

Exploring the regulatory roles of microRNAs in mammalian development

by

Grace Xinying Zheng

B.Sc. University of British Columbia
Vancouver, Canada, 2004

SUBMITTED TO THE PROGRAM OF COMPUTATIONAL AND SYSTEMS BIOLOGY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTORATE OF PHILOSOPHY

AT THE

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

MAY 2010

© 2010 Massachusetts Institute of Technology
All rights reserved

Signature of Author.....
Grace X.Y. Zheng
Computational and Systems Biology
May 21, 2010

Certified by.....
Phillip A. Sharp
Institute Professor of Biology
Thesis Supervisor

Certified by.....
Christopher B. Burge
Whitehead Career Development Associate Professor of Biology
Thesis Supervisor

Accepted by.....
Christopher B. Burge
Whitehead Career Development Associate Professor of Biology
Chair, CSB Graduate Committee

Exploring the regulatory roles of microRNAs in mammalian development

By

Grace Xinying Zheng

Abstract

microRNAs (miRNAs) are ~22-nt long short RNAs that regulate gene expression in organisms ranging from plants to animals. In mammals, miRNAs post-transcriptionally repress gene expression by primarily binding to the 3' untranslated region (3' UTR) of target mRNAs. Although hundreds of miRNAs have been discovered, targets of most miRNAs and the method by which they affect their biological function remain elusive. To better understand the role of miRNAs in fundamental cellular processes, we characterized enriched miRNA populations in three distinct murine developmental programs, T lymphocytes, embryonic stem cells, and the placenta.

We started exploring the role of miRNAs in T lymphocytes by globally characterizing short RNA expression during key developmental stages of T lymphocytes. Our results showed that a distinct set of miRNAs is enriched in each stage. In particular, miR-181 is elevated at the double positive (DP) stage, when thymocytes expressing both CD4 and CD8 undergo positive and negative selection. We found that miR-181 can repress the expression of Bcl-2, CD69, and the T cell receptor, all of which are involved in positive selection.

Analysis of short RNAs in T lymphocytes also revealed a novel miRNA cluster, the *Sfmbt2* miRNA cluster, named as such since it maps to an intron of the *Sfmbt2* gene, a Polycomb Group gene. Instead of studying this cluster in T lymphocytes, we decided to use embryonic stem (ES) cells as this cluster is also expressed in ES cells and the cells are more conducive to lab experimentation. This cluster contains several miRNA families, and we addressed the function of one miRNA family, miR-467a, as it shares target specificity with other highly abundant miRNAs in ES cells. Gain and loss of function assays showed that this family of miRNAs can promote cell survival by advancing the G1 to S phase transition. In addition, they target certain proapoptotic factors to buffer ES cells from apoptosis, especially in the context of genotoxic stress.

The *Sfmbt2* cluster is a mouse-specific miRNA cluster, and individual members have been uniquely amplified in the *Sfmbt2* locus. We developed a method to explore the impact of species-specific miRNAs on the evolution of 3' UTRs, and found that target sites of many miRNAs show positive selection. In particular, mouse target sites have evolved to specifically gain binding sites (mouse-specific targets) for some *Sfmbt2* miRNAs, several of which are enriched in the placenta. These mouse-specific targets are enriched in pathways regulating cell survival, implicating the *Sfmbt2* miRNA cluster as a possible promoter to placental growth.

Our studies in T lymphocytes, ES cells and the placenta have revealed important roles of miRNAs in shaping 3' UTR evolution, and mammalian development. Several novel miRNA targets we uncovered are important regulators of differentiation, cell cycle, and apoptosis. Understanding their functions will not only shed light on their roles in normal physiology, but also generate useful insights that can be applied to cancer and reprogramming.

Thesis Supervisors:

Phillip A. Sharp, Institute Professor of Biology

Christopher B. Burge, Whitehead Career Development Associate Professor of Biology

Table of Contents

Abstract	2
Table of Contents.....	3
Acknowledgments	7
CHAPTER 1: Introduction.....	10
Discovery of miRNA genes	12
Genomic organization of miRNAs	13
Evolution of miRNAs	14
Biogenesis of miRNAs	15
Regulation of miRNA biogenesis.....	17
Mechanisms of miRNA-directed silencing	19
<i>Argonautes and GW182 proteins.....</i>	<i>19</i>
<i>Translational inhibition by miRNAs</i>	<i>20</i>
<i>Degradation of mRNAs by miRNAs.....</i>	<i>21</i>
<i>P bodies and stress granules.....</i>	<i>23</i>
Principles of miRNA-mRNA interactions	24
Roles of miRNAs in mammalian development	27
<i>miRNAs in ES cells</i>	<i>29</i>
T lymphocyte development	32
Placental development.....	33
Figure 1. Differentiation of totipotent cells.	35
Figure 2. T lymphocyte development.	36
CHAPTER 2: Dynamic regulation of miRNA expression in ordered stages of cellular development	37
Abstract	38
Introduction	39
Results.....	41
<i>Short RNA Profiling Using Low Nanogram Amounts of Total RNA.....</i>	<i>41</i>
<i>Quantitative Validation of Clone Representation</i>	<i>43</i>
<i>Dynamic Regulation of miRNA Species During Thymocyte Development.....</i>	<i>44</i>
<i>Genomic Signatures of miRNA Families at Individual Stages of Thymocyte Development..</i>	<i>47</i>
<i>miR-181a Represses Expression of Genes with Roles in Thymocyte Maturation Through Their 3' UTR Elements.....</i>	<i>49</i>
<i>Dysregulation of miR-181a Targets in Dicer-deficient DP Thymocytes</i>	<i>51</i>
<i>Modulation of Potential miR-181a Targets by Transfection of Nucleic Acids into Dicer-deficient DP Thymocytes</i>	<i>52</i>
Discussion	52
Methods.....	58
<i>Mice.....</i>	<i>58</i>
<i>Thymocyte and RNA isolation.....</i>	<i>58</i>
<i>Short RNA Cloning.....</i>	<i>58</i>
<i>Direct miRNA quantitation and calculation of miRNA copy number per cell</i>	<i>59</i>
<i>Phylogenetic Analysis.....</i>	<i>60</i>
<i>qRT-PCR Analysis of ADAR Expression.....</i>	<i>60</i>
<i>Luciferase assays</i>	<i>60</i>

<i>Transfection of DP thymocytes</i>	61
Table 1. Short RNA cloning statistic	62
Table 2. Calculation of copy numbers of individual miRNAs in CD4 ⁺ CD8 ⁺ thymocytes.	63
Figure 1. Modified cloning protocol.....	64
Figure 2. Processing and annotation flowchart for cloned sequences.	65
Figure 3. Schematic of longitudinal miRNA quantitation experiments	67
Figure 4. Absolute copy numbers of miRNAs in Thymocyte populations.....	69
Figure 5. Measurements of total RNA content per cell.	70
Figure 6. miRNA expression is dynamically regulated during thymocyte development.	71
Figure 7. Genomic signatures in response to changes in miRNA family member expression.	72
Figure 8. Repression of predicted miR-181a target UTRs in a reporter assay.....	75
Figure 9. Dysregulation of CD69 and the TCR in Dicer-deficient DP thymocytes.	76
Figure 10. CD69 expression is inversely correlated with GFP intensity in <i>Dicer</i> - deficient and control DP thymocytes transfected with miR-181a.	77
CHAPTER 3: Multiple abundant miRNA families collaborate to buffer embryonic stem cells from apoptosis.....	78
Abstract	79
Introduction	80
Results.....	82
<i>Predicted targets of AAGUGC containing miRNAs are enriched in pathways regulating apoptosis.</i>	82
<i>Caspase 2 and Ei24, key apoptotic mediators, are direct targets of both the miR-290 and miR-467 families.</i>	83
<i>AAGUGC containing miRNAs buffer ES cells against genotoxic stress.</i>	85
<i>Deletion of miRNAs containing AAGUGC seed makes cells more susceptible to apoptosis upon DNA damage.</i>	87
Discussion	88
Methods.....	92
<i>ES Cell Culture</i>	92
<i>Oligos and siRNAs used in all the experiments</i>	92
<i>Generation of luciferase constructs, mESC transfection, and luciferase assays</i>	93
<i>Northern Blot analysis</i>	94
<i>Western Blot analysis</i>	95
<i>RT-PCR</i>	95
<i>Transfection and BrdU assays</i>	95
<i>Transfection and Casp3 assays</i>	96
<i>AnnexinV assays</i>	97
<i>Microarray analysis</i>	97
<i>Gene Ontology analysis</i>	98
<i>Statistical analyses</i>	98
Table 1. Description of miRNAs in the AAGUGC family.	99
Figure 1. Repressive effect of miR-467a-5p and miR-20a on their control reporter constructs.	100
Figure 2. mRNA targets of AAGUGC miRNAs show decreased stability in Dcr KO cells and their luciferase targets can be repressed.....	101
Figure 3. Repression of predicted AAGUGC miRNA targets in WT and Dcr KO ES cells.	103
Figure 4. Repression of Casp2 and Ei24 by AAGUGC miRNAs.....	105

Figure 5. Repression of Casp2 and Ei24 in WT ES cells, as well as miR-290-295 KO ES cells.....	107
Figure 6. AAGUGC miRNAs protect ES cells from radiation-induced apoptosis.....	108
Figure 7. AAGUGC miRNAs protect ES cells from doxorubicin-induced apoptosis.	109
Figure 8. Expression of AAGUGC miRNA family stays the same before and after doxorubicin treatment.....	110
Figure 9. Expression of AAGUGC miRNAs in 295 KO ES cells before and after genotoxic stress.....	111
Figure 10. AAGUGC miRNAs protect ES cells from radiation-induced apoptosis....	112
Figure 11. AAGUGC miRNAs protect ES cells from doxorubicin-induced apoptosis.	113
Figure 12. Comparison of WT and Dcr KO cells' apoptosis response to stress with AnnexinV assay.....	114
Figure 13. AAGUGC miRNAs can promote G1 to S phase transition of Dcr KO ES cells.....	115

CHAPTER 4: Characterization of the *Sfmbt2* miRNA cluster and its function in murine placental development116

Abstract	117
Introduction	118
Results.....	119
<i>Expression of repeat-derived Sfmbt2 miRNA cluster is upregulated in murine placenta...</i>	119
<i>Predicted targets of Sfmbt2 miRNAs are enriched in pathways regulating cell growth and apoptosis.....</i>	121
<i>Target sites of many species-specific miRNAs are positively selected to lose binding sites in their 3' UTRs.....</i>	123
Discussion	125
<i>ES Cell Culture.....</i>	129
<i>Oligos and siRNAs used in all the experiments.....</i>	129
<i>Generation of luciferase constructs, mESC transfection, and luciferase assays.....</i>	129
<i>Northern Blot analysis.....</i>	130
<i>Microarray analysis.....</i>	131
<i>Bioinformatics analysis.....</i>	131
<i>Analysis of positive and purifying selection on human and mouse 3' UTRs</i>	132
<i>Gene Ontology analysis.....</i>	132
<i>Statistical analysis.....</i>	132
Table 1. Description of the <i>Sfmbt2</i> miRNA cluster.....	134
Table 2. Target statistics of <i>Sfmbt2</i> miRNAs.....	135
Table 3. Summary statistics of positive selection signals detected from heptamer binding sites in mouse and human 3' UTRs.....	136
Table 4. Target sites of 32 human species-specific miRNAs are under positive selection for site loss.....	137
Figure 1. Genomic structure of the <i>Sfmbt2</i> miRNA cluster and its expression in placenta.....	138
Figure 2. Multiple sequence alignments of <i>Sfmbt2</i> miRNAs.....	139
Figure 3. Predicted targets of <i>Sfmbt2</i> miRNAs show decreased stability in d13.5 placenta.....	140
Figure 4. Repression of <i>Dedd2</i> by two of <i>Sfmbt2</i> miRNAs.....	141
Figure 5. Illustration of the positive selection analysis for site gain of a mouse-specific miRNA.....	142
Appendix: Exploring miRNA targets that are under purifying selection.....	144

<i>Table 1. Summary statistics of purifying selection signals detected from heptamer binding sites in mouse and human 3' UTRs</i>	145
<i>Figure 2. Illustration of the purifying selection analysis</i>	146

CHAPTER 5: Conclusions and future directions.....147

miRNAs in T lymphocyte development.....	149
miRNAs in ES cells	151
miRNAs in placental development	152
Figure 1. Caspase 3 activation in Mouse Sarcoma and ES cells before and after stress induction.	157

CHAPTER 6: References158

Acknowledgments

The work described in this dissertation would not have been possible without the guidance and support from my colleagues, friends and family. First, I want to thank my advisors Chris Burge and Phil Sharp. Chris, you are a brilliant computational scientist, and you taught me how to use the right controls to answer the right questions. Phil, thank you for giving me the opportunity to work in the lab, and allowing me the freedom to create and explore. Your dedication, perseverance, and passion for science will always be a source of inspiration for me.

To my colleagues in the Sharp lab. You have made the lab an especially fun and friendly place to work in. Joel, you have taught me all the experimental techniques I know now. You are like a third advisor to me, and I am indebted to you for my intellectual development. I want to thank my collaborators, Joel (studying miRNAs in T cell development), Arvind (studying miRNAs in ES cells and the placenta), Amanda, Andrew, and Anthony (Ago2-CLIP project), Margaret (miRNA repression threshold project), and Alla (studying 21Us and chromatin regulation). You are a group of the smartest and most dedicated people I have ever worked with. I have learned tremendously from you, and I hope we will continue to collaborate in our future scientific endeavors. I am grateful to three great baymates, Amy, Jesse and Ryan. Amy, thank you for all your constructive criticism, and helping me become a more thorough and methodical experimentalist. Jesse, for your friendship, and for helping me improve my thesis and my presentation. Ryan, for your company and investment tips. In addition, thank you Amanda, Anthony, Allan, Lourdes, and Margaret for reading over my thesis and helping me with my presentation. To Lourdes, for your friendship, and patience. I will miss our crafting sessions and wonderful lunch hours we shared. To Mary, for your wonderful skills in Photoshop and Illustrator, and to Margarita, for your encouragement, for helping me with my internship, and for keeping the lab running smoothly. We are incredibly fortunate to have you.

To my colleagues in the Burge lab, Brad, Robin, and Cydney. I have benefited from our conversations from generating controls, to figuring out the best ways to evaluate the significance of a result. Brad, you are a great friend and a wonderful colleague. Thank you for answering my questions on R and statistics, and for believing in me.

To the members of my thesis committee, Laurie Boyer, Dave Bartel and Victor Ambros, for their thoughtful advice and encouragement. Laurie, thank you for your encouragement. You are a great role model to me. I would also like to thank the following funding agencies and fellowships, NSERC and NIH, for funding my graduate work.

I am among the first class of CSB program, and I want to thank Bruce Tidor, the first director of the program, for giving me the opportunity to study at MIT, and for guidance on my career. To my fellow classmates, Darlene, and Bonnielee, for making sure all the deadlines are met.

To my wonderful friends at and outside of MIT, Emily, Carol, Amy, Jane, Sally, Chris, Song-hee, Andy, Stever, Sezen, and Ben, thank you! I am incredibly fortunate to

have found such an amazing group of friends who have been here for me through many trying times and who provided much laughter throughout our time together at MIT. Without you and your company, I would not have been able to survive graduate school! To Carol, for being my exercise buddy, and for your encouragement and for being such a reliable friend. To Chris, for your undying optimism, and for listening. To Emily, we make the tastest pizzas, and I will always remember your laughter. To Sezen, you are my first roommate, and thank you for your support, sense of humor, and friendship. My experience at MIT would have been very different without you. To my girlfriends Song-hee, Jane, Amy, and Sally for listening and cooking with me.

I am also full of gratitude to my best friend Jeremy. Your perfectionism has made this thesis much better. Your dedication to your work and friends have inspired me to become a better scientist and person. I feel incredibly fortunate to have you in my life.

This work would truly not have been possible without the support provided by my family. To my parents, and my extended family in China. I am truly at a loss for words to express how thankful I am for all the sacrifices you have made to give me the educational opportunities that I have been able to enjoy throughout my life. I am deeply touched and thankful for your unending support and love. I would never have been able to accomplish this huge undertaking without you. Let this dissertation be a testament to the love, sacrifice, and support you have poured into me for the last thirty years (almost).

To my parents

CHAPTER 1: Introduction

Short RNAs participate in almost every aspect of eukaryotic biology through translational repression, mRNA degradation and chromatin modification. miRNAs are a class of short RNAs that regulate gene expression in organisms ranging from plants to mammals. To date, hundreds of miRNAs have been identified, and they are shown to regulate a diverse array of cellular functions, such as differentiation, proliferation, apoptosis, and metabolism, by post-transcriptionally regulating gene expression. In Chapter 1, we review the key findings in miRNA biogenesis, mechanisms and principles of miRNA targeting, and focus on the important role of miRNAs in ES cells. At the end of the chapter we also include an introduction on T lymphocyte and placental development to provide appropriate background for the research described in the thesis.

Discovery of miRNA genes

Since the discovery of RNA interference (RNAi), efforts to identify endogenous small RNAs have led to the discovery of hundreds of miRNAs in plants and animals (Bartel 2004). Over 400 miRNAs have been confidently identified in humans, a number that approaches 2% of protein coding genes (Landgraf et al. 2007). To date, three types of approaches have been used to identify miRNA genes: genetic screening, short RNA cloning, and bioinformatics.

The first miRNAs, *lin-4* and *let-7*, were discovered through genetic screens, as their mutations led to a defect in the timing of *C. elegans* development (Lee et al. 1993; Reinhart et al. 2000). Genetic screens have revealed more miRNAs in several model organisms, notably *bantam* in *Drosophila*, and *Isy-6* in *C. elegans* (Brennecke et al. 2003; Johnston and Hobert 2003).

Meanwhile, massive cloning efforts have taken center stage in uncovering miRNAs as well as other classes of short RNAs. To clone endogenous small RNAs, size-fractionated RNA is ligated to 5' and 3' adaptor molecules, then reverse transcribed and amplified by PCR to construct the cDNA library for sequencing (Kim and Nam 2006). The recent introduction of deep sequencing technology has enabled simultaneous sequencing of up to millions of DNA molecules (Shendure and Ji 2008). After sequencing, raw sequence reads are filtered based on quality, and mapped to various non-coding databases as well as genomic sequences. This led to the discovery of not only miRNAs, but also piRNAs (Aravin et al. 2006; Girard et al. 2006; Grivna et al. 2006; Lau et al. 2006; Watanabe et al. 2006), endo-siRNAs (Tam et al. 2008; Watanabe et al. 2008; Lau et al. 2009), and TSSa-RNAs (Core et al. 2008; Preker et al. 2008; Seila et al. 2008), small RNAs that overlap with transcription start sites in mammalian cells.

Bioinformatics has helped identify novel miRNAs by detecting sequence

conservation and predicting hairpin structures (Bartel 2004). The first miRNA search algorithm was miRScan (Lim et al. 2003b), which searched for miRNA-like features and conservation patterns in hairpin sequences that showed homology in two nematode species. The method was also extended to successfully identify miRNAs in vertebrates (Lim et al. 2003a). Through comparative analysis of the human, mouse, rat and dog genomes, Xie *et. al.* catalogued common regulatory motifs in promoters and 3' UTRs (Xie et al. 2005). Many of the motifs in 3' UTRs are associated with miRNAs, leading them to predict 129 novel miRNAs, many of which have subsequently been validated. Berezikov *et. al.* used the characteristic conservation profile around miRNA genes (high conservation in the stem region relative to flanking sequences) to discover novel miRNAs through cross-species comparison (Berezikov et al. 2005). Bentwich *et. al.* used an integrative approach to combine bioinformatics predictions with microarray analysis and sequence-directed cloning (Bentwich et al. 2005). They identified 89 novel human miRNAs, of which the majority can be mapped to two nonconserved clusters. One of them is expressed only from chromosome 19 in the placenta, whereas the other is found on chromosome X, and has been implicated in regulating testis development and spermatogenesis (Zhang et al. 2007).

Genomic organization of miRNAs

miRNAs are derived from hairpin precursors that are encoded in the genome (Kim et al. 2009). Approximately 80% of miRNA genes are found in intronic regions of protein coding and non-coding transcriptional units, and the rest are distributed in exonic and intergenic regions (Kim et al. 2009). About half of known miRNAs are found in clusters, and they are transcribed as polycistronic primary transcripts (Kim et al. 2009). However, not all miRNAs in the same cluster are expressed at the same level, suggesting that miRNAs may be post-transcriptionally processed on an individual basis

(Calabrese et al. 2007; Neilson et al. 2007; Babiarz et al. 2008; Ventura et al. 2008).

Members of miRNA clusters tend to have similar sequences, which allow them to regulate a common set of transcripts (Ambros 2004; Bartel 2004). An example of a functionally related cluster is the miR-290-295 cluster, which is specifically expressed in ES cells (Houbaviy et al. 2003). This cluster's role in maintaining stem cell pluripotency and development will be discussed in detail in the section "miRNAs in ES cells".

Evolution of miRNAs

miRNA creation and expansion have been linked to major developmental innovations. Hertel and colleagues have documented three episodes of miRNA creation that coincide with metazoan evolution: the advent of bilaterians, the rise of vertebrates, and the emergence of placental mammals (Hertel et al. 2006). About 30 miRNAs, including *let-7* and miR-1, are shared among bilaterians (Hertel et al. 2006; Niwa and Slack 2007). Higher number of miRNAs in an organism is generally associated with a more complex body plan and development. For example, about 800 miRNAs have been identified in primates, and about half of them are primate-specific (Bentwich et al. 2005).

Three mechanisms have been proposed to explain the evolution of animal miRNAs (Shabalina and Koonin 2008). First, many miRNAs appear to be derived from repeats and transposons, which comprise an especially large fraction of mammalian genomes. About 20% of human miRNAs share sequences with transposable elements, and a recent systematic analysis revealed that 55 miRNAs originated from LINE and SINE elements (Aravin et al. 2001; Smalheiser and Torvik 2006; Piriyaopongsa et al. 2007). Second, many clustered miRNAs are thought to have evolved through duplication of an existing miRNA followed by mutations in the target recognition region (Shabalina and Koonin 2008). Last, some miRNAs could arise from random hairpin structures that are embedded in the transcribed part of genomes. Genome-wide bioinformatics screens

showed that the human genome encodes millions of potential hairpins (Bentwich et al. 2005). New miRNAs with target specificities can emerge from the pool of hairpin transcripts via random mutations. miRNAs that confer beneficial regulations to their target genes can then be maintained through purifying selection (Liu et al. 2008; Shabalina and Koonin 2008).

Biogenesis of miRNAs

Most miRNA genes are transcribed by RNA polymerase II to generate primary transcripts (pri-miRNAs) that are several kilobases long, and are 5' capped and 3' polyadenylated (Cai et al. 2004; Lee et al. 2004; Kim et al. 2009). A typical mammalian primary miRNA (pri-miRNA) contains a stem of ~33 bp, a terminal loop and two single-stranded flanking regions (ssRNA). The stem and ssRNA segments are recognized by DGCR8, a protein that contains double-stranded RNA (dsRNA) binding domains. The stem is cleaved ~11 bp away from the ssRNA-dsRNA junctions by DROSHA, an RNase III type protein (Lee et al. 2002; Lee et al. 2003; Zeng and Cullen 2005; Han et al. 2006). Recent evidence suggests that DROSHA-mediated cleavage of pri-miRNA and splicing can occur co-transcriptionally and are highly coordinated (Kim and Kim 2007; Morlando et al. 2008; Pawlicki and Steitz 2008). Pri-miRNA processing does not inhibit splicing, and in many cases precedes the splicing of the host intron.

There is also a small group of miRNA genes, mirtrons, that are embedded in short introns, and their biogenesis is independent of Drosha (Berezikov et al. 2007; Okamura et al. 2007; Ruby et al. 2007; Babiarz et al. 2008). After splicing, the intron can form a hairpin resembling a miRNA precursor (pre-miRNA), whose 5' and 3' ends are trimmed by exonucleases (Kim et al. 2009). Pre-miRNAs are then exported from the nucleus by exportin 5 with Ran-GTP (Yi et al. 2003; Lund et al. 2004).

Once in the cytoplasm, pre-miRNAs are cleaved near the terminal loop by

DICER, another RNase III enzyme (Bernstein et al. 2001; Grishok et al. 2001; Hutvagner et al. 2001; Ketting et al. 2001; Knight and Bass 2001; Yi et al. 2003; Lund et al. 2004). DICER is characterized by an amino-terminal DEXD/H-box domain, a DUF283 domain, a PAZ domain, and two RNase III domains as well as a dsRNA-binding domain. The PAZ domain binds to the 3' end of pre-miRNA, and active sites at each of the RNaseIII domains cleave one of the two strands, generating a miRNA duplex (mature miRNAs) with 5' phosphates and 2-nt 3' overhang (Song et al. 2003; Yan et al. 2003; Lingel et al. 2004; Ma et al. 2004; Du et al. 2008). In human cells, DICER interacts with two double-stranded RNA-binding domain proteins TRBP and PACT, which enhance DICER-mediated cleavage of pre-miRNAs (Chendrimada et al. 2005; Haase et al. 2005; Lee et al. 2006). The ends of mature miRNAs are often heterogeneous, which could be the result of imprecise cleavage by DROSHA or DICER (Aravin and Tuschl 2005; Ruby et al. 2006; Neilson et al. 2007; Ruby et al. 2007; Azuma-Mukai et al. 2008; Seitz et al. 2008). Variations at 5' ends change miRNA seeds (see below), which affect target specificity of miRNAs. Thus it is not surprising that 5' ends show less variability than 3' ends, which often contain additional untemplated nucleotides with a bias for adenosine or uracil (Aravin and Tuschl 2005; Ruby et al. 2006; Neilson et al. 2007; Ruby et al. 2007).

After Dicer cleavage, the miRNA duplex is loaded into the effector miRNA-containing ribonucleoprotein complex (miRNP) with the help of DICER, TRBP, and Argonaute proteins (described in the section under "Mechanisms of miRNA-directed silencing") (Chendrimada et al. 2005; Gregory et al. 2005; Maniataki and Mourelatos 2005; MacRae et al. 2008). The double-stranded duplex must be unwound, and studies indicate that the strand with relatively unstable base pairing at the 5' end in the duplex is preferentially loaded into miRNP (Aza-Blanc et al. 2003; Khvorova et al. 2003; Schwarz et al. 2003). This strand is called the guide strand, as it will guide miRNP to mRNA

targets. The other strand is called the passenger strand, and is released from miRNP and subsequently degraded. In the case where there is extensive complementarity along the hairpin stem of the miRNA duplex, Argonaute can cleave the passenger strand (Matranga et al. 2005; Miyoshi et al. 2005; Leuschner et al. 2006; Diederichs and Haber 2007). However most miRNA duplexes contain mismatches, and human Argonautes 1 and 2 were found capable of performing multiple rounds of strand dissociation (Gregory et al. 2005; Maniataki and Mourelatos 2005; MacRae et al. 2008; Kawamata et al. 2009; Wang et al. 2009).

Regulation of miRNA biogenesis

Regulation of miRNA biogenesis occurs at the level of transcription, editing, and processing by DROSHA and DICER. The transcription of miRNAs is controlled by transcription factors that regulate the production of pri-miRNAs in specific cell types during development or in response to different environmental cues. For example, key ES cell transcription factors such as SOX2 and NANOG are associated with promoters of miRNA genes that are preferentially expressed in ES cells (Marson et al. 2008). Following the onset of DNA damage, P53 activates the transcription of pri-miR-34, and expression of miR-34 family can induce cell cycle arrest (He et al. 2007). In addition, methylation of promoter sequences can silence expression of miRNA genes. Hypermethylation of tumor suppressor miRNA genes have been observed in metastatic cancer cells (Lujambio et al. 2008).

RNA editing by adenosine deaminases acting on RNA (ADARs) can change sequences of pri-miRNAs, which can affect their further processing as well as target recognition abilities. ADAR editing of pri-miR-142 prevents its processing by DROSHA, and leads to decreased expression of mature miR-142-5p and miR-142-3p in hematopoietic tissues (Yang et al. 2006; Neilson et al. 2007).

Processing by RNaseIII enzymes and their auxiliary proteins can be regulated for individual miRNAs. For example, hnRNPA1 binding to the loop region of pri-miR-18a facilitates its processing, but not other miRNAs that belong to the same miR-17-92 cluster (Michlewski et al. 2008). Arsenate-resistance protein 2 (ARS2) is expressed by proliferating haematopoietic cells and interacts with the nuclear cap-binding complex to promote processing of pri-miRNA transcripts (Gruber et al. 2009). In addition, Erk activation can phosphorylate TRBP, which enhances the stability of TRBP as well as Dicer (Paroo et al. 2009). This leads to increased expression of growth-promoting miRNAs as well as downregulation of the let-7 family, which suppresses cell growth.

LIN28 can affect the biogenesis of let-7 genes in multiple ways. Pri-let-7 is expressed in both undifferentiated and differentiated ES cells, but mature let-7 is only detected in differentiated cells. LIN28 can prevent DROSHA processing of pri-let-7 by binding to conserved bases in its terminal loop (Newman et al. 2008; Viswanathan et al. 2008). In addition, LIN28 can bind to pre-let-7 in the cytoplasm, and prevent its cleavage by DICER (Rybak et al. 2008). Moreover, LIN28 can induce uridylation at the 3' end of pre-let-7, leading to its degradation by nucleases (Heo et al. 2008; Hagan et al. 2009). Interestingly, mature let-7 can also target Lin28 post-transcriptionally (Wu and Belasco 2005; Kumar et al. 2007; Hagan et al. 2009), and the interplay between let-7 and Lin28 is important in the regulation of stem cell differentiation (Melton et al. 2010).

In contrast to our knowledge about the biogenesis of miRNAs, very little is known about the half-life and degradation of individual miRNAs. Mature miRNAs bound to AGO2 are relatively stable, where most of them have a half-life greater than 14 hours (Lee et al. 2003; Hwang et al. 2007). However, some miRNAs display faster degradation kinetics (Hwang et al. 2007; Pedersen et al. 2007), suggesting specific mechanisms may control individual miRNA turnover.

Mechanisms of miRNA-directed silencing

Argonautes and GW182 proteins

Argonaute proteins (AGOs) associate with miRNAs and are core components of miRNPs that repress protein translation and/or trigger degradation of target mRNAs. AGOs are multidomain proteins that contain an N-terminal domain, a PAZ domain, a PIWI domain, and a MID (middle) domain (Carthew and Sontheimer 2009). The PAZ domain can recognize 3' dinucleotide termini of ssRNAs, suggesting that the domain is important in guide strand binding (Song et al. 2003; Yan et al. 2003; Lingel et al. 2004). The 5' phosphate of the guide strand is buried in a phosphate-binding pocket at the interface between the MID domain and the PIWI domain. The 5' nucleotide of the guide strand is distorted and does not base pair with the target strand (Ma et al. 2005; Parker et al. 2005). Nucleotides 2-8, which are critical for target recognition, are arranged in a geometry resembling an A form helix which favors Watson-Crick base pairing with their mRNA targets (Mallory et al. 2004; Ma et al. 2005; Parker et al. 2005). The PIWI domain adopts an RNase H-like fold that can induce endonucleolytic cleavage of the mRNA target when its pairing is nearly completely complementary to the entire guide strand (Song et al. 2004; Parker et al. 2005). However, such interactions are very rare between mammalian miRNAs and their targets. In humans, all four Agos contribute to miRNA directed silencing, but only AGO2 demonstrates endonucleolytic activity (Liu et al. 2004; Meister et al. 2004b; Su et al. 2009).

Because of their importance in miRNA-directed silencing, it is not surprising that many pathways exist to regulate the stability and activity of Agos. For example, Ago2 can be hydroxylated at proline 700 (Qi et al. 2008), and phosphorylated at Serine 387 via the p38 pathway (Zeng et al. 2008). Both modifications stabilize Ago2, and enhance

its localization to P bodies (see section “P bodies and stress granules”) (Qi et al. 2008; Zeng et al. 2008). In addition, mouse LIN41 acts as an E3 ubiquitin ligase to facilitate Ago2 turnover, which affects global miRNA activity in stem cells (Rybak et al. 2008).

Besides Argonautes, genetic screens and biochemical purifications identified GW182 among other proteins that are required for miRNA expression and function. Depletion of GW182 relieves miRNA directed repression (Rehwinkel et al. 2005; Behm-Ansmant et al. 2006a; Behm-Ansmant et al. 2006b; Eulalio et al. 2008; Eulalio et al. 2009a; Eulalio et al. 2009b; Eulalio et al. 2009c). In addition, tethering of an mRNA reporter directly to GW182 can induce silencing of the reporter independently of AGOs, suggesting that GW182 functions in the same pathway, but downstream of AGOs (Behm-Ansmant et al. 2006a; Behm-Ansmant et al. 2006b). GW182 is enriched in glycine and tryptophan repeats, which are critical for interacting with AGOs (Liu et al. 2005b; Meister et al. 2005; Behm-Ansmant et al. 2006a; Till et al. 2007).

Translational inhibition by miRNAs

Eukaryotic translation of mRNAs consists of three steps: initiation, elongation and termination. The mechanisms by which miRNP regulates translation have been a subject of debate. Depending on the experimental system used, translation inhibition can happen at initiation as well as post-initiation steps (Filipowicz et al. 2008; Carthew and Sontheimer 2009). Currently there are three models to account for translation inhibition at initiation. The first two involve regulations at the 5' cap and poly-A binding steps of the initiation process. AGO2 was shown to bind to the 5' cap, and compete with eIF4E at the start of translation (Humphreys et al. 2005; Pillai et al. 2005; Kiriakidou et al. 2007; Zdanowicz et al. 2009; Djuranovic et al. 2010). In addition, miRNP can promote mRNA deadenylation, which prevents proper circularization of the mRNA (Behm-Ansmant et al. 2006a; Giraldez et al. 2006; Wu et al. 2006; Wakiyama et al. 2007), a process that is

important for translation initiation. The last model suggests that AGO2 can interact with eIF6 *in vitro*, and block the joining of 60S to 40S ribosomal subunits (Sanvito et al. 1999; Chendrimada et al. 2007).

Another line of evidence suggests that miRNAs can repress translation after the initiation step (Kim et al. 2004; Nelson et al. 2004; Maroney et al. 2006; Lytle et al. 2007; Vasudevan and Steitz 2007). Petersen *et. al.* showed IRES-containing reporters can be repressed, and the repressed mRNAs are associated with active polysomes (Petersen et al. 2006). When they inhibited translation initiation with a drug, they observed ribosomal drop-off in a miRNA dependent manner, suggesting miRNAs can promote premature ribosome dissociation from mRNAs (Petersen et al. 2006).

Degradation of mRNAs by miRNAs

Despite the focus of miRNA-directed translation inhibition by early studies, recent evidence shows that an increase in miRNA abundance is associated with a reduction of mRNAs that contain partial miRNA complementary sites (Bagga et al. 2005; Krutzfeldt et al. 2005; Lim et al. 2005; Behm-Ansmant et al. 2006a; Giraldez et al. 2006; Wu et al. 2006). Bagga *et. al.* used Northern analysis to show that *let-7* can decrease mRNA levels of its target *lin-41* (Bagga et al. 2005). *Lin-4* also had similar effects on mRNA level of its targets *lin-14* and *lin-28* (Bagga et al. 2005; Ding and Grosshans 2009). Moreover, transcriptomic studies suggest that miRNAs can regulate the stability of a large number of mRNAs. After over-expressing miRNAs in HeLa cells, Lim *et. al.* observed downregulation of sets of mRNAs that preferentially displayed miRNA binding sites in their 3' UTRs (Lim et al. 2005). Zebrafish miR-430 promotes clearance of hundreds of maternal mRNAs by inducing their deadenylation at the onset of zygotic transcription (Giraldez et al. 2006).

Mechanistic studies initially performed in *Drosophila* S cells demonstrate that

GW182 is required for mRNA decay (Behm-Ansmant et al. 2006a). Knockdown or depletion of GW182 led to upregulation of miRNA targets at the mRNA level (Behm-Ansmant et al. 2006a). The upregulation is also dependent on deadenylation (CCR4-NOT1) and decapping (DCP1 and DCP2) complexes (Behm-Ansmant et al. 2006a; Eulalio et al. 2007c). Depletion of the components of the CCR4-NOT1 complexes, as well as knock-down of DCP1 and DCP2 prevented miRNA-mediated mRNA downregulation (Rehwinkel et al. 2005; Behm-Ansmant et al. 2006a; Chu and Rana 2006; Eulalio et al. 2007c). These and experiments in *C. elegans* and mammalian cells suggest a model in which GW182 first recruits the deadenylase and decapping complexes to target mRNAs. The target mRNAs can then be subsequently degraded by exosomes (3'-5' exonuclease) or XRN1 (after the removal of the 5' cap) (Behm-Ansmant et al. 2006a; Wu et al. 2006; Eulalio et al. 2007c).

Altogether, miRNPs have been shown to elicit translational repression, mRNA decay, or both. Translational repression can take place independently of target deadenylation, as mRNAs without polyA tails are still repressed by miRNAs (Pillai et al. 2005; Giraldez et al. 2006; Wu et al. 2006; Eulalio et al. 2008; Eulalio et al. 2009b). These targets show only degradation at the protein level (Baek et al. 2008; Selbach et al. 2008). Likewise, mRNA decay is not dependent on translation, as it can still occur when translation is inhibited by cycloheximide (Eulalio et al. 2007c; Wakiyama et al. 2007). Nevertheless, translational repression is often linked to mRNA decay. Recent proteomic experiments showed that many translationally repressed genes also displayed detectable mRNA destabilization (Baek et al. 2008; Selbach et al. 2008). In fact, mRNA destabilization comprised the major component of repression for a certain set of highly repressed targets (Baek et al. 2008; Selbach et al. 2008).

P bodies and stress granules

P-bodies are discrete cytoplasmic foci enriched in proteins that participate in miRNA-directed mRNA degradation and translational repression (Filipowicz et al. 2008; Carthew and Sontheimer 2009). The core P-body components include enzymes responsible for decapping, deadenylation, and degradation, and they interact directly or indirectly with AGOs and Gw182 proteins (Anderson and Kedersha 2006; Eulalio et al. 2007a; Parker and Sheth 2007). Although depletion of decapping enzymes leads to ineffective inhibition of target mRNAs, P-bodies are not essential for this process (Jakymiw et al. 2005; Liu et al. 2005b; Meister et al. 2005). Knocking down components of P bodies has no effect on miRNA activities, suggesting that the formation of P-bodies is a consequence rather than the cause of silencing (Pillai et al. 2005; Eulalio et al. 2007b).

The structures of P-bodies are dynamic, and their sizes change in response to translational status of the cell (Anderson and Kedersha 2006; Eulalio et al. 2007a; Parker and Sheth 2007). A global translation initiation block leads to an increase in the size of P-bodies (Anderson and Kedersha 2006; Eulalio et al. 2007a; Parker and Sheth 2007). In contrast, inhibition of miRNA biogenesis or activity results in dispersal of visible P-bodies, suggesting that functional miRNA pathways are essential for the formation of large P-body aggregates (Pauley et al. 2006; Eulalio et al. 2007b).

Another type of mRNA-containing cytoplasmic aggregates is stress granules (SG), which accumulate in response to various stress conditions (Anderson and Kedersha 2006). Leung *et al.* showed that miRNA mimics and the repressed reporter mRNAs accumulate in SGs upon stress (Leung et al. 2006). Moreover, the localization of Ago proteins to SGs but not P-bodies is miRNA-dependent, and the exchange of AGOs at SGs is much faster than P-bodies (Leung et al. 2006).

Principles of miRNA-mRNA interactions

Since the first discovery of miRNA target *lin-14* in the early 1990s, genetic, biochemical, and bioinformatics analyses have revealed many functional target sites, most of which are in the 3' UTRs of mRNAs (Bartel 2009; Carthew and Sontheimer 2009). The 5' end of miRNA is critical in determining target specificity, and bases 2-7 of the miRNA are referred to as the miRNA seed. Comparative genomic analysis revealed that the 5' region of miRNAs is the most conserved portion of miRNAs (Lewis et al. 2003; Lim et al. 2003a; Lewis et al. 2005; Chen and Rajewsky 2006). Furthermore, seed binding sites in the 3' UTRs of mRNAs tend to be more highly conserved than expected by chance (Lewis et al. 2003; Lim et al. 2003a; Lewis et al. 2005). In addition, reporter assays showed that pairing to the miRNA 5' region is sufficient to effect repression, and mutations that disrupt the binding between the seed and target site relieve reporter repression (Lai 2002; Doench and Sharp 2004; Brennecke et al. 2005b; Kloosterman and Plasterk 2006). Recent large-scale transcriptomic and proteomic studies also demonstrated that a high proportion of transcripts downregulated in response to miRNA overexpression contain sequences complementary to the seed region (Lim et al. 2005; Baek et al. 2008; Selbach et al. 2008).

Although the presence of a 6-mer seed binding site is important, additional matches to the 5' end of the miRNA can improve target specificity. Analyses of transcript expression before and after the addition of a miRNA revealed that mRNAs that contained a seed match flanked by a Watson-Crick match to miRNA base 8 (M8 7mer) exhibited enhanced downregulation (Grimson et al. 2007; Nielsen et al. 2007). The presence of an adenosine at position 1 of the mRNA target site (A1 7mer) also led to more mRNA reduction (Grimson et al. 2007; Nielsen et al. 2007). Moreover, mRNAs that have a M8 7mer as well as an A1 7mer site (M8-A1 8mer) exhibited the greatest mRNA

downregulation (Grimson et al. 2007; Nielsen et al. 2007).

While “seed only” type targets represent the vast majority of all conserved miRNA targets, there is another class of targets that have weak 5' base-pairing (single-nucleotide bulge or mismatch) with the miRNA and depend on strong compensatory pairing to the 3' end of the miRNA (Bartel 2009). Some notable examples include the *let-7* binding sites in *lin-41*, miR-2 sites in *Grim* and *Sickle*, and miR-196 site in *Hoxb8* (Reinhart et al. 2000; Lewis et al. 2003; Stark et al. 2003; Yekta et al. 2004). It has been speculated that 3' pairing confers target specificity to miRNAs in the same family, which share a common seed region, but have different 3' ends (Brennecke et al. 2005b; Bartel 2009).

The context of the miRNA binding site is also important in determining target specificity. Increased AU content near the seed match, additional pairing to miRNA nucleotides 13-16, close proximity of another miRNA binding site, positioning of the miRNA binding site away from the center of long UTRs but at least 15-nt away from the stop codon, are all associated with increased mRNA downregulation (Grimson et al. 2007; Nielsen et al. 2007). These context features can increase accessibility and affinity of miRNA binding sites, resulting in more favorable interaction between silencing complexes (Grimson et al. 2007).

These and previous observations have been extended to predict mRNA targets of miRNAs (Bartel 2009), and more recent analysis suggests that over half of human genes are conserved miRNA targets (Friedman et al. 2009). In addition to conserved sites, many mRNAs have non-conserved binding sites for each miRNA. While most of the mRNAs with non-conserved sites are expressed in tissues where the miRNA is absent, some are co-expressed with the miRNA, suggesting that they can represent important species-specific targets (Farh et al. 2005; Stark et al. 2005).

Predicted targets of miRNAs can be supported by comparing the activity of a

reporter in systems that exhibit different expression levels of miRNAs. One commonly used reporter is the luciferase reporter, which has the coding sequence of luciferase fused to the 3' UTR of the target gene (Elbashir et al. 2001). Luciferase reporters of true targets will have a higher activity in the system where the miRNA is expressed at a lower level. Several ways exist to manipulate miRNA expression in different experimental systems. Firstly, miRNA overexpression can be achieved through expressing miRNA hairpin and flanking sequences (Chen et al. 2004), or transient transfection of miRNA mimics and small interfering RNAs (siRNAs). Secondly, miRNAs can be depleted with the use of chemically modified antisense oligonucleotides, or miRNA sponges (Hutvagner et al. 2004; Meister et al. 2004a; Krutzfeldt et al. 2005; Orom et al. 2006; Ebert et al. 2007). Lastly, the entire mature miRNA repertoire can be depleted by knocking out *DGCR8* or *Dicer*, whose gene products are important for the maturation of miRNAs. *Dgcr8* null mouse ES cells (mESC), *Dicer* null mESCs and *Dicer* null mouse embryonic fibroblasts (mEF) have all been successfully generated, and they provide an alternative to dissect miRNA functions (Kanellopoulou et al. 2005; Murchison et al. 2005; Calabrese et al. 2007; Nielsen et al. 2007; Wang et al. 2007).

Besides 3' UTR targets, targets have been found in the coding sequence (CDS) or 5' UTR of mRNAs. Comparative genomics analyses detected a significant signal above background in conserved miRNA seed matches in the CDS of mRNAs (Lewis et al. 2003; Brennecke et al. 2005b; Lewis et al. 2005). A handful of targets have been validated experimentally; they include *Nanog*, *Oct4*, and *Sox2* (targeted by miR-134, miR-296, and miR-470), *Dnmt3b* (targeted by miR-148), and *p16* (targeted by miR-24) (Lytle et al. 2007; Baek et al. 2008; Duursma et al. 2008; Selbach et al. 2008; Tay et al. 2008). In addition, proteomic studies have revealed miRNA binding sites in the CDS, although they are generally less effective than sites in 3' UTRs (Baek et al. 2008; Selbach et al. 2008). More recently, an effort to identify miRNA binding sites in RNAs

that crosslink to AGO2 also uncovered prevalent miRNA binding sites in CDS, suggesting that CDS targets may be more widespread than previously considered (Leung et al. 2006).

Roles of miRNAs in mammalian development

miRNAs exert vast impact on global expression and evolution of mammalian mRNAs. To date, around 400 miRNA genes have been cloned from humans, and more than 50% of human genes contain conserved miRNA binding sites (Bartel 2009). Systematic analyses of mRNA and miRNA expression profiles show that conserved targets of tissue-specific miRNAs are frequently expressed at lower levels than other tissues, and nonconserved targets are usually not expressed in the same tissue as the miRNA, suggesting that miRNAs facilitate state transitions and help maintain cell identity (Farh et al. 2005). A handful of targets whose activity can be strongly repressed by a miRNA to an inconsequential level are referred as “switch targets”, in contrast to the vast majority of targets, “fine-tuning targets”, whose expression is only moderately dampened by the miRNA (Bartel and Chen 2004). In addition, ubiquitously expressed genes and tissue specific genes tend to have shorter 3' UTRs so that they can avoid being targeted by miRNAs (anti-targets) (Farh et al. 2005; Stark et al. 2005). Additionally, proliferating cells are associated with widespread 3' UTR shortening, further supporting the global impact of miRNA on 3' UTR evolution (Sandberg et al. 2008; Mayr and Bartel 2009).

Since the identification of the first miRNAs in regulating the developmental timing of *C. elegans*, miRNAs have been implicated in fundamental cellular processes (Lee et al. 1993; Wightman et al. 1993). A close link exists between deregulation of normal cellular processes and tumorigenesis, and a growing body of evidence indicates that altered expression of miRNAs is involved in the pathogenesis of cancers (Croce and Calin 2005; Johnson et al. 2005; Lu et al. 2005; Costinean et al. 2006; Esquela-Kerscher

and Slack 2006; Hammond 2006; Voorhoeve et al. 2006; Kumar et al. 2007; Mayr et al. 2007). miR-15a and miR-16 can function as tumor suppressor genes by targeting Bcl-2, an oncogene that inhibits apoptosis (Cimmino et al. 2005). They map to a region of chromosome 13 that is commonly lost in patients with chronic lymphocytic leukemia (Calin et al. 2002). Subsequent expression profiling of different types of tumors revealed many more miRNAs that might function as tumor suppressors or oncogenes by targeting transcription factors, epigenetic machinery as well as existing tumor suppressors and oncogenes (Calin et al. 2004; Esquela-Kerscher and Slack 2006; Volinia et al. 2006; Chin et al. 2008; Croce 2009).

