# THE ROLE OF MICRORNAS IN EMBRYONIC STEM CELL DEVELOPMENT AND DIFFERENTIATION

**TAY MEI SIAN YVONNE**

**NATIONAL UNIVERSITY OF SINGAPORE**

**2008**

# THE ROLE OF MICRORNAS IN EMBRYONIC STEM CELL DEVELOPMENT AND DIFFERENTIATION

**TAY MEI SIAN YVONNE**

B.Sc. (Hons), NUS

**A THESIS SUBMITTED FOR THE DEGREE OF DOCTOR OF PHILOSOPHY**

NUS Graduate School for Integrative Sciences and Engineering

**NATIONAL UNIVERSITY OF SINGAPORE**

**2008**

**ACKNOWLEDGEMENTS**

First and foremost, I would like to express my deepest gratitude to my supervisor, A/P Bing Lim, for believing in me, providing me with an excellent environment to hone my technical expertise and intellectual prowess, and giving me the opportunity, independence and resources to conduct my research. His tremendous drive, intense results-oriented focus and passion for scientific discovery have been a great source of inspiration throughout these four years, and will be far into the future.

My sincere thanks and appreciation also go out to Dr Andrew Thomson and Dr Isidore Rigoutsos, for their thoughtful guidance and endless patience, and for keeping me sane. I wouldn't be here without you. To Wai Leong, Yen Sinn, Li Pin, Boon Seng, Huibin, Yin Loon, Minh, Sandy, Phil and all other past and present members of Bing's lab, thank you for your friendship, and for making the lab such a stimulating and fun place to work in.

I would also like to express heartfelt thanks to my parents for their unconditional love, support and encouragement, and to Mynn, for opening my eyes to the real world, showing me what love really is, and convincing me that nice guys do really exist, after all.

And last, but certainly not least, to God, for making all things possible in His time.

## TABLE OF CONTENTS


| | | |
|---|---|---|
| | Summary | i |
| | List of Tables | iii |
| | List of Figures | iv |
| | List of Symbols | vii |
| **1.** | **INTRODUCTION** | **1** |
| **1.1** | **Embryonic stem cells** | **1** |
| **1.1.1** | History | 1 |
| **1.1.2** | Properties & Potential | 2 |
| **1.1.3** | Maintaining Pluripotency: LIF, BMP & Wnt Signalling | 4 |
| **1.1.4** | Maintaining Pluripotency: Transcription Factors | 6 |
| **1.2** | **MicroRNAs** | **11** |
| **1.2.1** | Function | 11 |
| **1.2.2** | Identification | 13 |
| **1.2.3** | Biogenesis | 15 |
| **1.2.4** | Mechanism of action | 19 |
| **1.2.5** | Target prediction | 21 |
| **1.3** | **MicroRNAs in ESCs** | **24** |
| **2.** | **STATEMENT OF AIMS** | **29** |

| | | |
|---|---|---|
| **3.** | **MICRORNAS MODULATE MESC DIFFERENTIATION** | **30** |
| **3.1** | Introduction | 30 |
| **3.2** | Identification of microRNAs modulated during mESC differentiation | 32 |
| **3.3** | MicroRNAs can modulate *Oct4* and *Nanog* promoter activity | 35 |
| **3.4** | Expression profile of miR-134 during mESC differentiation | 40 |
| **3.5** | MiR-134 modulates mESC differentiation even in the presence of LIF | 41 |
| **3.6** | The mRNA expression patterns between RA-treated and miR-134-transfected mESCs demonstrate a high degree of correlation | 45 |
| **3.7** | MiR-134 enhances RA- and N2B27-mediated mESC differentiation | 48 |
| **3.8** | Discussion | 51 |
| **4.** | **MICRORNA TARGET PREDICTION** | **53** |
| **4.1** | Introduction | 53 |
| **4.2** | Validation of *rna22*, a microRNA target prediction algorithm | 54 |
| **4.3** | MiR-134 targets *Nanog* and *LRH1*, amongst other genes | 59 |
| **4.4** | Knockdown of miR-134 targets induces mESC differentiation | 65 |
| **4.5** | Discussion | 67 |
| **5.** | **MIR-134 IN NEURAL DEVELOPMENT** | **71** |
| **5.1** | Introduction | 71 |
| **5.2** | MiR-134 may target Chrdl and Dcx, amongst other genes | 72 |
| **5.3** | Expression profiling of miR-134 in embryos and adult tissues | 75 |
| **5.4** | Discussion | 78 |

| | | |
|---|---|---|
| **6.** | **MICRORNAS TARGETING OUTSIDE THE 3'UTR** | **80** |
| **6.1** | Introduction | 80 |
| **6.2** | MiR-296 targets the coding region of mouse Nanog | 81 |
| **6.3** | MiR-296 modulates mESC differentiation | 91 |
| **6.4** | MiR-134 targets the coding region of mouse Sox2 | 98 |
| **6.5** | Discussion | 103 |
| **7.** | **CONCLUSION** | **106** |
| **8.** | **MATERIALS AND METHODS** | **111** |
| **8.1** | **Cell culture, tissue preparation & cell-based assays** | **111** |
| **8.1.1** | Routine cell line maintenance | 111 |
| **8.1.2** | Differentiation of mESCs | 111 |
| **8.1.3** | Preparation of mouse tissues | 112 |
| **8.1.4** | Transfection | 112 |
| **8.1.5** | Immunostaining & Cellomics High Throughput Screening | 113 |
| **8.1.6** | *In situ* hybridization | 114 |
| **8.1.7** | Colony formation assay | 115 |
| **8.1.8** | pOct4/pNanog-Luciferase reporter assays | 115 |
| **8.1.9** | microRNA target validation assay | 116 |
| **8.2** | **DNA manipulation** | **117** |
| **8.2.1** | General DNA manipulation techniques | 117 |
| **8.2.2** | Construction of pOct4/pNanog-Luciferase reporters | 119 |
| **8.2.3** | Construction of microRNA overexpression plasmids | 120 |

| | | |
|---|---|---|
| **8.2.4** | Construction of pLuc-MRE plasmids | 122 |
| **8.2.5** | Construction of gene-specific RNAi plasmids | 123 |
| **8.2.6** | Construction of Nanog-CDS plasmids | 124 |
| **8.3** | **RNA and protein work** | **126** |
| **8.3.1** | RNA extraction & quantitative PCR | 126 |
| **8.3.2** | Northern blot | 127 |
| **8.3.3** | Microarray | 128 |
| **8.3.4** | Protein extraction & Western blot | 129 |
| **8.4** | **Statistical analysis** | **130** |
| **9.** | **BIBLIOGRAPHY** | **131** |
| **10.** | **APPENDICES** | **163** |
| **10.1** | Gene names & sequences of miR-375 target predictions tested | 163 |
| **10.2** | Gene names & sequences of miR-296 target predictions tested | 165 |
| **10.3** | Gene names & sequences of miR-134 target predictions tested | 167 |
| **10.4** | Luciferase results for predicted neural MREs | 175 |
| **10.5** | Summary of *rna22*'s predictions for four model genomes | 176 |
| **10.6** | Related publications | 177 |

**SUMMARY**

Hundreds of microRNAs are expressed in mammalian cells where they modulate gene expression by mediating transcript cleavage and/or regulation of translation. Functional studies to date have demonstrated that several of these microRNAs are important during development and disease. However, the role of microRNAs in the regulation of stem cell growth and differentiation is not well understood. It was shown, firstly, that microRNA (miR)-134 levels increase during retinoic acid- or N2B27-induced differentiation of mouse embryonic stem cells (mESCs). Secondly, elevation of miR-134 levels in mESCs enhances differentiation towards ectodermal lineages, an effect that is selectively blocked with a miR-134 antagonist. MiR-134's promotion of mESC differentiation is due, in part, to its direct translational attenuation of *Nanog*, *LRH1* and *Sox2*, known positive regulators of *Oct4/POU5F1* and mESC growth. Together, the data demonstrate that miR-134 alone can enhance the differentiation of mESCs to ectodermal lineages; additionally, they establish a functional role for miR-134 in modulating mESC differentiation through its potential to target and regulate multiple mRNAs.

Experimental validation of *rna22*, a method for identifying microRNA binding sites and their corresponding heteroduplexes, is presented. *rna22* does not rely upon cross-species conservation, is resilient to noise, and, unlike previous methods, it finds putative microRNA binding sites in the sequence of interest before identifying the targeting microRNA. In a luciferase reporter screen, average repressions of 30% or more for 168 of 226 tested 3'UTR targets are obtained. The analysis suggests that some microRNAs may have as many as a few thousand targets, and that between 74%

and 92% of the gene transcripts in four model genomes are likely under microRNA control.

Computational analyses by *rna22* suggests that fairly extensive microRNA regulation may be effected through the 5′ untranslated regions (UTRs) and coding sequences (CDSs) of gene transcripts in animals, in addition to 3′UTRs. To explore the possibility of microRNA targeting outside the 3′UTR of a transcript, two distinct, non-overlapping *rna22*-predicted targets for miR-296 in the CDS of Nanog were pursued experimentally. Reporter assays, quantitative PCR, and Western blot analyses demonstrated that miR-296 post-transcriptionally regulates Nanog by acting independently on each of these two binding sites. Silent mutations at these sites abolish Nanog's down-regulation by miR-296. To demonstrate that this is not an isolated incident of coding region targeting, similar experiments were performed to validate a single *rna22*-predicted target for miR-134 in the coding region of Sox2. Considered together, the results show that miR-296 and miR-134 repress the translation of Nanog and Sox2 mRNAs respectively via their interactions with specific CDS elements, and provide the first examples of animal microRNAs targeting genes in their coding regions.

The combined data imply that, by controlling specific genesets, microRNAs have a powerful influence on how mESCs sense and respond to their environment. This is further highlighted by the observation that each microRNA may potentially target hundreds or even thousands of genes. Additionally, the existing number of microRNAs, coupled with the continual discovery of novel microRNAs, suggests that they may be involved in many aspects of post-transcriptional regulation in stem cells.

## LIST OF TABLES


| | | |
|---|---|---|
| Table 1.1. | MicroRNA target prediction tools available when this study began. | 22 |
| Table 1.2. | Some of the microRNAs that are downregulated during ESC differentiation. | 26 |
| Table 1.3. | Some of the microRNAs that are upregulated during ESC differentiation. | 28 |
| Table 3.1. | Overexpression of miR-134 reduces the colony forming efficiency of mESCs. | 45 |

## LIST OF FIGURES


| | | |
|---|---|---|
| Figure 1.1. | Origin and differentiation potential of mESCs. | 2 |
| Figure 1.2. | MicroRNA biogenesis. | 17 |
| Figure 1.3. | Mechanism of microRNA action. | 19 |
| Figure 3.1. | MicroRNA expression levels change during EB differentiation of mESCs | 33 |
| Figure 3.2. | MicroRNA expression levels change during RA-induced mESC differentiation. | 34 |
| Figure 3.3. | Northern blot validation of microRNA microarray expression profiles. | 34 |
| Figure 3.4. | Schematic representations of pOct4-Luciferase and pNanog-Luciferase constructs. | 35 |
| Figure 3.5. | Luciferase assay demonstrating the efficacy of Anti-miRs and Pre-miRs. | 37 |
| Figure 3.6. | Overexpression of miR-134 downregulates Oct4 and Nanog promoter activities. | 39 |
| Figure 3.7. | Expression profile of miR-134 during mESC differentiation. | 40 |
| Figure 3.8. | miR-134 modulates the transcript levels of lineage-specific biomarkers, even in the presence of LIF. | 42 |
| Figure 3.9. | miR-134 downregulates protein levels of pluripotency markers and induces changes in mESC morphology indicative of differentiation, even in the presence of LIF. | 44 |
| Figure 3.10. | miR-134 induces a subset of genes similar to that induced by RA treatment of mESCs. | 47 |
| Figure 3.11. | miR-134 enhances the effect of RA on mESCs. | 48 |
| Figure 3.12. | miR-134 enhances the effect of N2B27 medium on mESCs. | 50 |
| Figure 4.1 | Flowchart depicting the various steps of the target prediction method used. | 54 |
| Figure 4.2. | Schematic representation of pLuc-MRE plasmid reporter construct. | 55 |

| | | |
|---|---|---|
| Figure 4.3. | Luciferase-based validation of predicted targets for miR-375 and miR-296. | 57 |
| Figure 4.4. | Luciferase-based validation of predicted targets for miR-134. | 58 |
| Figure 4.5. | Expression analysis of all *rna22*-predicted miR-134 target genes. | 59 |
| Figure 4.6 | *LRH1*, *FADD*, $G\alpha_o$ and *Nanog* are potential targets of miR-134. | 62 |
| Figure 4.7. | miR-134 reduces the protein levels of predicted pluripotency-associated targets *LRH1*, $G\alpha_o$ and *Nanog* without altering their mRNA levels. | 64 |
| Figure 4.8. | Knockdown of *Nanog*, *LRH1* and $G\alpha_o$ results in differentiation of mESCs. | 66 |
| Figure 5.1. | *BMP8b, Chrdl1, Dcx, Dtx4 and Hoxc10* are potential targets of miR-134. | 74 |
| Figure 5.2. | miR-134 expression increases during embryogenesis and is present in several adult tissues. | 76 |
| Figure 5.3. | Distribution of miR-134 expression in the E11.5 embryo. | 77 |
| Figure 6.1. | Nucleotide sequence of Nanog's CDS region, codons and the corresponding amino acid translation. | 82 |
| Figure 6.2. | Nanog-235p and Nanog-493p are potential targets of miR-296. | 84 |
| Figure 6.3. | Transfection of PmiR-296 reduces the amount of endogenous Nanog protein in mESCs and the amount of exogenous Nanog protein in 293T cells. | 85 |
| Figure 6.4. | Selection of 293T cells as a suitable cell line to study exogenous Nanog. | 86 |
| Figure 6.5. | Several MRE mutants are able to rescue the miR-296 induced reduction in luciferase activity of Nanog-235p and Nanog-493p. | 89 |
| Figure 6.6. | The 235-m4/493-m2 double mutant is able to rescue the miR-296 induced reduction in Nanog protein levels. | 90 |
| Figure 6.7. | miR-296 is upregulated during mESC differentiation, and reduces the alkaline phosphatase activity and colony forming efficiency of mESCs. | 92 |
| Figure 6.8. | miR-296 modulates the transcript levels of lineage-specific biomarkers in mESCs. | 94 |

| | | |
|---|---|---|
| Figure 6.9. | 235-m4/493-m2 double mutants are able to rescue miR-296-modulated mESC differentiation. | 97 |
| Figure 6.10. | Sox2-637p is a potential target of miR-134. | 99 |
| Figure 6.11. | The 637-m4 mutant is able to rescue the miR-134 induced reduction in Sox2 protein levels. | 102 |
| Figure 8.1. | Vector map of pIRES2-EGFP. | 120 |
| Figure 8.2. | Vector map of pLL3.7. | 121 |
| Figure 8.3. | Vector map of psiCHECK-2. | 122 |
| Figure 8.4. | Vector map of pSuper.puro. | 123 |
| Figure 8.5. | Vector map of pcDNA3.1. | 125 |

## LIST OF SYMBOLS

| | |
|---|---|
| Ago | Argonaute |
| Anti-miR | Anti-miR microRNA inhibitor |
| AP | Alkaline phosphatase |
| ARE | AU-rich element |
| bHLH | basic helix-loop-helix |
| BIO | 6-bromoindirubin-3V-oxime |
| BMP | Bone morphogenetic protein |
| CDS | Coding sequence |
| CHIP | Chromatin immunoprecipitation |
| CTL | Control |
| D | Day |
| DE | Distal enhancer |
| DGCR8 | DiGeorge syndrome critical region gene 8 |
| DPBS | Dulbecco's phospate-buffered saline |
| dsRBD | double-stranded RNA-binding domain |
| EB | Embryoid body |
| EC | Embryonic carcinoma |
| EEmiRC | Early embryonic microRNA cluster |
| ESC | Embryonic stem cell |
| FBS | Fetal bovine serum |
| FXR1 | Fragile X mental retardation-related protein 1 |
| GSK-3 | Glycogen synthase kinase-3 |
| HCS | High content screening |
| HDACs | Histone deacetylases |
| hESC | Human embryonic stem cell |
| ICM | Inner cell mass |
| Id | Inhibitor of differentiation |
| ISH | *In situ* hybridization |
| JAK | Janus-associated tyrosine kinase |
| LIF | Leukemia inhibitory factor |
| Luc | Luciferase |
| LIFR | LIF receptor |
| MCS | Multiple cloning site |
| MECPs | Methyl-CpG-binding proteins |
| mESC | Mouse embryonic stem cell |
| MHB | Midbrain-hindbrain isthmus/boundary |
| miR | MicroRNA |
| miRNA | MicroRNA |
| miRISCs | MicroRNA-containing RNA-induced silencing complexes |
| MRE | MicroRNA response element |
| mRNA | Messenger RNA |
| MT | Mock transfection |
| Mtpn | Myotrophin |
| Mut | Mutant |
| Non-Sil | Non-silencing |
| NP | Neural progenitor |
| nt | Nucleotide |
| ORF | Open reading frame |

| | |
|---|---|
| PAZ | Piwi Argonaute Zwille |
| PcG | Polycomb group protein |
| PE | Proximal enhancer |
| PGR | Photogenerated reagent |
| Pol | Polymerase |
| POU | Pit, Oct, Unc |
| PP | Proximal promoter |
| pre-microRNA | Precursor microRNA |
| Pre-miR | Pre-miR microRNA precursor |
| pri-microRNAs | Primary microRNA |
| PTW | PBS + 0.1% Tween-20 |
| RA | Retinoic acid |
| RBP | RNA-binding protein |
| RC | Reverse complement |
| RIIID | RNase III domain |
| RISC | RNA-induced silencing complex |
| RNA | Ribonucleic acid |
| RNAi | RNA interference |
| Scr | Scrambled oligomer |
| SHP2 | Src homology 2 |
| shRNA | Short hairpin RNA |
| SMAD | Mothers against dpp related |
| STAT | Signal transducer and activator of transduction |
| TGF | Transforming growth factor |
| TSS | Transcription start site |
| UTR | Untranslated region |
| WT | Wild-type |

# CHAPTER 1. INTRODUCTION

## 1.1 Embryonic stem cells

### 1.1.1 History

In the 1970s, the search for a cell culture platform to study early embryonic development led to the isolation of stem cells from teratocarcinomas. Teratocarcinomas are malignant gonadal tumors consisting of differentiated cell types from the three embryonic germ layers (endoderm, mesoderm and ectoderm), as well as a significant population of undifferentiated cells, termed embryonic carcinoma (EC) cells, which resemble early embryonic cells (Martin and Evans, 1975). EC cells could be expanded continuously in culture while retaining the capacity to differentiate into derivatives of all three germ layers (Kleinsmith & Pierce 1964, Martin & Evans 1975). However, these cancer-derived EC cells have an aneuploid karyotype (Martin, 1980), possibly due to uncontrolled selection pressures during tumour growth, and are thus incapable of undergoing meiosis to produce mature gametes (Smith, 2001).

