Small RNAs – their biogenesis, regulation and function in embryonic stem cells*

Joshua E. Babiarz1§ and Robert Blelloch1§, 1The Eli and Edythe Broad Center of Regeneration Medicine and Stem Cell Research, Center for Reproductive Sciences, and Department of Urology, University of California, San Francisco, California 94143, USA

Table of Contents

1. Introduction .............................................................................1
2. The Biogenesis of miRNAs and related small RNAs ............................................2
3. The role of small RNAs in embryonic stem cells ............................................3
4. Transcriptional regulation of the miRNA content of a cell ..........................4
5. Post-transcriptional regulation of the processing of miRNAs .....................5
6. miRNAs and induced pluripotency .................................................7
7. Non-canonical miRNAs and endogenous siRNAs ..................................8
8. Conclusions ............................................................................10
9. References ............................................................................10

1. Introduction

Mammalian development follows a carefully orchestrated unfolding of cell fate decisions whose time, place, and specialization must occur robustly and consistently to end with a healthy adult organism. How this feat is achieved time and time again remains a mystery and is of intense interest to the biological and medical communities. One of the central aspects of this feat is the rapid shifts in molecular programs during developmental transitions. These shifts must occur at multiple levels including the regulation of transcription, mRNA stability, protein translation, protein stability, and protein activity. Regulation occurs with the help of transcription factors (activators and suppressors), epigenetic enzymes, non-coding RNAs, RNA binding proteins, ubiquinating enzymes, kinases, phosphatases, and more. Of great interest are a specific group of non-coding RNAs, the microRNAs, which are only 18–24 nucleotides in length and function by both destabilizing and inhibiting translation of mRNAs. microRNAs are compelling regulators of developmental progression as their promiscuous nature allows a single miRNA to regulate hundreds of mRNAs, simultaneously allowing rapid shifts in the transcriptome and proteome of cells.


Copyright: © 2009 Joshua E. Babiarz and Robert Blelloch. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

§To whom correspondence should be addressed. E-mail: babiarzj@stemcell.ucsf.edu, blellochr@stemcell.ucsf.edu
2. The Biogenesis of miRNAs and related small RNAs

MicroRNAs undergo a series of processing events before maturing into a functional miRNA-silencer complex (Figure 1). Typically, miRNAs are initially expressed as part of a longer RNA polymerase II (pol II) driven transcript, the primary miRNA (pri-miRNA). These longer transcripts can be either coding (often with the miRNA within the introns), or non-coding (Cai et al., 2004; Lee et al., 2004). The miRNAs themselves are contained within hairpin folds along the length of the pri-miRNAs. These hairpins are recognized by the RNA binding protein, Dgcr8 (Denli et al., 2004; Gregory et al., 2004; Han et al., 2004; Han et al., 2006; Landthaler et al., 2004), which then directs the RNase III-containing enzyme, Drosha, to cleave the RNA at the base of the hairpin resulting in the precursor miRNA (pre-miRNA) (Lee et al., 2003). Dgcr8 and Drosha form the Microprocessor complex. Two pieces of evidence suggest that the Microprocessor processes the hairpins as the pri-miRNA is being transcribed, as opposed to following transcription. First, Drosha and RNA pol II can be cross-linked to DNA near the pre-miRNA hairpins, suggesting that the Microprocessor is in close proximity to the DNA (Morlando et al., 2008). Second intronic pre-miRNAs can be released by the Microprocessor prior to splicing of the surrounding exons (Kim and Kim, 2007).

Following production of the pre-miRNA hairpin by the Microprocessor, Exportin-5 transports the pre-miRNA to the cytoplasm (Bohsack et al., 2004; Lund et al., 2004; Yi et al., 2005; Yi et al., 2003). Once in the cytoplasm, a second RNase III-containing enzyme, Dicer, cleaves the pre-miRNA to liberate the mature miRNA and the complementary “star” sequence (Reviewed in (Hammond, 2005)). Similar to the Microprocessor, where the RNase III Drosha pairs with the RNA-binding protein Dgcr8, Dicer pairs with the human immunodeficiency virus-1 transactivating response RNA-binding protein (TRBP) (Chendrimada et al., 2005; Haase et al., 2005). However, unlike the Microprocessor, where Dgcr8 is essential for activity (Wang et al., 2007), TRBP appears to be dispensable for Dicer activity in vitro, but important for choosing the correct strand for transfer to the silencing complex and downstream gene silencing (Chendrimada et al., 2005; Gregory et al., 2005; Haase et al., 2005).

In addition to miRNA hairpins, Dicer is also capable of cleaving long, double stranded RNAs to small RNAs. The resulting small RNAs are named small interfering RNAs (siRNA). Endogenous siRNAs are found across many species and their function may be even more evolutionary conserved than miRNAs. Indeed, Dicer exists in the absence of the Microprocessor in many species outside of the metazoan lineage (Grimson et al., 2008). siRNAs are capable of directing heterochromatin in non-mammalian eukaryotes, such as *S. pombe* (Verdel and Moazed, 2005), *Tetrahymena*...
(Liu et al., 2007) and *Drosophila* (Riddle and Elgin, 2008). Whether there is a similar role for siRNAs in mammalian cells remains unclear. However, endogenous siRNAs have been recently discovered in mammals (Babiarz et al., 2008; Tam et al., 2008; Watanabe et al., 2006; Watanabe et al., 2008; Yang and Kazazian, 2006).