The global functional role of miRNAs in development can be inferred from animals that lack DICER, DGCR8, and AGO2, as all of them are important in the maturation and activity of miRNAs. Loss of Dicer leads to lethality early in development and depletion of stem cells in Dicer null embryos (Bernstein et al. 2003; Kanellopoulou et al. 2005). Tissue specific deletion of Dicer resulted in defects such as limb morphogenesis, lung development, and incomplete embryonic myogenesis (Harfe et al. 2005; Harris et al. 2006; O'Rourke et al. 2007). Similarly, Dgcr8 null and Ago2 null embryos show severe developmental delays and defects (Liu et al. 2004; Morita et al. 2007; Wang et al. 2007).

The role of individual miRNAs in mammalian development has been gradually revealed through expression profiling experiments and genetic studies. For example, miR-1 is highly expressed in cardiac and skeletal muscle cells in mouse (Zhao et al. 2005b). Overexpression of miR-1 in myoblasts promoted differentiation while reducing cell proliferation (Zhao et al. 2007). Deletion of *miR-1-2*, one of the two genes encoding miR-1 in the mouse genome, resulted in animals with defects in ventricular septum and cardiac rhythm disturbances (Zhao et al. 2007). Here we will focus our discussion on developmental roles of miRNAs in mouse ES cells.

miRNAs in ES cells

ES cells are derived from the inner cell mass of the blastocyst-stage embryo (Figure 1). During gastrulation, ES cells rapidly differentiate into the three primary germ layers of the developing fetus. However, ES cells can be cultured *in vitro*, and their capacity for self-renewal and differentiation potential can be maintained in the presence of the cytokine, leukemia inhibitory factor (Smith et al. 1988; Williams et al. 1988). Cultured ES cells can be injected into blastocysts and contribute to all tissues of the organism except those of trophoctoderm or primitive endoderm lineages (Beddington and Robertson 1989).

The self-renewing capacity of ES cells is controlled by both intercellular and intracellular mechanisms (Lin 2002). Intense investigation revealed a multitude of intracellular mechanisms to regulate gene expression at epigenetic, transcriptional, translational and post-translational levels. Two homeodomain transcription factors, OCT4 and NANOG, were the first proteins identified as essential for maintaining pluripotency in ES cells (Nichols et al. 1998; Chambers et al. 2003; Mitsui et al. 2003). Later, genome-wide binding sites of transcription factors revealed that OCT4, NANOG, and SOX2 share a substantial fraction of their target genes (Boyer et al. 2005; Loh et al. 2006). While some target genes encode key transcription factors for differentiation and development which are transcriptionally inactive, another set of active target genes are involved in the maintenance of pluripotency (Boyer et al. 2005). Interestingly, these three factors control one another's transcription in a large regulatory circuit consisting of specialized autoregulatory and feedforward loops (Boyer et al. 2005).

The ability to self-renew is inherently linked to the ability of cells to remain in a proliferative state. ES cells have an unusual cell cycle structure and rapid rate of cell division (Savatier et al. 1994; Stead et al. 2002). The cell cycle structure of mESCs

consists of a truncated G1 phase and a prolonged S phase, which explain why ES cells can divide approximately every 8-10 hours (Savatier et al. 1994; Stead et al. 2002). In contrast to cyclin expression patterns in regular mammalian cells, mESCs express elevated levels of all cyclins, which remain almost the same throughout the cell cycle (Savatier et al. 1994; Faast et al. 2004). Additionally, Cdk inhibitory genes such as the Ink family, p21, p27, and p57 have low expression levels which lead to the hyperphosphorylation of pRb and inactivation of the G1-S checkpoint (Savatier et al. 1994; Savatier et al. 1996; Faast et al. 2004).

The important role of miRNAs in ES cell regulation can be inferred from experiments that study the loss of *Dicer* and *Dgcr8* in ES cells. Although *Dicer*-null ES cells are viable, they proliferate more slowly than WT ES cells (Kanellopoulou et al. 2005; Murchison et al. 2005). While the null cells express about the same level of Oct4 as the WT cells, they cannot differentiate *in vitro*. Markers characteristic of endodermal (*Hnf4*) and mesodermal (*Brachyury*, *Bmp4*, and *Gata1*) lineages cannot be detected (Kanellopoulou et al. 2005; Murchison et al. 2005). *Dgcr8* knockout ES cells displayed similar growth and differentiation defects, although they do express some markers of differentiation, and embryoid bodies can continue to grow and differentiate after 16 days of being cultured (Wang et al. 2007).

miRNA expression profiles in mouse ES cells revealed that ES cells express a unique set of miRNAs, and that these miRNAs are downregulated as ES cells differentiate into embryoid bodies, providing further support that miRNAs play an important role in maintaining pluripotency of ES cells (Houbaviy et al. 2003; Suh et al. 2004; Houbaviy et al. 2005; Landgraf et al. 2007; Babiarz et al. 2008; Ciaudo et al. 2009). Members of miR-302 and miR-290-295 clusters along with their homologs are among the best characterized miRNAs in ES cells, and have been linked to regulating ES cell lineage maintenance, differentiation, and proliferation capacity. Benetti *et. al.* and

Sinkkonen *et al.* both showed that mESC specific miR-290-295 cluster can target Rbl2, which control the expression of Dnmt1, Dnmt3a, and Dnmt3b. This suggests that the miRNA cluster plays a role in regulating *de novo* DNA methylation. (Benetti *et al.* 2008; Sinkkonen *et al.* 2008). miR-302 has been shown to control germ layer specification by inhibiting lefty, an inhibitor of the Nodal pathway (Rosa *et al.* 2009). In addition, Melton *et al.* showed that the miR-290-295 cluster can indirectly increase the expression of Lin28 and c-Myc to maintain mESCs in their self-renewing state (Melton *et al.* 2010). Lastly, miR-290-295, miR-302, and miR-372 have been found to accelerate cell proliferation by promoting G1 to S phase transition through targets such as p21 and Lats2 (Voorhoeve *et al.* 2006; Wang *et al.* 2007; Card *et al.* 2008; Lee *et al.* 2008).

Recent studies have also linked ES cell specific transcription factors to miRNAs. Chip-seq data showed that NANOG, OCT4 and SOX2 bind to promoter regions of miR-290-295, miR-302, and other highly expressed miRNAs (Card *et al.* 2008; Marson *et al.* 2008). In addition, Polycomb Group protein SUZ12 occupies promoters of tissue specific miRNAs that are silenced in ES cells (Card *et al.* 2008; Marson *et al.* 2008). Interestingly, miRNAs can also control the expression of Nanog, Oct4 and Sox2. Upon retinoid-acid-induced differentiation, miR-134, miR-296, and miR-470 have been reported to repress the expression of each transcription factor by targeting its coding region, further illustrating the important interplay between miRNAs and transcription factors in regulating ES cell self-renewal and differentiation (Tay *et al.* 2008).

The abilities of ES cells to self-renewal and rapid division make them an important system to dissect the function of miRNAs. Understanding how miRNAs control their pluripotency potential and proliferation will not only help us better understand key aspects of mammalian development, but also provide insights that can be applied to reprogramming and cancer. In this thesis, we will focus on the role of miRNAs in regulating ES cell survival.

T lymphocyte development

T lymphocytes are part of the adaptive immune system that recognizes and eliminates specific foreign antigens. T lymphocytes arise in the bone marrow, and migrate to the thymus gland to mature into CD4 or CD8 T cells (Figure 2). Mature T cells express a unique antigen-binding molecule, the T-cell receptor (TCR) on their membrane, and can only recognize antigen that is bound to cell membrane proteins called major histocompatibility complex (MHC) molecules. T cells that recognize self-MHC molecules are selected for survival during positive selection (Starr et al. 2003). However, T cells that react too strongly with self-MHC are eliminated through negative selection (Starr et al. 2003).

Maturation of T cells consists of six major steps (Figure 2). Thymocytes early in development lack detectable CD4 and CD8, and are referred to as double negative (DN). DN T cells can be subdivided into four subsets (DN1-4) characterized by the presence or absence of cell surface molecules in addition to CD4 and CD8, such as CD44, an adhesion molecule, and CD24, the alpha chain of the IL-2 receptor. The cells that enter the thymus, DN1, are capable of giving rise to all subsets of T cells, and are phenotypically CD44^{hi}, and CD25⁻. Once DN1 cells encounter the thymic environment, they begin to proliferate and express CD25, becoming CD44^{low}, and CD25⁺. They are called DN2 cells, where rearrangement of genes for the TCR chains begins. As cells progress to DN3, the expression of CD44 is turned off and cells stop proliferating to start TCR β chain rearrangement. Upon its completion, the DN3 cells quickly progress to DN4, where the level of CD25 decreases.

Both CD4 and CD8 receptors are expressed in the double positive (DP) stage, where rapid cell division increases the diversity of the T-cell repertoire. After the rapid proliferation, TCR α chain rearrangement starts, which is then followed by positive and

negative selection. Cells that fail to make productive TCR gene arrangement or thymic selection are eliminated by apoptosis. The 2% that survive will develop into immature CD4 or CD8 thymocytes. These single-positive cells undergo additional negative selection and migrate to the medulla, where they pass from the thymus into the circulatory system.

Individual cell populations along the developmental series of T lymphocytes are enriched for cells undergoing fundamental cellular processes such as proliferation, differentiation, and apoptosis. Additionally, cells can be distinguished and easily isolated by their unique set of surface markers, making them an ideal experimental system to study the role of miRNAs in development.

Placental development

Mouse placental development begins in the blastocyst when the trophoblast layer is set aside from the inner cell mass at embryonic day 3.5 (Cross et al. 1994) (Figure 1). The placenta provides the fetus with nutrients, allows for gas and waste exchange, and protects the fetus from the maternal immune system as well as environmental stress (Sood et al. 2006).

Imprinted gene expression has been observed in the placenta, and linked to placental function (Kaneko-Ishino et al. 2003). Gynogenetic (two maternal/no paternal genomes) and androgenetic (two paternal/no maternal genomes) embryos both exhibit defects in trophoblast development (McGrath and Solter 1984). Gynogenetic embryos have very few trophoblast cells while androgenetic embryos are characterized by a mass of hypertrophic trophoblast (McGrath and Solter 1984). This observation is consistent with the parent-offspring conflict hypothesis that has been proposed to explain the evolution and maintenance of imprinting in mammals (Moore and Haig 1991). Paternally expressed genes are proposed to increase embryonic growth, thereby maximizing the

competitiveness of individual offspring bearing a particular paternal genome (Moore and Haig 1991). Maternally expressed genes are proposed to suppress fetal growth (Moore and Haig 1991). This would allow a more equal distribution of maternal resources to all offspring and increase transmission of the maternal genome to multiple offspring, which may have different paternal genomes (Moore and Haig 1991).

To date, over 50 imprinted genes have been discovered, and ~20% of these genes show placenta-specific imprinting (Wagschal and Feil 2006). Interestingly, most of them are expressed only from the maternal allele (Wagschal and Feil 2006). Many are involved in cellular proliferation and growth, although their precise roles in placental development and function remain largely unknown (Tycko and Morison 2002).

Two experimental systems have been used to study trophoblast differentiation and placental function. Culture conditions have been established for trophoblast stem cells (Quinn et al. 2006). In addition, cells from choriocarcinoma, a malignant cancer of the placenta, have been derived and cultured in labs (Pattillo and Gey 1968).

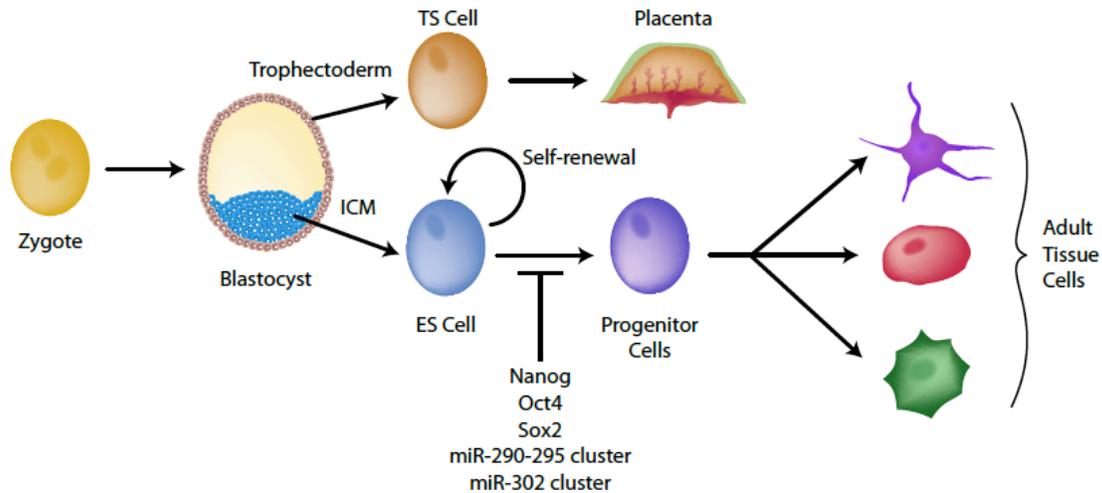
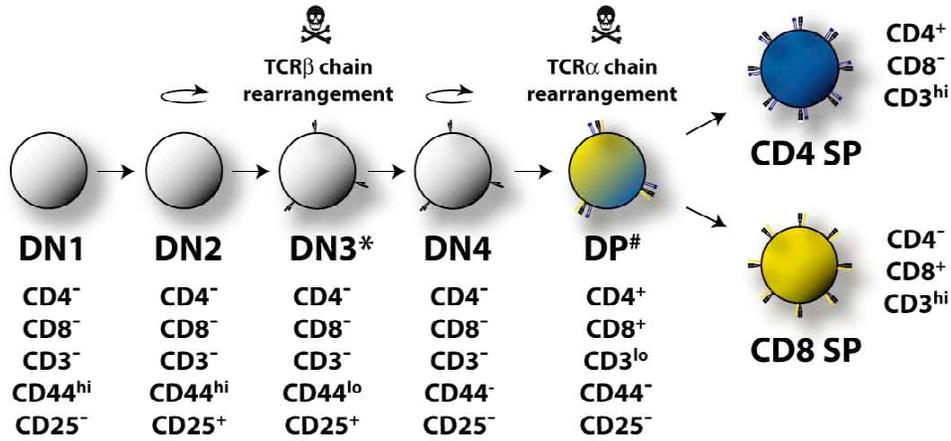


Figure 1. Differentiation of totipotent cells.

ES cells are derived from the inner cell mass (ICM) of the blastocyst-stage embryo. ES cells can be cultured *in vitro*, and their capacity for self-renewal and differentiation potential can be maintained in the presence of the cytokine, leukemia inhibitory factor (Smith et al. 1988; Williams et al. 1988). Cultured ES cells can be injected into blastocysts and differentiate into three germ layers. Transcription factors (Nanog, Oct4, and Sox2) and miRNAs (miR-290-295 and miR-302 clusters) are essential for maintaining pluripotency in ES cells. The outer layer of the blastocyst, termed the trophoctoderm, gives rise to trophoblast stem (TS) cells *in vitro* and populate the major structures of the placenta *in vivo*.



☠ = wired for apoptosis

↻ = proliferation

*Commitment to $\alpha\beta$ T lineage

#Positive and negative selection

Figure 2. T lymphocyte development.

T lymphocytes arise in the bone marrow, and migrate to the thymus gland to mature into CD4 or CD8 T cells. DN1 stage of development is a mix of lymphoid and myeloid progenitor cells. T lineage commitment is fixed upon rearrangement of the genetic locus encoding the β chain of the T cell receptor (TCR) at the DN3 stage. Cells that successfully rearrange this locus proceed to the DN4 stage. Following the DN4 stage, cells start expressing the CD4 and CD8 coreceptors to become DP cells, where the thymocytes rearrange the α chain of the TCR and undergo positive and negative selection. Those cells that are positively selected mature to CD4 SP or CD8 SP T lymphocytes and migrate to the circulatory system.

CHAPTER 2: Dynamic regulation of miRNA expression in ordered stages of cellular development

The material presented in this chapter was adapted, with permission, from the following publication:

Joel R. Neilson, Grace X.Y. Zheng, Christopher B. Burge and Philip A. Sharp (2007).
Dynamic regulation of miRNA expression in ordered stages of cellular development.
Genes Dev. 21, 578-589.

Experimental contributions:

Joel R. Neilson cloned the short RNAs and performed most of the experiments in the chapter. Grace X.Y. Zheng performed all the bioinformatics analysis.

Abstract

Short RNA expression in several distinct stages of T lymphocyte development was comprehensively profiled. The total number of miRNAs expressed per cell at different stages of development varies over nearly an order of magnitude in parallel with changes in total cellular RNA content, suggesting that global miRNA levels are coregulated with the translational capacity of the cell. However, individual miRNAs were dynamically regulated during T cell development, with at least one miRNA or miRNA family overrepresented at each developmental stage. MiRNA regulation in this developmental pathway is characterized by analog rather than switch-like behavior, with temporal enrichments at distinct stages of development observed against a background of constant, basal expression of the miRNA. Enrichments of these miRNAs are temporally correlated with depletions of the transcript levels of predicted targets, and have specific functional consequences. MiR-181a, which is specifically enriched at the CD4⁺CD8⁺ (DP) stage of thymocyte development, can represses the expression of Bcl-2, CD69, and the T cell receptor, all of which are coordinately involved in positive selection.

Introduction

Short RNAs are known to control gene expression at several different levels in organisms ranging from yeast to plants to mammals (Kloosterman and Plasterk 2006). In mammals, the best-characterized class of short RNA species is the microRNA (miRNA) class. These 22 nucleotide RNA species repress gene expression at the level of translation by binding to the 3' untranslated region (3' UTR) of target mRNAs. The precise mechanisms of this repression are probably multifold (Bagga et al. 2005; Pillai et al. 2005; Giraldez et al. 2006; Leung et al. 2006; Petersen et al. 2006).

Both bioinformatic and *in vitro* experimental data indicate that miRNA:mRNA recognition is predominantly mediated by an interaction between the 5' end of a miRNA and a complementary sequence in the mRNA target (Doench and Sharp 2004; Brennecke et al. 2005b; Lim et al. 2005). Bases 2-7 or 2-8 of the miRNA are primary contributors to target specificity and are referred to as the miRNA "seed" region (Doench and Sharp 2004; Brennecke et al. 2005b). Genomic experiments examining the effects of transfection of individual miRNAs in HeLa cells have demonstrated that a high proportion of the transcripts downregulated in response to a transfection of a miRNA contain sequences complementary to the seed region (Lim et al. 2005). Based on these observations, it is thought that miRNA family members that share common seed sequences have similar target specificity. These observations have been extended to predict mRNA targets of miRNAs (Lewis et al. 2003; Lewis et al. 2005) and to demonstrate a general reduction in the transcript levels of these targets in tissues where these miRNAs are expressed (Farh et al. 2005; Krutzfeldt et al. 2005).

The number of described mammalian miRNAs continues to increase through direct cloning efforts and bioinformatic prediction. Expression of several individual miRNAs such as miR-142 (lymphoid), miR-223 (myeloid), miR-1 (muscle), and miR-9

(neuronal), is largely restricted to single tissues or organ systems, consistent with these miRNAs having a role in the developmental specification of these cells (Lagos-Quintana et al. 2002; Chen et al. 2004). However, a large number of miRNAs exhibit a more wide-ranging pattern of expression, consistent with these regulators being involved in events such as growth and homeostasis. A systematic understanding of the roles of miRNAs in such events is incomplete as few direct studies of changes in mRNA and miRNA expression over the course of a single developmental pathway have been conducted.

Specific expression of individual miRNAs in several compartments of the mammalian hematopoietic system has been well described (Chen et al. 2004; Monticelli et al. 2005). However, a close examination of the dynamic regulation of miRNAs and the consequences of this regulation on global gene expression during sequential stages of development of a single cell lineage has not been performed. T lymphocyte development in the thymus has been extensively characterized (Starr et al. 2003). The DN1 stage of development is a mix of lymphoid and myeloid progenitor cells (Porritt et al. 2004). T lineage commitment is fixed upon rearrangement of the genetic locus encoding the beta chain of the antigen receptor at the DN3 stage; cells that successfully rearrange this locus proceed to the highly proliferative DN4 stage. Following the DN4 stage, expression of the CD4 and CD8 co-receptors defines the DP stage of thymocyte development, at which point the thymocytes rearrange the alpha chain of the T cell receptor and undergo positive and negative selection. Those cells that are positively selected mature to CD4 SP (helper) or CD8 SP (killer) T lymphocytes and egress the thymus to the periphery. Importantly, individual cell populations along this developmental series are enriched for cells undergoing fundamental cellular processes such as proliferation, differentiation, and apoptosis. Dynamic regulation of individual miRNAs within these stages might be expected to influence these processes without perhaps being critical for development of the T lineage *per se*.

Here we describe discovery-compatible profiling and quantitative measurement of short RNA expression from several ordered stages of T lymphocyte development. The pool of miRNAs at a given developmental stage correlates closely with the cytoplasmic volume as indicated by amount of total RNA. Although nearly all miRNAs that we observe in this process are constitutively expressed, specific miRNAs are enriched at distinct stages of thymocyte development. These enrichments can be correlated with depletion of predicted targets of the miRNAs at the genomic level and bear functional consequences; in reporter assays, the DP-enriched miR-181a confers repression through the 3' UTR elements of three predicted targets (CD69, Bcl-2, and the T cell receptor) known to be regulated at this stage of development.

Results

Short RNA Profiling Using Low Nanogram Amounts of Total RNA

We directly cloned short RNAs from developing thymocytes to enable profiling of known miRNAs while allowing for the discovery of additional small RNA species of potential interest. This involved developing a variation of existing protocols (Lagos-Quintana et al. 2001; Lau et al. 2001) to allow cloning of short RNA species from low nanogram amounts of total RNA (Figure 1).

We collected and analyzed 10,533 short RNA clones comprised of 3445 non-redundant sequences from six stages of T lymphocyte development, each sorted to >95% purity (Table 1, Figure 2). The sequences ranged in length from 16 to 33 nucleotides. Each sequence in the library was assigned a cloning frequency from each cell type. 957 of the sequences were observed only once, suggesting that the library is not close to saturation. The library was searched using the BLAST algorithm against full length miRNA hairpins in mirBase (Release 8.1), tRNA and NONCODE databases (Liu

et al. 2005a), and the mouse genome (mm8). While 67% of the clones matched precisely to sequences in above databases, an additional 19% of the distinct sequences could be annotated if we allowed a single internal mismatch or up to three mismatches at the end of the sequence. We also noticed that over 80% of the clones annotated as mismatches using the above criteria fell into the category of 3' end-mismatches with a bias for adenosine or uracil in these positions. This pattern could reflect biases in the cloning procedures or might reflect nucleotide modification or addition *in vivo* following processing by Drosha or Dicer. Indeed, similar observations have been reported in cloning efforts by other groups (Aravin and Tuschl 2005). Because of this, we elected to include these sequences in our analysis and were thus able to annotate 9092 (86.3%) of the clones using the above criteria.

As expected, the majority of clones were miRNAs. The average percentage of clones representing miRNA sequences from each cell type was 68%, ranging from a low of 56% in DN3 cells to a high of 82% in DP thymocytes (Table 1). The relative representation of individual sequence classes within each cell type varied similarly. Sequences corresponding to genomic regions annotated as repeats or to which ESTs have been mapped were generally the second most abundant classes, observed with an average cloning frequency of 12%. The third tier of abundance consisted of sequences associated with tRNAs or un-annotated regions of the genome, with average cloning frequencies of 6.6%. An exception was noted in the highly proliferative DN4 population, where tRNA-associated sequences were only exceeded in abundance by miRNAs. Sequences corresponding to ribosomal RNA, "other" non-tRNA and non-ribosomal ncRNA classes, CpG islands, and the RefSeq "Known Gene" category constituted the fourth tier of abundance with an average cloning frequency of 1.9%. The vast majority of clones associated with ncRNA loci corresponded to the sense strand of the specific ncRNAs. Interestingly, although 319 clones overlapped RefSeq "Known" genes, only

36% of these clones mapped to these annotations in the sense orientation. Fully half of the clones overlapping annotated genes mapped antisense to the protein-coding region of the gene. The remaining 14%, while associated with RefSeq genes, mapped to more than one region of the genome, so the precise location from which these sequences originated, and the strand polarity associated with these sequences, could not be determined.

Quantitative Validation of Clone Representation

We profiled short RNA cloning frequency from two independent preparations of DN4 thymocytes and observed a high degree of similarity (Pearson correlation coefficient = 0.925). This high degree of similarity demonstrated that the short RNA profiling method was reproducible. To determine whether the profile of short RNAs accurately reflected their relative abundance in the cell, a quantitative solution-hybridization based Trilogy® assay (Neely et al. 2006) was used to directly measure the abundance of seventeen individual miRNAs in the DP thymocyte population (Table 1, Figure 3). Reassuringly, there was a high degree of correlation between the calculated number of copies per cell and relative cloning frequency of each of these miRNAs (Pearson coefficient = 0.974). Surprisingly, this quantification revealed that miRNAs are expressed at much lower levels in DP thymocytes than in other mammalian cells in which miRNA expression has been directly quantitated. For example, miR-181a, the miRNA cloned with the highest frequency in this cell type, was expressed at roughly 810 copies per cell as compared to the up to 50,000 copies per cell reported for miRNAs in HeLa cells (Lim et al. 2003b).

To determine that the changes in cloning frequency of individual miRNAs throughout thymocyte development reflected the relative abundance of these miRNAs at each stage, we used the Trilogy® assay to calculate the copy number of each of 3-5

miRNAs independently for each cell stage except DN1 (the scarcity of these cells did not permit isolation of sufficient RNA to perform the Trilogy assay). The copy number for each miRNA in each of the cell types was independently determined and plotted against the cloning frequency of the miRNA (Figure 4). We observed generally high correlations between miRNA copy number measured by Trilogy assay and cloning frequency in each cell type, allowing the copy number of each miRNA species (and of the overall miRNA pool) to be estimated for each cell.

Surprisingly, there was a marked change in per-cell RNA levels during T lymphocyte development. In fact, the per-cell total RNA fluctuates by nearly an order of magnitude during this process, from a maximum of 6.8 pg/cell in highly proliferative DN4 thymocytes to a minimum of 0.7 pg/cell in DP thymocytes undergoing selection. The calculated miRNA pool was also highly dynamic throughout T lymphocyte development, ranging from about 5000 copies per cell in DP thymocytes to 33000 copies per cell in DN4 thymocytes. However, the estimated size of the miRNA pools covaried closely with changes in total RNA level at the various stages of T cell development (Figure 4, Figure 5). This covariation is reflected in a remarkably constant value for calculated number of miRNAs per fg of cellular RNA. Indeed, although the per-cell RNA and miRNA pool vary by 10- and 7-fold, respectively, the calculated miRNAs:total RNA ratio slowly and steadily increased by a factor of two between the immature DN3 and mature SP stages. In sum, these results strongly suggest that the total pool of miRNAs is tightly regulated relative to the levels of ribosomal and messenger RNA.

Dynamic Regulation of miRNA Species During Thymocyte Development

The possibility that each stage of T cell development might have a signature miRNA pattern was next investigated. We tested whether the ontogeny relationships of the developing thymocytes to one another could be predicted with the miRNA signature

of these cells. We performed hierarchical clustering of the cloning frequency of miRNAs to infer cell lineage relationships. A ranked correlation coefficient was calculated for all pairwise cell type combinations and used as a similarity matrix to generate a rooted tree (Figure 6A). A second rooted tree was generated using mRNA array data from Hoffman and colleagues, who clustered the same cell types (Hoffmann et al. 2003) (Figure 6B). A multiscale bootstrapping method was used to analyze the clustering in each case (Suzuki and Shimodaira 2006). The two dendrograms revealed strikingly similar clustering patterns. The immature DN3 and DN4 cells clustered together, the mature CD4 and CD8 SP cells clustered together, and DP thymocytes fell at an intermediate position. The DN1 population (which contains several different cell types) clustered in the same location in both dendrograms. These data suggest that the miRNA profile reflects the developmental relationships between individual cells to a similar degree as the mRNA profile.

Examination of the libraries indicated that the vast majority of the 136 known miRNAs that were cloned in this analysis were present at every stage of thymocyte development. However, there were a number of cases in which an individual miRNA's expression was significantly higher in one cell type than in others by Chi square analysis of the frequency data set. The relative abundance of 21 individual miRNAs varied among the six cell populations in a statistically significant fashion (Figure 6C). Statistically significant changes were heavily biased towards overrepresentation, with twenty events in which miRNAs were significantly enriched in a specific cell type compared to five events in which miRNAs were significantly depleted. Three of the five depletions observed (miR-15b, miR-16, and miR-181a) were at the DN1 stage, suggesting upregulation of these miRNAs upon commitment to the T lineage. The other two depletions were striking in that they occurred at intermediate stages of T lymphocyte development. Most striking was miR-142-3p, which was downregulated over six fold at

the DN3 stage relative to the temporally adjacent DN1 and DN4 stages. The uniqueness of this event in regards to global miRNA expression patterns led us to examine it more closely.

Expression of the miR-142 primary transcript (pri-miR-142) was first examined. Surprisingly, qRT-PCR analysis of this transcript revealed that there was no significant reduction in relative expression between the DN1 and DN3 stages (Figure 6D). Editing of the pri-miR-142 transcript by the p110 isoform of adenosine deaminase 1 (ADAR1) has been shown to inhibit processing of this transcript to the pre-miRNA hairpin by the Drosha enzyme (Yang et al. 2006). Since ADAR activity has been demonstrated in rat thymus (Paul and Bass 1998; Yang et al. 2003), we wondered whether we could attribute the specific downregulation of miR-142-3p at the DN3 stage to the activity of this enzyme. Quantitative RT-PCR analysis demonstrated an increase in expression of ADAR1p110 between the DN1 and DN3 stages of development, concomitant with a decrease in expression of ADAR1p150 and ADAR2 (Figure 6E). Strikingly, although pre-miR-142 expression was apparently reduced at the DN4 stage, we noticed an increase in miR-142 expression at this stage consistent with a re-upregulation of ADAR2 at this stage. Later increases in miR-142 expression correlated with an increase in pri-miR142 expression and/or a higher relative expression of ADAR2 or ADAR1p150 relative to the ADARp110 isoform.

The miRNA expression dynamics during T lymphocyte development are largely consistent with a mode of expression in which basally expressed miRNAs are specifically enriched at given stages. In contrast, stage-specific depletion is much more rare, and in the case examined appears to be controlled at the post-transcriptional level.

Genomic Signatures of miRNA Families at Individual Stages of Thymocyte Development

Several studies have reported that mRNAs containing seed matches to tissue-specific miRNAs tend to be expressed at lower levels in the tissue(s) where the miRNA is expressed than in other tissues (Brennecke et al. 2005a; Farh et al. 2005; Krutzfeldt et al. 2005; Stark et al. 2005). We analyzed the relationship between stage-specific enrichment of miRNAs during T lymphocyte development and mRNA expression in these stages using the microarray data of Hoffman et al (Hoffmann et al. 2003).

The mRNA array data were normalized to reflect the level of relative expression of a transcript in a particular stage as compared to the mean expression of all six stages examined. In this representation, one might expect that genes whose mRNA levels changed most between stages would be enriched for large-scale differences controlled at the level of transcription, obscuring the more subtle effects that have been observed for miRNAs in a global analysis. We therefore eliminated the five percent of genes exhibiting the greatest degree of enrichment and the five percent of genes exhibiting the greatest degree of depletion at each stage, examining the cumulative distribution of the remaining 90% of transcripts at each stage using a non-parametric Kolmogorov-Smirnov (KS) statistic. The hypothesis that relative expression of genes with UTR elements harboring seed matches was increased or decreased in each cell type was tested.

The largest enrichment in cloning frequency observed was for miR-181a at the DP stage of thymocyte maturation; we chose this miRNA for a preliminary analysis. The most significant depletion of predicted miR-181 family targets at the DP stage was observed when targets were defined as all transcripts harboring a sequence corresponding to positions 2-8 of the miR-181 family seed and compared to a set of control targets harboring a similar numbers of seed matches to a control sequence ($p = 0.003$, Figure 7A). Interestingly, there was also a significant enrichment of predicted

miR-181a targets at the DN3 stage, with no enrichment or depletion occurring at the DN1, DN4, CD4 SP, or CD8 SP stages.

We extended our analysis to other miRNA families that were statistically enriched or depleted at various stages of thymocyte development, using the same target definition for the broader analysis. Again, the aggregate expression of all miRNA family members in a given cell type (e.g. summing the expression of miR-15a, miR-15b, and miR-16, all of which share a common seed) was considered rather than the abundance of individual statistically enriched or depleted miRNAs. Of 91 seed families tested, 20 were demonstrated to vary significantly over the course of thymocyte development (Figure 7B).

When requiring a p value of less than 0.05 in the KS test, 26 enrichments or depletions were observed in a set of 120 comparisons (20 miRNA family aggregates in six stages) (Figure 7B). These enrichment and depletion events were identified for fourteen of twenty significantly changed miRNA families, with depletion of seed match-containing transcripts for at least one significantly enriched miRNA family in each cell type. (Figure 7B). Relative miRNA expression was ranked in each cell stage and enrichment and depletion events were examined in regards to these levels. Fully 50% of target depletion events were observed at the stage at which the miRNA was most highly expressed, while no enrichment events were observed at this stage (Figure 7C). Surprisingly, the majority of enrichment events for predicted targets were observed at the stage in which the miRNA was expressed at the second highest level. We took advantage of the temporal arrangement of the various thymocyte developmental stages to address this observation.

The general pattern of expression of miRNA families mirrored that of individual miRNAs, characterized by constitutive expression with significant enrichments at a single stage of thymocyte development. Examination of the enrichment and depletion of

predicted targets revealed an overrepresentation of depletion events at the cell stages prior to, at, or following which the highest miRNA expression was observed (Figure 7D). Closer examination of these adjacent stages revealed that target enrichment only occurred prior to, but never at or following the cell stage with the highest level of miRNA expression (Figure 7E).

The observed patterns of miRNA expression are consistent with three modes of activity. In the first case, (miR-181 family) depletion of predicted targets is observed at the stage where the miRNA is acutely enriched. In the second (miR-25,92 family), a more gradual upregulation of the miRNA results in depletion of predicted targets prior to the maximal expression of the miRNA. In the final case, enrichment of miR-142-5p at the DN1 and DN4 stages appears to hold predicted miR-142-5p targets at a background level, with depletion of this miRNA at DN3 allowing enrichment of predicted targets at this stage. Importantly, all three of these potential modes of action are consistent with negative effects on transcript expression at stages of high miRNA expression. We wondered whether these global effects could be correlated with the regulation of genes known to be involved in thymocyte maturation at specific stages.

miR-181a Represses Expression of Genes with Roles in Thymocyte Maturation Through Their 3' UTR Elements

We focused again on the miR-181 family; this family of miRNAs is strikingly upregulated at the DP stage of thymocyte development, and might thus be expected to most dramatically impact gene expression at this stage. MiR-181 has been reported to be highly expressed in thymocytes (Chen et al. 2004); however, the developmental specificity of this expression has not been described. The Targetscan 3.0 server (Lewis et al. 2005) was used to identify candidate targets of the miR-181 family that might play a role in the processes of positive or negative selection characteristic of this stage of T

cell development. Notably, CD69 and Bcl-2 are predicted targets of the miR-181a family. Bcl-2 is known to be selectively downregulated at the DP stage of thymocyte development (Gratiot-Deans et al. 1993), but expression levels of this gene are restored following positive selection to the CD4 or CD8 SP stage. CD69 expression is increased on DP cells that have undergone positive or negative selection; this marker has been shown to appear on the surface of stimulated T cells less than thirty minutes following stimulation (Hara et al. 1986). Interestingly, surface expression of the antigen receptor is also increased in DP thymocytes following positive selection; this is known to occur post-transcriptionally (Bonifacino et al. 1990; Maguire et al. 1990). Although neither the alpha nor the beta chain of T cell receptor is a predicted miRNA target on the Targetscan server, we noticed a seed match to the miR-181 family in the 3' UTR of the TCR α transcript. The precise location of this seed match is conserved to rat, and the human TCR α UTR features a miR-181 seed match nearby. We wondered whether miR-181 might control expression of Bcl-2, CD69, the TCR α chain, and other genes with known or putative roles in positive selection through interactions with the 3' UTR elements of their transcripts

The 3' UTR elements of several predicted targets of miR-181a were fused to a luciferase reporter, and we examined the ability of a synthetic miR-181a siRNA duplex to repress the expression of these constructs in HeLa cells. Relative to transfection of a control siRNA, miR-181a significantly downregulated reporters fused to the 3' UTRs of the TCR α chain, CD69, and Bcl2 (Figure 8). The downregulation of Bcl-2 was marginally enhanced but not dependent upon the presence of the AU-rich stability element (ARE) in the 3' UTR of this gene. Expression of reporters fused to the predicted miR-181 targets TOX, Runx1, EGR1, Bcl2AF1, and FoxP1 were not significantly downregulated by transfection of synthetic miR-181a (Figure 8), even though each of

these genes is a predicted target of the miR-181 family and contain similar seed matches. We concluded that miR-181a is able to selectively repress the expression of a reporter gene fused to the 3' UTR elements of the Bcl-2, TCR α , or CD69 genes.

Dysregulation of miR-181a Targets in Dicer-deficient DP Thymocytes

The finding that miR-181a was able to directly repress the expression of a reporter gene fused to the 3' UTR elements of CD69, Bcl-2, or the TCR α -chain suggested that the expression of these genes would increase if miR-181a levels were decreased in DP thymocytes. Thus, expression of these genes in *Dicer*-deficient DP thymocytes was examined. An *lck-cre* transgene (Hennet et al. 1995) was crossed into the floxed *Dicer* background (Harfe et al. 2005) to effect recombination of *Dcr^f* alleles at the DN3 stage of thymocyte development.

The overall pattern of CD69 expression was conserved in *Dicer*-deficient DP thymocytes. However, these cells exhibited a reproducibly higher median fluorescence intensity when compared to sex-matched littermate controls harboring a functional allele of *Dicer* (Figure 9A). Similarly, the T cell receptor (TCR) was expressed at a higher level in *Dicer*-deficient cells (Figure 9B). Interestingly, the relative increases in expression of CD69 and the TCR in *Dicer*-deficient DP thymocytes observed *in vivo* correlated with the relative repression observed in the above 3' UTR assays.

Finally, Bcl-2 expression in *Dicer*-deficient DP thymocytes was examined. *Dicer*-deficient DP thymocytes expressed levels of Bcl-2 protein markedly lower than those observed in control cells. Consistent with a previous study (Cobb et al. 2005), *Dicer*-deficient thymocytes were more apoptotic than control cells. We suggest that loss of *Dicer* function induces a transcriptional program in DP thymocytes that obscures any specific effect of miR-181a on levels of Bcl-2 protein expression.

Modulation of Potential miR-181a Targets by Transfection of Nucleic Acids into Dicer-deficient DP Thymocytes

The increased expression of CD69 and the TCR in *Dicer*-deficient DP thymocytes is consistent with the regulation of these gene products by miR-181a and miR-181b, both of which are specifically upregulated at the DP stage. However, *Dicer* deletion would be expected to reduce the levels of all miRNAs. Thus, the increase in expression of CD69 and the TCR in DP thymocytes might be due to reasons other than the reduction in levels of the miR-181a family. To demonstrate direct regulation of CD69 and the TCR in DP thymocytes by the miR-181 family *Dicer*-deficient DP thymocytes were isolated from mice and cotransfected with miR-181a or control siRNAs in addition to a GFP marker plasmid. There was a reproducible reduction in the surface levels of CD69 in GFP positive thymocytes when compared to GFP positive thymocytes that had been transfected with a control siRNA duplex (Figure 9C and D; Figure 10). A similar repression of the TCR complex was not observed within the time frame of the experiment, potentially due to the fact that the fully assembled T cell receptor has an exceptionally long half-life (Minami et al. 1987). We concluded that in *Dicer*-deficient DP cells, specific introduction of miR-181a is able to repress CD69 expression. The reporter assays demonstrate that miR-181a is able to effect this repression directly through the 3' UTR elements of these genes. Thus, it is likely that miR-181 family members directly regulate the levels of CD69 and the TCR through their 3' UTR elements at the DP stage of thymocyte development.

Discussion

We quantitatively profiled short RNA species from sequential stages of development of a single mammalian cell lineage, documenting dynamic regulation of individual miRNAs and miRNA families during this process. This regulation is correlated

to changes in gene expression at the genomic level in the processes of lineage commitment, stepwise maturation, and the developmental decision to mature to one of two related but highly different cell types. The functional significance of these dynamics is demonstrated in the context of miR-181a and its predicted targets.

The degree of variation in the total miRNA pool and total cellular RNA content across the T lymphocyte developmental progression is striking. The miRNA pool varies per cell from a minimum of 5,000 molecules per cell in DP thymocytes to 33,000 molecules per cell in DN4 cells. At the same time, total RNA content ranges from a low of 0.7 pg of RNA per cell in DP thymocytes to a maximum of 6.8 pg of RNA per cell in DN4 thymocytes. The major changes that we observe in per-cell miRNA pool and total RNA content stand in stark contrast to the constant, progressive increase in the ratio of the total miRNA pool to the total RNA pool. This progressive increase is consistent with previous studies in which overall miRNA levels are correlated to the level of differentiation of a cell (Lu et al. 2005), but our data also suggest that global miRNA levels are tightly coupled to cytoplasmic volume and/or ribosomal content. Cells that are rapidly proliferating (e.g. DN4 cells) are likely to be more transcriptionally and translationally active. The corresponding increase in the number of transcripts and level of translation would necessitate an increase in miRNA levels to maintain the same degree of control of the transcriptome. Although the absolute copy number of miR-181a is lower in DP cells than at any other developmental stage, it represents a higher fraction of the miRNA pool. It is only in the context of relative concentration that miR-181a mediates its biological effect at the DP stage of development. Thus, the concentration of an individual miRNA in relation to its targets is likely to be more important than its absolute copy number.

In relative terms, the vast majority of short RNAs were expressed at constant levels throughout thymocyte development. Of 136 miRNAs that were observed in the

study, only 21 were observed to vary by cell type. Consideration of the aggregate expression of miRNA family members yielded largely consistent observations in regards to miRNA expression dynamics. However, there were instances in which the aggregate miRNA family revealed additional enrichment (miR-23) or changed the stage at which significance was observed (miR-15 family). The significantly changed miRNAs and miRNA families were continually expressed throughout thymocyte development, with transitory enrichments at specific cell stages. Curiously, while a high level of enrichment for at least one miRNA or miRNA family was observed at each distinct stage of thymocyte development, the reverse pattern of specific depletion at a single stage was much more rare. A major depletion in the cloning frequency of both products of the *miR-142* gene at the DN3 stage is one exception (Figure 6, Figure 7). The relative expression of various ADAR family members at and adjacent to this developmental stage coupled with the lack of transcriptional downregulation of pre-miR-142 at the DN3 stage suggest a model in which ADAR2 and/or ADAR1p150 compete with ADAR1p110 to inhibit editing of pri-miR-142, ensuring proper maturation of the mature miR-142 miRNAs. Indeed, the perturbed pri-miR-142 editing patterns described in ADAR null mutants are consistent with this model (Yang et al. 2006). This may partially explain why this miRNA is an exception to the rule of basal expression with stage-specific enrichment observed for other miRNA species.

Statistically significant changes in the relative levels of expression of predicted mRNA targets were observed for 14 of 20 dynamically regulated miRNA families and for at least one miRNA family in each developmental stage examined. Consistent with a role in directly repressing the levels of proteins encoded by target mRNAs or reinforcing transcriptional downregulations, in most cases underrepresentation of predicted targets was observed around local maxima in miRNA family expression. This strong correlation indicates that regulation by miRNAs is a characteristic of many intermediate

developmental stages, and not solely of mature, differentiated cells or “one-way” developmental specification.

The dispensability of the RNAi pathway for thymocyte maturation has been suggested by conditional deletion of *Dicer* in the T lineage (Cobb et al. 2005; Muljo et al. 2005). However, it is unclear from these studies whether individual cells had completely lost Dicer activity and miRNA function at the point that positive selection and lineage choice occurred. Indeed, even using the early-acting *lck-cre* transgene, DP cells exhibit detectable levels of mature miRNAs (Cobb et al. 2005). The 10-fold reduction in thymocyte number in these mice at stages immediately following deletion of the floxed locus (Cobb et al. 2005), coupled with a marked increase in apoptosis of *Dicer*-deficient thymocytes both *in vivo* and *ex vivo* (Cobb et al. 2005), leads us to believe that elimination of *Dicer* is quite deleterious, manifesting after Dicer protein, and the miRNA pool have been depleted. Indeed, peripheral T cells in *lck-cre Dicer^{ff}* mice are enriched for non-recombined alleles of *Dicer^f*.

In any case, it is clear that even individual miRNAs can influence hematopoietic development. Mice reconstituted with cells transduced with a miR-181a overexpression vector exhibit a paucity of T lymphocytes (Chen et al. 2004). While this was interpreted as a facilitated commitment to the B lineage, the results of this study provide an alternative explanation. Disruption of CD69 signaling has been shown to negatively impact the egress of lymphocytes from lymphoid organs, including the thymus (Nakayama et al. 2002; Alfonso et al. 2006; Shioh et al. 2006). Forced expression of miR-181 past the DP stage of thymocyte development would be expected to decrease CD69 levels on positively selected thymocytes, resulting in retention of these cells in the thymus and an apparent decrease in peripheral T cells. Our data in regards to miR-181a's effect on the expression of Bcl-2 and the TCR *in vivo* are less complete. However, the effects of forced expression of miR-181 on these gene products would

further exacerbate the phenotype resulting from lowered CD69 expression. Reduced expression of the TCR would be expected to shift the threshold for positive and negative selection, while an inability to increase Bcl2 expression upon positive selection would result in cell death. Recent work by Li and colleagues suggest that miR-181 can increase the sensitivity of DP cells to stimulation of the T-cell receptor (TCR) (Li et al. 2007). Blocking miR-181 in DP cells suppresses both positive and negative selection. miR-181 represses the expression of a group of protein phosphatases, which are negative regulators of two TCR signaling molecules, Lck and Erk (Li et al. 2007). Thus the expression of miR-181 in DP cells increases the basal phosphorylation levels of Lck and Erk, reduces their activation threshold, and enhances TCR signaling strength (Li et al. 2007).

Returning to the more global observations of this study, the lower relative degree of enrichment or depletion of predicted targets and subtle effects on gene expression in this study, (particularly in regards to similar computational studies) at first seem unremarkable. However, other studies have compared gene expression profiles in tissues specifically expressing a given miRNA or from cell lines in which an ectopic miRNA is introduced. This analysis stands in stark contrast for the reason that nearly every miRNA that we observe is expressed throughout T lymphocyte development, with graded increases or decreases at specific times during this process. This pattern of constant expression with transient enrichment is not consistent with the “switch like” function that has been attributed to miRNAs with known developmental roles (Reinhart et al. 2000; Fazi et al. 2005; Giraldez et al. 2006), and suggests two things: First, any individual miRNA would be expected to exert some level of post-transcriptional control at all times, with more or less robust effects during times of significant enrichment or depletion, respectively. It follows that these miRNAs would not be expected to dramatically effect gene activation and repression, instead exerting most of their effects

to modulate gene expression at an intermediate level. Second, the role of individual miRNAs and miRNA families extends beyond the temporal window of their first expression, continuing to play an active role in fundamental processes in subsequent cell stages.