Nevertheless, studies with EC cells were of vital importance in establishing the technical expertise necessary for the derivation of embryonic stem cells (ESCs) (Evans and Kaufman, 1981; Martin, 1981; Stevens, 1970; Stevens *et al.*, 1977; Stevens LC, 1978). A crucial insight was the discovery that EC cells thrived and maintained a high differentiation capacity when co-cultured with mitotically inactivated embryonic fibroblast cells (Martin *et al*. 1977; Martin & Evans, 1975), but did poorly when cultured in isolation. As these fibroblasts appeared to be providing some essential nutrient or trophic factor, they were described as feeder cells (Friel *et al.*, 2005; Smith, 2001).

This discovery was instrumental in enabling the successful isolation and culture of ESCs from mouse blastocysts, described by two groups of scientists in 1981 (Evans and Kaufman, 1981; Martin, 1981). Embryos at the expanded blastocyst stage are first plated, either intact or after immunosurgical isolation of the inner cell mass (ICM), onto a layer of feeder cells (Smith, 2001). The mass of cells is dissociated and replated onto a fresh feeder layer several days later. Along with various types of differentiated colonies, colonies with a characteristic undifferentiated morphology arise that are individually dissociated, replated and expanded to establish ESC lines (Figure 1.1) (Robertson, 1987).

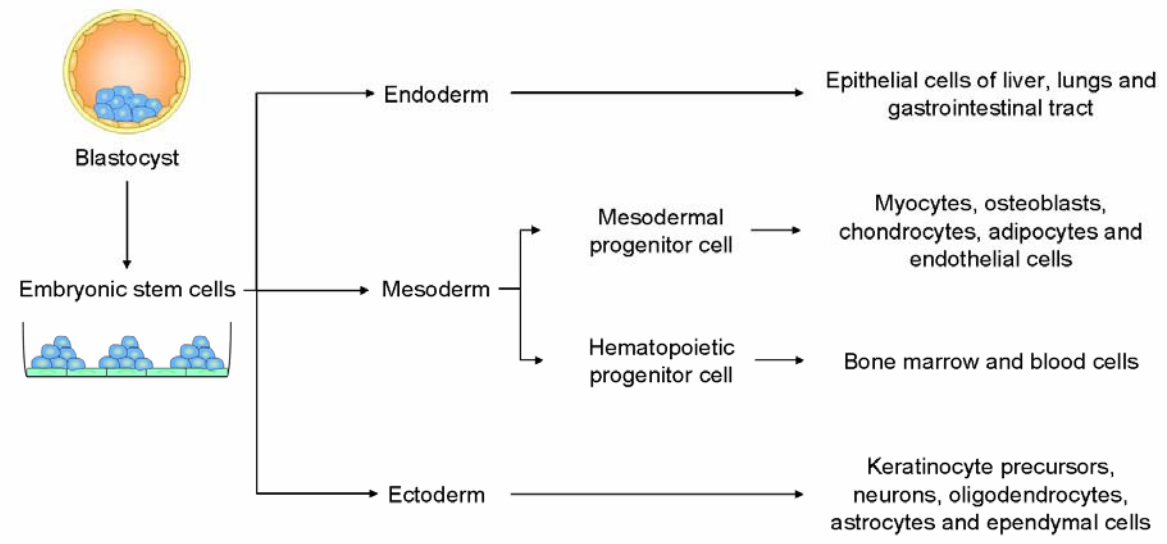



**Figure 1.1. Origin and differentiation potential of mESCs.**

### 1.1.2 Properties & Potential

ESCs are pluripotent, ie. they possess the dual properties of unlimited self-renewal without senescence and the ability to differentiate into cell types of all three germ layers *in vitro*. Furthermore, tumours generated from ESCs contain endodermal,

ectodermal and mesodermal tissue and cell types (Evans and Kaufman, 1983); and ESCs (unlike EC cells) are able to participate fully in fetal development when reintroduced into an embryo (Smith, 2001).

The drive to develop ESC-based systems stems from this potential of ESCs to differentiate into all cell types in the body. Although the clinical use of adult stem cells, which are present in multiple tissues in the mammalian body, is attractive due to the lack of allogenecity, adult stem cells are only able to differentiate into multiple cell types of a specific tissue, organ or physiological system (Erlandsson and Morshead, 2006; Mimeault and Batra, 2006; Serakinci and Keith, 2006). Aside from their limited differentiation potential, adult stem cells are also unable to self-renew indefinitely in culture and can be difficult to isolate (Erlandsson and Morshead, 2006; Mimeault and Batra, 2006; Serakinci and Keith, 2006). These properties limit their use as a scaleable, continous resource for generating multiple cell types for cell-based therapies.

ESCs have been instrumental in enabling groundbreaking research into drug discovery facilitated by high throughput screening, and hold much promise for cell-based therapy to treat a whole spectrum of degenerative diseases and injuries. Although the more recently-derived human ESCs (hESCs) are undisputedly a better disease model than mESCs, at the time this project began, the advantages of using mESCs as a model to elucidate the role and mechanisms of microRNAs in modulating pluripotency and differentiation outweighed those of hESCs. For example, mESCs can be used to generate gene knockouts which can be reintroduced into embryos to better elucidate the role of these genes of interest in development. They also have a

shorter doubling time, are more stable karyotypically, and do not require a feeder layer for indefinite self-renewal in culture.

### 1.1.3 Maintaining Pluripotency: LIF, BMP & Wnt Signalling

One of the major breakthroughs in ESC maintenance was reported in 1988, when leukemia inhibitory factor (LIF), a member of the IL-6 cytokine family, was identified as a major factor which enabled mESCs (mESCs) to self-renew indefinitely in an undifferentiated state without a feeder layer (Smith *et al.*, 1988, Williams *et al.*, 1988). This enabled the feeder-free culture of mESCs in growth medium supplemented with serum and recombinant LIF. LIF stimulates mESCs by binding to a heterodimeric LIF receptor (LIFR)-gp130 signaling complex that activates two major signaling pathways, the canonical JAK-STAT (Janus-associated tyrosine kinase, signal transducer and activator of transduction) pathway and the Src homology 2 (SHP2)-Erk pathway (Rao, 2004).

Activation of the JAK-STAT pathway results in JAK-mediated phosphorylation of STAT3, leading to the formation of homodimers which subsequently translocate to the nucleus where they regulates transcription of genes involved in the self-renewal of ESCs (Niwa *et al.*, 1998). This activation of STAT3 is necessary for LIF to maintain the self-renewal of mESCs (Niwa *et al*., 1998), and it alone is sufficient to prolong mESC self-renewal in the absence of LIF (Matsuda *et al*., 1999). Conversely, LIF-induced activation of the SHP2-Erk pathway in mESCs is a promoter of differentiation, and may be a negative regulatory mechanism for STAT3 (Burdon *et al.*, 1999; Cheng *et al.*, 1998; Liu *et al.*, 2006). In this context, the balance between

LIF-induced STAT3 activation and ERK signaling is a critical modulator of mESC self-renewal (Burdon *et al.*, 1999). Intriguingly, LIF signaling is not required for the maintenance of hESC pluripotency. The report by Nichols *et al.* that LIFR gp130 -/- mouse embryos can develop and be used to establish ESC lines, coupled with the observation that LIF is unable to sustain mESCs in the absence of serum, suggests that other pathways may play a role in maintaining mESC pluripotency (Nichols *et al.*, 2001; Pan and Thomson, 2007).

The factor present in serum which is essential for mESC self-renewal is likely to be bone morphogenetic protein (BMP), a member of the transforming growth factor (TGF)-β superfamily (Liu *et al.*, 2006). This is supported by a study by Ying *et al.* which demonstrated that successful derivation and maintenance of ESC lines was possible in serum-free medium supplemented with LIF and BMP (Wilson *et al.*, 1995; Ying *et al.*, 2003). BMP signaling activates cytoplasmic proteins called SMADs (mothers against dpp related), which subsequently induce the expression of Id (inhibitor of differentiation) genes. Id proteins antagonize neurogenic basic Helix-Loop-Helix (bHLH) transcription factors and block neural differentiation of ESCs (Ying *et al.*, 2003). Expression of Id1, Id2 or Id3 is able to compensate for the presence of BMP in mESC cultures supplemented with LIF (Ying *et al.*, 2003).

In addition to LIF and BMP signaling, recent studies have postulated a role for the Wnt pathway in the maintenance of ESC pluripotency. Wnt signaling is endogenously activated in mESCs, and is downregulated during differentiation (Sato *et al.*, 2004). Sato *et al.* show that Wnt pathway activation by a specific pharmacological inhibitor of glycogen synthase kinase-3 (GSK-3), 6-bromoindirubin-3V-oxime (BIO), is able to

maintain ESCs in an undifferentiated state and sustain expression of the pluripotency markers Oct4, Rex1 and Nanog (Sato *et al.*, 2004). Taken together, the contributions of the LIF, BMP and Wnt pathways to maintaining mESC pluripotency are suggestive of a complex network of interactions that can control the growth or differentiation of mESCs.

### 1.1.4 Maintaining Pluripotency: Transcription Factors

External signaling pathways such as the abovementioned LIF, BMP and Wnt eventually lead to the regulation of several genes that are critical for the maintenance of mESC self-renewal and pluripotency, such as Oct4, Sox2 and Nanog. The direct activation of these transcription factors have been found to influence ESC growth and differentiation.

The Pit, Oct, Unc (POU)-domain transcription factor Oct4 (also known as Oct3), which is encoded by *Pou5f1*, is an important regulator of pluripotency *in vivo* (Pan *et al.*, 2002). Oct4 expression, which begins at the four-cell stage during mouse embryogenesis, is restricted to totipotent and pluripotent cells and is downregulated in most adult tissues except the germ line (Pesce *et al.*, 1998; Yeom *et al.*, 1996). Mouse embryos lacking *Oct4* do not have pluripotent ICM and thus cannot develop past the blastocyst stage (Nichols *et al.*, 1998). This suggests that Oct4 is an essential modulator of pluripotency *in vivo*, and that it plays an important role in differentiation.

Oct4 also acts as a gatekeeper for *in vitro* ESC pluripotency at the crossroads between self-renewal and lineage specification (Nichols *et al.*, 1998; Stefanovic & Puceat, 2007). Its expression is high in undifferentiated mESCs, and decreases during differentiation (Pan and Thomson, 2007). Precise levels of Oct4 are required for the maintenance of pluripotent ESCs: Reduction of Oct4 expression to 50% or less induces trophectodermal differentiation, while overexpression causes differentiation to primitive endoderm and mesoderm (Yeom *et al.*, 1996; Niwa, 2001; Niwa *et al.*, 2000).

Oct4 binds to the DNA octamer motif ATGCAAAT in the promoter or enhancer regions of numerous ESC-specific genes to regulate their transcription (Puceat, Han). Target genes of Oct4 that have been identified thus far include *Fgf4*, *Hand1*, *Utf1*, *Opn*, *Rex1*/ *Zfp42*, *Fbx15*, and *Sox2* (Nishimoto *et al*., 1999; Tomioka *et al*., 2002; Zeng *et al*., 2004). Oct4 can cooperate with other transcription factors to activate or repress target genes (Guo *et al.*, 2002; Yuan *et al.*, 1995). One such transcription factor is the SRY-related HMG family member, Sox2 (Avilion *et al.*, 2003).

Sox2 is abundantly expressed in mESCs, where its knockdown induces differentiation into multiple lineages (Ivanova *et al.*, 2006). Interestingly, Sox2 expression *in vivo* differs to that of Oct4, where its earliest detection is at the morula stage (E2.5), continutes in the ICM (E3.5), epiblast (E6.5), extraembryonic lineages and throughout the neural plate; before it becomes restricted to stem cells, neural, gut and germ cells (Graham *et al.*, 2003; O'Shea, 2004). In addition, *Sox2*-null embryos die immediately after implantation (Avilion *et al.*, 2003). These data demonstrate that, similar to Oct4,

Sox2 plays an important role during development *in vivo* as well as in mESC pluripotency and differentiation.

However, the expression of Oct4 alone does not prevent ESC differentiation in the absence of LIF, suggesting that other factors may be important regulators of ESC pluripotency (Pan and Thomson, 2007). In 2003, a novel factor that is instrumental for maintaining ESC pluripotency was identified (Chambers *et al*., 2003; Mitsui *et al*., 2003). Nanog, a homeobox transcription factor named after the mythical land of the ever young Tir Na Nog, is expressed in mouse ES, EC and embryonic germ cells, is downregulated during differentiation, and is not expressed in adult tissues or differentiated cells (Chambers *et al.*, 2003, Mitsui *et al.*, 2003). Mouse ESCs overexpressing Nanog are able to self-renew in the absence of LIF, which suggests that Nanog may be a major regulator of pluripotency (Chambers *et al.*, 2003, Mitsui *et al.*, 2003). These two groups also found that although Nanog acts in concert with LIF, it does not modulate either the LIF or BMP signaling pathways. In addition, disruption of Nanog in ESCs causes differentiation into Gata-6 positive parietal endoderm-like cells (Mitsui *et al.*, 2003).

Nanog is also a critical regulator of cell fate *in vivo*: ICM cells in Nanog-null mice spontaneously differentiate into visceral and parietal endoderm (Mitsui *et al.*, 2003). Nanog expression can be detected in the morula, ICM, early germ cells and proximal epiblast at the location of the future primitive streak (Hart *et al.*, 2004; Mitsui *et al.*, 2003).

Genome-wide chromatin immunoprecipitation (CHIP), microarray expression profiling and RNA interference assays have identified numerous target genes of Nanog, Oct4 and Sox2 (Boyer *et al*., 2005, 2006; Ivanova *et al*., 2006; Loh *et al*., 2006; Rao and Orkin, 2006). In mESCs, Loh *et al.* describe 1083 and 3006 binding sites for Oct4 and Nanog respectively, with substantial overlap between the two gene sets. The core downstream targets include genes related to pluripotency, self-renewal and cell fate determination such as Oct4, Sox2 and FoxD3 (Loh *et al.*, 2006). Oct4, Nanog and Sox2 appear to regulate themselves, and each other, and form a transcriptional regulatory feedback circuit that is essential for the maintenance of ESC pluripotency (Boyer *et al*., 2005, 2006; Ivanova *et al*., 2006; Loh *et al*., 2006; Rao and Orkin, 2006).

