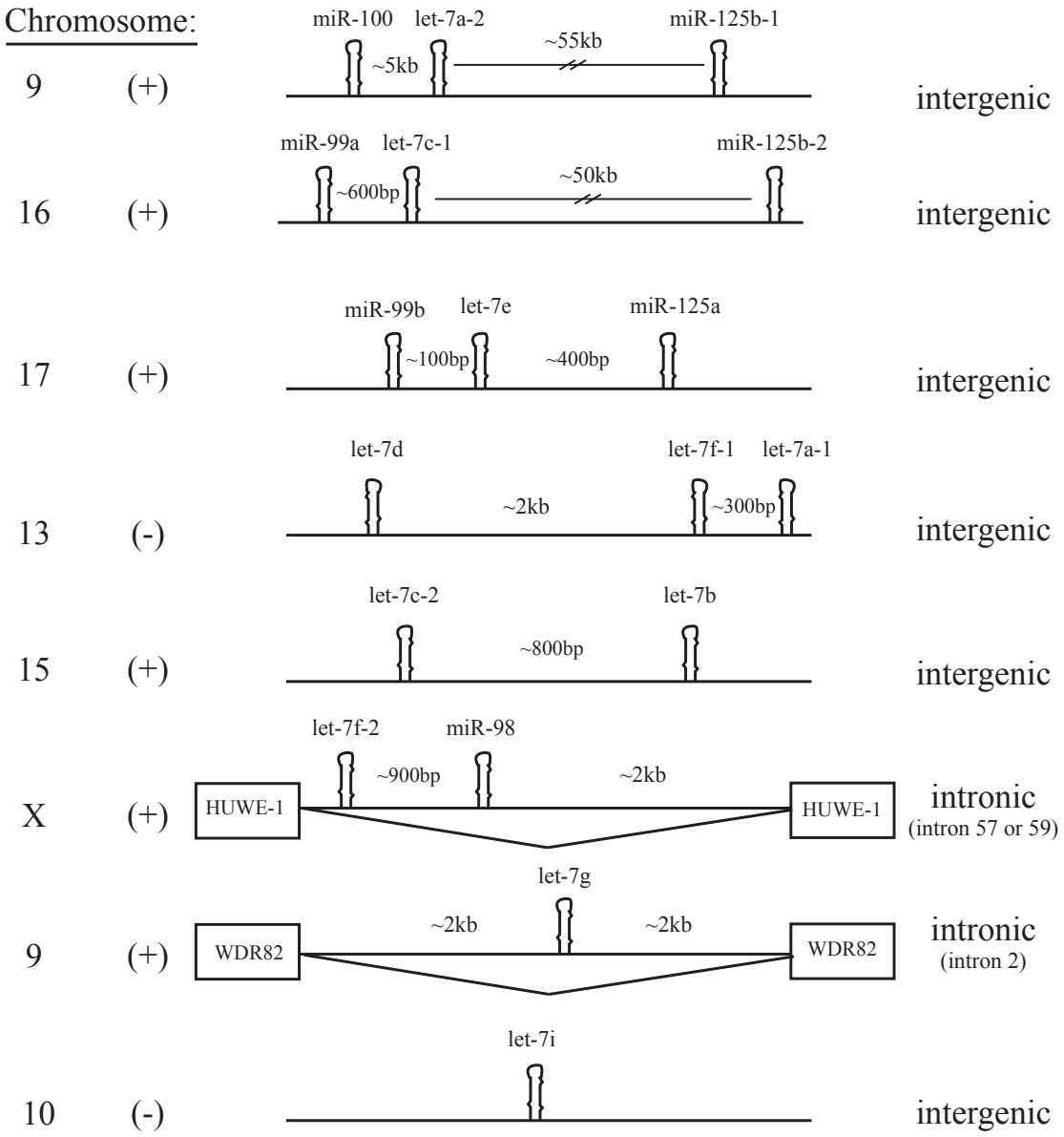


Figure 1.4: Genomic organization of the let-7 family of miRNAs in mice. Distances between miRNA stemloops are not drawn to scale. (-) and (+) denotes the chromosome strand from which the pri-miRNA transcript originates. White, labeled boxes denote exons of a protein coding transcript.

Lin28:

Lin28 is an ES-cell specific RNA binding protein that was originally identified in *C. elegans* as a heterochronic gene (Ambros & Horvitz, 1984); mutations in the *lin-28* gene caused precocious differentiation of many cell lineages implicating *lin-28* as a critical regulator of early development in the worm. It was later discerned that *lin-28* encodes an RNA binding protein with a cold-shock domain (CSD) and 2 CCHC zinc finger motifs, a unique domain architecture among known proteins (Moss, Lee, & Ambros, 1997b); furthermore, lin-28 mRNA is targeted in its 3' UTR by the lin-4 and let-7 miRNAs (Moss, Lee, & Ambros, 1997b; Reinhart et al., 2000). Several studies in mammalian cell culture revealed that the Lin28 protein associates with actively translated mRNA and promotes the translation of the IGF-2 mRNA suggesting it functions, in part, in the post-transcriptional regulation of genes in undifferentiated cells (Balzer & Moss, 2007; Polesskaya et al., 2007). Mammals possess two separate genes that encode to homologs of the C. elegans lin-28, Lin28a and Lin28b. Interestingly, not only does Lin28 play an important biological role in undifferentiated cells but also Lin28b is overexpressed in hepatocellular carcinoma (HCC), implicating mammalian Lin28a and Lin28b as potential oncoproteins (Guo et al., 2006). Amazingly, the ectopic expression of the lin28a gene in combination with several other ES-cell-specific factors causes "reprogramming" of somatic cells into induced pluripotent stem (iPS) cells (Yu et al., 2007).

Non-canonical poly (A) polymerases and the modulation of RNA stability and activity:

As some of the research described in this dissertation focuses on the uridylation of pre-miRNAs, it is worth discussing terminal uridyl transferases (or TUTases) here. TUTases are part of a diverse family of RNA-specific ribonucleotidyl transferase enzymes (Martin & Keller, 2007) which also includes the canonical Poly (A) polymerases (or P(A) Ps) found in animals that add a long tail of adenosines to mature mRNAs (Wilusz & Wilusz, 2008). The TUTases (sometimes, confusingly, called non-canonical P(A)Ps) differ from

the canonical P(A)Ps in that they add shorter stretches of adenosines or uridines to specific RNAs, depending on the TUTase (Wilusz & Wilusz, 2008). While little is known about the cellular function of these enzymes, they have been implicated in diverse processes, including U6 snRNA recycling (Trippe et al., 2006), histone mRNA degradation (Mullen & Marzluff, 2008), the uridylation of miRNA-mediated cleavage products in diverse organisms (Shen & Goodman, 2004) as well as in the stabilization of miR-122 (Katoh et al., 2009). Intriguingly, a recent study by Bartel and colleagues G&D (2010) revealed that many miRNA 3' strands ("3p" or "star" strands) possess 3' terminal untemplated uridines, possibly indicating the widespread uridylation of pre-miRNAs (Chiang et al., 2010).

CHAPTER 2:

Extensive post-transcriptional regulation of microRNAs and its implications for cancer[1].

Summary:

MicroRNAs (miRNAs) are short, non-coding RNAs that post-transcriptionally regulate gene expression. While hundreds of mammalian miRNA genes have been identified, little is known about the pathways that regulate the production of active miRNA species. Here we show that a large fraction of miRNA genes are regulated post-transcriptionally. During early mouse development, many miRNA primary transcripts, including the let-7 family, are present at high levels but are not processed by the enzyme Drosha. An analysis of gene expression in primary tumors indicates that the widespread down-regulation of miRNAs observed in cancer is due to a failure at the Drosha processing step. This data uncovers a novel regulatory step in miRNA function, and provides a mechanism for miRNA down-regulation in cancer.

Introduction:

The founding miRNA, *lin-4*, was discovered more than 20 years ago in the nematode *Caenorhabditis elegans* (R. C. Lee et al., 1993). Elegant studies demonstrated that the gene product of *lin-4* was a non-coding RNA that post-transcriptionally regulated several critical genes during development (Wightman et al., 1993). More recently a large number of miRNA genes have been identified in plant and animal genomes. To date over 300

---

1 Reprinted from Thomson et al., Genes Dev. 2006 Aug 15;20(16):2202-7 with permission from Cold Spring Harbor Press. This manuscript has been formatted for the dissertation.

miRNAs have been validated in the human genome, and computational predictions suggest that many hundreds more exist (Aravin & Tuschl, 2005). While it is clear that these small, non-coding RNAs have essential functions in mammalian biology, few miRNA genes have been functionally linked to specific cellular pathways. For example, miR-181 modulates hematopoietic differentiation, and miR-1 and miR-133 participate in skeletal and cardiac muscle development. The polycistronic cluster miR-17~92 promotes lymphomagenesis, while let-7 suppresses proliferation by targeting the oncogene Ras (see Wienholds & Plasterk, 2005 for a review).

For most miRNA genes, however, biological information is limited to expression analyses. In mouse and zebrafish, few miRNAs are expressed in early embryos (Thomson, Parker, Perou, & Hammond, 2004b; Wienholds et al., 2005). During mid to late embryonic development, large numbers of miRNAs are induced in temporal and spatial patterns (Kloosterman, Wienholds et al., 2006). This culminates in adult tissues, where a large fraction of known miRNA genes are expressed. Interestingly, most of these same miRNA genes are down-regulated in cancer, perhaps reflecting a loss of cellular differentiation (J. Lu et al., 2005; Takamizawa et al., 2004). The regulatory mechanism behind these expression changes is largely unknown, though the most obvious candidate is transcriptional control. It is difficult, however, to reconcile this model with the widespread, parallel reduction in miRNA expression observed in most types of human cancer.

A second possibility is that miRNA maturation is a regulated event. The biogenesis of miRNAs begins with a primary transcript, termed the pri-miRNA, which is generated by RNA polymerase II (V. N. Kim, 2005). The active species is contained in a stem-loop structure which is liberated by the nuclear ribonuclease III (RNase III) Drosha. This stem loop, termed the pre-miRNA or precursor, is exported from the nucleus in a Ran/GTP/ Exportin 5 mediated event. In the cytoplasm, the precursor is further processed by the

RNAseIII enzyme Dicer to generate the mature miRNA species. In a concerted reaction, this siRNA-like product is loaded into the RNA induced silencing complex (RISC) where it mediates target mRNA translational suppression (Du & Zamore, 2005; Gregory et al., 2005; Maniataki & Mourelatos, 2005).

In principal, any step during this maturation process could be regulated. The core enzymes are widely expressed, however, and no post-translational regulation of Drosha or Dicer has been reported. Discrepancies between the levels of primary transcript, precursor, and mature miRNA species have been reported, however no clear model is apparent (Eis et al., 2005; Suh et al., 2004). Here we report that a large fraction of miRNAs are regulated during the Drosha processing step, and this regulation has a major impact on miRNA expression during embryonic development and in cancer.

Results and Discussion:

A striking event during mouse development is the massive induction of Let-7 family miRNAs at approximately 10.5 days of gestation. This is exemplified by Let-7g. This miRNA is housed in the second intron of the WD40 repeat gene AK037620 (**Figure 2.1A**). Processing of this primary transcript by Drosha yields a 79 nucleotide stem-loop precursor. Further processing by Dicer leads to the 21 nucleotide mature species. We analyzed the expression of all three molecular forms by northern blotting (**Figure 2.1B**). As previously reported, mature Let-7g is undetectable in embryonic stem (ES) cells and P19 embryonal teratocarcinoma cells (Thomson et al., 2004). Expression is detectable at 10.5 day gestation and is high at 14.5 day. The precursor is detectable slightly earlier than the mature miRNA, suggesting a delay at the Dicer processing step. Similar results have been observed in mouse and Drosophila development (Hutvagner et al., 2001; Schulman, Esquela-Kerscher, & Slack, 2005). Surprisingly, the primary transcript is highly expressed throughout development, with

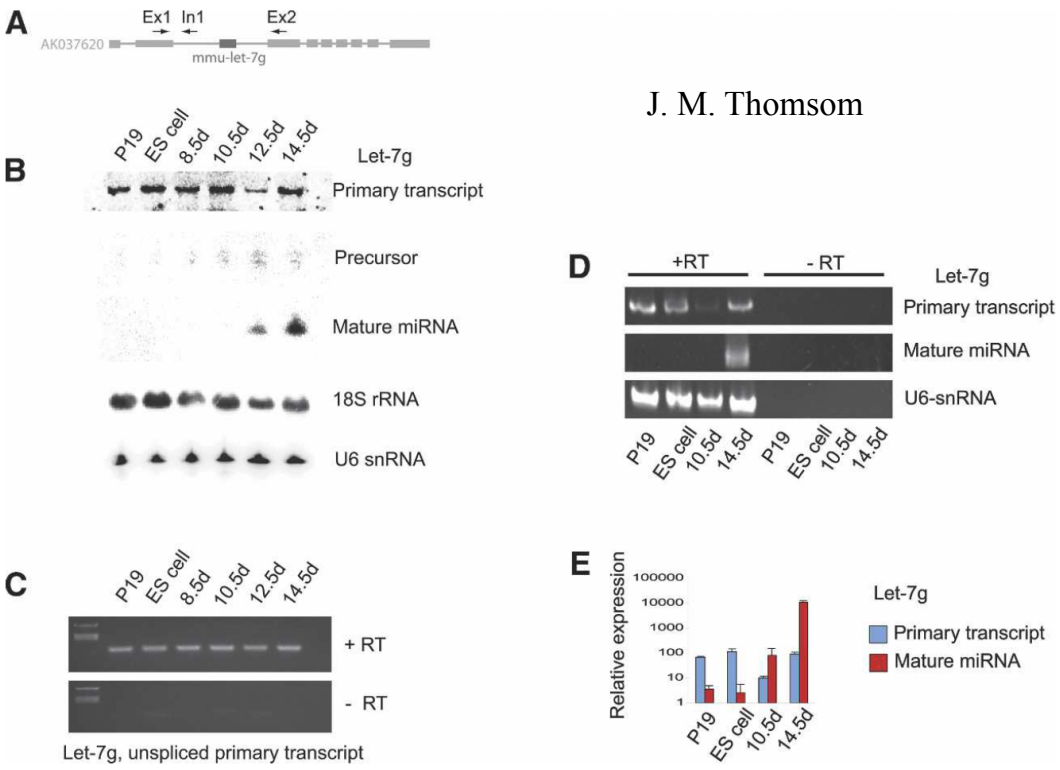


Figure 2.1. Expression of let-7g during mouse development (These experiments were carried out by J. Michael Thomson). (A) The genomic organization of Let-7g is shown. RT-PCR primer sites are indicated by arrows. Primer Ex1 and Ex2 were used for PCR of the primary transcript. Primer Ex1 and In1 were used for PCR of the unspliced primary transcript. (B) Northern blot analysis of the molecular species of Let-7g. 18S rRNA and U6 snRNA were used for loading controls for pri-miRNA and pre-miRNA/mature, respectively. (C) Non-quantitative RT-PCR of the unspliced primary transcript. Control reactions without the reverse transcription step are shown. (D) Non-quantitative RT-PCR of the let-7g primary transcript and mature species. The U6-snRNA reference is also shown. Right lanes are control reactions without reverse transcription step. (E) Real-time RT-PCR analysis of let-7g primary transcript and mature species. Reactions were performed in triplicate and normalized to U6 cycle threshold values. Expression is shown as relative values on a logarithmic scale.

little difference in expression between ES cells and 14.5 day embryos. This discrepancy is not resolved by the amount of unspliced pri-miRNA, since this is also essentially constant (**Figure 2.1C**).

We next quantitated the levels of primary and mature species by real time RT-PCR. We adapted a published procedure to measure the amount of both species from a single reverse transcription (RT) reaction (Shi & Chiang, 2005). (see Supplemental Figures 1 and 2 for validation; **Figure 2.1D** illustrates a non-quantitative example). U6 snRNA was used for normalization for mature and primary species. This RNA exhibited more consistent expression during embryonic development than more typical reference RNAs such as beta-2 microglobulin, glyceraldehyde-3-phosphate dehydrogenase, and beta-glucuronidase. Using this procedure, we determined that the amount of mature let-7g increases 4200 fold between ES cells and 14.5 day embryo. As observed by Northern blot analysis, expression levels of the primary transcript does not match the mature miRNA, but is relatively constant (**Figure 2.1E**, note log scale).

There are several possible explanations for the different expression levels of the pri-miRNA and the mature miRNA. There could be rapid turnover of the mature species in ES cells compared to late embryos. This is unlikely, since mature miRNAs from other gene families are present in ES cells (e.g. miR-290~295), and published data suggests that all miRNAs are loaded into RISC complexes non-discriminately (Liu et al., 2004). A second possibility is that the precursor for let-7g is retained in the nucleus, or otherwise made inaccessible to Dicer. In that circumstance we would expect to see an accumulation of precursor. Our data disagrees with this, since the precursor is essentially undetectable in ES cells. A third possibility, most congruous with our data, is that processing of let-7g is blocked at the Drosha step. Release of the block would enable production of mature let-7g. It is worth emphasizing that this is not simply a delay in processing during development, but

a complete block. ES cells are grown continuously in culture, express high levels of let-7g primary transcript, and never process a significant quantity to the precursor or mature species.

We next investigated processing of other miRNA genes that are expressed during development (**Figure 2.2**). Let-7 family miRNAs are located at eight genomic loci (Griffiths-Jones et al., 2006). Of these we were able to detect primary transcripts for four loci. Notable is the let-7f-2/ miR-98 cistron, which is located within an intron of the widely expressed ubiquitin ligase gene HUWE-1. All four primary transcripts were expressed at similar levels throughout development. Similar to let-7g, all mature miRNAs within these genes were elevated several thousand fold between ES cells and 14.5 day gestation. The correlation between primary transcript expression and mature miRNA expression was close to zero, with the exception of let-7i (**see Table 2.1**). For this miRNA, the primary transcript is induced during development, though there is still a large difference in processing efficiency (300 fold, **Table 2.1**). Two of the four let-7 clusters we analyzed were intronic, and two were exonic. Thus there is no correlation between regulated processing by Drosha and intron/exon location.

This Drosha regulatory step has not been observed in other organisms. *C. elegans let-7*, for example, is induced at the fourth larval stage and continues into the adult. Expression is controlled by defined enhancer elements in the promoter of the primary transcript (Johnson, Lin, & Slack, 2003a). Expression of the primary transcript is closely correlated with mature *let-7* (Bracht et al., 2004).

The developmental function of let-7 family members in mammals is not known. There have been reports of decreased let-7 expression in primary tumors (J. Lu et al., 2005; Takamizawa et al., 2004). This may have functional consequences. The oncogene Ras is a validated target of let-7 (Johnson et al., 2005). Thus, decreased let-7 expression would lead

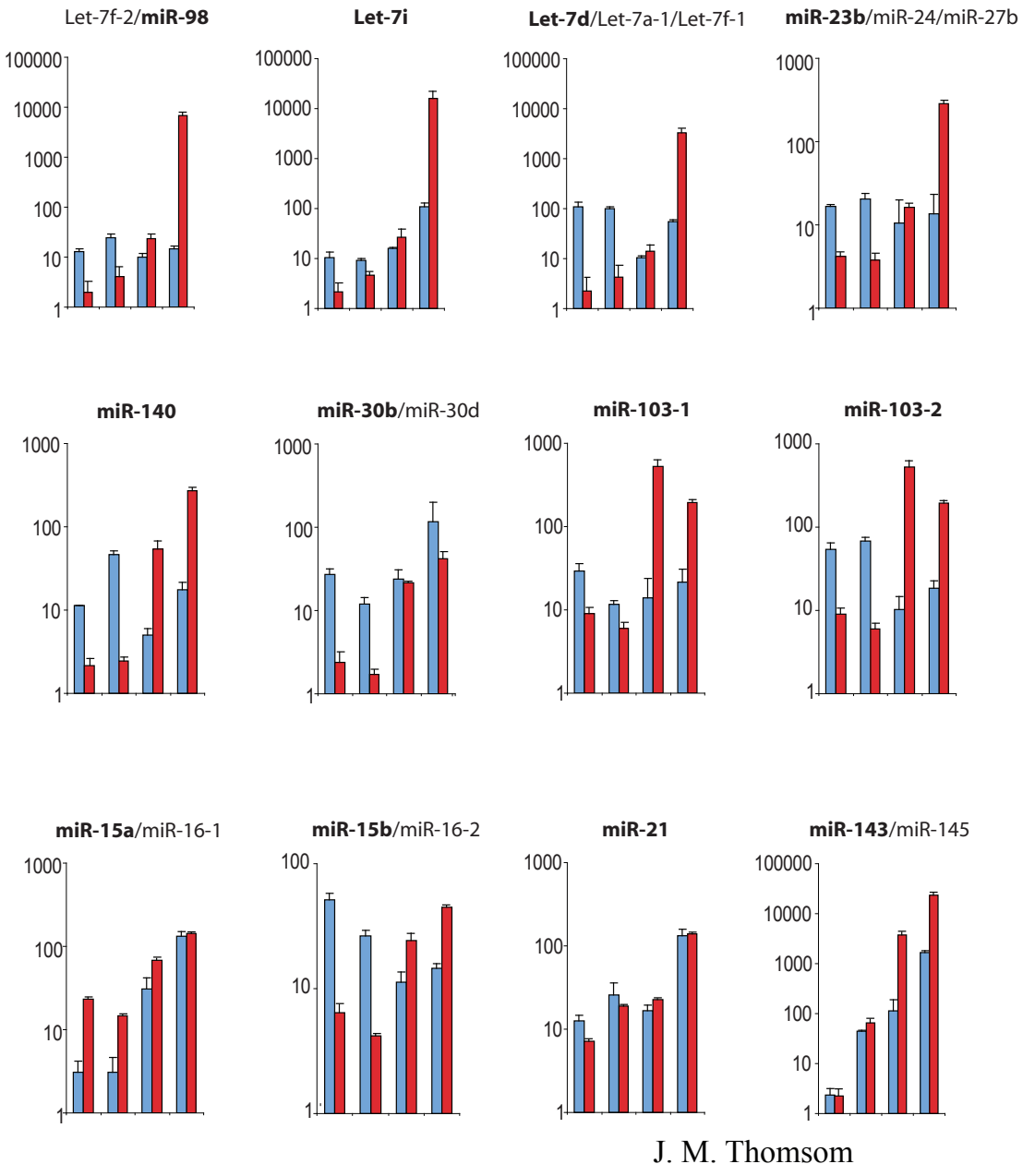


Figure 2.2. miRNA expression during mouse development (These experiments were carried out by J. Michael Thomson). The expression levels of primary transcript and mature species from 12 representative miRNAs are shown. Red bars represent mature miRNA and blue bars represent primary transcript. All analyses are performed with the following RNA samples, from left to right: P19 teratocarcinoma cell line; embryonic stem cell line; 10.5 day gestation mouse embryo; 14.5 day gestation mouse embryo. Reactions were performed in triplicate and normalized to U6 cycle threshold values. Expression is shown as relative values on a logarithmic scale. The bold font indicates the mature miRNA gene that was used for RT-PCR analysis. In all cases these are unique miRNA sequences in the mouse genome.

to increased Ras expression, which would promote survival of tumor cells. Our data raises the possibility that let-7 down-regulation in tumors is post-transcriptional (see below).

Several other families of miRNAs are up-regulated during mouse development. We were able to detect primary transcripts for nine of these miRNAs. Most cases exhibited regulated Drosha processing (**Figure 2.2, Table 2.1**). Exceptions were miR-21 and the miR-143/miR-145 cluster. Both primary transcripts matched expression of the mature species (correlation = 1.0 and 0.99, respectively). Another interesting example is the miR-15/miR-16 cistron, which is located at two genomic loci. Each primary transcript displays distinct expression kinetics. The chromosome 14 copy is located within an uncharacterized gene which is induced during development. The chromosome 3 copy, in contrast, is located within the condensin subunit SMC4L1 which is constitutively expressed. Thus, each pri-miRNA contributes to the mature species in a complex manner, with a post-transcriptionally regulated component (chromosome 3) and a transcriptionally regulated component (chromosome 14).

It should be noted that coordinate expression of the primary transcript and the mature miRNA does not prove that Drosha processing is not regulated. Such miRNAs (bottom panel, Table 1) may be regulated at transcription and Drosha processing. It remains a possibility that all miRNAs are regulated at the Drosha step, except for miRNAs that have demonstrated expression in the early mouse embryo. Interestingly, this has implications for experiments that utilize ectopic expression of miRNAs in ES cells. We predict that this would fail to produce the mature species, unless cis-regulatory elements were removed from the expression cassette.

Our data suggests that the differentiation events that occur during embryonic development activate Drosha processing of specific miRNAs. To directly test this we used

miRNA gene	pri-miRNA/mature correlation	D processing efficiency
let-7g	+0.29	5400
let-7f-2/miR-98	-0.12	2700
let-7i	+1.00	300
let-7d/let-7a-1/let-7f-1	-0.21	1400
miR-140	-0.20	290
miR-23b/miR-24-2/miR-27b	-0.29	110
miR-30b/miR-30d	+0.89	2.5
miR-103-1	-0.38	18
miR-103-2	-0.87	120
miR-15b/miR-16-2	-0.68	19
miR-15a/miR-16-1	+0.98	0.23
miR-21	+0.99	1.4
miR-143/miR-145	+0.99	9.3

J.M. Thsomson and S.M. Hammond

Table 2.1. Expression correlation for mature/pri-miRNA pairs (this analysis was carried out by J. Michael Thomson and Scott Hammond). Pearson correlations were calculated for each mature miRNA/pri-miRNA pair and are shown in the first column. The change in processing efficiency is defined as the ratio of mature to pri-miRNA in 14.5 day embryo divided by the ratio of mature to pri-miRNA in ES cells, and is shown in the right column. The table is divided into miRNAs regulated by Drosha (top) and miRNAs potentially unregulated by Drosha (bottom).

the teratocarcinoma cell line P19, which can be differentiated in culture into multiple cell types (Rudnicki et al., 1990). miRNA microarray expression analysis of untreated and differentiated P19 cells mirrors the expression changes observed during mouse development (data not shown). Embryonic miRNAs, such as the miR-290~295 cluster, are down-regulated during differentiation. In contrast, miRNAs that are up-regulated during mouse development, including let-7 family members, are increased upon differentiation. As in embryonic development, the increases in let-7 miRNAs are not coupled to transcription of the pri-miRNA (**see Figure 2.3**).

One experimental observation of cancer cells is the widespread alterations in miRNA expression. While some miRNAs are elevated, for example the oncogenic cluster miR-17~92, most miRNAs have significantly reduced expression (L. He et al., 2005; J. Lu et al., 2005). Our data raises the possibility that these reductions are a consequence of the Drosha processing block. To address this we analyzed existing expression data, comparing mature miRNA levels with the primary transcript. The most complete data set, from the Broad Institute, has Affymetrix mRNA expression data and Luminex miRNA data from a wide range of primary tumors and normal tissues (J. Lu et al., 2005; Ramaswamy et al., 2001). We combined the data set in the following manner: 68 tumors and 21 normal tissue samples had data for miRNA and mRNA expression. Of 217 miRNAs, 22 could be mapped to a primary transcript that was present on the Affymetrix microarray. When this restricted miRNA dataset was hierarchically clustered, the normal samples were located in a single node with some tumor samples, while most tumors clustered in a separate node (**Figure 2.4A**). The tumor samples had overall reduction in miRNA expression levels. This was qualitatively similar to the entire miRNA expression map, as previously published (J. Lu et al., 2005).