Methods

Mice

C57/BL6 mice were from Taconic or Jackson labs. The *Dicer*^{ff} mouse has been described (Harfe et al. 2005). The *lck-cre* transgenic mouse was from the Jackson laboratories. All mice were housed, handled, and euthanized in accordance with federal and institutional guidelines.

Thymocyte and RNA isolation

Thymocytes were isolated from female mice aged 5-10 weeks and stained with antibodies from BD-Pharmingen. Populations were designated as follows: CD4SP - CD4⁺CD8⁺TCR^{hi}, CD8SP - CD8⁺CD4⁻TCR^{hi}, DP - CD4⁺CD8⁺TCR^{med/lo}, DN - CD4⁻CD8⁻CD19⁻GR1⁻γδ⁻TER-119⁻CD11b⁻DX5⁻. DN cells were further divided as follows: DN1 - CD44⁺CD25⁻, DN3 - CD44⁻CD25⁺, DN4 - CD25⁻CD44⁻. Cells were sorted to > 95% purity and processed using Ambion's miRVana kit as per manufacturer's instructions to isolate both short and long fractions of RNA. For preparation of bulk DP cell RNA in the direct quantification experiments, we used a one step positive selection with biotinylated anti-CD8 and anti-biotin magnetic beads (Miltenyi). This routinely resulted in purities of greater than 95%.

Short RNA Cloning

Short RNAs were cloned essentially as described (Lagos-Quintana et al. 2001; Lau et al. 2001). We cloned directly from the short RNA fraction eluted from miRvana columns. There was no gel isolation between the 5' and 3' adapter ligations steps. Following RT (Superscript III, InVitrogen) and ten cycles of PCR amplification with AmpliTaq Gold

(Perkin Elmer), we digested our amplifications with the appropriate enzyme (Stu I or Pvu II, both from NEB) and gel isolated from a denaturing polyacrylamide gel using a 10 bp ladder (*In vitro*gen) and SyBR Gold (Molecular Probes). After the first round of amplification, we switched to Pfu Turbo, iterating amplification, digestion, and gel isolation until cloneable product could be identified. For this study, we amplified various libraries anywhere from 30 to 38 cycles in total. Bioinformatic extraction of individual short RNA clones from sequence reads was as described (Houbaviy et al. 2003).

3'linkers: Stu: 5'-CCTGTATCTGTGTATGGddC-3'; Pvu:5'-CTGG-TATCTGTGTATGGddC-3'.

5' linkers: Stu: 5'-ACCACAGAGAAACCGrArGrG-3'; Pvu: 5'-

ACCACAGAGAAACCGrCrArG-3'. 3'RTprime/PCRoligos: Stu: 5'-GACTAGCTTGGTGCC-ATACACAGATACAGG-3'; Pvu: 5'-GACTAGCTTGGTGCCATACACAGATACCAG-3'. 5'

PCR oligos: Stu: 5'-GAGCCAACAGGCACCACAGAGAAACCGAGG-3'; Pvu: 5'-GAGCCA-ACAGGCACCACAGAGAAACCGCAG-3'. A detailed protocol is available on request.

Direct miRNA quantitation and calculation of miRNA copy number per cell

Thymocytes were sorted to >95% purity, directly counted in a hemacytometer, and lysed in Trizol. Five fmol of miR-196a was spiked into each lysis and later detected as a recovery control. Initial experiments in DP thymocytes were performed without a recovery control. In later experiments, the average recovery for four independent isolations of DP thymocytes was calculated, and data from earlier DP quantitations was normalized to this value. RNA was processed as per manufacturer's instructions, quantitated on a Nanodrop spectrophotometer, and assayed using the Trilogy Assay (U.S. Genomics) essentially as described (Neely et al. 2006). The molarity of a given miRNA in each sample was determined by fitting the coincident events in the RNA sample to a curve generated with a synthetic miRNA template in a complex background.

Phylogenetic Analysis

We used the complete linkage method of hierarchical clustering in the R software package to generate rooted trees from miRNA and mRNA expression data. To eliminate noise, we omitted miRNAs cloned less than five times overall and set mRNA array values less than $\log_2(20)$ to $\log_2(20)$. Significance of clustering was analyzed with the Pvcust package (Suzuki and Shimodaira 2006).

qRT-PCR Analysis of ADAR Expression

RNA was extracted with Trizol, DNase-I digested, and reverse-transcribed using the SuperScript III kit. Quantitative Real-Time PCR was performed using a SYBR-Green kit (Applied Biosystems) on a ABI 7500 instrument using the following oligos: A1LF 5'-GACTAC-GCGTTGGGACTAGC-3'; A1LR 5'-TGCTGAAGCTGGAAACTCCT -3'; A1SF 5'-CTTGCC-GGCACTATGTCTC-3'; A1SR 5'-TGCTGAAGCTGGAAACTCCT-3'; AB1F 5'-CCAGTCAA-GAAGCCCTCAAA-3'; AB1R 5'-GCGGTA CTTGGAGTGACCAT-3'.

Changes in relative expression of ADARp110 and ADARp150 were similar whether TBP or Tubulin were used as a control. The pri-miR-142 oligos have been described (Yang et al. 2006).

Luciferase assays

HeLa Cells were transfected at 80% confluency in 24 well plates with 100 nM siRNA, 100 ng of pGL3, 700 ng of Renilla/UTR reporter. Cells were split 1:3 4-6 hours after transfection and assayed at 48h in a Dual Luciferase Assay (Promega). All results were normalized to the effect of siRNAs on pRL-TK-RenCX6X (Doench et al. 2003). Si181a is the predicted post-Dicer processing product of the miR-181a pre-miRNA hairpin. UTRs were amplified with the following oligos and subcloned into the Not I and Xho I

sites of pRL-TK-CX6x: ToxF 5'-AGCTAGCTCGAGCATGTGAGCTTGTGG- GTCAC-3',
ToxR 5'-ATGCATGCGGCCGC- AGCACTTAGCTAGCGCGTTC-3', RunxF 5'-
AGCTAGCTCGAGTGCATCTGGGTGGTCA- TTTA-3', RunxR 5'-
ATGCATGCGGCCGCT- TGGATCTTTGGGGTACAGC-3', CD69F 5'-
AGCTAGCTCGAGACTGTGCCATAGCACC-ACAG-3', CD69R 5'-
ATGCATGCGGCCGCA- CAGCTTAACTTTATAGTGGGTTTT-3', EGRF 5'-
AGCTAGCTCGAGCATCTGTGCCAT- GGATTTTG-3', EGRR 5'-
ATGCATGCGGCCGCTATCCCATGGGCAATAGAGC-3', Foxp1F 5'-
AGCTAGCTCGAGAGACCGAAGATTGGGGAAAA-3', Foxp1R 5'-ATGCATGCGGCC-
GCTG AGGTCAGAACTTAAA 5'-ATG-3', Bcl2af1F 5'-AGCTAGCTCGAGGCAAACATA-
AGGAGGACAGCTT-3', Bcl2af1R 5'-ATGCATGCGGCCGCAGGGGAGCATCATGCAA-
TAC-3', TcrbF 5'-AGCTAGCTCGAGTATGCATCCTGAGCCGTTCT-3', TcrbR 5'-
ATGCA-TGCGGCCGCTCCATGTTTTTATTGATTTAGTCTG-3', TcraF2 5'-
AGCTAGCTCGAG-GCAAGACTGACAGAGCCTGA-3', TcraR2 5'-
ATGCATGCGGCCGCGAATCACCTTTAA-TGATGTCATGG-3'.

Transfection of DP thymocytes

Whole thymocytes were isolated and nucleofected using the Primary T cell Nucleofection Kit (Amaxa) and program X-001. 5×10^6 cells were nucleofected with a 1.5-3 μ g of siRNA duplexes and 1 μ g pMaxGFP. Levels of surface markers were monitored at 12-16h post-nucleofection.

Table 1. Short RNA cloning statistic.

The number of clones falling into a particular annotation class is listed for each cell type. The representation of the class within the library derived from an individual cell type is listed in parentheses. Column totals do not reflect the sum of annotated clones or percentages in a given cell type due to overlap among the Known Gene, EST Repeat, and CpG classes in the UCSC annotation database. miRNA[#] denotes any fragment of a miRNA hairpin other than the mature miRNA, e.g. the miRNA* strand or loop.

Type	DN1	DN3	DN4	DP	CD4	CD8	Total
miRNA	1426 (58.0%)	404 (56.0%)	1764 (72.1%)	742 (82.4%)	873 (67.8%)	914 (71.4%)	6123
miRNA [#]	34 (1.4%)	18 (2.5%)	37 (1.5%)	7 (0.8%)	18 (1.4%)	15 (1.2%)	129
tRNA	161 (6.6%)	82 (11.4%)	221 (9.0%)	53 (5.9%)	83 (6.5%)	64 (5.0%)	664
rRNA	29 (1.2%)	7 (1.0%)	24 (1.0%)	3 (0.3%)	3 (0.2%)	0 (0.0%)	66
Other ncRNA	58 (2.4%)	20 (2.8%)	86 (3.5%)	9 (1.0%)	35 (2.7%)	12 (0.9%)	220
Known gene	151 (6.1%)	34 (4.7%)	41 (1.7%)	7 (0.8%)	41 (3.2%)	45 (3.5%)	319
EST	371 (15.1%)	113 (15.7%)	200 (8.2%)	67 (7.4%)	171 (13.3%)	169 (13.2%)	1091
Repeat	367 (14.9%)	118 (16.3%)	196 (8.0%)	66 (7.3%)	188 (14.6%)	188 (14.7%)	1123
CpG	37 (1.5%)	8 (1.1%)	8 (0.3%)	1 (0.1%)	11 (0.9%)	8 (0.6%)	73
Unannotated	289 (11.8%)	58 (8.0%)	84 (3.4%)	10 (1.1%)	65 (5.1%)	71 (5.5%)	577
Total	2458	722	2445	900	1287	1280	9092

Table 2. Calculation of copy numbers of individual miRNAs in CD4⁺CD8⁺ thymocytes.

miRNAs were quantitated by solution hybridization and direct detection using a U.S. Genomics Trilogy platform. In each individual experiment, a standard curve was plotted from triplicate standards of known concentration of a synthetic target in a complex background. Values obtained for triplicate samples of DP thymocyte RNA were fitted to this curve to determine moles per microgram, and then molecules per cell. In later experiments, we performed a recovery control by detecting a synthetic miR-196a that had been “spiked” into the Trizol during cell lysis. The average recovery from these latter experiments (~75%) was used to estimate the recovery of earlier RNA isolations, and initial measurements were scaled by this factor. The plot of copy number per cell vs. frequency of cloning in DP thymocytes in Figure 4 is a graphical representation of the above analysis. The data represent 43 individual experiments performed in triplicate from any one of seven distinct preparations of DP total RNA. Results of two independent measurements of miR-16 and miR-21 in HeLa cells are also shown.

DP thymocytes

miRNA	Mean copy #/cell	SEM (n)	Cloning Frequency
miR-132	0.0	0.0 (3)	0.000
miR-219	1.8	6.2 (3)	0.001
miR-28	2.5	1.2 (6)	0.001
miR-126	4.3	3.1 (3)	0.000
miR-301	7.5	0.3 (3)	0.002
miR-125b	10.5	4.4 (9)	0.000
miR-181a*	12.0	1.1 (3)	0.003
miR-24	34.3	2.5 (12)	0.004
miR-19a	52.6	3.4 (3)	0.007
miR-150	68.0	3.7 (9)	0.010
miR-21	70.0	8.3 (12)	0.016
miR-25	80.2	11.3 (12)	0.014
miR-93	81.8	9.1 (12)	0.013
miR15a	100.3	20.4 (12)	0.023
miR-15b	219.4	55.9 (6)	0.058
miR-16	403.4	68.9 (6)	0.059
miR-181a	809.9	69.7 (15)	0.155

HeLa Cells

miRNA	Mean copy #/cell	SEM (n)	Cloning Frequency
miR-16	1169.6	312.1 (4)	N/A
miR-21	19132.2	1770.0 (8)	N/A

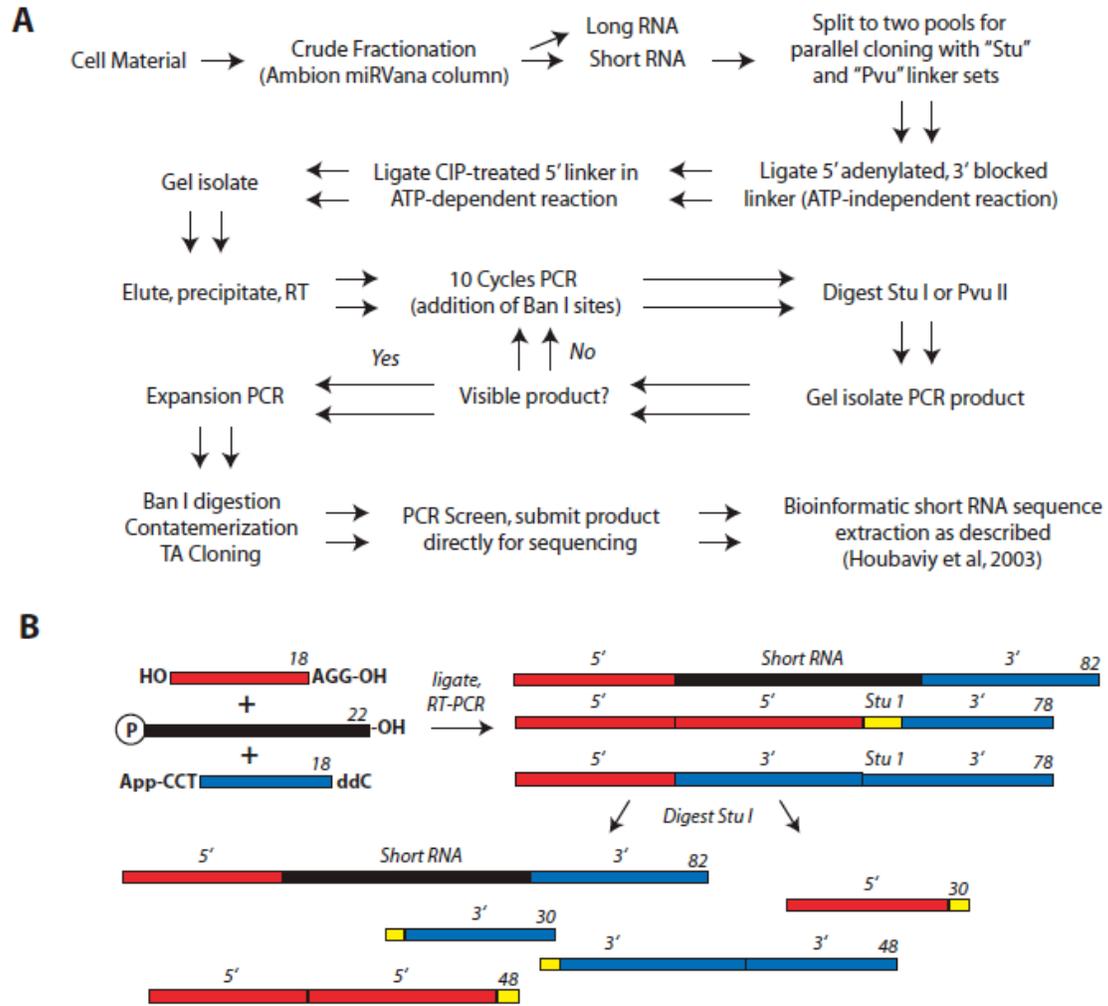


Figure 1. Modified cloning protocol.

(A) In order to generate short RNA libraries from limiting (ng) amounts of total RNA, we redesigned linker sets with “blunt-cutter” restriction half sites at their termini, such that ligation of a 5' linker to a 3' linker would result in a cleavable product. An iterated process of PCR amplification and enzyme digestion was used to obtain clonable product. This parallel strategy also offsets insert loss by single enzyme digestion and reduces library bias stemming from T4 RNA ligase end specificity. In addition, the two libraries can be compared prior to pooling for consistency. (B) Linker design and digestion schematics.

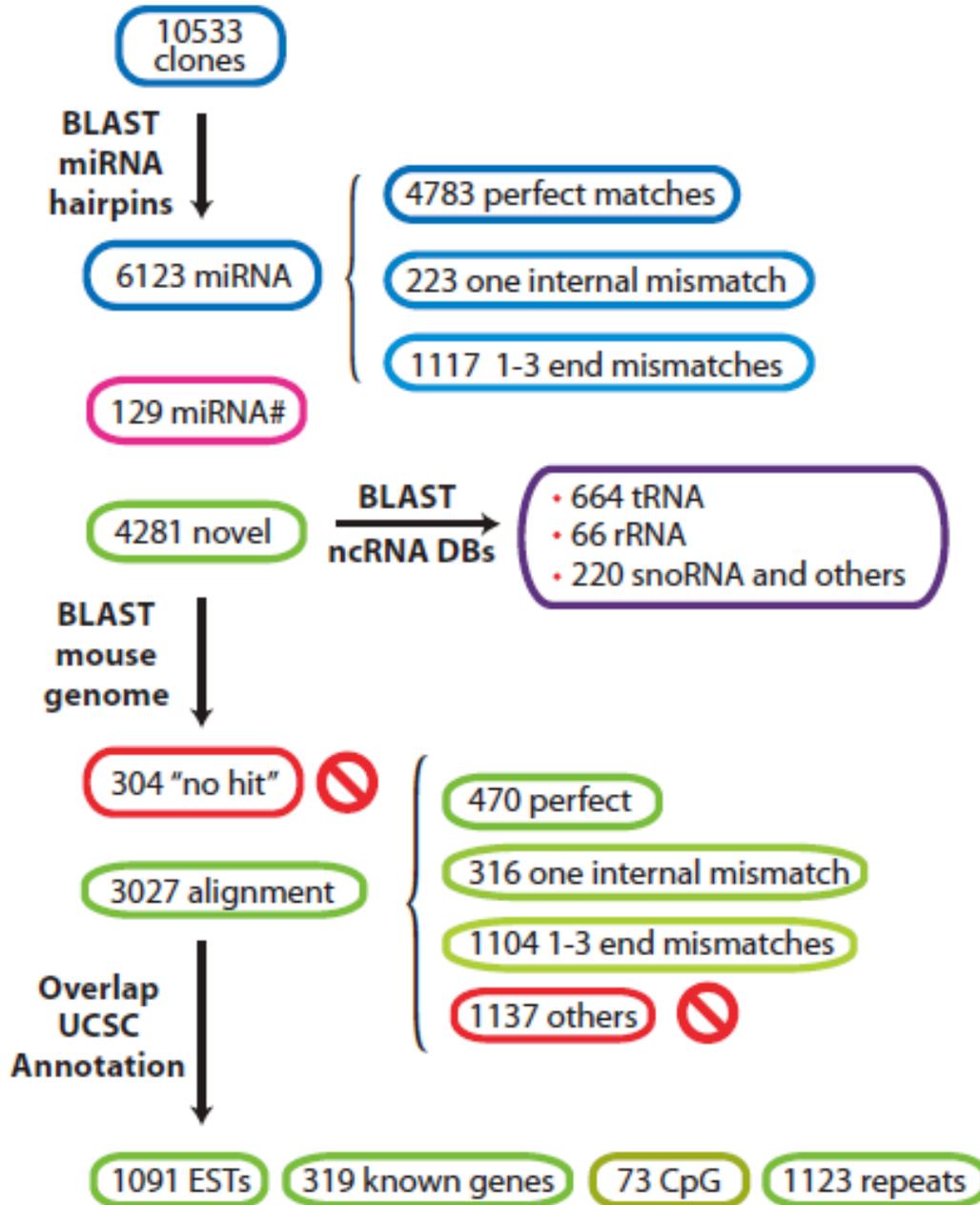


Figure 2. Processing and annotation flowchart for cloned sequences.

Cloning and sequencing of cDNA clones from six stages of T lymphocyte development yielded a total of 10533 clones ranging in length from 16 to 33 nucleotides. We used BLASTN to query these clones against miRNA mirBase (Release 8.1) (Griffiths-Jones 2006), Rfam (Griffiths-Jones et al. 2005), the NONCODE database (Liu et al. 2005a), and the mouse genome (mm8 assembly). In total, we were able to map 9092 (86.3%) of the clones to sequences in the above databases. Roughly two-thirds of these clones (67.3%) match precisely, whereas the balance of

of clones align to known sequences when up to 3 mismatches are allowed. The proportion of sequences not matching precisely is higher than what one would expect by chance, assuming a sequence of 22 nucleotides long, and overestimating the polymerase error rate at one percent. We elected to include sequences in our analysis aligning with one mismatch anywhere in the sequence or any sequence having up to three mismatches at the end. The latter concession allowed automated identification of clones derived from tRNA species, which are post-transcriptionally modified with a 3' CCA motif. In addition, the majority of mismatches observed in miRNA species appeared to be base additions to the 3' end, with a bias towards an adenosine or uracil. This observation has been previously described (Aravin and Tuschl 2005), and is suggestive of post-Dicer modification of miRNA species. We did not observe a bias in composition of the internal mismatches that would suggest known modes of RNA editing. The miRNA[#] class is composed of any fragment of a known miRNA hairpin except the mature miRNA (e.g. loop, *strand). The 1441 clones not aligning to the mouse genome according to our criteria and excluded from the above analysis were queried against the genomes of 369 bacterial species (downloaded from NCBI), *Bos taurus* (bosTau2 assembly) and *Homo sapiens* (hg18 assembly) to check for potential contamination introduced by cloning vehicles, bovine serum and human contact. In addition, we queried the sequence against one mammalian (*Canis familiaris*) and one non-mammalian (*Gallus gallus*) species as specificity controls.

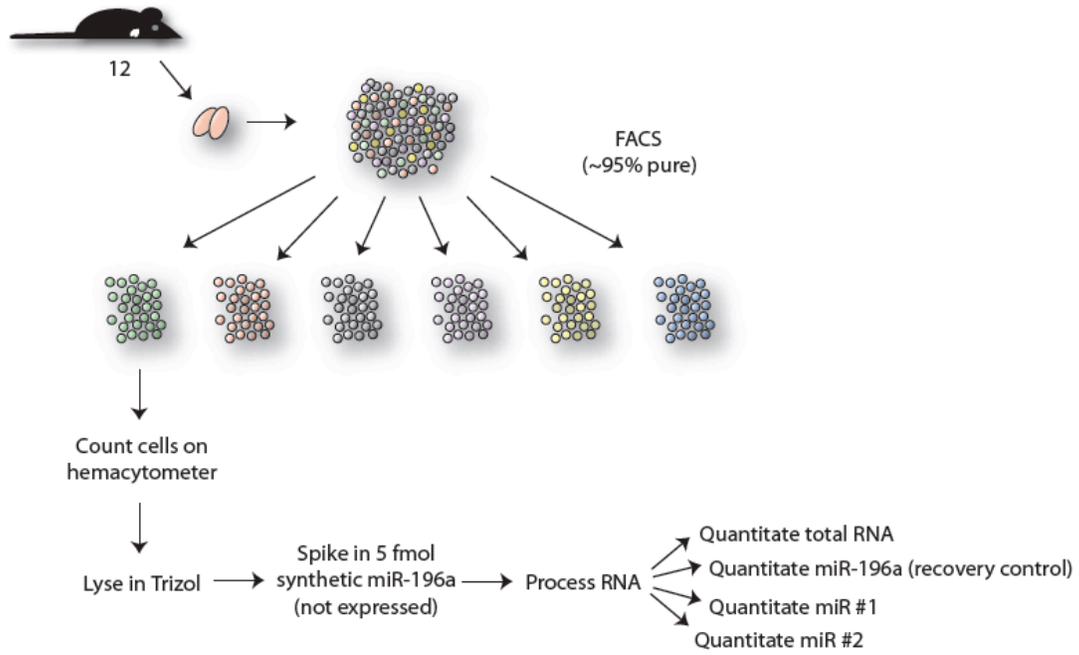


Figure 3. Schematic of longitudinal miRNA quantitation experiments.

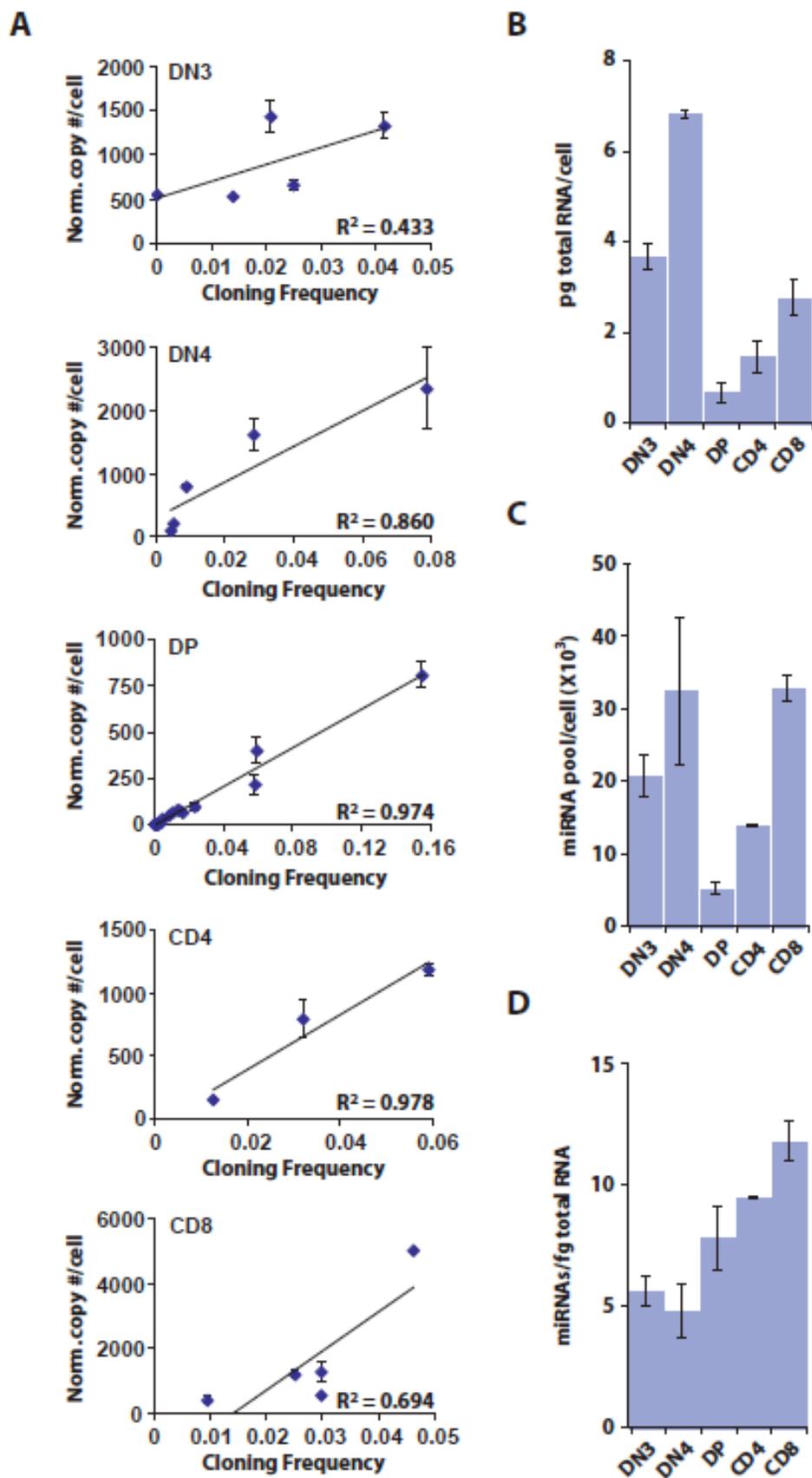


Figure 4. Absolute copy numbers of miRNAs in Thymocyte populations.

(A) Individual miRNAs were quantitated using the US Genomics Trilogy and Trilogy 2020 instruments. Measured values were fitted to a standard curve generated with synthetic miRNA to determine molarity. We used this information along with cell counts and recovery controls to calculate the copy number per cell for individual miRNAs in each cell type. (B) Measured pg total RNA per cell. (C) Calculated miRNA pool per cell. For each cell type, we used the average of the constants calculated for each data point in (A). (D) Calculated miRNAs/fg total RNA for each cell type. For correct propagation of error, error bars in (A) and (D) represent the SEM. Error bars in (B) and (C) represent the SD.

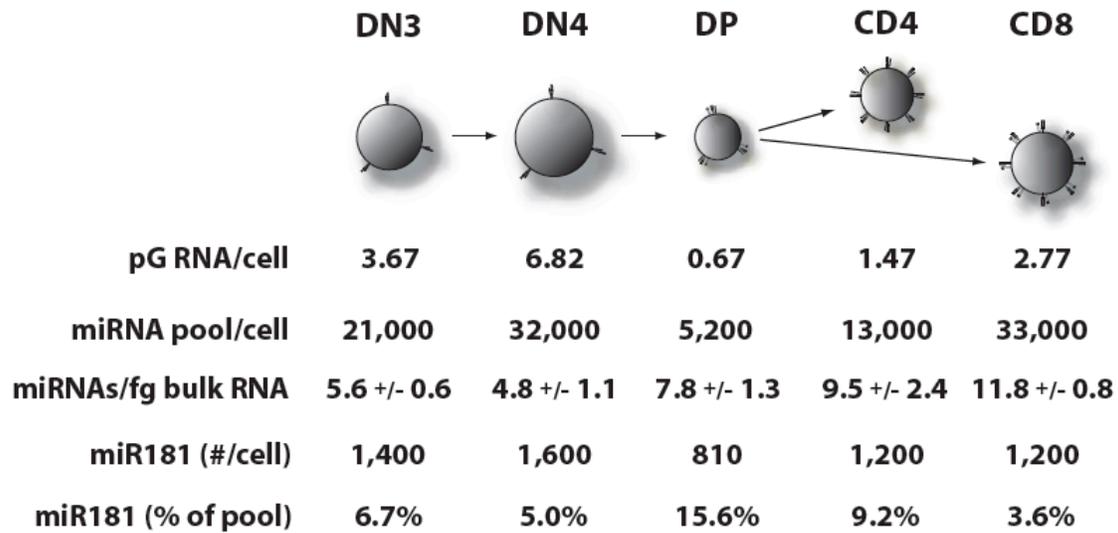


Figure 5. Measurements of total RNA content per cell.

Rooted trees were generated from miRNA and mRNA expression data derived from each thymocyte population. Values were derived from cell counts, total RNA quantification, and measurement of the concentration of individual miRNAs in each cell type normalized to a recovery control. The calculated miRNA pool represents an estimate based on the average of constants relating miRNA copy number to cloning frequency. We were not able to obtain enough DN1 cells to perform miRNA quantitation analysis.

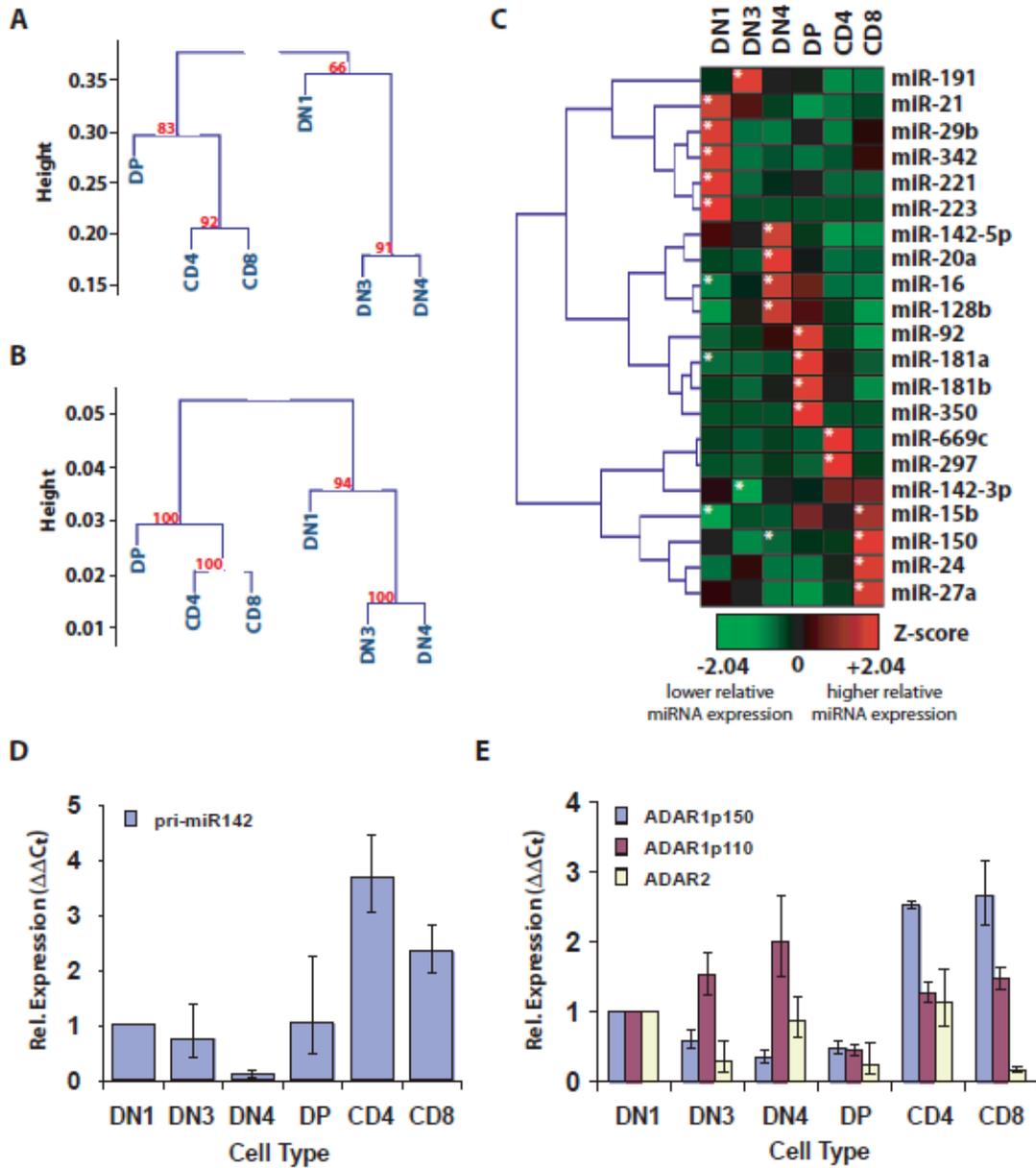


Figure 6. miRNA expression is dynamically regulated during thymocyte development.

(A) Relationship of cells determined by Spearman correlation of miRNA profiles. (B) Relationship of cells determined by Spearman correlation of mRNA profiles (Hoffmann et al. 2003). Values at branch points in (A) and (B) denote multiscale bootstrapping significance values. (C) Heat map representing relative expression of the 21 miRNAs that were identified as enriched or depleted during thymocyte development by the Chi square test. Asterisks reflect statistically significant enrichment or depletion. (D) qRT-PCR analysis of pri-miR-142 in thymocyte populations. (E) qRT-PCR analysis of ADAR family members in thymocyte populations.

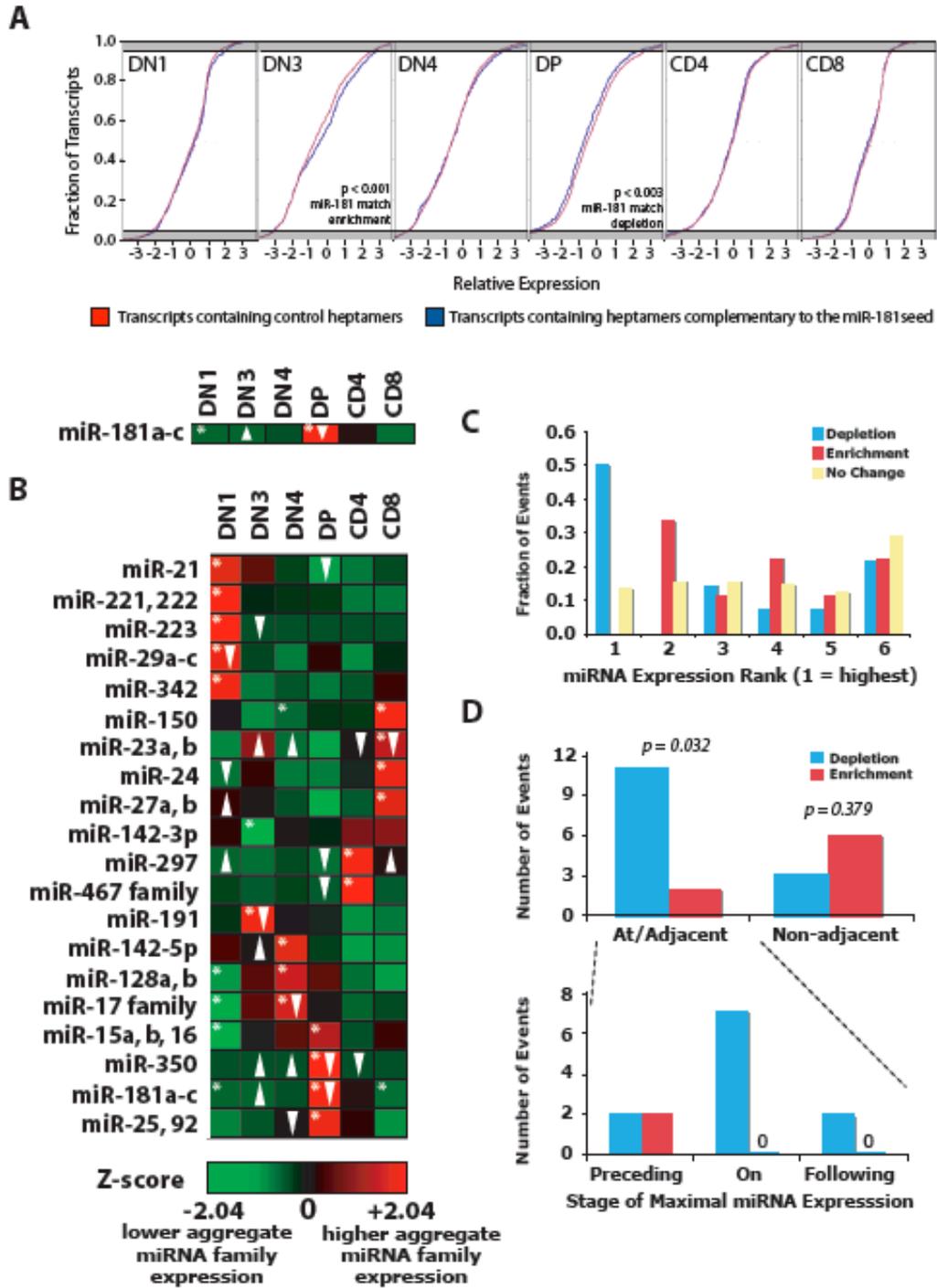


Figure 7. Genomic signatures in response to changes in miRNA family member expression.

(A) Cumulative distribution plots of predicted target expression for miR-181 family members versus control heptamers at each stage of thymocyte development. (B) Chi square analysis of aggregate expression of miRNA family members. Arrows indicate significant ($p < 0.05$)

enrichment (up) or depletion (Lu et al.) of targets relative to control heptamers. Asterisks reflect statistically significant enrichment or depletion of the indicated miRNA family. **(C)** Enrichments, depletions, and no change of predicted targets at each miRNA expression ranking level for the 20 miRNA families shown. **(D)** Temporal analysis of enrichments and depletions during thymocyte development relative to stage of maximum miRNA family aggregate expression. Significance of these enrichments when compared to 1000 randomizations of the data set is shown.

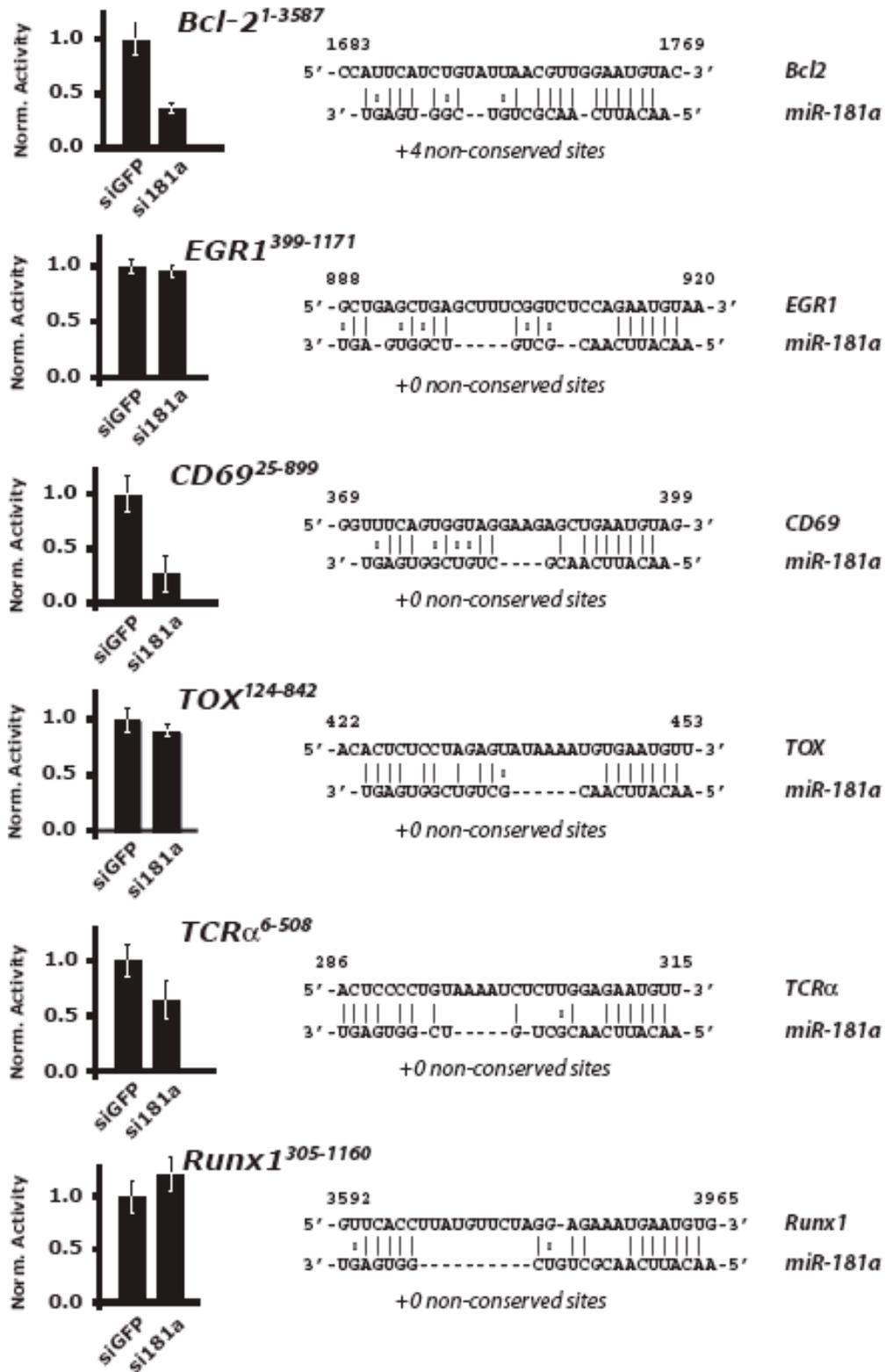


Figure 8. Repression of predicted miR-181a target UTRs in a reporter assay.

The 3' UTR elements of several genes identified by the Targetscan 3.0 server (Lewis et al. 2005) as potential miR-181a targets were tested in a dual-luciferase assay. An alignment of the conserved seed matches are indicated for each predicted target. Graphs indicate expression of the construct in HeLa cells transfected with a control siRNA duplex versus a miR-181a duplex normalized to the effect of each siRNA on a control UTR in a dual luciferase assay. Results are representative of a minimum of three independent experiments performed in triplicate, Error bars represent one SD.

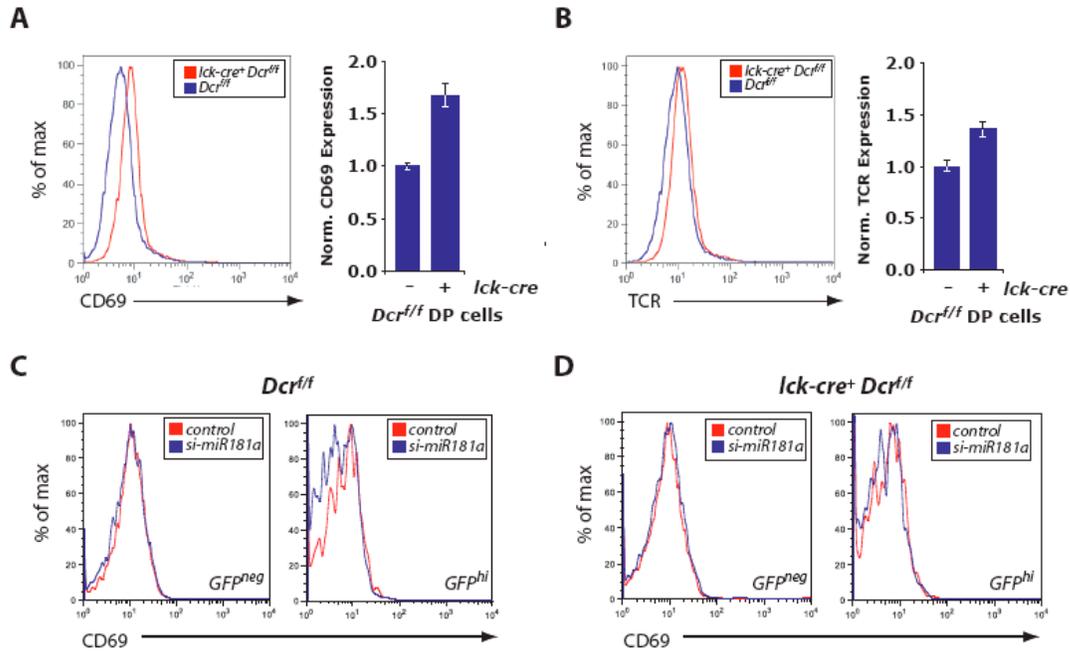


Figure 9. Dysregulation of CD69 and the TCR in Dicer-deficient DP thymocytes.

CD69 (**A**) and TCR (**B**) staining on electronically gated DP thymocytes from *lck-cre*⁺ and control *Dicer*^{ff/ff} mice. Derepression was normalized via median fluorescence intensity (bar graphs). (**C**) CD69 levels on GFP^{hi} and GFP^{neg} *Dicer*^{ff/ff} DP thymocytes transfected with a miR-181a or control siRNA. (**D**) As in C, examining *lck-cre*⁺ *Dicer*^{ff/ff} DP thymocytes. Median fluorescence intensities for each parameter are provided in Figure 10. Results are representative of a minimum of three individual experiments.