Following processing by Dicer, one strand of the duplex miRNA enters into the silencing complex, the RNA-induced silencing complex (RISC). RISC contains the Argonaute proteins and plays a central role in executing miRNA based silencing. The mechanism for choice of strand and removal of the opposing strand is not entirely clear. However, it appears the energetically less stable 5′ end of the duplex is preferentially loaded (Kvorova et al., 2003; Schwarz et al., 2003), while the passenger strand is cleaved by Argonaute-2 (Ago2) (Rand et al., 2005). Ago2 is one of four argonaute proteins (Ago1 to 4) in mammals (Carmell et al., 2002). However, it is only one with cleavage (or “slicer”) activity (Liu et al., 2004; Meister et al., 2004; Rand et al., 2005) and is the only one essential for development (Alisch et al., 2007; Liu et al., 2004; Morita et al., 2007). Although Ago2 is essential for normal development, it appears that the four Argonautes may functionally overlap in miRNA silencing. Deletion of all four Argonautes in mouse ES cells results in a loss of silencing and apoptosis (Su et al., 2009). Reintroduction of any of the four Argonautes in these cells results in a rescue of both post-transcriptional silencing and apoptosis (Su et al., 2009).

The miRNA loaded RISC decreases protein levels encoded by the target mRNAs, both by destabilizing the mRNA and by inhibiting translation. Although the precise mechanism for how a miRNA finds a particular target and inhibits translation is the subject of intense investigation, many rules are already known (Bartel, 2009). The major determinant for miRNA targeting is the “seed” sequence – the 7mer sequence present at positions 2 through 8 of the mature miRNA (Lewis et al., 2003). However, there are other components of the miRNA sequence that appear to influence target repression, including nucleotides 12 to 17 in the miRNA and the position of the target site in the 3′ UTR of the target gene (Grimson et al., 2007). Roughly one-third of human genes contain conserved miRNA binding-sites in their 3′ UTR (Lewis et al., 2005), and the number of recognized conserved sites has recently increased with more sensitive detection methods (Friedman et al., 2009). Furthermore, recent data show that miRNAs can also act through binding in the ORF of coding genes (Baek et al., 2008; Selbach et al., 2008; Tay et al., 2008), thereby further increasing the degree of miRNA control of the transcriptome.

Advances in quantitative mass spectrometry have enabled the study of the impact of miRNAs on the proteome as a whole, rather than selected targets. Using mass spectrometry, two studies compared the impact of exogenously added miRNAs on the proteome of HeLa cell nuclei (Baek et al., 2008; Selbach et al., 2008). Additionally, these groups analyzed the impact of removing a miRNA, either by deletion in the mouse (Baek et al., 2008) or by endogenous miRNA knockdown in HeLa cells (Selbach et al., 2008). These complimentary approaches revealed several important findings: First, bioinformatic analysis of downregulated genes confirmed that the most important predictor of target recognition is a sequence match to the miRNA seed sequence in the 3′ UTR. Second, translational repression is more often than not associated with a parallel decrease in mRNA levels. Third, miRNA effects on their targets are rarely dramatic, instead resulting in small changes in levels across many proteins (Baek et al., 2008; Selbach et al., 2008).

A surprising additional role for microRNAs activating, rather than suppressing, translation has been reported. Specifically, when cells in culture were induced to exit the cell cycle by serum starvation, miR-369–3 and let-7a bound the 3′ UTRs and upregulated translation of TNFα and HMGA2 (Vasudevan et al., 2007). Under cell growth conditions, these same miRNAs suppressed the targets suggesting a switch between miRNA induced suppression and activation of translation depending on the cell cycle status (Vasudevan et al., 2007). Transcriptional activation has also been proposed as an alternative role for miRNAs. Following introduction of exogenously synthesized dsRNAs directed to their promoters, expression of E-cadherin, p21, VEGF, and the progesterone receptor were increased (Janowski et al., 2007; Li et al., 2006). The mechanism for such upregulation may involve the dsRNAs targeting antisense transcripts produced at these promoters and the recruitment of Ago1 (Schwartz et al., 2008). An endogenous miRNA, miR-373, may act in a similar fashion to these exogenously introduced dsRNAs suggesting a potential physiological role for such a mechanism (Place et al., 2008). It will be interesting to see how robust these novel mechanisms of miRNA action are in vivo.