We then created an expression map of primary transcript expression levels (**Figure**

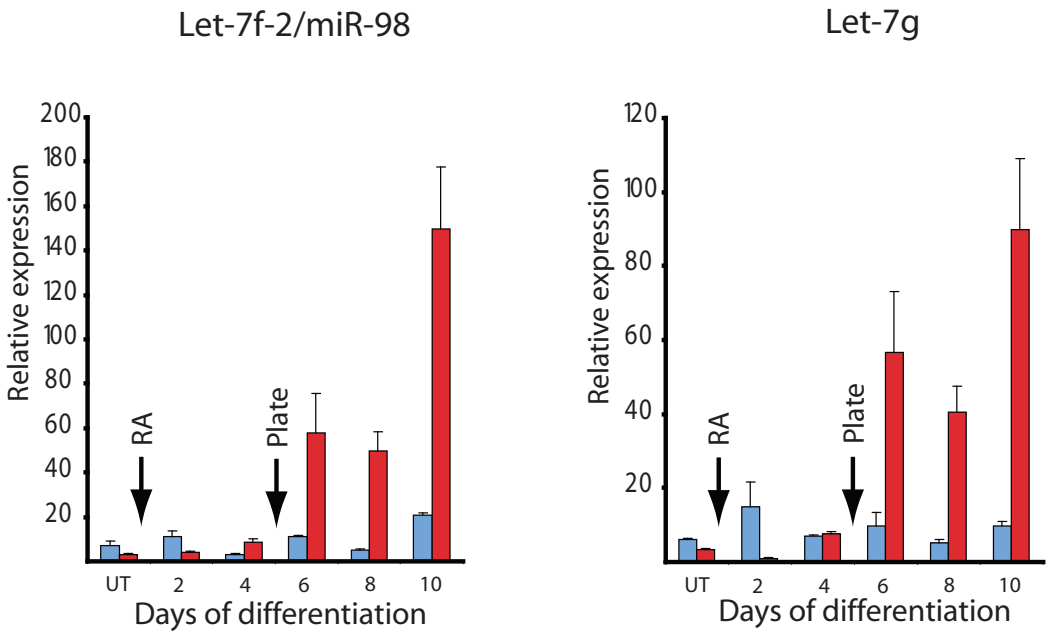


Figure 2.3. miRNA expression during P19 cell differentiation. P19 teratocarcinoma cells were differentiated by forming embryoid bodies for four days in the presence of all-trans retinoic acid (RA), followed by plating without RA. Arrows indicate RA treatment and plating time. Primary transcript and mature miRNA species were quantitated by real-time RT-PCR. Reactions were performed in triplicate and normalized to U6 cycle threshold values. Expression is shown as relative values on a linear scale. Red bars indicate mature miRNA and blue bars indicate pri-miRNA.

2.4A). The samples and genes are arranged in the same order as for the miRNA expression map. The high and low clusters of expression are completely lost in the primary transcript expression map. We performed correlation analysis between each mature/pri-miRNA expression pair, for each sample. The data is represented in the histogram in **Figure 2.4B**. There is no correlation between pri-miRNA and mature expression in the tumor samples, while the normal tissue samples had positive correlation (KS test, $p < 0.01$). This demonstrates that the miRNA alterations that occur in tumors, for this limited set of genes, are not due to mis-regulated transcription.

Our data suggests a multistep model for the control of miRNA expression. Transcription of the pri-miRNA can be regulated, as has been demonstrated for tissue specific miRNAs (J. F. Chen et al., 2006; Fazi et al., 2005; O'Donnell et al., 2005; Zhao, Samal, & Srivastava, 2005). Processing at the Dicer step can be delayed or inhibited (Hutvagner et al., 2001; Obernosterer et al., 2006; Schulman et al., 2005). We demonstrate that further suppression of miRNA production is achieved at the Drosha step. This may be necessary for early development, where inappropriate expression of even small amounts of let-7 may promote differentiation, with disastrous consequences. This model also allows miRNAs to be located in essential housekeeping genes that are ubiquitously expressed.

The biochemical nature of Drosha regulation is unknown. We measured the expression levels of Drosha and its binding partner DGCR8, during the P19 differentiation time course (Supplemental Figure 3). No induction occurred concordant with the production of mature miRNAs. It is possible these proteins are regulated by post-translational modification. Alternatively, additional regulatory binding proteins may be required for specific miRNA processing. Recently, adensosine deaminase (ADAR) editing of specific pri-miRNAs has been reported (W. Yang et al., 2006). This editing event leads to decreased processing of the miRNA by Drosha and increased turnover by the Tudor-SN nuclease. We

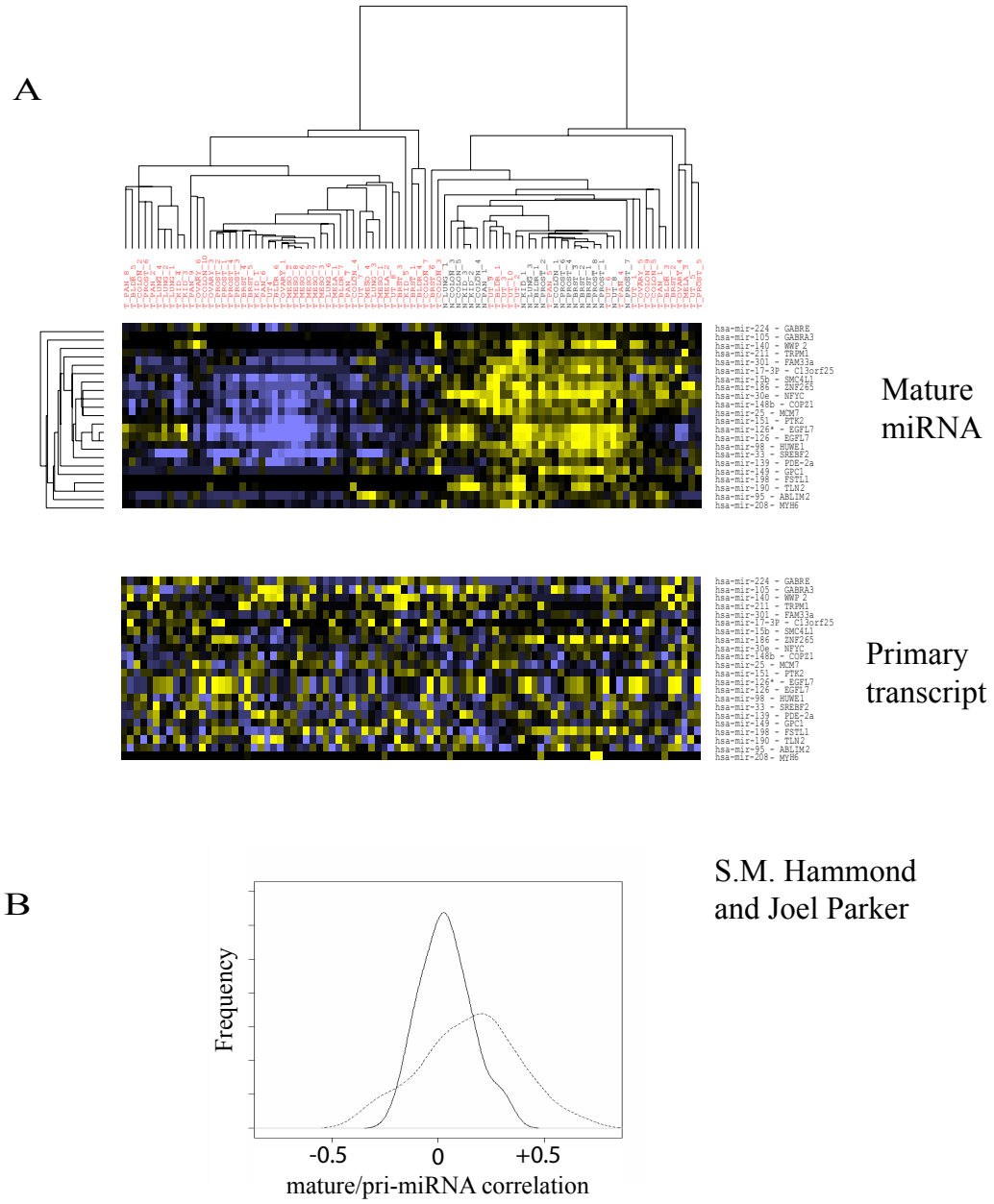


Figure 2.4. miRNA expression in primary tumors (This data analysis was carried out by Scott Hammond and Joel Parker). (A) Normalized miRNA expression data from (Lu et al. 2005) was clustered hierarchically in both dimensions and is displayed as an expression map. Yellow indicates increased expression and blue indicates decreased expression, relative to the median. Primary transcript expression data from (Ramaswamy et al. 2001) was mapped in the same order as the mature miRNA expression map. Mature miRNA/primary transcript gene pairs are indicated on the right. Tumor samples are indicated in red and normal tissues indicated in black. (B) Pearson correlation values for each mature/pri-miRNA value pair were calculated. Values for normal samples (dashed line) and tumor samples (solid line) are plotted as a histogram.

tested whether this mechanism is responsible for the Drosha block during embryogenesis. We failed to find evidence of A-I editing of let-7 pri-miRNAs in ES cells (not shown). Our data does not, however, discount the possibility of other structural changes in the pri-miRNA that block processing. An alternative mechanism is the sequestration of regulated pri-miRNAs away from the processing apparatus. The exact nature of the regulatory mechanism awaits further investigation.

Materials and Methods.

Cell lines.

The feederless ES cell line E14Tg2A.4 was grown on gelatin-coated plates in Glasgow MEM (Sigma) supplemented with 2 mM glutamine, 1 mM pyruvate, 1X non-essential amino acids, 15% fetal bovine serum (Hyclone), 0.1 mM b-mercaptoethanol, and 10^3 units/ml leukocyte inhibitory factor (Chemicon). The P19 teratocarcinoma cell line was grown in DMEM (Sigma) supplemented with 7.5% calf serum (Hyclone) and 2.5% fetal bovine serum. P19 cells were differentiated as follows: Cells were seeded on non-adhesive plates in a-MEM (Sigma) supplemented with 5% fetal bovine serum and 500 nM all-trans retinoic acid (Sigma). After four days the aggregated cells were plated on standard tissue culture dishes in DMEM supplemented with 7.5% calf serum and 2.5% fetal bovine serum, and grown for the indicated number of days.

RNA isolation

Dissected whole embryos from pregnant CD1 mice were flash frozen in liquid nitrogen and pulverized into powder. Total RNA was extracted from pooled embryos of two to eight embryos or from cell lines with Trizol Reagent (Invitrogen, Carlsbad, CA, USA)

according to manufacturer's instructions.

Northern blot analysis of mRNA

5 µg total RNA was resolved on a 1% agarose formaldehyde gel and transferred overnight to positively charged nylon membrane by capillary transfer with 10X SSC. The membrane was washed briefly in 2X SSC, dried, and UV cross-linked, and incubated overnight with 10^7 cpm of a random-primed probe directed against the mouse Let-7g primary transcript in Church and Gilbert high stringency buffer. PCR primers specific for the mouse Let-7g primary transcript (let-7g-probe-F 5'-TTTAGTCCCAATGGCGAGAC-3'; let-7g-probe-R 5'-CATCACCACACCCTTGAATG-3') were used to make the probe. The membrane was washed 1 x 15 min at 420C in 2X SSC, 0.1% SDS and then 2 x 15 min 650C in 0.2X SSC, 0.1% SDS. Phosphorimaging using the Storm system (Molecular Dynamics) revealed the let-7g primary transcript. The size of pri-Let-7g (3.7 kb) was confirmed against 18S and 28S rRNA markers. The size of pre-Let-7g was confirmed against an in vitro transcribed pre-Let-7g RNA. All primers were purchased from MWG Biotech (Highpoint, NC, USA).

Northern blot analysis of pre-miRNA and mature miRNA

5 ug total RNA was resolved on 15% (mature) or 7.5%(precursor) acrylamide-8M urea-TAE gels as described previously (Hammond, Bernstein, Beach, & Hannon, 2000). After electroblotting to Hybond N+ membranes, blots were probed with an end-labeled Locked Nucleic Acid (LNA) probe ACTgTaCaAaCtAcTaCcTcA (small letters are LNA, caps are DNA). Hybridization was performed in Church and Gilbert high stringency buffer at 65 C, followed by 3 washes in 0.1X SSC at 65 C.

Quantitative RT-PCR

Total RNA, 20 µg, was DNase I (Promega, Madison, WI, USA) treated according

to manufacturer's instructions, phenol:chloroform extracted, ethanol precipitated and dissolved in DEPC-treated dH_2O. RNA, 10µg, was polyadenylated using Poly(A) polymerase (Ambion) according to manufacturer's instructions, phenol:chloroform extracted, ethanol precipitated and dissolved in DEPC-treated dH_20. A modified cDNA was made as follows; 10 µg of polyadenylated RNA was reverse transcribed using Superscript II reverse transcriptase (Invitrogen, Carlsbad, CA, USA) using 2.5 ug random hexamers and 500 ng oligo-dT adapter primer (5'- GCGAGCACAGAATTAATACGACTCACTATAGGTTTTTTTTTTTTVN-3') according to manufacturers instructions. The reaction was terminated by incubation at 70^0C for 10 min and diluted into 2 ml dH2O (5 µg/ml). Quantitative PCR was used to measure both the mature miRNA and the host mRNA transcript as follows; 5 µl cDNA was mixed with 5 pmol of both the forward and reverse primers in a final volume of 12.5 µl and mixed with 12.5 µl of 2x SYBR® Green PCR master mix (Applied Biosystems) . Primer sequences and other pri-miRNA information is in Supplemental Table 1. All reactions were run in triplicate on a DNA Engine Opticon® 2 (MJ Research) utilizing two amplification protocols. A standard protocol was performed for the pri-miRNA PCR, 20 s at 94^0C, 30 s at 59^0C and 20 s at 72^0C for 40 cycles. Mature miRNA PCR was performed according to the high-stringency protocol of (Shi and Chiang 2005) except the reverse primer Mir-qPCR-3-3' (5'-GCGAGCACAGAATTAATACGACTCAC

-3') was used in conjunction with an exact sequence-specific primer to each miRNA. Mature and pri-miRNA expression both utilized the reference gene U6 snRNA (U6-F 5'-CGCTTCGGCAGCACATATAC-3' and U6-R 5'-TTCACGAATTTGCGTGTCAT-3'). Expression was calculated using the formula $2^{-\Delta CT}$, with $\Delta C_T = (C_T$ miRNA - C_T reference RNA) for mature miRNA or $\Delta C_T = (C_T$ pri-miRNA - C_T reference RNA) for pri-miRNA. Non-quantitative PCR conditions for the Let-7g pri-miRNA and mature miRNA, and U6 reference gene were as mentioned above, except the cycle number was 40, 35, and 30 respectively. The same conditions were performed on an equal amount of DNAse I-treated RNA as a negative control. Reactions, 5 µl, were resolved by electrophoresis on 12% native

polyacrylamide TBE gels and stained with ethidium bromide. Unspliced pri-miRNA was analyzed by RT-PCR at 50 cycles. The same conditions were performed on an equal amount of DNAse I-treated RNA as a negative control. For analysis of Drosha, DGCR-8, and Oct-4, qRT-PCR was performed similarly, using b2-macroglobulin as a reference RNA.

Microarray gene expression analysis.

Normalized, log transformed expression data was downloaded from http://www.broad.mit.edu/cgi-bin/cancer/datasets.cgi. 89 samples were common to the miRNA and mRNA datasets. 22 miRNA/pri-miRNA pairs were found in the combined data sets. This reduced set of miRNA data was median centered by gene and hierarchically clustered in both dimensions using Cluster software (Stanford University). An expression map was generated using Treeview software (Stanford University). The pri-miRNA data set was arranged in the same order as the miRNA clustered data, median centered by gene, and a non-clustered expression map was generated. Pearson correlation of normalized, log transformed intensities was calculated for each mature/pri-miRNA pair. Pairwise correlations were tabulated across tumor or normal samples, and a smoothed histogram was generated.

Chapter 3:

Lin-28 interaction with the Let-7 precursor loop mediates regulated microRNA processing[2]

Summary

A hallmark of mammalian embryonic development is the widespread induction of microRNA (miRNA) expression. Surprisingly, the transcription of many of these small, noncoding RNAs is unchanged through development; rather, a post-transcriptional regulatory event prevents accumulation of the mature miRNA species. Here we present a biochemical framework for the regulated production of the Let-7 family of miRNAs. Embryonic cells contain a Drosha Inhibitor that prevents processing of the Let-7 primary transcript. This inhibitor specifically binds to conserved nucleotides in the loop region of the Let-7 precursor, and competitor RNAs that mimic the binding site restore Let-7 processing. We have identified the Drosha Inhibitor as the embryonic stem cell specific protein Lin-28. Lin-28 has been previously implicated in developmental regulatory pathways in *Caenorhabditis elegans*, and promotes reprogramming of human somatic cells into pluripotent stem cells. Our findings outline a microRNA post-transcriptional regulatory network and establish a novel role for the miRNA precursor loop in the regulated production of mature Let-7.

Introduction:

---

2 Reprinted from Newman et al., RNA. 2008 Aug;14(8):1539-49 with permission from Cold Spring Harbor Press. This manuscript has been reformatted for the dissertation.

It is now generally appreciated that diverse forms of double stranded RNA (dsRNA) can act as triggers of RNA interference (RNAi) or related homology dependent gene silencing pathways (Chapman & Carrington, 2007). Among these triggers are microRNAs (miRNAs), which are single stranded but fold into stable stem-loop structures, providing the essential double stranded feature. miRNAs are encoded in the genomes of most metazoans and function in a post-transcriptional layer of gene regulation (see (Bartel, 2004) for a review).

The founding miRNA, *lin-4*, was discovered in *C. elegans* as a mutant that displayed heterochronic, or developmental timing defects (R. C. Lee et al., 1993; Wightman et al., 1993). The activity of this small RNA is mediated largely through repression of two well established target mRNAs, *lin-14* and *lin-28* (Ambros, 1989). A second miRNA, *let-7*, was later identified as a heterochronic mutant (Reinhart et al., 2000). Surprisingly, this miRNA has complete nucleotide conservation from *C. elegans* to humans, suggesting an ancient biological role (Pasquinelli et al., 2000). More recently, thousands of miRNAs have been identified across many phyla. While few validated mRNA targets have been assigned to these miRNAs, computational predictions suggest that each miRNA has tens to hundreds of targets, underscoring their immense potential for controlling gene expression (Bartel, 2004).

The initiation of miRNA dependent gene regulation is the transcription of a primary transcript, or pri-miRNA (see (V. N. Kim, 2005) for a review). This RNA is typically thousands of nucleotides long and is often capped, spliced, and poly-adenylated. The stem-loop structure is excised by the ribonuclease enzyme Drosha, liberating the precursor, or pre-miRNA. After export out of the nucleus, the precursor is further processed by the enzyme Dicer. The resultant siRNA-like molecule is loaded into the RNAi effector complex RISC, where it directs nucleolytic degradation and translational repression of target mRNAs.

While hundreds of miRNAs have been identified in the human and mouse genomes, the biological role of most miRNAs is unknown. However, numerous studies have linked miRNA function to the regulation of cell growth and differentiation (see (Esquela-Kerscher

& Slack, 2006) for a review). For example, the miRNA cluster miR-17-92 is highly expressed in a wide range of cancers, and ectopic expression of this candidate oncogene promotes cancer in several mouse models. This miRNA cluster is also highly expressed in early mouse development, and at least one study has confirmed its role in preventing the differentiation of progenitor cells (Y. Lu et al., 2007). In contrast to these "oncomiRs", a large number of miRNAs, including the Let-7 family, is depleted in cancer. Let-7 itself has been shown to target the oncogenes Ras, Myc, and HMGA2, and expression of this miRNA inversely correlates with disease severity (Esquela-Kerscher & Slack, 2006; Y. S. Lee & Dutta, 2007; Mayr, Hemann, & Bartel, 2007). Interestingly, global downregulation of all miRNAs promoted disease in a mouse model for lung carcinoma, suggesting that the overall tumor-suppressive functions of miRNAs are more important than oncogenic functions of, for example, miR-17-92 (Kumar, Lu, Mercer, Golub, & Jacks, 2007).

Results:

Regulated microRNA biogenesis:

The widespread alterations in miRNA expression in cancer prompted us to investigate the regulatory mechanisms responsible for their production. Our initial studies focused on the Let-7 family, where we previously demonstrated that induction of this miRNA family during mouse embryogenesis occurs at a post-transcriptional stage (Thomson et al., 2006). Since we observed a large amount of primary transcript in the absence of precursor and mature species, we predicted that a block was in place at the Drosha endonuclease processing step. We tested this directly using an established cell-free assay that reports Drosha activity (Y. Lee et al., 2002). We compared the processing efficiency of a regulated miRNA, Let-7g, with a miRNA that is readily processed in all known cell types, miR-17. We incubated radiolabeled primary transcripts for these miRNAs in nuclear extracts prepared from the mouse embryonal carcinoma cell line P19. We had previously shown that this cell line

exhibits regulated processing of Let-7, as it has abundant pri-Let-7 but no detectable mature Let-7 (Thomson et al., 2006). As a control we performed assays in extracts derived from Hela cells, which have abundant mature Let-7, thus are competent for Let-7 processing. Nuclear extracts derived from Hela cells were efficient at processing both Let-7g and miR-17, with approximately equal amounts of each precursor product being generated (**Figure 3.1**). While several RNA species are generated after the Drosha reaction, we were able to confirm the identity of the correct precursor reaction products by Northern blot hybridization of unlabeled products (Supplemental Figure 1).

While the control nuclear extract was competent for both Let-7g and miR-17 processing, extracts derived from undifferentiated P19 cells were inefficient at processing Let-7g (**Figure 3.1**). We calculated the ratio of Let-7g product to miR-17 product to generate a Let-7 processing efficiency. Undifferentiated P19 cells were ~10 fold less efficient at processing Let-7g compared to miR-17.

This data demonstrates that Drosha processing of Let-7 is less efficient in embryonic cells, but it does not discriminate between an activator present in the Hela extract or an inhibitor present in the P19 extract. Therefore we performed the same assay in a 1:1 mixed extract of P19 and Hela cells. The mixed extract assay yielded product ratios similar to the P19 extract alone, indicating that the regulatory factor is dominant in P19 cells and is therefore an inhibitor of Let-7 processing (**Figure 3.1**). To rule out the possibility that the regulatory event was due to the modification of Drosha itself, we immunoprecipitated the Drosha protein from undifferentiated P19 nuclear extracts. This purified protein was no longer subject to regulated processing as it was fully competent for Let-7 processing (**Figure 3.1**).

While our data demonstrates a regulatory point at the Drosha processing step, it does not preclude regulation at other steps in miRNA biogenesis. Specifically, regulation of Let-7 biogenesis at the Dicer processing step has been reported (Wulczyn et al., 2007).

microRNA regulatory sequences:

Since Drosha protein that had been immunopurified had lost the regulatory factor, the most logical conclusion of our data was that a Drosha Inhibitor, present in the P19 extract, was interacting directly with the Let-7 primary transcript. Alignment of the stem-loop precursors for the Let-7 family revealed highly conserved nucleotides within the loop region (**Figure 3.2A**). These sequence elements do not contribute to mRNA targeting; therefore, we reasoned they maintained conservation due to a regulatory role. We undertook several approaches to test whether the loop sequences of Let-7 were essential for regulated Drosha processing. Our first strategy was to build chimeric Let-7g primary transcripts that contained loop regions from other, unregulated miRNAs. We did detect weak processing in P19 cells, though it was unclear if these altered stem-loops were interacting properly with the processing machinery (data not shown). As an alternative we employed site-directed mutagenesis to alter conserved residues in the loop (see **Figure 3.2A** for sites). Importantly, these mutations did not affect folding of the stem-loop structure based on computational folding algorithms (mFOLD). Pri-Let-7g transcripts containing the SD1 or SD2 mutations were partially released from the Drosha processing block as evidenced by increased pri-Let-7g cleavage in P19 nuclear extracts (**Figure 3.2B**). The SD3 and SD4 mutations had no effect (data not shown). Mutation at both SD1 and SD2 sites further released the processing block. Importantly, these mutations did not impair general Drosha cleavage as processing was unaffected in Hela extracts.