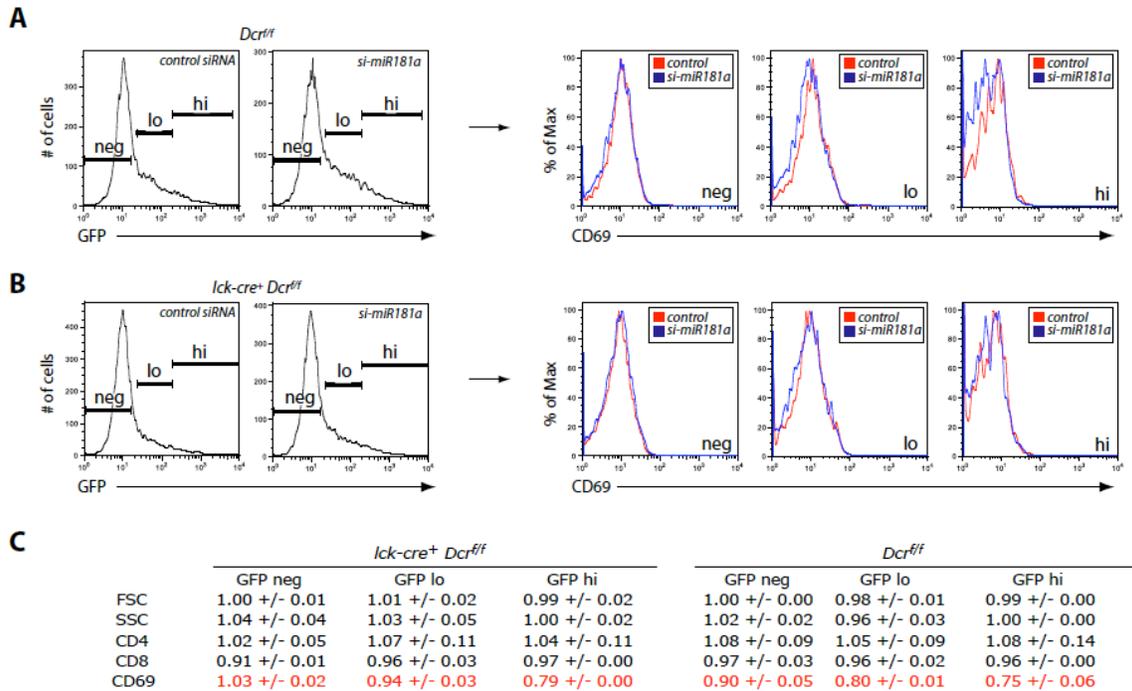


Figure 10. CD69 expression is inversely correlated with GFP intensity in *Dicer*-deficient and control DP thymocytes transfected with miR-181a.

(A) Representative gating of forward- and side-scatter gated GFP negative, low, and high populations in *Dcr^{fl/fl}* thymocytes transfected with control or miR-181a or control siRNAs are shown. The CD69 staining profiles of DP thymocytes from each population are overlaid to the right. (B) Analysis of transfected *lck-cre⁺ Dcr^{fl/fl}* DP thymocytes as in (A). (C) Normalized median fluorescence intensity of populations transfected with si-miR-181a in (A) and (B). The values for each parameter are set relative to the same populations in cells transfected with the control siRNA.

CHAPTER 3: Multiple abundant miRNA families collaborate to buffer embryonic stem cells from apoptosis

The material presented in this chapter was adapted, with permission, from the following manuscript:

Grace X.Y. Zheng, Arvind Ravi, Christopher B. Burge and Philip A. Sharp. Multiple abundant miRNA families collaborate to buffer embryonic stem cells from apoptosis. (*in preparation*)

Experimental contributions:

Arvind Ravi cloned all the luciferase constructs, and performed all the luciferase assays. Grace Zheng performed all the bioinformatics analysis and conducted the cell cycle and apoptosis experiments.

Abstract

MicroRNAs are an important class of short RNAs that play critical roles in post-transcriptional gene regulation. Generated by cleavage of longer hairpin transcripts by the enzyme DICER, they are emerging as key regulators of various mammalian cell types, including embryonic stem cells (ESCs). In order to better understand their functions, we attempted to identify targets of the most highly expressed of these noncoding RNAs in mouse ESCs, which share the common seed “AAGUGC”, nucleotides 2-7 of the miRNA, and include members of the miR-290-295 cluster, the miR-302 cluster, and the miR-467 family. After identifying potential targets by combining bioinformatic predictions with experimental microarray data from both Dicer null (Dcr KO) and miR-290-295 cluster null (295 KO) mouse ESCs, an unbiased gene ontology search suggested that these miRNAs modulate key players in apoptosis. We confirmed this prediction by validating two targets – Caspase 2 and Ei24 – both involved in the mitochondrial apoptosis pathway. Notably, their corresponding AAGUGC miRNAs were protective against apoptosis when transfected into Dcr KO and 295 KO in a state-specific manner, specifically following exposure to doxorubicin or gamma irradiation. These data implicate AAGUGC miRNAs as buffers of a critical ESC decision point between DNA repair and cell death.

Introduction

miRNAs are endogenous ~22nt RNAs that regulate gene expression post-transcriptionally. In animals, the ability of miRNAs to accomplish this regulation depends on complementarity between mature miRNA sequences and their mRNA targets. Most commonly, partial binding of miRNAs leads to destabilization of mRNA transcripts and/or inhibition of productive translation, although in some cases perfect (or near perfect) complementarity instead leads to target cleavage. Both *in vitro* experiments and bioinformatics have shown that matches for positions 2-7 of the miRNA, referred to as the miRNA “seed,” are generally required for effective miRNA-directed mRNA downregulation (Behm-Ansmant et al. 2006b; Bartel 2009).

Since the discovery of the first miRNA gene in the early 1990s, hundreds of miRNAs have been identified across various mammalian cell types through cloning and bioinformatics (Bartel 2009). The roles of miRNAs in mESCs have been of particular interest in recent years, as this knowledge may shed light on key aspects of mammalian development as well as generate useful insights that can be applied to reprogramming and cancer, both of which recapitulate aspects of an ES expression state (Ben-Porath et al. 2008; Lin et al. 2008; Bosnali et al. 2009; Judson et al. 2009). In addition, mESCs can survive in the absence of Dicer, and serve as a unique model system for dissecting miRNA functions (Kanellopoulou et al. 2005; Murchison et al. 2005).

Recent studies have linked several miRNAs to ES cell lineage maintenance, differentiation, and proliferation capacity. Benetti *et. al.* (Benetti et al. 2008) and Sinkkonen *et. al.* (Sinkkonen et al. 2008) both showed that mESC specific miR-290-295 cluster can target Rbl2, which control Dnmt expression, suggesting that the miRNA cluster plays a role in regulating *de novo* DNA methylation. miR-302 has been shown to control germ layer specification by inhibiting lefty, an inhibitor of the Nodal pathway

(Rosa et al. 2009). In addition, Melton *et. al.* showed that miR-290-295 cluster can indirectly increase the expression of Lin28 and c-Myc to maintain mESCs in their self-renewing state (Lee et al. 2008). Lastly, miR-290-295, miR-302, and miR-372, have been found to accelerate cell proliferation by promoting G1 to S phase transition through targets such as p21 and Lats2 (Voorhoeve et al. 2006; Card et al. 2008; Wang et al. 2008).

Overlapping roles for the miR-290-295 cluster and the miR-302 cluster are not surprising, given their common seed AAGUGC (Table 1). However, other related miRNAs include rodent specific miR-467a, c, and d, which have been cloned from T cells (Neilson et al. 2007) as well as ES cells (Mineno et al. 2006; Landgraf et al. 2007), remain largely uncharacterized. We will refer to these families collectively as the “AAGUGC miRNAs” to emphasize the commonality of their seed sequence. Members of the miR-17-92 cluster also share this hexamer, though it corresponds to positions 3-8 rather than 2-7 of the mature sequence. Because this shift is expected to significantly reduce the activities of the miR-17-92 family in regulation of the AAGUGC seed targets (Figure 1), we will not include the family in our study.

Several large-scale sequencing datasets (Mineno et al. 2006; Calabrese et al. 2007; Landgraf et al. 2007; Babiarz et al. 2008; Leung et al. 2010) have revealed that AAGUGC miRNAs constitute the most highly expressed miRNAs in ES cells (Table 1), underscoring their important regulatory roles in this cell type. Although a subset of targets of AAGUGC miRNAs has been identified, we hypothesized that an unbiased approach might uncover novel roles of these miRNAs in ES cells. To this end, we attempted to identify endogenous targets by combining existing target prediction data with microarrays of mESCs before and after miRNA loss (Calabrese et al. 2007), as well as before and after specific deletion of the miR-290-295 cluster (Jaenisch 2008). Initial analysis using this data suggested strong enrichment of targets involved in apoptosis, a

phenotype has to date not been linked to specific miRNAs in ES cells. Through gain and loss of function studies, we show that AAGUGC miRNAs can protect ES cells from apoptosis, especially during exposure to genotoxic stress. We validated 2 candidate targets, Casp2 and Ei24, and propose that the anti-apoptotic property of these miRNAs is mediated in part by the repression of these target genes.

Results

Predicted targets of AAGUGC containing miRNAs are enriched in pathways regulating apoptosis.

To better understand the roles of AAGUGC miRNAs, we attempted to identify their endogenous targets by combining existing target prediction data with microarrays of mESCs before and after miRNA loss. Using a previously characterized floxed Dicer mESC line (Harfe et al. 2005; Calabrese and Sharp 2006; Leung et al. 2006), we compared wild type samples to those 5 days after Dicer deletion, having confirmed that a majority of miRNAs was lost by this time. In both cases, we calculated a cumulative density function (cdf) plot comparing expression differences for the set of all potential 7mer or 8mer targets (ie, transcripts containing at least an A1-7 or m2-8 match) for miRNAs with the seed AAGUGC, shown in Table 1. Relative to a control set of genes (“control”) that lacked the AAGUGC binding sites, but were matched for 3’ UTR length, dinucleotide composition, and expression level, the AAGUGC target set (“target”) was more derepressed upon Dicer loss (Figure 2A). Conserved AAGUGC target genes (“conserved_target”) as predicted by TargetsScan 5.1 (Friedman et al. 2009) showed an even larger derepression, suggesting further enrichment of genuine targets in this set (Figure 2A).

We next performed Gene Ontology Analysis on this candidate set. Of all ES-expressed genes (defined as WT expression \geq 16 on the microarray), we looked for GO

category enrichment in the genes that increase on Dicer loss (defined as an 1.2 fold upregulation in Dicer KO cells). From this analysis, top statistically significant categories include regulation of apoptosis, regulation of cell cycle, and regulation of transcription. We further refined our candidate list using array data of a specific miR-290-295 cluster deletion (miR-290-295 KO) ES line, which also showed cdf plot signature changes for AAGUGC containing seeds (Jaenisch 2008). In all, 806 candidates were identified as Targetscan-predicted AAGUGC targets that showed an 1.2 fold upregulation in knockout populations from both datasets (Figure 2B).

To validate some of predicted targets, we picked several candidate genes based on the degree of upregulation in Dcr KO and miR-290-295 KO ES lines, as well as their functional annotations. Their 3' UTRs were cloned into luciferase constructs, and expression level ratios between Dicer WT and Dcr KO cells were evaluated (Figure 3A). All candidates tested displayed at least mild repression relative to a control construct lacking miRNA target sites, as did the previously identified miR-295 targets, *Lats2* and *p21* (Wang et al. 2008). Additional transfection studies confirmed that repression could be conferred specifically by either of two representative AAGUGC miRNAs, miR-295 or miR-467a, in a Dcr KO background (Figure 3B).

Caspase 2 and Ei24, key apoptotic mediators, are direct targets of both the miR-290 and miR-467 families.

We chose to further characterize the most downregulated target, Caspase 2, as it suggested a novel link between embryonic stem cell-specific microRNAs and cell survival. An initiator of apoptosis in response to genotoxic stress (Li and Yuan 2008), Caspase 2 has four AAGUGC binding sites in its 3'UTR, though much of the repression can be conferred by the first two sites alone (Figure 4B). RT-PCR demonstrated an approximately five-fold increase in Caspase 2 transcript levels in Dcr KO cells,

suggesting that miRNAs may significantly destabilize this transcript (Figure 5C).

Transfection of either miR-295 or miR-467a strongly repressed an intact Caspase 2 reporter in these cells, but not a reporter in which the four target sites were mutated (Figure 4A). In support of these reporter data, a decrease in endogenous Caspase 2 protein on transfection of these two miRNAs was also observed (Figure 5A).

Given the extent to which Caspase 2 is regulated by miR-295 and miR-467a, we decided to test whether AAGUGC miRNAs regulate other proapoptotic factors. Bim, a BH3 only Bcl-2 family protein previously identified as a target of the miR-17-92 cluster (Ventura et al. 2008), has potentially 3 AAGUGC target sites. Therefore, we additionally designed luciferase reporters for it as well as Ei24/PIG8, a direct p53 transcriptional target that binds Bcl-2 (Gu et al. 2000; Zhao et al. 2005a), and contains one 7mer miR-295 site. While Bim reporter was not repressed in WT mESCs, Ei24 reporter displayed a two-fold repression. More specifically, Ei24 appeared to be more responsive to miR-295 than miR-467a, as the transfection of miR-295 in Dcr KO cells led to greater repression of the Ei24 reporter (Figure 4A). The luciferase data were additionally supported by RT-PCR data showing decreased levels of Ei24 in WT relative to Dcr KO cells (RT-PCR was performed in place of Western Blot due to the lack of a good antibody against Ei24) (Figure 5C).

To determine the extent to which the miR 290-295 cluster alone contributes to repression of these targets, we additionally examined our reporter constructs in the miR-290-295 KO line relative to its wild-type counterpart. The repression observed on both Casp2 and Ei24 reporters was approximately half as seen in the Dcr WT and KO systems (Figure 4C). This suggests that the miR-467 and miR-302 families of miRNAs incompletely compensate for miR-290-295 cluster loss, despite having many overlapping seeds (Figure 4C). As before, exogenous miR-295 or miR-467a could repress protein levels for both targets in the miR-290-295 KO line (Figure 5B).

AAGUGC containing miRNAs buffer ES cells against genotoxic stress.

Since two of our validated targets are key players in cell death, we wondered if miRNAs with the seed AAGUGC are involved in regulating apoptosis of mESCs. We compared the apoptosis rate of WT and Dcr KO ES cells in a 24-hr period by staining them with antibodies against cleaved Caspase-3 (Casp3), and analyzing them by flow cytometry. Neither cell types exhibited substantial apoptosis, although Dcr KO ES cells showed higher apoptosis rate than WT ES cells. (Figure 6A). As embryonic stem cells are highly sensitive to DNA damage (Tichy and Stambrook 2008), we hypothesized that AAGUGC miRNAs may be more protective against apoptosis in the context of genotoxic stress. To test this, we examined the effect of overexpressing miR-295 and miR-467 in cells exposed to either doxorubicin treatment, or gamma-irradiation. Doxorubicin inhibits topoisomerase II, and produces double stranded breaks. Gamma irradiation induces DNA damage, and activates ATM and p53 in a manner similar to doxorubicin. Both sources of stress activate intrinsic apoptosis pathways, and result in the cleavage of Casp3. Both Dcr KO and WT ES cells showed minimal response immediately after 5Gy-gamma radiation or 100nM doxorubicin treatment (Figure 6A). However, there was a significant difference in their responses 24 hours after the treatment. While 5% of WT cells became apoptotic, more than 25% of Dcr KO population exhibited Casp3 activity (Figure 6A). Similar results were seen with using AnnexinV as a marker, a complementary assay for detecting early apoptosis (Figure 12).

In order to examine if the phenotype is specific to miRNAs with the AAGUGC seed, we transfected miR-295 and miR-467 into Dcr KO cells respectively, and quantified their apoptosis rate immediately, and then 24 hours after radiation. Overall, there were very small differences between Dcr KO cells transfected with miR-467 or miR-295 and a pool of control siRNAs immediately following radiation. However, there

was a stark difference in cells' response to apoptosis 24 hours after the treatment. Relative to control siRNAs, the overexpression of miR-467 or miR-295 drastically decreased the apoptosis response of Dcr KO cells to radiation treatment (Figure 6B). The reduction in apoptosis is AAGUGC seed specific, as siRNAs with changes in the seed region failed to rescue Dcr KO ES cells from apoptosis (Figure 6B).

Next we wanted to see if any of the targets we validated were responsible for AAGUGC miRNA specific apoptosis regulation. We applied siRNAs specific to each target, and quantified the apoptosis rate 24 hours after exposing cells to radiation. A higher level of apoptosis was observed with the transfection of Bim, Casp2, and Ei24 siRNAs immediately after treatment (Figure 6C). Although Bim is not targeted by the AAGUGC miRNA family, we thought to use Bim siRNA as a positive control, as Bim, a pro-apoptotic factor, is important in inducing cell death in mESCs (Su et al. 2009). We believed that this was primarily due to the toxicity elicited upon siRNAs transfection. In order to account for differences in transfection-specific toxicity, we decided to look at the difference in Casp3 activation between 0 and 24hr timepoints (although the general trends remain unchanged). After the transfection of Bim, Casp2, or Ei24 siRNAs, Dcr KO cells exhibited a decrease of 5 to 10% in Casp3 activation 24 hours after radiation, a level similar to miR-467a overexpression in Dcr KO cells (Figure 6D). In addition, the combination of the three siRNAs (BCEsiRNAs) reduced the difference in apoptosis rate even further to almost the same level seen in the WT ES cells (Figure 6C and D).

A similar set of apoptosis responses was observed when cells were treated with 100nM doxorubicin (Figure 7A). Transfection of miR-467a or siRNAs against the target genes led to the reduction of apoptosis in Dcr KO cells (Figure 7B). However, transfection of miR-295 was not as effective as miR-467a in protecting Dcr KO cells from doxorubicin-induced apoptosis as it was with gamma-radiation (Figure 7B). Notably, the 8th nucleotide of miR-295 seed is different from that of miR-467. It is possible that miR-

467 regulates a slightly different set of targets or has an increased affinity for a subset of common targets, and is therefore more effective in suppressing doxorubicin-induced apoptosis. To test this, we repeated the assays with miR-290-3p, a miRNA in the miR-290 cluster that has the same 7-mer seed as miR-467. Interestingly, miR-290-3p is equally as effective as miR-467 in buffering Dcr KO cells from doxorubicin-induced apoptosis, suggesting that the 8th nucleotide is critical in silencing targets that are in the doxorubicin-induced cell death pathways (Figure 7B). It is also important to note that doxorubicin treatment did not affect the expression of mature AAGUGC miRNA family in WT mESCs (Figure 8).

Deletion of miRNAs containing AAGUGC seed makes cells more susceptible to apoptosis upon DNA damage.

We exploited miR-290-295 cluster KO ES cells to assess if the loss of the majority of the miRNAs with the AAGUGC seed can make cells more susceptible to apoptosis upon exposure to DNA damaging agents. Northern blot analysis revealed that the deletion of the miR-290-295 cluster not only results in the loss of the cluster, but also reduces the expression of the miR-302 and miR-467 clusters. Interestingly, the expression of let-7, a marker of ES cell differentiation, also increased in miR-290-295 KO cells, suggesting that some of the miR-290-295 cluster KO ES cells may be undergoing differentiation (Figure 9). We irradiated miR-290-295 KO and WT ES cells, and measured the cleaved caspase-3 activity 0 and 24 hours after the treatment. As expected, miR-290-295 KO cells are much more sensitive to radiation than their WT counterpart (Figure 10A). Overexpression of either of the two of the miRNAs in the cluster, miR-290-3p, and miR-295, as well as miR-467, significantly reduced the rate of apoptosis (Figure 10B). In addition, knocking down each of the 2 identified apoptosis targets, Casp2, and Ei24, can partially rescue cells from apoptosis caused by radiation

(Figure 10B). A similar set of observations was made with doxorubicin treatment, with the exception of miR-295, which did not have a protective effect (consistent with our observation in Dcr KO ES cells) (Figure 11).

Discussion

We provide the first demonstration of a group of miRNAs in ES cells controlling the induction of cell death. We have shown that miRNAs with the hexamer seed AAGUGC can suppress apoptosis of mESCs through target genes that are key players in cell death.

We propose that AAGUGC miRNAs affect their anti-apoptotic response by down-regulating the expression of Caspase 2 and Ei24, proapoptotic factors that are direct targets of the miRNAs. Although the exact mechanisms are still emerging, Caspase 2 and Ei24 are involved in pathways that converge at the mitochondria, and ultimately lead to the release of cytochrome c, formation of apoptosome, and activation of effector caspases (Jin and El-Deiry 2005). Ei24 resides in the ER, and can bind to Bcl-2 to initiate ER-stress induced apoptosis (Gu et al. 2000). Caspase 2, one of the most conserved caspases, has also been linked to DNA damage and ER stress response (Krumshabel et al. 2009). Mouse oocytes that lack Caspase 2 were found to be resistant to cell death following exposure to chemotherapeutic drugs (Bergeron et al. 1998). This phenotype is reminiscent of the genotoxic-stress induced apoptosis, response of WT mESCs, and KO mESCs when transfected with Caspase 2 siRNAs.

Our results showed that while downregulation of one of the three targets can partially reduce the rate of stress-induced apoptosis of KO mESCs, a simultaneous reduction of all the targets was able to suppress apoptosis to the same extent seen with overexpression of AAGUGC seed containing miRNAs in Dcr KO mESCs. This suggests that multiple apoptosis pathways are activated in mESCs upon exposure to genotoxic

stress, and that this family of miRNAs controls most if not all of key players. The intricate network of apoptosis pathways not only underscores the importance of apoptosis regulation in mESCs, but also implies that there could be more miRNA target genes involved in suppressing apoptosis of mESCs.

miRNAs have the ability to confer robustness upon biological systems (Hornstein and Shomron 2006). Because miRNAs can fine-tune gene expression post-transcriptionally, they can correct leaky transcription of target genes and promote rapid developmental transitions. Often, these roles are related to their position within feed-forward networks, where they stabilize gene expression (Hornstein and Shomron 2006; Tsang et al. 2007). In particular, this stabilization can help buffer cells from environmental stress. For instance, a direct experimental link between miRNAs and robustness was first provided by Li and colleagues (Li et al. 2009), who demonstrated that miR-7 was necessary for the proper development of *Drosophila* sensory organs specifically under conditions of temperature fluctuations.

Our results provide another example of how miRNAs can protect ES cells from genotoxic stress. It is intriguing that AAGUGC miRNAs suppress apoptosis of ES cells following exposure to DNA damaging agents. Intuitively, we would expect ES cells to readily induce apoptosis in the event of genotoxic stress, as any change in their DNA can be quickly amplified in other cell types. When the genome of a mESC is damaged, it has several options: DNA repair, apoptosis, and differentiation (Cervantes et al. 2002). ES cells are quick in eliciting DNA repair response, especially homologous recombination-mediated repair and mismatch repair (Tichy and Stambrook 2008). They also readily undergo apoptosis to eliminate cells with damaged DNAs (Cervantes et al. 2002). In addition, some studies suggest that DNA damage can cause downregulation of Nanog, and induce differentiation of ES cells into other cell types that are no longer pluripotent (Lin et al. 2005; Fujita et al. 2008). Thus, it is important to control the

induction of cell death in the reproductive capacity of an organism. Interestingly, the anti-apoptotic role of AAGUGC miRNAs may not be limited to ESCs, as the miR-467 family is also upregulated in the transition from Double Positive to Single Positive thymocytes, a time when apoptosis must be tightly calibrated for cells to escape negative selection (Neilson et al. 2007).

In addition to identifying this novel anti-apoptotic phenotype, we were able to confirm the results of other groups linking AAGUGC miRNAs to more rapid cell cycle progression (Wang et al. 2008; Wang and Blelloch 2009), with miR-467 overexpression in Dcr KO cells accelerating the G1/S transition similarly to previous reports for miR-295 (Figure 13). Thus, we can broadly identify these miRNAs as pro-proliferative, a feature that likely extends to humans as well given the high expression of the homologous clusters in human ESCs (Landgraf et al. 2007; Bar et al. 2008). Physiologically, the rapid rodent-specific expansion of the miR-467 family may further increase the effective dosage of this seed. This change may have important consequences for the reproductive differences between rodents and humans, which in part depend on cellular decisions between survival and death; tolerating higher rates of mutational stress may allow rodents to maintain greater fecundity. Caspase 2, which has already been implicated in the reproductive axis as Casp2 null mice harbor an excess of oocytes, is likely differentially regulated by miRNAs in these species as well. Of the four sites we identified in the Caspase 2 mouse 3'UTR, one was conserved in rat and none were conserved in human (TargetsScan 5.1).

Beyond these roles in normal physiology, the proliferative effects of AAGUGC miRNAs also appear to have important consequences for cancer, which often share the rapid growth pattern of ESCs. For instance, the miR-467 family and the miR-302 cluster have been cloned from neuroblastoma and teratocarcinoma cell lines (Landgraf et al. 2007) respectively. We hypothesize that these miRNAs may be serving similar roles to

promote rapid growth and survival in physiological as well as neoplastic contexts.

Methods

ES Cell Culture

Feeder-free *Dicer1^{ff}* and *Dicer1*^{-/-} mouse embryonic stem cells (mESCs) were generated and maintained on gelatin as described previously (Calabrese et al. 2007). mESCs cells containing a floxed and excised miR-290 cluster were generated in a similar manner (Jaenisch 2008).

Oligos and siRNAs used in all the experiments

siRNA	Sequence (5' → 3' unless otherwise noted)
miR-295	5'- AAAGUGCUACUACUUUUGAGUCU -3' 3'- UCUUUCACGAUGAUGAAAACUCA -3'
miR-295 seed mutant	5'- AAAGACGUACUACUUUUGAGUCU -3' 3'- UCUUUCUGCAUGAUGAAAACUCA -3'
miR-467a	5'- UAAGUGCCUGCAUGUAUAUGCG -3' 3'- GCAUUCACGGACGUACAUAUAC -5'
miR-467a seed mutant	5'- UAAGACGCUGCAUGUAUAUGCG -3' 3'- GCAUUCUGCGACGUACAUAUAC -5'
miR-290-3p	5'- AAAGUGCCGCCUAGUUUUAAGCCC -3' 3'- CCUUUCACGGCGGAUCAAAAUUCG -5'
Control siRNA	(from Dharmacon, Accell Non-targeting pool)
si-p21	(from Dharmacon, Smartpool)
si-bim	(from Dharmacon, Smartpool)
si-casp2	(from Dharmacon, Smartpool)
si-ei24	(from Dharmacon, Smartpool)
RT-PCR Primers	
Casp2 Forward	GCAGGGTCACTTGGAAGACT
Casp2 Reverse	GAAGACAGGGAGGACCATCA
Ei24 Forward	TCTCTTCCCCATCCATCTT
Ei24 Reverse	TAACGTAACGACACTCCTTTC
β-actin Forward	GACGAGGCCCCAGAGCAAGAGAGG
β-actin Reverse	GGTGTTGAAGGTCTCAAACATG
3'UTR Primers	
Isgf3g Forward	AATAACTCGAGCGCGTCTCCATGGAAATAGA
Isgf3g Reverse	AATAAGGGCCCTTTAATTTGGAGCTCACATTTCT
Casp2 Forward	AATAACTCGAGCCGCCTGCTATTCTGCT
Casp2 Reverse	AATAAGGGCCCTCAACATTTATTTGGCACCTG

Pax6 Forward	AATAACTCGAGAGAGAGAAGGAGAGAGCATGTG
Pax6 Reverse	AATAAGGGCCCAAATCATTCTGAGGATTCTAGGG
p21 Forward	AATAACTCGAGCCTCTTCTGCTGTGGGTCA
p21 Reverse	AATAAGCGGCCGCAATCATCGAGAAGTATTTATTGAGC
Thbs1 Forward	AATAACTCGAGTCATCAGCTGCCAATCATAA
Thbs1 Reverse	AATAAGGGCCCTTCCATATGATTTATTGTTGTTCCCTT
Itgav Forward	AATAACTCGAGCCACTTCTGTCCGCTCCA
Itgav Reverse	AATAAGGGCCCGAAGTCAACTGTAGTGTAATGTGTACC
Irak3 Forward	AATAACTCGAGATCCACCAGAAGATCAAGCAA
Irak3 Reverse	AATAAGCGGCCGCTTTTATATAACAATTGGAATGCCACAG
Lats2 Forward	AAACTCGAGCGAGGAAACCCAAAATGAGA
Lats2 Reverse	AAAGGGCCCTCCAACAAAACACCACAAATG
Mutagenic Primers	
Ei24 Site 1	GACCAGAGTTTTCCAGCTGTTTTTTTACGTCTTGCCAGCTCCTGT
Bim Site 1	GACCAGAGTTTTCCAGCTGTTTTTTTACGTCTTGCCAGCTCCTGT
Bim Site 2	AGCTTCCATTATGCCGAGTAAACGTCTTGTCTTCCACAAGATGTCT
Bim Site 3	CACAGCCTGGTGGAGGACGTCTTTCTAACCTGTGGAG
Casp2 Site 1	CCTTACTGTGGCTTCTGCATCGTCTTACACTGTACTTGACGGC
Casp2 Site 2	GTACCATATGTGATATAACCTAGAACGTCTTGTCTCTGCTCTTATGAAACTTG
Casp2 Site 3	GTGCTTACTGCAGGCTGTAATGCGTCTTTTGCTTGTTCACTTGTTTC
Casp2 Site 4	CTTACTTACTGATATCCAGTAACTGCGTCTTACTAGGTCTTCATGAATGTTTC
Northern LNA Oligos	
Let-7g	AACTGTACAAACTACTACCTCA
miR-16	GCCAATATTTACGTGCTGCTA
miR-295	AGACTCAAAAGTAGTAGCACTTT
miR-302d	CACTCAAACATGGAAGCACTTA
miR-467a-5p	CGCATATACATGCAGGCACTTA
2x-reporter oligos	Subscript ₂ means the sequence in the bracket was present twice.
2x-bulged miR-20a	Forward: TCGAG(CTACCTGCACTAAAGCACTTTA) ₂ GGGCC Reverse: C(TAAAGTGCGGATAGTGCAGGTAG) ₂ C
2x-bulged miR467	Forward: TCGAG(CATATACATGCAGGCACTTA) ₂ GGGCC Reverse: C(TAAGTGCCTTATGGTATATG) ₂ C

Generation of luciferase constructs, mESC transfection, and luciferase assays

MicroRNA mediated repression of each candidate gene was tested by cloning PCR amplified products corresponding to the entire 3'UTR (or in the case of Bim, 2kb containing the 3 AAGUGC hexamer binding sites) into the 3' UTR of a pRL-CMV Renilla luciferase reporter as described previously (Doench and Sharp 2004). Nucleotides 5-7

of Casp2, Bim, and Ei24 binding sites were mutated by Quickchange site-directed mutagenesis.

Digests were performed using either XhoI or Sall to give the 5' site and ApaI or NotI to give the 3' site. Firefly luciferase (pGL3) was used as a transfection control. Data shown are summaries of three or more independent trials.

24 hours before transfection $1e^5$ mESC cells were plated/well of gelatinized 24-well plate. Cells were transfected with 2 μ l Lipofectamine 2000 (Invitrogen), 0.1 μ g of CMV-GFP plasmid (Invitrogen), 0.7 μ g of pWS (carrier plasmid), and 50nM of siRNAs in 300 μ l of Opti-MEM (Invitrogen). 4 hours after transfection, transfection mix was removed from cells and replaced with ESC media.

24 hours after transfection, cells were lysed with 1X Passive Lysis Buffer (Promega) and Dual luciferase was measured using Dual Luciferase reporter assay system (Promega) according to manufacturer's instructions.

Northern Blot analysis

Total RNA was isolated from *Dcr^{-/-}* and *Dcr^{+/+}* ES cells using Trizol (Invitrogen), following the standard protocol. Approximately 50 μ g of each RNA was loaded onto a 15% denaturing MOPS gel, according to the Northern Blot protocol outlined previously (Seila et al. 2008). Membrane probed with Gln-tRNA was exposed to phosphorimager for 6 hours before being scanned. The rest was exposed to phosphorimager for 1 day before being scanned. Prior to hybridizing with a different probe, membranes were stripped by

incubating the membrane in boiling 0.1% SDS for 30 minutes and loss of signal was confirmed prior to rehybridization.

Western Blot analysis

24 hours after transfection with short RNAs, *Dicer1*^{-/-}, *Dicer1*^{fllox/fllox}, *miR290-295*^{-/-}, or *miR290-295*^{fllox/fllox} cells were lysed in RIPA buffer (1% NP40, 0.5% sodium deoxycholate, 0.1% SDS, in pH 7.4 PBS) containing protease inhibitors. 30–50 µg lysate was loaded onto 8-12% Bis-Tris gels (Invitrogen) and wet-transferred at 4°C to Westran PVDF membranes for 2h at 70V. After 1 hr blocking at room temperature in 5% milk-TBST, membranes were probed overnight at 4°C with 1:2000 mouse anti-vinculin (Santa Cruz Biotechnology), 1:200 rat anti-Caspase 2 (Millipore, 10C6), 1:200 goat anti-Ei24 (Santa Cruz, H-20), or 1:200 rabbit anti-Bim (Assay Designs, AAP-330). After 2x 10 min. TBST washes, membranes were probed for 1hr at room temperature with 1:2000 corresponding HRP-conjugated secondary, washed an additional 2x 10 min. in TBST, and visualized using Wester Lightning Plus ECL (PerkinElmer).

RT-PCR

Trizol (Qiagen) was used to extract RNA from *Dicer1*^{fllox/fllox} and *Dicer1*^{-/-} cells. A Superscript III kit (Invitrogen) was used to reverse transcribe 1 µg RNA following DNase treatment with the Turbo-DNA free kit (Ambion), and real time PCR was performed with the primer sequences listed, using beta actin for normalization.

Transfection and BrdU assays

24 hours before transfection 1e⁵ mESC cells were plated/well of gelatinized 24-well plate. Cells were transfected with 2µl Lipofectamine 2000 (Invitrogen), 0.1µg of CMV-

GFP plasmid (Invitrogen), 0.7 μ g of pWS (carrier plasmid), and 50nM of siRNAs in 300 μ l of Opti-MEM (Invitrogen). 4 hours after transfection, transfection mix was removed from cells and replaced with ESC media.

24 hours after transfection, cells were pulsed labeled with BrdU for 10 min. APC BrdU Flow Kit (BD Biosciences) was used to analyze cell cycle profile. Only GFP positive cells were used in our data analysis.

Transfection and Casp3 assays

24 hours before transfection $2e^5$ mESC cells were plated/well of gelatinized 12-well plates. Cells were transfected with 4 μ l Lipofectamine 2000 (Invitrogen), 0.2 μ g pCAGGS-mCherry plasmid, 1.4 μ g of pWS, and 50nM of siRNA in 600 μ l of Opti-MEM (Invitrogen). 4 hours after transfection, transfection mix was removed from cells and replaced with ESC media.

24 hours after transfection, cells were exposed to 5-Gy gamma radiation or 100nM doxorubicin. Immediately after exposure, one plate of cells were trypsinized and fixed with 1x BD Perm buffer. Cells were stained with Rabbit Anti-Casp3 antibody (BD Biosciences) at 1:100 for 20 min at room temperature. Following washing, cells were incubated with Alexa-488-conjugated secondary antibody (diluted 1:250) (Invitrogen) for 60 min at room temperature, washed, and resuspended in BD FACS buffer containing 1:5000 Hoechst stain. 24 hours after the treatment, another plate of cells was trypsinized and treated with the same protocol for FACS analysis.

Casp3 assays were also performed on Dcr KO and WT mESCS without transfection. 24 hours before collecting cells for 0hr time point for Casp3 assay, 2×10^5 mESC were plated/well of gelatinized 6-well plates. In the context of genotoxic stress, 4×10^5 mESCs were plated/well of gelatinized 6-well plates. 24 hours after plating, cells were treated with 5-Gy radiation or 100nM doxorubicin. Casp3 assays were performed at 0hr and 24hr after the treatment following the same protocol described above.

AnnexinV assays

4×10^5 mESCs were plated/well of gelatinized 6-well plates. 24 hours after plating, cells were exposed to 100nM doxorubicin. Cells were trypsinized 0hr and 24hr after the treatment for Annexin V detection, following Annexin V-FITC apoptosis detection kit (BD Biosciences).

Microarray analysis

Microarray analysis was performed 5 days following transfection of *Dicer1*^{flox/flox} wild-type cells with either GFP alone or GFP and cre recombinase, and data was analyzed using biological triplicates. Microarrays for the miR290-295 cluster deletion were performed on two deletion and two wild-type lines independently derived. Spot replicates were condensed using geometric means.

The log fold change (LFC) value for WT/Dcr_KO was defined as the difference between the mean log expression in WT cells and the mean log expression in *Dcr*^{-/-} cells. The conserved set of targets were downloaded from TargetScanMouse5.1 website (http://www.targetscan.org/mmu_50/). To identify targets predicted for the AAGUGC seed family, we looked at all miRNAs that contain AAGUGC in their seed region. More

specifically, they include “miR-291b-3p/519a/519b-3p/519c-3p”, “miR-290-3p/292-3p/467a”, “miR-467cd”, “miR-106/302”, and “miR-467b”. We excluded all the targets of “miR-302ac/520f”, as well as T1A 7mer targets of “miR-467b”, as they do not contain the 6mer match to AAGUGC. Targets with top 25% of branch length scores were considered “conserved”.

Gene Ontology analysis

Gene Set Analysis Toolkit (<http://bioinfo.vanderbilt.edu/webgestalt/>) was used to perform GO analysis. Targets and controls were generated as described in the text.

Statistical analyses

All test statistics were calculated using R (<http://www.r-project.org>). The Wilcoxon rank sum test was used because it does not assume normality of the underlying distributions. T-tests and Kolmogorov–Smirnov (KS) test using these data gave generally similar results.

Table 1. Description of miRNAs in the AAGUGC family.

Sequences and cloning statistics of miRNAs in the AAGUGC family are listed. The hexamer seed is highlighted in red. ES cell lines from Leung *et. al.* were used for all the experiments described in the chapter.

Cluster	Chr	miRNA	Sequence	% cloned in different ES cell lines		
				Leung <i>et. al.</i> (a)	Babiarz <i>et. al.</i> (b)	Ciaudo <i>et. al.</i> (c)
miR-290-295 cluster	chr7	miR-290-3p	AAAGTGCCGCCTAGTTTTAAGCC	0.01	6.31	0.03
		miR-291a-3p	AAAGTGCTTCCACTTTGTGTG	1.28	8.36	16.71
		miR-291b-3p	AAAGTGCAATCCATTTGTTTG	0.03	0.88	0.00
		miR-292-3p	AAAGTGCCGCCAGGTTTTGAGTG	2.66	12.47	10.10
		miR-294	AAAGTGCTTCCCTTTGTGTG	13.56	10.82	12.72
		miR-295	AAAGTGCTACTACTTTTGAGTC	25.01	13.70	8.89
miR-302 cluster	chr3	miR-302a	TAAGTGCTTCCATGTTTGGTG	1.55	0.14	0.04
		miR-302b	TAAGTGCTTCCATGTTTAGTA	1.27	0.16	0.00
		miR-302d	TAAGTGCTTCCATGTTTGAGTG	1.63	0.32	0.02
miR-467 cluster	chr2	miR-467a	TAAGTGCTGCATGTATATGC	0.00	0.08	0.33
		miR-467c	TAAGTGCGTGCATGTATATGT	0.68	0.00	0.03
		miR-467d	TAAGTGCGCGCATGTATATGC	0.21	0.00	0.36
AAGUGC miRNA family / total miRNA reads				47.89	53.23	49.22

(a) A. Leung, A. Young, AJ Bhukar, G. Zheng, A. Bosen, and P.A. Sharp. 2010 *Submitted*.

(b) J.E. Barbiarz, J.G. Ruby, Y. Wang, D.P. Bartel, and R. Blelloch. *Genes Dev*, 2008. 22 (20).

(c) C. Ciaudo *et. al.*, *Plos Genetics* 5, 2009, e1000620.

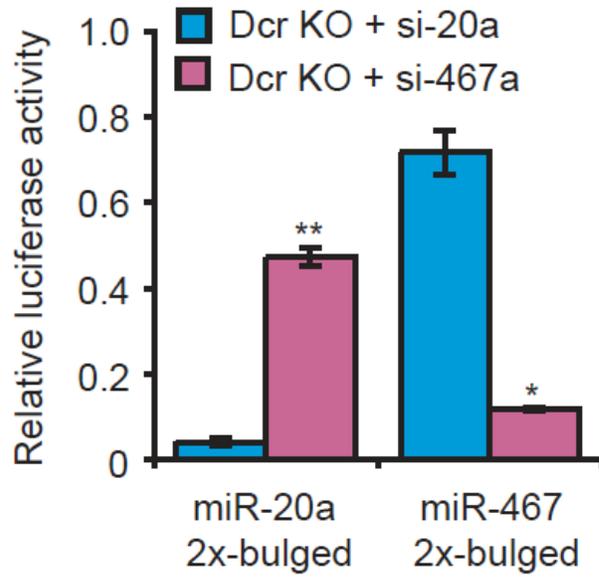
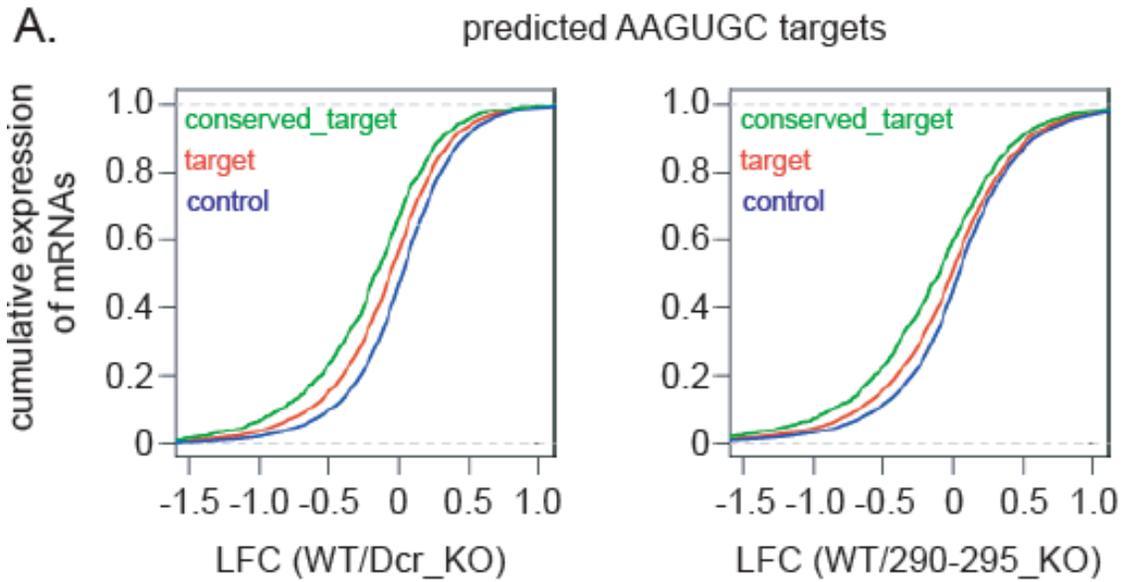


Figure 1. Repressive effect of miR-467a-5p and miR-20a on their control reporter constructs.

Luciferase assays of miR-20a and miR-467a-5p reporters with two bulged binding sites in Dcr KO ES cells. Luciferase activity of the reporter was normalized to a reporter with no binding site to miRNAs. 20nM of si-20a and si-467a-5p were transfected 24 hours before the reading of luciferase activity. n=2, and results are shown as mean \pm S.E.M. P-values were results of t-tests, * denotes $p \leq 0.05$, and ** denotes $p \leq 0.01$.



B.

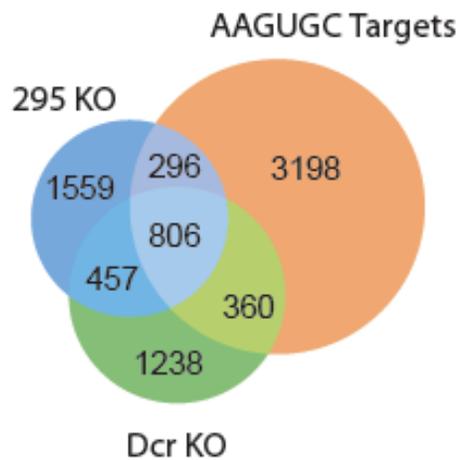


Figure 2. mRNA targets of AAGUGC miRNAs show decreased stability in Dcr KO cells and their luciferase targets can be repressed.

(A) CDFs (cumulative distribution functions) of \log_2 fold change (LFC) in mRNA expression between the wild type (WT) and Dcr KO ES cells (the left panel), and the wild type and 290-295 KO ES cells (the right panel), are plotted. Plots include *conserved_target* (green line), *all*

predicted AAGUGC target (red), and *Control mRNAs* (blue). All predicted AAGUGC targets include ~3000 predicted TargetScan targets that contain a 7-mer or 8-mer match to AAGUGC in their 3' UTRs. *Conserved target* set contains ~500 top 10% of all predicted AAGUGC targets that were ranked by branch length scores. The *Control mRNA* set was selected to match the predicted targets in seed match type and count. Targets are depressed in both Dcr KO as well as 295 KO mESCs ($p < 2.2e^{-16}$ and $p < 2.2e^{-16}$ by rank sum test respectively). (B) Venn diagram of microarray and target prediction data used to generate AAGUGC seed candidates. AAGUGC = genes with at least an A1-7 or M2-8 match to the AAGUGC seed; 295 KO = genes showed a 1.2 fold upregulation on miR290-295 cluster loss. Dcr KO = genes showed a 1.2 fold upregulation on Dicer loss. Only ES-expressed genes (i.e., genes with an expression of at least 16 in the wild-type arrays) were considered for analysis.