### 3. The role of small RNAs in embryonic stem cells

The first miRNA, lin-4, was identified based on the alterations in cell lineage specification of *C. elegans* (Lee et al., 1993). Analysis of many other organisms has confirmed a central role of miRNAs in normal metazoan development. In mammals, multiple components of the miRNA pathway are essential and their deletion results in early death.
embryonic lethality. This embryonic lethality has been observed for the deletion of Dicer (Bernstein et al., 2003), Dgcr8 (Wang et al., 2007), and Ago2 (Alisch et al., 2007; Liu et al., 2004; Morita et al., 2007). To gain insight into early development, embryonic stem cells that contained null alleles of Dgcr8 and Dicer were generated (Kanellopoulou et al., 2005; Murchison et al., 2005; Wang et al., 2007). ES cells lacking either Dgcr8 or Dicer have defects in proliferation and differentiation. The DicerΔ/Δ ES cells appeared to have more severe phenotypes than the Dgcr8Δ/Δ ES cells, including a stricter block in differentiation, a severer proliferation defect immediately following loop out, and an inability to form teratomas. These differences suggested the presence of Dgcr8-independent, Dicer-dependent small RNAs – a conclusion that has been recently confirmed (Babiarz et al., 2008).

Although these phenotypes occurred in the absence of all canonical miRNAs, it remained possible that they could be explained at least in part by one or a few miRNAs that are highly expressed in ES cells. For example, one important difference between wild type and Dgcr8Δ/Δ ES cells is that the ES cells accumulated in the G1 phase of the cell cycle in mutants (Wang et al., 2007). Initial attempts at uncovering miRNAs responsible for this defect in a wild-type background failed, presumably due to redundancy among the ES cell miRNAs. Supporting this notion, relatively few individual miRNA knockouts in C. elegans have defects in development or viability (Miska et al., 2007). To circumvent the issue of redundancy an alternative approach was taken. Starting with the Dgcr8Δ/Δ ES cells lacking all canonical miRNAs, individual miRNAs were added back in the form of miRNA mimics, chemically synthesized double stranded miRNAs representing the post-Dicer miRNA duplex. This approach enabled the study of each miRNA in isolation from the normal milieu. A screen using miRNA mimics representing 266 of the 461 annotated miRNAs (Griffiths-Jones et al., 2008) revealed a large family of miRNAs, sharing the same sequence that could rescue the G1 defect in Dgcr8Δ/Δ ES cells (Wang et al., 2008). This family includes a subset of the mir-290 cluster of miRNAs, which are specifically expressed in ES cells (Houbaviy et al., 2003; Wang et al., 2008). These miRNAs were coined ESCC for ES cell–specific cell cycle–regulating miRNAs. Further analysis uncovered targets for the ESCC miRNAs that included multiple inhibitors of the G1/S transition, regulators of the cyclin/cdk2 pathway, including p21 (Cdkn1a), Rb, Rbl2, and Lats2 (Wang et al., 2008). These results correlated nicely with previous results with p21, which showed that while its mRNA levels changed little, its protein levels rapidly increased upon differentiation coincidently with the lengthening of the G1 phase of the cell cycle (Sabapathy et al., 1997). Taken together, these findings show that in ES cells, the mir-290 cluster is highly expressed, members of which suppress inhibitors of the G1/S transition promoting an abbreviated G1 phase and hence a shortened cell cycle. Upon differentiation, the mir-290 cluster is rapidly down regulated, allowing the translation of the G1/S inhibitors and the lengthening of G1.

In addition to the role the mir-290 cluster plays in the unique ES cell cycle, two groups identified an indirect control of DNA methyltransferase by the mir-290 cluster (Sinkkonen et al., 2008). Similar to Wang et al, the authors found that the mir-290 cluster controlled Rbl2. The authors also found that members of the DNA methyltransferase complex, Dnmt3a and 3b, were downregulated in DicerΔ/Δ relative to wild type. Rbl2 had previously been shown to transcriptionally suppress Dnmt3b (Litovchick et al., 2007) and the authors showed that inhibition of Rbl2 in the knockout cells rescues Dnmt3a and 3b levels. The authors suggest that the diminished Dnmt levels explain the deficiency of Dicer null cells to differentiate as de novo DNA methylation is important in suppressing pluripotency genes (Feldman et al., 2006). However, unlike the Dgcr8 and Dicer knockouts, differentiation appears to initiate normally and pluripotency genes are downregulated in the absence of Dnmt3a and 3b, with a fraction of cells reactivating the pluripotency program when returned to ES cell conditions (Feldman et al., 2006). Furthermore, the mir-290 cluster members do not rescue differentiation (Melton and Blelloch, unpublished observation). In either case, the de novo DNMTs appear to be important indirect targets of the mir-290 cluster in ES cells.

Taken together, the study of ES cells that lacked miRNAs specifically (dgcr8Δ/Δ) or even broader classes of small RNAs (dicerΔ/Δ) have been used successfully to uncover important roles for the mir-290 cluster of miRNAs: ES cell cycle regulation and DNA methylation. Although earlier studies had identified the mir-290 cluster as highly enriched in ES cells and the early embryo (Houbaviy et al., 2003), the precise function of these abundant miRNAs had not been appreciated. The combination of global miRNA knockout models along with miRNA mimics and inhibitors should continue to identify miRNAs responsible for specific phenotypes in ES cells and other differentiated cell types.