While this data suggests the requirement for these loop nucleotides, it does not formally prove that the Drosha Inhibitor interacts with the loop. To establish this we designed competitor RNAs based on Let-7 stem loop sequences. If the Drosha Inhibitor was binding to the loop region, we reasoned that these competitor RNAs should divert the Inhibitor away from the Let-7 primary transcript, restoring processing in P19 cells. We tested competitors corresponding to all 11 Let-7 family members (not including miR-98). Competitors that matched Let-7 stem loop sequences restored Drosha processing of Let-7g without affecting production of miR-17 (data not shown). The most effective competitor

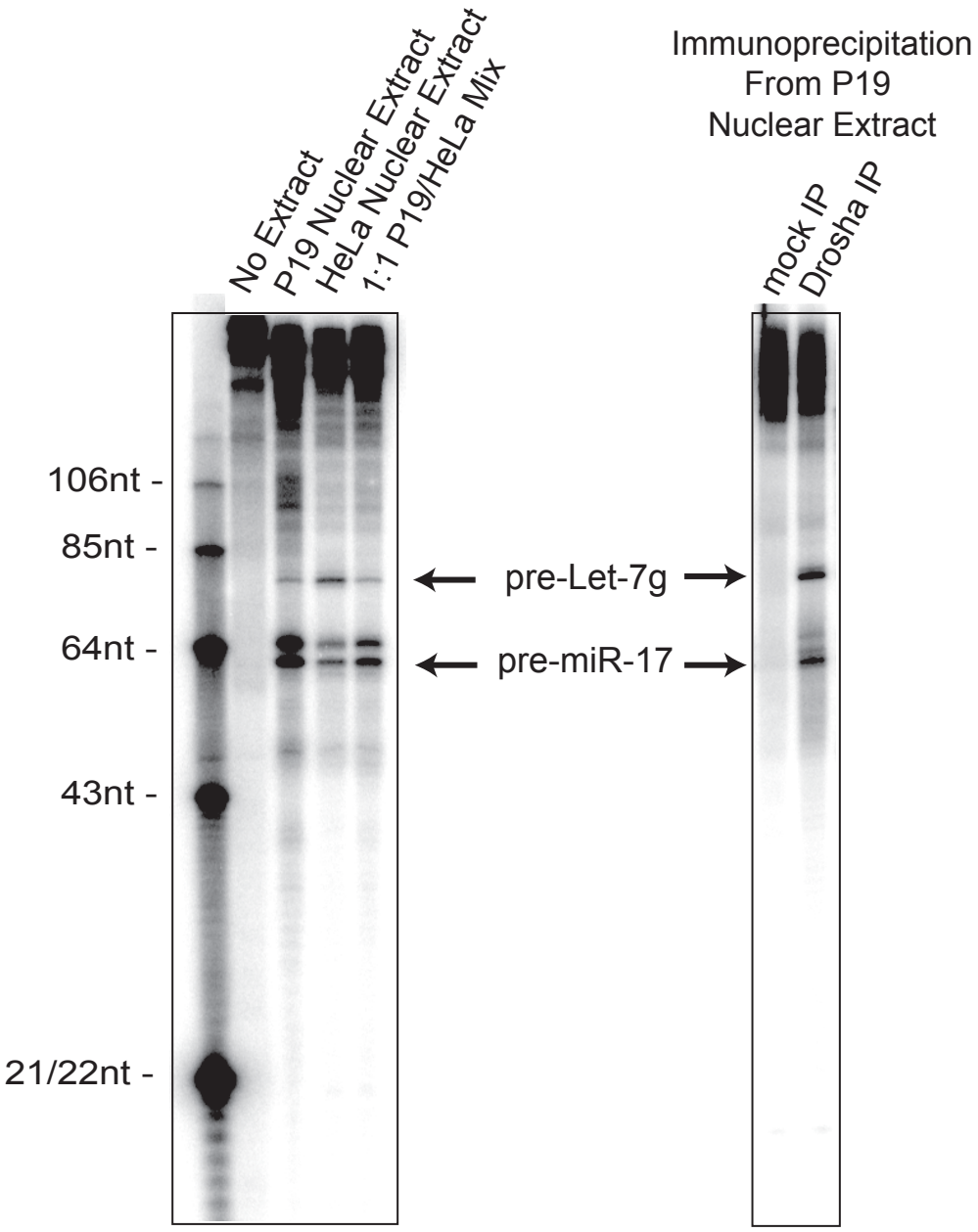


Figure 3.1. Embryonic cells contain a Drosha Inhibitor that specifically regulates Let-7. Radiolabeled pri-miRNA substrates corresponding to Let-7g and miR-17 were incubated in P19 or Hela nuclear extracts, as indicated. Drosha protein, or mock, was immunoprecipitated from P19 nuclear extracts with a polyclonal antibody. Immobilized protein was incubated with Let-7g and miR-17 pri-miRNA substrates, as indicated. Drosha endonuclease products were resolved on a denaturing polyacrylamide gel. The production of Let-7g and miR-17 precursors is indicated by arrows. A labeled RNA oligonucleotide ladder is shown for size reference.

was based on Let-7d. **Figure 3.2C** illustrates restoration of Let-7g processing in P19 nuclear extracts in the presence of Let-7d stem loop competitor. Competitors based on an unregulated miRNA, miR-17, had no effect on Let-7 processing (**Figure 3.2C**). Interestingly, an oligonucleotide corresponding to the loop alone was highly effective, suggesting that a stem loop structure is not essential for interaction with regulatory proteins (**Figure 3.2C**).

The Drosha Inhibitor:

We next investigated the nature of the Drosha Inhibitor. We designed an affinity purification strategy using the Let-7d loop as a capture reagent. We confirmed that a linear sequence, comprised of loop sequence only, fully 2'-O-methyl modified (for stability), and linked to a biotin moiety, was able to restore Drosha processing of Let-7g when used as a competitor (**Figure 3.2C**). We bound this capture probe to streptavidin-agarose and isolated binding proteins from P19 nuclear extracts. Proteins were identified by peptide mass fingerprinting. A large number of proteins specifically bound to the Let-7 loop probe, as shown in **Figure 3.3A**. Many of the proteins belong to the hnRNP family. This family of RNA binding proteins has diverse roles in RNA splicing and has been implicated in the regulation of miR-18a processing (Guil & Caceres, 2007). Interestingly, we also captured the RNA binding proteins Lin-28 and Lin-28B. This protein family was originally identified in *C. elegans* as a genetic mutant that exhibited abnormal cell lineage (Ambros & Horvitz, 1984). Family members are characterized as having a cold shock domain and two zinc-finger domains; a unique domain organization among all known RNA binding proteins (Moss, Lee, & Ambros, 1997a). The biochemical activity of *lin-28* is not well characterized, though it has been suggested to regulate translation of specific mRNAs (Polesskaya et al., 2007).

We confirmed the interaction between Lin-28 and the Let-7 loop by UV-crosslink analysis. Radiolabeled Let-7 stem loop RNA specifically crosslinked a protein in P19

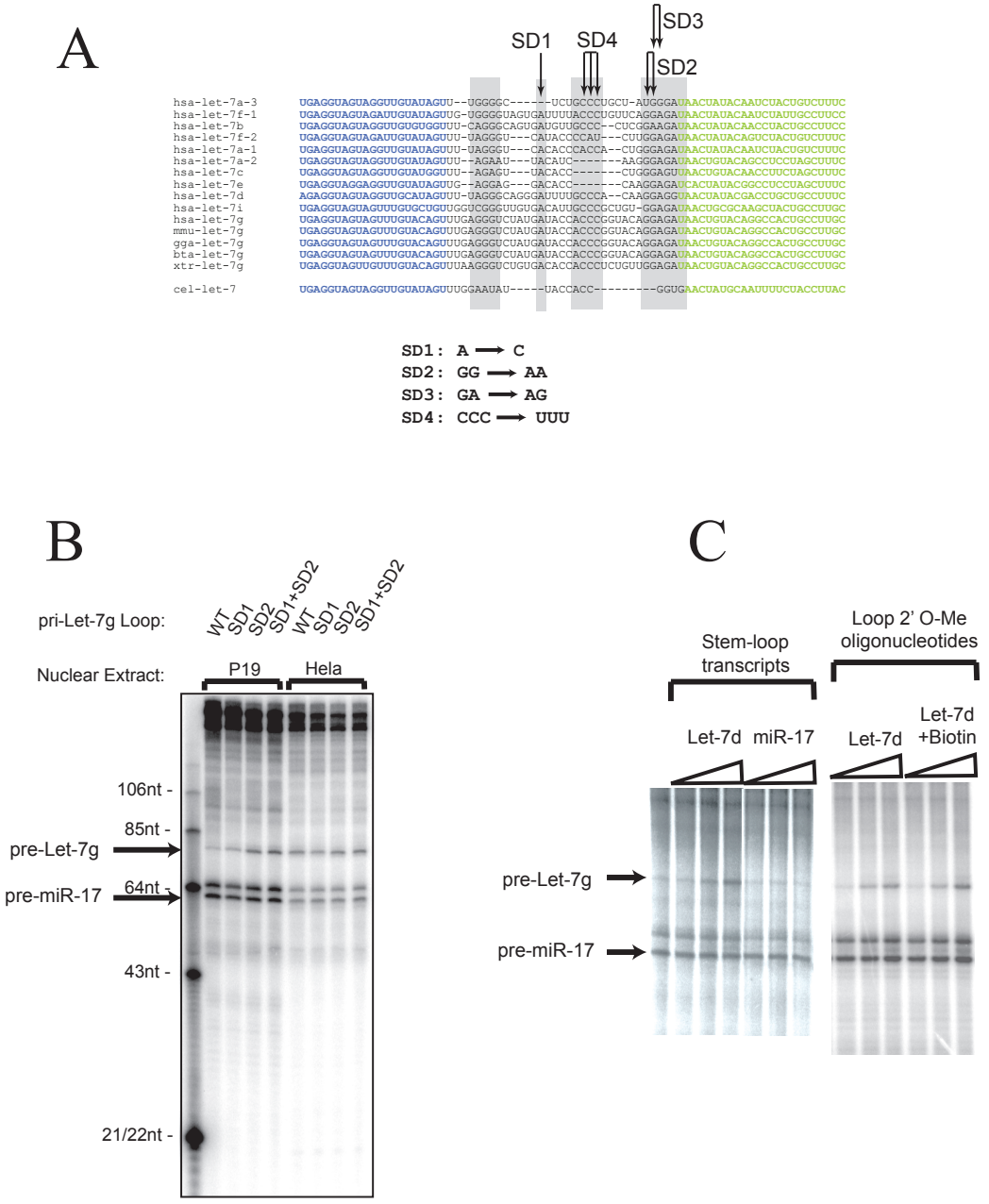


Figure 3.2. The loop region of Let-7 interacts with the Drosha Inhibitor. (A) The sequence alignment of Let-7 family members is shown. Blue text indicates the mature miRNA sequence. Green text indicates the complementary stem strand (not exactly the star strand). Gray boxes indicate regions of homology. Arrows indicate nucleotides that were mutagenized. Sequence changes are indicated below the alignment. (B) Wild type or mutant pri-Let-7g substrates were combined with the pri-miR-17 substrate and incubated in a P19 or Hela nuclear extract. Drosha products were resolved and are indicated. (C) Pri-miRNA substrates for Let-7g and miR-17 were combined and incubated in P19 nuclear extracts. Competitor RNA transcripts corresponding to the loop plus 12 nucleotides of each stem, or competitor 2'-O methyl oligonucleotides, were included at 10, 50, and 250 nM final concentration. In one case the oligonucleotide had a 3' Biotin moiety. The left lane had no competitor. Drosha products wereresolved on a denaturing polyacrylamide gel. Precursor products are indicated by arrows.

nuclear extracts that matched the size of Lin-28 (**Figure 3.3B**). We confirmed the identity of the protein by immunoprecipitation of crosslink reactions with an antibody to Lin-28. As evident in **Figure 3.3B**, one labeled protein of the correct size was recovered from Let-7 crosslink reactions, but no detectable protein was immunoprecipitated from the control reaction. Importantly, the Let-7/Lin-28 interaction was sensitive to the presence of excess competitor Let-7 stem loop RNA. Control competitors did not disrupt the Lin-28/Let-7 interaction.

While this data demonstrated interaction of Lin-28 with the Let-7 stem-loop, it did not demonstrate a regulatory role. To test this we introduced Lin-28 and Lin-28B into NIH-3T3 cells by retroviral transduction and measured steady state levels of mature miRNAs. Microarray analysis indicated that most miRNAs were unchanged, except for a cluster of miRNAs reduced in cells expressing either Lin-28 or Lin-28B (**Figure 3.4A**). As predicted, this cluster included most Let-7 family members. We confirmed miRNA expression changes by RT-PCR (**Figure 3.4B**). This method is more sensitive and provides greater discrimination among the highly related Let-7 family members. Let-7f, Let-7g, and Let-7i were repressed 240, 90, and 195 fold, respectively, with other family members repressed to a lesser degree. We measured steady state levels of unrelated miRNAs and found no significant change. Importantly, the reduction in Let-7g was at a post-transcriptional step, as the level of the primary transcript was not decreased to the same magnitude (2.8 fold). Northern blot analysis indicated no accumulation of Let-7g precursor, consistent with a block at the Drosha processing step (**Figure 3.4C**). While it is formally possible that Lin-28 promotes a block at the Dicer processing step, this would require an increase in precursor turnover at the same rate as precursor production, since no alteration in precursor steady state levels is detected.

To further confirm a role for Lin-28 in the repression of Let-7 expression, we performed RNAi knockdown of Lin-28 in P19 embryonic cells. Two siRNAs that reduced protein levels >90% released the processing block and led to increased mature Let-7 (**Figure**

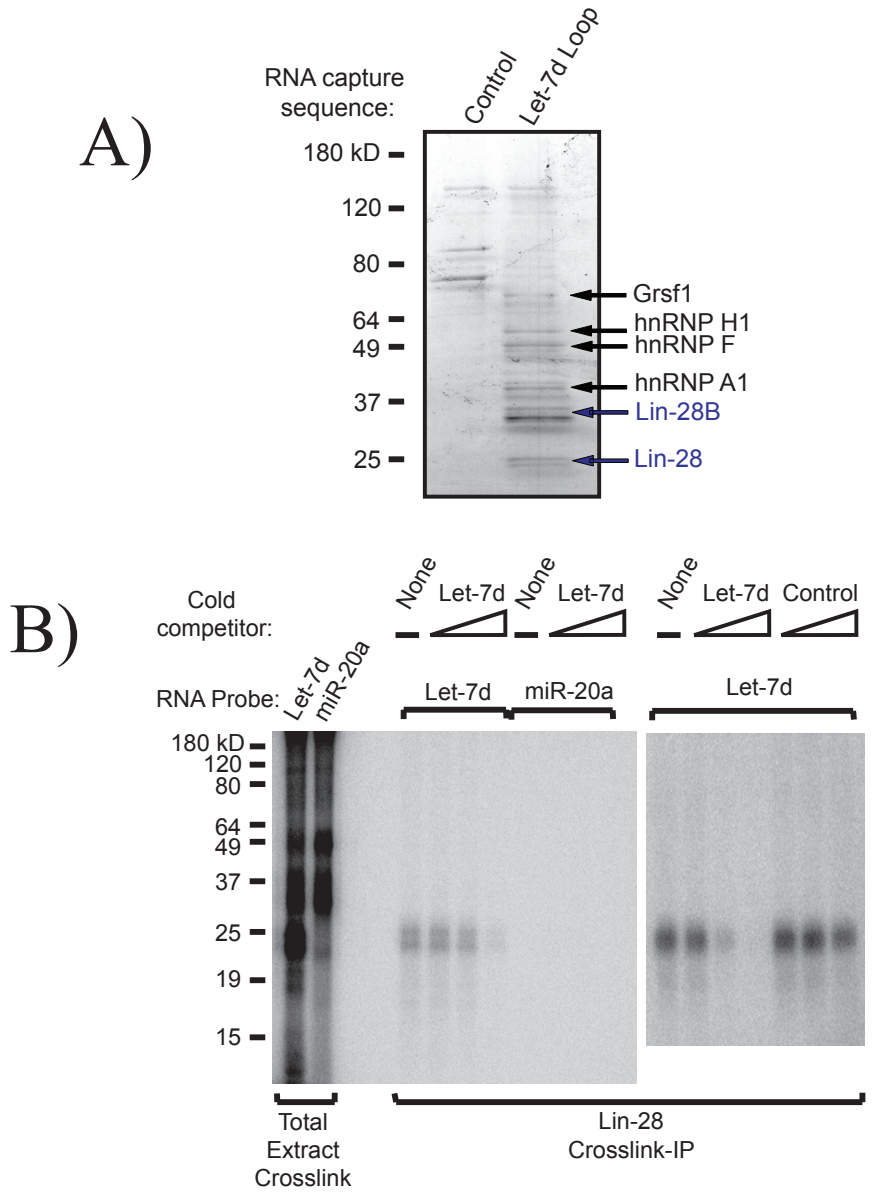


Figure 3.3. The RNA binding protein Lin-28 specifically binds to the Let-7 loop region. (A) Oligonucleotide capture probes corresponding to the Let-7d loop or a random (control) sequence, fully 2'-O-methyl modified, 3' Biotin linked, were bound to streptavidin agarose. Proteins were captured from P19 nuclear extracts, were resolved on a 4-20% polyacrylamide gel, and stained with coomassie blue. Proteins were isolated and identified by MALDI-TOF fingerprinting. (B) Radiolabeled RNA probes corresponding to the Let-7d stem-loop or miR-20a stem-loop were incubated with P19 nuclear extracts. Non-labeled Let-7d loop or control oligonucleotide competitors were included as indicated at 12.5, 125, or 1250 nM. Proteins were crosslinked to probes with UV light. Lin-28 was immunoprecipitated from crosslink reactions with a polyclonal antibody. Total extract (IP input) and immunoprecipitates were resolved, as indicated, on a polyacrylamide gel.

3.4D, Supplemental Figure 3). As above, Let-7 family members had differential response to siRNA knockdown, with Let-7g, Let-7i, Let-7a, and Let-7f the most affected.

Our final goal was to demonstrate the direct regulation of Let-7 processing by Lin-28. We generated recombinant Lin-28 and an unrelated RNA binding protein NF-45 in *E. coli* and introduced these purified proteins into our cell free Drosha assay (see Supplemental Figure 2 for recombinant proteins). As described above, Hela extracts will process Let-7g with similar efficiency as miR-17. As shown in Figure 5A, Lin-28 inhibited processing of Let-7g without affecting processing of miR-17. This was not a non-specific affect of RNA binding as the control protein (NF-45) had no effect of processing of either substrate. To further support our conclusions, we removed endogenous Lin-28 from the P19 embryonic cell extract by immunodepletion. This enabled processing of Let-7 (**Figure 3.5B**).

Discussion:

Our results outline a regulatory mechanism for the production of the Let-7 family of miRNAs. The primary transcripts for these miRNAs are uniformly expressed during embryonic development. The interaction of a Drosha Inhibitor with the loop structure of the pri-miRNAs leads to a block at the Drosha processing step. We have identified conserved nucleotides within the loop that are essential for the Drosha processing block. This is the first report of cis-regulatory elements within the Let-7 primary transcript. While our work focused on the loop region of Let-7, it is interesting that sequence conservation extends outside the stem-loop precursor. **Figure 3.6** illustrates that extensive vertebrate conservation is present in sequences flanking the precursor. This sequence may be important for splicing of the host gene Wdr82. However, the proximity to the stem-loop suggests these elements play a regulatory role in some aspect of miRNA function. In further support of this, Let-7 family members that are not intronic (for example Let-7i, as shown in **Figure 3.6**) have conserved sequence elements flanking the stem-loop.

We identify Lin-28 as a protein that interacts with the loop region of Let-7. This

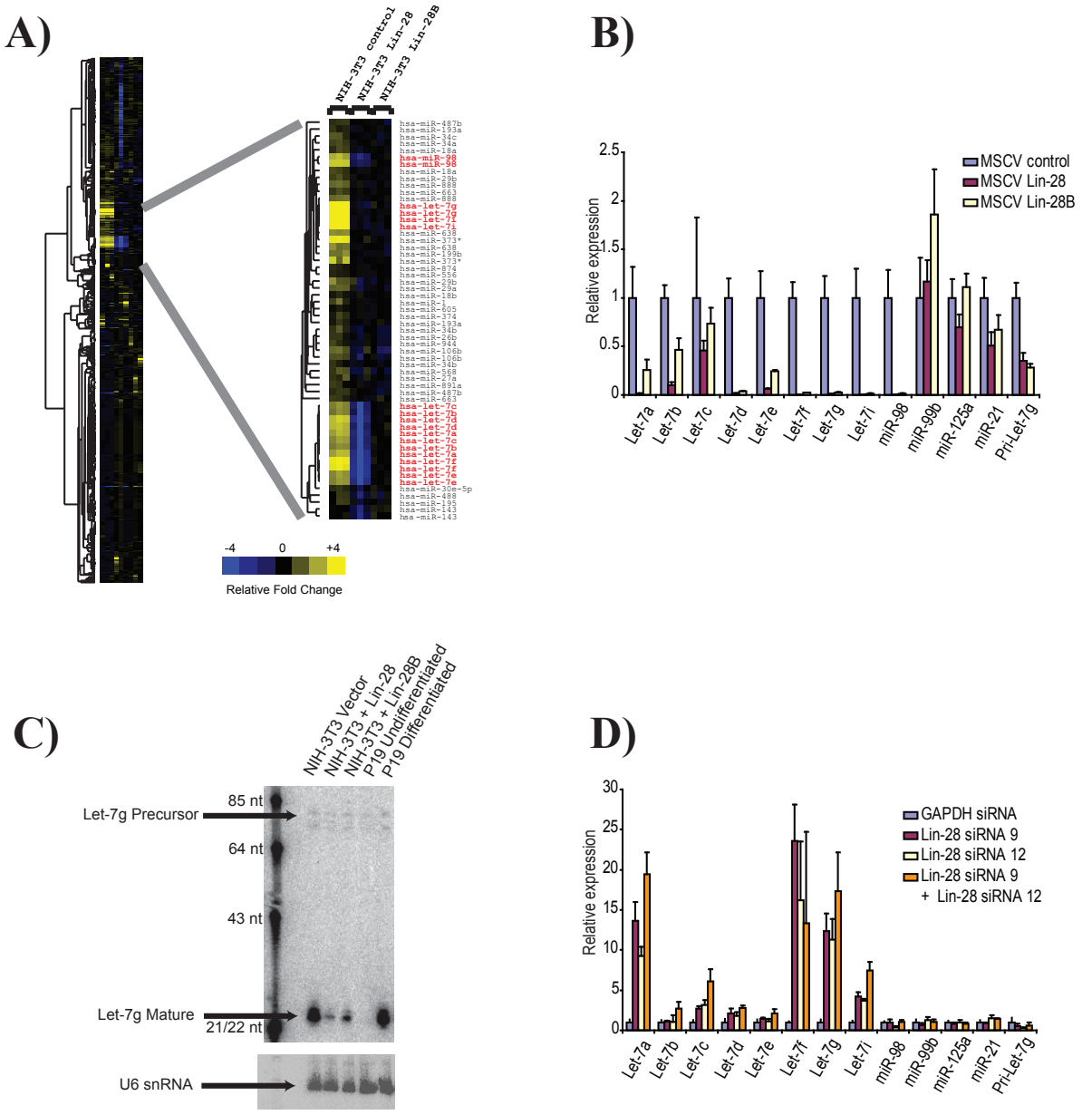


Figure 3.4. Lin-28 expression blocks production of Let-7. (A) NIH-3T3 cells were transduced with MSCV retroviral constructs that drive expression of Lin-28 or Lin-28B or control, as indicated. Steady state miRNA expression levels were quantitated using a custom microarray platform 10 days post-infection. Normalized measurements were hierarchically clustered and are plotted as a heat map. Yellow indicates high expression and blue low, relative to the mean. Let-7 family members are indicated in red font. (B) Steady state miRNA expression levels from NIH-3T3 cells expressing Lin-28, Lin-28B, or control were quantitated by real time RT-PCR. U6 snRNA was used as the reference. Expression of pri-Let-7g was also quantitated by real time RT-PCR. β2-microglobulin was used as the reference. (C) Let-7g precursor and mature species in NIH-3T3 cells expressing Lin-28, Lin-28B, or control were analyzed by northern blotting. (D) P19 cells were transfected with siRNAs targeting Lin-28. Two effective siRNAs were used alone or in combination. 5 days post-transfection mature-miRNA levels were measured by real time RT-PCR. U6 was used as a reference. Expression of pri-Let-7g was also quantitated by real time RT-PCR. β2-microglobulin was used as the reference.

interaction is specific for Let-7 family members has a dissociation constant in the high nanomolar range, based on competitor assays (**Figure 3.3B**) and fluorescent anisotropy binding assays (not shown). The interaction of Lin-28 with Let-7 inhibits Drosha processing. This RNA binding protein is highly expressed in early development and has been proposed as a marker for ES cells (Richards et al., 2004; D. H. Yang & Moss, 2003). Differentiation of P19 cells with retinoic acid leads to robust induction of Let-7; Lin-28 protein decays at the time point that Let-7 processing is enabled (Y. S. Lee et al., 2005; Wu & Belasco, 2005). We propose that a similar decay of this protein during embryonic development is what allows Drosha processing of Let-7 and eventual production of the mature species.

While it is clear that Let-7 processing is negatively regulated by Lin-28, is this the primary mode of regulation for the production of this miRNA family? We have previously reported that Let-7 processing efficiency increases >1000 fold during mammalian development (Thomson et al., 2004). Our overexpression studies, however, perturb Let-7 levels ~100 fold. It is probable that we are not achieving high enough expression levels to fully recapitulate embryonic cells, and western blot analysis confirms this (Supplemental Figure 4). Similarly, our siRNA knockdown studies do not fully eliminate Lin-28 expression and it is possible that remaining Lin-28 protein has some repressive function (Supplemental Figure 3). These points suggest we have identified the major regulatory point in Let-7 production, though we cannot rule out contributions at Dicer processing or at pri-Let-7 transcription, both of which have been reported (Chang et al., 2008; Wulczyn et al., 2007).

Recently, an independent study has demonstrated that Lin-28 selectively inhibits processing of Let-7 (Viswanathan, Daley, & Gregory, 2008). Similarly, this work suggested inhibition at the Drosha processing step, lending further support to our model proposed

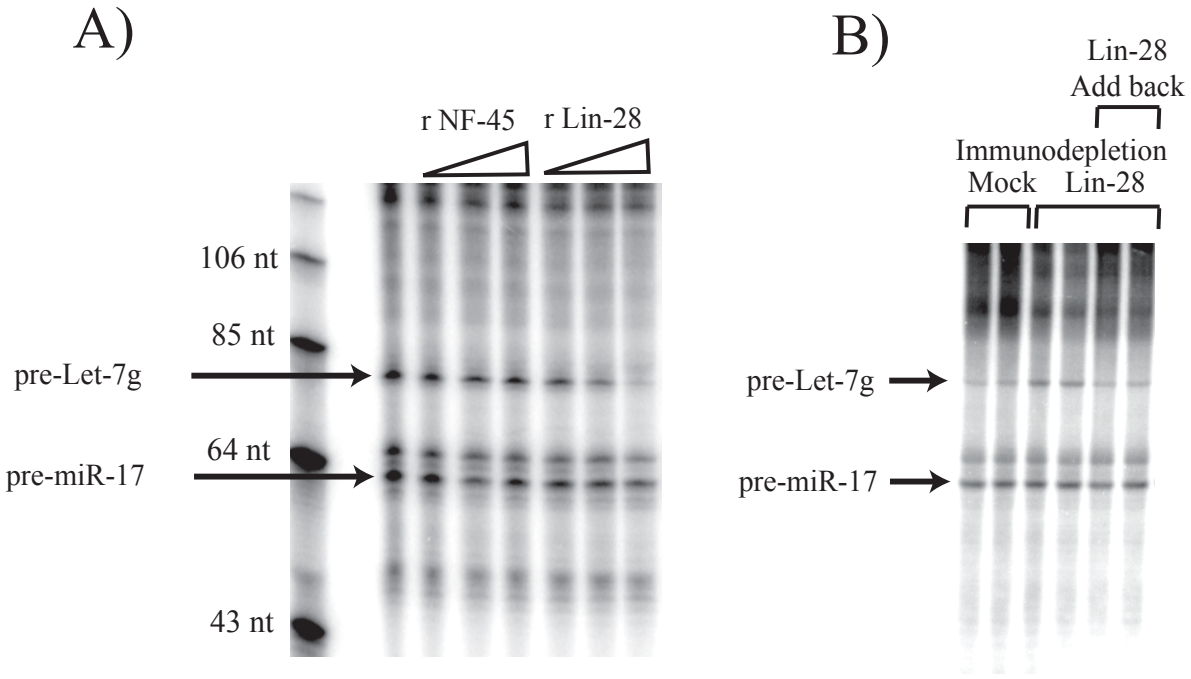


Figure 3.5. Lin-28 is necessary and sufficient for regulated Let-7 processing. (A) Pri-miRNA substrates for Let-7g and miR-17 were combined and incubated in Hela nuclear extracts. Purified, recombinant NF-45 (control) or Lin-28 were included at 2, 20, and 200 ng per reaction. Drosha products were resolved on a denaturing polyacrylamide gel. Precursor products are indicated by arrows. Recombinant protein was produced in E. coli. (B) Polyclonal Lin-28 antibody, or mock, was bound to protein A sepharose. P19 nuclear extracts were incubated with immobilized antibody. Resultant immunodepleted extracts were incubated with pri-miRNA substrates for Let-7g and miR-17. In right lanes, recombinant Lin-28 was added back to immunodepleted reactions at 100 ng/µl final concentration. Drosha products were resolved on a denaturing polyacrylamide gel. Precursor products are indicated by arrows. Recombinant protein was produced in HEK-293 cells.

herein.