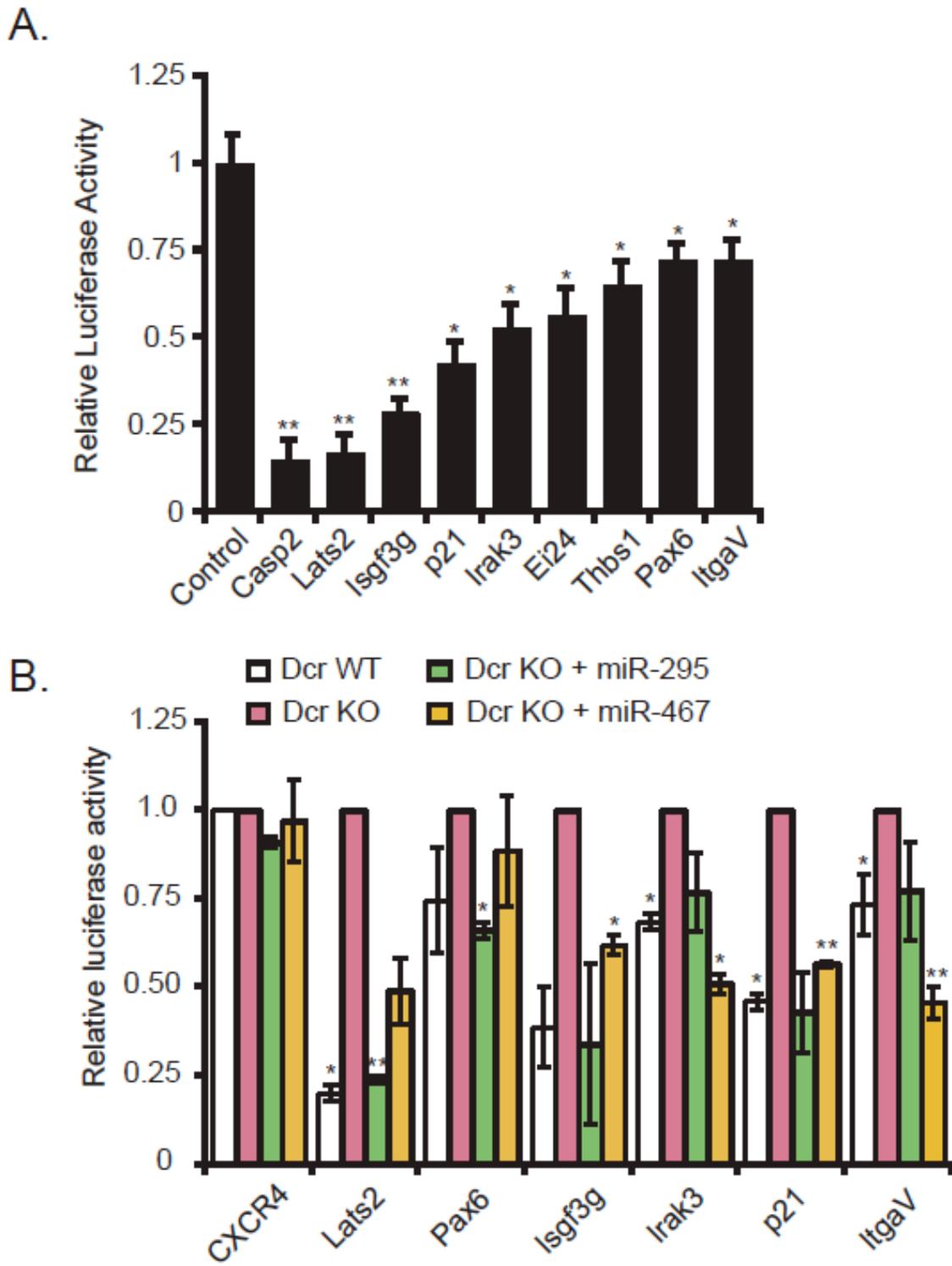


Figure 3. Repression of predicted AAGUGC miRNA targets in WT and Dcr KO ES cells.

(A) Activity of luciferase reporters of predicted AAGUGC miRNA targets were assayed in WT and Dcr KO ES cells. Luciferase reporters contain full length 3' UTRs of predicted targets. Relative luciferase activity is the ratio of the reporter's activity in WT ES cells and Dcr KO ES cells. (B) Activity of luciferase reporters of predicted AAGUGC miRNA targets were assayed in WT, Dcr KO ES cells, as well as in Dcr KO ES cells after over expression of 20nM of miR-295 or 20nM of miR-467. $n \geq 3$, and results are shown as mean \pm S.E.M. P-values were results of t-tests, * denotes $p \leq 0.05$, and ** denotes $p \leq 0.01$.

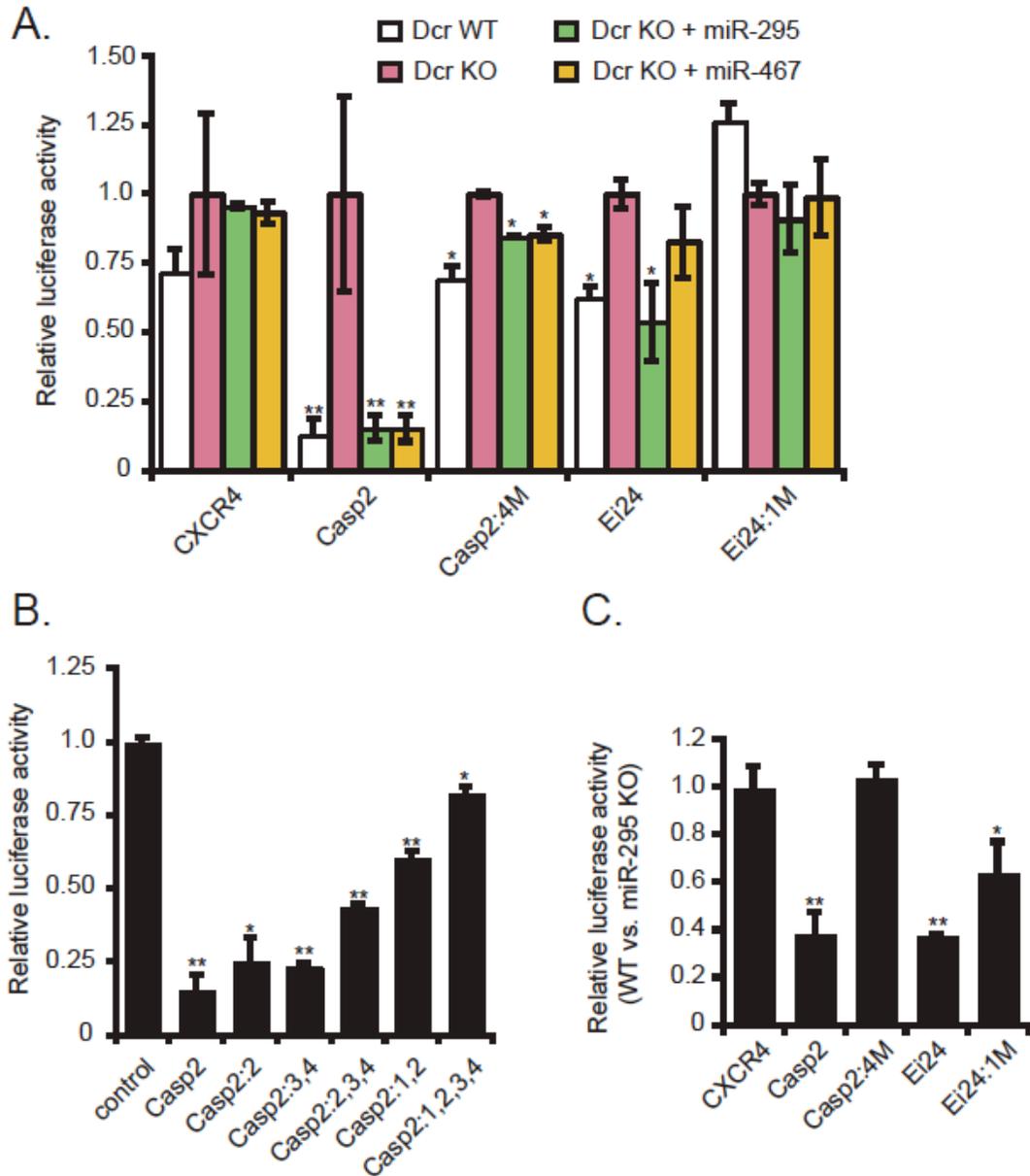


Figure 4. Repression of Casp2 and Ei24 by AAGUGC miRNAs.

(A) Luciferase reporters with full length Casp2 3' UTR, Ei24 3' UTR, as well as their seed mutant versions were assayed in WT and Dcr KO ES cells. Casp2:4M has all 4 AAGUGC seed binding sites mutated, and Ei24:1M has 1 AAGUGC seed binding site mutated. 20nM of miR-295 and miR-467a-5p were transfected in Dcr KO ES cells to test if the repression of luciferase reporters is specifically due to AAGUGC miRNAs. (B) Casp2 luciferase reporters bearing different combinations of AAGUGC seed binding sites mutations were tested in WT and Dcr KO ES cells. Casp2:2, 2nd AAGUGC binding site was mutated; Casp2:3,4, 3rd and 4th binding sites were mutated; Casp2: 2,3,4, 2nd, 3rd, and 4th binding sites were mutated; Casp2:1,2, 1st and 2nd binding sites were mutated; Casp2: 1,2,3,4, all binding sites were mutated. (C) Luciferase reporters with full length Casp2 3' UTR, Ei24 3' UTR, as well as their seed mutant versions were assayed in WT

and miR-290-295 KO ES cells. $n \geq 3$, and results are shown as mean \pm S.E.M. P-values were results of t-tests, * denotes $p \leq 0.05$, and ** denotes $p \leq 0.01$.

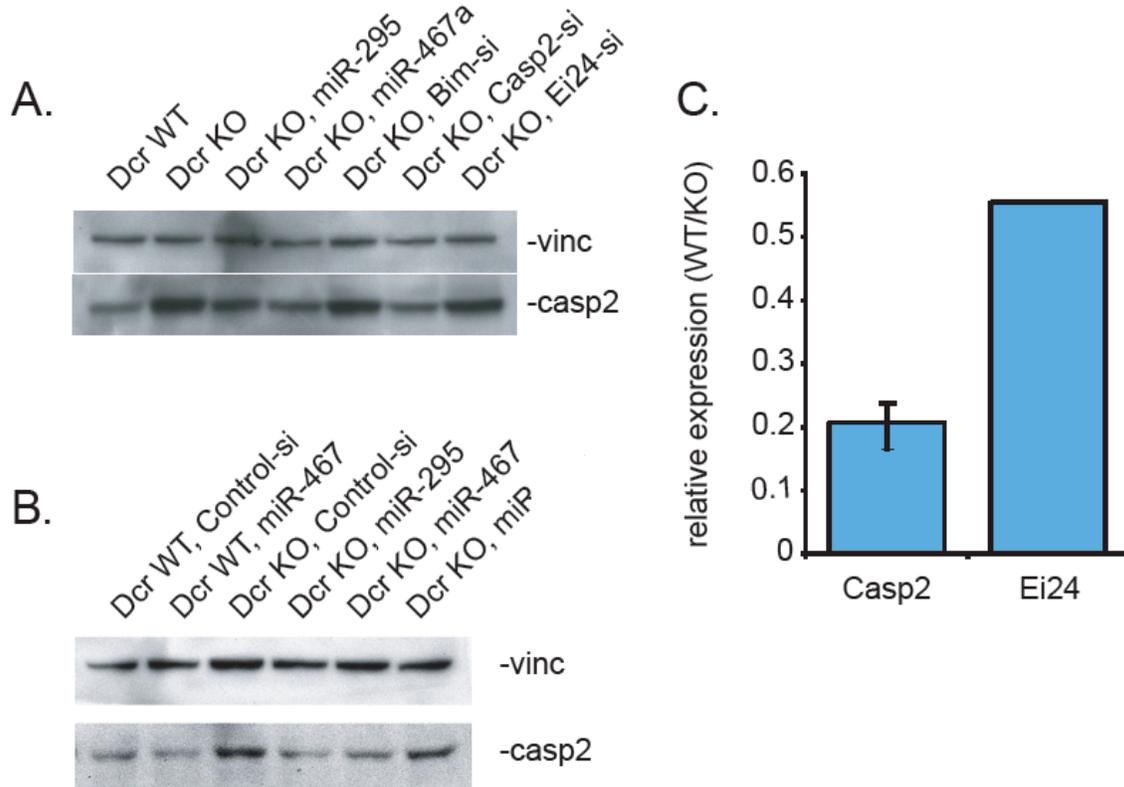


Figure 5. Repression of Casp2 and Ei24 in WT ES cells, as well as miR-290-295 KO ES cells.

(A) Western blot of Casp2 in WT and Dcr KO ES cells. 50nM of miR-295, miR-467a-5p, Bim siRNA, Casp2 siRNA as well as Ei24 siRNA were transfected into Dcr KO ES cells, and Casp2 protein expression was assayed 24 hours after the transfection. (B) Western blot of Casp2 in WT and miR-290-295 KO ES cells. (C) RT-PCR of Casp2 and Ei24 in WT and Dcr KO ES cells. n=3 for Casp2, and n=1 for Ei24.

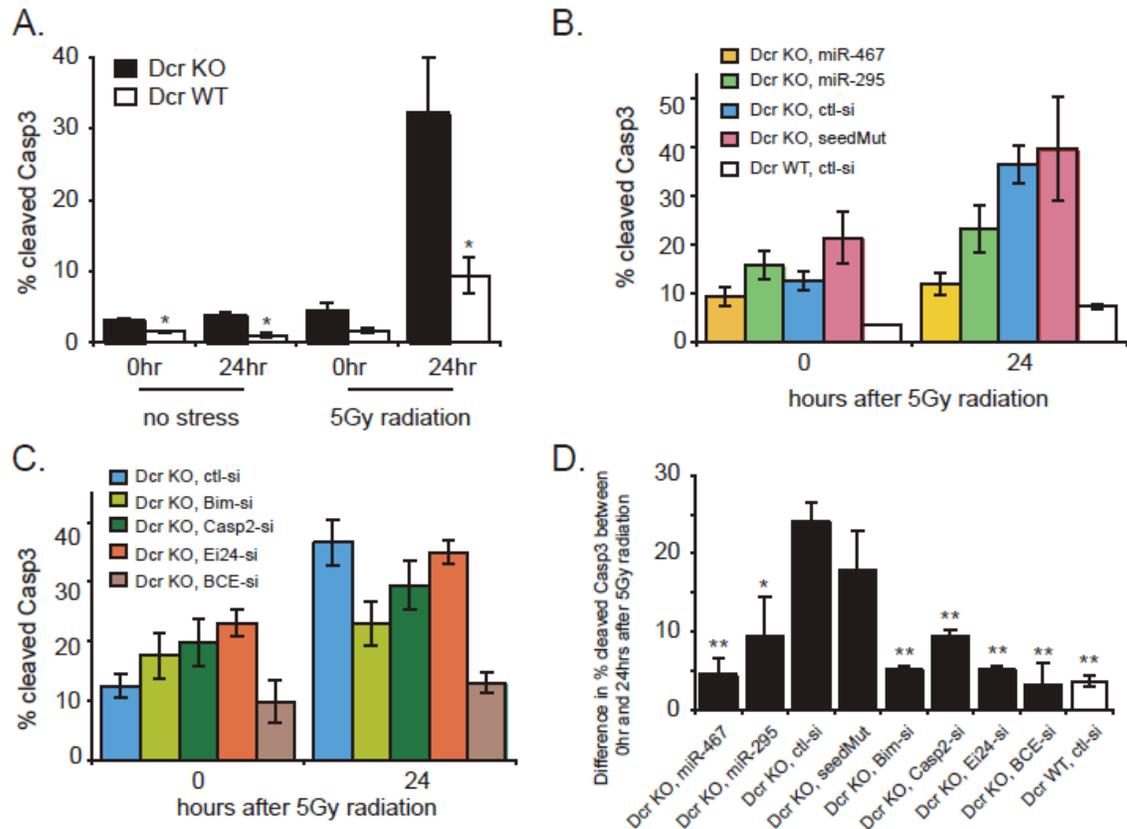


Figure 6. AAGUGC miRNAs protect ES cells from radiation-induced apoptosis.

(A) The percentage of cleaved Casp3 in WT and Dcr KO ES cells under normal culturing conditions (0 and 24 hours after plating) and after exposure to 5-Gy radiation (0 and 24 hours after radiation treatment). Cleaved Caspase-3 was assayed by flow cytometry, and was used to estimate apoptosis response. Apoptosis rate of Dcr KO cells was shown in black bars, and that of WT cells was shown in white bars. (B) Dcr KO cells were treated with 5-Gy radiation 24 hours after transfection of 50nM of miR-467a or miR-295. Caspase-3 activity was assayed 0 and 24hr after the treatment. Transfection of seed mutants and control siRNAs (50nM) into Dcr KO cells, and overexpression of control siRNAs (50nM) into WT cells served as controls. (C) Dcr KO cells were treated with 5-Gy radiation 24 hours after transfection of 50nM siRNAs against Bim, Casp2, and Ei24, or a combination of the three. Caspase-3 activity was assayed 0 and 24hr after the treatment. (D) Results of B and C were summarized here, where the difference in apoptosis rate between 0 and 24hr time points was shown. $n \geq 3$ for all experiments. Results are shown as mean \pm S.E.M. (standard error of the mean). P-values were results of Mann-Whitney tests, * denotes $p \leq 0.05$, and ** denotes $p \leq 0.01$.

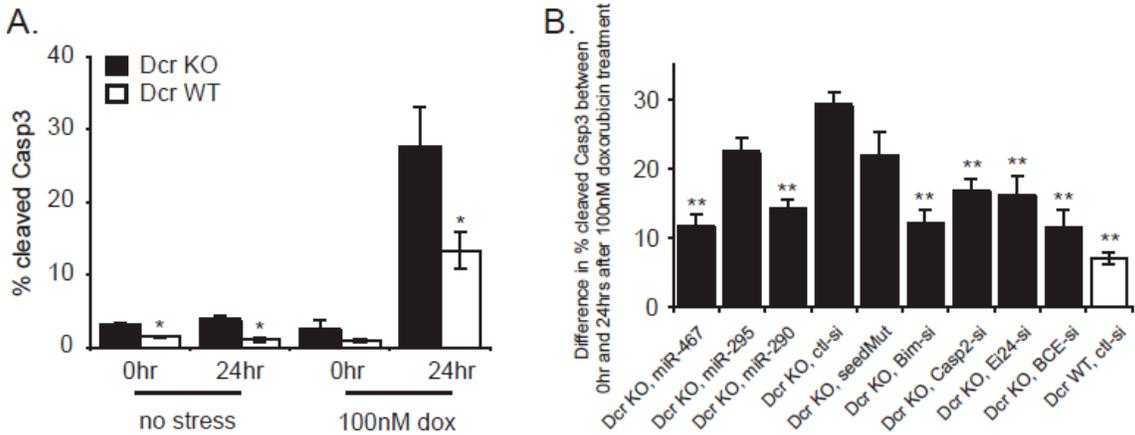


Figure 7. AAGUGC miRNAs protect ES cells from doxorubicin-induced apoptosis.

(A) The percentage of cleaved Casp3 in WT and Dcr KO ES cells under normal culturing conditions (0 and 24 hours after plating) and after exposure to 100nM doxorubicin (0 and 24 hours after radiation treatment). Cleaved Caspase-3 was assayed by flow cytometry, and was used to estimate apoptosis response. Apoptosis rate of Dcr KO cells was shown in black bars, and that of WT cells was shown in white bars. (B) Cells were transfected with 50nM siRNAs as shown, and difference in apoptosis response of WT and Dcr KO ES cells 24 hours after exposure to 100nM doxorubicin was plotted. $n \geq 3$ for all experiments. Results are shown as mean \pm S.E.M. (standard error of the mean). P-values were results of Mann-Whitney tests, * denotes $p \leq 0.05$, and ** denotes $p \leq 0.01$.

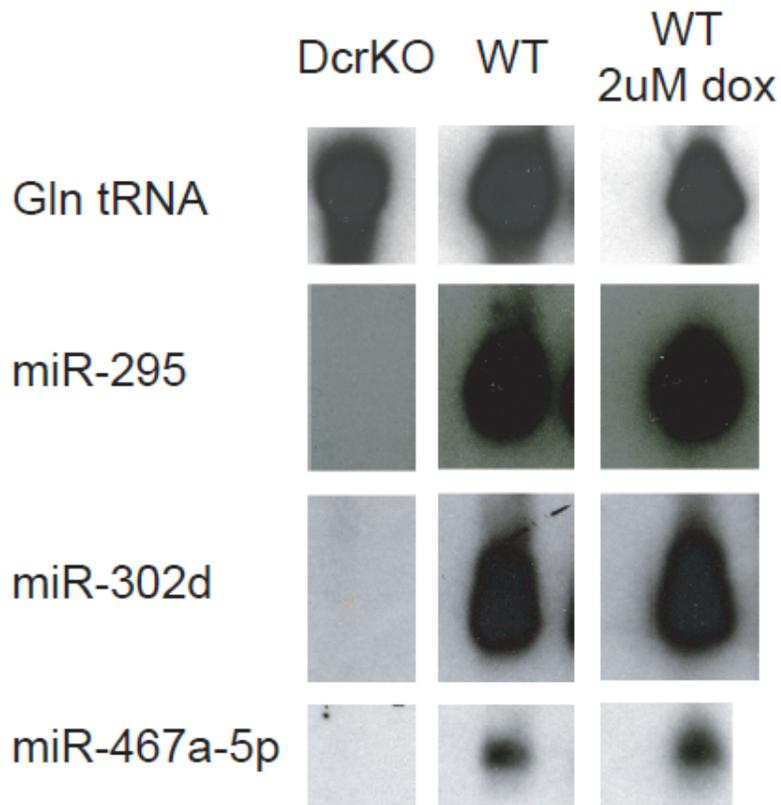


Figure 8. Expression of AAGUGC miRNA family stays the same before and after doxorubicin treatment.

Northern analysis for miR-295, miR-302d, and miR-467a-5p in Dcr KO ES cells, WT ES cells, and WT ES cells 6 hours after 2uM doxorubicin treatment.

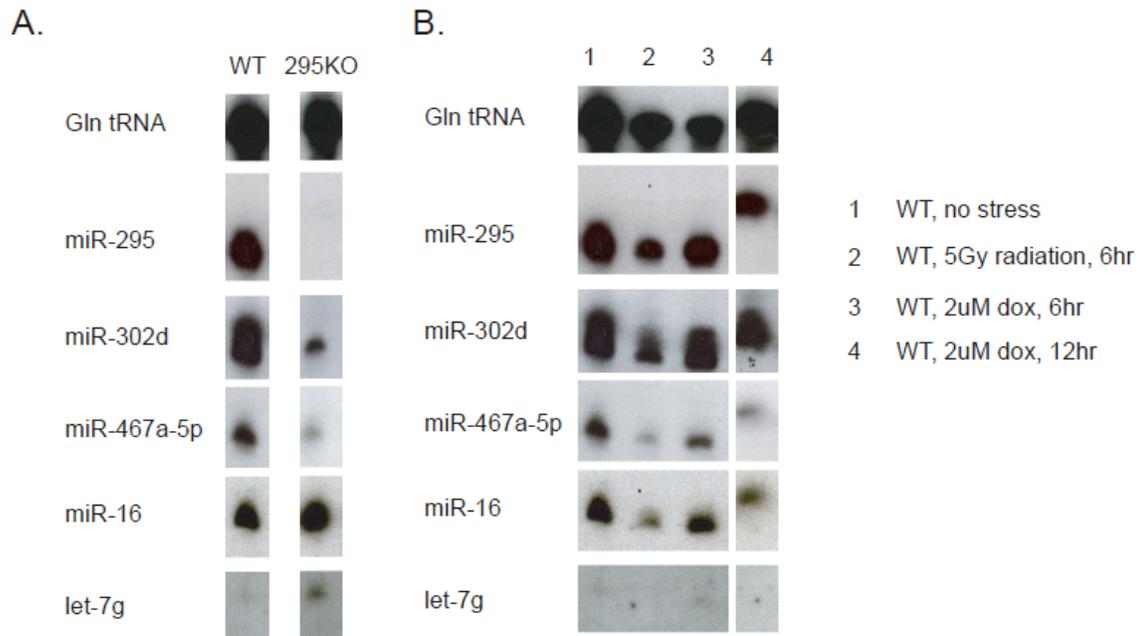


Figure 9. Expression of AAGUGC miRNAs in 295 KO ES cells before and after genotoxic stress.

(A) Northern analysis of miR-295, miR-302d, and miR-467a-5p in WT ES cells and miR-290 KO ES cells. Gln tRNA was probed as a loading control, whereas miR-16 and let-7g were probed as negative controls. (B) Northern analysis of miR-295, miR-302d, and miR467a-5p in miR-295 KO ES cells before and after stress. Lane 1, no stress; 2, 6 hours after 5-Gy radiation; 3, 6 hours after 2uM doxorubicin; 4, 12 hours after 2uM doxorubicin.

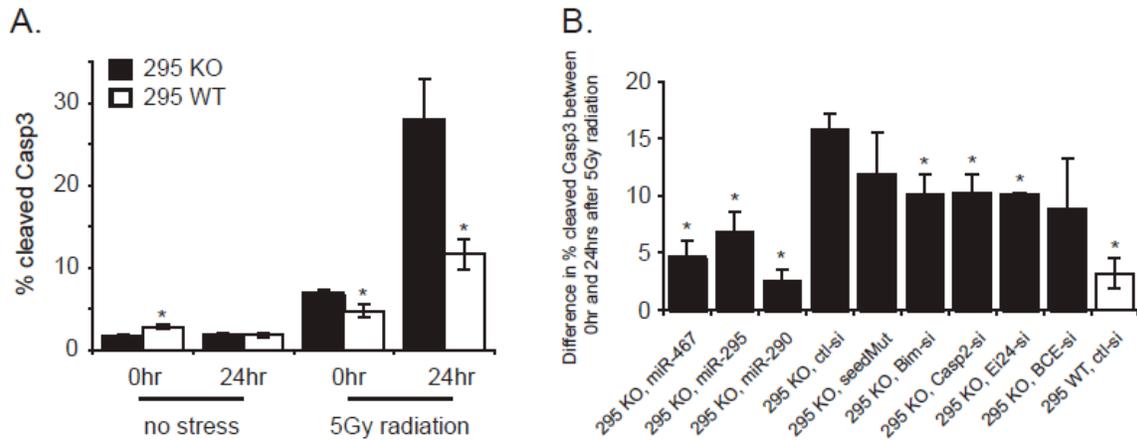


Figure 10. AAGUGC miRNAs protect ES cells from radiation-induced apoptosis.

(A) The percentage of cleaved Casp3 in WT and miR-290-295 KO ES cells under normal culturing conditions (0 and 24 hours after plating) and after exposure to 5-Gy radiation (0 and 24 hours after radiation treatment). Cleaved Caspase-3 was assayed by flow cytometry, and was used to estimate apoptosis response. Apoptosis of KO cells was shown in black bars, and that of WT cells was shown in white bars. (B) Cells were transfected with 50nM siRNAs as shown, and difference in apoptosis response of WT and 295KO ES cells 24 hours after exposure to 5-Gy radiation was plotted. $n \geq 3$ for all experiments. Results are shown as mean \pm S.E.M. (standard error of the mean). P-values were results of Mann-Whitney tests, and * denotes $p \leq 0.05$.

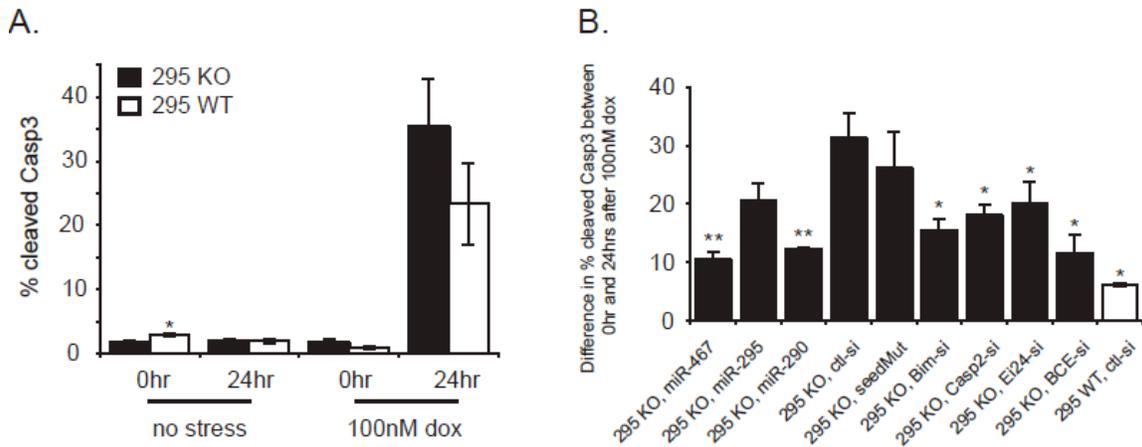


Figure 11. AAGUGC miRNAs protect ES cells from doxorubicin-induced apoptosis.

(A) The percentage of cleaved Casp3 in WT and miR-290-295 KO ES cells under normal culturing conditions (0 and 24 hours after plating) and after exposure to 100nM doxorubicin (0 and 24 hours after radiation treatment). Cleaved Caspase-3 was assayed by flow cytometry, and was used to estimate apoptosis response. Apoptosis of KO cells was shown in black bars, and that of WT cells was shown in white bars. (B) Cells were transfected with 50nM siRNAs as shown, and difference in apoptosis response of WT and KO ES cells 24 hours after exposure to 100nM doxorubicin was plotted. $n \geq 3$ for all experiments. Results are shown as mean \pm S.E.M. (standard error of the mean). P-values were results of Mann-Whitney tests, * denotes $p \leq 0.05$, and ** denotes $p \leq 0.01$.

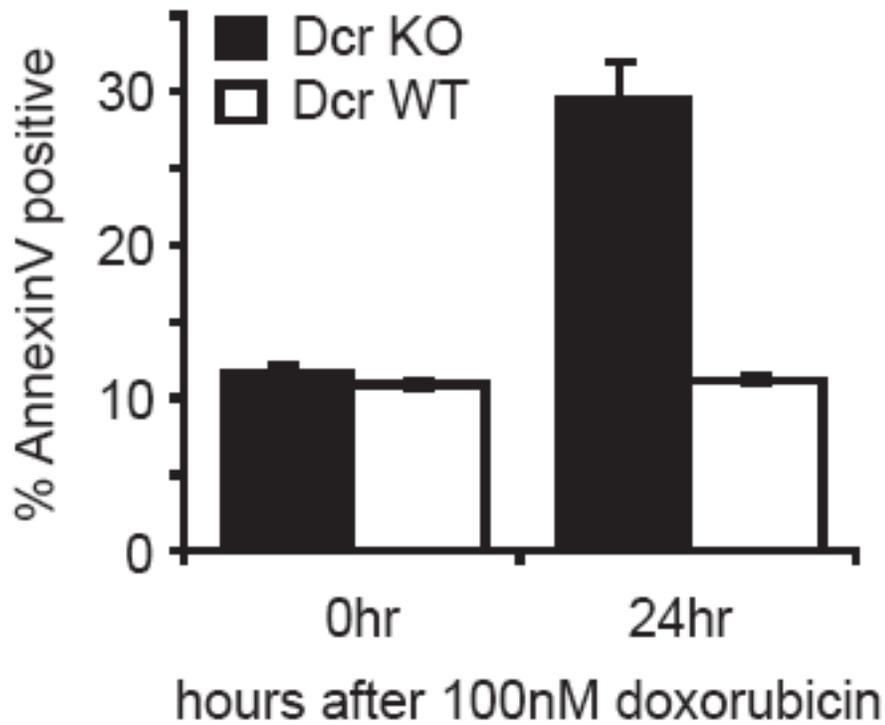


Figure 12. Comparison of WT and Dcr KO cells' apoptosis response to stress with AnnexinV assay.

AnnexinV positive cells were assayed by flow cytometry immediately or 24hr after exposure to 100nM doxorubicin. n=2. Results are shown as mean \pm S.E.M. (standard error of the mean).

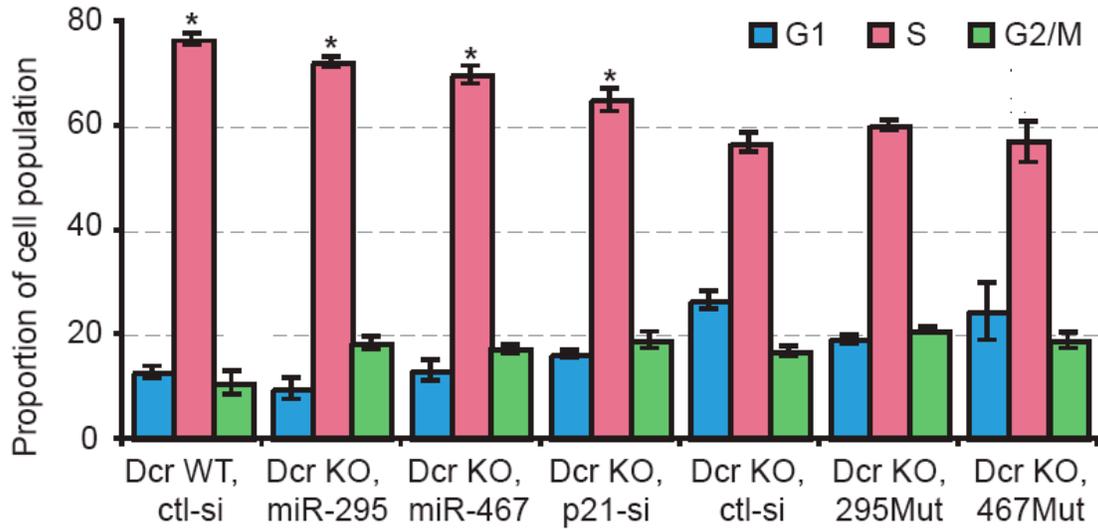


Figure 13. AAGUGC miRNAs can promote G1 to S phase transition of Dcr KO ES cells.

Dcr KO ES cells were transfected with 50 nM of miR-295, miR-467, and other siRNAs as shown. 24 hours after the transfection, cells were incubated with BrdU for 10 min, and BrdU positive cells were analyzed with flow cytometry. Assays with miR-295 seed mutant, miR-467 seed mutant, and control siRNAs serve as negative controls. Results are percentages in each stage of the cell cycle and are shown as mean \pm S.E.M. (standard error of the mean). $n \geq 3$ for miR-295, miR-467, and ctl transfections. $n = 2$ for miR-295 seed mutant, miR-467 seed mutant, and p21 siRNA transfections. P-values were results of Mann-Whitney tests, and * denotes $p \leq 0.05$.

CHAPTER 4: Characterization of the *Sfmbt2* miRNA cluster and its function in murine placental development

The work included in the chapter is a preliminary characterization of the *Sfmbt2* cluster. Much of the analysis on the evolutionary impact of species-specific miRNAs is ongoing, and needs to be extended before final conclusions can be reached. In this chapter, we will point out limitations of the study, and propose directions that need to be explored in the future.

Experimental contributions:

Arvind Ravi cloned all the luciferase constructs, and performed all the luciferase assays. Arvind Ravi and Grace Zheng discussed all the computational analysis, and Grace Zheng performed all the computational analysis.

Abstract

miRNAs are endogenous ~22nt RNAs that post transcriptionally regulate gene expression and control fundamental cellular processes. While targets and functions of some well conserved miRNAs have been characterized through comparative genomic analysis, studies about species-specific miRNAs have proved challenging. Here we characterized a novel, mouse-specific miRNA cluster that is upregulated in murine placenta. Placental-expressed, mouse-specific targets of the cluster are enriched in pathways regulating growth and apoptosis, suggesting they play an important role in promoting placental growth. In addition, we compared mouse and human 3' UTRs to show that many target sites of the cluster show positive selection. We have begun to extend this analysis to other species-specific miRNAs.

Introduction

miRNAs are key regulators of gene expression, and play important roles in development, cell growth, differentiation, and metabolism (Bartel 2009). Hundreds of miRNAs have been identified in mammals, and approximately 50% of known miRNAs are found in clusters, transcribed as polycistronic primary transcripts (Kim et al. 2009). While many miRNA clusters are conserved, recent efforts in large-scale sequencing have uncovered poorly conserved miRNA clusters that have tissue-specific expression (Zhang et al. 2008).

The *Sfmbt2* miRNA cluster maps to chromosome 2 of the mouse genome. The *Sfmbt2* cluster is localized in an intron of *Sfmbt2* (Figure 1A), a poorly characterized Polycomb Group gene. Although the coding region of the *Sfmbt2* gene is highly conserved among vertebrates, the intron which harbors the miRNA cluster bears little similarity to the equivalent intron outside of rodent species. Five of the 42 mouse *Sfmbt2* miRNAs can be mapped to the homologous intron in rat, but none can be aligned to the corresponding intron in human. Expression of the *Sfmbt2* cluster was initially detected in murine T cells, and subsequently cloned from mouse ES cells (Table 1).

We are interested in further characterizing the *Sfmbt2* cluster in the placenta for two reasons. First, the host gene, *Sfmbt2*, was recently identified as an imprinted gene expressed preferentially from the paternal allele in early embryos and in later stage extraembryonic tissues (Kuzmin et al. 2008). Imprinted gene expression has been observed in the placenta, and linked to placental functions (Kaneko-Ishino et al. 2003). Paternally expressed genes are proposed to increase embryonic growth, while maternally expressed genes are proposed to suppress fetal growth (Moore and Haig 1991). Consistent with this theory, Cattanaach *et. al.* showed that paternal duplication of proximal chromosome 2 (which includes the *Sfmbt2* gene) resulted in placental growth

enhancement, whereas maternal disomy resulted in fetal and placental growth reduction, implicating *Sfmbt2* as a candidate for the placental growth effect (Cattanach et al. 2004). Secondly, a survey of miRNA expression across many mouse tissues revealed that the *Sfmbt2* miRNA cluster is upregulated in the placenta relative to other tissues (Landgraf et al. 2007). Interestingly, we have identified one of the miRNA families in the cluster, miR-467a, as a miRNA family promoting proliferation because it advances the G1 to S phase transition and suppresses apoptosis (see Chapter 3). This is consistent with the hypothesized pro-growth role of the *Sfmbt2* miRNAs as a maternally imprinted placental regulator.

In this study we investigated the possible role of *Sfmbt2* miRNA cluster in contributing to murine placental growth. We first characterized the *Sfmbt2* miRNA cluster and confirmed its expression in murine placenta. Mouse-specific targets of the *Sfmbt2* miRNA cluster are enriched in pathways regulating cell survival, implicating the *Sfmbt2* miRNA cluster as a possible promoter to placental growth. While many genes are subject to selection during evolution to enrich for or avoid miRNA binding sites by changes in 3' UTR length and in site density (Farh et al. 2005; Stark et al. 2005), it is unclear if recently evolved and species-specific miRNAs have similar effects on their target genes. We therefore explored the impact of the *Sfmbt2* miRNA cluster and other species-specific miRNAs on shaping the evolution of 3' UTRs. Sequence comparison between mouse and human 3' UTRs revealed that target sites of many human or mouse-specific miRNAs might be under positive selection.

Results

Expression of repeat-derived Sfmbt2 miRNA cluster is upregulated in murine placenta.

The *Sfmbt2* miRNAs collectively map to intron 9 of the *Sfmbt2* gene, with some

of the individual miRNAs also mapping to other sites in the mouse genome (Figure 1A). The intron spans 50kb, and largely consists of simple repeats and B4 SINE repeat elements. The miRNA cluster can be classified into four groups based on their sequence similarities: miR-297s, miR-466s, miR-467s, and miR-669s. While the seed regions of miR-297s, miR-466s, and miR-467s are well conserved within each group, there is little similarity among the 5' ends of sequences in the miR-669 family (Figure 2). This observation suggests that each miRNA precursor in the miR-669 family could have multiple stem-pairing configurations, resulting in different mature miRNA sequences. miR-297s are mapped to two ends of the intron, and miRNAs from miR-466s, miR-467s and miR-669s are part of a 2kb region that tiles across the intron 12 times (Figure 1A).

Short RNA sequencing data from ES cells revealed that the *Sfmbt2* miRNA cluster consists of 42 miRNAs (Table 1). Although most datasets suggest the miRNA cluster is expressed at low levels in ES cells, Calabrese *et al.* showed that the cluster represents ~30% of expressed miRNAs in ES cells (Calabrese *et al.* 2007; Babiarz *et al.* 2008; Ciaudo *et al.* 2009; Leung *et al.* 2010). A closer comparison of the miRNA expression profile in ES cells between Calabrese *et al.* and other studies revealed that the *Sfmbt2* cluster's upregulation is accompanied by a downregulation of miR-290-295 miRNA cluster in Calabrese *et al.* (Calabrese *et al.* 2007; Babiarz *et al.* 2008; Ciaudo *et al.* 2009; Leung *et al.* 2010). This suggests that the ES cells used by Calabrese *et al.* may be more differentiated towards the trophoblast lineage.

The cluster contains 25 miRNA seeds, defined as nucleotides 2-7 of the mature miRNA sequence, and the majority are not found in human miRNAs. One exception is the miR-467a family of miRNAs, which shares the seed AAGUGC with other conserved miRNA clusters, such as the miR-290-295 and miR-302 clusters. Their function to promote cell growth in mouse ES cells has been examined in Chapter 3. The potential functions of other miRNA seeds in the *Sfmbt2* cluster remain unclear.

Publicly available microarray results show that the relative expression of the *Sfmbt2* gene in the placenta is two fold higher than in ES cells (Wu et al. 2009). To test if the same expression trend would be observed with the miRNA cluster, Northern Blot analysis was performed on placental RNA obtained from day 11.5 (d11.5) and day 13.5 (d13.5) embryos, along with RNA obtained from wild type (WT) and Dcr knockout (KO) ES cells (Figure 1B). Expression of the three *Sfmbt2* miRNAs probed was quantified to be five fold higher in the placenta than in WT ES cells. Mature *Sfmbt2* miRNAs are absent from Dcr KO ES cells, confirming that their expression is dependent on Dicer, an RNase III enzyme required to process miRNA precursors to mature miRNAs (Figure 1B). Published microarray data also show that the relative expression of host gene *Sfmbt2* is 1.5 fold higher at d11.5 than d13.5 (Lee et al. 2009). However, our Northern analysis was not sensitive enough to detect a similar expression difference of the *Sfmbt2* miRNAs between the two timepoints. qRT-PCR or short RNA sequencing from the placenta will need to be performed to further examine the difference in the *Sfmbt2* miRNA expression between d11.5 and d13.5. The data will also allow us to assess the relative abundance of individual *Sfmbt2* miRNAs in the placenta.

Predicted targets of Sfmbt2 miRNAs are enriched in pathways regulating cell growth and apoptosis.

To better understand the roles of *Sfmbt2* miRNAs in placental development, we attempted to identify their endogenous targets by combining the predicted target dataset with microarrays of murine placenta at d11.5 and d13.5. The expression of host gene *Sfmbt2* showed a 1.5 fold increase at d11.5 (relative to d13.5) (Lee et al. 2009). Despite the small change, we hypothesized that predicted targets of *Sfmbt2* miRNAs may be more destabilized at d11.5. When comparing the gene expression profiles at d11.5 and d13.5, predicted targets (defined as mRNAs that have at least a A1-7mer or M8-7mer

match) of 13 *Sfmbt2* miRNAs showed a significant decrease in expression at d11.5 compared to control genes matched for 3' UTR length and dinucleotide composition (Figure 3 and Table 2).

Since most *Sfmbt2* miRNA seeds are mouse-specific, we wondered if their mouse-specific targets also displayed a decrease in expression at d11.5. We selected mRNAs that contained the *Sfmbt2* miRNA target sites exclusively in mouse, but not in orthologous positions in dog or human. To focus on potentially real mouse-specific targets, we required the mutation rate at each of the target sites to be significantly higher than that of adjacent regions (defined as the 80-nt 3' UTR sequences upstream and downstream) (See the next section and Methods for more details of the analysis). We found that mouse-specific targets of 12 *Sfmbt2* miRNAs were derepressed in the comparison between d13.5 and d11.5 relative to a control set (generated as above), and five of them showed even greater derepression than all predicted targets (Figure 3 and Table 2).

We next performed gene ontology (GO) analysis on mouse-specific *Sfmbt2* targets. Of all genes expressed in the placenta (defined as expression at d11.5 \geq 8), we looked for GO category enrichment in 330 upregulated targets (defined by a 1.2 fold increase in expression at d13.5 relative to d11.5). The top statistically significant categories were regulation of cell proliferation, negative regulation of cellular processes, and establishment of localization ($p = 1.1e^{-4}$, $8.3e^{-6}$, and $1.8e^{-5}$ respectively). Pathway enrichment (KEGG) analysis also revealed that upregulated targets were enriched in pathways that regulate cell survival, such as MAPK signaling pathways, tumorigenesis, and apoptosis ($p = 3.9e^{-12}$, $7.2e^{-12}$, and $1.4e^{-6}$ respectively). Interestingly, many upregulated targets overlap with validated AAGUGC miRNA targets (such as P21, Lats2, and Casp2) that have been shown to regulate cell cycle and apoptosis of ES cells (Chapter 3).

To validate predictions of mouse-specific targets, we picked several candidate genes based on the degree of downregulation at d11.5 (relative to d13.5), as well as their functional annotations. The candidate genes' 3' UTRs were cloned into luciferase constructs, and expression level was evaluated in Dcr KO ES cells, before and after the transfection of specific *Sfmbt2* miRNAs. Dcr KO ES cells lack mature miRNAs, which allowed us to test the specificity of *Sfmbt2* targets by expressing exogenous miRNAs. Luciferase reporter of *Dedd2*, a proapoptotic factor, contained a binding site for miR-297a-5p, and a binding site for miR-466a-5p. Our preliminary analysis showed that *Dedd2* reporter was repressed 1.5-fold by miR-297a-5p and miR-466a-5p respectively. We are actively cloning and testing other luciferase constructs (Figure 4).

Target sites of many species-specific miRNAs are positively selected to lose binding sites in their 3' UTRs.

As evolutionary variations in miRNA genes contribute to the generation of new regulatory specificities, the unique presence of the *Sfmbt2* miRNA cluster in rodents has prompted us to look for potential mouse 3' UTRs that responded to the creation of the cluster (Bartel 2009). We hypothesized that comparative analysis of 3' UTRs would reveal two groups of variants: A) mRNAs that gained binding sites to *Sfmbt2* miRNAs in their 3' UTRs for downregulation; B) mRNAs that were selected to lose target sites to *Sfmbt2* miRNAs to maintain expression levels in certain cellular states.

We tested for the presence of each group by analyzing aligned 3' UTRs from mouse and human in two steps. For group A, if the *Sfmbt2* miRNA targets are under positive selection for site gain, then we would expect more sites gained in mouse 3' UTRs than control heptamers, and higher variation at the binding site than their adjacent sequences. Accordingly, in Step I, we counted the number of times a *Sfmbt2* miRNA binding site (M8-7mer match) is present in a mouse 3' UTR, but absent in its human

counterpart, and calculated the fraction of seed sites gained in mouse (relative to human) (Figure 5A). To assess if the fraction was higher than what we would expect by chance, we performed the same analysis with control heptamers of similar composition. miRNA target sites with a significantly higher fraction of sites gained in mouse relative to control heptamers were considered for further analysis in Step II, where the variation at the miRNA binding site was compared to that of adjacent sequences (defined as the 80-nt 3' UTR sequences upstream and downstream) (Figure 5B and C). Of the 25 *Sfmbt2* miRNA seeds, target sites of four seeds showed a significant gain signal in mouse (Step I). They are miR-297a-5p, miR-466a-5p, miR-466k, and miR-466l-5p, all of which displayed a distinctly higher mutation rate at the miRNA binding sites than flanking sequences (Step II).

We tested for group B in a similar manner, but looking for sites lost in mouse instead. Therefore we first counted the number of times a *Sfmbt2* miRNA binding site is present in human, but absent in mouse 3' UTRs (Step I). We then focused on target sites with a significant loss signal, and compared their variation to adjacent sequences (Step II). Overall, we found that binding sites of miR-297a-5p and miR-466a-5p showed significant loss signals in mouse. Combined results from groups A and B showed that four out of 25 *Sfmbt2* miRNA seed binding sites displayed positive selection in mouse when compared to human 3' UTRs.

We then extended the analysis to other miRNA seeds in the human and mouse genomes to test if their target sites are also under positive selection. Since the *Sfmbt2* cluster is mouse-specific, we focused on human or mouse-specific miRNAs. Out of 595 human-specific miRNA seeds, we found 32 seeds (5.38%) whose target sites showed significant loss signals (Table 3 and Table 5). This ratio is significantly higher than that of shuffled control heptamers (2.12%, $p=0.0005$ by Chi-square test) (Table 3), suggesting that the positive selection signals from targets of 32 human-specific miRNAs are

statistically significant (Table 5).