Although this transcriptional regulatory network is crucial for keeping ESCs in a undifferentiated state, other factors may also be important for the maintenance of pluripotency. One such example may be epigenetic processes such as the modification of DNA, histones or chromatin structure, as transcription factor activity is dependent on the accessibility of target genes (Jaenisch and Bird, 2003; Niwa *et al*., 2000; Mitsui *et al*., 2003; Chambers *et al*., 2003; Boyer *et al*., 2005; Niwa *et al*., 2005; Boyer *et al*., 2006; Meshorer and Misteli, 2006). Chromatin modification factors such as histone deacetylases (HDACs), methyl-CpG-binding proteins (MECPs) and polycomb group proteins (PcG) are differentially expressed as ESCs differentiate, and may be crucial modulators of self-renewal and differentiation (Rao, 2004).

Another factor which may be involved in maintaining the pluripotent state of ESCs is a class of recently-discovered small non-coding RNAs, microRNAs, which have been

shown to play vital roles in gene regulation (Bartel, 2004). Genome-wide CHIP analyses in mESCs by Loh *et al.* demonstrated that Nanog bound to sites within 30 kb of 5 microRNA genes, and that Oct4 and Nanog co-occupied sites near 2 of these genes (Loh *et al.*, 2006). Boyer *et al.* showed that Oct4, Nanog and Sox2 were associated with 14 microRNA genes in hESCs, and co-occupied the promoters of 2 of these genes (Boyer *et al.*, 2005). These results imply that microRNA genes are likely to be regulated by Oct4, Sox2 and Nanog in both mESCs and hESCs, and may thus be important regulators of pluripotency and self-renewal. Furthermore, the network of regulatory interactions that exists in ESCs as suggested by the CHIP data offers the intriguing possibility that these transcription factors may in turn be regulated by microRNAs, adding to the complexity of environmental sensing and gene regulation controlling growth and differentiation in ESCs.

## 1.2 MicroRNAs

### 1.2.1 Function

MicroRNAs, a family of small (~22 nucleotides long), noncoding RNAs similar to the siRNAs involved in RNA silencing, originate from stem-loop precursors in the genome. They have been shown to play important roles in diverse processes including apoptosis, fat metabolism, cancer, major signaling pathways, tissue morphogenesis and development.

For example, Bantam and miR-14 have been implicated in programmed cell death in *Drosophila*. Bantam inhibits apoptosis by regulating the proapoptotic gene hid (Brennecke *et al*., 2003), while miR-14 suppressESC death by acting on a yet unknown cellular target (Xu *et al*., 2003). Intriguingly, miR-14 mutants are also phenotypically obese with elevated levels of triacylglycerol. This suggests that miR-14 may be involved in fat metabolism (Xu *et al*., 2003).

MicroRNA expression signatures are associated consistently with several types of cancers and cancer cell lines (McManus, 2003; Metzler *et al.*, 2004; Takamizawa *et al.*, 2004; Lu *et al.*, 2005; Miska, 2005). Calin *et al.* demonstrated that the region on chromosome 13q14 containing miR-15 and miR-16 is deleted in the majority of chronic lymphocytic leukemia cases (Calin *et al*., 2002). Moreover, microRNA expression profiling in cancer patients has potential prognostic value as expression levels of miR-155 in B cell lymphoma patients and let-7 in lung cancer are indicative of patient survival (Kloosterman and Plasterk, 2006).

A number of microRNAs exhibit distinct spatial and temporal expression patterns during development (Aboobaker *et al*., 2005; Ason *et al*., 2006; Kloosterman *et al*., 2006; Wienholds *et al*., 2005). Additionally, some microRNA expression patterns show species conservation, eg. miR-1 in muscles, miR-124 in the central nervous system and miR-10 in anterior-posterior patterning (Kloosterman and Plasterk, 2006). These observations indicate that microRNAs may be involved in the specification and maintenance of tissue identity and other facets of development. This is supported by studies which show that animals without mature microRNAs are not viable, eg. Dicer-deficient mice die at embryonic day 7.5 and lack multipotent stem cells (Bernstein *et al*., 2003; Ketting *et al*., 2001; Wienholds *et al*., 2003).

Studies in invertebrate model systems have identified lsy-6, the first microRNA found to play a role in neuronal patterning (Johnston and Hobert, 2003), and miR-9a, which ensures the generation of the precise number of neuronal precursor cells during development (Li *et al.*, 2006). In vertebrate models, the restoration of a single microRNA (miR-430) in zebrafish modified to prevent production of endogenous microRNAs ameliorated deficits in neuroectodermal development and neuronal differentiation (Giraldez *et al.*, 2005). MicroRNA regulation of *Hox* expression modulates developmental patterning processes to allow the generation of asymmetric morphology (Mansfield *et al.*, 2004; Yekta *et al.*, 2004).

In mammals, specific microRNAs have been shown to regulate B cell differentiation (Chen *et al.*, 2004), adipocyte differentiation (Esau *et al.*, 2004), and insulin secretion (Poy *et al.*, 2004). MicroRNAs have also been found to play key roles during neural differentiation *in vitro* (Krichevsky *et al.*, 2006; Smirnova *et al.*, 2005), and in

vertebrate central nervous system development (Giraldez *et al.*, 2005; Krichevsky *et al.*, 2004; Miska *et al.*, 2004). In particular, miR-134 has been recently identified as a potential regulator of dendritic spine volume and synapse formation in mature rat hippocampal neurons *in vitro* through the localized repression of a protein kinase LimK1 (Schratt *et al.*, 2006). The mouse homologue of miR-134, which demonstrates conservation across rodents and primates, was originally identified by cloning from the mouse cortex (Lagos-Quintana *et al.*, 2002) and is located in a large imprinted microRNA gene cluster at the mouse Dlk1-Gtl2 domain (Seitz *et al.*, 2004).

Thus far, microRNAs have been found in diverse species including *Arabidopsis thaliana, Caenorhabditis elegans, Drosophila melanogaster, Danio rerio, Mus musculus, Homo sapiens* and even the Epstein Barr virus (miRBase; Griffiths-Jones, 2006). As of December 2007, miRBase, a searchable database of published microRNA sequences and annotations, contained a total of 5395 entries. New microRNAs which have been identified and validated will be added to this repository.

### 1.2.2 Identification

Thus far, three approaches have been used successfully for microRNA discovery: forward genetics, the cloning and sequencing of size-fractionated cDNA libraries, and computational prediction. The first microRNA genes, *lin-4* and *let-7*, were identified using forward genetics. *Lin-4* was found to be responsible for a defective cell lineage in a *Caenorhabditis elegans* mutant (Lee *et al.*, 1993). It had a large ~ 60 nt form which folded into a hairpin structure, and a small ~ 22 nt form which originated from the stem of the hairpin and repressed *lin-14* gene expression via imperfect pairing

with its 3' untranslated region (UTR) (Wightman *et al.*, 1993). The discovery of *let-7* in *C. elegans* generated considerable interest in the microRNA field as it was conserved among a diverse range of phylogenetic taxa (Pasquinelli *et al.*, 2000; Reinhart *et al.*, 2000; Slack *et al.*, 2000). This suggested that gene regulation by microRNAs may be more widespread and pervasive than previously thought. Four other microRNAs, *bantam*, *miR-14*, *miR-278* and *lsy-6*, have also been identified by forward genetics (Brennecke *et al.*, 2003; Johnston and Hobert, 2003; Teleman *et al.*, 2006; Xu *et al.*, 2003). However, as a result of factors such as the small size of microRNAs, their tolerance to mutations and redundancy, forward genetics is a relatively inefficient method of discovering microRNAs (Abbott *et al.*, 2005).

Another approach useful for microRNA identification is the sequencing of size-fractionated cDNA libraries. This protocol, which was originally used to clone small interfering RNA molecules (Elbashir *et al.*, 2001), has been adapted by various groups for the successful identification of the majority of the microRNAs known today. Briefly, following size-fractionation of an RNA sample in a denaturing polyacrylamide gel, 5' and 3' adapters are added to the 20-25 nt fraction. RT-PCR is performed next, followed by the optional concatamerization of cDNAs into large fragments which increases the amount of sequence information obtainable (Berezikov *et al.*, 2006, Lagos-Quintana *et al.*, 2001; Lau *et al.*, 2001; Pfeffer *et al.*, 2003). These fragments are then cloned into vectors, sequenced and analyzed. However, microRNAs that have low, temporal or cell-type specific expression levels, and microRNAs that have specific sequence composition or post-transcriptional modifications may not be detected using this method (Luciano *et al.*, 2004; Yang *et al.*, 2006; Yang *et al.*, 2006).

These cloning approaches provided enough information for scientists to recognize several distinctive properties of microRNAs, and begin to develop computer algorithms for microRNA prediction (Bentwich, 2005; Berezikov and Plasterk, 2005). All microRNA prediction approaches use secondary structure information as the hairpin loop is a defining microRNA characteristic. They also rely on one or more of the following: (1) Phylogenetic conservation of sequence and structure, (2) Thermodynamic stability of hairpins, (3) Similarity to known microRNAs in terms of sequence and structure, (4) Genomic location, as many microRNAs are found in clusters or in close proximity (Altuvia *et al.*, 2005; Berezikov *et al.*, 2006; Lau *et al.*, 2001; Seitz *et al.*, 2004). All predicted candidate microRNAs need to be validated experimentally. Expression of the ~22 nt long mature microRNAs can be demonstrated using techniques such as northern blot analysis, primer extension, microRNA QUANTITATIVE PCR and/or *in situ* hybridization.

### 1.2.3 Biogenesis

Most mammalian microRNA genes (~70%) are found in defined transcription units, with about half of known microRNAs located in close proximity to other microRNAs (Lagos-Quintana *et al.*, 2001; Lau *et al.*, 2001; Mourelatos *et al.*, 2002; Rodriguez *et al.*, 2004). Many microRNAs are in the introns of protein-coding transcription units, while others are in the introns or exons of non-coding transcription units (Kim, 2005; Rodriguez *et al.*, 2004). Although RNA polymerase (Pol) III was initially thought to mediate microRNA transcription as it transcribes most small RNAs, some microRNA precursors contain stretches of five or more uracils, which is a termination sequence

for Pol III (Lee *et al.*, 2002). Increasing evidence suggests that microRNA gene transcription is mediated mainly by RNA Pol II: (1) Some microRNA precursors contain both cap structures and poly(A) tails; (2) MicroRNA transcription activity demonstrates sensitivity to conditions that specifically inhibit Pol II and not Pol I or III; (3) CHIP analyses show the physical association of Pol II with a microRNA promoter (Cai *et al.*, 2004; Lee *et al.*, 2004).

MicroRNA transcription produces primary microRNA transcripts (pri-microRNAs), which contain a hairpin structure and may be up to several kilobases in length (Figure 1.2) The nuclear RNase III enzyme Drosha cleaves the stem-loop to release precursor microRNAs (pre-microRNAs) (Lee *et al.*, 2003). Drosha, a large protein which is evolutionarily conserved in animals, contains two RNase III domains (RIIIDs) and a double-stranded RNA-binding domain (dsRBD) that are essential for its function (Filippov *et al.*, 2000; Fortin *et al.*, 2002; Han *et al.*, 2004; Wu *et al.*, 2000). Drosha forms a complex with the double-stranded-RNA-binding protein DiGeorge syndrome critical region gene 8 (DGCR8, also known as Pasha in *D. melanogaster* and *C. elegans*) (Denli *et al.*, 2004; Han *et al.*, 2004; Gregory *et al.*, 2004; Landthaler *et al.*, 2004). Gregory *et al.* have shown via knock-down *in vivo* and reconstitution *in vitro* studies that this Microprocessor complex is necessary and sufficient for the genesis of microRNAs from pri-microRNAs (Gregory *et al.*, 2004). Drosha cleavage creates a short ~2 nucleotide overhang at the 3' end, which is recognized by the downstream biogenesis factors and also generates one end of the mature microRNA (Lee *et al.*, 2003; Lund *et al.*, 2004).

After Drosha processing, pre-microRNAs are exported out of the nucleus through nuclear pore complexes (Nakielny *et al.*, 1999). Pre-microRNA export is mediated by exportin-5, a nuclear transport receptor, in a process requiring the hydrolysis of GTP to GDP (Bohnsack *et al.*, 2004; Lund *et al.*, 2004; Yi *et al.*, 2003). Cullen *et al.* utilized mutational analyses to demonstrate that an RNA stem of more than 16 base pairs and a short 3' overhang are important requirements for the export of pre-microRNAs (Zeng and Cullen, 2004).

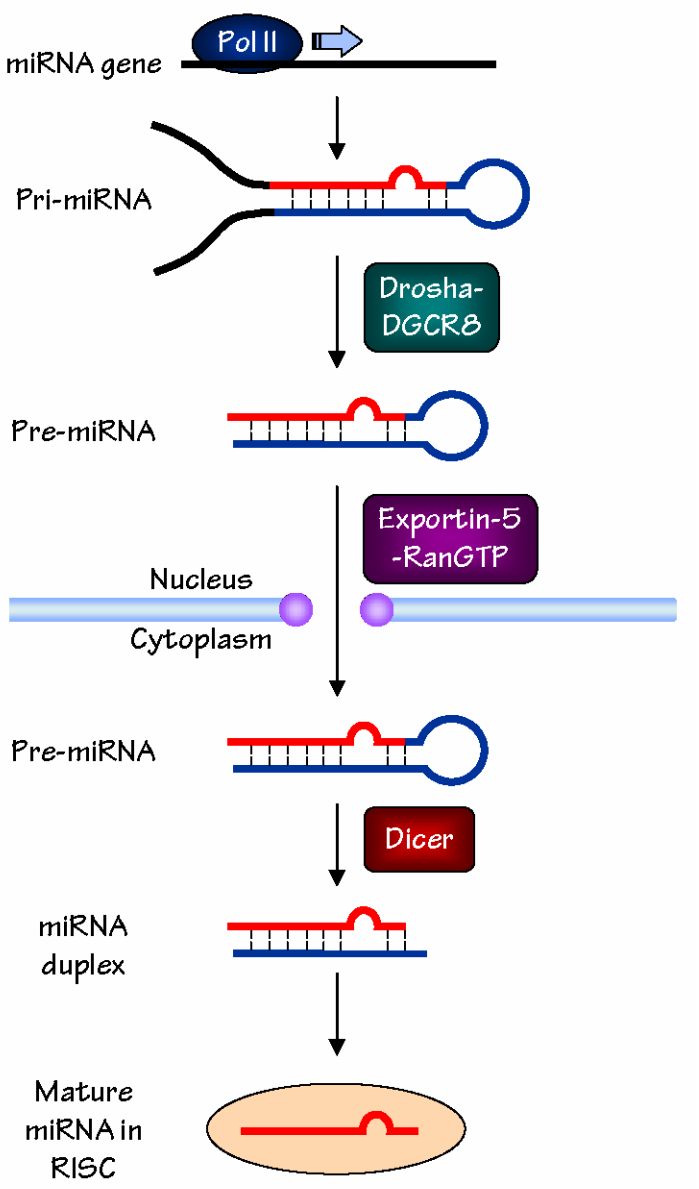



**Figure 1.2. MicroRNA biogenesis.** See text for details.

Pre-microRNAs are next cleaved by the cytoplasmic RNase III enzyme Dicer, which is also involved in siRNA genesis, to produce ~22 nucleotide microRNA duplexes (Bernstein *et al.*, 2001; Grishok *et al.*, 2001; Hammond *et al.*, 2000; Hutvagner *et al.*, 2001; ketting *et al.*, 2001; Knight *et al.*, 2001). Like Drosha, Dicer is highly conserved evolutionarily and contains two RIIIDs and a dsRBD. In addition, it contains a Piwi Argonaute Zwille (PAZ) domain which binds to the protruding 3' ends of small RNAs (Lingel *et al.*, 2004; Ma *et al.*, 2004; Song *et al.*, 2003).

Dicer interacts with other proteins such as RDE-4 in *C. elegans*, R2D2 and FMR1 in *D. melanogaster*, and the PAZ domain-containing Argonaute (Ago) proteins (Caudy *et al.*, 2002; Carmell *et al.*, 2002; Hammond *et al.*, 2001; Ishizuka *et al.*, 2002; Jin *et al.*, 2004; Tabara *et al.*, 2002). These proteins may not be involved directly in the cleavage reaction, but are important for microRNA stability and effector complex formation and action (Kim, 2005; Liu *et al.*, 2003; Zhang *et al.*, 2004). Interestingly, they are also known to regulate mRNA stability and translation rates. For instance, Ago2 was originally described as a translation enhancer protein (Carmell *et al.*, 2002). In the microRNA context, human AGO2 has been shown to function as the 'slicer' enzyme that mediates target mRNA cleavage (Meister *et al.*, 2004; Song *et al.*, 2004).

After Dicer processing, one strand of the microRNA duplex is usually degraded while the other persists as a mature microRNA (Kim, 2005). The strand that has a less thermodynamically stable 5' end is thought to be incorporated into effector complexes called microRNA-containing RNA-induced silencing complexes (miRISCs) (Khvorova *et al.*, 2003; Schwarz *et al.*, 2003). These miRISCs recognize and bind to target mRNAs to modulate their expression.