The exact mechanism whereby Lin-28 inhibits Drosha processing is not known. Its high affinity binding to Let-7 may simply block access of Drosha or DGCR8 to the pri-miRNA. This binding event may promote turnover of the pri-miRNA, as there is no accumulation of this RNA species under conditions where processing is inhibited. Alternatively, pri-miRNAs may have a short half life under all conditions and therefore are degraded if not rapidly processed by the Drosha machinery. A final possibility is that the Lin-28/Let-7 complex is exported out of the nucleus and shuttled to P-bodies for degradation. Lin-28 is known to shuttle between the nucleus and cytoplasm, and has been localized to P-bodies, lending support to this hypothesis (Balzer & Moss, 2007b).

The regulation and developmental expression of the mammalian Lin-28 closely parallels that of the *C. elegans* ortholog, suggesting it performs a similar function in that organism. Expression of *let-7* during *C. elegans* larval development is known to be regulated at transcription (Johnson, Lin, & Slack, 2003b). In *lin-28* mutants, however, the production of mature *let-7* is accelerated by one larval stage. Similarly, in *lin-4* mutants, which have elevated expression of *lin-28*, the production of mature *let-7* is delayed (Johnson, Lin, & Slack, 2003a). While it is not known if these effects are a result of direct interaction between the *lin-28* protein and the *let-7* primary transcript, it is possible that the mechanism we report here is conserved from *C. elegans* to mammals.

The role of Lin-28 in mammalian development is not known. However, a combination of Lin-28, Nanog, Oct-3/4, and Sox2 are sufficient to reprogram human somatic cells to pluripotent stem cells (Yu et al., 2007). The role of Nanog, Oct-3/4, and Sox2 in maintaining pluripotency is well established (Pan & Thomson, 2007).
The additional requirement for Lin-28 in these studies suggests that restoration of the Let-7 processing block is an essential step for reprogramming stem cells. These findings strongly implicate miRNA expression patterns as important determinants of stem cell fate.

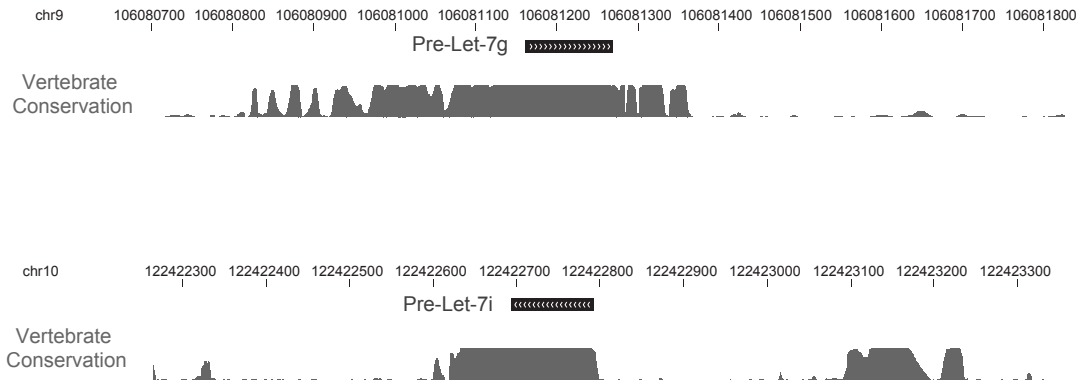


Figure 3.6. Extensive sequence conservation of pri-Let-7. The genomic locus of the mouse pre-Let-7g and pre-Let-7i stem-loop sequence are shown, with 600 nucleotides flanking on each side included. The conservation plot of 30 vertebrate species is shown. Data was obtained from the UCSC genome browser July 2007 release.

Materials and methods:

Pri-microRNA substrates:

Substrate constructs (stem-loop with ~10nts of ssRNA flanking region) were created by the polymerase chain reaction (PCR) using oligonucleotides as the DNA templates:

pri-Let-7g: a5'-TGCCTGATTCCAGGCTGAGGTAGTCGTTTGTACAGTTTGAGGGTCT ATGATACCACCCGGTACAGGAGATAACTGTACAGGCAACTGCCTTGCCAGGAACA GCGCG-3'

pri-miR-17: -;v5'-GTCAGAATAATGTCAAAGTGCTTACAGTGCAGGTAGTGATATGTG CATCTACTGCAGTGAAGGCACTTGTAGCATTATGGTGAC-3'

Forward and reverse primers added ssRNA flanks as well as 5' Bgl-II and 3' Xho-I restriction sites.

Primers for pri-Let-7g:

Forward 5'-GTCAAGATCTCGTTTCCTTTTGCCTGATTCCAGGCTGA-3'

Reverse 5'-GTCACTCGAGGGCAGCTGGCGCGCTGTTCCTGGC-3'

Primers for pri-miR-17:

Forward - 5'-GTCAAGATCTATTGTGACCAGTCAGAATAATGTCAAAGTGCTTACAG -3'

Reverse 5'-GTCACTCGAGCGAGGCAGCTGTCACCATAATGCTACAAGTGCCT-3'

Resulting PCR products were digested and cloned into a MSCV-splice-donor/splice-acceptor vector (SDSA3.0; J.M. Thomson and S.M. Hammond, unpublished results) based on the MSCV-puro vector (Clontech).

Pri-microRNA transcription templates were created from the pri-microRNA

constructs described above; PCR was employed to add on a 5' T7 promoter (5'-TCGTAATACGACTCACTATAGGGTCCGCTAGCCTAGCTACTACCA-3' and 5'-ATAAGTATGATATTGTCAAGGAAACCC-3'). Radiolabeled pri-microRNAs were created using T7 RNA polymerase (NEB) in accordance with the manufacturer's instructions. The resulting pri-microRNA substrates were (pri-microRNA sequences are in bold type; flanking vector sequences are italicized):

Pri-Let-7g:

TCCGCTAGCCTAGCTACTACCAGGTGAGTGGAGATCTCGTTTCCTTTTGCCTGAT-TCCAGGCTGAGGTAGTCGTTTGTACAGTTTGAGGGTCTATGATACCACCCGGTA-CAGGAGATAACTGTACAGGCAACTGCCTTGCCAGGAACAGCGCGCCAGCTGC-CCTCGAGGTTTAAACGAATTCAGGGTTTCCTTGACAATATCATACTTAT

Pri-miR-17:

TCCGCTAGCCTAGCTACTACCAGGTGAGTGGAGATCTATTGTGACCAGTCA-GAATAATGTCAAAGTGCTTACAGTGCAGGTAGTGATATGTGCATCTACTGCAGT-GAAGGCACTTGTAGCATTATGGTGACAGCTGCCTCGCTCGAGGTTTAAACGAAT-TCAGGGTTTCCTTGACAATATCATACTTAT.

Cell culture and nuclear extract preparation:

P19 cells were differentiated for 10 days in retinoic acid, as described(Thomson et al., 2006). Undifferentiated and differentiated P19 cells, and Hela S3 nuclear extracts were prepared as previously described(Dignam, 1990). In brief, cell pellets, washed once in phosphate buffered saline (PBS), were resuspended in 2.5 volumes of Buffer A [10mM HEPES, pH 7.9, 10mM KCl, 1.5mM MgCl2, 0.2mM EDTA, 1mM dithiotreitol (DTT), 1mM phe-

nylmethylsulfonyl fluoride (PMSF), 20% glycerol and Complete protease inhibitors (CPI, Roche)] and incubated on ice for 10 minutes. Cells were then lysed using 5-10 passes in a dounce homogenizer. Nuclei were centrifuged at 1200rpm for 5 minutes, followed by centrifugation 15,000 x g for 15 minutes and the resulting supernatant was discarded. Pelleted nuclei were extracted in buffer C [20mM HEPES, pH 7.9, 420mM NaCl, 1.5mM MgCl2, 0.2mM EDTA, 1mM DTT, 1mM PMSF, 20% glycerol and CPI] at 4oC for 30 minutes (3ml of buffer C was used per 1x109 cells). Nuclear debris was then pelleted by centrifugation at 15,000rpm for 30 minutes and the resulting nuclear extract was dialyzed for 5 hours against >50 volumes of buffer D [20mM HEPES, pH 7.9, 100mM KCl, 1.5mM MgCl2, 0.2mM EDTA, 1mM DTT, 1mM PMSF and 20% glycerol).

Cell free Drosha processing assay:

Drosha assay reactions were carried out as previously described with some modifications (Lee et al., 2002). 1x105cpm each of pri-Let-7g and pri-miR-17 transcripts were combined and incubated with 100mg of nuclear extract (where indicated, either 100mg P19 extract, 100mg of differentiated P19 extract, 100mg of Hela extract or 50mg P19 + 50mg HeLa extract) + 0.1ml RNasin (Promega) for 30 minutes at 37oC. RNA was isolated using Trizol and precipitated in cold isopropanol. RNA pellets were washed once in 70% ethanol and resuspended in 15ml of formamide loading buffer. Processed pri-microRNA products were resolved on a 12.5% acrylamide-8M urea sequencing gel and visualized by autoradiography using a Storm phosphoimager (GE Life Sciences); pre-microRNA band intensities were quantitated using Imagequant software (GE Life Sciences). A molecular weight ladder of radiolabeled, concatamerized RNA oligos was used to determine the size of the pre-microRNA products.

Drosha immunoprecipitation:

For Drosha assays using immunoprecipitated Drosha, we first generated rabbit polyclonal anti-Drosha antibody using the peptide sequence CPEEAEDIKK. 250µl of P19

nuclear extract was precleared with 25µl protein-A sepharose and incubated either with 40µl of affinity-purified Drosha antibody and 45µl of packed protein-A sepharose or with protein-A sepharose alone. Igepal CA-630 (Sigma) was added to 0.5% v/v and immunoprecipitation reactions were incubated for 1 hour at 4oC. Beads were then washed three times in buffer D + 0.5% Igepal. Individual Drosha reactions were carried out as described above using 10µl of immobilized protein from each immunoprecipitation reaction, pri-miRNA substrates individually or combined, and buffer D and 1µl RNasin.

Drosha assay northern blot:

To identify pre-microRNA products, Drosha assays were carried out as described above except that 0.15mg of unlabeled pri-microRNA was individually incubated in Hela extract. Separated RNA products were then transferred onto Nylon (Amersham) which was then UV cross-linked and hybridized to locked nucleic acid (LNA) probes (Integrated DNA Technologies) that were labeled on the 5' end with T4 polynucleotide kinase (NEB). Hybridizations occurred overnight at 65oC, were washed three times at 65oC in 0.4xSSC + 0.2%SDS and cross-reacted bands were visualized by autoradiography.

Northern probe sequences (lower case denotes LNA; upper case denotes DNA):

Let-7g: 5'- ACTgTaCaAaCgAcTaCcTcA -3'

miR-17-3p: 5'-ACAaGtGcCtTcAcTgCaGt -3'

Stem-loop RNA competitor assay:

Drosha reactions were pre-incubated on ice for 30 minutes with unlabeled microRNA stemloop RNAs added to final concentrations of 12.5nM, 125nM and 1250nM, respectively. The competitor RNAs were either 2'-O-methlyated RNA oligos (Dharmacon), or were in vitro transcribed from long double stranded DNAs (dsDNAs) encoding a T7 promoter, the upper 10 basepairs of stem and the loop of the indicated pri-microRNA. The dsDNAs were created by annealing a T7 adapter primer to a longer microRNA primer, followed by a brief

PCR extension step with Pfu DNA polymerase (Promega) (95oC, 5 minutes, 59oC, 1 minute, 72oC, 10 minutes for 1 cycle).

T7 adapter:

5’- gatgTAATACGACTCACTATAGGG-3’

>hsa-let-7d

5’-caggtcgtatagttacctccttgtgggcaaaatccctgccctaaaactatgcaacccccctatagtgagtcgtattaCATC -3’

>hsa-let-7a-2

5’-gaggctgtacagttatctcccttgatgtaattctaaactatacaacccccctatagtgagtcgtattaCATC -3’

>hsa-let-7f-1

5’-agattgtatagttatctcctgaacagggtaaaatcactaccccacaactatacaatcccctatagtgagtcgtattaCATC -3’

>hsa-let-7c

5’-aaggttgtacagttaactcccagggtgtaactctaaaccatacaacccccctatagtgagtcgtattaCATC -3’

>hsa-let-7a-3

5’-tagattgtatagttatcccatagcagggcagagccccaaactatacaacccccctatagtgagtcgtattaCATC -3’

>hsa-let-7i

5’-agcttgcgcagttatctccacagcgggcaatgtcacaacccgaccaacagcacaaaccccctatagtgagtcgtattaCATC -3’

>hsa-let-7f-2

5’tagactgtatagttatctccaagatggggtatgaccctaaaactatacaatcccctatagtgagtcgtattaCATC -3’

>hsa-let-7g

5’ggcctgtacagttatctcctgtaccgggtggtatcatagaccctcaaactgtacaaacccctatagtgagtcgtattaCATC -3’

>hsa-let-7a-1

5’tagattgtatagttatctcccagtggtgggtgtgaccctaaaactatacaacccccctatagtgagtcgtattaCATC -3’

>hsa-let-7e

5’gaggccgtatagtgatctccttgggtgtcctcctcaactatacaacccccctatagtgagtcgtattaCATC -3’

>hsa-let-7b

5’aggttgtatagttatcttccgaggggcaacatcactgccctgaaaccacacaaccccctatagtgagtcgtattaCATC -3’

>hsa-mir-17

5’tcactgcagtagatgcacatatcactacctgcactgtccctatagtgagtcgtattaCATC -3’

>hsa-mir-20a

5’ctcataatgcagtagataactaaacactacctgcactatccctatagtgagtcgtattaCATC -3’

Lin28 immunodepletion Drosha assay:

Protein A-sepharose beads were washed 3 times in “Hi DTT” immunoprecipitation (IP) buffer [20mM HEPES, pH 7.6, 2mM MgCl2, 150mM NaCl, 10mM DTT, 1mM PMSF, 0.5% NP-40] followed by 5 washes in IP buffer [20mM HEPES, pH 7.6, 2mM MgCl2, 150mM NaCl, 1mM DTT, 1mM PMSF, 0.5% NP-40]. 15mL of anti-Lin-28 antibody (Ab-cam) was prebound to 30mL protein A sepharose beads (Sigma) for 1 hour at 25oC in IP buffer, followed by extensive washing with buffer D + 0.5% NP-40. 100mL of P19 nucelar lysate (5mg/ml) was diluted 2-fold in buffer D + 2ml of RNasin and rotated overnight at 4oC with either the antibody-beads mixture or beads alone as a mock.

Recombinant protein competitor assay:

Drosha assays were carried out as described above except that, where indicated, 2ng, 20ng and 200ng of recombinant protein was added to Hela extract/pri-microRNA mixtures before the 37oC incubation step.

RNA affinity pulldown:

30ml of 100mM Let-7d Loop biotinylated 2’-O-methylated RNA oligonucleotide, (5’UUmAmGmGmGmCmAmGmGmGmAmUmUmUmUmGmCmCmCmAmCmAmAmG-mGmAmGmGmU-18s-Bi 3’, Dharmacon) or a non-specific control oligo (5’Bi-18S-18S-mAmUmAmAmGmUmAmUmGmAmUmAmUmUmGmUmC-3’, Dharmacon) were bound to 150ml of streptdavidin-agarose beads (Fluka) for 1 hour at 37oC in “high salt” buffer D

(+1M KCl). Bead-oligo mixtures were washed 3 times in buffer D+T. 15mg of P19 cell nuclear extract (equivalent to 1x109 cells) was diluted up to 20ml in buffer D+T, pre-cleared with streptavidin agarose beads and incubated with bead-oligo mixtures at 25oC for 3 hours. Beads were then washed 6 times in Buffer D+T and bound proteins were eluted using SDS protein loading buffer + 50mM DTT. Proteins were separated on a 4-20% pre-cast polyacrylamide gel (Jule Biotechnologies) and stained with Coomassie brilliant blue G-250 (Bio-Rad); appropriate bands were excised and submitted to the UNC-Duke Michael Hooker Proteomics Facility for tryptic digestion and MALDI-TOF footprinting.

UV crosslinking:

Cross-linking experiments were carried out essentially as previously described (Myer et al., 1997). In brief, 50mg of P19 nuclear lysate, 3mg of yeast tRNA and 5x105 cpm of hsa-Let-7d or hsa-miR-20a loop probes (same competitor RNA sequences as in the "stem-loop RNA competitor assay" above) were combined and adjusted to a volume of 14ml with buffer D [20mM HEPES, pH 7.9, 20% glycerol, 100mM KCl, 1.5mM MgCl2, 0.2mM EDTA, 1mM DTT, 1mM PMSF]. Reaction mixtures were incubated for 30 minutes at 25oC, followed by irradiation with 1 Joule of UV light. Each sample was then incubated with 10mg RNase A for 30 minutes at 37oC. The samples were then either prepared for immunoprecipitation (see below) or boiled in SDS protein loading buffer for 5 minutes ("input" samples). Where indicated, 2'-O-methyl oligonucleotides were added to the final concentration of 12.5nM, 125nM or 1250nM to the reaction mixture and pre-incubated at 25oC for 30 minutes prior to the addition of the labeled probe.

2'-O-methyl competitor sequences:

Let-7d loop 2'-O-methyl: 5'mUmUmAmGmGmGmCmAmGmGmGmAmUmUmUmUmG-mCmCmCmAmCmAmAmGmGmAmGmGmU3'

Non-specific 2'-O-methyl:

5’mAmCmCmAmAmCmAmGmGmCmCmGmGmGmAmCmAmAmGmUmGmCmAmmAU-mAmAmC3’

Crosslinking Immunoprecipitations:

RNase A-treated samples were boiled in TSD buffer [50mM Tris-Cl, pH 7.5, 1% SDS, 5mM DTT] for 10 minutes and diluted 10-fold in TNN buffer [50mM Tris-Cl, pH 7.5, 250mM NaCl, 5mM EDTA, 0.5% NP-40 and complete protease inhibitor (Roche)]. Diluted lysates were pre-cleared with protein-A sepharose beads (Sigma) followed by immunoprecipitation overnight at 4oC with anti-Lin28 antibody (Abcam) and protein A-beads or with protein-A beads alone (“mock” immunoprecipitation). Protein-bound beads were washed extensively with cold TNN buffer; proteins were eluted by boiling SDS protein loading buffer and separated on a 12% SDS-PAGE gel which was subsequently fixed and dried down. Radiolabeled proteins were visualized by autoradiography.

Recombinant proteins:

cDNAs were amplified from full-length ESTs coding for NFAT-45 and Lin-28 as well as a partial EST for Lin28B (Open Biosystems) using the primers

Lin-28:
5’- tacgaattcACCATGGACTACAAAGACGATGACGACAAGGGCTCGGTGTCCAACCA-3’
5’-catgcggccgcTCAATTCTGGGCTTCTGGG-3’

Lin28B:
5’- AGCCAGAAAAACTGCCCGGGCTGGCAGAGGACGAACCCCAGGTTCTGCATG-GC-3’,
5’- TGACGACAAGGCCGAAGGCGGGGCAAGCAAAGGTGAAGAGCCAGAAAAACT-GCCCG-3’

5’-tacgaattcACCATGGACTACAAAGACGATGACGACAAGGCCGAAG-3’ and 5’- catgcggc-cgcCTAAGTCTTTTTCCGTTTCTGAATCA -3’were used to rebuild the 5’ end of the Lin28B gene by overlap extension PCR.

NFAT 45:

5’- tacgaattcACCATGGACTACAAAGACGATGACGACAAGAGGGGTGACAGAGGAC-3’

5’- catgcggccgcTCACTCCTGAGTCTCCATGC-3’

Flag-tagged cDNAs were cloned into pcDNA 3.0 using 5’ EcoRI and 3’ NotI restriction sites. cDNAs were subsequently subcloned into pMSCV-puro-IRES-GFP(He et al., 2005). For over-expression studies, NIH-3T3 cells were transduced using virus generated with the LinXE ecotropic packaging line. Cells were selected with puromycin. 10 days post-infection total RNA and protein samples were isolated using Trizol and SDS loading buffer, respectively. microRNA expression levels were examined by quantitative real-time PCR (qRT-PCR) as well as by microRNA microarray as previously reported(Thomson et al., 2004; Thomson et al., 2006). Protein over-expression was verified by western blot with a Lin-28 antibody (Abcam).

Purification of His-tagged Lin-28 and NF-45:

The following primers were used to amplify mouse coding sequences for NF-45 and Lin-28; 6xHis-NF45 5’-ggccatcatcatcatcacaggggtgacagagga-3’; BsmFI_6xHis_NF45_push 5’- cagtgggacgctgtctcaccatgggccatcatcatc-3’; XhoI_NF45 5’- catgctcgagtcactcctgagtctccatgc-3’; 6x_His_Lin28 5’-ggccatcatcatcatcacggctcggtgtccaac-3’; NcoI_6xHis_Lin28_push 5’- atac-catgggccatcatcatcatcac-3’; Lin28_XhoI 5’- catctcgagtcaattctgggcttctggg-3’. PCR products were inserted into the pET28b vector (Novagen) after NcoI/XhoI (vector; Lin28) or BsmFI/XhoI (NF45) digestion. BL21 (DE3) pRIL E. coli (Stratagene) served as the host for protein expres-sion. Briefly, positive clones were grown in 1 liter of LB to a density of 0.6 OD600 at which time recombinant protein was induced by the addition of 0.25mM IPTG and allowed to grow

for an additional 3 hours. Cells were harvested by centrifugation and cell pellets suspended in 30 ml Talon resin equilibration/wash buffer 50mM sodium phosphate, pH 7.0; 300mM NaCl (Clontech). Cells were lysed by sonication and cleared by centrifugation. The cell lysate was incubated with 5 ml of equilibrated Talon resin beads and allowed to bind at 40C for 1 hour. The resin was batch washed 3 x 20 minutes in 30 ml of equilibration/wash buffer and applied to a disposable column. The protein was eluted along a 50 ml gradient of equilibration buffer containing 0-200mM imidazole and fractions collected. Pure fractions were combined and dialysed in buffer D; 20mM Hepes, pH 7.9; 100mM KCl; 0.2mM EDTA; 1.5mM MgCl2; 1mM PMSF; 1mM DTT; and 20% Glycerol. Aliquots were stored at -800C.

Purification of Flag-tagged recombinant Lin-28:

Twenty 10cm plates of HEK293 cells were transiently transfected with either Flag-tagged NFAT-45, Lin28 or Lin28B over-expression constructs using Fugene 6 (Roche) according to the manufacturer's instructions. After 48 hours, cells were harvested by scraping in cold PBS, centrifuged at 1200rpm for 5 minutes and lysed in IP buffer. Lysates were pre-cleared with 50ml of protein A beads for 1 hour followed by immunoprecipitation with 50ml of M2 flag agarose beads (Sigma) overnight at 4oC. Beads were then pelleted and washed 10 times in cold buffer D followed by elution with Flag peptide (Sigma) at 400 mg/ml in buffer D. Purity of eluted proteins was verified by SDS-PAGE followed by coomassie blue staining of the gel.

Lin-28 knockdown/overexpression studies:

P19 cells were seeded in six-well plates at ~40% confluency and 2 hours later were transfected with siRNAs to either GAPDH or Lin-28 (Dharmacon) using Lipofectamine 2000 (Invitrogen) as instructed by the manufacturer. Total RNA and protein was isolated approximately 72 hours later using Trizol and SDS loading buffer, respectively. microRNA expression levels were examined by quantitative real-time PCR (qRT-PCR)(Thomson et al., 2006).

siRNA sequences:

si#2: 5’ ggagacaggugcuacaacuuu3’

si #9: 5’ ugacguaucuugugcguuuuu 3’

si#12: 5’ aaaugugucucacggguuuuu 3’

Lin-28 and Lin-28B was ectopically expressed using MSCV retroviral expression constructs. MSCV retrovirus was transduced into NIH-3T3 cells and cells were selected with puromycin. 10 days post-infection cells were harvested and miRNA expression levels were characterized by microarray analysis and real time RT-PCR essentially as described (Thomson et al., 2004; Thomson et al., 2006). Northern blot analysis was performed as described, using a radiolabeled LNA probe for Let-7g (Thomson et al., 2006).

Chapter 4:
An unprecedented view of pre-microRNA uridylation and degradation in murine embryonic carcinoma cells.

Summary:

MicroRNAs (miRNAs) play crucial roles in orchestrating gene-expression patterns during development as well as in the onset of many diseases. However, before these small non-coding RNAs can bind to cognate sequences in target mRNAs, they must first be processed from larger RNA precursors in a well-characterized biogenesis pathway. Surprisingly, in early development, as well as in poorly differentiated tumors, many mature miRNAs are largely absent due to a post-transcriptional block in their biogenesis. Recently, it was shown that Lin28, an embryonic cell-specific RNA binding protein, blocks the processing of the let-7 family of tumor suppressor miRNAs. While the biochemical nature of the Lin28 block isn't entirely clear, recent intriguing evidence suggests that Lin28 recruits the terminal uridyl transferase, TUT4 to let-7 precursor miRNAs (pre-miRNAs) leading to oligo-uridylation and, presumably, degradation of the pre-miRNA. This observation raises the possibility that the global repression of mature miRNAs occurs, in part, at the level of pre-miRNA uridylation. However, current technical limitations in pre-miRNA detection have made this hypothesis difficult, if not impossible to address. Here, we describe the development of a method termed "RM-seq" (for Reverse-transcription and Multiplexed PCR in conjunction with deep sequencing) to assess the abundance, uridylation and degradation of over 40 pre-miRNAs simultaneously in P19 embryonic carcinoma (EC) cells. Using this technique, we demonstrate that many pre-miRNAs, in addition to those of the let-7 family, are 3' oligo-uridylated, possibly by a mechanism independent of the Lin28/TUT4 pathway.

Furthermore, we show that the degradation of uridylated pre-miRNAs is governed, in part, by the 3'-5' exosome pathway.

Introduction:

microRNAs (miRNAs) are small non-coding RNAs with important functions in embryonic development as well as in the pathogenesis of various diseases including cancer (Bushati & Cohen, 2007). It is now clear that the precise control of miRNA abundance relies not only on the transcription of miRNA genes but also depends on the post-transcriptional regulation (Newman & Hammond, 2010). Many steps in the miRNA biogenesis pathway after transcription are regulated but, to date, the most intensely studied phenomena of regulated miRNA processing center on the RNaseIII enzymes Drosha and Dicer and the modulation of their accessibility and activity towards particular miRNA precursors. In one instance, effector proteins of signal transduction pathways either increase or decrease Drosha processing of target pri-miRNAs, either by binding directly to pri-miRNA sequences (Davis et al., 2010) or by modulating the interaction of Drosha with RNA-helicases, p68 and/or p72 (Suzuki et al., 2009; Yamagata et al., 2009). Another phenomenon is the binding of regulatory proteins to conserved sequences in the loop region of pri- and pre-miRNAs, leading either to activation or repression of Drosha and Dicer processing. The best-studied example of this second kind of regulation is the early embryonic processing block of let-7 miRNA precursors by the embryonic-stem (ES) cell-specific RNA binding protein Lin28. Lin28 recognizes the conserved GGAG motif in the 3' region of the loop (Heo et al., 2009). While Lin28 was originally shown to block Drosha processing of pri-let-7 *in vitro* (Newman, Thomson, & Hammond, 2008; Viswanathan et al., 2008), subsequent studies have suggested that Lin28 also blocks pre-let-7 processing by Dicer, and, even more intriguingly, facilitates the turnover of pre-let-7 (Heo et al., 2008; Rybak et al., 2008). Lin28-mediated turnover of pre-let-7 is thought to occur by the recruitment of the terminal uridyl transferase TUT4 to

the pre-let-7-Lin28 complex in the cytosol. TUT4 adds an oligo-uridine (oligo-U) "tail" to the 3' end of pre-let-7, which is thought to block Dicer processing while presumably leading to degradation of the pre-miRNA (Heo et al., 2009). The exonuclease that might carry out decay of uridylated pre-let-7 is unknown.