Our method also revealed that target sites of seven out of 201 (3.48%) mouse-specific miRNA seeds showed significant loss signals (Table 3). However, this ratio is not significantly different from that of shuffled control heptamers ($p=0.21$ by Chi-square test). Similarly, while we detected gain signals from target sites of 23 human or mouse-specific miRNAs, it is possible that many of them are false positives ($p=0.38$ and $p=0.21$ for human and mouse-specific miRNA targets by Chi-square tests respectively). The results of our analysis may be updated as miRNA and 3' UTR sequences of mouse and human become better annotated. Furthermore, we are actively refining the two-step procedure used to detect positive selection signals among miRNA targets. Improved sensitivity of our method may reveal more miRNA target sites that are under positive selection.

Discussion

Here we provided a preliminary characterization of a novel murine-specific miRNA cluster, the *Sfmbt2* cluster, and investigated its function in regulating placental development. The cluster is upregulated in the placenta, and some mRNAs appear to have evolved binding sites for this cluster after the divergence of mouse from human or dog. Collectively these mRNAs are enriched in pathways that regulate cell proliferation and apoptosis, suggesting that the *Sfmbt2* cluster promotes murine placental growth. Our method also allowed us to detect a total of 32 human-specific miRNAs whose targets have been positively selected to lose miRNA binding sites since the divergence of human from mouse or dog.

The host *Sfmbt2* gene is imprinted in the placenta, and has been implicated in promoting placental growth (Kuzmin et al. 2008). We suggest that the *Sfmbt2* miRNA cluster within an intron of this gene might facilitate this role. However, more genetic

evidence and biochemical assays will be needed to delineate the specific functions of the miRNA cluster and the host gene. SFMBT2 is a PolyComb Group protein, and has been shown to interact with YY1 (Yy1 is the mammalian ortholog of Pho, an important component of *Drosophila* Pho-RC complex) *in vitro* (Kuzmin et al. 2008). It would not be surprising if SFMBT2 regulates the transcription of target genes with YY1 to promote placental growth, in parallel with the proposed miRNA regulation from its intron. In fact, functional interactions between intronic miRNAs and their host genes have been recently demonstrated in tumorigenesis. MCM7 and its intronic cluster, miR-106b~25, cooperate to drive transformation, while overexpression of either individual gene is not sufficient to do so (Poliseno et al.).

miRNA creation and expansion have been linked to major developmental innovations (Lee et al. 2007). Hertel *et. al.* observed that tandem duplications of miRNAs are strongly overrepresented in the vertebrate ancestor, and at the origin of placental mammals (Hertel et al. 2006). One notable example is the creation of paternally-imprinted miR-134 cluster (located at the mouse *Dkl-Gtl2* domain in eutherian mammals (Seitz et al. 2004). Here we reported a maternally-imprinted miRNA cluster that appears to have arisen through de novo creation after the separation of primates and rodents, and undergone expansion through tandem duplication in mouse. Evidence from previous studies on the *Sfmbt2* host gene as well as our study on the miRNA cluster suggests that the miRNA cluster contributes to placental growth in mouse. The speculated role of the *Sfmbt2* cluster is consistent with the parent-offspring conflict hypothesis that has been proposed to explain the evolution and maintenance of imprinting in mammals (Moore and Haig 1991). Paternally expressed genes are proposed to increase embryonic growth, thereby maximizing the competitiveness of individual offspring bearing a particular paternal genome (Moore and Haig 1991). Maternally expressed genes are proposed to suppress fetal growth (Moore and Haig

1991). This would allow a more equal distribution of maternal resources to all offspring and increase transmission of the maternal genome to multiple offspring (Moore and Haig 1991).

We speculate that the unique creation of the miRNA cluster is linked to the heavy reproductive demand on mouse. Murine gestation time is ~19-21 d and average litter sizes are in the range of 12 (in contrast to humans, who have a gestation time of ~9 months and largely singleton pregnancies) (Knox and Baker 2008). Thus it is not surprising that the mouse genome has evolved to ensure successful pregnancy and full development of fetuses within a short period. In fact, mouse mature placenta is dramatically enriched for rodent-specific genes, including prolactin-like hormones and pregnancy-specific glycoproteins (Knox and Baker 2008).

Besides mouse, human and all other eutherians' placenta have evolved to facilitate the broad range of reproductive strategies displayed by respective species (Knox and Baker 2008). Studies have shown that the human placenta is enriched for primate-specific genes, including a miRNA cluster that is exclusively expressed in the placenta (Bentwich et al. 2005; Knox and Baker 2008). This miRNA cluster (Chr19 miRNA cluster) consists of over 50 miRNAs, and is found on human chromosome 19, downstream of an imprinted gene *Znf331*, and upstream of the miR-371-372 cluster. Members of the miR-371-372 cluster have been linked to promoting tumorigenesis, and its mouse homolog miR-290-295 cluster can promote proliferation and suppress apoptosis (Voorhoeve et al. 2006; Wang and Blelloch 2009). Interestingly, several miRNAs from the Chr19 miRNA cluster share a seed sequence with the miR-372 and miR-295 family, suggesting that this novel miRNA cluster may be involved in growth regulation and human placental development.

Not only do miRNAs contribute to early mammalian development, they also influence the evolution of 3' UTRs. Genes are under evolutionary pressure to maintain or

avoid miRNA complementary sites in their 3' UTRs (Farh et al. 2005; Stark et al. 2005). Additionally, recent evidence suggests that proliferating cells tend to shorten their 3' UTRs to escape global miRNA regulation (Sandberg et al. 2008; Mayr and Bartel 2009). Evolving miRNA-mRNA 3' UTR interactions have also been linked to phenotypic variation among species. For example, Texel sheep have evolved to acquire a miR-1 and miR-126 binding site in the 3' UTR of the myostatin gene (*Gdf8*), whose repression leads to the muscular hypertrophy of Texel sheep (Clop et al. 2006). While previous studies have all focused on the interactions between conserved miRNAs and their conserved targets (and anti-targets, mRNAs that lack conserved binding sites), our study suggests for the first time that species-specific miRNA target sites can be positively selected. In particular, our analysis suggests that 3' UTRs of human genes have undergone selection to avoid regulation by some of human-specific miRNAs. As miRNA and 3' UTR sequences of more species become available and better annotated, we speculate that our methods will reveal more miRNA target sites that are under positive selection. Such method would become an invaluable tool to understand regulatory roles of miRNAs on species-specific developmental processes.

Methods

ES Cell Culture

Feeder-free *Dicer1*^{-/-} mouse embryonic stem cells (mESCs) were generated and maintained on gelatin as described previously (Leung et al. 2010).

Oligos and siRNAs used in all the experiments

siRNA	Sequence (5' → 3' unless otherwise noted)
miR-297a-5p	5'- AUGUAUGUGUGCAUGUGCAUGU -3' 3'- UGUACAUACACACGUACACGUA -3'
miR-466a-5p	5'- UAUGUGUGUGUACAUGUACAUAU -3' 3'- UAAUACACACACAUGUACAUGUA -3'
Control siRNA	(from Dharmacon, Accell Non-targeting pool)
3'UTR Primers	
Dedd2 Forward	AATAACTCGAGGGGAGGCATAACCCCCTGC
Dedd2 Reverse	AATAAGGGCCCCCACCTGTGCCCTTTCCA
Northern LNA Oligos	
miR-297-5p	ACATGCACATGCACACATACAT
miR-466a-5p	ATGTACATGTACACACACATA
miR-467a-5p	CGCATATACATGCAGGCACTTA

Generation of luciferase constructs, mESC transfection, and luciferase assays

MicroRNA mediated repression of each candidate gene was tested by cloning PCR amplified products corresponding to the full length of 3' UTR that contained the miRNA binding sites into the 3' UTR of a pRL-CMV Renilla luciferase reporter as described previously (Doench and Sharp 2004).

Digests were performed using either XhoI or Sall to give the 5' site and ApaI or NotI to give the 3' site. Firefly luciferase (pGL3) was used as a transfection control.

24 hours before transfection $1e^5$ mESC cells were plated/well of gelatinized 24-well plate. Cells were transfected with 2 μ l Lipofectamine 2000 (Invitrogen), 0.1 μ g of CMV-GFP plasmid (Invitrogen), 0.7 μ g of pWS (carrier plasmid), and 50nM of siRNAs in 300 μ l of Opti-MEM (Invitrogen). 4 hours after transfection, transfection mix was removed from cells and replaced with ESC media.

24 hours after transfection, cells were lysed with 1X Passive Lysis Buffer (Promega) and Dual luciferase was measured using Dual Luciferase reporter assay system (Promega) according to manufacturer's instructions.

Northern Blot analysis

Total placental RNA from d11.5 and d13.5 was obtained from Lee's lab. Approximately 36 μ g of each RNA was loaded onto a 12% denaturing UREA gel, according to the Northern Blot protocol outlined previously (Calabrese et al. 2007). Membrane probed with Gln-tRNA was exposed to phosphoimager for 3 hours before being scanned; miR-297a-5p membrane was exposed for 16 hours; miR-466a-5p 24 hours; and miR-467a-5p 24 hours. Prior to hybridizing with a different probe, membranes were stripped by incubating the membrane in boiling 0.1% SDS for 30 minutes and loss of signal was confirmed prior to rehybridization.

Microarray analysis

Microarray data was obtained from and processed according to Lee *et. al.* (Lee et al. 2009). The log₂ fold change (LFC) value for d11.5/d13.5 was defined as the difference between the mean log expression in d11.5 cells and the mean log expression in d13.5 cells. Targets of a *Sfmbt2* miRNA were selected based on at least one match to a T1-A or M2-8 7mer of the mature miRNA sequence. A target was defined as mouse-specific if the orthologous heptamer binding site in human has mutated, and if the mutation rate is significantly higher than that of flanking sequences. Controls were selected to match targets in 3' UTR length and composition, and did not overlap with target sets.

Bioinformatics analysis

Sfmbt2 miRNAs were aligned against mouse (mm9), human (hg18), and rat (rn4) genomes by running BLAST analysis. The genomic sequences were downloaded from UCSC database (<http://genome.ucsc.edu/>).

Human and dog mature miRNA sequences were obtained from miRBase Release 15 (Griffiths-Jones et al. 2008). Mouse mature miRNA sequences were obtained from Chiang *et. al.* (Chiang et al.). miRNA seeds were defined as nucleotides 2-8 of the mature sequence. A miRNA seed was considered shared between mouse and human if it existed in both human and mouse. However, a miRNA seed was defined as mouse specific if it existed in mouse, but not in dog or human. Likewise, a miRNA seed was defined as human specific if it existed in human, but not in dog or mouse.

Sfmbt2 miRNA precursors coordinates were obtained from mirBase (release 15) (Griffiths-Jones et al. 2008), and sequences were extracted from mm9. ClustalW was used to obtain multiple sequence alignments among precursors (Larkin et al. 2007).

Analysis of positive and purifying selection on human and mouse 3' UTRs

Aligned human, mouse, and dog 3' UTRs were obtained from TargetScan 5.1 (Friedman et al. 2009). 17840 3' UTRs were used for the analysis. In Step I, the control distribution for group A (or group B) was generated from heptamers that had a similar number of site matches in mouse (or human) as the heptamer of interest (for example, a miRNA seed). In Step II, the mutation rate of each heptamer binding site was calculated as the fraction of mismatched nucleotides in a 7-nt window. T test was used to check if the mutation rate around the seed binding site was significantly higher than that of neighboring sequences. The mutation around the seed binding site was calculated as the average of the mutation rates of 7 heptamers around the seed binding site (the seed binding site, as well as three heptamers upstream and downstream). The mutation rate of neighboring sequences was calculated as the average of the mutation rates of the rest of heptamers in the 80-nt window.

Gene Ontology analysis

Gene Set Analysis Toolkit (<http://bioinfo.vanderbilt.edu/webgestalt/>) was used to perform GO analysis. *Sfmbt2* targets and mouse specific targets were generated as described in the text. Control genes were defined as all genes that are expressed in the mouse placenta (detected by the microarray analysis).

Statistical analysis

All test statistics were calculated using R (<http://www.r-project.org>). The Wilcoxon rank sum test was used because it does not assume normality of the underlying distributions. T-tests and Kolmogorov–Smirnov (KS) test using these data gave generally similar

results. Chi-square test was used in positive and purifying selection analyses (Steps I and II) to check if a specific category was over-represented relative to the control.

Table 1. Description of the *Sfmbt2* miRNA cluster.

Sequences and cloning statistics of *Sfmbt2* miRNAs in mouse ES cells are listed. The cloning statistics were taken from previous studies in mouse ES cells. ES cells from Leung *et. al.* were used for experiments presented in the chapter.

Id	Sequence	Length	Seed	% cloned in different ES cell lines			
				Calabrese et. al. (a)	Leung et. al. (b)	Babiarz et. al. (c)	Ciaudo et. al. (d)
miR-297a	AUGUAUGUGUGCAUGUGCAUGU	22	UGUAUG	0.8743	0.0003	0.0318	0.0087
miR-297b-5p	AUGUAUGUGUGCAUGAACAUGU	22	UGUAUG	1.6289	0.0013	0.0164	0.0000
miR-297c	AUGUAUGUGUGCAUGUACAUG	21	UGUAUG	1.3791	0.0000	0.0000	0.0000
miR-467a-5p	UAAGUGCCUGCAUGUAUUGCG	22	AAGUGC	4.8136	0.0000	0.0761	0.3291
miR-467b	GUAAGUGCCUGCAUGUAUUAU	20	UAAGUG	0.2295	0.0000	0.0642	0.0532
miR-467c	UAAGUGCGUGCAUGUAUUAUG	21	AAGUGC	0.5354	0.6753	0.0000	0.0310
miR-467d	UAAGUGCGCGCAUGUAUUAUG	21	AAGUGC	0.4091	0.2147	0.0000	0.3605
miR-467e	AUAAGUGUGAGCAUGUAUUAUG	21	UAAGUG	0.4555	0.8718	0.0000	0.0705
miR-467a-3p	AUAUACAUAACACACCCUACAC	22	UAUACA	5.1394	0.0000	0.0761	0.0199
miR-467b*	AUAUACAUAACACACCCAACAC	22	UAUACA	0.0070	0.0000	0.0642	0.0000
miR-467e*	AUAUACAUAACACACCCUAUA	21	UAUACA	0.4856	0.0000	0.0000	0.0000
miR-466a-5p	UAUGUGUGUGUACAUGUACA	21	AUGUGU	0.2457	0.0008	0.0645	0.0000
miR-466b-5p	GAUGUGUGUGUACAUGUACA	21	AUGUGU	0.0765	0.0002	0.0000	0.0000
miR-466c-5p	GAUGUGUGUGUGCAUGUACA	21	AUGUGU	0.2098	0.0000	0.0000	0.0000
miR-466d-5p	UGUGUGUGCGUACAUGUACA	21	GUGUGU	0.0035	0.0000	0.0000	0.0074
miR-466e-5p	GAUGUGUGUGUACAUGUACA	21	AUGUGU	0.0765	0.0000	0.0000	0.0000
miR-466f-5p	UACGUGUGUGUGCAUGUGCAUG	22	ACGUGU	0.0000	0.0000	0.0000	0.020
miR-466h	UGUGUGCAUGUGCUUGUGUGU	24	GUGUGC	0.0598	0.0000	0.0000	0.0000
hp2288	GUGUGCAUGUGGAUGUAUGU	20	UGUGCA	0.0133	0.0000	0.0000	0.0000
miR-466a-3p	UAUACAUAACACGCACACUAAG	22	AUACA	1.0639	0.0000	0.0645	0.0099
miR-466b-3p	UAUACAUAACGCACACUAAGA	23	AUACA	2.1641	0.0000	0.0000	0.0099
miR-466c-3p	AAUACAUAACGCACACUAAG	22	AUACA	0.0070	0.0121	0.0000	0.0000
miR-466d-3p	UAUACAUAACGCACACUA	20	AUACA	2.1208	0.0000	0.0000	0.0050
miR-466f-3p	CAUACAACACACAUACACA	20	AUACAC	0.0127	0.0000	0.0000	0.1683
miR-466g	AUACAGACACAUGCACACAC	20	UACAGA	0.0603	0.0005	0.0000	0.0000
miR-466i	UAUAAAUAACAUACACAUUU	22	AUAAA	0.0399	0.0042	0.0000	0.0000
hp2090	UACAUAACACACAUACACGCA	22	ACAUAC	0.0565	0.0000	0.0000	0.0000
miR-297a*	UAUACAUAACACACAUACCCAU	21	AUACA	1.7557	0.0000	0.0000	0.0000
miR-669a-5p	AGUUGUGUGUGCAUGUUAUGU	22	GUUGUG	0.3488	0.0000	0.1318	0.2117
miR-669b-5p	AGUUUUGUGUGCAUGUGCAUGU	22	GUUUUG	0.1928	0.0000	0.0554	0.8918
miR-669c-5p	AUAGUUGUGUGUGGAUGUGUGU	22	UAGUUG	0.0931	0.2609	0.3251	0.0756
miR-669d	ACUUGUGUGUGCAUGUAUUAUGU	22	CUUGUG	0.6050	0.0000	0.0000	0.0000
miR-669e	UGUCUUGUGUGUGCAUGUUAUGU	22	GUCUUG	0.1363	0.0000	0.0000	0.0000
miR-669g	UGCAUUGUAUGUGUACAUGAU	23	GCAUUG	0.0033	0.0000	0.0000	0.0000
miR-669h	AUGCAUGGUGUAUAGUUGAGUGC	24	UGCAUG	0.0432	0.0000	0.0000	0.0000
miR-669a-3p	ACAUAACAUAACACACACGUAU	23	CAUAAC	4.4878	0.0000	0.1318	0.0730
miR-669b-3p	AUAUACAUAACACACAAACAUAU	22	UAUACA	0.2959	0.0000	0.0554	0.0000
miR-669f	CAUAUAACAUAACACACGUAU	23	AUAUAC	0.5053	0.0002	0.0000	0.0000
miR-669i	UGCAUAUAACACACAUGCAUAC	21	GCAUUA	0.0332	0.0000	0.0000	0.0000
miR-669j	UGCAUAUAACACACAUGCAAACA	22	GCAUAU	0.0199	0.0000	0.0000	0.0000
miR-669k	UAUGCAUAUAACACGCAUGCAA	21	AUGCAU	0.0266	0.0000	0.0000	0.0000
hp2252	UAUGCAUAUAACACACAUGUACA	22	AUGCAU	0.0066	0.0000	0.0000	0.0000
sfmbt2 miRNA cluster / total miRNA reads (%)				30.7213	2.0424	1.1576	2.3451

(a) J. M. Calabrese, A. C. Seila, G. W. Yeo, and P. A. Sharp, Proc Natl Acad Sci, 2007. 104 (46).
 (b) A. Leung, A. Young, AJ Bhukar, G. Zheng, A. Boson, and P.A. Sharp. 2010Submitted.
 (c) J.E. Babiarz, J.G. Ruby, Y. Wang, D.P. Bartel, and R. Blelloch. Genes Dev, 2008. 22 (20).
 (d) C. Ciaudo et al., Plos Genetics 5, 2009, e1000620.

Table 2. Target statistics of *Sfmbt2* miRNAs.

LFC of predicted targets (between d11.5 and d13.5) of each *Sfmbt2* miRNA were compared to those of controls by Wilcoxon test (*p1*), and LFCs of mouse specific *Sfmbt2* miRNA targets were compared to those of controls by Wilcoxon test (*p2*). “total targets” – number of all predicted targets of a specific *Sfmbt2* miRNA seed; “mouse specific targets” – number of mouse specific targets of a *Sfmbt2* miRNA seed.

miRNA seed	miRNA	<i>p1</i>	<i>p2</i>	total targets	mouse specific targets
CUUGUGU	miR-669d	0.00002	0.01846	921	399
UAUACAU	miR-467a-3p/467b*/467e*/669b-3p	0.00005	0.00181	1407	519
GUGUGCA	miR-466hj	0.00044	0.00134	1196	526
GUCUUGU	miR-669e	0.00066	0.12920	678	258
AUGCAUA	miR-669h-3p/669k	0.00108	0.00642	865	239
UGUAUGU	miR-297a/297b-5p/297c	0.00147	0.02438	1033	553
AAGUGCC	miR-467a	0.00219	0.06870	781	239
GCAUUGU	miR-669g	0.00347	0.01990	578	212
GUGUGUG	miR-466d-5p/466k	0.00607	0.11072	1616	977
AUAUACA	miR-669f	0.01177	0.04886	2014	659
UAAGUGU	miR-467e	0.01722	0.00427	731	328
UAGUUGU	miR-669c	0.03293	0.70988	489	202
GUUUUGU	miR-669b	0.04674	0.00062	1329	615
GCAUAUA	miR-669ij	0.08557	0.38816	636	213
ACGUGUG	miR-466f-5p	0.10080	0.32340	365	202
UGCAUGG	miR-669h-5p	0.11350	0.07416	854	359
UAAGUGC	miR-467b	0.11701	0.07844	627	173
AUACAUA	miR-297b-3p/466a-3p/466b-3-3p/466c-3p/466d-3p/466e-3p/467g	0.14642	0.03100	1548	665
AUACACA	miR-466f-3p	0.17663	0.00965	1908	850
AAGUGCG	miR-467c/d	0.22560	0.35483	165	62
AUAAAUA	miR-466l	0.22908	0.05234	2832	500
UGUGCAU	hp2288	0.24434	0.00252	1080	432
CAUAACA	miR-669a-3p	0.27217	0.15256	849	357
AUGUGUG	miR466a-5p/466b-5p/466c-5p/466e-5p	0.32233	0.55551	1237	688
GUUGUGU	miR-669a	0.32867	0.56666	596	292
ACAUACA	hp2090	0.37034	0.30598	1827	887
UACAGAC	miR-466g	0.49446	0.87167	977	304

Table 3. Summary statistics of positive selection signals detected from heptamer binding sites in mouse and human 3' UTRs.

The analysis was performed in two steps, and each step consisted of two tests. Here we explained the details of the analysis by searching for targets that have specifically gained binding sites to a mouse-specific miRNA seed. The same analysis can be extended to other human or mouse-specific miRNA seeds to look for gain or loss signals in target 3' UTRs.

In Step I, we compared mouse to human 3' UTRs to check if the mouse-specific miRNA seed is preferentially gained in mouse 3' UTRs. First, we calculated the fraction of the seed site gained in mouse. Then we compare the fraction to that of control heptamers. The seed sites that displayed a significantly higher fraction of site gain than heptamer controls were considered significant ($p \leq 0.05$). In Step II, we compared mouse to human 3' UTRs to find the variation at the target site as well as variations along its 80-nt flanking sequences. The percentage of nucleotides changed (% mutation) was calculated for the miRNA seed, as well as for heptamers that were 40-nt upstream, and 40-nt downstream. If the mutation rate at the miRNA binding site was much higher than those of adjacent heptamers, target sites of the miRNA seed was considered to be under positive selection for site gain ($p \leq 0.05$).

“count” represents the number of heptamers that were significant in each test. “total” represents the total number of possible heptamers in the specific test category. *P-values* were the result of Chi-square tests. Numbers colored in red represent the ratios that were significantly higher than the controls (the ratios from “all heptamers” categories). There are 16384 total heptamers, and 1117 distinct M2-8 miRNA seeds in mouse and human miRNAs combined. There are 796 mouse or human specific miRNAs (dog miRNAs were used as an outgroup).

Step I				
gain events				
	count	total	count/total (%)	<i>p</i>
all heptamers	765	16384	4.67	
human-specific miRNA seeds	38	595	6.39	0.077
mouse-specific miRNA seeds	13	201	6.47	0.31
human + mouse summary	51	796	6.41	0.036
loss events				
	count	total	count/total (%)	<i>p</i>
all heptamers	765	16384	4.67	
human-specific miRNA seeds	49	595	8.24	0.001
mouse-specific miRNA seeds	12	201	5.97	0.39
human + mouse summary	61	796	7.66	0.001

Step II				
gain events				
	count	total	count/total (%)	<i>p</i>
all heptamers	347	16384	2.12	
human-specific miRNA seeds	16	595	2.69	0.38
mouse-specific miRNA seeds	7	201	3.48	0.21
human + mouse summary	23	796	2.89	0.16
loss events				
	count	total	count/total (%)	<i>p</i>
all heptamers	347	16384	2.12	
human-specific miRNA seeds	32	595	5.38	0.0005
mouse-specific miRNA seeds	7	201	3.48	0.21
human + mouse summary	39	796	4.90	0.001

Table 4. Target sites of 32 human species-specific miRNAs are under positive selection for site loss.

miRNA seed	miRNA	direction of target site selection
CAUAGCC	hsa-miR-103-as	site loss in human
GGAUGGU	hsa-miR-1193	site loss in human
GGAUGAG	hsa-miR-1255a hsa-miR-1255b	site loss in human
GGGUGGU	hsa-miR-1293	site loss in human
CAUGGGU	hsa-miR-1308	site loss in human
AUAGGGA	hsa-miR-135a*	site loss in human
UGUAGGG	hsa-miR-135b*	site loss in human
UGUGUGG	hsa-miR-147	site loss in human
CCCCACA	hsa-miR-1975	site loss in human
GACAGCG	hsa-miR-2277	site loss in human
GGGUAGA	hsa-miR-3132	site loss in human
UAGGGAG	hsa-miR-3162	site loss in human
AGGACUG	hsa-miR-3169	site loss in human
UGGGGUU	hsa-miR-3170	site loss in human
AGUGAGU	hsa-miR-3174	site loss in human
GAAGAAG	hsa-miR-3185	site loss in human
GGGACUG	hsa-miR-3199	site loss in human
GGGUGGA	hsa-miR-363*	site loss in human
UAGGAGG	hsa-miR-4266	site loss in human
UAGGGGG	hsa-miR-4278	site loss in human
CCCCACU	hsa-miR-4286	site loss in human
CCCACUA	hsa-miR-4301	site loss in human
GGGUAAG	hsa-miR-555	site loss in human
GAGUGUG	hsa-miR-574-5p	site loss in human
ACACGGG	hsa-miR-602	site loss in human
GGGGUGG	hsa-miR-608	site loss in human
GGGUGUU	hsa-miR-609	site loss in human
GGGGGAA	hsa-miR-625	site loss in human
GCUGUCU	hsa-miR-626	site loss in human
GGUUGGG	hsa-miR-92a-1*	site loss in human
GGUGGGG	hsa-miR-92a-2*	site loss in human
UGACUGU	hsa-miR-943	site loss in human

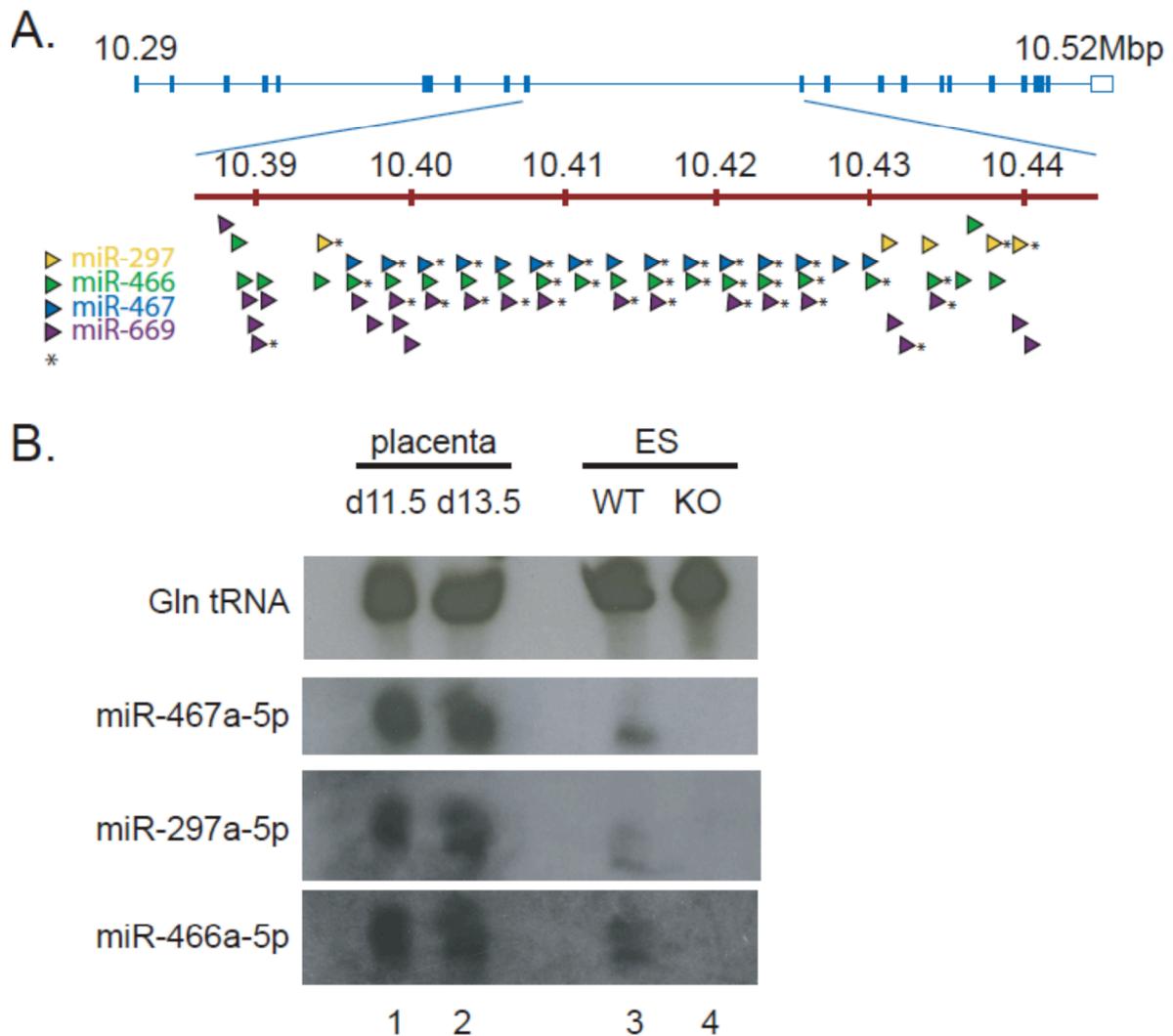


Figure 1. Genomic structure of the *Sfmbt2* miRNA cluster and its expression in placenta.

(A) Precursors of *Sfmbt2* miRNAs were mapped to the 9th intron of the *Sfmbt2* gene by BLAST. The intron spans from 10.39Mbp to 10.44Mbp on Chromosome 2. The *Sfmbt2* miRNAs were color coded, and * refers to miRNAs with both 5' and 3' sequences mapped to the locus. (B) Northern blot analysis of miR-467a-5p, miR-297a-5p, and miR-466a-5p in placental and ES cells. Gln tRNA was probed as a loading control. Lane 1, placental d11.5; 2, placental d13.5; 3, WT ES cells; 4, KO ES cells.

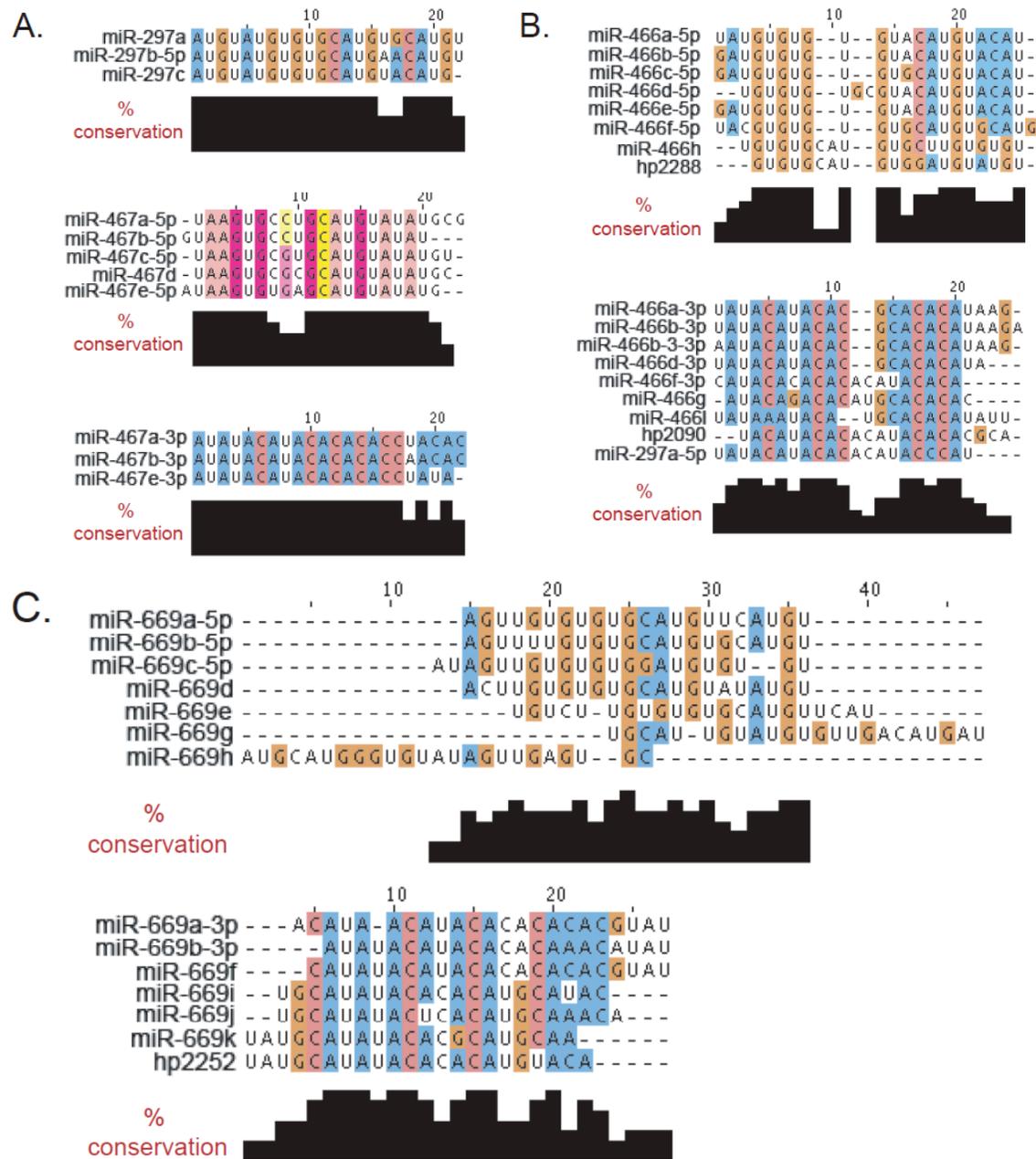


Figure 2. Multiple sequence alignments of *Sfmbt2* miRNAs.

Precursors of *Sfmbt2* miRNAs were aligned with ClustalW. Alignments of mature sequences and % conservation calculated along the mature sequences were shown for each miRNA family found in *Sfmbt2* miRNA cluster. (A) miR-297 and miR-467 families. (B) miR-466 families. (C) miR-669 families.

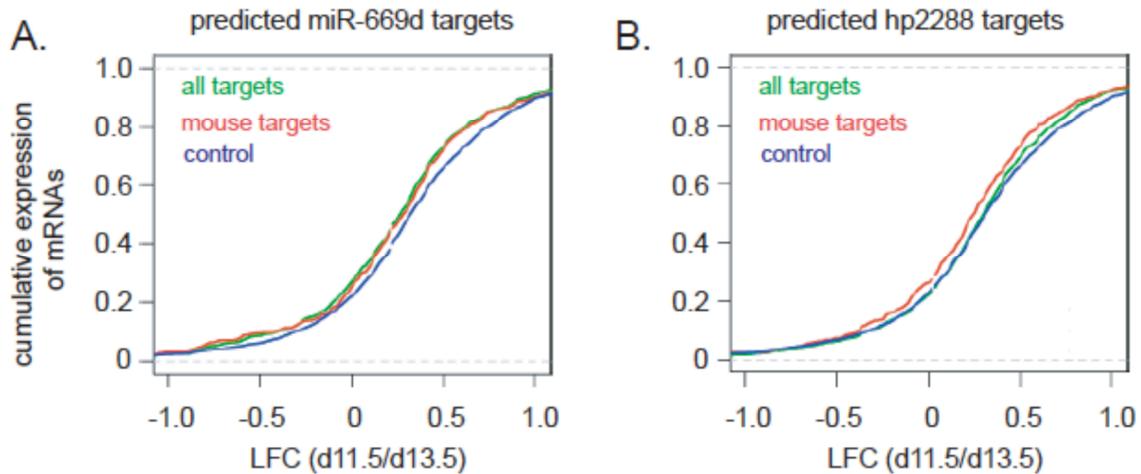


Figure 3. Predicted targets of *Sfmbt2* miRNAs show decreased stability in d13.5 placenta.

Representative CDFs (cumulative distribution functions) of log₂ fold change (LFC) in mRNA expression between d11.5 and d13.5 placenta cells are plotted for two miRNA target sets that showed the most significant difference between “*all targets*” and “*control*” (A) and “*mouse targets*” and “*control*” (B). (A) “*all targets*” (green) include 399 predicted targets of miR-669d, and “*mouse targets*” (red) include 921 mouse specific targets. The control mRNA (blue) was selected to match the predicted targets in di-nucleotide composition and 3' UTR length. “*all targets*” and “*mouse targets*” are both de-repressed in d13.5 placental cells ($p=2.0e^{-5}$ and $p=1.85e^{-2}$ by rank sum test respectively). (B) CDFs were plotted for 1080 predicted targets and 432 mouse specific targets of hp2288, an unannotated miRNA from the *Sfmbt2* cluster. Although “*all targets*” are not derepressed relative to the control mRNAs at d13.5, “*mouse specific*” targets are ($p=2.52e^{-3}$ by rank sum test).

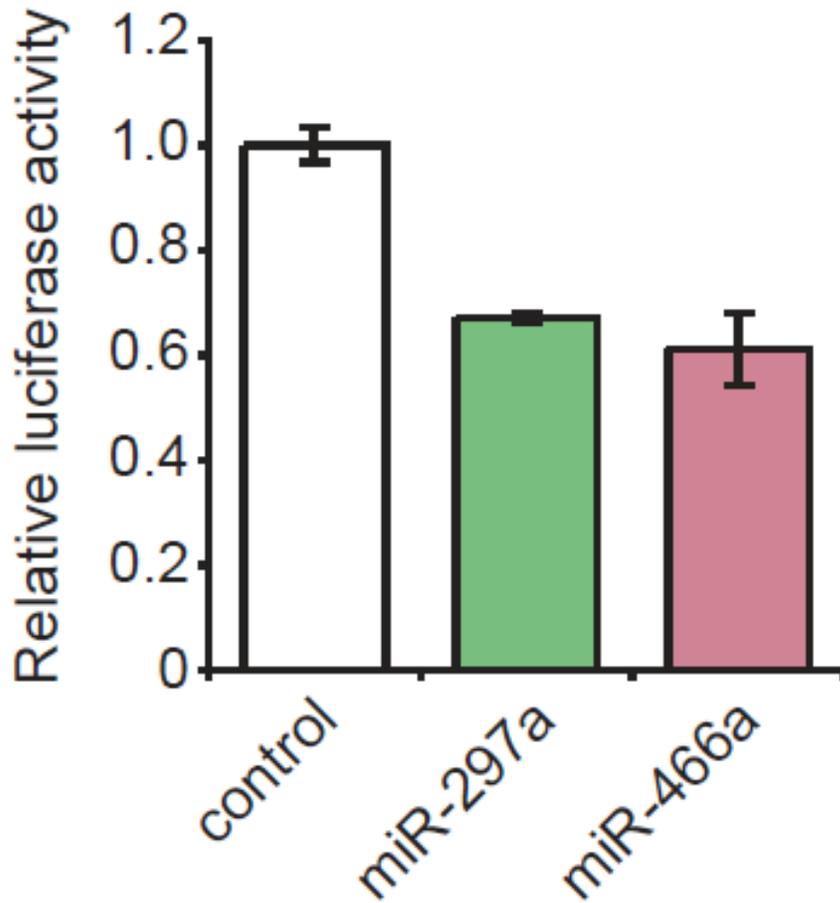


Figure 4. Repression of *Dedd2* by two of *Sfmbt2* miRNAs.

Luciferase reporter with full length *Dedd2* 3' UTR was assayed in Dcr KO ES cells. 20nM of control siRNA, miR-297a-5p, and miR-466a-5p were transfected in Dcr KO ES cells. n=3 and results are shown as mean ± S.E.M.

A. Steps taken to test for site gain of a mouse-specific miRNA

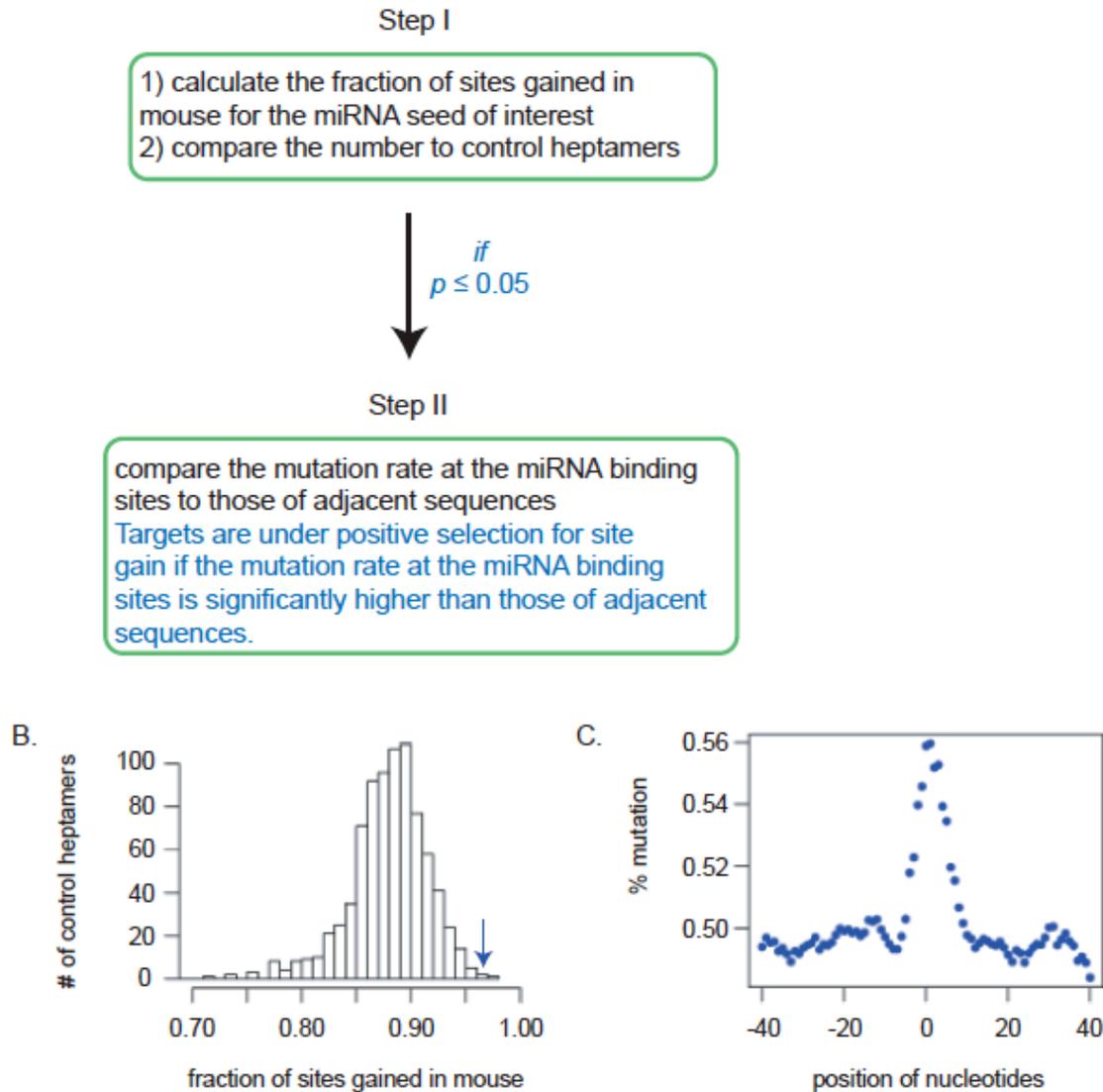


Figure 5. Illustration of the positive selection analysis for site gain of a mouse-specific miRNA.

(A) A flow chart of the analysis. The analysis was divided into 2 steps. In Step I, the fraction of sites gained in mouse was calculated for the mouse-specific miRNA. Then the fraction was compared to those of control heptamers. If the fraction of the sites gained for the mouse-specific miRNA was significantly higher than that of controls, the miRNA seed was considered for Step II of the analysis, where the mutation rate at the miRNA binding sites was compared to those of adjacent sequences. Targets were considered to be positively selected for site gain if the mutation rate at the miRNA binding site is significantly higher than those of adjacent sequences. (B) Histogram of the fraction of sites gained in mouse for control heptamers. The blue arrow points at

the fraction of sites gained for the miRNA seed of interest. (C) Mutation rates of the miRNA binding site and its flanking sequences. % mutation = fraction of mismatched nucleotides in the 7-nt window. “0” - the miRNA binding site, “40” – 40-nt downstream of the miRNA binding site, “-40” – 40-nt upstream of the miRNA binding site.

Appendix: Exploring miRNA targets that are under purifying selection

Comparative genomic analysis revealed that the 5' region of miRNAs is the most conserved portion of miRNAs (Lewis et al. 2003; Lim et al. 2003a; Lewis et al. 2005; Chen and Rajewsky 2006). Seed binding sites in the 3' UTRs tend to be more highly conserved than expected by chance (Lewis et al. 2003; Lim et al. 2003a; Lewis et al. 2005). As a proof of principle, we extended our method to look for purifying signals of miRNA target sites in two steps. If target sites of a miRNA are under purifying selection, then we would expect more sites maintained between human and mouse 3' UTRs, and lower variation at the binding site than their adjacent sequences. Accordingly, in Step I, we counted the number of times a miRNA binding site (M8-7mer match) is present in the human 3' UTR, but absent in its mouse counterpart (Figure 1). The fraction of sites gained in human was calculated, and compared to that of controls. Target sites that showed a conservation signal should have a significantly lower number of sites gained in human than expected. In Step II, we focused only on miRNAs target sites with significant conservation signals, and assessed their variation relative to adjacent sequences (defined as the 80-nt UTR sequences upstream and downstream) (Figure 1). Of the 1117 miRNA seeds, target sites of 71 (6.36%) of them showed a significant conservation signal between human and mouse. This is significantly higher than the proportion of control heptamers with conserved target sites (2.56%, $p=0.00050$ by Chi-squareTest) (Table 1). Furthermore, target sites of almost 20% of conserved miRNAs (defined as shared between mouse and human) are under purifying selection, a result that is consistent with results from previous studies (Lewis et al. 2003; Lim et al. 2003a; Lewis et al. 2005; Friedman et al. 2009) (Table 1).

Table 1. Summary statistics of purifying selection signals detected from heptamer binding sites in mouse and human 3' UTRs.

The analysis was performed in two steps. In Step I, we compared human to mouse 3' UTRs to find heptamers that showed significant conservation signals between human and mouse. In Step II, we compared human to mouse 3' UTRs to find the variation at the target site as well as variations along its 80-nt flanking sequences. The percentage of nucleotides changed (% mutation) was calculated for the heptamer, as well as for heptamers that were 40-nt upstream, and 40-nt downstream. If the mutation rate of the heptamer was much lower than those adjacent heptamers, the heptamer was considered to be under purifying selection ($p \leq 0.05$).