### 1.2.4 Mechanism of action

MicroRNAs modulate target expression in two different ways: by directing transcript degradation or inhibiting translation (Bartel, 2004) (Figure 1.3). In plants and very rarely in animals, microRNAs bind to highly complementary microRNA binding sites in target mRNAs to guide sequence-specific cleavage. This process is similar to RNA interference (Peters and Meister, 2007). In animals, microRNAs bind to partially complementary microRNA binding sites, usually in the 3' UTRs of target mRNAs, and repress translation. This repression is achieved by interfering with translation or by guiding degradation processes that are initiated by mRNA deadenylation and decapping (Pillai *et al*., 2007). In contrast with sequence-specific RNA cleavage which is well characterized, the molecular mechanisms behind microRNA-mediated translational repression are poorly understood.

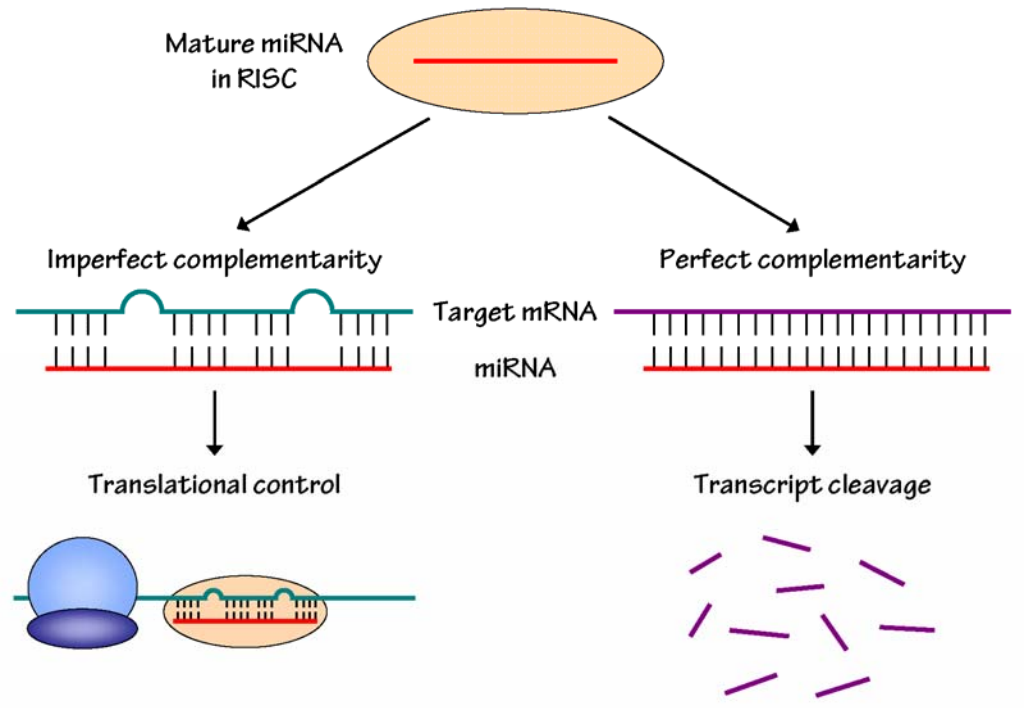



**Figure 1.3. Mechanism of microRNA action.** MicroRNAs act via either translational control or transcript cleavage, depending on the degree of sequence complementarity between the microRNA and its mRNA target.

Recently, new insights into these mechanisms have been gained via the use of cell-free *in vitro* systems (Mathonnet *et al*., 2007; Thermann and Hentze, 2007; Wakiyama *et al*., 2007; Wang *et al*., 2006). Wang *et al.* showed that microRNA-mediated translational inhibition requires a functional m7G-cap and a poly(A) tail (Wang *et al*., 2006). Mathonnet *et al.* found that microRNAs inhibit ribosome recruitment to their target mRNA, and interfere with translational initiation by targeting the mRNA cap structure. They also suggest that inhibition of translation is an early event in microRNA-guided gene silencing that may be followed by mRNA degradation (Mathonnet *et al*., 2007).

Ago proteins possess a highly conserved motif, containing two amino acids that specifically bind the m7G-cap, which is similar to the m7G-cap-binding motif of eIF4E. Kiriakidou *et al.* demonstrate that mutation of these two critical phenylalanines interferes with Ago2's ability to interact with the m7G-cap, without affecting its ability to cleave target mRNAs (Kiriakidou *et al.*, 2007). They propose a model in which Ago proteins and eIF4E compete for m7G-cap binding. eIF4E cannot access the cap once an Ago protein has bound to it. This results in repression of translational initiation (Kiriakidou *et al.*, 2007).

Intriguingly, Vasudevan *et al.* recently showed that human miR-369-3 activates translation by directing the association of Ago and fragile X mental retardation-related protein 1 (FXR1) with AU-rich elements (AREs) (Vasudevan *et al.*,2007). They also demonstrate that two other microRNAs, Let-7 and the synthetic miRcxcr4, induce translation up-regulation of their target mRNAs on cell cycle arrest

(Vasudevan *et al.*, 2007). These results provide the first evidence for microRNAs upregulating translation.

Other studies in *C. elegans* and mammals have shown cosedimentation of microRNAs with polyribosomes, suggesting a role for microRNAs in regulating translational elongation (Maroney *et al*., 2006; Nottrott *et al*., 2006; Olsen and Ambros, 1999; Petersen *et al*., 2006; Seggerson *et al*., 2002). In addition, Petersen *et al.* reported that microRNA binding to the 3'UTR causes ribosomes to drop off mRNAs. This has led to a ribosome drop-off model of microRNA function (Petersen *et al*., 2006).

Although there is increasing support for a model in which microRNAs regulate translational initiation, other evidence suggests that they may also function during the later steps of translation. MicroRNAs may also function in spatially and temporally distinct ways on different mRNAs.

### 1.2.5 Target prediction

Mature microRNAs regulate gene expression by binding to microRNA response elements (MREs) in their target mRNAs. Target prediction algorithms have been developed to identify microRNA targets, as target recognition is based partly on sequence complementarity between a microRNA and its MRE. However, at the time this study began, only a few such algorithms had been published, and only a small number of validated predictions had been reported (Table 1.1).

| Reference/*Species* | No. of miRs tested | No. of predicted targets | No. of targets validated | Method basis |
|---|---|---|---|---|
| **Lewis *et al.*, 2003 /** *H. sapiens, m. musculus, r. norvegicus, f. rubripes* | 8 | 15 | 11 | **TargetScan** |
| **Enright *et al.*, 2003 /** *Drosophila melanogaster* | Computational recovery of a few previously reported targets | | None | **MiRanda** |

**Table 1.1. MicroRNA target prediction tools available when this study began.**

MicroRNA target prediction in plants is relatively straightforward as MREs, which are usually found in coding regions, have extensive complementarity to their corresponding microRNAs. Thus, comparatively simple bioinformatics screen have been sufficient to identify many plant microRNAs binding sites (Rhoades *et al.*, 2002; Schwab *et al.*, 2005).

Prediction of animal microRNAs binding sites, which have been identified mostly in 3'UTRs, has proven to be a significantly more challenging task as they usually exhibit imperfect complementarity with their mature microRNAs (Bentwich, 2005; Rajewsky, 2006; Segupathy *et al.*, 2006). These microRNAs thus have the potential to be active against many mRNAs with little sequence homology. Although this ingenious mechanism allows the coordinate control of many genes, relatively complex algorithms will be required for prediction of these targets.

In animals, the majority of target discovery approaches to date have focused almost exclusively on 3'UTRs as microRNA targets (Bentwich, 2005; Rajewsky, 2006). This

may be influenced by the fact that the founding members of the microRNA class were shown to act on the 3′UTRs of *lin-14* and *lin-41*. 3'-UTRs of mRNAs also tend to be longer than 5'-UTRs, and are known to direct mRNA stability, translation efficiency and localization. However, the 5'-UTR also directs mRNA translation (Muckenthaler *et al.*, 1998; Rouault, 2005; Thomson *et al.*, 2005). No studies have yet been published on the targeting of endogenous 5'-UTR or coding regions by microRNAs in animals. However, it is interesting to note that other components of the post-transcriptional machinery, the RNA-binding proteins (RBPs), are known to bind all along mRNAs and regulate their translation and degradation (George and Tenenbaum, 2006). If the microRNA mode of action is reliant on incorporation in RBP-mRNA complexes, then one may surmise that microRNAs may exhibit activity along most of a mRNA, not just its 3'UTR.

To date, despite the considerable effort expended on developing prediction algorithms, the number of confirmed heteroduplexes remains small. The number of target sites that have been validated *in vivo* under endogenous conditions and by mutagenesis is even smaller (Chen and Rajewsky, 2007). There is also little overlap between predictions made by various algorithms (Rajewsky 2006). These observations underscore the challenging nature of this field. Undoubtedly, breakthroughs in microRNA target prediction and validation will be instrumental in advancing our understanding of microRNA function and potential therapeutic applications.

## 1.3 MicroRNAs in ESCs

In the context of mESCs, the loss of mature microRNAs in Dicer1 null mESCs results in a failure of mESCs cells to differentiate (Kanellopoulou *et al.*, 2005). Furthermore, DGCR8, an RNA-binding protein involved in microRNA processing, is essential for microRNA biogenesis and silencing of mESC self-renewal (Wang *et al.*, 2007). These data highlight the importance of regulated microRNA expression in controlling ESC growth and differentiation.

ESC-specific microRNAs have been identified in murine and human ESCs (Houbaviy *et al.*, 2003; Suh *et al.*, 2004), however, their functional significance has not been evaluated (Table 1.2). The expression patterns of microRNAs in ESCs can be classified into five groups (1) microRNAs that are expressed in ESCs as well as in embryonic carcinoma (EC) cells, which may have conserved roles in mammalian pluripotent stem cells (2) microRNAs expressed specifically in ESCs but not in other cells including EC cells. These may have functions specific to ESCs (3) microRNAs that are rare in ESCs and increase upon differentiation, which may be involved in the differentiation process (4) microRNAs that are present in ESCs and remain at a constant level during differentiation. These may be involved in general aspects of cell physiology (5) microRNAs that increase or decrease transiently during ESC differentiation. These may module ESC differentiation into specific cell types.

MicroRNAs were first identified in ESCs using cDNA cloning (Table 1.2 and 1.3). Houbaviy *et al.* describe 53 microRNAs in mESCs, of which 15 are novel (Houbaviy *et al.*, 2003). Although the levels of many previously described microRNAs remain constant or increase upon differentiation, eight of the novel microRNAs (miR-290,

miR-291s, miR-291as, miR-292, miR-292as, miR-293, miR-294, miR-295) appear to be ESC or early embryo specific by four criteria: (1) their sequences are distinct from those of previously described microRNAs, including microRNAs cloned from adult mouse organs; (2) they cannot be detected in adult mouse organs by Northern analyses; (3) they are repressed during ESC differentiation *in vitro*, (4) all ESTs that map within the cluster are derived from ESCs or preimplantation embryos (Houbaviy *et al.*, 2005).

MiR-290 to miR-295 is a cluster of partially homologous pre-microRNA hairpins encoded by genomic loci clustered within 2.2kb of each other (Houbaviy *et al.*, 2003). This entire Early Embryonic microRNA Cluster (EEmiRC) is spanned by a spliced, capped and polyadenylated primary transcript, and transcription is directed by a conserved promoter element containing a TATA box. Sequence analysis shows that the EEmiRC transcription unit is remarkably variable and can only be identified bioinformatically in placental (eutherian) mammals (Houbaviy *et al.*, 2005). The only conserved regions within the locus are the pre-microRNA hairpins and the putative minimal promoter. The number and precise sequences of the pre-microRNAs, their distance from the promoter and the polyadenylation sites, the regions flanking the hairpins, and the types, positions and numbers of repetitive element insertions vary in species belonging to different mammalian orders (Houbaviy *et al.*, 2005).

Experimental support for the *in silico* prediction of an EEmiRC counterpart in the human genome was provided by the cDNA cloning of the corresponding microRNA homologs (miR-371, miR-372, miR-373 and miR-373*) from hESCs (Suh *et al.*, 2004). Although mouse and human EEmiRC microRNAs are sufficiently different

from each other to warrant different numerical designations, multiple sequence alignment reveals related sequences both within each cluster and across species. Furthermore, sequence conservation extends beyond the mature microRNAs to the entire pre-microRNA hairpin sequences (Houbaviy *et al.*, 2003).

| Species | MicroRNA | Type of differentiation | Function | Reference |
|---|---|---|---|---|
| Mouse | **miR-290** | Embryoid body formation (+/- RA) | MESC-specific microRNA cluster, hypothesized to have ESC-specific functions | Houbaviy *et al.*, 2003 |
| | **miR-291-s/as** | | | |
| | **miR-292-s/as** | | | |
| | **miR-293** | | | |
| | **miR-294** | | | |
| | **miR-295** | | | |
| Human | **miR-302a/a*** | Embryoid body formation | Homologous cluster, also expressed in EC cells | Suh *et al.*, 2004 |
| | **miR-302b/b*** | | | |
| | **miR-302c/c*** | | | |
| | **miR-302d** | | | |
| | **miR-367** | | | |
| Human | **miR-371** | Embryoid body formation | Human ESC-specific microRNA cluster, homologous to the mESC-specific cluster above | Suh *et al.*, 2004 |
| | **miR-372** | | | |
| | **miR-373/3*** | | | |

**Table 1.2. Some of the microRNAs that are downregulated during ESC differentiation.**

The miR-371 to miR-373* cluster is located within a 1050 bp region on chromosome 19 (Suh *et al.*, 2004). Intriguingly, they are also highly expressed in cancer cell lines, and miR-372 and miR-373 have the ability to protect cells from oncogenic stress and transform primary human cells (Kim, 2005; Voorhoeve *et al.*, 2006). This suggests that they may operate key regulatory networks conserved in pluripotent stem cells and cancer cells, and serve as molecular markers for undifferentiated hESC and specific cancers. Suh *et al.* report another cluster of eight highly related microRNAs (miR-302b, miR-302b*, miR-302c, miR-302c*, miR-302a, miR-302a*, miR-302d and miR-367) located within a ~700bp region on chromosome 4. MiR-302b, miR-302c and miR-302d appear to be the close homologues of miR-302 that was cloned from mESCs. Although these are the most abundant microRNAs in hESCs, their murine homologue miR-302 appears to be less abundant (Suh *et al.*, 2004).

It is intriguing that two microRNA clusters are conserved and specifically expressed in both mouse and human ESCs. Although the variation in the numbers and sequences of the homologues may implicate divergence of the conserved regulatory pathways, these conserved microRNAs are likely to play central roles in the regulation of mammalian ESCs. Aside from these two clusters, there are considerable differences between the microRNAs cloned from human and mESCs. This may imply fundamental differences between the regulatory networks in hES and mESCs, or may simply be due to the cloning techniques employed, which may not be sensitive enough to identify the complete set of microRNAs in these cell lines. This limitation has been overcome by recently developed microRNA microarray techniques which have enabled global microRNA expression profiling. Microarray analysis of 124 mammalian microRNAs has identified 6 additional microRNAs, miR-214, miR-134,

miR-25, miR-182, miR-204, miR-132, which demonstrate mESC-specific expression (Thomson *et al.*, 2004).

Despite the relative abundance of expression data, at the time this study began, comparatively few microRNAs have been functionalized in ESCs (Tables 1.2, 1.3). Thus, more comprehensive expression profiling, in conjunction with in-depth functional studies, will enhance current knowledge about microRNAs and should also increase our capacity to manipulate and develop ESCs as therapeutics for regenerative medicine.

| MicroRNA | Type of differentiation | Function | Reference |
|---|---|---|---|
| **Let-7a-3** | Neural | | Lee *et al.*, 2005 |
| **miR-9/9*** | Neural | Affects neural lineage differentiation, may act in the STAT3 signaling pathway | Krichevsky *et al.*, 2006 |
| **miR-21** | Embryoid body formation (+/- RA) | | Houbaviy *et al.*, 2003 |
| **miR-22** | Embryoid body formation (+/- RA) | | Houbaviy *et al.*, 2003 |
| | Neural | | Krichevsky *et al.*, 2006 |
| **miR-124a** | Neural | Affects neural lineage differentiation, may act in the STAT3 signaling pathway | Krichevsky *et al.*, 2006 |
| **miR-125b** | Neural | | Lee *et al.*, 2005 |
| | Neural | | Krichevsky *et al.*, 2006 |

**Table 1.3. Some of the microRNAs that are upregulated during mESC differentiation.**

## CHAPTER 2. STATEMENT OF AIMS

Due to the paucity of knowledge about the functions of microRNAs in ESCs and the limited target prediction algorithms available at the time of this study, the two initial aims were to identify and characterize additional microRNAs which may be involved in the modulation of ESC pluripotency and differentiation, and to aid the development of a robust and sensitive microRNA prediction algorithm with higher positive predictive rates than those previously available.