The observation of pre-let-7 uridylation by Lin28/TUT4 is an exciting finding which bolsters the relatively new idea that certain TUTases add an oligo-U tail to the 3' end of RNAs, ultimately leading to turnover (Mullen & Marzluff, 2008; Wilusz & Wilusz, 2008). However, many non-let-7 miRNAs are also post-transcriptionally repressed during embryonic development and in poorly differentiated tumors (J. Lu et al., 2005b; Thomson et al., 2006); furthermore, recent evidence suggests that Lin28/TUT4 might also regulate other pre-miRNAs in addition to pre-let-7. Thus, it will be crucial to determine whether non-let-7 pre-miRNAs are uridylated and/or degraded. Answering this question is very challenging due to the technical limitations in our ability to detect pre-miRNAs. Pre-miRNAs are generally far less abundant than mature miRNAs, making detection by northern blot very difficult. Other, well-established, techniques such as primer extension, RNase protection assays (RPA;(Cai & Cullen, 2007)) and intramolecular circularization followed by RT-PCR (cRT-PCR; (Basyuk et al., 2003)) are more sensitive but are laborious and cannot be used to simultaneously monitor many pre-miRNAs.

The emergence of deep sequencing technology has provided researchers in the miRNA field with a powerful tool for studying mature miRNA abundance (Hafner et al., 2008). Preparation of a cDNA library for deep sequencing entails the ligation of RNA "linker" oligonucleotides (oligos) to the 5' and 3' ends of total cellular RNA or an enriched pool of a certain type of RNA (e.g. poly-adenylated mRNAs or miRNAs); the ligation is followed by RT-PCR using the linker sequences (Hafner et al., 2008). However, this general library preparation scheme is not amenable to pre-miRNAs for several reasons. First, pre-miRNAs are highly base-paired and contain a 2 nucleotide 3' overhang; this makes the 5' end of the pre-miRNA largely inaccessible for ligation of the 5' linker. Furthermore, RNAs such

as ribosomal RNA (rRNA) and tRNAs are so abundant that they are the predominant RNAs to which the linker oligos are ligated; therefore, low abundance RNAs such as pre-miRNAs, are rarely, if ever, observed in deep sequencing data. Here, we report the development of a new technique, based on high-throughput sequencing, termed "**RM-seq**" (for **R**everse-transcription and **M**ultiplexed PCR in conjunction with deep **seq**uencing) which we have used to extensively monitor the uridylation and degradation status of over 40 microRNAs simultaneously in embryonic carcinoma (EC) cells.

Results:

RM-seq method:

The preparation of an RM-seq cDNA libarary is outlined in **Figure 4.1**. To eliminate most abundant RNAs, we isolated pre-miRNA-sized RNA from total RNA using denaturing gel electrophoresis (STEP 1). There is only one ligation reaction in the RM-seq protocol (as opposed to two ligation steps in conventional library preparation); a linker oligo is ligated to the 3' end of the RNA, which corresponds to the 5' sequencing primer during the deep sequencing reaction (STEP 2). This eliminates the abundant "linker-only" dimer PCR product that can greatly reduce the efficiency of cellular RNA sequence counts during the deep sequencing run. Following reverse-transcription with a linker specific primer (STEP 3), a miRNA-specific multiplex PCR simultaneously enriches 41 pre-microRNAs sequences and adds the second linker sequence required for cluster amplification and sequencing-by-synthesis (STEP 4). The resulting cDNA is then subjected to deep sequencing (STEP 5). In order to analyze the sequence data, we first separated the reads by barcode, which is the first 3nts of each read. The barcodes were then removed, yielding a 33-mer pre-miRNA-specific sequence (**Figure 4.1A**). The RM-seq cDNA library is relatively complex; because pre-

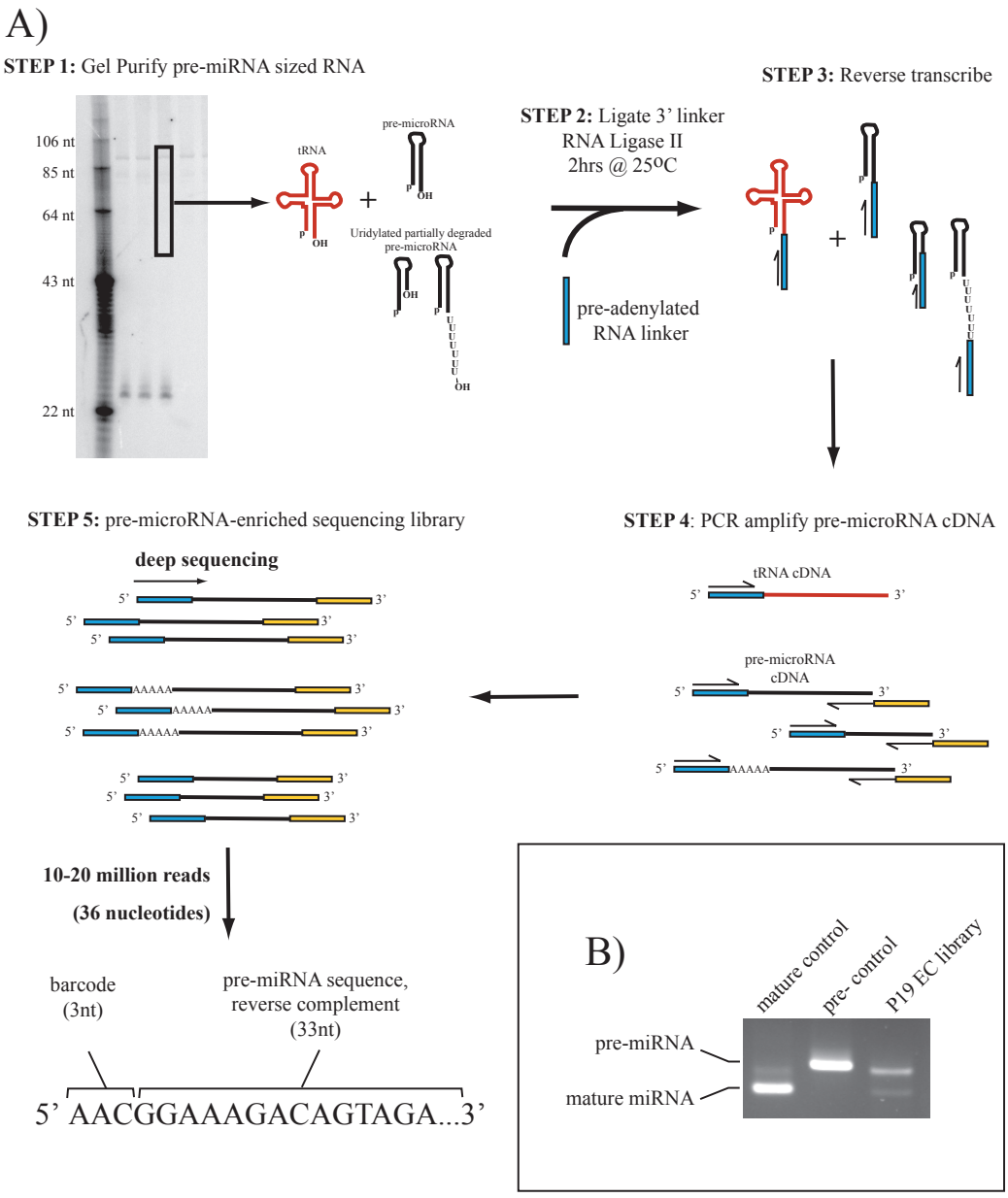


Figure 4.1: The RM-seq cDNA library protocol. A) Schematic for the construction of RM-seq deep sequencing library as described in the beginning of the "Results" section. B) Control for effective amplification of pre-miRNAs out of total RNA from P19 embryonic carcinoma (EC) cells. For size controls the protocol in A) was carried using either a synthetic mature let-7a microRNA (lane 1) or synthetic pre-let-7a-2 (lane 2); a majority of the resulting P19 library PCR product is within the pre-miRNA size-range, suggesting an enrichment of pre-miRNAs over mature miRNAs.

miRNAs are amplified from the 5' end during the PCR reaction, any information regarding the 3' end heterogeneity should be retained. As different pre-miRNAs in a given cell type are likely to vary greatly in abundance, we limited the multiplex PCR step to 10 cycles (followed by 19 additional cylces using only the linker sequences). To ensure that contamination of the cDNA library by mature miRNAs was minimal, we prepared parallel RM-seq samples using synthetic mature and pre-miRNAs as size standards. In our RM-seq library prepared from P19 EC cell RNA, the predominant PCR product comigrated with the product created using the synthetic pre-miRNA, suggesting that the enrichment of pre-miRNAs was successful (**Figure 4.1B**).

Examining pre-miRNA levels during cell differentiation using RM-seq:

It is possible that inherent short-comings of multiplex-PCR could skew the apparent abundance of different pre-miRNAs; factors like the variation in annealing temperatures between miRNA-specific primers and different melting temperatures of the pre-miRNA cDNA hairpins themselves could cause variation in amplification efficiency of certain pre-miRNAs. To test whether changes of pre-miRNA levels between samples using RM-seq reflect previously observed changes, we examined a differentiation time course of P19 EC cells treated with all-*trans* retinoic acid (RA). It is known that ES cell specific miRNAs (miR-290 and miR-302 families) are transcriptionally downregulated during ES cell differentiation (Marson et al., 2008); consistent with this data, we observed a sharp decrease in ES-specific pre-miRNA levels during P19 differentiation. Less is known about the transcriptional control of let-7 genes during embryogenesis; several groups have reported conflicting results using northern blotting to detect let-7 pre-miRNAs: in northern blots using staged mouse embryo RNA, pre-let-7 sized bands were detected only in later development (Schulman et al., 2005; Thomson et al., 2006). In another instance, pre-let-7-sized bands

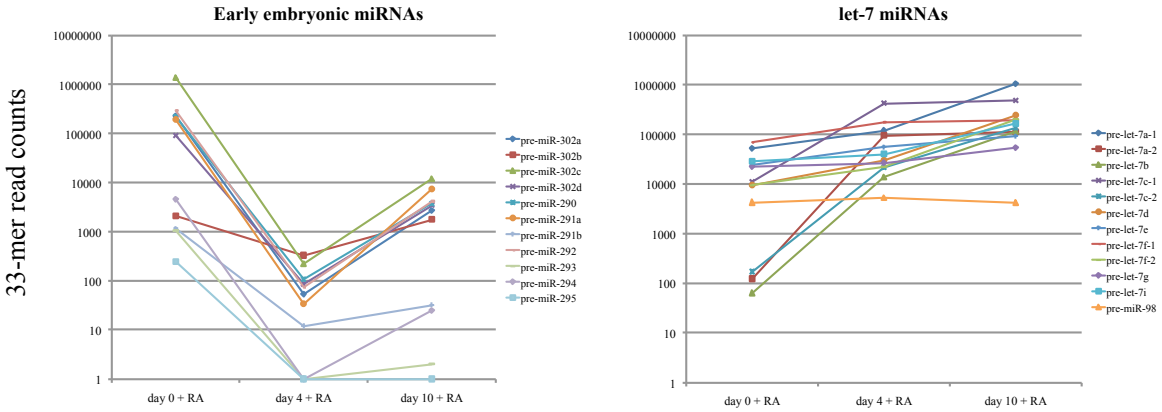
were observed in both ES cells and differentiated tissue RNA (Rybak et al., 2008; Wulczyn et al., 2007). In neither instance were the identities of the pre-let-7-sized bands verified. Using RM-seq, we observed that 11 out of 12 let-7 pre-miRNAs (as well as miR-125 pre-miRNAs [lin-4 homologs], data not shown) were greatly increased in differentiated P19 cells. (**Figure 4.2A**). Northern blotting for a let-7a and miR-302c shows that changes in mature miRNAs largely correlate with pre-miRNA levels during P19 cell differentiation (**Figure 4.2B**).

Persistent uridylation of let-7 pre-miRNAs in differentiated P19 cells:

We next investigated our RM-seq data for pre-miRNA 3' end heterogeneity. For each pre-miRNA, we counted the number of 33nt sequence reads that perfectly aligned to the complete pre-miRNA sequence that is annotated in miRBase (http://www.mirbase.org/) as well as truncated reads and reads with 3' untemplated nucleotides. Consistent with the possibility of the Lin28/TUT4-mediated uridylation of pre-let-7, we observed an over-representation of 3' untemplated uridines on 11 of 12 let-7 pre-miRNAs. Previously, two let-7 pre-miRNAs were shown to be uridylated *in vivo* (pre-let-7a-1 and pre-let-7d(Heo et al., 2008)); we found pre-let-7a-1 and pre-let-7d to be 1.8% and 4.3% uridylated, respectively (**Figure 4.3A**). Unexpectedly, many let-7 family members remained substantially uridylated during P19 differentiation (**Figure 4.3A**). For pre-let-7a-2, pre-let-7c-1, pre-let-7d and pre-let-7e, uridylated reads were not greatly enriched over other modified reads (0.9x, 1.2x, 1.5x and 0.5x fold more uridylated reads over all other modified reads combined, data not shown), suggesting these uridylated reads might be due to sequence error. However, the remaining let-7 pre-mIRNAs had a clear over-representation of uridylated reads (7a-1: 7.1x / 7b: 11.0x / 7c-2: 16.2x / 7f-1: 5.2x / 7f-2: 8.2x / 7g: 6.8x / 7i: 8.1x / miR-98: 5.6x fold more uridylated reads than all other modified reads combined, data not shown), raising the possibility that many let-7 pre-miRNAs are U-tailed in differentiated cells. To rule out the possibility that

Figure 2

A) RM-seq (pre-miRNA)



P19 differentiation timecourse

B) Northern blot (mature miRNA)

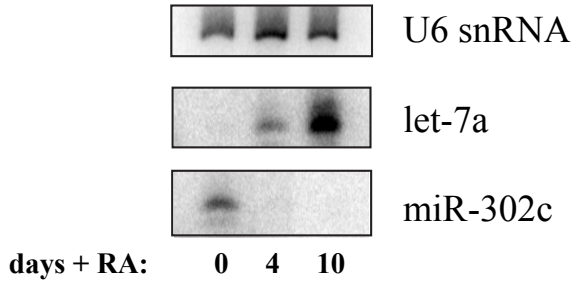


Figure 4.2: Full-length pre-miRNA levels during neuronal differentiation of P19 cells. A) Read counts (y-axis) of pre-miRNA were plotted against time points during P19 differentiation. Only 33nt sequences aligning perfectly to the annotated 3' end of pre-miRNAs were used. B) Northern blot for mature miRNA using the total RNA from the experiment in (A).

the untemplated uridines were randomly distributed across the precursor sequence, we aligned both perfectly matched ("No Mod") and uridylated ("Uridylated," **Figure 4.3B**) reads of pre-let-7a-1 and pre-let-7g to their full-length pre-miRNA sequence. There was a clear concentration of uridylated reads just beyond the 3' Drosha cut site (**Figure 4.3C**), indicative of TUTase activity.

Diverse pre-miRNAs are uridylated in P19 EC cells:

Surprisingly, we found that many other pre-microRNAs, including those corresponding to the tumor suppressor miR-15a (14.5% +U) and the early embryonic miR-302C (9.4% +U), were uridylated in P19 cells (**Figure 4.4A**). Other pre-miRNAs that were notably uridylated were pre-miR-103-1, pre-miR-23a, pre-miR-23b and pre-miR-30e (data not shown). This is consistent with a recent report in which the authors observed widespread untemplated uridylation on 3' mature miRNA strands and proposed the possibility of pre-miRNA uridylation similar to that mediated by Lin28 (Chiang et al., 2010). Inspection of the most abundant reads at each position revealed that these pre-miRNAs possessed bona fide oligo-uridine tails (**Figure 4.4B and 4.4C**); therefore, most untemplated uridines were not due to sequencing error.

Accumulation of uridylated pre-miRNAs upon exosome disruption:

In eukaryotes, oligo-uridylation has been shown to mediate RNA degradation by facilitating the recruitment of exonucleases. Therefore, we hypothesized that depletion of the RNA degradation machinery would cause an accumulation of uridylated pre-microRNAs that are normally targeted for turnover. We used RM-seq to examine pre-miRNAs in P19

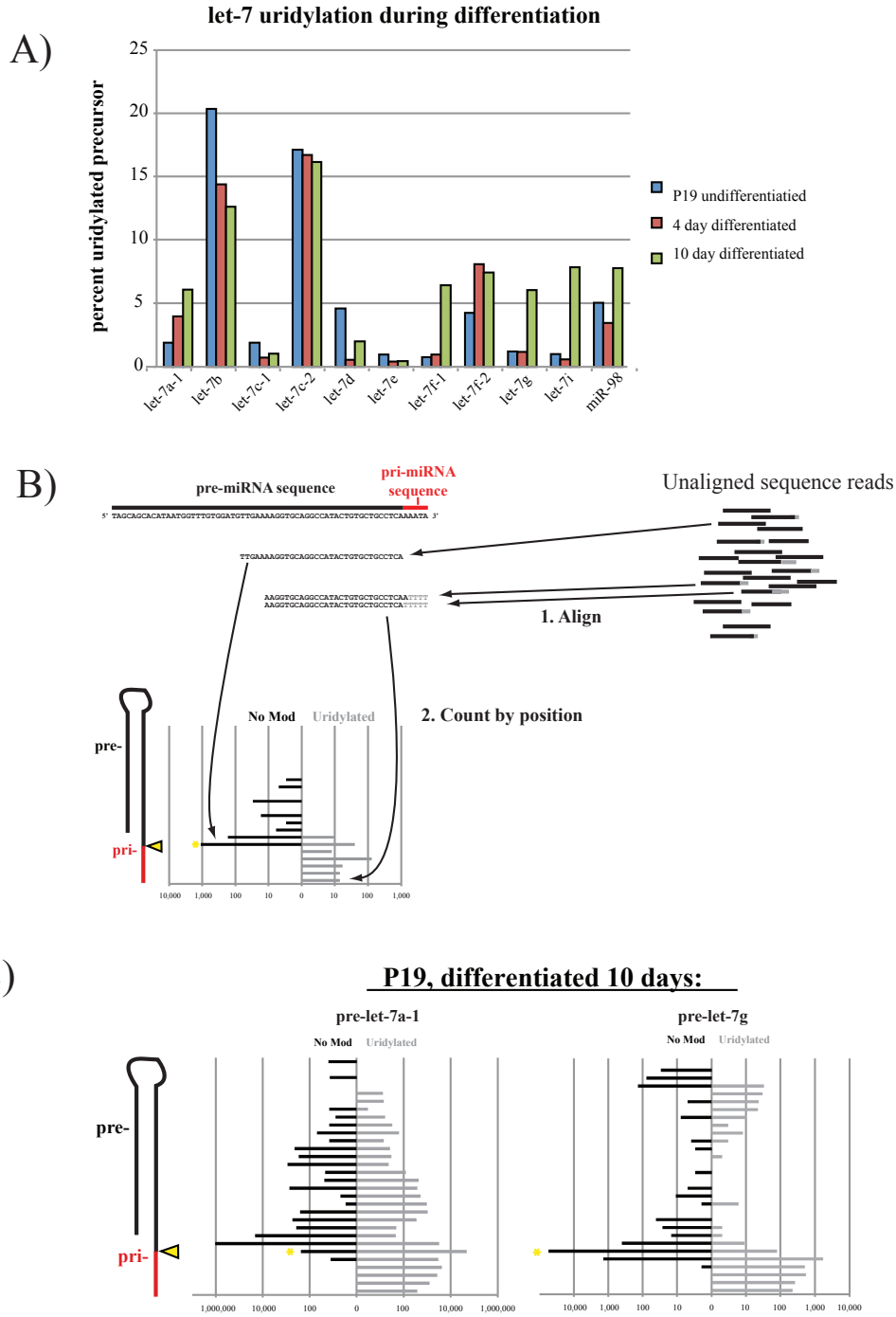


Figure 4.3: let-7 pre-miRNAs are uridylated in differentiated P19 EC cells. A) bar graph displaying the percent uridylated pre-miRNA reads during neuronal differentiation of P19 cells. B) Schematic depicting strategy used to graphically map distribution of uridylated versus perfectly matching reads. Reads are aligned to the annotated pre-miRNA sequence that has been extended by 5 nts into the pri-miRNA sequence. A read is scored as either "No Mod" (perfect match) or "Uridylated" (contains one or more untemplated uridines (grey letters). It should be noted that the scored uridylated reads are not necessarily homogeneous (see the two aligned uridylated reads in the same position); however, a large majority of uridylated reads in a given position is one sequence. C) position of perfect and uridylated reads in 10 days differentiated P19 cells for pre-let-7a-1 and pre-let-7g.

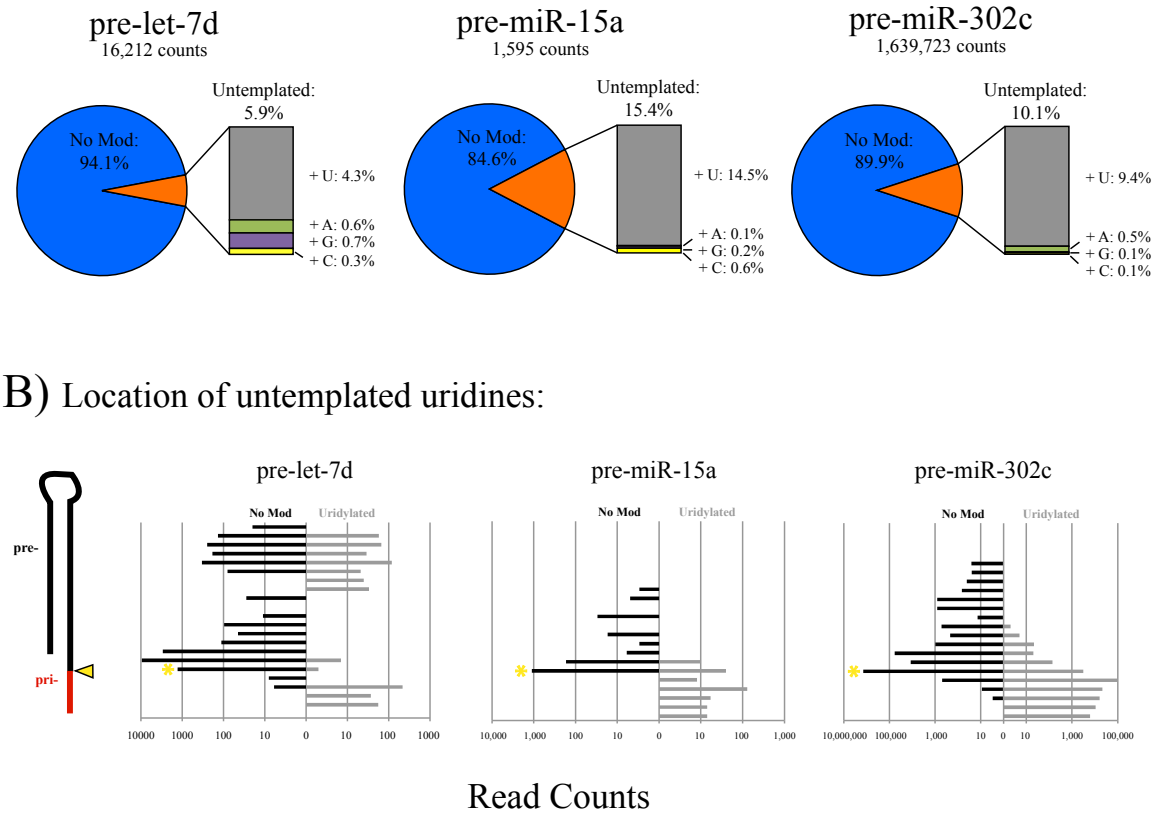


Figure 4.4: Uridylation and degradation of diverse pre-miRNAs in P19 EC cells. A) Pie charts representing the overall percentages of reads for each pre-miRNA: No Mod – reads aligning perfectly to any part of the loop or 3' arm of the pre-miRNA; +U/+A/+G/+C – reads with a non-templated U, A, G or C (respectively) within the first 8nts of the sequence read (corresponding to within 8nts from the 3' end of the cloned pre-miRNA). B) Abundance of perfect reads (x-axis, left) and urdylated reads (x-axis, right) were plotted by position along the 3' end of the pre-miRNA (y-axis); the yellow asterisk denotes reads that align perfectly to annotated 3' end of the pre-miRNA. C) Manual alignment of the most abundant reads in each position ("RM-seq reads") against the annotated pre-miRNA sequence ("pre-miR").

cells after siRNA-mediated depletion of RRP40 which was previously shown to disrupt 3'-5' exosome function (Preker et al., 2008). Indeed, when compared with a control knockdown using an siRNA to Luciferase ("Luc siRNA"), we observed an accumulation of uridylated sequence reads for many pre-miRNAs; the most drastic example of exosome-targeted pre-miRNAs were pre-miR-103-1 (Luc siRNA: 1080/15654 (6.9%) vs. RRP 40 siRNA: 4837/7005 (69.1%)) and pre-miR-23a (Luc siRNA: 151/3241 (4.7%) vs. RRP40 siRNA: 2578/5616 (45.9%), **Figure 4.5A, B and C**).

Other pre-miRNAs showed a robust accumulation of uridylated reads, including pre-miR-103-2 (Luc siRNA: 126/5582 (2.2%) vs. RRP40 siRNA 331/719 (46%)), pre-miR-16-1 (Luc siRNA: 69/6447 (1.1%) vs. RRP40 siRNA: 671/7096 (9.5%)), pre-miR-16-2 (Luc siRNA: 23/887 (2.6%) vs. RRP40 siRNA: 172/1723 (10.0%)) and pre-miR-302c (Luc siRNA: 235706/1221369 (19.3%) vs. RRP40 siRNA: 268299/873783 (30.7%).

Discussion:

In this study, we devised a new method to look at pre-miRNA levels, 3' end untemplated nucleotide addition and degradation. Levels of let-7 pre-miRNAs and early embryonic pre-miRNAs largely correlated with corresponding mature miRNAs; this suggests that the post-transcriptional block of let-7 microRNAs does not occur solely at the point of Dicer processing. Thus, the robust increase of many pre-let-7 miRNAs upon P19 differentiation may be due to a transcriptional increase as well as a release of any block at the Drosha processing step. Interestingly, different let-7 pre-miRNAs increase in strikingly different patterns. For instance, pre-let-7a-1 increases less than 2-fold during the first 4 days of differentiation while pre-let-7a-2 increases ~1000-fold. While the mechanism behind such differences aren't known, it is clear that these two closely related pre-miRNAs regulated differently during cell differentiation.