“count” represents the number of heptamers that were significant in each test. “total” represents the total number of possible heptamers in the specific test category. Numbers colored in red represent the ratios that were significantly higher than the controls (the ratios from “all heptamers” categories). There are 16384 total heptamers, and 1117 distinct M2-8 miRNA seeds in mouse and human miRNAs combined. 218 are shared between mouse and human.

Step I

purifying selection			
	count	total	count/total (%)
all heptamers	788	16384	4.81
all miR seeds	98	1117	8.77
miR seeds shared by human and mouse	54	218	24.77

Step II

purifying selection			
	count	total	count/total (%)
all heptamers	419	16384	2.56
all miR seeds	71	1117	6.36
miR seeds shared by human and mouse	43	218	19.72

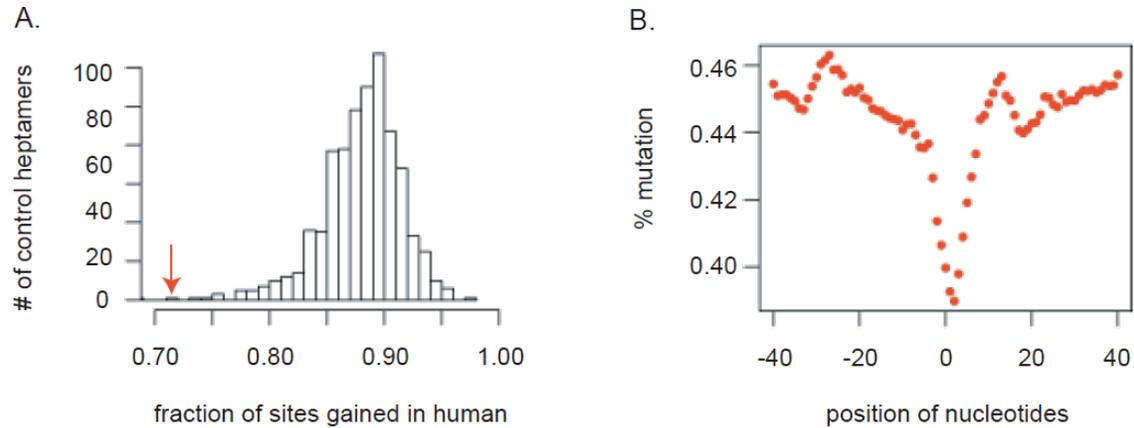


Figure 2. Illustration of the purifying selection analysis.

The analysis was divided into 2 steps. In Step I, the fraction of sites gained in human was calculated for each heptamer. If the heptamer gained fewer sites in human than expected ($p \leq 0.05$), it was considered for further analysis in Step II, where the mutation rate of the miRNA binding sites was compared to those of heptamers in the neighboring 80-nt sequence. (A) Histogram of the fraction of sites gained for control heptamers. The red arrow points at the fraction of sites gained for the miRNA seed of interest. (B) Mutation rates of the miRNA binding site and its flanking sequences. % mutation = fraction of mismatched nucleotides in the 7-nt window. “0” - the miRNA binding site, “40” – 40-nt downstream of the miRNA binding site, “-40” – 40-nt upstream of the miRNA binding site.

CHAPTER 5: Conclusions and future directions

miRNAs are key modulators of animal development. In this thesis, we have studied the regulatory roles of miRNAs in the development of murine T lymphocytes, ES cells and the placenta. In each chapter, we focused on miRNAs characterized by enriched expression during development. Cloning microRNAs from various stages of T cell development revealed variable expression of miRNAs during this process. miR-181 was significantly enriched at the CD4⁺CD8⁺ (DP) stage of T lymphocytes. We identified several targets that implicated miR-181 in the process of thymocyte positive selection. In ES cells, we focused on a family of miRNAs sharing the AAGUGC seed sequence. This family, the most abundant in embryonic stem cells, impacted ES cell division in the G1 to S phase transition and suppressed apoptosis. Lastly, we characterized a mouse-specific *Sfmbt2* miRNA cluster and speculated in regards to its role in promoting placental growth. While these studies were undertaken in three distinct developmental systems, they only scratch the surface in understanding the complex developmental networks that are regulated by miRNAs.

miRNAs in T lymphocyte development

Cloning statistics of short RNAs in T cells have shown that the expression of miR-181 is elevated at least 10-fold at the DP stage, when thymocytes undergo positive and negative selection to become mature T cells. Studies from our and other labs suggest that miR-181 can enhance TCR signaling strength (Li et al. 2007; Neilson et al. 2007). However, our current understanding of miR-181's role in T cell development is likely incomplete. Firstly, while we are able to associate the functions of a few validated targets of miR-181 to TCR signaling, we do not know the effects of the remaining targets *in vivo*. Secondly, it is not clear the extent to which miR-181 influences the expression of these targets, and the relative importance of each target in T cell maturation. It is also possible that miR-181 regulates additional targets since T cells express more than 40 tyrosine phosphatases and other negative regulators of TCR signaling (Mustelin et al. 2005). Finally, besides TCR signaling, miR-181a may play a role in other pathways such as the costimulatory pathways suggested by Li *et. al.* (Li et al. 2007). All three questions can be addressed by creating a conditional knockout of miR-181 in the T cell lineage, or a miR-181 knockout mouse. The miR-181 knockout mouse will also allow us to examine its role in B lymphoid cells. miR-181 is highly expressed in B lymphoid cells, and its overexpression can bias haematopoietic cell development towards the B cell lineage (Chen et al. 2004).

While miR-181a is upregulated more than 10-fold in DP cells, the mechanism of upregulation remains unknown. It is known that miR-181a and miR-181b are located in tandem on chromosome 2 and they exhibit coordinate upregulation in DP cells. These facts suggest that the mechanism of regulation is more likely transcriptional than post-transcriptional. It has been proposed that the miR-181 family is transcriptionally regulated by MYCN in glioblastoma (Nicoloso and Calin 2008). Understanding how miR-

miR-181 expression is transcriptionally regulated has the potential to not only lend more insight into its role in T cell development, but also its role in oncogenesis as overexpression of miR-181 has been detected in breast, pancreas, and prostate cancers (Calin and Croce 2006).

Beyond our observations regarding miR-181, our short RNA sequencing data throughout T cell development raises several interesting questions that might be better addressed through deep sequencing of short RNAs in T cells. First of all, many of noncoding short RNA sequences mapped to genomic regions annotated as repeats or to which ESTs have been mapped. Additionally, among the clones that overlapped RefSeq “known” genes, less than half of the clones mapped to these annotations in the sense orientation. Fully half of the clones overlapping annotated genes mapped antisense to the protein-coding region of the gene. It is curious what these noncoding RNAs are, and we speculate that some of these clones may be endo-siRNAs, which have been detected in mouse oocytes and ES cells (Babiarz et al. 2008; Tam et al. 2008; Watanabe et al. 2008). The precursors of mouse endo-siRNAs are transcripts that contain long hairpin structures or dsRNAs that are derived from sense-antisense pairs (Tam et al. 2008; Watanabe et al. 2008). Increased depth of coverage in these libraries might determine whether any of the non-coding short RNAs cluster in the genome, and if their genomic sequences can form hairpin-like structures.

Secondly, deep sequencing might help us to better understand whether the process of active RNA editing impacts miRNA function in T cells. In Chapter 2, we attributed the decreased expression of miR-142 at DN3 to ADAR editing. However, we were not able to observe ADAR-edited sequences in our cloning data. While this is in part due to the observation that endogenous ADAR editing sites were often outside the mature miRNA, increased depth of coverage would also allow comprehensive profiling of precursor and mature miRNA sequences at the DN3 stage of thymocyte development.

Lastly, deep sequencing might enable us to uncover other miRNAs and noncoding RNAs playing important roles in T cell development. We first cloned miRNAs from the *Sfmbt2* cluster in developing T lymphocytes, and the expression of members of this cluster were found to be specifically elevated in mature CD4 and CD8 cells. However, the relative expression of *Sfmbt2*-derived miRNA sequences uncovered in T cells is greatly underrepresented relative to their cloning statistics in ES cells, making it difficult to know whether their presence is simply a result of stochastic processes, or signals their functional importance in T cells. Deep sequencing data will allow us to ascertain a non-random presence, and provide a better direction for future investigation of the role of the *Sfmbt2* miRNA cluster in T cells.

miRNAs in ES cells

Our functional assays on Dcr KO and WT ES cells suggest that AAGUGC miRNAs are involved in buffering ES cells from apoptosis, especially when cells are subjected to genotoxic stress. This is consistent with miRNAs' ability to confer robustness upon biological systems. Recent additional work from our lab provides more experimental evidence in multiple systems that demonstrate the central role of miRNAs in apoptosis regulation. Upon oxidative stress, *Dicer* null Sarcoma cells show a much stronger Caspase 3 activation than their wild type counterparts (Figure 1) (Arvind Ravi, unpublished data). Interestingly, sarcoma cells express a very different set of miRNAs from ES cells, suggesting that miRNAs can control the induction of programmed cell death in a cell-autonomous manner. Our preliminary data suggested that ES cells undergo apoptosis in a p53-dependent manner. However, the sarcoma cells under study lack p53 function, implying that miRNAs can control multiple signaling pathways activated by different stress stimuli. Identification of downstream signaling networks upon Caspase 3 activation in ES and Sarcoma cells will help us better understand the

role of miRNAs in regulating stress and apoptosis.

Our work has also shown that AAGUGC miRNAs can affect apoptosis by downregulating the expression of certain proapoptotic factors (Casp2 and Ei24) that are direct targets of the miRNAs. Recent studies suggest an interesting link between apoptosis and ES cell differentiation. Upon DNA damage, induced P53 was found to suppress Nanog transcription, and induce differentiation (Lin et al. 2005). Additionally, activated Casp3 can induce the cleavage of NANOG protein (Larsen et al. ; Fujita et al. 2008). These observations present an alternative for ES cells to maintain genetic stability by differentiating into other cell types. However, future studies will need to address how cells decide to undergo apoptosis or differentiation upon DNA damage.

miRNAs in placental development

The *Sfmbt2* miRNA cluster is a maternally-imprinted, mouse-specific miRNA cluster that has affected the evolution of mouse 3' UTRs. Many mRNAs that have been positively selected to gain target sites to this cluster are enriched in pathways regulating cell growth and apoptosis, suggesting that the miRNA cluster can promote murine placental growth. However, our study opens up more questions on the role of the *Sfmbt2* miRNA cluster in placental development.

Although we have tested a few mouse-specific *Sfmbt2* miRNA targets in Dcr KO ES cells, further experiments are required to demonstrate their relevance in regulating placental growth. First, target repression by *Sfmbt2* miRNAs needs to be confirmed in trophoblast stem cells or choriocarcinoma cell lines. Since our preliminary data suggest that mouse-specific target sites confer weak repression, it will be useful to focus on mRNAs that have multiple miRNA binding sites with favorable 3' UTR context. Gain and loss of function experiments can then be performed in cell cultures to address their potential function in promoting cell proliferation and suppressing apoptosis.

Sfmbt2 has been implicated in promoting placental growth, because paternal duplication of *Sfmbt2*-containing regions enhanced placental growth, while maternal duplication reduced placental growth (Cattanach et al. 2004). It will be important to establish a direct functional link between the *Sfmbt2* gene and placental growth. To this end, we have obtained insertional gene trap mutants from the Bay Genomics consortium to disrupt the expression of the *Sfmbt2* locus. To parse the function of the miRNA cluster from the *Sfmbt2* gene, we utilized two types of insertion. An insertion upstream of the cluster should disrupt the function of the gene and the miRNA cluster, while an insertion downstream of the cluster should only disrupt the gene, leaving the expression of the miRNA cluster intact. Preliminary evidence showed that while the loss of the gene alone did not have any observable developmental defect, loss of the gene and the miRNA cluster resulted in abnormal blastocysts resembling the *Cdx2* knockout phenotype (Joel Neilson, unpublished data). *Cdx2* is a key transcription factor required for the establishment of a functional trophectoderm. This suggests a link between the *Sfmbt2* locus and trophoblast development (Beck et al. 1995; Strumpf et al. 2005). It will be interesting to further characterize the insertional mutants to ascertain the effect and relative contribution of the *Sfmbt2* gene and miRNA cluster in placental development.

Noncoding RNAs (ncRNAs) are often associated with imprinted gene clusters. In addition to the *Sfmbt2* miRNA cluster, another large imprinted miRNA cluster has been identified at the *Dlk1-Gtl2* domain, a well known maternally imprinted locus that includes growth-promoting genes such as *Peg10* (Seitz et al. 2004). Interestingly, the miRNA cluster is transcribed from the antisense strand, and is only expressed from the maternal allele. Several members of the cluster can bind to *Peg11* mRNA (another maternally imprinted gene) and induce its cleavage (Davis et al. 2005). This observation, along with other examples (such as the interaction between *Igf2* and *Igf2R*) is consistent with the parent-conflict theory, which states that positive growth effects from paternally

expressed genes are usually suppressed by maternally expressed imprinted genes (Moore and Haig 1991; Reik et al. 2003). These examples raise an interesting possibility that a maternally imprinted *Sfmbt2* locus could be targeted by a paternally imprinted gene to balance out its growth-promoting effect in placenta. Most placenta-imprinted genes are paternally-imprinted, and they will serve as a good starting point to look for the proposed interactions (Wagschal and Feil 2006).

The placenta is a hallmark of mammalian development (Jaenisch 1997). It is designed uniquely to provide the fetus with nutrients, gas and waste exchange, and protection from the maternal immune system and environmental stress (Sood et al. 2006). Since miRNAs play an important role in maintaining homeostasis and buffer cells from aberrant environmental stimuli, we would expect miRNAs to be a central player in placental development. However, our current understanding of miRNAs' role in the placenta is in its infancy. There is a lack of functional studies on placental miRNAs (Maccani and Marsit 2009). In addition, existing miRNA expression data has been generated from mature placenta, which provides a starting point for understanding their role in placental development (Landgraf et al. 2007). A comprehensive set of miRNA (and other short RNAs) expression profiles at key developmental time points of the placenta would add another dimension of information about gene expression.

Such a comprehensive set of miRNA expression profiles, combined with other datasets can greatly facilitate our understanding of placental function. This combination of data can be used to address several outstanding questions. First of all, we can use miRNA data to better understand the expression dynamics of the *Sfmbt2* cluster and its role throughout placental development. By examining the expression level of individual miRNAs, we can begin to tease out the relative contribution of members of the cluster, and look into differential regulation of the individual miRNA expression. The analysis can be extended to other miRNA clusters expressed in the placenta. Additionally, we can

focus on miRNAs that show enriched or dynamic expression during placental development, and use transcriptomic data to explore their functions. Secondly, ncRNAs have been implicated in epigenetic regulation, which is prevalent in the placenta (Maltepe et al.; Maccani and Marsit 2009). The deep sequencing data of short RNAs, coupled with methylation and chromatin immunoprecipitation data would enable us to explore the role of short RNAs in regulating DNA methylation and histone modifications.

It will also be interesting to extend the study of miRNAs and other short RNAs to human placenta, which will allow us to understand human placental development, and gain insight into mechanisms that affect immediate and long-term health of the fetus. Given the vast difference between murine and primate reproductive strategies, we would expect to uncover many primate-specific miRNAs that play important roles in their placental development. In fact, a primate-specific miRNA cluster was found to be uniquely expressed in the placenta (Bentwich et al. 2005). The availability of transcriptomic and noncoding RNA data during human placental development would greatly facilitate our understanding of their functions.

In addition to the primate-specific cluster and the *Sfmbt2* miRNA cluster discussed above, an increasing number of species-specific miRNAs have been revealed by deep sequencing studies. Our analysis in Chapter 4 revealed that targets of 32 human-specific miRNAs have evolved to preferentially lose binding sites. While some of these miRNAs are likely to have important functions, the challenge will be to identify a developing system in which the miRNAs are relevant. This task will become increasingly easier as short RNAs are sequenced in multiple organisms across multiple tissues.

Overall, miRNAs are key players in all aspects of mammalian development. They have also been shown to be involved in tumorigenesis and many developmental defects (Stefani and Slack 2008). Since there is a close link between deregulation of normal developmental processes and pathogenesis, our understanding of the regulatory roles of

miRNAs in developmental networks will generate useful insights that can be applied to the treatment of diseases.

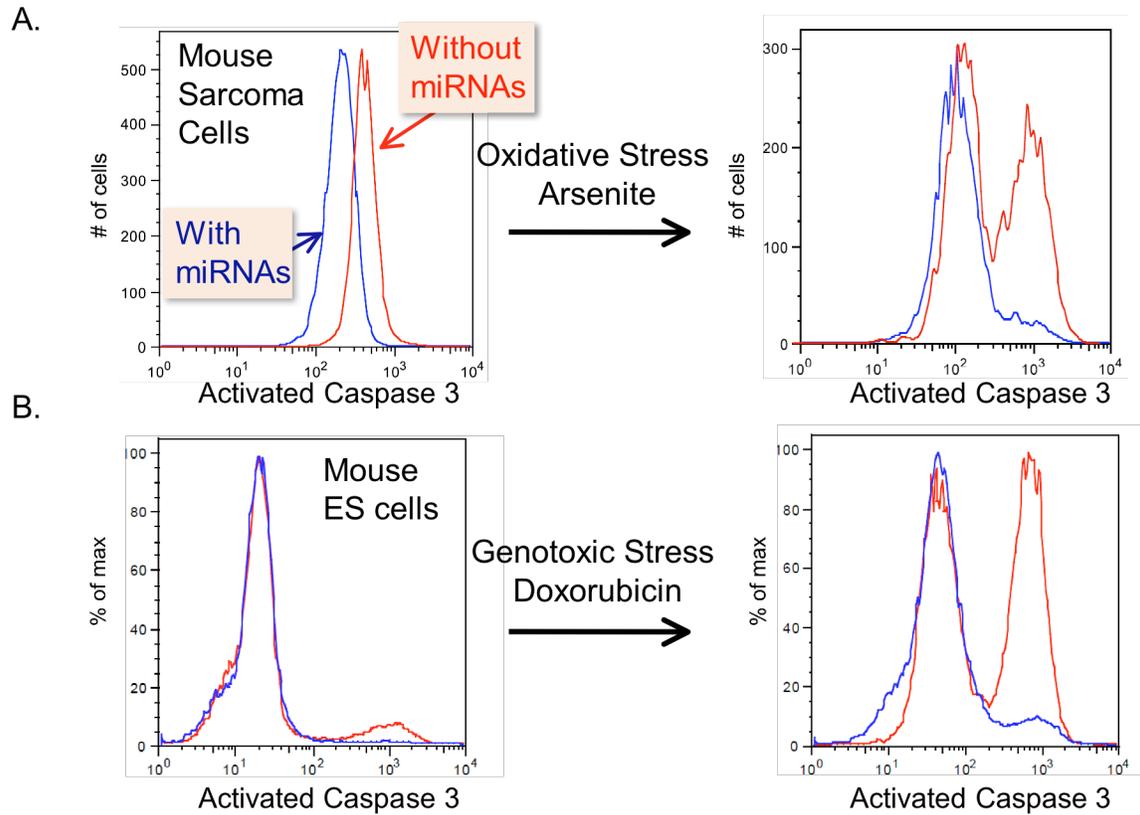


Figure 1. Caspase 3 activation in Mouse Sarcoma and ES cells before and after stress induction.

(A) The percentage of cleaved Casp3 in WT (blue) and Dicer null (red) Sarcoma cells under normal culturing conditions (left) and 19 hours after 25nM Sodium Arsenite treatment (right). (B) The percentage of cleaved Casp3 in WT (blue) and Dicer null (red) ES cells under normal culturing conditions (left) and 24 hours after 100nM doxorubicin treatment (right).

CHAPTER 6: References

- Alfonso, C., McHeyzer-Williams, M.G., and Rosen, H. 2006. CD69 down-modulation and inhibition of thymic egress by short- and long-term selective chemical agonism of sphingosine 1-phosphate receptors. *Eur J Immunol* **36**(1): 149-159.
- Ambros, V. 2004. The functions of animal microRNAs. *Nature* **431**(7006): 350-355.
- Anderson, P. and Kedersha, N. 2006. RNA granules. *J Cell Biol* **172**(6): 803-808.
- Aravin, A., Gaidatzis, D., Pfeffer, S., Lagos-Quintana, M., Landgraf, P., Iovino, N., Morris, P., Brownstein, M.J., Kuramochi-Miyagawa, S., Nakano, T., Chien, M., Russo, J.J., Ju, J., Sheridan, R., Sander, C., Zavolan, M., and Tuschl, T. 2006. A novel class of small RNAs bind to MILI protein in mouse testes. *Nature*.
- Aravin, A. and Tuschl, T. 2005. Identification and characterization of small RNAs involved in RNA silencing. *FEBS Lett* **579**(26): 5830-5840.
- Aravin, A.A., Naumova, N.M., Tulin, A.V., Vagin, V.V., Rozovsky, Y.M., and Gvozdev, V.A. 2001. Double-stranded RNA-mediated silencing of genomic tandem repeats and transposable elements in the *D. melanogaster* germline. *Curr Biol* **11**(13): 1017-1027.
- Aza-Blanc, P., Cooper, C.L., Wagner, K., Batalov, S., Deveraux, Q.L., and Cooke, M.P. 2003. Identification of modulators of TRAIL-induced apoptosis via RNAi-based phenotypic screening. *Mol Cell* **12**(3): 627-637.
- Azuma-Mukai, A., Oguri, H., Mituyama, T., Qian, Z.R., Asai, K., Siomi, H., and Siomi, M.C. 2008. Characterization of endogenous human Argonautes and their miRNA partners in RNA silencing. *Proc Natl Acad Sci U S A* **105**(23): 7964-7969.
- Babiarz, J.E., Ruby, J.G., Wang, Y., Bartel, D.P., and Blelloch, R. 2008. Mouse ES cells express endogenous shRNAs, siRNAs, and other Microprocessor-independent, Dicer-dependent small RNAs. *Genes Dev* **22**(20): 2773-2785.
- Baek, D., Villen, J., Shin, C., Camargo, F.D., Gygi, S.P., and Bartel, D.P. 2008. The impact of microRNAs on protein output. *Nature* **455**(7209): 64-71.
- Bagga, S., Bracht, J., Hunter, S., Massirer, K., Holtz, J., Eachus, R., and Pasquinelli, A.E. 2005. Regulation by let-7 and lin-4 miRNAs results in target mRNA degradation. *Cell* **122**(4): 553-563.
- Bar, M., Wyman, S.K., Fritz, B.R., Qi, J., Garg, K.S., Parkin, R.K., Kroh, E.M., Bendoraitis, A., Mitchell, P.S., Nelson, A.M., Ruzzo, W.L., Ware, C., Radich, J.P., Gentleman, R., Ruohola-Baker, H., and Tewari, M. 2008. MicroRNA discovery and profiling in human embryonic stem cells by deep sequencing of small RNA libraries. *Stem Cells* **26**(10): 2496-2505.
- Bartel, D.P. 2004. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* **116**(2): 281-297.
- . 2009. MicroRNAs: target recognition and regulatory functions. *Cell* **136**(2): 215-233.
- Bartel, D.P. and Chen, C.Z. 2004. Micromanagers of gene expression: the potentially widespread influence of metazoan microRNAs. *Nat Rev Genet* **5**(5): 396-400.
- Beck, F., Ertler, T., Russell, A., and James, R. 1995. Expression of Cdx-2 in the mouse embryo and placenta: possible role in patterning of the extra-embryonic membranes. *Dev Dyn* **204**(3): 219-227.
- Beddington, R.S. and Robertson, E.J. 1989. An assessment of the developmental potential of embryonic stem cells in the midgestation mouse embryo. *Development* **105**(4): 733-737.
- Behm-Ansmant, I., Rehwinkel, J., Doerks, T., Stark, A., Bork, P., and Izaurralde, E. 2006a. mRNA degradation by miRNAs and GW182 requires both CCR4:NOT deadenylase and DCP1:DCP2 decapping complexes. *Genes Dev* **20**(14): 1885-1898.
- Behm-Ansmant, I., Rehwinkel, J., and Izaurralde, E. 2006b. MicroRNAs silence gene

- expression by repressing protein expression and/or by promoting mRNA decay. *Cold Spring Harb Symp Quant Biol* **71**: 523-530.
- Ben-Porath, I., Thomson, M.W., Carey, V.J., Ge, R., Bell, G.W., Regev, A., and Weinberg, R.A. 2008. An embryonic stem cell-like gene expression signature in poorly differentiated aggressive human tumors. *Nat Genet* **40**(5): 499-507.
- Benetti, R., Gonzalo, S., Jaco, I., Munoz, P., Gonzalez, S., Schoeftner, S., Murchison, E., Andl, T., Chen, T., Klatt, P., Li, E., Serrano, M., Millar, S., Hannon, G., and Blasco, M.A. 2008. A mammalian microRNA cluster controls DNA methylation and telomere recombination via Rbl2-dependent regulation of DNA methyltransferases. *Nat Struct Mol Biol* **15**(3): 268-279.
- Bentwich, I., Avniel, A., Karov, Y., Aharonov, R., Gilad, S., Barad, O., Barzilai, A., Einat, P., Einav, U., Meiri, E., Sharon, E., Spector, Y., and Bentwich, Z. 2005. Identification of hundreds of conserved and nonconserved human microRNAs. *Nat Genet* **37**(7): 766-770.
- Berezikov, E., Chung, W.J., Willis, J., Cuppen, E., and Lai, E.C. 2007. Mammalian mirtron genes. *Mol Cell* **28**(2): 328-336.
- Berezikov, E., Guryev, V., van de Belt, J., Wienholds, E., Plasterk, R.H., and Cuppen, E. 2005. Phylogenetic shadowing and computational identification of human microRNA genes. *Cell* **120**(1): 21-24.
- Bergeron, L., Perez, G.I., Macdonald, G., Shi, L., Sun, Y., Jurisicova, A., Varmuza, S., Latham, K.E., Flaws, J.A., Salter, J.C., Hara, H., Moskowitz, M.A., Li, E., Greenberg, A., Tilly, J.L., and Yuan, J. 1998. Defects in regulation of apoptosis in caspase-2-deficient mice. *Genes Dev* **12**(9): 1304-1314.
- Bernstein, E., Caudy, A.A., Hammond, S.M., and Hannon, G.J. 2001. Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* **409**(6818): 363-366.
- Bernstein, E., Kim, S.Y., Carmell, M.A., Murchison, E.P., Alcorn, H., Li, M.Z., Mills, A.A., Elledge, S.J., Anderson, K.V., and Hannon, G.J. 2003. Dicer is essential for mouse development. *Nat Genet* **35**(3): 215-217.
- Bonifacino, J.S., McCarthy, S.A., Maguire, J.E., Nakayama, T., Singer, D.S., Klausner, R.D., and Singer, A. 1990. Novel post-translational regulation of TCR expression in CD4+CD8+ thymocytes influenced by CD4. *Nature* **344**(6263): 247-251.
- Bosnali, M., Munst, B., Thier, M., and Edenhofer, F. 2009. Deciphering the stem cell machinery as a basis for understanding the molecular mechanism underlying reprogramming. *Cell Mol Life Sci* **66**(21): 3403-3420.
- Boyer, L.A., Lee, T.I., Cole, M.F., Johnstone, S.E., Levine, S.S., Zucker, J.P., Guenther, M.G., Kumar, R.M., Murray, H.L., Jenner, R.G., Gifford, D.K., Melton, D.A., Jaenisch, R., and Young, R.A. 2005. Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* **122**(6): 947-956.
- Brennecke, J., Hipfner, D.R., Stark, A., Russell, R.B., and Cohen, S.M. 2003. bantam encodes a developmentally regulated microRNA that controls cell proliferation and regulates the proapoptotic gene hid in Drosophila. *Cell* **113**(1): 25-36.
- Brennecke, J., Stark, A., and Cohen, S.M. 2005a. Not miR-ly muscular: microRNAs and muscle development. *Genes Dev* **19**(19): 2261-2264.
- Brennecke, J., Stark, A., Russell, R.B., and Cohen, S.M. 2005b. Principles of microRNA-target recognition. *PLoS Biol* **3**(3): e85.
- Cai, X., Hagedorn, C.H., and Cullen, B.R. 2004. Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. *Rna* **10**(12): 1957-1966.
- Calabrese, J.M., Seila, A.C., Yeo, G.W., and Sharp, P.A. 2007. RNA sequence analysis defines Dicer's role in mouse embryonic stem cells. *Proc Natl Acad Sci U S A*

- 104**(46): 18097-18102.
- Calabrese, J.M. and Sharp, P.A. 2006. Characterization of the short RNAs bound by the P19 suppressor of RNA silencing in mouse embryonic stem cells. *RNA* **12**(12): 2092-2102.
- Calin, G.A. and Croce, C.M. 2006. MicroRNA signatures in human cancers. *Nat Rev Cancer* **6**(11): 857-866.
- Calin, G.A., Dumitru, C.D., Shimizu, M., Bichi, R., Zupo, S., Noch, E., Aldler, H., Rattan, S., Keating, M., Rai, K., Rassenti, L., Kipps, T., Negrini, M., Bullrich, F., and Croce, C.M. 2002. Frequent deletions and down-regulation of micro- RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. *Proc Natl Acad Sci U S A* **99**(24): 15524-15529.
- Calin, G.A., Sevignani, C., Dumitru, C.D., Hyslop, T., Noch, E., Yendamuri, S., Shimizu, M., Rattan, S., Bullrich, F., Negrini, M., and Croce, C.M. 2004. Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. *Proc Natl Acad Sci U S A* **101**(9): 2999-3004.
- Card, D.A., Hebbar, P.B., Li, L., Trotter, K.W., Komatsu, Y., Mishina, Y., and Archer, T.K. 2008. Oct4/Sox2-regulated miR-302 targets cyclin D1 in human embryonic stem cells. *Mol Cell Biol* **28**(20): 6426-6438.
- Carthew, R.W. and Sontheimer, E.J. 2009. Origins and Mechanisms of miRNAs and siRNAs. *Cell* **136**(4): 642-655.
- Cattanach, B.M., Beechey, C.V., and Peters, J. 2004. Interactions between imprinting effects in the mouse. *Genetics* **168**(1): 397-413.
- Cervantes, R.B., Stringer, J.R., Shao, C., Tischfield, J.A., and Stambrook, P.J. 2002. Embryonic stem cells and somatic cells differ in mutation frequency and type. *Proc Natl Acad Sci U S A* **99**(6): 3586-3590.
- Chambers, I., Colby, D., Robertson, M., Nichols, J., Lee, S., Tweedie, S., and Smith, A. 2003. Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. *Cell* **113**(5): 643-655.
- Chen, C.Z., Li, L., Lodish, H.F., and Bartel, D.P. 2004. MicroRNAs modulate hematopoietic lineage differentiation. *Science* **303**(5654): 83-86.
- Chen, K. and Rajewsky, N. 2006. Natural selection on human microRNA binding sites inferred from SNP data. *Nat Genet* **38**(12): 1452-1456.
- Chendrimada, T.P., Finn, K.J., Ji, X., Baillat, D., Gregory, R.I., Liebhaber, S.A., Pasquinelli, A.E., and Shiekhattar, R. 2007. MicroRNA silencing through RISC recruitment of eIF6. *Nature* **447**(7146): 823-828.
- Chendrimada, T.P., Gregory, R.I., Kumaraswamy, E., Norman, J., Cooch, N., Nishikura, K., and Shiekhattar, R. 2005. TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. *Nature* **436**(7051): 740-744.
- Chiang, H.R., Schoenfeld, L.W., Ruby, J.G., Auyeung, V.C., Spies, N., Baek, D., Johnston, W.K., Russ, C., Luo, S., Babiarz, J.E., Blelloch, R., Schroth, G.P., Nusbaum, C., and Bartel, D.P. Mammalian microRNAs: experimental evaluation of novel and previously annotated genes. *Genes Dev.*
- Chin, L.J., Ratner, E., Leng, S., Zhai, R., Nallur, S., Babar, I., Muller, R.U., Straka, E., Su, L., Burki, E.A., Crowell, R.E., Patel, R., Kulkarni, T., Homer, R., Zelterman, D., Kidd, K.K., Zhu, Y., Christiani, D.C., Belinsky, S.A., Slack, F.J., and Weidhaas, J.B. 2008. A SNP in a let-7 microRNA complementary site in the KRAS 3' untranslated region increases non-small cell lung cancer risk. *Cancer Res* **68**(20): 8535-8540.
- Chu, C.Y. and Rana, T.M. 2006. Translation repression in human cells by microRNA-induced gene silencing requires RCK/p54. *PLoS Biol* **4**(7): e210.
- Ciudo, C., Servant, N., Cognat, V., Sarazin, A., Kieffer, E., Vville, S., Colot, V., Barillot,

- E., Heard, E., and Voinnet, O. 2009. Highly dynamic and sex-specific expression of microRNAs during early ES cell differentiation. *PLoS Genet* **5**(8): e1000620.
- Cimmino, A., Calin, G.A., Fabbri, M., Iorio, M.V., Ferracin, M., Shimizu, M., Wojcik, S.E., Aqeilan, R.I., Zupo, S., Dono, M., Rassenti, L., Alder, H., Volinia, S., Liu, C.G., Kipps, T.J., Negrini, M., and Croce, C.M. 2005. miR-15 and miR-16 induce apoptosis by targeting BCL2. *Proc Natl Acad Sci U S A*.
- Clop, A., Marcq, F., Takeda, H., Pirottin, D., Tordoir, X., Bibe, B., Bouix, J., Caiment, F., Elsen, J.M., Eychenne, F., Larzul, C., Laville, E., Meish, F., Milenkovic, D., Tobin, J., Charlier, C., and Georges, M. 2006. A mutation creating a potential illegitimate microRNA target site in the myostatin gene affects muscularity in sheep. *Nat Genet* **38**(7): 813-818.
- Cobb, B.S., Nesterova, T.B., Thompson, E., Hertweck, A., O'Connor, E., Godwin, J., Wilson, C.B., Brockdorff, N., Fisher, A.G., Smale, S.T., and Merckenschlager, M. 2005. T cell lineage choice and differentiation in the absence of the RNase III enzyme Dicer. *J Exp Med* **201**(9): 1367-1373.
- Core, L.J., Waterfall, J.J., and Lis, J.T. 2008. Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science* **322**(5909): 1845-1848.
- Costinean, S., Zanesi, N., Pekarsky, Y., Tili, E., Volinia, S., Heerema, N., and Croce, C.M. 2006. Pre-B cell proliferation and lymphoblastic leukemia/high-grade lymphoma in E(mu)-miR155 transgenic mice. *Proc Natl Acad Sci U S A* **103**(18): 7024-7029.
- Croce, C.M. 2009. Causes and consequences of microRNA dysregulation in cancer. *Nat Rev Genet* **10**(10): 704-714.
- Croce, C.M. and Calin, G.A. 2005. miRNAs, cancer, and stem cell division. *Cell* **122**(1): 6-7.
- Cross, J.C., Werb, Z., and Fisher, S.J. 1994. Implantation and the placenta: key pieces of the development puzzle. *Science* **266**(5190): 1508-1518.
- Davis, E., Caiment, F., Tordoir, X., Cavaille, J., Ferguson-Smith, A., Cockett, N., Georges, M., and Charlier, C. 2005. RNAi-mediated allelic trans-interaction at the imprinted Rtl1/Peg11 locus. *Curr Biol* **15**(8): 743-749.
- Diederichs, S. and Haber, D.A. 2007. Dual role for argonautes in microRNA processing and posttranscriptional regulation of microRNA expression. *Cell* **131**(6): 1097-1108.
- Ding, X.C. and Grosshans, H. 2009. Repression of *C. elegans* microRNA targets at the initiation level of translation requires GW182 proteins. *EMBO J* **28**(3): 213-222.
- Djuranovic, S., Zinchenko, M.K., Hur, J.K., Nahvi, A., Brunelle, J.L., Rogers, E.J., and Green, R. 2010. Allosteric regulation of Argonaute proteins by miRNAs. *Nat Struct Mol Biol* **17**(2): 144-150.
- Doench, J.G., Petersen, C.P., and Sharp, P.A. 2003. siRNAs can function as miRNAs. *Genes Dev* **17**(4): 438-442.
- Doench, J.G. and Sharp, P.A. 2004. Specificity of microRNA target selection in translational repression. *Genes Dev* **18**(5): 504-511.
- Du, Z., Lee, J.K., Tjhen, R., Stroud, R.M., and James, T.L. 2008. Structural and biochemical insights into the dicing mechanism of mouse Dicer: a conserved lysine is critical for dsRNA cleavage. *Proc Natl Acad Sci U S A* **105**(7): 2391-2396.
- Duursma, A.M., Kedde, M., Schrier, M., le Sage, C., and Agami, R. 2008. miR-148 targets human DNMT3b protein coding region. *Rna* **14**(5): 872-877.
- Ebert, M.S., Neilson, J.R., and Sharp, P.A. 2007. MicroRNA sponges: competitive inhibitors of small RNAs in mammalian cells. *Nat Methods* **4**(9): 721-726.

- Elbashir, S.M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K., and Tuschl, T. 2001. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* **411**(6836): 494-498.
- Esquela-Kerscher, A. and Slack, F.J. 2006. Oncomirs - microRNAs with a role in cancer. *Nat Rev Cancer* **6**(4): 259-269.
- Eulalio, A., Behm-Ansmant, I., and Izaurralde, E. 2007a. P bodies: at the crossroads of post-transcriptional pathways. *Nat Rev Mol Cell Biol* **8**(1): 9-22.
- Eulalio, A., Behm-Ansmant, I., Schweizer, D., and Izaurralde, E. 2007b. P-body formation is a consequence, not the cause, of RNA-mediated gene silencing. *Mol Cell Biol* **27**(11): 3970-3981.
- Eulalio, A., Helms, S., Fritsch, C., Fauser, M., and Izaurralde, E. 2009a. A C-terminal silencing domain in GW182 is essential for miRNA function. *Rna* **15**(6): 1067-1077.
- Eulalio, A., Huntzinger, E., and Izaurralde, E. 2008. GW182 interaction with Argonaute is essential for miRNA-mediated translational repression and mRNA decay. *Nat Struct Mol Biol* **15**(4): 346-353.
- Eulalio, A., Huntzinger, E., Nishihara, T., Rehwinkel, J., Fauser, M., and Izaurralde, E. 2009b. Deadenylation is a widespread effect of miRNA regulation. *Rna* **15**(1): 21-32.
- Eulalio, A., Rehwinkel, J., Stricker, M., Huntzinger, E., Yang, S.F., Doerks, T., Dorner, S., Bork, P., Boutros, M., and Izaurralde, E. 2007c. Target-specific requirements for enhancers of decapping in miRNA-mediated gene silencing. *Genes Dev* **21**(20): 2558-2570.
- Eulalio, A., Triteschler, F., Buttner, R., Weichenrieder, O., Izaurralde, E., and Truffault, V. 2009c. The RRM domain in GW182 proteins contributes to miRNA-mediated gene silencing. *Nucleic Acids Res* **37**(9): 2974-2983.
- Faast, R., White, J., Cartwright, P., Crocker, L., Sarcevic, B., and Dalton, S. 2004. Cdk6-cyclin D3 activity in murine ES cells is resistant to inhibition by p16(INK4a). *Oncogene* **23**(2): 491-502.
- Farh, K.K., Grimson, A., Jan, C., Lewis, B.P., Johnston, W.K., Lim, L.P., Burge, C.B., and Bartel, D.P. 2005. The widespread impact of mammalian MicroRNAs on mRNA repression and evolution. *Science* **310**(5755): 1817-1821.
- Fazi, F., Rosa, A., Fatica, A., Gelmetti, V., De Marchis, M.L., Nervi, C., and Bozzoni, I. 2005. A minicircuitry comprised of microRNA-223 and transcription factors NFI-A and C/EBPalpha regulates human granulopoiesis. *Cell* **123**(5): 819-831.
- Filipowicz, W., Bhattacharyya, S.N., and Sonenberg, N. 2008. Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? *Nat Rev Genet* **9**(2): 102-114.
- Friedman, R.C., Farh, K.K., Burge, C.B., and Bartel, D.P. 2009. Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res* **19**(1): 92-105.
- Fujita, J., Crane, A.M., Souza, M.K., Dejosez, M., Kyba, M., Flavell, R.A., Thomson, J.A., and Zwaka, T.P. 2008. Caspase activity mediates the differentiation of embryonic stem cells. *Cell Stem Cell* **2**(6): 595-601.
- Giraldez, A.J., Mishima, Y., Rihel, J., Grocock, R.J., Van Dongen, S., Inoue, K., Enright, A.J., and Schier, A.F. 2006. Zebrafish MiR-430 promotes deadenylation and clearance of maternal mRNAs. *Science* **312**(5770): 75-79.
- Girard, A., Sachidanandam, R., Hannon, G.J., and Carmell, M.A. 2006. A germline-specific class of small RNAs binds mammalian Piwi proteins. *Nature*.
- Gratiot-Deans, J., Ding, L., Turka, L.A., and Nunez, G. 1993. bcl-2 proto-oncogene expression during human T cell development. Evidence for biphasic regulation. *J Immunol* **151**(1): 83-91.

- Gregory, R.I., Chendrimada, T.P., Cooch, N., and Shiekhattar, R. 2005. Human RISC couples microRNA biogenesis and posttranscriptional gene silencing. *Cell* **123**(4): 631-640.
- Griffiths-Jones, S. 2006. miRBase: the microRNA sequence database. *Methods Mol Biol* **342**: 129-138.
- Griffiths-Jones, S., Moxon, S., Marshall, M., Khanna, A., Eddy, S.R., and Bateman, A. 2005. Rfam: annotating non-coding RNAs in complete genomes. *Nucleic Acids Res* **33**(Database issue): D121-124.
- Griffiths-Jones, S., Saini, H.K., van Dongen, S., and Enright, A.J. 2008. miRBase: tools for microRNA genomics. *Nucleic Acids Res* **36**(Database issue): D154-158.
- Grimson, A., Farh, K.K., Johnston, W.K., Garrett-Engele, P., Lim, L.P., and Bartel, D.P. 2007. MicroRNA targeting specificity in mammals: determinants beyond seed pairing. *Mol Cell* **27**(1): 91-105.
- Grishok, A., Pasquinelli, A.E., Conte, D., Li, N., Parrish, S., Ha, I., Baillie, D.L., Fire, A., Ruvkun, G., and Mello, C.C. 2001. Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control *C. elegans* developmental timing. *Cell* **106**(1): 23-34.
- Grivna, S.T., Beyret, E., Wang, Z., and Lin, H. 2006. A novel class of small RNAs in mouse spermatogenic cells. *Genes Dev* **20**(13): 1709-1714.
- Gruber, J.J., Zatechka, D.S., Sabin, L.R., Yong, J., Lum, J.J., Kong, M., Zong, W.X., Zhang, Z., Lau, C.K., Rawlings, J., Cherry, S., Ihle, J.N., Dreyfuss, G., and Thompson, C.B. 2009. Ars2 links the nuclear cap-binding complex to RNA interference and cell proliferation. *Cell* **138**(2): 328-339.
- Gu, Z., Flemington, C., Chittenden, T., and Zambetti, G.P. 2000. ei24, a p53 response gene involved in growth suppression and apoptosis. *Mol Cell Biol* **20**(1): 233-241.
- Haase, A.D., Jaskiewicz, L., Zhang, H., Laine, S., Sack, R., Gagnon, A., and Filipowicz, W. 2005. TRBP, a regulator of cellular PKR and HIV-1 virus expression, interacts with Dicer and functions in RNA silencing. *EMBO Rep* **6**(10): 961-967.
- Hagan, J.P., Piskounova, E., and Gregory, R.I. 2009. Lin28 recruits the TUTase Zcchc11 to inhibit let-7 maturation in mouse embryonic stem cells. *Nat Struct Mol Biol* **16**(10): 1021-1025.
- Hammond, S.M. 2006. MicroRNAs as oncogenes. *Curr Opin Genet Dev* **16**(1): 4-9.
- Han, J., Lee, Y., Yeom, K.H., Nam, J.W., Heo, I., Rhee, J.K., Sohn, S.Y., Cho, Y., Zhang, B.T., and Kim, V.N. 2006. Molecular basis for the recognition of primary microRNAs by the Drosha-DGCR8 complex. *Cell* **125**(5): 887-901.
- Hara, T., Jung, L.K., Bjorndahl, J.M., and Fu, S.M. 1986. Human T cell activation. III. Rapid induction of a phosphorylated 28 kD/32 kD disulfide-linked early activation antigen (EA 1) by 12-o-tetradecanoyl phorbol-13-acetate, mitogens, and antigens. *J Exp Med* **164**(6): 1988-2005.
- Harfe, B.D., McManus, M.T., Mansfield, J.H., Hornstein, E., and Tabin, C.J. 2005. The RNaseIII enzyme Dicer is required for morphogenesis but not patterning of the vertebrate limb. *Proc Natl Acad Sci U S A* **102**(31): 10898-10903.
- Harris, K.S., Zhang, Z., McManus, M.T., Harfe, B.D., and Sun, X. 2006. Dicer function is essential for lung epithelium morphogenesis. *Proc Natl Acad Sci U S A* **103**(7): 2208-2213.
- He, L., He, X., Lim, L.P., de Stanchina, E., Xuan, Z., Liang, Y., Xue, W., Zender, L., Magnus, J., Ridzon, D., Jackson, A.L., Linsley, P.S., Chen, C., Lowe, S.W., Cleary, M.A., and Hannon, G.J. 2007. A microRNA component of the p53 tumour suppressor network. *Nature* **447**(7148): 1130-1134.
- Hennet, T., Hagen, F.K., Tabak, L.A., and Marth, J.D. 1995. T-cell-specific deletion of a polypeptide N-acetylgalactosaminyl-transferase gene by site-directed

- recombination. *Proc Natl Acad Sci U S A* **92**(26): 12070-12074.
- Heo, I., Joo, C., Cho, J., Ha, M., Han, J., and Kim, V.N. 2008. Lin28 mediates the terminal uridylation of let-7 precursor MicroRNA. *Mol Cell* **32**(2): 276-284.
- Hertel, J., Lindemeyer, M., Missal, K., Fried, C., Tanzer, A., Flamm, C., Hofacker, I.L., and Stadler, P.F. 2006. The expansion of the metazoan microRNA repertoire. *BMC Genomics* **7**: 25.
- Hoffmann, R., Bruno, L., Seidl, T., Rolink, A., and Melchers, F. 2003. Rules for gene usage inferred from a comparison of large-scale gene expression profiles of T and B lymphocyte development. *J Immunol* **170**(3): 1339-1353.
- Hornstein, E. and Shomron, N. 2006. Canalization of development by microRNAs. *Nat Genet* **38 Suppl**: S20-24.
- Houbaviy, H.B., Dennis, L., Jaenisch, R., and Sharp, P.A. 2005. Characterization of a highly variable eutherian microRNA gene. *Rna* **11**(8): 1245-1257.
- Houbaviy, H.B., Murray, M.F., and Sharp, P.A. 2003. Embryonic stem cell-specific MicroRNAs. *Dev Cell* **5**(2): 351-358.
- Humphreys, D.T., Westman, B.J., Martin, D.I., and Preiss, T. 2005. MicroRNAs control translation initiation by inhibiting eukaryotic initiation factor 4E/cap and poly(A) tail function. *Proc Natl Acad Sci U S A* **102**(47): 16961-16966.
- Hutvagner, G., McLachlan, J., Pasquinelli, A.E., Balint, E., Tuschl, T., and Zamore, P.D. 2001. A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. *Science* **293**(5531): 834-838.
- Hutvagner, G., Simard, M.J., Mello, C.C., and Zamore, P.D. 2004. Sequence-specific inhibition of small RNA function. *PLoS Biol* **2**(4): E98.
- Hwang, H.W., Wentzel, E.A., and Mendell, J.T. 2007. A hexanucleotide element directs microRNA nuclear import. *Science* **315**(5808): 97-100.
- Jaenisch, R. 1997. DNA methylation and imprinting: why bother? *Trends Genet* **13**(8): 323-329.
- . 2008. Unpublished work. In.
- Jakymiw, A., Lian, S., Eystathioy, T., Li, S., Satoh, M., Hamel, J.C., Fritzler, M.J., and Chan, E.K. 2005. Disruption of GW bodies impairs mammalian RNA interference. *Nat Cell Biol* **7**(12): 1267-1274.
- Jin, Z. and El-Deiry, W.S. 2005. Overview of cell death signaling pathways. *Cancer Biol Ther* **4**(2): 139-163.
- Johnson, S.M., Grosshans, H., Shingara, J., Byrom, M., Jarvis, R., Cheng, A., Labourier, E., Reinert, K.L., Brown, D., and Slack, F.J. 2005. RAS is regulated by the let-7 microRNA family. *Cell* **120**(5): 635-647.
- Johnston, R.J. and Hobert, O. 2003. A microRNA controlling left/right neuronal asymmetry in *Caenorhabditis elegans*. *Nature* **426**(6968): 845-849.
- Judson, R.L., Babiarz, J.E., Venere, M., and Blalock, R. 2009. Embryonic stem cell-specific microRNAs promote induced pluripotency. *Nat Biotechnol* **27**(5): 459-461.
- Kaneko-Ishino, T., Kohda, T., and Ishino, F. 2003. The regulation and biological significance of genomic imprinting in mammals. *J Biochem* **133**(6): 699-711.
- Kanellopoulou, C., Muljo, S.A., Kung, A.L., Ganesan, S., Drapkin, R., Jenuwein, T., Livingston, D.M., and Rajewsky, K. 2005. Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. *Genes Dev* **19**(4): 489-501.
- Kawamata, T., Seitz, H., and Tomari, Y. 2009. Structural determinants of miRNAs for RISC loading and slicer-independent unwinding. *Nat Struct Mol Biol* **16**(9): 953-960.
- Ketting, R.F., Fischer, S.E., Bernstein, E., Sijen, T., Hannon, G.J., and Plasterk, R.H.