4. Transcriptional regulation of the miRNA content of a cell

Growing evidence suggests that, like the transcriptome and proteome, the miRNA content of an individual cell is unique (Houbaviy et al., 2003; Marson et al., 2008). For instance, ES cells are predominated by the mir-290
cluster (Houbaviy et al., 2003; Marson et al., 2008) and a more differentiated cell type, such as neural precursor cells (NPCs), are predominated by the let-7 family (Marson et al., 2008). Several recent studies have begun to shed light into regulation of the miRNA content of a cell (Marson et al., 2008; Ozsolak et al., 2008). Not surprisingly, many of the mechanisms appear similar to the regulation of other developmental genes. Much of the work has focused on ES cells, but the themes of the findings presented thus far are likely operating in other tissue types.

For instance, Marson and colleagues recently published evidence that the core pluripotency transcription factors, Oct4, Sox2, Nanog, and Tcf3, are present at the promoter of the ES-cell specific mir-290 and mir-302 clusters in mouse ES cells. Additionally, these authors showed that Oct4 is present at the human ES-cell specific mir-302 cluster. Intellectually, this result is very satisfying, as it ties the mir-290 cluster into the ES cell transcriptional network. Indeed, knockdown of Oct4 results in decreased expression of the mir-290 cluster. Furthermore, during ES cell differentiation, the down-regulation of the mir-290 cluster is coincident with the loss of Oct4, Nanog, Sox2, and Tcf3.

Beyond identifying how the core ES-cell transcription factor machinery controls certain miRNAs, Marson and colleagues also identified mechanisms for how certain miRNAs are epigenetically silenced in ES cells. Previous work in mouse ES cells had identified a surprising epigenetic domain, coined a bivalent domain, which contained both active (Histone H3K4me3) and repressive (Histone H3K27me3) marks in the same genomic region (Bernstein et al., 2006). These bivalent domains were predominantly found at the promoters of developmentally important transcription factors that were silent in ES cells, but became activated with differentiation in a lineage dependent fashion (Bernstein et al., 2006). The bivalent domains appear to “resolve” with differentiation: genes that are activated retain the H3K4me3 and lose the H3K27me3 mark, while genes that are suppressed lose H3K4me3 and retain H3K27me3 (Bernstein et al., 2006; Mikkelsen et al., 2007). By examining the H3K4me3 and H3K27me3 marks at miRNA promoters in ES cells and more differentiated cell types, Marson and colleagues were able to show that these bivalent domains also occur at miRNA genes that are silent in ES cells and are activated in a lineage specific fashion (Marson et al., 2008). The most compelling example the authors provide is that of mir-9, which is present in NPCs but absent in ES cells and embryonic fibroblasts (MEFs). The promoter of mir-9 is bivalent in ES cells and is resolved to H3K4me3 in NPCs, versus H3K27me3 in MEFs.

Taken together, these data suggest that regulation of miRNAs during development follows many of the same rules as other developmentally important genes. The unique content of transcription factors and epigenetic modifications found in any particular cell will define the diversity of miRNAs present. In turn, these miRNAs feed back to regulate the signaling pathways, transcription factors and epigenetic status of genes resulting in both negative and positive feedback loops (Marson et al., 2008). Therefore, during cell fate transitions these programs will change in parallel reinforcing the shift from one cell type to another.

5. Post-transcriptional regulation of the processing of miRNAs

Beyond transcription, there are additional layers of regulation of miRNA content of cells. In particular, there is growing evidence for regulation occurring at the level of miRNA biogenesis (Table 1). This regulation, could occur at any one of the steps in their biogenesis pathway including Microprocessor cleavage, Exportin-5 nuclear/cytoplasmic export, Dicer cleavage, transfer to RISC, and RISC silencing. At the global level, it is now clear that the levels of the Microprocessor are very tightly regulated. This regulation occurs through a combination of a negative and a positive feedback loop (Han et al., 2009). The negative feedback loop consists of the Microprocessor cleaving hairpins in the 5’UTR and coding sequence of Dcr8, thereby destabilizing its miRNA. The positive feedback loop involves Dgcr8 protein stabilizing the Drosha protein through an unclear mechanism. The end result is that any transient shifts in the Microprocessor levels will be rapidly corrected. However, Microprocessor levels can differ between cells both during normal development (Shiohama et al., 2003) and in disease (Ambs et al., 2008; Merritt et al., 2008; Muralidhar et al., 2007), suggesting a resetting of the homeostatic Microprocessor levels.