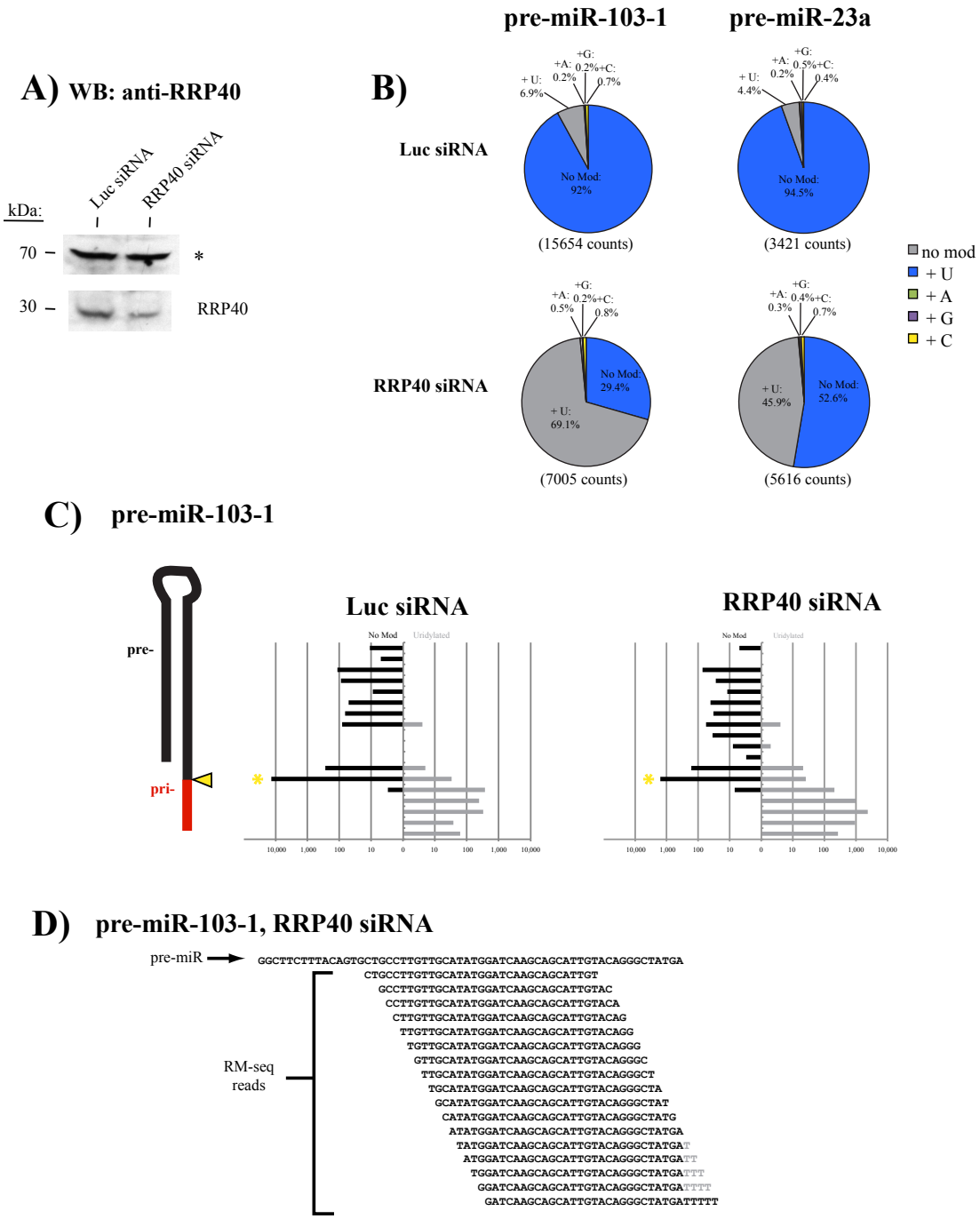


Figure 4.5: Disruption of the 3'-5' exosome pathway causes accumulation of uridylated pre-miRNAs. A) Western blot for knockdown of RRP40; asterisk denotes a cross-reacting band on the anti-RRP40 blot that was used as a loading control. B) Pie charts showing the overall percentages of perfectly-aligned versus reads with untemplated nucleotides. C) Abundance of pre-miR-103-1 reads by position for cell transfected with a control siRNA versus cells transfected with an RRP40 siRNA. D) Manual alignment of the most abundant RM-seq reads against the annotated pre-miRNA.

We observed that pre-let-7d was acted upon by a processive uridylase activity *in vivo*, corroborating a previous report (Heo et al., 2008); however, it is still not known whether this oligo-U tail is reflective of the Lin28/TUT4 pathway or whether other TUTases are able to uridylated pre-let-7. We found that it was difficult to knock down TUT4 consistently, as was previously reported (Heo et al., 2009). Creation of transgenic TUT4 knockout mouse ES cells will be required to definitively answer this question.

Surprisingly many other pre-miRNAs were significantly uridylated in P19 EC cells. For many of these pre-miRNAs, the corresponding mature miRNAs are post-transcriptionally repressed during murine embryonic development (Thomson et al., 2006). Thus, it is possible that the post-transcriptional regulation of many miRNAs occurs, in part, through pre-miRNA uridylation and degradation. Whether these pre-miRNAs are turned over by the Lin28/TUT4 pathway is not known. Lin28 binds to pre-miRNA loop sequences containing a GGAG motif (or close variations) and this motif must be precisely positioned within the pre-miRNA loop region in order for TUT4 uridylation to occur *in vitro* (Heo et al., 2009); however, this recognition site is absent from all of the uridylated non-let-7 pre-miRNAs that we observed. An especially unexpected observation is the extensive uridylation of miR-302c. As miRNAs of the miR-302 family are highly expressed in these cells, it doesn't seem likely that Lin28 blocks pre-miR-302c processing. We were also surprised to observe that many let-7 pre-miRNAs were uridylated in differentiated P19 cells. Taken together, this data points to a Lin28-independent pre-miRNA uridylation pathway in embryonic cells.

We also observed a robust accumulation of uridylated pre-miRNAs upon the disruption of 3'-5' exosome activity. This is the first evidence that the exosome pathway is, in part, responsible for the turnover of uridylated pre-miRNAs. Some pre-miRNAs showed a drastic increase in the percent of uridylated reads but a lower overall read count. While it isn't entirely clear why this occurred, we cannot rule out that the 5'-3' exonucleases, Xrn-1 and Xrn-2 also function in pre-miRNA degradation. It is possible that the partial ablation of

exosome activity leads to increased 5'-3' degradation which would account for the decreased abundance of certain pre-miRNAs.

It should be noted that we did not actively search for U-tails longer than 5nts; because our deep sequence reads were only 33nts, we couldn't simultaneously search for longer U-tails and match sequence reads sufficiently to distinguish closely related pre-miRNAs. Thus, it is entirely possible that we cloned many pre-miRNAs with longer U-tails (such as let-7 pre-miRNAs with 10-30 U's as was previously observed, (Heo et al., 2008)) but could not successfully curate them due to the confined sequence space. While we observed up to 9 U's on many partially-degraded pre-miRNAs (data not shown), these reads were relatively rare suggesting that a pre-miRNAs with this many U's are highly unstable.

It leads to reason that RM-seq has many potentially useful applications; while this study makes use of RNAi in mammalian cell culture, RM-seq would also be particularly powerful in model organisms that are much more amenable to genetic manipulation such as *C. elegans* (worms) and *D. melanogaster* (flies). Of course, RM-seq could be employed to study the degradation and untemplated nucleotide addition of many other groups of RNAs such as (e.g.) small nuclear RNAs (snRNAs), tRNAs and groups of candidate mRNAs that may be turned over by uridylation (e.g. histone mRNAs).

Taken together, this data suggests that RM-seq can be a powerful tool for approximating fully intact pre-miRNAs levels as well as for the thorough analysis of pre-miRNA degradation and uridylation on a much larger scale than is possible with conventional techniques for RNA detection.

Methods and Materials:

Cell Culture and transfection:

P19 embryonic carcinoma (EC) cells were cultured, transfected with siRNAs and differentiated with all-trans retinoic acid (RA) as previously described (RNA paper); to enhance knockdown efficiency, P19 cells were transfect with siRNAs on three consecutive days. Cells were harvested for RNA isolation 48 hours after the final transfection.

RM-Seq cDNA library preparation:

RNA preparation:

P19 total RNA was obtained using Trizol. High molecular weight (MW) RNA from 50ug of total RNA was partially depleted by precipitation with 5% PEG_{8000} / 0.5M NaCl; the lower MW fraction was separated by electrophoresis on an 8% denaturing acrylamide/urea gel. 50-100nt RNA was isolated and eluted by incubation of gel slices overnight in 0.3M Sodium Acetate along with 50uL of phenol-chloroform. Eluted RNA was then precipitated with 4uL of glycogen and an equal volume of isopropanol.

RNA cloning linker and pre-adenylation:

Cloning linker RNAs were purchased from IDT. The RNAs contained 5' phosphates, were blocked on the 3' end by di-deoxy-cytidine and had the following sequences (barcodes are italicized and underlined; it should be noted that an oligo starting with the 5' sequence "CCG…" was also purchased but failed to be pre-adenylated):

Barcode ***AAC***:

5' p*GUU*GAUCGUCGGACUGUAGAACUCUGAACddC 3'

Barcode ***TTC***:

5' p*GAA*GAUCGUCGGACUGUAGAACUCUGAACddC 3'

Barcode ***CCC***:

5' p*GGG*GAUCGUCGGACUGUAGAACUCUGAACddC 3'

Linkers were pre-adenylated as previously described with some modifications (Vigneault, Sismour, & Church, 2008). 800 pmols of RNA linker was incubated for 5 hours at 37°C in a reaction containing 200 units of T4 RNA ligase I (New England Bio labs (NEB)), 1x RNA ligase I buffer [50mM Tris-HCl pH 7.8 / 10mM $MgCl_2$ / 1mM ATP / 10mM DTT], 4 units of Ribolock RNase A inhibitor (Fermentas) and 25% DMSO. The RNA was then extracted with Phenol-Chloroform and electrophoresed on an 8% denaturing gel. Adenylated RNA oligo (which migrates roughly 1-2 nucleotides higher than un-adylated oligo) was gel purified away from un-adenylated oligo and eluted as described above.

Ligation, reverse transcription and PCR amplification:

500ng of gel-purified RNA was ligated to 10 pmol pre-adenylated linker for 2 hours at 25°C in a reaction containing 200 units RNA ligase II, truncated (NEB), 1x PNK buffer [70mM Tris-HCl pH 7.6 / 1mM $MgCl_2$ / 0.5 mM DTT] and 2 units of Ribolock. RNA was then extracted with phenol-chloroform and precipitated with ethanol, 0.3M sodium acetate and 2uL glycogen. Ligated RNA was reverse transcribed using the primer, **5' GTTCAGAGTTCTACAGTCCGA 3'** and Superscript II reverse transcriptase (Invitrogen) according to the manufacturer's instructions.

The resulting cDNA was then amplified in two PCR reactions:

1) miRNA-specific, 10 cycles: 4 uL of undiluted cDNA was amplified with the Solexa small RNA forward primer, **5' AATGATACGGCGACCACCGACAGGTTCAGAGTTCTACAGTCCGA 3' (2.5uM)**, a mixture of 42 reverse primers; these primers contained the Solexa small RNA reverse primer appended to the sense sequence of the 5' arm of 42 different miRNAs, **5' CAAGCAGAAGACGGCATACGA[sense sequence of pre-microRNA 5' arm] 3' (1.25mM; see supplemental data for full list of**

primer sequences) and Phusion DNA polymerase (Finnzymes) according to the manufacturer's instructions.

2) Solexa primers only, 19 cycles: 1% of the first PCR reaction was amplified with the forward primer above and the solexa small RNA reverse primer, **5' CAAGCAGAAGACGGCATACGA 3'** and Phusion DNA polymerase.

Thermo-cycling profile was as follows: 98°C for 30 seconds followed by cycles of 98°C for 10 seconds, 59°C for 30 seconds and 72°C for 15 seconds.
PCR reactions were Phenol extracted and ethanol precipitated.

High-throughput DNA sequencing and data analysis:

36 nucleotide sequence reads were generated on the Solexa genome analyzer II (Illumina). The data was then analyzed using custom scripts (S.M.H and M.A.N.). Briefly, sequence reads were separated by barcode and the barcodes were removed.

Sequence reads that completely matched pre-miRNA sequences or contained non-templated uridines (read as an "A" in the raw sequence) were annotated by their alignment to a particular position along a given pre-miRNA sequence using a "register" number. Register numbers were established whereby the a register number of "1" aligned to the end of the pre-microRNA and contained 6 additional nucleotides beyond the Drosha cleavage site (either matching sequences or, most often, containing non-templated nucleotides).

i) Pie charts were created in Microsoft Excel. "Tornado" charts were created using Microsoft Excel. Briefly, unmodified and uridylated sequence counts by register count in two separate bar graphs which were juxtaposed by dividing 1/X (where X = count of uridylated reads) and making the x-axis logarithmic. Decimal values were converted to integers in Adobe Illustrator.

cRT-PCR:

Analysis of 5' and 3' pre-miRNA ends was carried out as previously described (Basyuk et al., 2003). Briefly, total RNA was treated in a dilute RNA ligase I reaction on ice overnight to promote circularization of individual RNAs (200uL reaction volume containing 5ug RNA, 1x RNA ligase I buffer, 2 units of Ribolock and 20 units of RNA ligase I). The ligated RNA was then phenol-chloroform extracted and 1ug was used in a conventional reverse transcription (RT) reaction using Superscript II. The resulting RT reactions were diluted to 100uL and 5uL was used per PCR reaction (25uL reaction containing 1x Titanium Taq buffer [40mM Tricine, pH 9.2 / 15mM KCl / 3.5mM MgCl2 / 3.75ug/mL BSA], 0.15mM dNTPs, 1.5 units of Hotstar Taq DNA polymerase (Qiagen) and 0.1uM primers). The primers used were as follows:

Pre-let-7d-cRT CTCCTTGTGGGCAAAATC

Pre-let-7d-cRTPCR_F TCCTTGTGGGCAAAATCTCT

Pre-let-7d-cRTPCR_R TTAACTATACGACCTGCTGCCT

Pre-miR-103-1_cRT, CTTGATCCATATGCAACAAG

Pre-miR-103-1_cRTPCR_F, GCAACAAGGCAGCACTGTAA

Pre-miR-103-1_cRTPCR_R, CAGCATTGTACAGGGCTATG

Pre-miR-302c_cRT, TTGCTTTTGTTTAACACAGCA

Pre-miR-302c_cRTPCR_F, CACAGCAGGTAACCCCAT

Pre-miR-302c_cRTPCR_R, GTGCTTCCATGTTTCAGTG

Pre-miR-15b_cRT, CTGTAGTATGTAAACCATGATG

Pre-miR-15b_cRTPCR_F, ATGTAAACCATGATGTGCTGCT

Pre-miR-15b_cRTPCR_R, GCGAATCATTATTTGCTGCTC

Northern Blotting:

20ug PEG-treated RNA was separated on a 12.5% denaturing urea-acrylamide gel

and transferred to a Hybond N+ nylon membrane (Amersham). Membranes were hybridized overnight at 37oC with [P32] 5'-end labeled DNA probes of the following sequences: 5' - ACT ATA CAA CCT ACT ACC TCA – 3' (let-7a), 5' – CCA CTG AAA CAT GGA AGC ACT TA – 3' (miR-302c) and 5' – TTC ACG AAT TTG CGT GTC AT -3' (U6 snRNA). Membranes were then washed 3x at 37oC with 2xSSC / 0.05% SDS and 2x at 25oC with 0.1x SSC / 0.1% SDS;
washed blots were then exposed to a phosphor-imaging screen and images were recorded on a Typhoon scanner (GE Life Sciences).

Western blotting:

Control or RRP40 knockdown P19 cells were lysed in 2x loading buffer; proteins were then separated on a 12% SDS-Tris-Glycine Lamelli gel slab (???) and transferred to a Hybond ECL nitrocellulose membrane (GE Life Sciences). The membrane was blotted with a rabbit polyclonal antibody to ExoSC3 (Abcam, catalog # ab76605) diluted 1:1000.

CHAPTER 5:

Emerging paradigms of regulated microRNA processing and future directions[1].

Summary:

Early studies by several labs found that mature miRNA expression does not always correlate with expression of the pri-miRNA (Blenkiron et al., 2007; Obernosterer et al., 2006; Thomson et al., 2006; Wulczyn et al., 2007). Thus, miRNAs themselves must be post-transcriptionally regulated. In fact, regulation at multiple biogenesis steps and at turnover of the mature miRNA has now been established (Pawlicki & Steitz, 2010). The majority of discoveries on miRNA regulation can be distilled down to two contrasting paradigms based on the biochemical point of regulation: at the Microprocessor complex and at the terminal loop of specific microRNA precursors.

Multiple regulatory events converge on the Microprocessor complex:

As the first processing step in miRNA biogenesis, the Microprocessor is positioned to play a pivotal role in the regulation of mature microRNA abundance. This complex is minimally composed of two proteins; the double stranded RNA binding protein DGCR8 (Pasha) and the RNase III enzyme Drosha (Denli et al., 2004; Gregory et al., 2004; Landthaler, Yalcin, & Tuschl, 2004). These two proteins represent the essential requirements for the first processing step, defined by reconstitution of activity with purified recombinant proteins. In human cell extracts, however, Drosha has been described to reside in multiple

---

1 Parts of this chapter are reprinted from Newman and Hammond, Genes Dev. 2010 Jun 1;24(11):1086-92.

complexes (Gregory et al., 2004). A complex that is solely comprised of Drosha and DGCR8 has robust pri-miRNA processing activity. A second, larger complex, has multiple accessory proteins, yet also has pri-miRNA processing activity. It is not clear whether this large complex is assembled around RNAs, since many of the accessory factors contain RNA interaction motifs, and RNase treatment of extracts has been shown to shift the size of the Drosha complex (Han et al., 2004). The accessory proteins in this large Drosha complex include the EWSR1, Fus, numerous hnRNP proteins, and the DEAD box helicases p68(DDX5) and p72(DDX17). While the exact biochemical composition of the cellular Microprocessor is unknown, a recent collection of studies implicates p68 and p72 as important cofactors of the Microprocessor *in vivo*.

Integrating signal transduction with Drosha processing: the role of p68 and p72.

The p68 and p72 helicases are conserved across eukaryotes and are implicated in diverse RNA processing pathways (Fuller-Pace, 2006). In cell free assays, both proteins have moderately processive ATP dependent RNA unwinding and annealing activities. In cells the proteins have been linked to several pathways, most notably splicing and transcription. p68 was isolated as a component of the spliceosome, and specifically is required for proper assembly-disassembly of the U1 5' splice site. Interestingly, p68 and p72 interact with several components of the transcriptional machinery, including RNA polymerase II and CBP/p300. This raises the interesting possibility that one role of p68 is to coordinate/co-localize the splicing machinery with the transcription apparatus. Independent of this, p68 acts as a transcriptional cofactor for several factors, including CBP/p300, p53, and the Estrogen Receptor ERa. For example, etoposide mediated induction of p53 responsive genes is dependent on p68. Remarkably, this requirement is not dependent on p68 helicase activity, as an ATPase mutant allele of p68 rescued induction by p53. RNA binding, however, was still required. Similarly, ERa coactivation is not dependent on helicase activity. This suggests a model whereby p68 and possibly p72 mediate assembly of transcription complexes at

promoters, recruit the splicing machinery to the promoter, thus facilitating co-transcriptional splicing events.

Recently, p68 and p72 were implicated as regulators of processing for many miRNAs (Fukuda et al., 2007). This was first demonstrated in studies on mice that are homozygous null for either p68 or p72. For both genes, loss of function is lethal; p68 null mice die at embryonic day 11.5, and p72 null mice are postnatal lethal. miRNA expression analysis of null embryos indicated reduced mature miRNA levels. Whether or not the reduction of these miRNAs was due to loss of function in the respective helicase, or whether the mutant phenotype led to aberrant miRNA expression due to developmental alterations, remains to be addressed. However, fibroblasts (MEFs) derived from p72 null embryos had a reduction in at least two miRNAs (miR-16 and mir-145), but not all miRNAs, and this reduction could be rescued by introduction of ectopic p72. Pri-miRNA expression was unaffected, demonstrating an alteration in processing of the miRNA precursors. Interestingly, ATPase mutant alleles of p72 did not rescue production of either miR-16 or miR-145. This is in contrast to the transcriptional co-activation role for p68, which was not dependent on helicase activity. Therefore, for at least these two miRNAs, a role for helicase activity (or possibly ATP dependent conformation change) is required. Furthermore, cell free extracts that lack either p68 or p72 are inefficient at conversion of pri-miR-16 to the corresponding precursor miRNA. RNA immunoprecipitation (RIP) experiments in the same study showed that Drosha recruitment to pri-miR-199a and pri-miR-214 was lost upon RNAi-mediated depletion of p68 or in p72 knockout cells, consistent with the observation that mature miR-199a and miR-214 levels are lower in p68/p72 null embryos. Thus, *in vivo*, these helicases seem to be required to properly recruit the Microprocessor to some pri-miRNAs. As discussed below, it is now clear that multiple critical cellular signaling pathways utilize the p68 and p72 association with Microprocessor to effect regulation of pri-microRNA processing.

TGF-beta/BMP.

The first study to report a signal transduction-mediated change in miRNA biogenesis demonstrated that the transforming growth factor (TGF-beta) and bone morphogenic protein (BMP) pathways specifically promote processing of pri-miR-21 (Davis et al., 2008). The contractile phenotype of human smooth muscle cells treated with TGF-beta or BMP was found to be dependent on the rapid up-regulation of miR-21. Interestingly, this rapid increase in miR-21 occurred post-transcriptionally; quantitative real-time PCR (qRT-PCR) experiments revealed that pre-miR-21 and mature miR-21 levels quickly increased upon TGF-beta and BMP treatment while levels of pri-miR-21 were unchanged. The authors then reasoned that SMAD proteins might be the intermediaries responsible for connecting these signaling pathways with miR-21 processing as they are translocated into the nucleus upon TGF-beta or BMP signaling, and are known to bind DNA. Furthermore, a previous study reported that SMAD1 interacts directly with the RNA helicase p68 (Warner et al., 2004). Accordingly, depletion of receptor-specific SMADs (SMAD1, SMAD5) or p68 abrogated the BMP and TGF-beta-driven increase in pri-miR-21 processing. Some initial mechanistic insight was also provided; GST-pulldowns with recombinant SMADs in nuclear extracts indicated that p68 binds to the MH2 domain of R-SMADs while the MH1 domain binds to pri-miR-21 either directly or through other unknown factors.

The full complement of miRNAs that are regulated by this pathway is presently unknown. Since the TGF-beta and BMP signaling pathways regulate many biological phenomena, it is possible that these pathways stimulate processing of other pri-miRNAs; indeed the authors also observed that pri-miR-199a processing is regulated in this manner. This is not surprising, since these signaling pathways converge on miRNA processing via p68, and this helicase has been shown to modulate a number of miRNAs.

The p53 DNA damage pathway.

Unexpectedly, a subsequent study demonstrated that the tumor suppressor p53 also promotes pri-miRNA processing. It was noted that many microRNAs depleted in p72 knockout mice, coincidentally, are also upregulated by the DNA damaging agent Doxorubicin (Dox). As Dox strongly stimulates p53 activity, leading to induction of p53 responsive genes, it was perhaps not surprising that p53 might also induce the expression of miRNAs. In fact, miR-34 had been well established as a direct p53 transcriptional target (X. He, He, & Hannon, 2007). However, after measuring pri-miRNA expression levels of induced miRNAs, the authors realized that these miRNAs were not being transcriptionally induced (Suzuki et al., 2009). Rather, processing at the Drosha step was triggered. In hindsight this was not so surprising: p68 was known to regulate pri-miRNA processing, and the helicase had been shown to interact with p53 as a co-activator; *ergo*, p53 might be able to directly modulate miRNA processing (Bates et al., 2005). Accordingly, co-immunoprecipitation experiments as well as *in vitro* Drosha processing assays demonstrated that p53 associated with the large Drosha complex upon treatment of cells with Dox. Intriguingly, p53 alleles with known oncogenic activity *decreased* processing efficiency of p53-targeted microRNAs by disrupting p68's association with Drosha, thus displacing the Drosha complex selected pri-miRNAs. Again, the biochemical basis of p53's association with p68 and the subsequent increase in target pri-microRNA processing remains obscure. Detailed structure-function studies will be needed to understand how p68 binding to p53 (both wildtype and oncogenic alleles) modulates Drosha's ability to process pri-miRNAs.

Estrogen receptor alpha/Estrogen.

Another interesting report revealed how a signaling cascade can *negatively* regulate pri-miRNA processing. Kato and colleagues noticed that ERa was previously shown to interact with p68 and p72 (Endoh et al., 1999). Augmentation of the ERa signaling pathway produced a striking pattern of changes in certain microRNAs; many microRNAs were increased in ERa -/- mice (Yamagata et al., 2009). Furthermore, administration of estradiol

(E2) to ovariectomized mice reduced the same miRNAs within the uterus. Importantly, this effect of E2/ ERa could be recapitulated in the human breast cancer cell line MCF7. The physiological importance of ERa-regulated microRNA biogenesis was evident from the observation that the 3' untranslated region (3' UTR) of VEGF, an ERa target gene, is targeted by ERa-repressed miRNAs. The authors then demonstrated *in vitro* that E2-bound ERa could directly inhibit Drosha processing of ERa-target pri-microRNAs. This study delved into more biochemical detail than the reports mentioned above: it was determined that p68 and/ or p72 bridge the interaction between Drosha and E2-bound ERa; this interaction requires the C-terminal domain of Drosha and the N-terminus of ERa. Thus, when ERa is recruited to the large Drosha complex in an E2-dependent manner, Drosha is dissociated from ERa-targeted pri-miRNA loci. It will be interesting to determine whether ERa/E2 weakens the overall integrity of the large Drosha complex or simply its affinity for certain pri-miRNAs.

A pri-miRNA processing holoenzyme? Facts and future directions.

As mentioned above, Drosha isolated from mammalian extracts exists in a large complex with many different RNA-binding proteins. The role of most of these Drosha-associated factors in pri-microRNA processing is not clear. It seems entirely possible that many of the factors purified in the large cellular Drosha complex may be the result of a non-physiological RNA-protein complex assembled upon cell lysis. In support of this, while the dsRNA binding proteins NFAT-90 and NFAT-45 were previously shown to be Drosha associated factors (Gregory et al., 2004), one recent study showed that this heterodimer does not interact with Drosha and, in fact, inhibits processing of certain pri-microRNAs (Gregory et al., 2004; Sakamoto et al., 2009). This study underscores the need to more carefully analyze the functional interactions between Drosha-DGCR8 and the many auxiliary factors that are contained within the large complex. As with p68 and p72, there may indeed be important roles for other Drosha-associated factors, as general Microprocessor co-factors or as factors that integrate specific cell -signaling pathways with pri-miRNA processing.