Chapter 3 describes work to identify putative microRNAs whose attenuation or overexpression may induce ESC differentiation, and to determine if these microRNAs are able to modulate specific lineage differentiation.

Chapter 4 will describe briefly a microRNA target prediction algorithm, *rna22*, developed by our collaborator Isidore Rigoutsos, and experimental validation of 226 *rna22*-predicted targets. Several of these will be further investigated to determine whether they mediate the effect of microRNAs in modulating mESC differentiation.

As results from chapter 3 suggest that miR-134 may modulate *in vitro* mESC differentiation towards neural lineages, chapter 5 will investigate the expression profile *in vivo* of miR-134, and provide preliminary evidence of possible mechanisms by which miR-134 may modulate neural differentiation *in vivo*.

Chapter 6 challenges the dogma that animal microRNAs only target the 3'UTRs of genes. This will be investigated in the context of microRNAs targeting the coding regions of transcription factors which are important regulators of pluripotency.

# CHAPTER 3. MICRORNAS MODULATE MESC DIFFERENTIATION

## 3.1 Introduction

MicroRNA expression profiling in mESCs has been performed previously by Houbaviy *et al.* by cloning and sequencing size-fractionated cDNA libraries of undifferentiated and differentiated mESCs. As most sequences (73%) were observed only once, Houbaviy *et al.* surmised that their dataset was probably not representative of the complete pool of microRNAs present in undifferentiated and differentiated mESCs (Houbaviy *et al.*, 2003). Additionally, as mentioned earlier, this method of detection may not detect microRNAs that have low expression levels, specific sequence composition or post-transcriptional modifications (Luciano *et al.*, 2004; Yang *et al.*, 2006; Yang *et al.*, 2006).

A later study by Krichevsky *et al.* describes microRNA expression profiling during mESC-derived neurogenesis *in vitro* with probes specific for 135 microRNAs (Krichevsky *et al.*, 2005). They showed that a number of microRNAs are simultaneously co-induced during the differentiation of neural progenitor (NP) cells to neurons and astrocytes. Seventy of the 135 vertebrate microRNAs tested showed a greater than two-fold change between any two stages of the five-step neural differentiation protocol. MiR-124a and miR-9, which are almost exclusively expressed in the brain, affect neural lineage differentiation of the mESC cultures. Krichevsky *et al.* hypothesize that early overexpression of miR-124a in NPs prevents gliogenesis whereas miR-9 expression contributes to neurogenesis (Krichevsky *et al.*, 2005). These data demonstrate a functional role for distinct microRNAs in the determination of neural fates during ESC differentiation.

The microarray analysis by Krichevsky *et al.* profiled microRNA expression during a defined 5-stage protocol that induces neural differentiation of mESCs. Differentiation of mESCs can be induced in many ways, including embryoid body (EB) formation, induction with retinoic acid (RA), culture in serum-free medium containing N2 and B27 supplements, co-culture with specific tissues or cell lines, promoter trapping and the forced expression of certain developmental control genes (Bain *et al*., 1995; Barberi *et al.*, 2003; Finley *et al.*, 1999; Friedrich and Soriano, 1991; Fujikura *et al*., 2002; Leahy *et al.*, 1999; Niwa *et al*., 2005; Okada *et al.*, 2004; Ying *et al.*, 2003).

EBs, cell aggregates formed from ESCs grown in suspension, are useful models of gene expression and differentiation in early development as they resemble the early postimplantation embryo in size, differentiation capacity and gene expression profile (O'Shea, 2004). EBs have been shown to differentiate into derivatives of all three germ layers (Leahy *et al.*, 1999). As such, many differentiation protocols start with several days of EB formation in the absence of LIF (Doetschman *et al.*, 1985; Kanai-Azuma *et al.*, 2002; Lickert *et al.*, 2002; Martin and Evans, 1975; Ying *et al.*, 2003).

Treatment with RA is another commonly used and robust method of mESC differentiation. RA is an important morphogen which is required during development of the central nervous system, lung, and kidneys, and during proximodistal patterning and limb generation (Aliotta *et al.*, 2005; Burrow, 2000; Maden, 2001; Yashiro *et al.*, 2004). In the absence of LIF, RA enhances both neural differentiation and caudalization of mESCs in a concentration-dependent manner (Bain *et al*., 1995; Okada *et al.*, 2004). Furthermore, certain concentrations of RA have been shown to enhance cardiomyocyte differentiation from mESCs (Wobus *et al.*, 1997).

In an attempt to obtain a more comprehensive profiling of microRNA expression during mESC differentiation that was not restricted to any particular lineage, microarray analysis was performed using chips containing probes for 461 microRNAs. For this study, the two methods of differentiation used were EB formation and RA induction.

## 3.2 Identification of microRNAs modulated during mESC differentiation

To examine the relative changes in microRNA expression during mESC differentiation, microarray analysis was performed on two sets of samples: Total RNA isolated from E14 mESCs subject to EB differentiation for 0, 3, 6, 9, 12 and 15 days, and total RNA isolated from mESCs ± RA (100 nM) at 0, 2, 4 and 6 days post-treatment. The difference in time durations selected was due to the fact that RA induction results in more rapid mESC differentiation than EB formation.

Out of the 461 microRNAs probed, 201 were differentially expressed during EB differentiation; the expression pattern of several of these is shown in Figure 3.1A and B. Additionally, 134 microRNAs that were differentially expressed following RA treatment were identified; the expression pattern of several of these is shown after four days of RA treatment (Figure 3.2A, B). Northern blot analysis was performed to confirm the microarray profiling for a representative sample of these differentially expressed microRNAs (Figure 3.3A-C).

Microarray analysis also demonstrates that members of the miR-290 to 295 microRNA cluster are highly expressed in undifferentiated mESCs and decrease during both EB and RA differentiation (Figure 3.1A, 3.2B) This is consistent with the previous report by Houbaviy *et al.* and affirms the reliability of the cell culture and microarray platforms used.

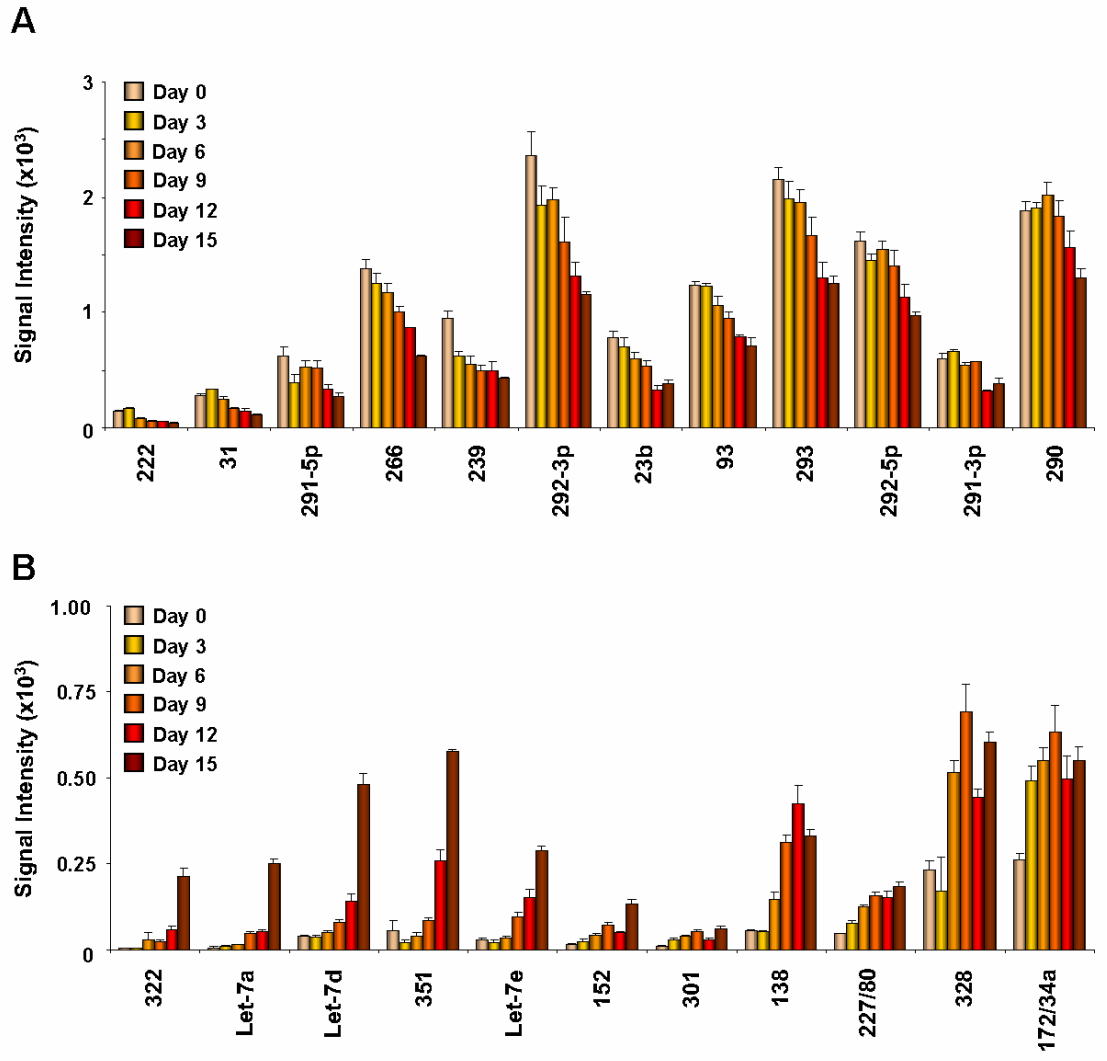



**Figure 3.1. MicroRNA expression levels change during EB differentiation of mESCs (A,B)** The expression levels of representative microRNAs which **(A)** decreased or **(B)** increased during EB differentiation compared to day 0 mESCs. Total RNA extracted from mESCs (control) or EBs (day 3 to day 15) was size fractionated and hybridized to microRNA specific probes on a microarray as described in methods. The total array profile has MIAME accession number E-MEXP-977.

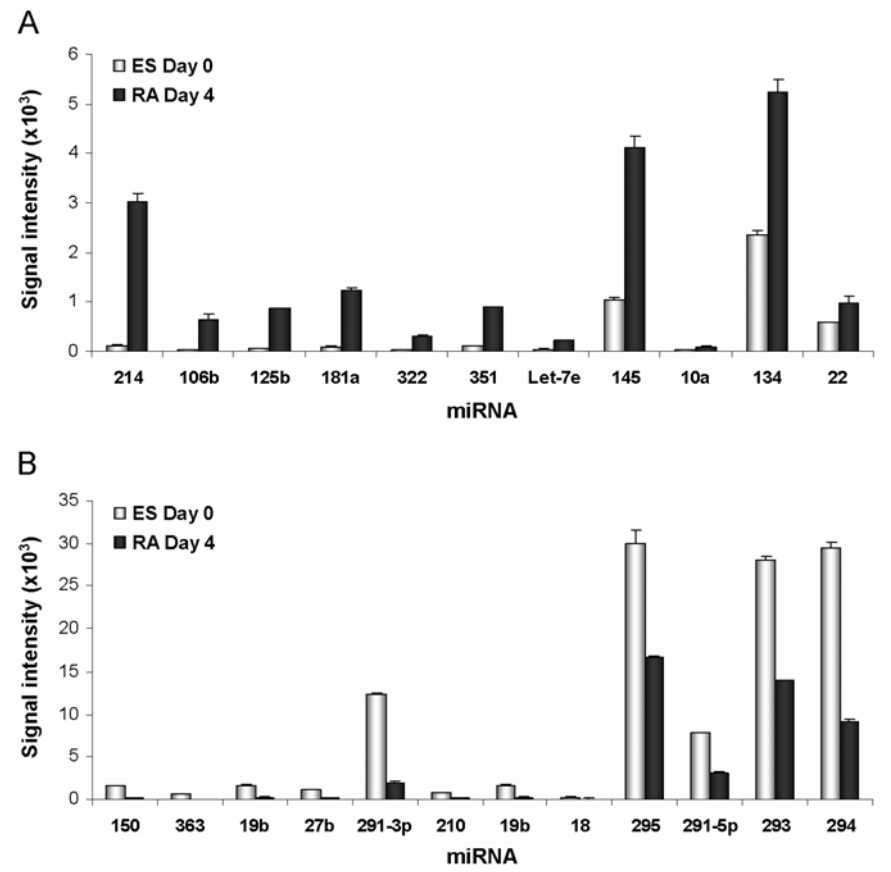



**Figure 3.2. MicroRNA expression levels change during RA-induced mESC differentiation (A,B)** The expression levels of representative microRNAs which **(A)** increased or **(B)** decreased with RA treatment (day 4) compared to day 0 mESCs. Total RNA extracted from mESCs (control) ± 100 nM RA (day 0 to day 4) was size fractionated and hybridized to microRNA specific probes on a microarray as described in methods. The total array profile has MIAME accession number E-MEXP-977.

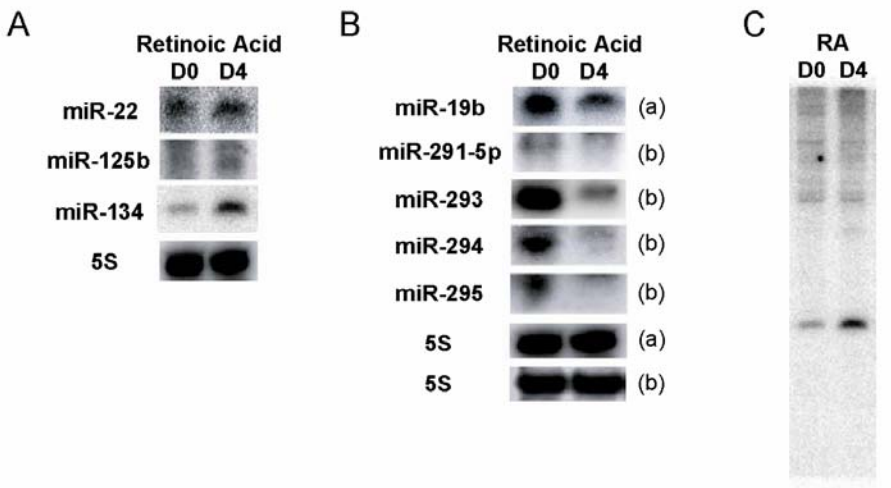



**Figure 3.3. Northern blot validation of microRNA microarray expression profiles (A,B)** Northern blot analyses demonstrating **(A)** upregulation or **(B)** downregulation of representative microRNAs by RA treatment (day 0 to day 4) of E14 mESCs, confirming the data obtained from the microRNA microarray analysis (**C)** Expanded Northern blot analyses of miR-134 upregulation during RA-induced mESC differentiation to demonstrate probe specificity.

## 3.3 MicroRNAs can modulate *Oct4* and *Nanog* promoter activity

Next, the effect of dysregulation of these differentially expressed microRNAs on self-renewing mESCs in the presence of serum and LIF was investigated. The transcription factors Oct4, Sox2 and Nanog are known to play a central role in ESC maintenance and differentiation, and the levels and activity of these transcription factors are indicators of ESCs' pluripotent- or differentiation-status (Loh *et al.*, 2006). Constructs containing either the *Oct4* promoter or the *Nanog* promoter that contain Oct4/Sox2 binding sites driving transcription of a luciferase reporter gene were used to measure the transcription activity of these promoters and provide a sensitive method to detect loss of pluripotency and onset of differentiation (Figure 3.4A, B). A stringent threshold of $p<0.0001$ was set based on the luciferase readout for each transfection relative to control across two mESC-lines (E14 and D3) and the two reporter constructs.