At the individual miRNA level, there is mounting evidence for both positive and negative regulation of biogenesis. For example, analysis of TGF-β and BMP4 activation of primary human pulmonary artery smooth muscle cells revealed that miR-21 levels are increased post-transcriptionally following growth factor addition (Davis et al., 2008). Because the SMAD proteins are direct targets of TGF-β and BMP4, the authors initially investigated whether receptor-specific SMADs would induce transcription of the pri-miR-21 transcript. However, the pri-miR-21 levels remained unchanged while the pre-, and mature miR-21 levels rose following TGF-β or BMP4 treatment. To confirm a lack of transcription was modulating miRNA levels, the cells were treated with both BMP4 and α-amanitin, which inhibits RNA polymerase II. In the presence of α-amanitin, there was a similar rise in pre-miR-21 and mature miR-21 levels, but an absence
Table 1. miRNAs exhibiting post-transcriptionally regulated processing

<table>
<thead>
<tr>
<th>miRNA</th>
<th>Biogenesis Step</th>
<th>Type of Regulation</th>
<th>Factor</th>
<th>Cell Type</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>miR-21</td>
<td>Microprocessor</td>
<td>Promotes</td>
<td>SMADs</td>
<td>primary human pulmonary artery smooth muscle cells</td>
<td>(Davis et al., 2008)</td>
</tr>
<tr>
<td>miR-18a</td>
<td>Microprocessor</td>
<td>Promotes</td>
<td>hnRNP A1</td>
<td>(Guil and Caceres 2007; Michlewski et al. 2008)</td>
<td></td>
</tr>
<tr>
<td>let-7 family</td>
<td>Microprocessor</td>
<td>Inhibits</td>
<td>Lin-28</td>
<td>mES and mEC</td>
<td>(Newman et al. 2008; Viswanathan et al. 2008)</td>
</tr>
<tr>
<td>let-7 family</td>
<td>Dicer</td>
<td>Inhibits</td>
<td>Lin-28</td>
<td>Hepatocellular</td>
<td>(Rybak et al. 2008)</td>
</tr>
<tr>
<td>let-7 family</td>
<td>Dicer</td>
<td>Inhibits</td>
<td>Lin-28</td>
<td>Hepatocellular</td>
<td>(Heo et al. 2008)</td>
</tr>
<tr>
<td>miR-221, miR-222</td>
<td>RISC translational Inhibition of p27</td>
<td>Inhibits</td>
<td>Dnd1</td>
<td>Human Germ Cells</td>
<td>(Kedde et al. 2007)</td>
</tr>
<tr>
<td>miR-138</td>
<td>Dicer</td>
<td>Inhibits</td>
<td>Unknown</td>
<td>HeLa</td>
<td>(Obernosterer et al. 2006)</td>
</tr>
<tr>
<td>miR-101-1, let-7a-1, miR-379, miR-31</td>
<td>Microprocessor</td>
<td>Promotes</td>
<td>Unknown</td>
<td>In vitro</td>
<td>(Michlewski et al. 2008)</td>
</tr>
</tbody>
</table>

of transcription of a known downstream BMP4 target. SMAD1 was previously shown to interact with p68 (Warner et al., 2004), an RNA-helicase present in the Microprocessor complex. Furthermore, it was shown that the loss of p68 causes a decrease in miR-21 in E9.5 embryos (Fukuda et al., 2007). Therefore, Davis et al. tested and showed that the SMADs interact with pri-mir-21, recruit the Microprocessor via the p68 subunit, thereby enhance pri- to pre-mir-21 processing (Davis et al., 2008). These results showed that signaling molecules are able to direct post-transcriptional increases in miRNA processing, which enables rapid miRNA mediated responses.

Another example of positive post-transcriptional regulation of miRNAs is the ribonuclear protein, hnRNP A1, which regulates the processing of miR-18a. Using CLIP (cross-linking and IP) (Ule et al., 2005; Ule et al., 2003) to identify targets of hnRNP A1, Guil et al. pulled down miR-18a and showed that binding of hnRNP A1 to the pri-miR-18a transcript facilitates its processing (Guil and Caceres, 2007). Depletion of hnRNP A1 in HeLa cells resulted in a reduction in the amounts of miR-18a (Guil and Caceres, 2007). A follow-up study showed that hnRNP A1 binds to the stem of pri-miR-18a, altering the secondary structure of the hairpin, which facilitates cleavage by the Microprocessor (Michlewski et al., 2008). The authors suggest this type of regulation may be widespread, as they found evolutionary conservation in the stem loops of a variety of pre-miRNA hairpins suggesting selection to keep the loop intact, and hence a likely regulatory role. Furthermore, anti-sense oligonucleotides against the conserved loop of the pre-miRNA hairpins often diminished processing in vitro (Michlewski et al., 2008). These results suggest that there may be an entire class of proteins that bind the loop of the pri-miRNA hairpins in order to promote Microprocessor cleavage. The lack of a Microprocessor facilitating protein may allow pri-miRNAs to accumulate and then be rapidly processed to mature miRNAs following the expression of that facilitator.