So does Drosha/DGCR8 function within the context of a larger holoenzyme *in vivo*? The studies mentioned above suggest that, of the Drosha-associated factors initially identified, p68 and p72 are the first to be assigned as important co-factors in Drosha processing of certain pri-microRNAs. Taken together, the aforementioned studies suggest a model shown in **Figure 5.1**. Drosha processing is known to occur co-transcriptionally (V. N. Kim, Han, & Siomi, 2009b). p68 has been proposed to link mRNA transcription to splicing (Fuller-Pace & Ali, 2008). Therefore, p68 and/or p72 might perform a similar function with miRNAs. That is, p68 might act as a bridge between the Microprocessor and specific miRNAs, at the RNA polymerase holoenzyme during transcription. This effect might be independent of helicase activity, as is the role of p68 in mRNA transcriptional events. However, p72 activation of miR-16 and miR-145 processing is ATP dependent, suggesting a second, helicase dependent role for p72 (Fukuda et al., 2007). This is further supported by the fact that p72 modulates Microprocessor function in cell free extracts, which is unlinked to transcription. It is important to note that all these cell free assays are based on crude extracts or on partially purified components. Without an exact understanding of the protein composition of the assays, it is difficult to fully understand the role of these accessory proteins.

II. It's all in the loop: specific interactions between regulatory proteins and the terminal loop of microRNA precursors

As discussed above, one mode of regulated microRNA biogenesis occurs through modulation of Microprocessor activity. A second regulatory phenomenon occurs when RNA binding proteins directly bind to specific sequences within pri-miRNAs, diverting the RNA away from biogenesis events, and often leading to degradation of the RNA.

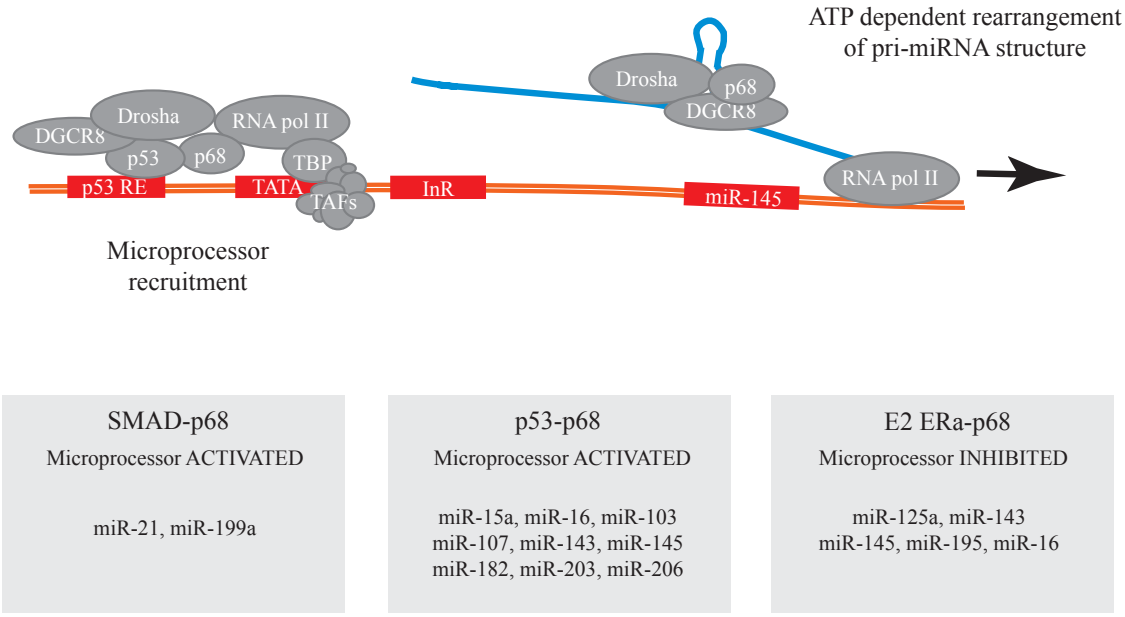


Figure 5.1. Regulation of miRNA biogenesis at the Microprocessor. Multiple signaling pathways converge on the Microprocessor via the helicases p68 and p72. These helicases have ATP dependent and independent functions. We present a model whereby both functions impinge on miRNA production. The ATP independent functions have been linked with recruitment of processing activities to RNA polymerase. This could be utilized by the miRNA pathway, as a manner of recruitment of the Microprocessor to transcription sites. The ATP dependent role of p68 has been described as maintenance of proper RNA secondary structure to facilitate Drosha cleavage. The components of the RNA polymerase preinitiation complex and Microprocessor are simplified.

Lin28 as a master regulator of let-7 production.

let-7 is one of the most abundant miRNA families in mammals, with high expression in essentially all adult tissues. In embryonic cells, in contrast, mature let-7 is present at ~1000 fold lower levels. Unexpectedly, pri-let-7 expression is constant throughout development, thus providing a dramatic illustration of post-transcriptional regulation of a miRNA (Thomson et al., 2006). Research published over the past two years has demonstrated that the RNA binding protein Lin28 blocks let-7 microRNA maturation in early embryonic cells, and the related Lin28B contributes to let-7 reduction in cancer (reviewed in (Viswanathan & Daley, 2010)). Initial studies demonstrated that Lin28 represses Drosha and Dicer processing of let-7 *in vitro* by binding to conserved sequences in the precursor loop. Furthermore, over-expression of Lin28 in differentiated cells and knockdown of Lin28 in embryonic cells resulted in the depletion and accumulation of mature let-7, respectively. The exact mechanism whereby Lin28 blocks processing steps is unknown. What was clear is that the blocked pri-miRNA/pre-miRNA was not accumulating in cells, suggesting a turnover mechanism. We now know, in the case of the pre-let-7, that Lin28 directly recruits a poly(U) polymerase, TUT4 to the precursor RNA (Heo, Joo, Kim, Ha, Yoon, Cho, Yeom, Han, & Kim, 2009a). Oligo-uridylation of the precursor occurs, which triggers degradation. The fate of the blocked pri-let-7 is still unknown (see **Figure 5.2**).

The story of Lin28 thus comes full circle. This embryonic stem (ES) cell-specific RNA binding protein was one of the earliest-recognized let-7 targets (Reinhart et al., 2000). In embryonic development, its expression is reciprocal to that of mature let-7. Concordantly, Lin28 over-expression, in combination with several ES-specific transcription factors, is sufficient to reprogram somatic cells into induced pluripotent stem (iPS) cells (Yu et al., 2007).

KSRP binds let-7 loop sequences and activates microRNA processing:

Recently, a role was uncovered for the RNA binding protein KSRP in *promoting* microRNA biogenesis in mammals (Trabucchi et al., 2009). This protein specifically binds to 5'guanosine-rich patches on the loop region of several microRNA precursors, including let-7. Knockdown of KSRP significantly decreases the expression of these miRNAs, in part by preventing efficient recruitment of Drosha and Dicer to the pri- and pre-microRNAs, respectively. The obvious question is how can Lin28 and KSRP bind to the let-7 loop, but with opposing terminal outcomes? Interestingly, while recombinant fragments of KSRP bind to the let-7 terminal loop with high affinity *in vitro*, KSRP cannot bind to pri-let-7 in embryonic cells, which have abundant Lin28. This suggests a model whereby Lin28 and KSRP bind in a mutually exclusive manner. Specifically, Lin28 may sterically hinder KSRP binding even thought the two proteins bind to unique sequences on opposites sides of the let-7 loop. Alternatively, an unidentified factor in embryonic cells could bind cooperatively with Lin28 to let-7 loops, thereby preventing KSRP binding.

Regulation of microRNA processing by hnRNP proteins:

While the above examples point to regulators with specific functions, there are other examples of RNA binding proteins with a pleiotropic function in RNA biogenesis. One such example is the highly abundant RNA binding protein hnRNP A1. This protein has been well established as a component of many heterogenous nuclear RNA complexes (hnRNP) and performs essential functions in many RNA processing and transport pathways (He and Smith 2009). Using cross-linking immunoprecipitation (CLIP, (Guil & Caceres, 2007; Ule et al., 2005)). This miRNA is one of six miRNA stem-loops on the oncogenic polycistron miR-17-92. The role for hnRNP A1 in miRNA processing was demonstrated when HeLa extracts depleted of hnRNP A1 by RNAi failed to process pri-miR-18a from pri-miR-17-92 while the other pri-miRNAs in the polycistron were still processed. Interestingly, when recombinant hnRNP A1 was added back to the depleted extract, pri-miR-18a processing

was not restored; this suggests that hnRNP A1 might have to be "pre-complexed" with other RNA-binding factors in order to act specifically on pri-miR-18a. Of course, it is also possible that bacterially purified hnRNP A1 is missing post-translational modifications that are required for the positive regulation of pri-miR-18a processing.

Further studies demonstrated that hnRNP A1 binding to the loop and stem regions of pri-miR-18a relaxes the stem region and facilitates cleavage by Drosha. Bioinformatic analysis, RNA affinity purification using microRNA loop sequences as well as *in vitro* Drosha assays revealed that many hnRNP proteins could potentially form an extensive interaction network with microRNA loops (Michlewski, Guil, Semple, & Caceres, 2008). Thus, hnRNPs may broadly control the levels of microRNAs at the post-transcriptional level. The general physiological functions of hnRNP-microRNA loop interactions remain to be elucidated; ultimately, it will be important to understand the global distribution of hnRNP proteins between microRNAs and non-microRNA-containing substrates.

III. Future directions:

Follow up RM-seq experiments:

There are several outstanding questions from the intial RM-seq study. First, where in the cell does the non-let-7 pre-miRNA uridylation occur? Crude nuclear/cytoplasmic fractionation will reveal possible selective compartmentalization of the oligo-uridylation of certain pre-miRNAs. Secondly, it is possible that uridylation of very abundant pre-miRNAs (e.g. pre-miR-302c) is reflective of the saturation of miRNA biogenesis machinery; excess pre-miRNA substrate might be deferred to a turnover pathway in the nucleus and/or cytoplasm. This hypothesis is supported by an intriguing study showing that massively overexpressed artificial pre-miRNAs (as a vehicle for gene therapy) caused a broad reduction in mature miRNA expression due to inundation of nuclear export factor, Exportin-5 (Exp-5)

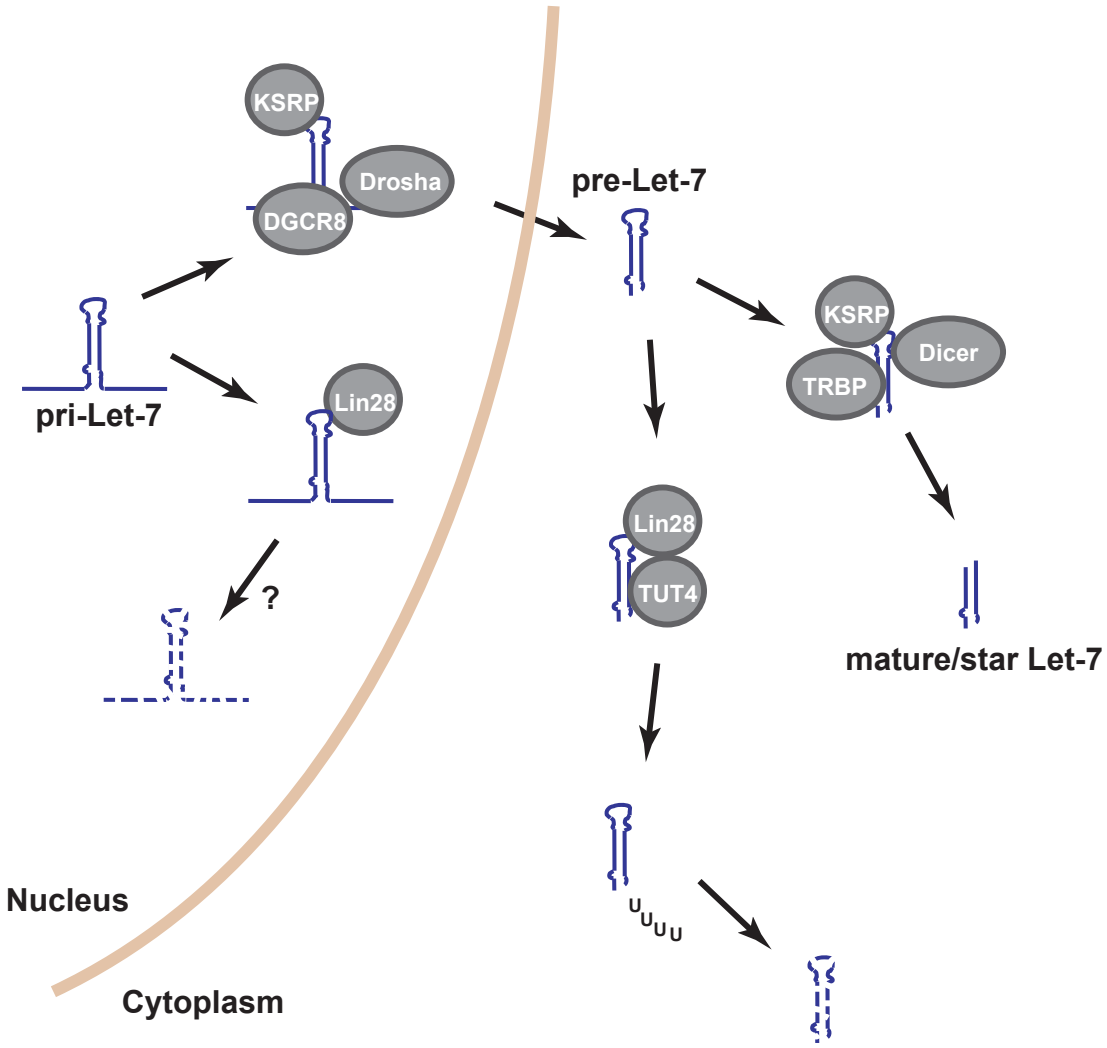


Figure 5.2. Regulation of miRNA biogenesis by loop binding proteins. The processing of let-7 is shown, with contrasting effects of an inhibitor (Lin28) and an activator (KSRP). Both regulators have been linked to the Drosha and Dicer step. In addition, Lin28 promotes degradation of the precursor. Degradation of pri-let-7 has not been demonstrated; however, pri-let-7 does not accumulate in the presence of Lin28 blockade, and therefore may be actively degraded.

(Grimm et al., 2006). To test this possibility, we will deplete Exp-5 in P19 cells and analyze the overall degree of pre-miRNA uridylation in comparison to control cells. If Exp-5 causes an increase in the uridylation of abundant pre-miRNAs this would suggest that one function of Lin28-independent uridyalation is to mark "unexportable" pre-miRNAs for turnover.

Because miRNAs are globally down regulated in poorly differentiated cancers, we will conduct a large-scale RM-seq project comparing the uridylation and degradation status of pre-miRNAs in poorly differentiated tumor samples versus non-cancerous adjacent tissue from UNC cancer patients. If we do find that pre-miRNAs are globally uridylated in the cancer samples it might provide a major mechanistic insight in to how some of the most aggressive cancers seemingly regress into an undifferentiated state. This could also lead to the development of diagnostic tools to characterize patient tumors; for example, it would be relatively simple to devise a rapid and efficient qRT-PCR protocol to detect any highly uridylated pre-miRNAs using a very small amount of RNA from a resected tumor.

We have preliminary data indicating that knockdown of TUT4 causes a marked increase in the levels of pre-let-7 and a corresponding decrease in terminal uridylation, which was expected. Surprisingly, knockdown of RRP40 revealed the same phenotype; this suggests the exciting possibility that the TUT4/Lin28/pre-let-7 complex might actually require the direct physical association of 3'-5' exosome components for TUT4 uridylation of pre-let-7 to occur. This raises the question: is there a Lin28/TUT4/Exosome-containing "holoenzyme" that selectively ensures the efficient turnover of let-7 pre-miRNAs? To address this possibility we will knockdown other crucial components of the exosome and look for the same phenotype. The goal would be to identify as many components of this putative "pre-let-7 degradosome" complex as possible. We would eventually move *in vitro* assays using P19 cell extracts in order to biochemically purify a partial or complete degradosome complex.

The unexplored potential of natural chemical modifications of small non-coding RNAs (ncRNAs) and the role of these modifications in ncRNA function:

The best-known example of RNA modification in the RNAi field is the 2'-O-methylation of piRNAs and (in plants) by the RNA methyltransferse HEN1. This protects against uridylation and degradation (Li, Yang, Yu, Liu, & Chen, 2005).

I previously investigated the possibility that pre-let-7 might accumulate in cells based on previously published fluorescence *in situ* hybridizaiton (FISH) experiments (Rybak et al., 2008) and that this accumulation might be due to 3' end modification, thereby preventing turnover. We attempted to follow the fate of pre-let-7 by Northern blotting total RNA from P19 cells after pre-treatment of the RNA with sodium periodate and alkaline Borax solution (Z. Yang et al., 2007) which results in the removal of the 3' most nucleotide (and a slight downward shift) of any unmodified RNA by a beta-elimination reaction (that is, the 3' most nucleotide has a hydroxyl moieties in the 2' and 3' positions). While the strong pre-let-7-sized band we observed was most likely an artifact caused by the cross-reaction of our locked nucleic acid (LNA) probe with one or more tRNAs (which are close in size to pre-let-7), it exhibited a strange behavior when subjected to beta-elimination: only a fraction on the pre-let-7-sized band shifted downward; pre-treating by extraction with phenol, pH 8.0 causes a full shift downward. In conclusion, we most likely witnessed the hydrolysis of one or more aminoacylted tRNAs, which are notoriously labile under even mild alkaline conditions (Stepanov & Nyborg, 2002); Trizol reagent, used for RNA islolation, contains acidified phenol (pH 4.0) does not have such a reactivity).

Regardless of the biochemical nature of this northern blot artifact, this series of studies peaked my interest in the possibility that there are many as-yet undiscovered chemical modifications of RNA and that this might have profound implications for the function of ncRNAs in vivo. My interest was compounded by recent groundbreaking research carried out by David R. Liu and colleagues, who showed in an elegant series of mass spectrometry

experiments that many small RNAs (<200nts) are covalently linked on their 5' end to small metabolites such as acetyl-CoA and NAD (Y. G. Chen et al., 2009; Kowtoniuk et al., 2009). This phenomenon was observed in bacteria, albeit two very divergent species. Thus, it will be important to investigate whether these modified RNAs exist in eukaryotes. I will persue this avenue of research in the laboratory of Dr. Gary Ruvkun at Harvard Medical School.

REFERENCES

Altuvia, Y., Landgraf, P., Lithwick, G., Elefant, N., Pfeffer, S., Aravin, A., et al. (2005). Clustering and conservation patterns of human microRNAs. *Nucleic Acids Research, 33*(8), 2697-2706. doi:10.1093/nar/gki567

Ambros, V. (1989). A hierarchy of regulatory genes controls a larva-to-adult developmental switch in C. elegans. *Cell, 57*(1), 49-57.

Ambros, V., & Horvitz, H. R. (1984). Heterochronic mutants of the nematode caenorhabditis elegans. *Science (New York, N.Y.), 226*(4673), 409-416.

Ameres, S. L., Horwich, M. D., Hung, J. H., Xu, J., Ghildiyal, M., Weng, Z., et al. (2010). Target RNA-directed trimming and tailing of small silencing RNAs. *Science (New York, N.Y.), 328*(5985), 1534-1539. doi:10.1126/science.1187058

Aravin, A., & Tuschl, T. (2005). Identification and characterization of small RNAs involved in RNA silencing. *FEBS Letters, 579*(26), 5830-5840. doi:10.1016/j.febslet.2005.08.009

Baek, D., Villen, J., Shin, C., Camargo, F. D., Gygi, S. P., & Bartel, D. P. (2008). The impact of microRNAs on protein output. *Nature, 455*(7209), 64-71. doi:10.1038/nature07242

Balzer, E., & Moss, E. G. (2007a). Localization of the developmental timing regulator Lin28 to mRNP complexes, P-bodies and stress granules. *RNA Biology, 4*(1), 16-25.

Bartel, D. P. (2004). MicroRNAs: Genomics, biogenesis, mechanism, and function. *Cell, 116*(2), 281-297.

Bartel, D. P. (2009). MicroRNAs: Target recognition and regulatory functions. *Cell, 136*(2), 215-233. doi:10.1016/j.cell.2009.01.002

Basyuk, E., Suavet, F., Doglio, A., Bordonne, R., & Bertrand, E. (2003). Human let-7 stem-loop precursors harbor features of RNase III cleavage products. *Nucleic Acids Research, 31*(22), 6593-6597.

Bates, G. J., Nicol, S. M., Wilson, B. J., Jacobs, A. M., Bourdon, J. C., Wardrop, J., et al. (2005). The DEAD box protein p68: A novel transcriptional coactivator of the p53 tumour suppressor. *The EMBO Journal, 24*(3), 543-553. doi:10.1038/sj.emboj.7600550

Blenkiron, C., Goldstein, L. D., Thorne, N. P., Spiteri, I., Chin, S. F., Dunning, M. J., et al. (2007). MicroRNA expression profiling of human breast cancer identifies new markers of tumor subtype. *Genome Biology, 8*(10), R214. doi:10.1186/gb-2007-8-10-r214

Bohnsack, M. T., Czaplinski, K., & Gorlich, D. (2004). Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. *RNA (New York, N.Y.), 10*(2), 185-191.

Bracht, J., Hunter, S., Eachus, R., Weeks, P., & Pasquinelli, A. E. (2004). Trans-splicing and polyadenylation of let-7 microRNA primary transcripts. *RNA (New York, N.Y.), 10*(10), 1586-1594. doi:10.1261/rna.7122604

Bushati, N., & Cohen, S. M. (2007). microRNA functions. *Annual Review of Cell and Developmental Biology, 23*, 175-205. doi:10.1146/annurev.cellbio.23.090506.123406

Bussing, I., Slack, F. J., & Grosshans, H. (2008). Let-7 microRNAs in development, stem cells and cancer. *Trends in Molecular Medicine, 14*(9), 400-409. doi:10.1016/j.molmed.2008.07.001

Cai, X., & Cullen, B. R. (2007). The imprinted H19 noncoding RNA is a primary microRNA precursor. *RNA (New York, N.Y.), 13*(3), 313-316. doi:10.1261/rna.351707

Cai, X., Hagedorn, C. H., & Cullen, B. R. (2004). Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. *RNA (New York, N.Y.), 10*(12), 1957-1966. doi:10.1261/rna.7135204

Calin, G. A., Dumitru, C. D., Shimizu, M., Bichi, R., Zupo, S., Noch, E., et al. (2002). Frequent deletions and down-regulation of micro- RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. *Proceedings of the National Academy of Sciences of the United States of America, 99*(24), 15524-15529. doi:10.1073/pnas.242606799

Cazalla, D., Yario, T., & Steitz, J. A. (2010). Down-regulation of a host microRNA by a herpesvirus saimiri noncoding RNA. *Science (New York, N.Y.), 328*(5985), 1563-1566. doi:10.1126/science.1187197

Chang, T. C., Yu, D., Lee, Y. S., Wentzel, E. A., Arking, D. E., West, K. M., et al. (2008). Widespread microRNA repression by myc contributes to tumorigenesis. *Nature Genetics, 40*(1), 43-50. doi:10.1038/ng.2007.30

Chapman, E. J., & Carrington, J. C. (2007). Specialization and evolution of endogenous small RNA pathways. *Nature Reviews.Genetics, 8*(11), 884-896. doi:10.1038/nrg2179

Chatterjee, S., & Grosshans, H. (2009). Active turnover modulates mature microRNA activity in caenorhabditis elegans. *Nature, 461*(7263), 546-549. doi:10.1038/nature08349

Chen, J. F., Mandel, E. M., Thomson, J. M., Wu, Q., Callis, T. E., Hammond, S. M., et al. (2006). The role of microRNA-1 and microRNA-133 in skeletal muscle proliferation and differentiation. *Nature Genetics, 38*(2), 228-233. doi:10.1038/ng1725

Chen, Y. G., Kowtoniuk, W. E., Agarwal, I., Shen, Y., & Liu, D. R. (2009). LC/MS analysis of cellular RNA reveals NAD-linked RNA. *Nature Chemical Biology, 5*(12), 879-881. doi:10.1038/nchembio.235

Chendrimada, T. P., Gregory, R. I., Kumaraswamy, E., Norman, J., Cooch, N., Nishikura, K., et al. (2005). TRBP recruits the dicer complex to Ago2 for microRNA processing and gene silencing. *Nature, 436*(7051), 740-744. doi:10.1038/nature03868

Chiang, H. R., Schoenfeld, L. W., Ruby, J. G., Auyeung, V. C., Spies, N., Baek, D., et al. (2010). Mammalian microRNAs: Experimental evaluation of novel and previously annotated genes. *Genes & Development, 24*(10), 992-1009. doi:10.1101/gad.1884710

Christodoulou, F., Raible, F., Tomer, R., Simakov, O., Trachana, K., Klaus, S., et al. (2010). Ancient animal microRNAs and the evolution of tissue identity. *Nature, 463*(7284), 1084-1088. doi:10.1038/nature08744

Davis, B. N., Hilyard, A. C., Lagna, G., & Hata, A. (2008). SMAD proteins control DROSHA-mediated microRNA maturation. *Nature, 454*(7200), 56-61. doi:10.1038/nature07086

Davis, B. N., Hilyard, A. C., Nguyen, P. H., Lagna, G., & Hata, A. (2010). Smad proteins

bind a conserved RNA sequence to promote microRNA maturation by drosha. *Molecular Cell, 39*(3), 373-384. doi:10.1016/j.molcel.2010.07.011

Denli, A. M., Tops, B. B., Plasterk, R. H., Ketting, R. F., & Hannon, G. J. (2004). Processing of primary microRNAs by the microprocessor complex. *Nature, 432*(7014), 231-235. doi:10.1038/nature03049

Du, T., & Zamore, P. D. (2005). microPrimer: The biogenesis and function of microRNA. *Development (Cambridge, England), 132*(21), 4645-4652. doi:10.1242/dev.02070

Eis, P. S., Tam, W., Sun, L., Chadburn, A., Li, Z., Gomez, M. F., et al. (2005). Accumulation of miR-155 and BIC RNA in human B cell lymphomas. *Proceedings of the National Academy of Sciences of the United States of America, 102*(10), 3627-3632. doi:10.1073/pnas.0500613102

Endoh, H., Maruyama, K., Masuhiro, Y., Kobayashi, Y., Goto, M., Tai, H., et al. (1999). Purification and identification of p68 RNA helicase acting as a transcriptional coactivator specific for the activation function 1 of human estrogen receptor alpha. *Molecular and Cellular Biology, 19*(8), 5363-5372.