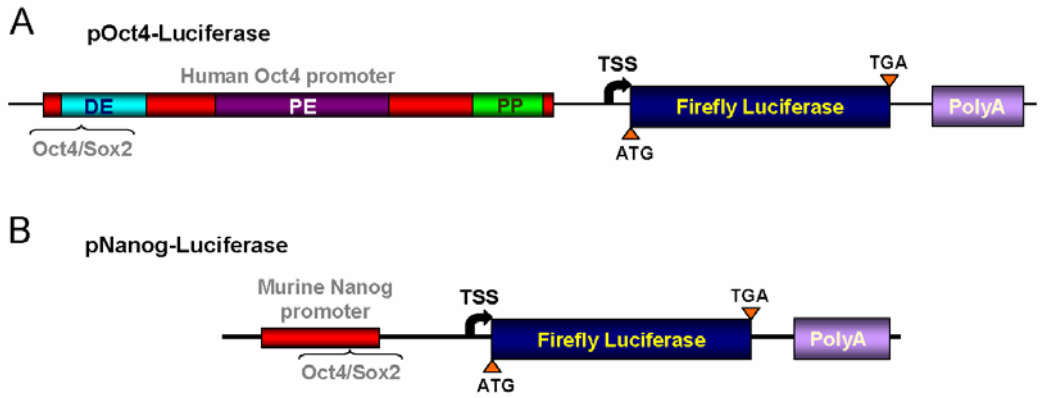
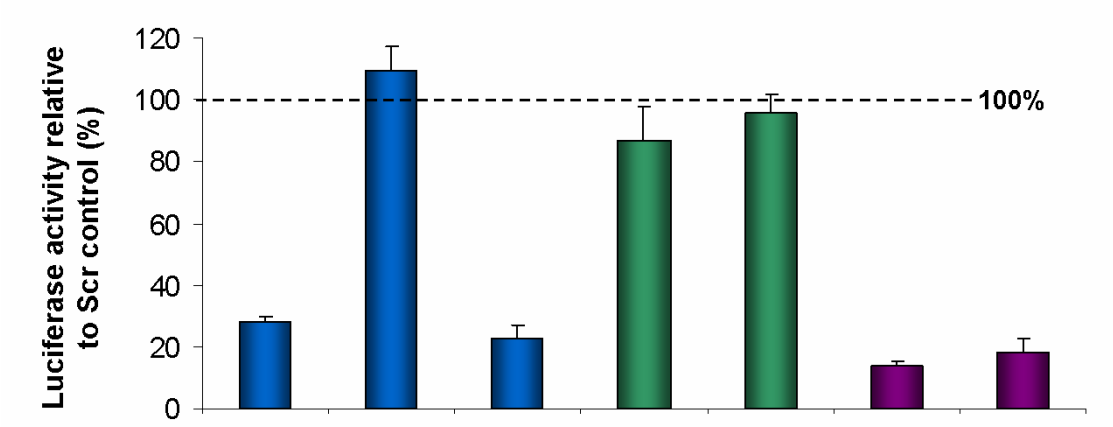



**Figure 3.4. Schematic representations of pOct4-Luciferase and pNanog-Luciferase constructs (A)** Schematic representation of pOct4-Luciferase plasmid reporter construct, which was used as a measure of *Oct4* promoter activity. pOct4-Luciferase comprises 3 kb of the human *Oct4* promoter, where the Oct4-Sox2 binding site is indicated (DE, distal enhancer; PE, proximal enhancer; PP, proximal promoter), upstream of the transcription start site (TSS) for *Photinus pyralis luc* (firefly luciferase) (ATG, translation start codon; TGA, translation stop codon; PolyA, polyadenylation signal) **(B)** Schematic representation of pNanog-Luciferase plasmid reporter construct, which was used as a measure of *Nanog* promoter activity. pNanog-Luciferase comprises 500 bp of the mouse *Nanog* promoter, where the Oct4-Sox2 binding site is indicated.

Anti-miR™ microRNA inhibitors (Anti-miRs) or Pre-miR microRNA™ precursors (Pre-miRs) from Ambion were used to repress or overexpress the selected microRNAs. An experiment was first carried out in 293T cells to demonstrate the efficacy of these synthetic oligomers. As the effectiveness of target repression by a microRNA is dependent on the degree of sequence complementarity it has with its target sequence, it would exhibit maximal repression of a perfectly complementary target. The reverse complements (RCs) of miR-291-5p and miR-291-3p were cloned downstream of a luciferase reporter to serve as the readouts for this experiment.

Relative to the Scrambled oligomer control transfection, transfection of Pre-miR-291 resulted in a very significant repression of luciferase activity of both 291-5p and 291-3p RC targets (Figure 3.5). This was not affected by the co-transfection of the non-complementary Anti-miR 19b (Figure 3.5). Transfection of Pre-miR-22 did not repress the luciferase activity of the 291-5p RC target relative to Scrambled control transfection (Figure 3.5). These results indicate that Pre-miRs are processed within transfected cells to yield mature microRNAs which are able to act specifically on their mRNA targets.

This luciferase reporter assay was also used to investigate the effectiveness of Anti-miRs. As depicted in Figure 3.5, the co-transfection of Pre-miR-291 (which generates both miR-291-5p and miR-291-3p) and Anti-miR-291-5p was able to specifically rescue the effect of Pre-miR-291 on the 291-5p RC target but not the 291-3p target, demonstrating that the Anti-miRs were functional as microRNA inhibitors in the cell culture platform used (Figure 3.5).




| RC Target | 291-5p | 291-5p | 291-5p | 291-5p | 291-5p | 291-3p | 291-3p |
|---|---|---|---|---|---|---|---|
| Pre-miR | 291 | 291 | 291 | 22 | 22 | 291 | 291 |
| Anti-miR | Scr | 291-5p | 19b | Scr | 291-5p | Scr | 291-5p |

**Figure 3.5. Luciferase assay demonstrating the efficacy of Anti-miRs and Pre-miRs.** Graph showing the effect on luciferase activity in 293T cells of the listed Pre-miR and Anti-miR cotransfections with the respective reverse complement (RC) target. Luciferase experiments were performed in quadruplicate thrice (n=12), where error bars denote standard error.

In addition to the luciferase assays, Northern blots were also performed after transfection of Pre-miRs and Anti-miRs into mESCs. Anti-miR-293 was able to reduce the amount of the abundantly expressed miR-293 in mESCs (Figure 3.6B). Other studies have also reported a reduction in microRNAs with Anti-miR treatment, however, the mechanism remains unknown (Flynt *et al.*, 2007). Furthermore, Pre-miR-134 was processed to generate mature miR-134, which could be sustained for at least 5 days post-transfection (Figure 3.6D).

Having confirmed the efficacy of both Anti-miRs and Pre-miRs, they were next co-transfected with the *Oct4*-luciferase reporter construct in mESCs to provide an initial screen for microRNAs which may modulate the loss of mESC pluripotency. In all experiments, co-transfection of E14 and D3 mESCs with *Oct4* RNAi (positive

control) and the *Oct4*-luciferase reporter resulted in at least a 60% reduction in promoter activity (Figure 3.6A, C, E).

For the microRNAs that were downregulated during differentiation, the *Oct4*-luciferase reporter was co-transfected with Anti-miRs to ascertain whether reduction in the levels of these microRNAs alone could result in a loss of mESC pluripotency. However, compared to the Scrambled control transfection, transfection of any one of the Anti-miRs did not result in any significant change in luciferase activity across the 2 mESC-lines tested (Figure 3.6A). It is possible that the simultaneous downregulation of several of these microRNAs may be needed to observe a significant effect.

For the microRNAs that were upregulated during differentiation, the *Oct4*-luciferase reporter was co-transfected with Pre-miRs to ascertain whether upregulation in the levels of these microRNAs alone could result in a loss of mESC pluripotency. Of the eleven microRNAs tested, only miR-134 resulted in a significant reduction in *Oct4*-promoter activity (Figure 3.6C) relative to the Scrambled oligomer control in both E14 and D3 mESCs.

To further investigate whether miR-134 could induce mESC differentiation as quantified by a decrease in transcription activity of a pluripotency marker, Pre-miR-134 was next co-transfected with the *Nanog*-luciferase reporter. Consistent with the *Oct4*-luciferase reporter results, this also resulted in significantly reduced *Nanog*-promoter activity relative to the Scrambled oligomer control in both mESC-lines (Figure 3.6E).

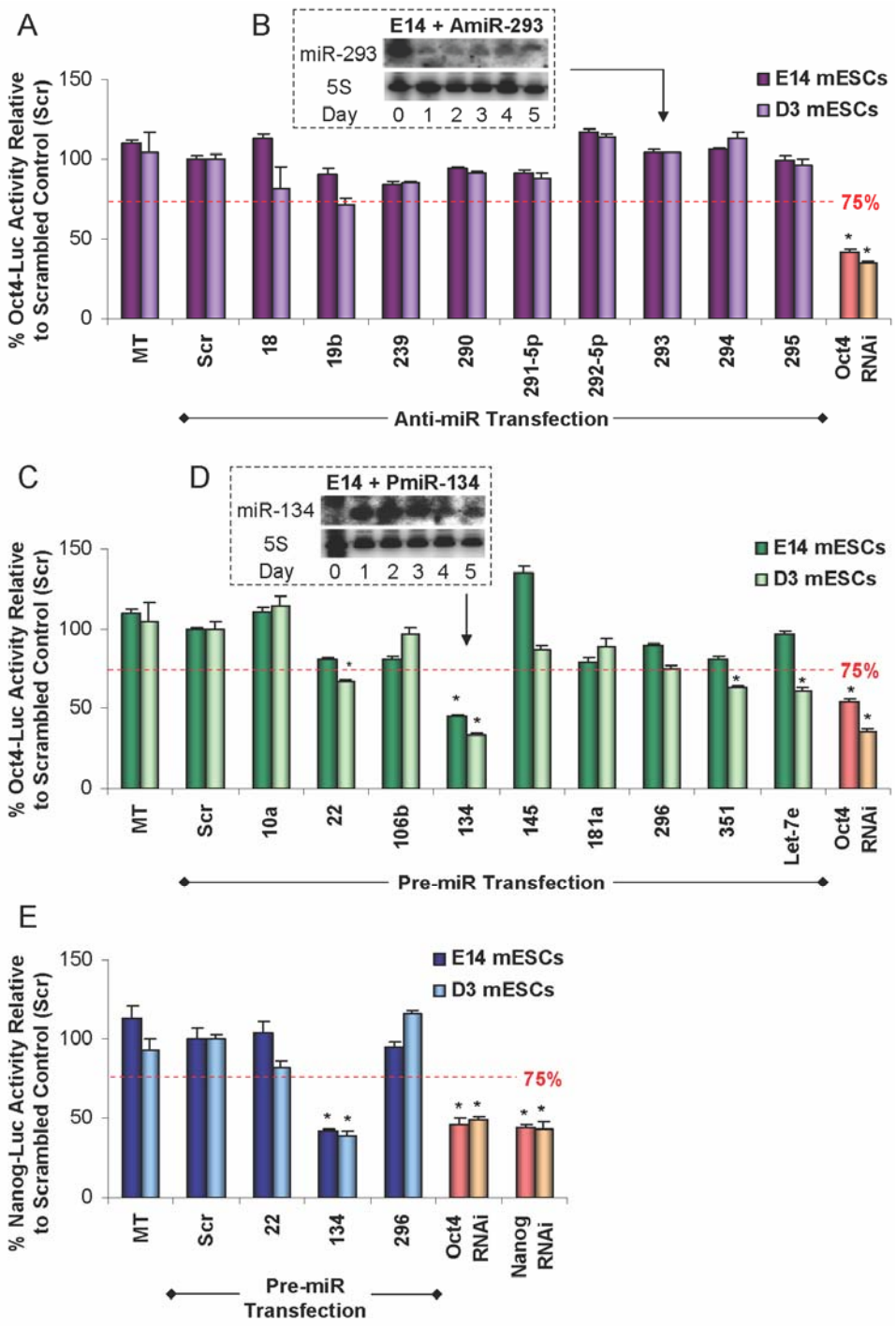



**Figure 3.6. Overexpression of miR-134 downregulates Oct4 and Nanog promoter activities (A)** Relative to the Scr control transfection, none of the Anti-miRs transfected were able to induce a significant downregulation in *Oct4* promoter activity in E14 or D3 mESCs **(C,E)** Relative to the Scr control transfection, PmiR-134 transfection induced a significant downregulation in *Oct4* promoter **(C)** and *Nanog* promoter **(E)** activities in both E14 and D3 mESCs. Experiments were performed in triplicate thrice (n=9), where error bars denote standard error, and asterisk (*) denotes significance of difference to scrambled RNA oligomer control at $p<0.0001$ **(B,D)** Northern blot analyses showing that Anti-miR-293 transfection into E14 mESCs led to reduced levels of mature miR-293 **(B)**, while PmiR-134 transfection led to enhanced levels of mature miR-134 **(D)**, relative to non-transfected control (where 0 denotes no transfection, D1 to D5 denote day 1 through 5 post-transfection). Experiments were performed in triplicate twice (n=6), and a representative blot is shown.

### 3.4 Expression profile of miR-134 during mESC differentiation

As the results of the *Oct4* and *Nanog* promoter screens suggested that miR-134 may be able to modulate mESC differentiation, its expression during mESC differentiation was explored more thoroughly. Quantitative PCR analyses corroborated the microarray data which demonstrated that miR-134 expression increases during RA differentiation, reaching a peak at day 4 before subsiding by day 6 (though still remaining at an elevated level relative to undifferentiated mESCs) (Figure 3.7).

When a less heterogeneous neuroectodermal lineage differentiation promoting regime (N2B27) was used (Aiba *et al.*, 2005; Ying *et al.*, 2003), the levels of endogenous miR-134 increased significantly within two days (3 times greater induction compared to RA), and again the levels of miR-134 subsided over the time course but remained elevated above control (Figure 3.7). Interestingly, and consistent with the microarray data, miR-134 levels did not change during EB formation (Figure 3.5).

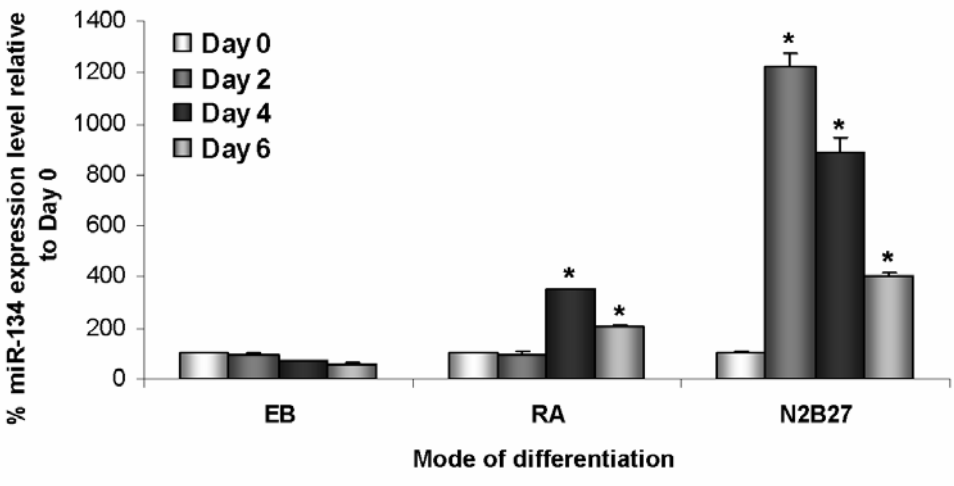



**Figure 3.7. Expression profile of miR-134 during mESC differentiation.** MicroRNA quantitative PCR analyses showing that, relative to expression in untreated E14 mESCs, miR-134 expression increases during both RA and N2B27 treatment of E14 mESCs, and not EB differentiation. Experiments were performed in duplicate thrice (n=6), where error bars denote standard error, and asterisk (*) denotes significance of difference to day 0 at $p<0.0001$.

## 3.5 MiR-134 modulates mESC differentiation even in the presence of LIF

To further investigate the possible role of miR-134 in modulating mESC differentiation, the effect of miR-134 alone on changes in the mRNA levels of key marker genes associated with pluripotency and differentiation was examined. Quantitative PCR of total RNA samples taken from E14 and D3 mESCs 3 days post-transfection with Pre-miR-134 showed that the *Oct4* mRNA level was significantly reduced, consistent with the reduction of promoter activity (Figure 3.8.). *Nanog* transcript levels were reduced significantly in D3, but not E14, mESCs transfected with Pre-miR-134. *Sox2* levels were reduced significantly in both D3 and E14 mESCs after Pre-miR-134 transfection. Additionally, elevated expression levels of early primitive ectoderm marker (*Fgf5*), neuroectoderm marker (*Nestin*) were observed, with no change or a down-regulation of endoderm (*Gata4, Sox17*) and mesoderm (*Bmp4, Nkx2.5*) markers (Figure 3.8.). These findings indicate that elevated miR-134 levels alone in undifferentiated mESCs promote a transcriptional expression profile suggestive of ectodermal differentiation.