In contrast to the hnRNP story, another protein, Lin-28, has been shown to block miRNA processing. The exact step where the processing is blocked has been controversial as two studies identified a block at the Microprocessor step (Newman et al., 2008; Viswanathan et al., 2008) while two others identified the block at the Dicer step (Heo et al., 2008; Rybak et al., 2008). Importantly though, they all agree that Lin-28 inhibits the maturation of a particular subset of miRNAs, the let-7 miRNAs. The search for Lin-28 began with the finding that while mature let-7 was absent in ES cells and the early embryo, the pri-miRNA was being transcribed (Thomson et al., 2006; Wulczyn et al., 2007). With ES cell differentiation or embryonic maturation, mature let-7 levels rose with little change in the steady state pri-let-7 levels. This finding suggested a block in processing somewhere between the pri- and mature miRNA. As the pri-let-7 miRNAs were present in the ES cells, two groups hypothesized that the block had to be at the pri- to pre-miRNA processing step (Newman et al., 2008; Viswanathan et al., 2008). They determined that there was indeed a protein present in ES, but not differentiated cell extracts that could block pri- to pre-miRNA processing. Therefore, they used affinity columns along with an in vitro Microprocessor activity assay to identify the responsible protein(s). The affinity columns consisted of beads linked to either pre-let-7g (Viswanathan et al., 2008) or the let-7d loop (Newman
et al., 2008). Purified extracts from these columns retained inhibitory activity and, therefore, were analyzed by mass spectrometry. From this analysis and follow-up experiments, both groups identified Lin-28, and the closely related Lin-28B protein as the inhibitory factors. Indeed, ectopic expression of Lin-28 in a cell line that has high mature let-7 levels and low Lin-28 resulted in a decrease in mature let-7, while Lin-28 knockdown in ES cells resulted in elevated mature let-7 levels (Newman et al., 2008; Viswanathan et al., 2008). These findings led the authors to conclude that Lin-28 bound to the let-7 hairpins in the pri-miRNA preventing Microprocessor cleavage.

However, Lin-28 is predominantly a cytoplasmic protein (Balzer and Moss, 2007), while pri- to pre- processing occurs in the nucleus. Work by two other groups may have shed some light on this conundrum. Rybak and colleagues had previously suggested that pre-, not pri-let-7 family members accumulated in ES cells (Wulczyn et al., 2007). In vitro assays using the pre-let-7 hairpin as opposed to the longer pri-let-7, revealed that Lin-28 blocks mature let-7 accumulation at the Dicer step (Rybak et al., 2008). Additionally, Heo et al. found that Lin-28B blocks Dicer cleavage of pre-let-7 by promoting poly-uridinylation of the pre-let-7. This uridinylation blocks Dicer cleavage and simultaneously targets the RNA for degradation. Further, they found that the let-7 inhibitory function was present primarily in the cytoplasm, consistent with Lin-28’s subcellular localization.

Recent evidence has emerged that suggests Lin-28, and by extension, let-7 family members, are integrated into the pluripotency circuit. The Lin-28 promoter is occupied by the core pluripotency transcription factors Oct4, Nanog, Sox2, and Tcf3 (Marson et al., 2008). The high level of Lin-28 expression prevents the accumulation of mature let-7 family miRNAs. Surprisingly, the promoter of at least one member of the let-7 family, let-7g, is also bound by Oct4, Nanog, Sox2, and Tcf3, likely driving high levels of transcription of pri-let-7g in ES cells (Marson et al., 2008). Upon the induction of differentiation, Oct4, Nanog, and Sox2 are downregulated. This results in the downregulation of Lin-28 and concomitant upregulation of the processing of the accumulated pri- or pre-let-7g to mature let-7g. This upregulation of let-7 further downregulates Lin-28 by targeting the Lin-28 3′ UTR (John et al., 2004; Moss and Tang, 2003; Nelson et al., 2004; Reinhart et al., 2000; Wu and Belasco, 2005) producing a positive feedback loop or “bistable switch” (Heo et al., 2008). This type of regulatory circuit allows the powerful and rapid transition from a pluripotent stem cell that lacks let-7 to a more differentiated cell that has mature let-7.

Finally, a recent report revealed that following processing and RISC loading, at least one RNA binding protein is capable of inhibiting the silencing effect of RISC on certain mRNAs. Noting the a high level of conservation of the 3′ UTR of p27, outside of the miR-221 and miR-222 seeds, Kedde and colleagues speculated that an RNA-binding protein may recognize this sequence and modulate miR-221/222 silencing of p27 (Kedde et al., 2007). Through co-transfections of candidate RNA-binding proteins and miR-221 in HEK-293 cells, the authors found Dead end 1 (Dnd1) prevented miR-221 silencing of p27. Surprisingly, mutating the miR-221 site to a perfect match, showed that Dnd1 could not prevent the siRNA slicing of the transcript. This finding suggests that Dnd1 does not merely function through steric hindrance, although it may prevent a stable, long-term association of RISC with the 3′ UTR. The authors conclusively showed that Dnd1 inhibits miR-221 silencing of p27. It will be interesting to see if Dnd1 or similar proteins influence other mRNA/miRNA interactions.

Taken together, the findings discussed above show a broad range of control of miRNAs at nearly every step in their biogenesis as well as execution of their function. At this point, descriptions of only a handful miRNAs controlled at any particular step are published. It will be fascinating to see how many miRNAs are controlled at each of these steps, and how many proteins are capable of these types of regulation.