Esquela-Kerscher, A., & Slack, F. J. (2006). Oncomirs - microRNAs with a role in cancer. *Nature Reviews.Cancer, 6*(4), 259-269. doi:10.1038/nrc1840

Eulalio, A., Huntzinger, E., & Izaurralde, E. (2008). Getting to the root of miRNA-mediated gene silencing. *Cell, 132*(1), 9-14. doi:10.1016/j.cell.2007.12.024

Farazi, T. A., Juranek, S. A., & Tuschl, T. (2008). The growing catalog of small RNAs and their association with distinct Argonaute/Piwi family members. *Development (Cambridge, England), 135*(7), 1201-1214. doi:10.1242/dev.005629

Fazi, F., Rosa, A., Fatica, A., Gelmetti, V., De Marchis, M. L., Nervi, C., et al. (2005). A minicircuitry comprised of microRNA-223 and transcription factors NFI-A and C/EBPalpha regulates human granulopoiesis. *Cell, 123*(5), 819-831. doi:10.1016/j.cell.2005.09.023

Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E., & Mello, C. C. (1998). Potent and specific genetic interference by double-stranded RNA in caenorhabditis elegans. *Nature, 391*(6669), 806-811. doi:10.1038/35888

Fukuda, T., Yamagata, K., Fujiyama, S., Matsumoto, T., Koshida, I., Yoshimura, K., et al. (2007). DEAD-box RNA helicase subunits of the drosha complex are required for processing of rRNA and a subset of microRNAs. *Nature Cell Biology, 9*(5), 604-611. doi:10.1038/ncb1577

Fuller-Pace, F. V. (2006). DExD/H box RNA helicases: Multifunctional proteins with important roles in transcriptional regulation. *Nucleic Acids Research, 34*(15), 4206-4215. doi:10.1093/nar/gkl460

Fuller-Pace, F. V., & Ali, S. (2008). The DEAD box RNA helicases p68 (Ddx5) and p72 (Ddx17): Novel transcriptional co-regulators. *Biochemical Society Transactions, 36*(Pt 4), 609-612. doi:10.1042/BST0360609

Ghildiyal, M., & Zamore, P. D. (2009). Small silencing RNAs: An expanding universe. *Nature Reviews.Genetics, 10*(2), 94-108. doi:10.1038/nrg2504

Gregory, R. I., Chendrimada, T. P., Cooch, N., & Shiekhattar, R. (2005). Human RISC

couples microRNA biogenesis and posttranscriptional gene silencing. *Cell, 123*(4), 631-640. doi:10.1016/j.cell.2005.10.022

Gregory, R. I., Yan, K. P., Amuthan, G., Chendrimada, T., Doratotaj, B., Cooch, N., et al. (2004). The microprocessor complex mediates the genesis of microRNAs. *Nature, 432*(7014), 235-240. doi:10.1038/nature03120

Griffiths-Jones, S., Grocock, R. J., van Dongen, S., Bateman, A., & Enright, A. J. (2006). miRBase: MicroRNA sequences, targets and gene nomenclature. *Nucleic Acids Research, 34*(Database issue), D140-4. doi:10.1093/nar/gkj112

Grimm, D., Streetz, K. L., Jopling, C. L., Storm, T. A., Pandey, K., Davis, C. R., et al. (2006). Fatality in mice due to oversaturation of cellular microRNA/short hairpin RNA pathways. *Nature, 441*(7092), 537-541. doi:10.1038/nature04791

Guil, S., & Caceres, J. F. (2007). The multifunctional RNA-binding protein hnRNP A1 is required for processing of miR-18a. *Nature Structural & Molecular Biology, 14*(7), 591-596. doi:10.1038/nsmb1250

Guo, Y., Chen, Y., Ito, H., Watanabe, A., Ge, X., Kodama, T., et al. (2006). Identification and characterization of lin-28 homolog B (LIN28B) in human hepatocellular carcinoma. *Gene, 384*, 51-61. doi:10.1016/j.gene.2006.07.011

Hafner, M., Landgraf, P., Ludwig, J., Rice, A., Ojo, T., Lin, C., et al. (2008). Identification of microRNAs and other small regulatory RNAs using cDNA library sequencing. *Methods (San Diego, Calif.), 44*(1), 3-12. doi:10.1016/j.ymeth.2007.09.009

Hammond, S. M., Bernstein, E., Beach, D., & Hannon, G. J. (2000). An RNA-directed nuclease mediates post-transcriptional gene silencing in drosophila cells. *Nature, 404*(6775), 293-296. doi:10.1038/35005107

Han, J., Lee, Y., Yeom, K. H., Kim, Y. K., Jin, H., & Kim, V. N. (2004). The drosha-DGCR8 complex in primary microRNA processing. *Genes & Development, 18*(24), 3016-3027. doi:10.1101/gad.1262504

Han, J., Lee, Y., Yeom, K. H., Nam, J. W., Heo, I., Rhee, J. K., et al. (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

He, L., Thomson, J. M., Hemann, M. T., Hernando-Monge, E., Mu, D., Goodson, S., et al. (2005). A microRNA polycistron as a potential human oncogene. *Nature, 435*(7043), 828-833. doi:10.1038/nature03552

He, X., He, L., & Hannon, G. J. (2007). The guardian's little helper: MicroRNAs in the p53 tumor suppressor network. *Cancer Research, 67*(23), 11099-11101. doi:10.1158/0008-5472.CAN-07-2672

Heo, I., Joo, C., Cho, J., Ha, M., Han, J., & Kim, V. N. (2008). Lin28 mediates the terminal uridylation of let-7 precursor MicroRNA. *Molecular Cell, 32*(2), 276-284. doi:10.1016/j.molcel.2008.09.014

Heo, I., Joo, C., Kim, Y. K., Ha, M., Yoon, M. J., Cho, J., et al. (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

Hutvagner, G., McLachlan, J., Pasquinelli, A. E., Balint, E., Tuschl, T., & Zamore, P. D.

(2001). A cellular function for the RNA-interference enzyme dicer in the maturation of the let-7 small temporal RNA. *Science (New York, N.Y.), 293*(5531), 834-838. doi:10.1126/science.1062961

Hwang, H. W., Wentzel, E. A., & Mendell, J. T. (2007). A hexanucleotide element directs microRNA nuclear import. *Science (New York, N.Y.), 315*(5808), 97-100. doi:10.1126/science.1136235

Johnson, S. M., Grosshans, H., Shingara, J., Byrom, M., Jarvis, R., Cheng, A., et al. (2005). RAS is regulated by the let-7 microRNA family. *Cell, 120*(5), 635-647. doi:10.1016/j.cell.2005.01.014

Johnson, S. M., Lin, S. Y., & Slack, F. J. (2003a). The time of appearance of the C. elegans let-7 microRNA is transcriptionally controlled utilizing a temporal regulatory element in its promoter. *Developmental Biology, 259*(2), 364-379.

Katoh, T., Sakaguchi, Y., Miyauchi, K., Suzuki, T., Kashiwabara, S., Baba, T., et al. (2009). Selective stabilization of mammalian microRNAs by 3' adenylation mediated by the cytoplasmic poly(A) polymerase GLD-2. *Genes & Development, 23*(4), 433-438. doi:10.1101/gad.1761509

Kim, V. N. (2005). MicroRNA biogenesis: Coordinated cropping and dicing. *Nature Reviews. Molecular Cell Biology, 6*(5), 376-385. doi:10.1038/nrm1644

Kim, V. N., Han, J., & Siomi, M. C. (2009a). Biogenesis of small RNAs in animals. *Nature Reviews.Molecular Cell Biology, 10*(2), 126-139. doi:10.1038/nrm2632

Kim, Y. K., & Kim, V. N. (2007). Processing of intronic microRNAs. *The EMBO Journal, 26*(3), 775-783. doi:10.1038/sj.emboj.7601512

Kloosterman, W. P., Wienholds, E., de Bruijn, E., Kauppinen, S., & Plasterk, R. H. (2006). In situ detection of miRNAs in animal embryos using LNA-modified oligonucleotide probes. *Nature Methods, 3*(1), 27-29. doi:10.1038/nmeth843

Knight, S. W., & Bass, B. L. (2001). A role for the RNase III enzyme DCR-1 in RNA interference and germ line development in caenorhabditis elegans. *Science (New York, N.Y.), 293*(5538), 2269-2271. doi:10.1126/science.1062039

Kowtoniuk, W. E., Shen, Y., Heemstra, J. M., Agarwal, I., & Liu, D. R. (2009). A chemical screen for biological small molecule-RNA conjugates reveals CoA-linked RNA. *Proceedings of the National Academy of Sciences of the United States of America, 106*(19), 7768-7773. doi:10.1073/pnas.0900528106

Krol, J., Busskamp, V., Markiewicz, I., Stadler, M. B., Ribi, S., Richter, J., et al. (2010). Characterizing light-regulated retinal microRNAs reveals rapid turnover as a common property of neuronal microRNAs. *Cell, 141*(4), 618-631. doi:10.1016/j.cell.2010.03.039

Kumar, M. S., Lu, J., Mercer, K. L., Golub, T. R., & Jacks, T. (2007). Impaired microRNA processing enhances cellular transformation and tumorigenesis. *Nature Genetics, 39*(5), 673-677. doi:10.1038/ng2003

Landthaler, M., Yalcin, A., & Tuschl, T. (2004). The human DiGeorge syndrome critical region gene 8 and its D. melanogaster homolog are required for miRNA biogenesis. *Current Biology : CB, 14*(23), 2162-2167. doi:10.1016/j.cub.2004.11.001

Lee, R. C., Feinbaum, R. L., & Ambros, V. (1993). The C. elegans heterochronic gene lin-4

encodes small RNAs with antisense complementarity to lin-14. *Cell, 75*(5), 843-854.

Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., et al. (2003). The nuclear RNase III drosha initiates microRNA processing. *Nature, 425*(6956), 415-419. doi:10.1038/nature01957

Lee, Y., Jeon, K., Lee, J. T., Kim, S., & Kim, V. N. (2002). MicroRNA maturation: Stepwise processing and subcellular localization. *The EMBO Journal, 21*(17), 4663-4670.

Lee, Y. S., & Dutta, A. (2007). The tumor suppressor microRNA let-7 represses the HMGA2 oncogene. *Genes & Development, 21*(9), 1025-1030. doi:10.1101/gad.1540407

Lee, Y. S., Kim, H. K., Chung, S., Kim, K. S., & Dutta, A. (2005). Depletion of human micro-RNA miR-125b reveals that it is critical for the proliferation of differentiated cells but not for the down-regulation of putative targets during differentiation. *The Journal of Biological Chemistry, 280*(17), 16635-16641. doi:10.1074/jbc.M412247200

Li, J., Yang, Z., Yu, B., Liu, J., & Chen, X. (2005). Methylation protects miRNAs and siRNAs from a 3'-end uridylation activity in arabidopsis. *Current Biology : CB, 15*(16), 1501-1507. doi:10.1016/j.cub.2005.07.029

Liu, J., Carmell, M. A., Rivas, F. V., Marsden, C. G., Thomson, J. M., Song, J. J., et al. (2004). Argonaute2 is the catalytic engine of mammalian RNAi. *Science (New York, N.Y.), 305*(5689), 1437-1441. doi:10.1126/science.1102513

Lu, J., Getz, G., Miska, E. A., Alvarez-Saavedra, E., Lamb, J., Peck, D., et al. (2005). MicroRNA expression profiles classify human cancers. *Nature, 435*(7043), 834-838. doi:10.1038/nature03702

Lu, Y., Thomson, J. M., Wong, H. Y., Hammond, S. M., & Hogan, B. L. (2007). Transgenic over-expression of the microRNA miR-17-92 cluster promotes proliferation and inhibits differentiation of lung epithelial progenitor cells. *Developmental Biology, 310*(2), 442-453. doi:10.1016/j.ydbio.2007.08.007

Lund, E., Guttinger, S., Calado, A., Dahlberg, J. E., & Kutay, U. (2004). Nuclear export of microRNA precursors. *Science (New York, N.Y.), 303*(5654), 95-98. doi:10.1126/science.1090599

Maniataki, E., & Mourelatos, Z. (2005). A human, ATP-independent, RISC assembly machine fueled by pre-miRNA. *Genes & Development, 19*(24), 2979-2990. doi:10.1101/gad.1384005

Marson, A., Levine, S. S., Cole, M. F., Frampton, G. M., Brambrink, T., Johnstone, S., et al. (2008). Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells. *Cell, 134*(3), 521-533. doi:10.1016/j.cell.2008.07.020

Martin, G., & Keller, W. (2007). RNA-specific ribonucleotidyl transferases. *RNA (New York, N.Y.), 13*(11), 1834-1849. doi:10.1261/rna.652807

Martinez, J., Patkaniowska, A., Urlaub, H., Luhrmann, R., & Tuschl, T. (2002). Single-stranded antisense siRNAs guide target RNA cleavage in RNAi. *Cell, 110*(5), 563-574.

Mayr, C., Hemann, M. T., & Bartel, D. P. (2007). Disrupting the pairing between let-7 and Hmga2 enhances oncogenic transformation. *Science (New York, N.Y.), 315*(5818), 1576-1579. doi:10.1126/science.1137999

Michlewski, G., Guil, S., Semple, C. A., & Caceres, J. F. (2008). Posttranscriptional regulation of miRNAs harboring conserved terminal loops. *Molecular Cell, 32*(3), 383-393. doi:10.1016/j.molcel.2008.10.013

Morlando, M., Ballarino, M., Gromak, N., Pagano, F., Bozzoni, I., & Proudfoot, N. J. (2008). Primary microRNA transcripts are processed co-transcriptionally. *Nature Structural & Molecular Biology, 15*(9), 902-909.

Moss, E. G., Lee, R. C., & Ambros, V. (1997a). The cold shock domain protein LIN-28 controls developmental timing in C. elegans and is regulated by the lin-4 RNA. *Cell, 88*(5), 637-646.

Mullen, T. E., & Marzluff, W. F. (2008). Degradation of histone mRNA requires oligouridylation followed by decapping and simultaneous degradation of the mRNA both 5' to 3' and 3' to 5'. *Genes & Development, 22*(1), 50-65. doi:10.1101/gad.1622708

Newman, M. A., & Hammond, S. M. (2010). Emerging paradigms of regulated microRNA processing. *Genes & Development, 24*(11), 1086-1092. doi:10.1101/gad.1919710

Newman, M. A., Thomson, J. M., & Hammond, S. M. (2008). Lin-28 interaction with the let-7 precursor loop mediates regulated microRNA processing. *RNA (New York, N.Y.), 14*(8), 1539-1549. doi:10.1261/rna.1155108

Obernosterer, G., Leuschner, P. J., Alenius, M., & Martinez, J. (2006b). Post-transcriptional regulation of microRNA expression. *RNA (New York, N.Y.), 12*(7), 1161-1167. doi:10.1261/rna.2322506

O'Donnell, K. A., Wentzel, E. A., Zeller, K. I., Dang, C. V., & Mendell, J. T. (2005). c-myc-regulated microRNAs modulate E2F1 expression. *Nature, 435*(7043), 839-843. doi:10.1038/nature03677

Pan, G., & Thomson, J. A. (2007). Nanog and transcriptional networks in embryonic stem cell pluripotency. *Cell Research, 17*(1), 42-49. doi:10.1038/sj.cr.7310125

Pasquinelli, A. E., Reinhart, B. J., Slack, F., Martindale, M. Q., Kuroda, M. I., Maller, B., et al. (2000). Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. *Nature, 408*(6808), 86-89. doi:10.1038/35040556

Pawlicki, J. M., & Steitz, J. A. (2008). Primary microRNA transcript retention at sites of transcription leads to enhanced microRNA production. *The Journal of Cell Biology, 182*(1), 61-76. doi:10.1083/jcb.200803111

Pawlicki, J. M., & Steitz, J. A. (2010). Nuclear networking fashions pre-messenger RNA and primary microRNA transcripts for function. *Trends in Cell Biology, 20*(1), 52-61. doi:10.1016/j.tcb.2009.10.004

Polesskaya, A., Cuvellier, S., Naguibneva, I., Duquet, A., Moss, E. G., & Harel-Bellan, A. (2007). Lin-28 binds IGF-2 mRNA and participates in skeletal myogenesis by increasing translation efficiency. *Genes & Development, 21*(9), 1125-1138. doi:10.1101/gad.415007

Poliseno, L., Salmena, L., Zhang, J., Carver, B., Haveman, W. J., & Pandolfi, P. P. (2010). A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. *Nature, 465*(7301), 1033-1038. doi:10.1038/nature09144

Preker, P., Nielsen, J., Kammler, S., Lykke-Andersen, S., Christensen, M. S., Mapendano, C. K., et al. (2008). RNA exosome depletion reveals transcription upstream of active

human promoters. *Science (New York, N.Y.), 322*(5909), 1851-1854. doi:10.1126/science.1164096

Ramaswamy, S., Tamayo, P., Rifkin, R., Mukherjee, S., Yeang, C. H., Angelo, M., et al. (2001). Multiclass cancer diagnosis using tumor gene expression signatures. *Proceedings of the National Academy of Sciences of the United States of America, 98*(26), 15149-15154. doi:10.1073/pnas.211566398

Reinhart, B. J., Slack, F. J., Basson, M., Pasquinelli, A. E., Bettinger, J. C., Rougvie, A. E., et al. (2000). The 21-nucleotide let-7 RNA regulates developmental timing in caenorhabditis elegans. *Nature, 403*(6772), 901-906. doi:10.1038/35002607

Richards, M., Tan, S. P., Tan, J. H., Chan, W. K., & Bongso, A. (2004). The transcriptome profile of human embryonic stem cells as defined by SAGE. *Stem Cells (Dayton, Ohio), 22*(1), 51-64. doi:10.1634/stemcells.22-1-51

Rudnicki, M. A., Sawtell, N. M., Reuhl, K. R., Berg, R., Craig, J. C., Jardine, K., et al. (1990). Smooth muscle actin expression during P19 embryonal carcinoma differentiation in cell culture. *Journal of Cellular Physiology, 142*(1), 89-98. doi:10.1002/jcp.1041420112

Rybak, A., Fuchs, H., Smirnova, L., Brandt, C., Pohl, E. E., Nitsch, R., et al. (2008a). A feedback loop comprising lin-28 and let-7 controls pre-let-7 maturation during neural stem-cell commitment. *Nature Cell Biology, 10*(8), 987-993. doi:10.1038/ncb1759

Sakamoto, S., Aoki, K., Higuchi, T., Todaka, H., Morisawa, K., Tamaki, N., et al. (2009). The NF90-NF45 complex functions as a negative regulator in the microRNA processing pathway. *Molecular and Cellular Biology, 29*(13), 3754-3769. doi:10.1128/MCB.01836-08

Schulman, B. R., Esquela-Kerscher, A., & Slack, F. J. (2005). Reciprocal expression of lin-41 and the microRNAs let-7 and mir-125 during mouse embryogenesis. *Developmental Dynamics : An Official Publication of the American Association of Anatomists, 234*(4), 1046-1054. doi:10.1002/dvdy.20599

Seitz, H. (2009). Redefining microRNA targets. *Current Biology : CB, 19*(10), 870-873. doi:10.1016/j.cub.2009.03.059

Selbach, M., Schwanhausser, B., Thierfelder, N., Fang, Z., Khanin, R., & Rajewsky, N. (2008). Widespread changes in protein synthesis induced by microRNAs. *Nature, 455*(7209), 58-63. doi:10.1038/nature07228

Sempere, L. F., Freemantle, S., Pitha-Rowe, I., Moss, E., Dmitrovsky, E., & Ambros, V. (2004). Expression profiling of mammalian microRNAs uncovers a subset of brain-expressed microRNAs with possible roles in murine and human neuronal differentiation. *Genome Biology, 5*(3), R13. doi:10.1186/gb-2004-5-3-r13

Shen, B., & Goodman, H. M. (2004). Uridine addition after microRNA-directed cleavage. *Science (New York, N.Y.), 306*(5698), 997. doi:10.1126/science.1103521

Shi, R., & Chiang, V. L. (2005). Facile means for quantifying microRNA expression by real-time PCR. *BioTechniques, 39*(4), 519-525.

Stepanov, V. G., & Nyborg, J. (2002). Thermal stability of aminoacyl-tRNAs in aqueous solutions. *Extremophiles : Life Under Extreme Conditions, 6*(6), 485-490. doi:10.1007/s00792-002-0285-4

Suh, M. R., Lee, Y., Kim, J. Y., Kim, S. K., Moon, S. H., Lee, J. Y., et al. (2004). Human embryonic stem cells express a unique set of microRNAs. *Developmental Biology, 270*(2), 488-498. doi:10.1016/j.ydbio.2004.02.019

Suzuki, H. I., Yamagata, K., Sugimoto, K., Iwamoto, T., Kato, S., & Miyazono, K. (2009). Modulation of microRNA processing by p53. *Nature, 460*(7254), 529-533. doi:10.1038/nature08199

Takamizawa, J., Konishi, H., Yanagisawa, K., Tomida, S., Osada, H., Endoh, H., et al. (2004). Reduced expression of the let-7 microRNAs in human lung cancers in association with shortened postoperative survival. *Cancer Research, 64*(11), 3753-3756. doi:10.1158/0008-5472.CAN-04-0637

Thomson, J. M., Newman, M., Parker, J. S., Morin-Kensicki, E. M., Wright, T., & Hammond, S. M. (2006a). Extensive post-transcriptional regulation of microRNAs and its implications for cancer. *Genes & Development, 20*(16), 2202-2207. doi:10.1101/gad.1444406

Thomson, J. M., Parker, J., Perou, C. M., & Hammond, S. M. (2004a). A custom microarray platform for analysis of microRNA gene expression. *Nature Methods, 1*(1), 47-53. doi:10.1038/nmeth704

Trabucchi, M., Briata, P., Garcia-Mayoral, M., Haase, A. D., Filipowicz, W., Ramos, A., et al. (2009). The RNA-binding protein KSRP promotes the biogenesis of a subset of microRNAs. *Nature, 459*(7249), 1010-1014. doi:10.1038/nature08025

Trippe, R., Guschina, E., Hossbach, M., Urlaub, H., Luhrmann, R., & Benecke, B. J. (2006). Identification, cloning, and functional analysis of the human U6 snRNA-specific terminal uridylyl transferase. *RNA (New York, N.Y.), 12*(8), 1494-1504. doi:10.1261/rna.87706

Ule, J., Jensen, K., Mele, A., & Darnell, R. B. (2005). CLIP: A method for identifying protein-RNA interaction sites in living cells. *Methods (San Diego, Calif.), 37*(4), 376-386. doi:10.1016/j.ymeth.2005.07.018

Vigneault, F., Sismour, A. M., & Church, G. M. (2008). Efficient microRNA capture and bar-coding via enzymatic oligonucleotide adenylation. *Nature Methods, 5*(9), 777-779. doi:10.1038/nmeth.1244

Viswanathan, S. R., & Daley, G. Q. (2010). Lin28: A microRNA regulator with a macro role. *Cell, 140*(4), 445-449. doi:10.1016/j.cell.2010.02.007

Viswanathan, S. R., Daley, G. Q., & Gregory, R. I. (2008). Selective blockade of microRNA processing by Lin28. *Science (New York, N.Y.), 320*(5872), 97-100. doi:10.1126/science.1154040

Wang, Y., Baskerville, S., Shenoy, A., Babiarz, J. E., Baehner, L., & Blelloch, R. (2008). Embryonic stem cell-specific microRNAs regulate the G1-S transition and promote rapid proliferation. *Nature Genetics, 40*(12), 1478-1483. doi:10.1038/ng.250

Wang, Y., Medvid, R., Melton, C., Jaenisch, R., & Blelloch, R. (2007). DGCR8 is essential for microRNA biogenesis and silencing of embryonic stem cell self-renewal. *Nature Genetics, 39*(3), 380-385. doi:10.1038/ng1969

Warner, D. R., Bhattacherjee, V., Yin, X., Singh, S., Mukhopadhyay, P., Pisano, M. M., et al. (2004). Functional interaction between smad, CREB binding protein, and p68 RNA

helicase. *Biochemical and Biophysical Research Communications, 324*(1), 70-76. doi:10.1016/j.bbrc.2004.09.017

Wienholds, E., Kloosterman, W. P., Miska, E., Alvarez-Saavedra, E., Berezikov, E., de Bruijn, E., et al. (2005). MicroRNA expression in zebrafish embryonic development. *Science (New York, N.Y.), 309*(5732), 310-311. doi:10.1126/science.1114519

Wienholds, E., & Plasterk, R. H. (2005). MicroRNA function in animal development. *FEBS Letters, 579*(26), 5911-5922. doi:10.1016/j.febslet.2005.07.070

Wightman, B., Ha, I., & Ruvkun, G. (1993). Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans. *Cell, 75*(5), 855-862.

Wilusz, C. J., & Wilusz, J. (2008). New ways to meet your (3') end oligouridylation as a step on the path to destruction. *Genes & Development, 22*(1), 1-7. doi:10.1101/gad.1634508

Wu, L., & Belasco, J. G. (2005). Micro-RNA regulation of the mammalian lin-28 gene during neuronal differentiation of embryonal carcinoma cells. *Molecular and Cellular Biology, 25*(21), 9198-9208. doi:10.1128/MCB.25.21.9198-9208.2005

Wulczyn, F. G., Smirnova, L., Rybak, A., Brandt, C., Kwidzinski, E., Ninnemann, O., et al. (2007). Post-transcriptional regulation of the let-7 microRNA during neural cell specification. *The FASEB Journal : Official Publication of the Federation of American Societies for Experimental Biology, 21*(2), 415-426. doi:10.1096/fj.06-6130com

Yamagata, K., Fujiyama, S., Ito, S., Ueda, T., Murata, T., Naitou, M., et al. (2009). Maturation of microRNA is hormonally regulated by a nuclear receptor. *Molecular Cell, 36*(2), 340-347. doi:10.1016/j.molcel.2009.08.017

Yang, D. H., & Moss, E. G. (2003). Temporally regulated expression of lin-28 in diverse tissues of the developing mouse. *Gene Expression Patterns : GEP, 3*(6), 719-726.

Yang, W., Chendrimada, T. P., Wang, Q., Higuchi, M., Seeburg, P. H., Shiekhattar, R., et al. (2006). Modulation of microRNA processing and expression through RNA editing by ADAR deaminases. *Nature Structural & Molecular Biology, 13*(1), 13-21. doi:10.1038/nsmb1041

Yang, Z., Vilkaitis, G., Yu, B., Klimasauskas, S., & Chen, X. (2007). Approaches for studying microRNA and small interfering RNA methylation in vitro and in vivo. *Methods in Enzymology, 427*, 139-154. doi:10.1016/S0076-6879(07)27008-9

Yi, R., Qin, Y., Macara, I. G., & Cullen, B. R. (2003). Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes & Development, 17*(24), 3011-3016. doi:10.1101/gad.1158803

Yu, J., Vodyanik, M. A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J. L., Tian, S., et al. (2007). Induced pluripotent stem cell lines derived from human somatic cells. *Science (New York, N.Y.), 318*(5858), 1917-1920. doi:10.1126/science.1151526

Zhao, Y., Ransom, J. F., Li, A., Vedantham, V., von Drehle, M., Muth, A. N., et al. (2007). Dysregulation of cardiogenesis, cardiac conduction, and cell cycle in mice lacking miRNA-1-2. *Cell, 129*(2), 303-317. doi:10.1016/j.cell.2007.03.030

Zhao, Y., Samal, E., & Srivastava, D. (2005). Serum response factor regulates a muscle-specific microRNA that targets Hand2 during cardiogenesis. *Nature, 436*(7048),

214-220. doi:10.1038/nature03817