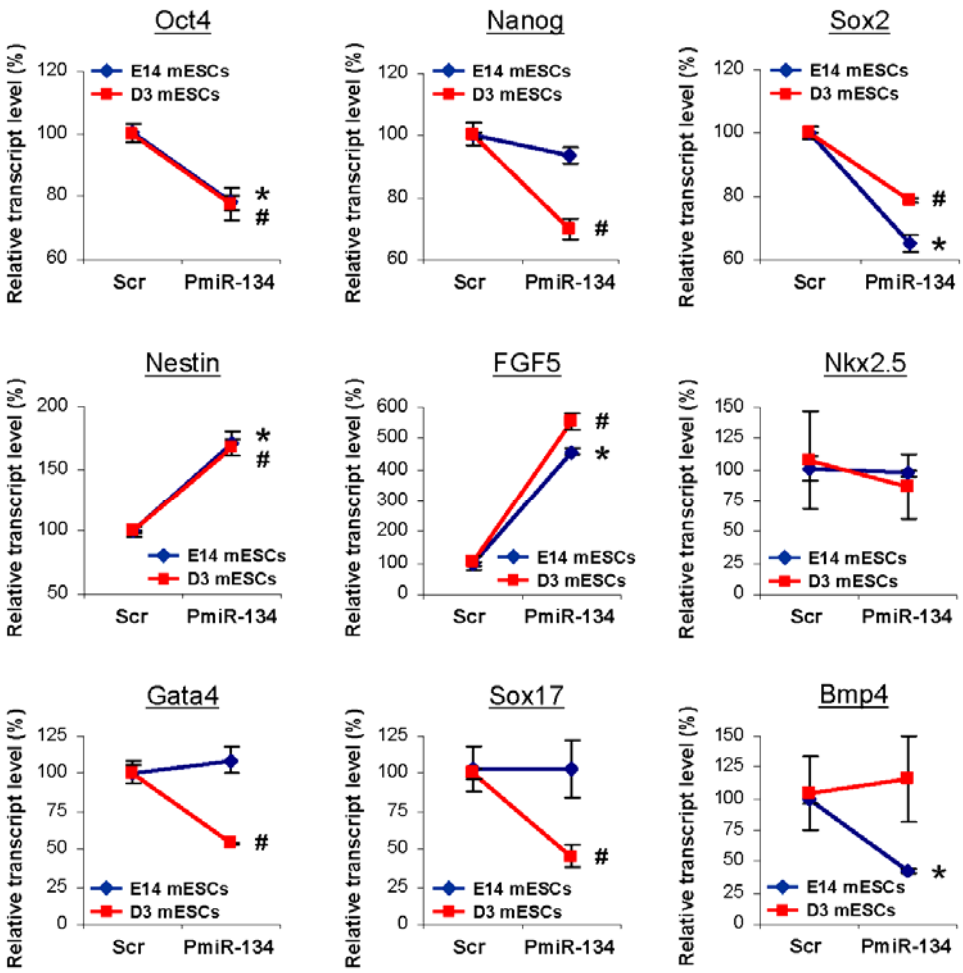



**Figure 3.8. miR-134 modulates the transcript levels of lineage-specific biomarkers, even in the presence of LIF.** Relative to Scr control transfection, transfection of PmiR-134 into E14 and D3 mESCs led to an appreciable upregulation of ectodermal markers *Nestin* and *Fgf5* after 3 days, with a concomitant reduction in endodermal, mesodermal and pluripotency-associated markers. Experiments were performed in triplicate thrice (n=9), where error bars denote standard error, and significance of difference to scrambled RNA oligomer control at $p<0.001$ are denoted as asterisk (*) and hex ($^{\#}$) for E14 and D3 mESCs respectively.

As transcript level is not always an accurate indication of the amount of a particular protein present, the effect of miR-134 on the protein levels of several pluripotency and differentiation markers was also explored. Western blot analyses showed that miR-134 elevation consistently resulted in a significant downregulation of Oct4, Nanog and Sox2 proteins in both E14 and D3 mESCs, relative to Scr control, during the second and third day following transfection with Pre-miR-134 (Figure 3.9A). Importantly, this effect was completely abolished by co-transfection of Pre-miR-134

with Anti-miR-134, miR-134's antagonist, thus demonstrating the specificity of miR-134 in the downregulation of these proteins (Figure 3.9A). Concomitant with changes in mRNA levels in E14 mESCs, Pre-miR-134 transfection led to increased protein levels of Nestin and decreased protein levels of BMP4. Gata4 protein levels were not affected (Figure 3.9B).

Having demonstrated at the molecular level that miR-134 induces mESC differentiation, miR-134's effect on mESCs at the cellular level was next investigated. E14 mESCs transfected with the scrambled RNA oligomer (Scr) maintained the characteristic domed colony structures of mESCs with no morphological changes for up to 4 days post-transfection (Figure 3.9C). However, transfection with Pre-miR-134 alone induced visible morphological changes in these cells within 4 days post-transfection (Figure 3.9C), with the cells acquiring a flattened epithelial-like morphology typical of differentiation. This data is consistent with reduced alkaline phosphatase activity in Pre-miR-134 treated mESCs (Figure 3.9C), where alkaline phosphatase is expressed at higher levels in pluripotent ESCs compared to differentiated cells.

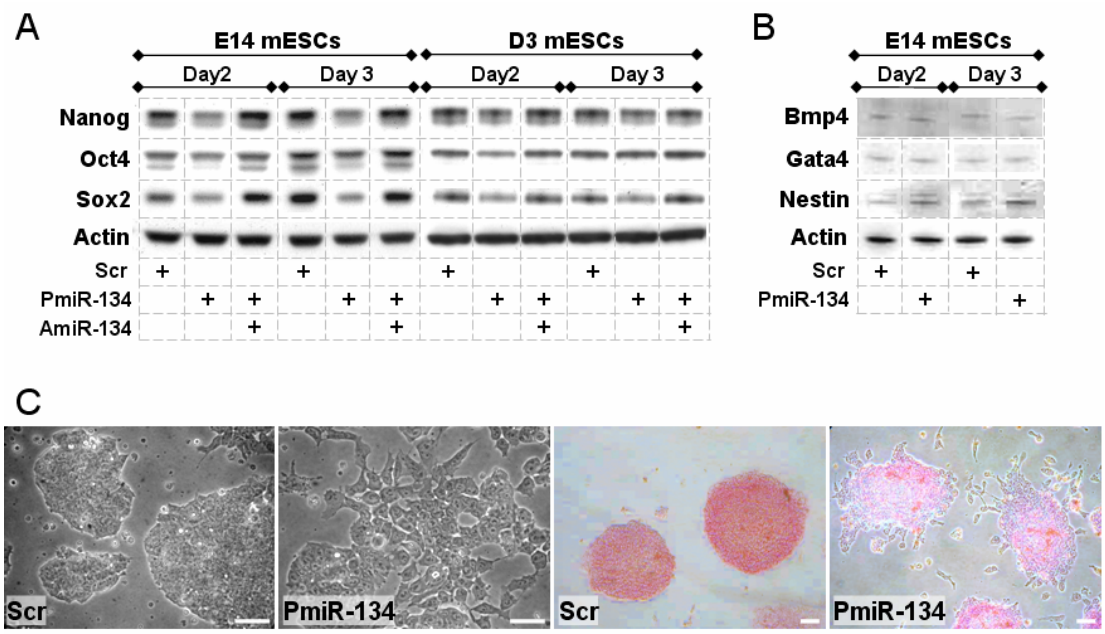



**Figure 3.9. miR-134 downregulates protein levels of pluripotency markers and induces changes in mESC morphology indicative of differentiation, even in the presence of LIF (A)** Western blot analyses showing that PmiR-134 induced significant downregulation of Nanog, Oct4 and Sox2 protein levels relative to Scr control in E14 and D3 mESCs. This effect was blocked by Anti-miR-134, over three days post-transfection. Experiments were performed in triplicate twice (n=6), and a representative blot is shown **(B)** Western blot analyses showing that PmiR-134 induced significant upregulation of Nestin protein relative to Scr control in E14 mESCs. Experiments were performed in triplicate twice (n=6), and a representative blot is shown **(C)** Photographs showing the morphology (left panel) and alkaline phosphatase (AP) staining (right panel) of E14 mESCs four days post-transfection with PmiR-134 or Scr control; where PmiR-134, unlike Scr, induced visible phenotypic changes to a flattened, epithelial-like morphology, and reduced AP activity relative to Scr control. Scale bar = 10 microns.

Colony forming unit assays, which involve the replating of mESCs 2 days post-transfection with Pre-miRs, Anti-miRs, Scrambled or water onto mEF feeder layers, were also performed. The number of secondary colonies obtained is a measure of the pluripotent state of the respective cultures. Colony formation assay results demonstrated that there was an ~30% decrease in the ability of mESCs treated for 2 days with Pre-miR-134 to form mESC colonies compared to mock transfection control (Table 3.1). This provided further evidence that miR-134 alone could modulate mESC differentiation.

| | Treatment | Cells seeded | | Unpaired *t*-test (vs Mock Transfection) |
|---|---|---|---|---|
| | | 200 | 400 | |
| 2° replating units | Mock Transfection | 138 ± 1 | 233 ± 16 | p=1.00 |
| | Scrambled | 129 ± 1 | 226 ± 10 | p=0.50 |
| | PmiR-134 | 92 ± 8 | 177 ± 2 | p=0.002 |
| | AmiR-134 | 119 ± 0 | 228 ± 14 | p=0.19 |
| | PmiR-134 + AmiR-134 | 124 ± 6 | 220 ± 4 | p=0.25 |

**Table 3.1. Overexpression of miR-134 reduces the colony forming efficiency of mESCs.** Colony formation assay results demonstrating that relative to mock transfection control, transfection of PmiR-134 into E14 mESCs significantly reduced the proportion of undifferentiated, colony forming mESCs in the cell population. This effect was effectively blocked by Anti-miR-134. The data is expressed as mean number of colony forming units ± standard error. Experiments were performed in triplicate thrice (n=9).

## 3.6 The mRNA expression patterns between RA-treated and miR-134-transfected mESCs demonstrate a high degree of correlation

As quantitative PCR and Western blot analyses suggested that miR-134 may enhance ectodermal differentiation, the effect of miR-134 during RA-induced differentiation of mESCs was next investigated by comparing the microarray profile of mESCs treated with Pre-miR-134 to that of RA treatment alone.

For RA-induced differentiation, total RNA was collected from untreated mESCs, and from mESCs after two and four days of RA treatment (in the absence of LIF); whereas for Pre-miR-induced differentiation, total RNA was collected three days after transfection of mESCs (in the presence of LIF) with Pre-miR-134, water only (mock transfection, MT), Scrambled oligomer, or Pre-miR-Let-7i, a microRNA that neither induces mESC differentiation nor is induced by RA. RNA samples were labeled,

hybridized to Illumina microarray chips and heat maps were generated to compare the spectrum of transcript levels altered by RA or Pre-miR-134 treatment.

There was a high degree of correlation between the expression pattern of genes altered after 4 days of RA treatment in mESCs and all Pre-miR-134 responsive genes (Figure 3.10A, B). Prior to Pearson correlation analysis, the intensity data were $\log_2$ transformed and subtracted from the medium intensity. The mean Pearson correlation coefficient between these two expression profiles (and replicates) was ~ 0.57, whereas the mean Pearson correlation coefficient between Pre-miR-134 transfection/4 days of RA treatment and other treatments (ESC/MT/Scr/Pre-Let-7i) was ~ -0.36. When comparing the means of the above Pearson correlation coefficients between RAd4 and miR-134 to the corresponding means between miR-134/RAd4 and ESC/MT/Scr/Pre-Let-7i, all p-values were < 0.001.

Analysis of mRNAs upregulated in response to either RA or Pre-miR-134 indicated increased levels of transcripts associated with ectodermal differentiation (Figure 3.10C) (Choi *et al.*, 2005; Hay *et al.*, 2004; Palmqvist *et al.*, 2005; Luo *et al.*, 2006; Tolkunova *et al.*, 2006; Yamada *et al.*, 2002), supporting the evidence presented in Figure 3.8. In contrast, levels of non-ectoderm markers were not significantly altered by either RA or Pre-miR-134 (Choi *et al.*, 2005; Hay *et al.*, 2004; Palmqvist *et al.*, 2005; Luo *et al.*, 2006; Tolkunova *et al.*, 2006; Yamada *et al.*, 2002). Transfection of Pre-Let-7i resulted in a very different pattern of gene expression compared with Pre-miR-134 treatment (Figure 3.10A).

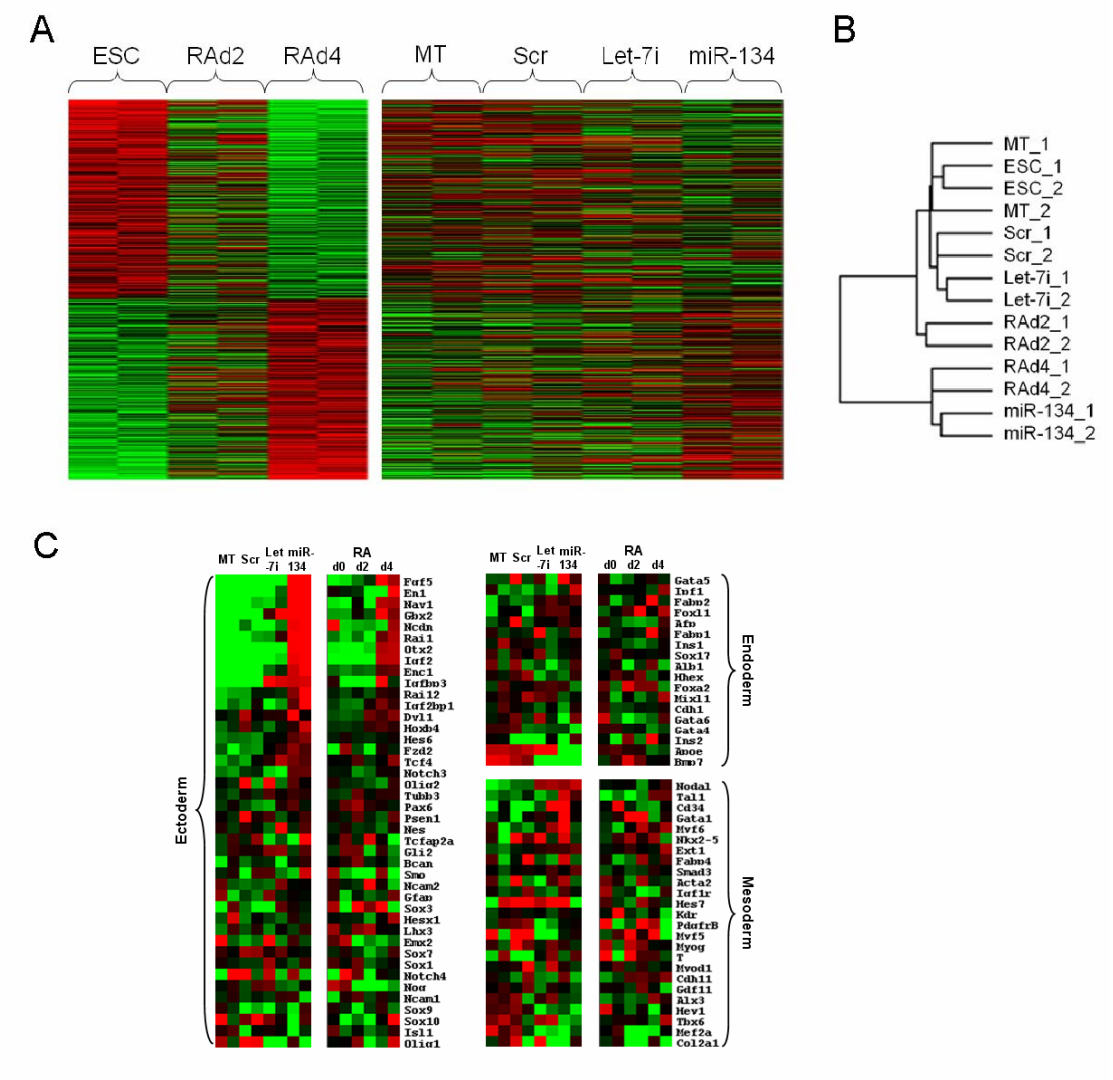



**Figure 3.10. miR-134 induces a subset of genes similar to that induced by RA treatment of mESCs (A)** Heat map depicting the expression profile of all genes with changes ≥2 fold after 4 days of RA treatment (left panel). The RAd4 dataset was ordered based on descending average intensity and the expression profile of these genes in the PmiR-134-treated dataset is shown in the same order (right panel) (**B)** Dendrogram obtained from hierarchical sample clustering showing the relatedness of all the samples profiled in (A) (**C)** Heat map depicting the expression profile of ectoderm, endoderm and mesoderm markers from RA-treated versus PmiR-134 transfected datasets. Each of the lineage markers genes are mean-centered separately in the PmiR-134 and RA-treated datasets and ordered based on the fold change in the RA-treated dataset.

## 3.7 MiR-134 enhances RA- and N2B27-mediated mESC differentiation

This microarray data which suggested that overexpression of miR-134 alone could induce the expression of a subset of genes similar to that induced by RA treatment of mESCs led to a study of whether Pre-miR-134 could enhance the RA-induced ectodermal differentiation of mESCs. In E14 mESCs, transfection of Pre-miR-134 into RA-treated cells resulted in accelerated acquisition of the early primitive ectoderm marker, *FgF5*, and the neuroectoderm markers *Nestin*, and *Neurogenin2* (Figure 3.11A), with a concomitant increase in *Sox1*, a promoter of neurogenesis (Ying *et al.*, 2003), relative to RA-treated cells transfected with Scrambled oligomer. The enhancement of these mRNAs by RA was blocked by Anti-miR-134 (Figure 3.11A). Pre-miR-134 enhanced the RA-induced decrease in the pluripotent marker *Oct4*, however, unlike the neuroectoderm markers, this effect was not blocked by Anti-miR-134 (Figure 3.11A). There were no major changes in the transcript levels of mesendoderm (*Gata4*, *Bmp4*) markers during these experiments (Figure 3.11A). MESCs co-treated with both Pre-miR-134 and RA exhibited lower protein levels of Oct4, Sox2 and Nanog, compared to treatment with Scrambled, Anti-miR-134 or Anti-miR-18 controls in the presence of RA (Figure 3.11B).