6. miRNAs and induced pluripotency

The finding that a combination of transcription factors can reprogram differentiated adult cells into a pluripotent state has revolutionized the regenerative medicine field. Interestingly, alternative combinations of viral expressed factors have been used to successfully reprogram mouse and human ES cells (Takahashi et al., 2007; Takahashi and Yamanaka, 2006; Yu et al., 2007). The original mix of factors used to dedifferentiate mouse somatic cells to induced pluripotent stem cells was a combination of Oct4, Sox2, c-Myc, and Klf4 (Takahashi and Yamanaka, 2006). This same set of factors was later shown to produce human iPS cells as well (Lowry et al., 2008; Park et al., 2008; Takahashi et al., 2007). However, another group identified a different combination of factors that could reprogram human cells consisting of Oct4, Sox2, Nanog, and Lin-28 (Yu et al., 2007). From a “miRNA-centric” view of reprogramming, these reprogramming factors are extremely intriguing. Oct4 and Sox2 (and Nanog) are important for driving the ES cell-specific miRNA clusters mir-290 (mouse) and mir-302 (mouse and human) (Marson et al., 2008). Lin-28 inhibits let-7 biogenesis, and, therefore, may support reprogramming by depleting the cells of mature
let-7. Finally, c-Myc controls the transcription of many miRNAs, in both a positive and negative fashion. c-Myc can drive the expression of the mir-17–92 cluster (O’Donnell et al., 2005), which is one of the more highly expressed families in ES cells (Calabrese et al., 2007; Marson et al., 2008). Additionally, c-Myc can downregulate a number of miRNAs, including let-7 family members (Chang et al., 2008). Therefore, the reprogramming factors may be acting in part by altering the miRNA content of the somatic cell to that of a pluripotent stem cell in order to induce dedifferentiation.

Indeed, a recent report has shown that the transfection of a subset of the mir-290 cluster miRNAs can replace Myc in reprogramming mouse fibroblasts to induced pluripotent stem cells (Judson et al., 2009). These were the same ESCC miRNAs that were previously shown to promote the ES cell cycle (Wang et al., 2008). This finding suggests that the transcriptional activation of the mir-290 cluster and possibly the establishment of the unique ES cell cycle structure are critical steps in de-differentiation. Indeed, both c-Myc and n-Myc bind the promoter of the mir-290 cluster, suggesting that one of the critical roles of Myc in reprogramming is to activate these miRNAs. However, there must be an intermediate step since the promoter of the mir-290 cluster is epigenetically silenced in fibroblasts (H3K27me3 positive, H3K4me3 negative). Importantly, the ESCC miRNAs do not recapitulate all the downstream affects of Myc. In particular, unlike Myc, the ESCC miRNAs do not promote proliferation of MEFs and they produce a homogenous population of fully reprogrammed iPSC colonies. Furthermore, teratomas from Myc-iPS, but not the ESCC miRNA-iPS cells, had a tendency to be invasive. Together, these findings show how microRNAs can play a central role in induced pluripotency.

7. Non-canonical miRNAs and endogenous siRNAs

While most miRNAs require both the Microprocessor and Dicer for their biogenesis, exceptions have been recently uncovered (Figure 2). For example, deep sequencing of small RNAs in Drosophila uncovered a novel subclass of miRNAs, the mirtrons (Okamura et al., 2007; Ruby et al., 2007). The mirtrons are generated from short introns that fold into pre-miRNA hairpins following debranching of the intron lariat. These mirtrons were discovered soon after in mammals (Babiarz et al., 2008; Berezikov et al., 2007). Babiarz and colleagues discovered the mirtrons in mouse ES cells by analyzing all small RNAs whose biogenesis relies on Dicer, but not the Microprocessor (Babiarz et al., 2008). However, the mirtrons made up a small fraction of small RNAs and as such, their specific role remains unclear. Most
of the Dicer-dependent, Dgcr8-independent small RNAs were derived from intergenic regions of the genome. The two most abundant of these miRNAs, miR-320 and miR-484, had been previously annotated and thought to be canonical miRNAs. Careful analysis of these two loci, as well as comparison to another related locus, suggest that miR-320 and miR-484 may be transcribed directly as short hairpin RNAs, similar to exogenously introduced short hairpins used experimentally by many labs (Dykxhoorn and Lieberman, 2005; Elbashir et al., 2001; Hannon and Rossi, 2004). However, it cannot be ruled out that there is a yet unknown nuclease(s) that act on the transcripts from these loci to produce the pre-miRNA. The Dgcr8-independence of these miRNAs was recently confirmed in mouse embryonic skin (Yi et al., 2009). These findings make it clear that there is a diversity of biogenic mechanisms in place in metazoan cells that can produce miRNAs. Whether these miRNAs act in the same manner as the canonical miRNAs is an open question.

Interestingly, both miR-320 and miR-484 are found upstream of genes and expressed in the antisense direction to those genes. Therefore, it was hypothesized they may be regulating the promoter regions of their respective genes. Indeed, in HeLa cells, where miR-320 levels are relatively high, H3K27me3 was found in the promoter of the adjacent gene. Additionally, the exogenous introduction of miR-320 into HEK-293 cells, which normally have undetectable miR-320, enhances H3K27me3 within the promoter of its adjacent gene, suppressing expression (Kim et al., 2008). However, analysis of genome-wide H3K27me3 in mouse ES cells, embryonic fibroblasts, and neural precursor cells (Mikkelsen et al., 2007) revealed no enrichment of H3K27me3 at this locus even though miR-320 is clearly present (Babiarz and Blelloch, unpublished observation). Therefore, it is unclear whether this role of miR-320 may be cell context dependent.