2001. Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in *C. elegans*. *Genes Dev* **15**(20): 2654-2659.
- Khvorova, A., Reynolds, A., and Jayasena, S.D. 2003. Functional siRNAs and miRNAs exhibit strand bias. *Cell* **115**(2): 209-216.
- Kim, J., Krichevsky, A., Grad, Y., Hayes, G.D., Kosik, K.S., Church, G.M., and Ruvkun, G. 2004. Identification of many microRNAs that copurify with polyribosomes in mammalian neurons. *Proc Natl Acad Sci U S A* **101**(1): 360-365.
- Kim, V.N., Han, J., and Siomi, M.C. 2009. Biogenesis of small RNAs in animals. *Nat Rev Mol Cell Biol* **10**(2): 126-139.
- Kim, V.N. and Nam, J.W. 2006. Genomics of microRNA. *Trends Genet* **22**(3): 165-173.
- Kim, Y.K. and Kim, V.N. 2007. Processing of intronic microRNAs. *Embo J* **26**(3): 775-783.
- Kiriakidou, M., Tan, G.S., Lamprinak, S., De Planell-Saguer, M., Nelson, P.T., and Mourelatos, Z. 2007. An mRNA m7G cap binding-like motif within human Ago2 represses translation. *Cell* **129**(6): 1141-1151.
- Kloosterman, W.P. and Plasterk, R.H. 2006. The diverse functions of microRNAs in animal development and disease. *Dev Cell* **11**(4): 441-450.
- Knight, S.W. and Bass, B.L. 2001. A role for the RNase III enzyme DCR-1 in RNA interference and germ line development in *Caenorhabditis elegans*. *Science* **293**(5538): 2269-2271.
- Knox, K. and Baker, J.C. 2008. Genomic evolution of the placenta using co-option and duplication and divergence. *Genome Res* **18**(5): 695-705.
- Krumschnabel, G., Sohm, B., Bock, F., Manzl, C., and Villunger, A. 2009. The enigma of caspase-2: the laymen's view. *Cell Death Differ* **16**(2): 195-207.
- Krutzfeldt, J., Rajewsky, N., Braich, R., Rajeev, K.G., Tuschl, T., Manoharan, M., and Stoffel, M. 2005. Silencing of microRNAs in vivo with 'antagomirs'. *Nature* **438**(7068): 685-689.
- Kumar, M.S., Lu, J., Mercer, K.L., Golub, T.R., and Jacks, T. 2007. Impaired microRNA processing enhances cellular transformation and tumorigenesis. *Nat Genet* **39**(5): 673-677.
- Kuzmin, A., Han, Z., Golding, M.C., Mann, M.R., Latham, K.E., and Varmuza, S. 2008. The PcG gene *Sfmbt2* is paternally expressed in extraembryonic tissues. *Gene Expr Patterns* **8**(2): 107-116.
- Lagos-Quintana, M., Rauhut, R., Lendeckel, W., and Tuschl, T. 2001. Identification of novel genes coding for small expressed RNAs. *Science* **294**(5543): 853-858.
- Lagos-Quintana, M., Rauhut, R., Yalcin, A., Meyer, J., Lendeckel, W., and Tuschl, T. 2002. Identification of tissue-specific microRNAs from mouse. *Curr Biol* **12**(9): 735-739.
- Lai, E.C. 2002. Micro RNAs are complementary to 3' UTR sequence motifs that mediate negative post-transcriptional regulation. *Nat Genet* **30**(4): 363-364.
- Landgraf, P., Rusu, M., Sheridan, R., Sewer, A., Iovino, N., Aravin, A., Pfeffer, S., Rice, A., Kamphorst, A.O., Landthaler, M., Lin, C., Socci, N.D., Hermida, L., Fulci, V., Chiaretti, S., Foa, R., Schliwka, J., Fuchs, U., Novosel, A., Muller, R.U., Schermer, B., Bissels, U., Inman, J., Phan, Q., Chien, M., Weir, D.B., Choksi, R., De Vita, G., Frezzetti, D., Trompeter, H.I., Hornung, V., Teng, G., Hartmann, G., Palkovits, M., Di Lauro, R., Wernet, P., Macino, G., Rogler, C.E., Nagle, J.W., Ju, J., Papavasiliou, F.N., Benzing, T., Lichter, P., Tam, W., Brownstein, M.J., Bosio, A., Borkhardt, A., Russo, J.J., Sander, C., Zavolan, M., and Tuschl, T. 2007. A mammalian microRNA expression atlas based on small RNA library sequencing. *Cell* **129**(7): 1401-1414.
- Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam,

- H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J., and Higgins, D.G. 2007. Clustal W and Clustal X version 2.0. *Bioinformatics* **23**(21): 2947-2948.
- Larsen, B.D., Rampalli, S., Burns, L.E., Brunette, S., Dilworth, F.J., and Megeney, L.A. Caspase 3/caspase-activated DNase promote cell differentiation by inducing DNA strand breaks. *Proc Natl Acad Sci U S A* **107**(9): 4230-4235.
- Lau, N.C., Lim, L.P., Weinstein, E.G., and Bartel, D.P. 2001. An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. *Science* **294**(5543): 858-862.
- Lau, N.C., Robine, N., Martin, R., Chung, W.J., Niki, Y., Berezikov, E., and Lai, E.C. 2009. Abundant primary piRNAs, endo-siRNAs, and microRNAs in a *Drosophila* ovary cell line. *Genome Res* **19**(10): 1776-1785.
- Lau, N.C., Seto, A.G., Kim, J., Kuramochi-Miyagawa, S., Nakano, T., Bartel, D.P., and Kingston, R.E. 2006. Characterization of the piRNA complex from rat testes. *Science* **313**(5785): 363-367.
- Lee, C.T., Risom, T., and Strauss, W.M. 2007. Evolutionary conservation of microRNA regulatory circuits: an examination of microRNA gene complexity and conserved microRNA-target interactions through metazoan phylogeny. *DNA Cell Biol* **26**(4): 209-218.
- Lee, E.Y., Yuan, T.L., Danielian, P.S., West, J.C., and Lees, J.A. 2009. E2F4 cooperates with pRB in the development of extra-embryonic tissues. *Dev Biol* **332**(1): 104-115.
- Lee, N.S., Kim, J.S., Cho, W.J., Lee, M.R., Steiner, R., Gompers, A., Ling, D., Zhang, J., Strom, P., Behlke, M., Moon, S.H., Salvaterra, P.M., Jove, R., and Kim, K.S. 2008. miR-302b maintains "stemness" of human embryonal carcinoma cells by post-transcriptional regulation of Cyclin D2 expression. *Biochem Biophys Res Commun* **377**(2): 434-440.
- Lee, R.C., Feinbaum, R.L., and 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., Lee, J., Provost, P., Radmark, O., Kim, S., and Kim, V.N. 2003. The nuclear RNase III Drosha initiates microRNA processing. *Nature* **425**(6956): 415-419.
- Lee, Y., Hur, I., Park, S.Y., Kim, Y.K., Suh, M.R., and Kim, V.N. 2006. The role of PACT in the RNA silencing pathway. *Embo J* **25**(3): 522-532.
- Lee, Y., Jeon, K., Lee, J.T., Kim, S., and Kim, V.N. 2002. MicroRNA maturation: stepwise processing and subcellular localization. *Embo J* **21**(17): 4663-4670.
- Lee, Y., Kim, M., Han, J., Yeom, K.H., Lee, S., Baek, S.H., and Kim, V.N. 2004. MicroRNA genes are transcribed by RNA polymerase II. *Embo J* **23**(20): 4051-4060.
- Leung, A.K., Calabrese, J.M., and Sharp, P.A. 2006. Quantitative analysis of Argonaute protein reveals microRNA-dependent localization to stress granules. *Proc Natl Acad Sci U S A* **103**(48): 18125-18130.
- Leung, A.K., young, A.G., Bhukar, A.J., Zheng, G.X.Y., Boson, A., and Sharp, P.A. 2010. Genome-wide identification of endogenous miRNA-binding sites at 3' untranslated and coding regions. In *submitted*.
- Leuschner, P.J., Ameres, S.L., Kueng, S., and Martinez, J. 2006. Cleavage of the siRNA passenger strand during RISC assembly in human cells. *EMBO Rep* **7**(3): 314-320.
- Lewis, B.P., Burge, C.B., and Bartel, D.P. 2005. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets.

- Cell* **120**(1): 15-20.
- Lewis, B.P., Shih, I.H., Jones-Rhoades, M.W., Bartel, D.P., and Burge, C.B. 2003. Prediction of mammalian microRNA targets. *Cell* **115**(7): 787-798.
- Li, J. and Yuan, J. 2008. Caspases in apoptosis and beyond. *Oncogene* **27**(48): 6194-6206.
- Li, Q.J., Chau, J., Ebert, P.J., Sylvester, G., Min, H., Liu, G., Braich, R., Manoharan, M., Soutschek, J., Skare, P., Klein, L.O., Davis, M.M., and Chen, C.Z. 2007. miR-181a is an intrinsic modulator of T cell sensitivity and selection. *Cell* **129**(1): 147-161.
- Li, X., Cassidy, J.J., Reinke, C.A., Fischboeck, S., and Carthew, R.W. 2009. A microRNA imparts robustness against environmental fluctuation during development. *Cell* **137**(2): 273-282.
- Lim, L.P., Glasner, M.E., Yekta, S., Burge, C.B., and Bartel, D.P. 2003a. Vertebrate microRNA genes. *Science* **299**(5612): 1540.
- Lim, L.P., Lau, N.C., Garrett-Engle, P., Grimson, A., Schelter, J.M., Castle, J., Bartel, D.P., Linsley, P.S., and Johnson, J.M. 2005. Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. *Nature* **433**(7027): 769-773.
- Lim, L.P., Lau, N.C., Weinstein, E.G., Abdelhakim, A., Yekta, S., Rhoades, M.W., Burge, C.B., and Bartel, D.P. 2003b. The microRNAs of *Caenorhabditis elegans*. *Genes Dev* **17**(8): 991-1008.
- Lin, H. 2002. The stem-cell niche theory: lessons from flies. *Nat Rev Genet* **3**(12): 931-940.
- Lin, S.L., Chang, D.C., Chang-Lin, S., Lin, C.H., Wu, D.T., Chen, D.T., and Ying, S.Y. 2008. Mir-302 reprograms human skin cancer cells into a pluripotent ES-cell-like state. *RNA* **14**(10): 2115-2124.
- Lin, T., Chao, C., Saito, S., Mazur, S.J., Murphy, M.E., Appella, E., and Xu, Y. 2005. p53 induces differentiation of mouse embryonic stem cells by suppressing Nanog expression. *Nat Cell Biol* **7**(2): 165-171.
- Lingel, A., Simon, B., Izaurralde, E., and Sattler, M. 2004. Nucleic acid 3'-end recognition by the Argonaute2 PAZ domain. *Nat Struct Mol Biol* **11**(6): 576-577.
- Liu, C., Bai, B., Skogerbo, G., Cai, L., Deng, W., Zhang, Y., Bu, D., Zhao, Y., and Chen, R. 2005a. NONCODE: an integrated knowledge database of non-coding RNAs. *Nucleic Acids Res* **33**(Database issue): D112-115.
- Liu, J., Carmell, M.A., Rivas, F.V., Marsden, C.G., Thomson, J.M., Song, J.J., Hammond, S.M., Joshua-Tor, L., and Hannon, G.J. 2004. Argonaute2 is the catalytic engine of mammalian RNAi. *Science* **305**(5689): 1437-1441.
- Liu, J., Rivas, F.V., Wohlschlegel, J., Yates, J.R., 3rd, Parker, R., and Hannon, G.J. 2005b. A role for the P-body component GW182 in microRNA function. *Nat Cell Biol* **7**(12): 1161-1166.
- Liu, N., Okamura, K., Tyler, D.M., Phillips, M.D., Chung, W.J., and Lai, E.C. 2008. The evolution and functional diversification of animal microRNA genes. *Cell Res* **18**(10): 985-996.
- Loh, Y.H., Wu, Q., Chew, J.L., Vega, V.B., Zhang, W., Chen, X., Bourque, G., George, J., Leong, B., Liu, J., Wong, K.Y., Sung, K.W., Lee, C.W., Zhao, X.D., Chiu, K.P., Lipovich, L., Kuznetsov, V.A., Robson, P., Stanton, L.W., Wei, C.L., Ruan, Y., Lim, B., and Ng, H.H. 2006. The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. *Nat Genet* **38**(4): 431-440.
- Lu, J., Getz, G., Miska, E.A., Alvarez-Saavedra, E., Lamb, J., Peck, D., Sweet-Cordero, A., Ebert, B.L., Mak, R.H., Ferrando, A.A., Downing, J.R., Jacks, T., Horvitz, H.R., and Golub, T.R. 2005. MicroRNA expression profiles classify human

- cancers. *Nature* **435**(7043): 834-838.
- Lujambio, A., Calin, G.A., Villanueva, A., Ropero, S., Sanchez-Cespedes, M., Blanco, D., Montuenga, L.M., Rossi, S., Nicoloso, M.S., Faller, W.J., Gallagher, W.M., Eccles, S.A., Croce, C.M., and Esteller, M. 2008. A microRNA DNA methylation signature for human cancer metastasis. *Proc Natl Acad Sci U S A* **105**(36): 13556-13561.
- Lund, E., Guttinger, S., Calado, A., Dahlberg, J.E., and Kutay, U. 2004. Nuclear export of microRNA precursors. *Science* **303**(5654): 95-98.
- Lytle, J.R., Yario, T.A., and Steitz, J.A. 2007. Target mRNAs are repressed as efficiently by microRNA-binding sites in the 5' UTR as in the 3' UTR. *Proc Natl Acad Sci U S A* **104**(23): 9667-9672.
- Ma, J.B., Ye, K., and Patel, D.J. 2004. Structural basis for overhang-specific small interfering RNA recognition by the PAZ domain. *Nature* **429**(6989): 318-322.
- Ma, J.B., Yuan, Y.R., Meister, G., Pei, Y., Tuschl, T., and Patel, D.J. 2005. Structural basis for 5'-end-specific recognition of guide RNA by the *A. fulgidus* Piwi protein. *Nature* **434**(7033): 666-670.
- Maccani, M.A. and Marsit, C.J. 2009. Epigenetics in the placenta. *Am J Reprod Immunol* **62**(2): 78-89.
- MacRae, I.J., Ma, E., Zhou, M., Robinson, C.V., and Doudna, J.A. 2008. In vitro reconstitution of the human RISC-loading complex. *Proc Natl Acad Sci U S A* **105**(2): 512-517.
- Maguire, J.E., McCarthy, S.A., Singer, A., and Singer, D.S. 1990. Inverse correlation between steady-state RNA and cell surface T cell receptor levels. *Faseb J* **4**(13): 3131-3134.
- Mallory, A.C., Reinhart, B.J., Jones-Rhoades, M.W., Tang, G., Zamore, P.D., Barton, M.K., and Bartel, D.P. 2004. MicroRNA control of PHABULOSA in leaf development: importance of pairing to the microRNA 5' region. *EMBO J* **23**(16): 3356-3364.
- Maltepe, E., Bakardjiev, A.I., and Fisher, S.J. The placenta: transcriptional, epigenetic, and physiological integration during development. *J Clin Invest* **120**(4): 1016-1025.
- Maniataki, E. and Mourelatos, Z. 2005. A human, ATP-independent, RISC assembly machine fueled by pre-miRNA. *Genes Dev* **19**(24): 2979-2990.
- Maroney, P.A., Yu, Y., Fisher, J., and Nilsen, T.W. 2006. Evidence that microRNAs are associated with translating messenger RNAs in human cells. *Nat Struct Mol Biol* **13**(12): 1102-1107.
- Marson, A., Levine, S.S., Cole, M.F., Frampton, G.M., Brambrink, T., Johnstone, S., Guenther, M.G., Johnston, W.K., Wernig, M., Newman, J., Calabrese, J.M., Dennis, L.M., Volkert, T.L., Gupta, S., Love, J., Hannett, N., Sharp, P.A., Bartel, D.P., Jaenisch, R., and Young, R.A. 2008. Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells. *Cell* **134**(3): 521-533.
- Matranga, C., Tomari, Y., Shin, C., Bartel, D.P., and Zamore, P.D. 2005. Passenger-strand cleavage facilitates assembly of siRNA into Ago2-containing RNAi enzyme complexes. *Cell* **123**(4): 607-620.
- Mayr, C. and Bartel, D.P. 2009. Widespread shortening of 3'UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells. *Cell* **138**(4): 673-684.
- Mayr, C., Hemann, M.T., and Bartel, D.P. 2007. Disrupting the pairing between let-7 and Hmga2 enhances oncogenic transformation. *Science* **315**(5818): 1576-1579.
- McGrath, J. and Solter, D. 1984. Completion of mouse embryogenesis requires both the

- maternal and paternal genomes. *Cell* **37**(1): 179-183.
- Meister, G., Landthaler, M., Dorsett, Y., and Tuschl, T. 2004a. Sequence-specific inhibition of microRNA- and siRNA-induced RNA silencing. *Rna* **10**(3): 544-550.
- Meister, G., Landthaler, M., Patkaniowska, A., Dorsett, Y., Teng, G., and Tuschl, T. 2004b. Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs. *Mol Cell* **15**(2): 185-197.
- Meister, G., Landthaler, M., Peters, L., Chen, P.Y., Urlaub, H., Luhrmann, R., and Tuschl, T. 2005. Identification of novel argonaute-associated proteins. *Curr Biol* **15**(23): 2149-2155.
- Melton, C., Judson, R.L., and Blueloch, R. 2010. Opposing microRNA families regulate self-renewal in mouse embryonic stem cells. *Nature* **463**(7281): 621-626.
- Michlewski, G., Guil, S., Semple, C.A., and Caceres, J.F. 2008. Posttranscriptional regulation of miRNAs harboring conserved terminal loops. *Mol Cell* **32**(3): 383-393.
- Minami, Y., Weissman, A.M., Samelson, L.E., and Klausner, R.D. 1987. Building a multichain receptor: synthesis, degradation, and assembly of the T-cell antigen receptor. *Proc Natl Acad Sci U S A* **84**(9): 2688-2692.
- Mineno, J., Okamoto, S., Ando, T., Sato, M., Chono, H., Izu, H., Takayama, M., Asada, K., Mirochnitchenko, O., Inouye, M., and Kato, I. 2006. The expression profile of microRNAs in mouse embryos. *Nucleic Acids Res* **34**(6): 1765-1771.
- Mitsui, K., Tokuzawa, Y., Itoh, H., Segawa, K., Murakami, M., Takahashi, K., Maruyama, M., Maeda, M., and Yamanaka, S. 2003. The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell* **113**(5): 631-642.
- Miyoshi, K., Tsukumo, H., Nagami, T., Siomi, H., and Siomi, M.C. 2005. Slicer function of Drosophila Argonautes and its involvement in RISC formation. *Genes Dev* **19**(23): 2837-2848.
- Monticelli, S., Ansel, K.M., Xiao, C., Socci, N.D., Krichevsky, A.M., Thai, T.H., Rajewsky, N., Marks, D.S., Sander, C., Rajewsky, K., Rao, A., and Kosik, K.S. 2005. MicroRNA profiling of the murine hematopoietic system. *Genome Biol* **6**(8): R71.
- Moore, T. and Haig, D. 1991. Genomic imprinting in mammalian development: a parental tug-of-war. *Trends Genet* **7**(2): 45-49.
- Morita, S., Horii, T., Kimura, M., Goto, Y., Ochiya, T., and Hatada, I. 2007. One Argonaute family member, Eif2c2 (Ago2), is essential for development and appears not to be involved in DNA methylation. *Genomics* **89**(6): 687-696.
- Morlando, M., Ballarino, M., Gromak, N., Pagano, F., Bozzoni, I., and Proudfoot, N.J. 2008. Primary microRNA transcripts are processed co-transcriptionally. *Nat Struct Mol Biol* **15**(9): 902-909.
- Muljo, S.A., Ansel, K.M., Kanellopoulou, C., Livingston, D.M., Rao, A., and Rajewsky, K. 2005. Aberrant T cell differentiation in the absence of Dicer. *J Exp Med* **202**(2): 261-269.
- Murchison, E.P., Partridge, J.F., Tam, O.H., Cheloufi, S., and Hannon, G.J. 2005. Characterization of Dicer-deficient murine embryonic stem cells. *Proc Natl Acad Sci U S A* **102**(34): 12135-12140.
- Mustelin, T., Vang, T., and Bottini, N. 2005. Protein tyrosine phosphatases and the immune response. *Nat Rev Immunol* **5**(1): 43-57.
- Nakayama, T., Kasprovicz, D.J., Yamashita, M., Schubert, L.A., Gillard, G., Kimura, M., Didierlaurent, A., Koseki, H., and Ziegler, S.F. 2002. The generation of mature, single-positive thymocytes in vivo is dysregulated by CD69 blockade or overexpression. *J Immunol* **168**(1): 87-94.
- Neely, L.A., Patel, S., Garver, J., Gallo, M., Hackett, M., McLaughlin, S., Nadel, M.,

- Harris, J., Gullans, S., and Rooke, J. 2006. A single-molecule method for the quantitation of microRNA gene expression. *Nat Methods* **3**(1): 41-46.
- Neilson, J.R., Zheng, G.X., Burge, C.B., and Sharp, P.A. 2007. Dynamic regulation of miRNA expression in ordered stages of cellular development. *Genes Dev* **21**(5): 578-589.
- Nelson, P.T., Hatzigeorgiou, A.G., and Mourelatos, Z. 2004. miRNP:mRNA association in polyribosomes in a human neuronal cell line. *RNA* **10**(3): 387-394.
- Newman, M.A., Thomson, J.M., and Hammond, S.M. 2008. Lin-28 interaction with the Let-7 precursor loop mediates regulated microRNA processing. *Rna* **14**(8): 1539-1549.
- Nichols, J., Zevnik, B., Anastassiadis, K., Niwa, H., Klewe-Nebenius, D., Chambers, I., Scholer, H., and Smith, A. 1998. Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. *Cell* **95**(3): 379-391.
- Nicoloso, M.S. and Calin, G.A. 2008. MicroRNA involvement in brain tumors: from bench to bedside. *Brain Pathol* **18**(1): 122-129.
- Nielsen, C.B., Shomron, N., Sandberg, R., Hornstein, E., Kitzman, J., and Burge, C.B. 2007. Determinants of targeting by endogenous and exogenous microRNAs and siRNAs. *Rna* **13**(11): 1894-1910.
- Niwa, R. and Slack, F.J. 2007. The evolution of animal microRNA function. *Curr Opin Genet Dev* **17**(2): 145-150.
- O'Rourke, J.R., Georges, S.A., Seay, H.R., Tapscott, S.J., McManus, M.T., Goldhamer, D.J., Swanson, M.S., and Harfe, B.D. 2007. Essential role for Dicer during skeletal muscle development. *Dev Biol* **311**(2): 359-368.
- Okamura, K., Hagen, J.W., Duan, H., Tyler, D.M., and Lai, E.C. 2007. The mirtron pathway generates microRNA-class regulatory RNAs in *Drosophila*. *Cell* **130**(1): 89-100.
- Orom, U.A., Kauppinen, S., and Lund, A.H. 2006. LNA-modified oligonucleotides mediate specific inhibition of microRNA function. *Gene* **372**: 137-141.
- Parker, J.S., Roe, S.M., and Barford, D. 2005. Structural insights into mRNA recognition from a PIWI domain-siRNA guide complex. *Nature* **434**(7033): 663-666.
- Parker, R. and Sheth, U. 2007. P bodies and the control of mRNA translation and degradation. *Mol Cell* **25**(5): 635-646.
- Paroo, Z., Ye, X., Chen, S., and Liu, Q. 2009. Phosphorylation of the human microRNA-generating complex mediates MAPK/Erk signaling. *Cell* **139**(1): 112-122.
- Pattillo, R.A. and Gey, G.O. 1968. The establishment of a cell line of human hormone-synthesizing trophoblastic cells in vitro. *Cancer Res* **28**(7): 1231-1236.
- Paul, M.S. and Bass, B.L. 1998. Inosine exists in mRNA at tissue-specific levels and is most abundant in brain mRNA. *Embo J* **17**(4): 1120-1127.
- Pauley, K.M., Eystathioy, T., Jakymiw, A., Hamel, J.C., Fritzler, M.J., and Chan, E.K. 2006. Formation of GW bodies is a consequence of microRNA genesis. *EMBO Rep* **7**(9): 904-910.
- Pawlicki, J.M. and Steitz, J.A. 2008. Primary microRNA transcript retention at sites of transcription leads to enhanced microRNA production. *J Cell Biol* **182**(1): 61-76.
- Pedersen, I.M., Cheng, G., Wieland, S., Volinia, S., Croce, C.M., Chisari, F.V., and David, M. 2007. Interferon modulation of cellular microRNAs as an antiviral mechanism. *Nature* **449**(7164): 919-922.
- Petersen, C.P., Bordeleau, M.E., Pelletier, J., and Sharp, P.A. 2006. Short RNAs repress translation after initiation in mammalian cells. *Mol Cell* **21**(4): 533-542.
- Pillai, R.S., Bhattacharyya, S.N., Artus, C.G., Zoller, T., Cougot, N., Basyuk, E., Bertrand, E., and Filipowicz, W. 2005. Inhibition of translational initiation by Let-7

- MicroRNA in human cells. *Science* **309**(5740): 1573-1576.
- Piriyapongsa, J., Marino-Ramirez, L., and Jordan, I.K. 2007. Origin and evolution of human microRNAs from transposable elements. *Genetics* **176**(2): 1323-1337.
- Poliseno, L., Salmena, L., Riccardi, L., Fornari, A., Song, M.S., Hobbs, R.M., Sportoletti, P., Varmeh, S., Egia, A., Fedele, G., Rameh, L., Loda, M., and Pandolfi, P.P. Identification of the miR-106b~25 MicroRNA Cluster as a Proto-Oncogenic PTEN-Targeting Intron That Cooperates with Its Host Gene MCM7 in Transformation. *Sci Signal* **3**(117): ra29.
- Porritt, H.E., Rumpfelt, L.L., Tabrizifard, S., Schmitt, T.M., Zuniga-Pflucker, J.C., and Petrie, H.T. 2004. Heterogeneity among DN1 prothymocytes reveals multiple progenitors with different capacities to generate T cell and non-T cell lineages. *Immunity* **20**(6): 735-745.
- Preker, P., Nielsen, J., Kammler, S., Lykke-Andersen, S., Christensen, M.S., Mapendano, C.K., Schierup, M.H., and Jensen, T.H. 2008. RNA exosome depletion reveals transcription upstream of active human promoters. *Science* **322**(5909): 1851-1854.
- Qi, H.H., Ongusaha, P.P., Myllyharju, J., Cheng, D., Pakkanen, O., Shi, Y., Lee, S.W., and Peng, J. 2008. Prolyl 4-hydroxylation regulates Argonaute 2 stability. *Nature* **455**(7211): 421-424.
- Quinn, J., Kunath, T., and Rossant, J. 2006. Mouse trophoblast stem cells. *Methods Mol Med* **121**: 125-148.
- Rehwinkel, J., Behm-Ansmant, I., Gatfield, D., and Izaurralde, E. 2005. A crucial role for GW182 and the DCP1:DCP2 decapping complex in miRNA-mediated gene silencing. *Rna* **11**(11): 1640-1647.
- Reik, W., Constancia, M., Fowden, A., Anderson, N., Dean, W., Ferguson-Smith, A., Tycko, B., and Sibley, C. 2003. Regulation of supply and demand for maternal nutrients in mammals by imprinted genes. *J Physiol* **547**(Pt 1): 35-44.
- Reinhart, B.J., Slack, F.J., Basson, M., Pasquinelli, A.E., Bettinger, J.C., Rougvie, A.E., Horvitz, H.R., and Ruvkun, G. 2000. The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*. *Nature* **403**(6772): 901-906.
- Rosa, A., Spagnoli, F.M., and Brivanlou, A.H. 2009. The miR-430/427/302 family controls mesendodermal fate specification via species-specific target selection. *Dev Cell* **16**(4): 517-527.
- Ruby, J.G., Jan, C., Player, C., Axtell, M.J., Lee, W., Nusbaum, C., Ge, H., and Bartel, D.P. 2006. Large-scale sequencing reveals 21U-RNAs and additional microRNAs and endogenous siRNAs in *C. elegans*. *Cell* **127**(6): 1193-1207.
- Ruby, J.G., Jan, C.H., and Bartel, D.P. 2007. Intronic microRNA precursors that bypass Drosha processing. *Nature* **448**(7149): 83-86.
- Rybak, A., Fuchs, H., Smirnova, L., Brandt, C., Pohl, E.E., Nitsch, R., and Wulczyn, F.G. 2008. A feedback loop comprising lin-28 and let-7 controls pre-let-7 maturation during neural stem-cell commitment. *Nat Cell Biol* **10**(8): 987-993.
- Sandberg, R., Neilson, J.R., Sarma, A., Sharp, P.A., and Burge, C.B. 2008. Proliferating cells express mRNAs with shortened 3' untranslated regions and fewer microRNA target sites. *Science* **320**(5883): 1643-1647.
- Sanvito, F., Piatti, S., Villa, A., Bossi, M., Lucchini, G., Marchisio, P.C., and Biffo, S. 1999. The beta4 integrin interactor p27(BBP/eIF6) is an essential nuclear matrix protein involved in 60S ribosomal subunit assembly. *J Cell Biol* **144**(5): 823-837.
- Savatier, P., Huang, S., Szekely, L., Wiman, K.G., and Samarut, J. 1994. Contrasting patterns of retinoblastoma protein expression in mouse embryonic stem cells and embryonic fibroblasts. *Oncogene* **9**(3): 809-818.
- Savatier, P., Lapillonne, H., van Grunsven, L.A., Rudkin, B.B., and Samarut, J. 1996.

- Withdrawal of differentiation inhibitory activity/leukemia inhibitory factor up-regulates D-type cyclins and cyclin-dependent kinase inhibitors in mouse embryonic stem cells. *Oncogene* **12**(2): 309-322.
- Schwarz, D.S., Hutvagner, G., Du, T., Xu, Z., Aronin, N., and Zamore, P.D. 2003. Asymmetry in the assembly of the RNAi enzyme complex. *Cell* **115**(2): 199-208.
- Seila, A.C., Calabrese, J.M., Levine, S.S., Yeo, G.W., Rahl, P.B., Flynn, R.A., Young, R.A., and Sharp, P.A. 2008. Divergent transcription from active promoters. *Science* **322**(5909): 1849-1851.
- Seitz, H., Ghildiyal, M., and Zamore, P.D. 2008. Argonaute loading improves the 5' precision of both MicroRNAs and their miRNA* strands in flies. *Curr Biol* **18**(2): 147-151.
- Seitz, H., Royo, H., Bortolin, M.L., Lin, S.P., Ferguson-Smith, A.C., and Cavaille, J. 2004. A large imprinted microRNA gene cluster at the mouse Dlk1-Gtl2 domain. *Genome Res* **14**(9): 1741-1748.
- Selbach, M., Schwanhauser, B., Thierfelder, N., Fang, Z., Khanin, R., and Rajewsky, N. 2008. Widespread changes in protein synthesis induced by microRNAs. *Nature* **455**(7209): 58-63.
- Shabalina, S.A. and Koonin, E.V. 2008. Origins and evolution of eukaryotic RNA interference. *Trends Ecol Evol* **23**(10): 578-587.
- Shendure, J. and Ji, H. 2008. Next-generation DNA sequencing. *Nat Biotechnol* **26**(10): 1135-1145.
- Shiow, L.R., Rosen, D.B., Brdickova, N., Xu, Y., An, J., Lanier, L.L., Cyster, J.G., and Matloubian, M. 2006. CD69 acts downstream of interferon-alpha/beta to inhibit S1P1 and lymphocyte egress from lymphoid organs. *Nature* **440**(7083): 540-544.
- Sinkkonen, L., Hugenschmidt, T., Berninger, P., Gaidatzis, D., Mohn, F., Artus-Revel, C.G., Zavolan, M., Svoboda, P., and Filipowicz, W. 2008. MicroRNAs control de novo DNA methylation through regulation of transcriptional repressors in mouse embryonic stem cells. *Nat Struct Mol Biol* **15**(3): 259-267.
- Smalheiser, N.R. and Torvik, V.I. 2006. Alu elements within human mRNAs are probable microRNA targets. *Trends Genet* **22**(10): 532-536.
- Smith, A.G., Heath, J.K., Donaldson, D.D., Wong, G.G., Moreau, J., Stahl, M., and Rogers, D. 1988. Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides. *Nature* **336**(6200): 688-690.
- Song, J.J., Liu, J., Tolia, N.H., Schneiderman, J., Smith, S.K., Martienssen, R.A., Hannon, G.J., and Joshua-Tor, L. 2003. The crystal structure of the Argonaute2 PAZ domain reveals an RNA binding motif in RNAi effector complexes. *Nat Struct Biol* **10**(12): 1026-1032.
- Song, J.J., Smith, S.K., Hannon, G.J., and Joshua-Tor, L. 2004. Crystal structure of Argonaute and its implications for RISC slicer activity. *Science* **305**(5689): 1434-1437.
- Sood, R., Zehnder, J.L., Druzin, M.L., and Brown, P.O. 2006. Gene expression patterns in human placenta. *Proc Natl Acad Sci U S A* **103**(14): 5478-5483.
- Stark, A., Brennecke, J., Bushati, N., Russell, R.B., and Cohen, S.M. 2005. Animal MicroRNAs confer robustness to gene expression and have a significant impact on 3'UTR evolution. *Cell* **123**(6): 1133-1146.
- Stark, A., Brennecke, J., Russell, R.B., and Cohen, S.M. 2003. Identification of Drosophila MicroRNA targets. *PLoS Biol* **1**(3): E60.
- Starr, T.K., Jameson, S.C., and Hogquist, K.A. 2003. Positive and negative selection of T cells. *Annu Rev Immunol* **21**: 139-176.
- Stead, E., White, J., Faast, R., Conn, S., Goldstone, S., Rathjen, J., Dhingra, U., Rathjen, P., Walker, D., and Dalton, S. 2002. Pluripotent cell division cycles are

- driven by ectopic Cdk2, cyclin A/E and E2F activities. *Oncogene* **21**(54): 8320-8333.
- Stefani, G. and Slack, F.J. 2008. Small non-coding RNAs in animal development. *Nat Rev Mol Cell Biol* **9**(3): 219-230.
- Strumpf, D., Mao, C.A., Yamanaka, Y., Ralston, A., Chawengsaksophak, K., Beck, F., and Rossant, J. 2005. Cdx2 is required for correct cell fate specification and differentiation of trophectoderm in the mouse blastocyst. *Development* **132**(9): 2093-2102.
- Su, H., Trombly, M.I., Chen, J., and Wang, X. 2009. Essential and overlapping functions for mammalian Argonautes in microRNA silencing. *Genes Dev* **23**(3): 304-317.
- Suh, M.R., Lee, Y., Kim, J.Y., Kim, S.K., Moon, S.H., Lee, J.Y., Cha, K.Y., Chung, H.M., Yoon, H.S., Moon, S.Y., Kim, V.N., and Kim, K.S. 2004. Human embryonic stem cells express a unique set of microRNAs. *Dev Biol* **270**(2): 488-498.
- Suzuki, R. and Shimodaira, H. 2006. Pvcust: an R package for assessing the uncertainty in hierarchical clustering. *Bioinformatics* **22**(12): 1540-1542.
- Tam, O.H., Aravin, A.A., Stein, P., Girard, A., Murchison, E.P., Cheloufi, S., Hodges, E., Anger, M., Sachidanandam, R., Schultz, R.M., and Hannon, G.J. 2008. Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes. *Nature* **453**(7194): 534-538.
- Tay, Y., Zhang, J., Thomson, A.M., Lim, B., and Rigoutsos, I. 2008. MicroRNAs to Nanog, Oct4 and Sox2 coding regions modulate embryonic stem cell differentiation. *Nature* **455**(7216): 1124-1128.
- Tichy, E.D. and Stambrook, P.J. 2008. DNA repair in murine embryonic stem cells and differentiated cells. *Exp Cell Res* **314**(9): 1929-1936.
- Till, S., Lejeune, E., Thermann, R., Bortfeld, M., Hothorn, M., Enderle, D., Heinrich, C., Hentze, M.W., and Ladurner, A.G. 2007. A conserved motif in Argonaute-interacting proteins mediates functional interactions through the Argonaute PIWI domain. *Nat Struct Mol Biol* **14**(10): 897-903.
- Tsang, J., Zhu, J., and van Oudenaarden, A. 2007. MicroRNA-mediated feedback and feedforward loops are recurrent network motifs in mammals. *Mol Cell* **26**(5): 753-767.
- Tycko, B. and Morison, I.M. 2002. Physiological functions of imprinted genes. *J Cell Physiol* **192**(3): 245-258.
- Vasudevan, S. and Steitz, J.A. 2007. AU-Rich-Element-Mediated Upregulation of Translation by FXR1 and Argonaute 2. *Cell* **128**(6): 1105-1118.
- Ventura, A., Young, A.G., Winslow, M.M., Lintault, L., Meissner, A., Erkeland, S.J., Newman, J., Bronson, R.T., Crowley, D., Stone, J.R., Jaenisch, R., Sharp, P.A., and Jacks, T. 2008. Targeted deletion reveals essential and overlapping functions of the miR-17 through 92 family of miRNA clusters. *Cell* **132**(5): 875-886.
- Viswanathan, S.R., Daley, G.Q., and Gregory, R.I. 2008. Selective blockade of microRNA processing by Lin28. *Science* **320**(5872): 97-100.
- Volinia, S., Calin, G.A., Liu, C.G., Ambs, S., Cimmino, A., Petrocca, F., Visone, R., Iorio, M., Roldo, C., Ferracin, M., Prueitt, R.L., Yanaihara, N., Lanza, G., Scarpa, A., Vecchione, A., Negrini, M., Harris, C.C., and Croce, C.M. 2006. A microRNA expression signature of human solid tumors defines cancer gene targets. *Proc Natl Acad Sci U S A* **103**(7): 2257-2261.
- Voorhoeve, P.M., le Sage, C., Schrier, M., Gillis, A.J., Stoop, H., Nagel, R., Liu, Y.P., van Duijse, J., Drost, J., Griekspoor, A., Zlotorynski, E., Yabuta, N., De Vita, G., Nojima, H., Looijenga, L.H., and Agami, R. 2006. A genetic screen implicates miRNA-372 and miRNA-373 as oncogenes in testicular germ cell tumors. *Cell*

- 124**(6): 1169-1181.
- Wagschal, A. and Feil, R. 2006. Genomic imprinting in the placenta. *Cytogenet Genome Res* **113**(1-4): 90-98.
- Wakiyama, M., Takimoto, K., Ohara, O., and Yokoyama, S. 2007. Let-7 microRNA-mediated mRNA deadenylation and translational repression in a mammalian cell-free system. *Genes Dev* **21**(15): 1857-1862.
- Wang, B., Li, S., Qi, H.H., Chowdhury, D., Shi, Y., and Novina, C.D. 2009. Distinct passenger strand and mRNA cleavage activities of human Argonaute proteins. *Nat Struct Mol Biol* **16**(12): 1259-1266.
- Wang, Y., Baskerville, S., Shenoy, A., Babiarz, J.E., Baehner, L., and Blelloch, R. 2008. Embryonic stem cell-specific microRNAs regulate the G1-S transition and promote rapid proliferation. *Nat Genet* **40**(12): 1478-1483.
- Wang, Y. and Blelloch, R. 2009. Cell cycle regulation by MicroRNAs in embryonic stem cells. *Cancer Res* **69**(10): 4093-4096.
- Wang, Y., Medvid, R., Melton, C., Jaenisch, R., and Blelloch, R. 2007. DGCR8 is essential for microRNA biogenesis and silencing of embryonic stem cell self-renewal. *Nat Genet* **39**(3): 380-385.
- Watanabe, T., Takeda, A., Tsukiyama, T., Mise, K., Okuno, T., Sasaki, H., Minami, N., and Imai, H. 2006. Identification and characterization of two novel classes of small RNAs in the mouse germline: retrotransposon-derived siRNAs in oocytes and germline small RNAs in testes. *Genes Dev* **20**(13): 1732-1743.
- Watanabe, T., Totoki, Y., Toyoda, A., Kaneda, M., Kuramochi-Miyagawa, S., Obata, Y., Chiba, H., Kohara, Y., Kono, T., Nakano, T., Surani, M.A., Sakaki, Y., and Sasaki, H. 2008. Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. *Nature* **453**(7194): 539-543.
- Wightman, B., Ha, I., and 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.
- Williams, R.L., Hilton, D.J., Pease, S., Willson, T.A., Stewart, C.L., Gearing, D.P., Wagner, E.F., Metcalf, D., Nicola, N.A., and Gough, N.M. 1988. Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells. *Nature* **336**(6200): 684-687.
- Wu, C., Orozco, C., Boyer, J., Leglise, M., Goodale, J., Batalov, S., Hodge, C.L., Haase, J., Janes, J., Huss, J.W., 3rd, and Su, A.I. 2009. BioGPS: an extensible and customizable portal for querying and organizing gene annotation resources. *Genome Biol* **10**(11): R130.
- Wu, L. and Belasco, J.G. 2005. Micro-RNA regulation of the mammalian *lin-28* gene during neuronal differentiation of embryonal carcinoma cells. *Mol Cell Biol* **25**(21): 9198-9208.
- Wu, L., Fan, J., and Belasco, J.G. 2006. MicroRNAs direct rapid deadenylation of mRNA. *Proc Natl Acad Sci U S A* **103**(11): 4034-4039.
- Xie, X., Lu, J., Kulbokas, E.J., Golub, T.R., Mootha, V., Lindblad-Toh, K., Lander, E.S., and Kellis, M. 2005. Systematic discovery of regulatory motifs in human promoters and 3' UTRs by comparison of several mammals. *Nature* **434**(7031): 338-345.
- Yan, K.S., Yan, S., Farooq, A., Han, A., Zeng, L., and Zhou, M.M. 2003. Structure and conserved RNA binding of the PAZ domain. *Nature* **426**(6965): 468-474.
- Yang, J.H., Luo, X., Nie, Y., Su, Y., Zhao, Q., Kabir, K., Zhang, D., and Rabinovici, R. 2003. Widespread inosine-containing mRNA in lymphocytes regulated by ADAR1 in response to inflammation. *Immunology* **109**(1): 15-23.
- Yang, W., Chendrimada, T.P., Wang, Q., Higuchi, M., Seeburg, P.H., Shiekhattar, R.,

- and Nishikura, K. 2006. Modulation of microRNA processing and expression through RNA editing by ADAR deaminases. *Nat Struct Mol Biol* **13**(1): 13-21.
- Yekta, S., Shih, I.H., and Bartel, D.P. 2004. MicroRNA-directed cleavage of HOXB8 mRNA. *Science* **304**(5670): 594-596.
- Yi, R., Qin, Y., Macara, I.G., and Cullen, B.R. 2003. Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes Dev* **17**(24): 3011-3016.
- Zdanowicz, A., Thermann, R., Kowalska, J., Jemielity, J., Duncan, K., Preiss, T., Darzynkiewicz, E., and Hentze, M.W. 2009. Drosophila miR2 primarily targets the m7GpppN cap structure for translational repression. *Mol Cell* **35**(6): 881-888.
- Zeng, Y. and Cullen, B.R. 2005. Efficient processing of primary microRNA hairpins by Drosha requires flanking nonstructured RNA sequences. *J Biol Chem* **280**(30): 27595-27603.
- Zeng, Y., Sankala, H., Zhang, X., and Graves, P.R. 2008. Phosphorylation of Argonaute 2 at serine-387 facilitates its localization to processing bodies. *Biochem J* **413**(3): 429-436.
- Zhang, R., Peng, Y., Wang, W., and Su, B. 2007. Rapid evolution of an X-linked microRNA cluster in primates. *Genome Res* **17**(5): 612-617.
- Zhang, R., Wang, Y.Q., and Su, B. 2008. Molecular evolution of a primate-specific microRNA family. *Mol Biol Evol* **25**(7): 1493-1502.
- Zhao, X., Ayer, R.E., Davis, S.L., Ames, S.J., Florence, B., Torchinsky, C., Liou, J.S., Shen, L., and Spanjaard, R.A. 2005a. Apoptosis factor EI24/PIG8 is a novel endoplasmic reticulum-localized Bcl-2-binding protein which is associated with suppression of breast cancer invasiveness. *Cancer Res* **65**(6): 2125-2129.
- Zhao, Y., Ransom, J.F., Li, A., Vedantham, V., von Drehle, M., Muth, A.N., Tsuchihashi, T., McManus, M.T., Schwartz, R.J., and Srivastava, D. 2007. Dysregulation of Cardiogenesis, Cardiac Conduction, and Cell Cycle in Mice Lacking miRNA-1-2. *Cell* **129**(2): 303-317.
- Zhao, Y., Samal, E., and Srivastava, D. 2005b. Serum response factor regulates a muscle-specific microRNA that targets Hand2 during cardiogenesis. *Nature* **436**(7048): 214-220.