In addition to identifying Microprocessor-independent miRNAs, the deep sequencing of small RNAs in the different mutant background uncovered endogenous siRNAs (endo-siRNAs) in mES cells (Babiarz et al., 2008) (Figure 3). The presence of endo-siRNAs in mammalian cells was thought unlikely, because the long, double stranded precursors to endo-siRNAs would trigger the interferon response to dsRNAs (Stark et al., 1998). However, there are at least two known cell types where the interferon response does not occur: oocytes and ES cells (Wianny and Zernicka-Goetz, 2000; Yang et al., 2001). Indeed, along with the Babiarz et al. study in ES cells, two other reports identified mammalian endo-siRNAs, but in mouse oocytes (Tam et al., 2008; Watanabe et al., 2008). In the mouse oocytes, endo-siRNAs arose from two primary sources: retrotransposons, both LINE L1 and LTR elements, and

Figure 3. The endo-siRNA biogenesis pathway. Hairpin siRNAs are transcribed and fold into a long hairpin, which is cleaved multiple times by Dicer. cis-siRNAs are derived from transcription on both strands of DNA, resulting in a two-strand dsRNA with perfect complementarity. trans-siRNAs are derived from gene/pseudogene pairs from distant sites in the genome. Because of mutations in the pseudogene, complementarity of the dsRNA is not perfect. hp-siRNAs have been found in mouse oocytes and ES cells. trans-siRNAs and cis-siRNAs have only been found in oocytes. Adapted from Babiarz et al., 2008.
gene/pseudogene pairs. A loss of Dicer in oocytes results in an increase in retrotransposons that have endo-siRNAs associated with them and an increase in transcripts associated with the gene/pseudogene pairs (Tam et al., 2008; Watanabe et al., 2008). However, it is currently unclear if the increase in transcripts is due to a loss of Dicer degrading the dsRNAs, or if it is a secondary effect of siRNAs normally feeding back to destabilize their transcripts of origin.

In the mES cells, a very different population of endo-siRNAs was uncovered (Babiarz et al., 2008). The ES cells did not contain any endo-siRNAs arising from gene/pseudogene pairs; instead, the endo-siRNAs arose from SINE B1/Alu elements, most of which could be mapped back to two loci within the genome (Babiarz et al., 2008). These loci contained two tandem, inverted SINE B1/Alu elements, that when transcribed form a long hairpin. Dicer could then processively cleave this hairpin with an approximate 21 nucleotide (nt) periodicity, giving rise to a number of unique small RNAs. The role of these endo-siRNAs is unclear; the most straightforward mechanism of endo-siRNA function, degradation of SINE RNAs, does not appear to occur in ES cells. The loss of Dicer in mES cells does not result in an increase in SINE transcripts (Calabrese et al., 2007).

Mammalian endo-siRNAs are developmentally regulated. As described above, there was no overlap between the species of endo-siRNAs found in oocytes and ES cells (Babiarz et al., 2008). Whether the levels of the different endo-siRNAs are regulated transcriptionally and/or post-transcriptionally is currently unknown. It is possible that demethylation during germ cell development enable expression of the pseudogenes resulting in siRNA production which then play a critical role in the oocyte to suppress the function of the psuedogene’s partner. Then, as the psuedogenes are methylated and silenced during early embryogenesis, the gene partners are stabilized and translated. Similarly, the presence of LINE and LTR siRNAs in oocytes but not ES cells may represent patterns in their expression. Interestingly, LINE L1 and IAP mRNAs are increased in 

8. Conclusions

The study of small RNAs has been occurring at a fast pace with many exciting discoveries being uncovered. The importance of miRNAs in development is clear, with severe phenotypes associated with their loss (Bernstein et al., 2003; Kanellopoulou et al., 2005; Wang et al., 2007). However, the mechanism of action of miRNAs in mammalian development remains largely unknown and hotly pursued. Although a great deal has been learned about how miRNAs are generated, many questions remain unanswered. Chief among those is what are the precise roles of individual miRNAs in any particular cell. How are miRNA levels controlled? Are miRNAs in the same family redundant, or are there non-overlapping functions of the different family members. How central is the miRNA content of a cell in determining the cell’s identity? What is the developmental role of endogenous siRNAs? Tackling these and other questions will keep the miRNA field a productive area of study for years to come.

9. References


Small RNAs – their biogenesis, regulation and function in embryonic stem cells


Small RNAs – their biogenesis, regulation and function in embryonic stem cells


Small RNAs – their biogenesis, regulation and function in embryonic stem cells


Small RNAs – their biogenesis, regulation and function in embryonic stem cells


Small RNAs – their biogenesis, regulation and function in embryonic stem cells


Small RNAs – their biogenesis, regulation and function in embryonic stem cells