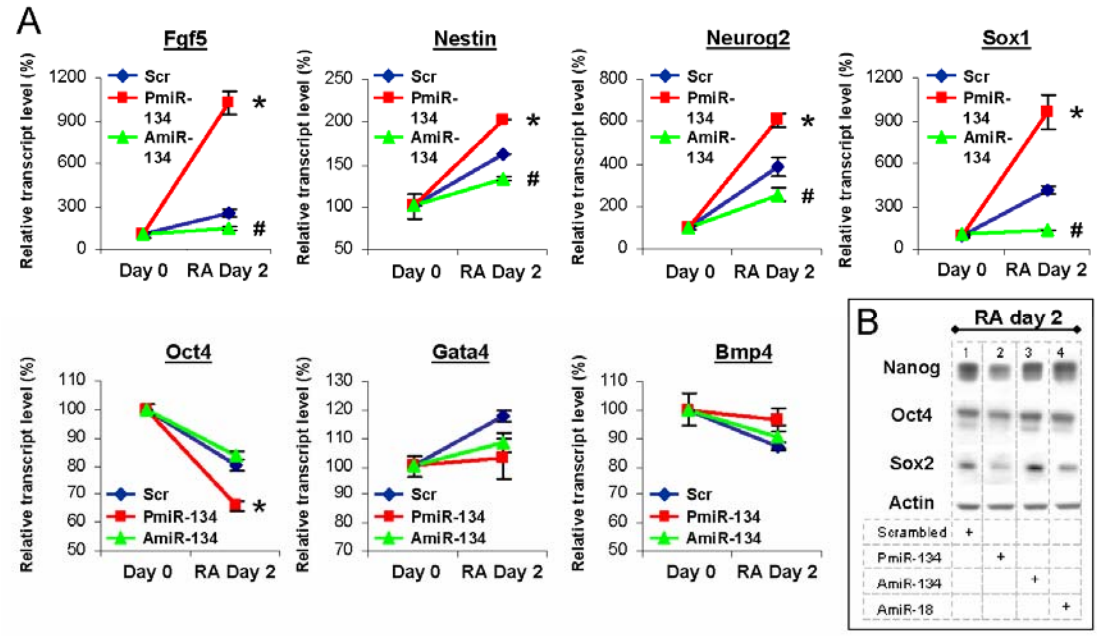



**Figure 3.11. miR-134 enhances the effect of RA on mESCs (A)** Transfection of PmiR-134 into RA-treated E14 mESCs led to an appreciable upregulation of ectodermal markers, with a concomitant reduction in the pluripotency associated marker Oct4, above that induced by transfection of Scr control into RA-treated E14 mESCs. Transcript levels of biomarkers were normalized to internal β-actin and expressed relative to day 0 control levels. Experiments were performed in triplicate thrice (n=9), where error bars denote standard error, and significance of difference to scrambled RNA oligomer control on RA day 2 at $p<0.001$ are denoted as asterisk (*) and hex ($^{\#}$) for PmiR-134 and Anti-miR-134 respectively **(B)** Western blot analyses showing that PmiR-134 enhances the RA-induced downregulation of Nanog, Oct4 and Sox2 protein levels relative to Scr control in E14 mESCs. Experiments were performed in triplicate twice (n=6), and a representative blot is shown.

As miR-134 expression also increased during N2B27 differentiation, the effect of Pre-miR-134 in the presence of N2B27 medium was tested next. Mouse ESCs were subject to Pre-miR-134 transfection for one day before N2B27 treatment for three or five days. Cell cultures transfected with Pre-miR-134 contained a larger percentage of cells positive for the neuronal marker β-III-Tubulin relative to scrambled control (Figure 3.12A, B). In addition, Cellomics analysis determined that the intensity of β-III-Tubulin immunostaining was significantly stronger in neuronal cell bodies and neurites in Pre-miR-134 treated cultures relative to Scr treated control cultures (Figure

3.12B). These data are indicative of accelerated neural induction under these conditions.

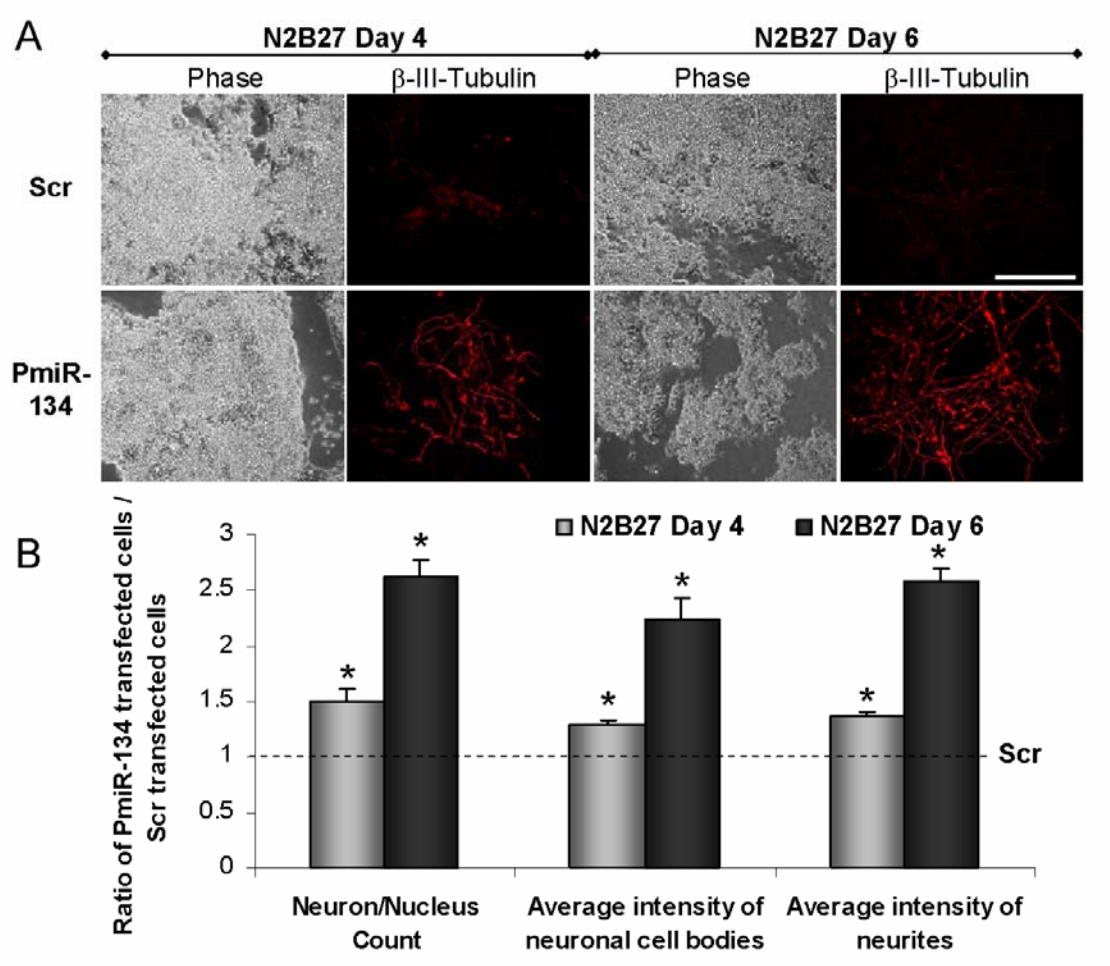



**Figure 3.12. miR-134 enhances the effect of N2B27 medium on mESCs (A)** Immunostaining of β-III-tubulin 4 and 6 days after the transfection of PmiR-134 into N2B27-treated E14 mESCs depicting that PmiR-134 transfection increases the number of β-III-tubulin positive cells compared to the Scr control transfection. Scale bar = 50 microns (**B)** Quantification of the immunostaining depicted in (C). Data shown are presented as the ratio of the respective readings for PmiR-transfected cells / Scr transfected cells, and are an average of values from 16 random fields from 3 replicate wells in two independent experiments. Error bars denote standard error, and asterisk (*) denotes significance of difference to scrambled RNA oligomer control at $p<0.0001$.

## 3.8 Discussion

This study has provided a comprehensive profiling of microRNA expression during both EB and RA-induced mESC differentiation. Although a large number of microRNAs are either up or downregulated during the differentiation of mESCs by these two methods, overexpression or reduction of expression of most of these microRNAs did not result in any significant phenotype change in mESCs. This may be explained by the hypothesis that the majority of microRNAs function in vertebrates by subtly modulating cell types instead of being inducers or regulators of differentiation (Flynt *et al.*, 2007).

Of all the microRNAs tested, miR-134 was found to have the most significant effect on mESCs. MiR-134 was identified as one of many microRNAs upregulated during RA- and N2B27-induced neural differentiation of mESCs. Interestingly, the data showed that miR-134 was not upregulated during EB differentiation. As EB formation results in a random differentiation of ESCs to endodermal, mesodermal and ectodermal lineages, the heterogeneity of cell types present in EBs may explain why an overall increase in miR-134 levels was not detected, even though its expression may increase in the population of cells differentiating towards ectoderm. This data also suggested that upregulation of miR-134 and its effect on target mRNAs is not an absolute requirement for mESC differentiation into lineages from all three germ layers, but may have a role in modulating neural lineage differentiation.

Significantly, miR-134 was the only one of the microRNAs tested that on its own modulated differentiation of mESCs, as determined by the profiling of pluripotency and differentiation markers at both transcript and protein levels, and cellular assays

such as colony forming efficiency and alkaline phosphatase staining. Furthermore, miR-134 was able to enhance RA and N2B27-induced transition of mESCs to neuro-ectodermal lineages, as measured by microarray profiling, transcript analysis and immunostaining. The observation that Anti-miR-134 was able to attenuate the RA-induced changes in *Sox1, Nestin* and *Neurogenin2* mRNA levels highlights the specificity of the results obtained and lends further support to the hypothesis that miR-134 is an enhancer of ectodermal lineage specification. However, Anti-miR-134 could not attenuate the RA-induced *Oct4* mRNA downregulation, suggesting that the elevation of miR-134 is not an absolute requirement for differentiation.

The data thus far showed that miR-134, a microRNA which was upregulated during RA- and N2B27-induced neural differentiation of mESCs, was able to modulate the differentiation of undifferentiated mESCs and enhance RA and N2B27-induced ectodermal lineage specification. These observations led to an investigation into the mechanisms by which miR-134 exerted its effects on mESCs.

# CHAPTER 4. MICRORNA TARGET PREDICTION

## 4.1 Introduction

To investigate the potential mechanisms through which miR-134 modulates mESC differentiation, it is necessary to identify potential targets which may mediate its action. As mentioned previously, the prediction of animal microRNAs binding sites, which usually exhibit imperfect complementarity with their mature microRNAs, has been a challenging task (Bentwich, 2005; Rajewsky, 2006; Segupathy *et al.*, 2006). Furthermore, at the time this study was conducted, there were only a few target prediction algorithms available. The number of validated predictions that had been reported was also small, and there was little overlap between predictions made by various algorithms (Rajewsky, 2006).

Thus, to further the dual objectives of identifying miR-134 targets which may aid in elucidating its mechanism of action, and to contribute to the development of a robust and sensitive microRNA prediction algorithm, we worked with collaborators at IBM, led by Isidore Rigoutsos, who were in the process of developing *rna22*, a method for identifying microRNA binding sites and their corresponding heteroduplexes.

*rna22* does not rely upon cross-species conservation, is resilient to noise, and, unlike previous methods, it first finds putative microRNA binding sites in the sequence of interest, then identifies the targeting microRNA. Computationally, the Rigoutsos lab has shown that *rna22* is able to predict most of the currently known heteroduplexes. A schematic depicting the various steps of *rna22* is presented in Figure 4.1.

## 4.2 Validation of *rna22*, a microRNA target prediction algorithm

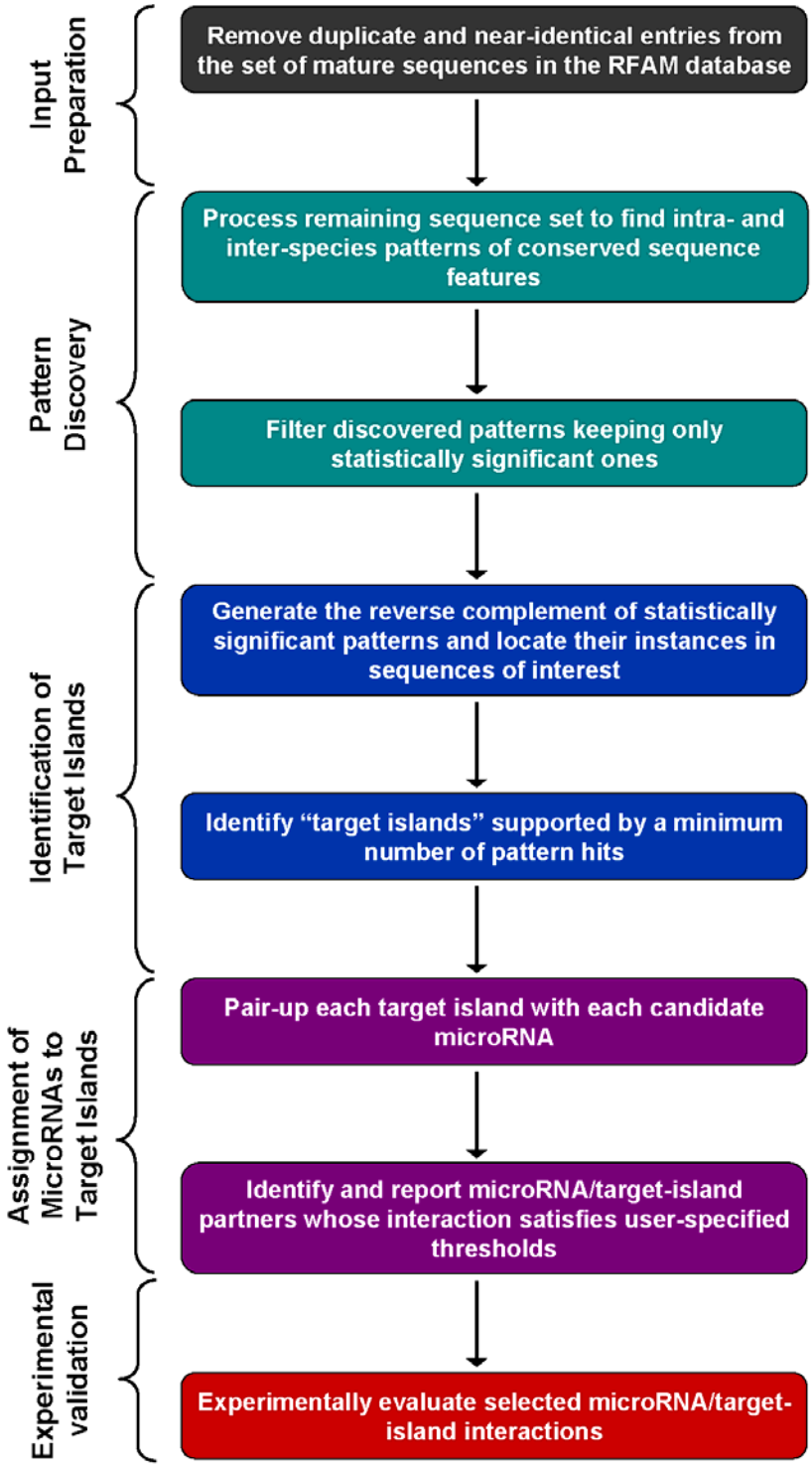



**Figure 4.1. Flowchart depicting the various steps of the target prediction method used** (adapted from Figure 1 of Miranda *et al.*, 2006)**.**

To provide experimental support for the *in silico* predictions generated by *rna22*, luciferase-reporter assays were used to evaluate randomly selected *rna22*-predicted microRNA binding sites. This was selected as the method of choice due to its high sensitivity, robustness and amenability to scale-up for high-throughput screening.

Each predicted microRNA binding site, also known as microRNA response element (MRE), was inserted as a single copy directly downstream of a Renilla luciferase open reading frame (ORF) (Figure 4.2.) Examining one predicted target at a time was an important component of the stringent evaluation approach so that any reduction in luciferase activity could be attributed to a single source. The relative luciferase activity of the control transfection (Luciferase-MRE plasmid co-transfected with scrambled RNA oligo or empty plasmid vector – represented as 100%) was compared to the corresponding activity when the cognate microRNA was added.

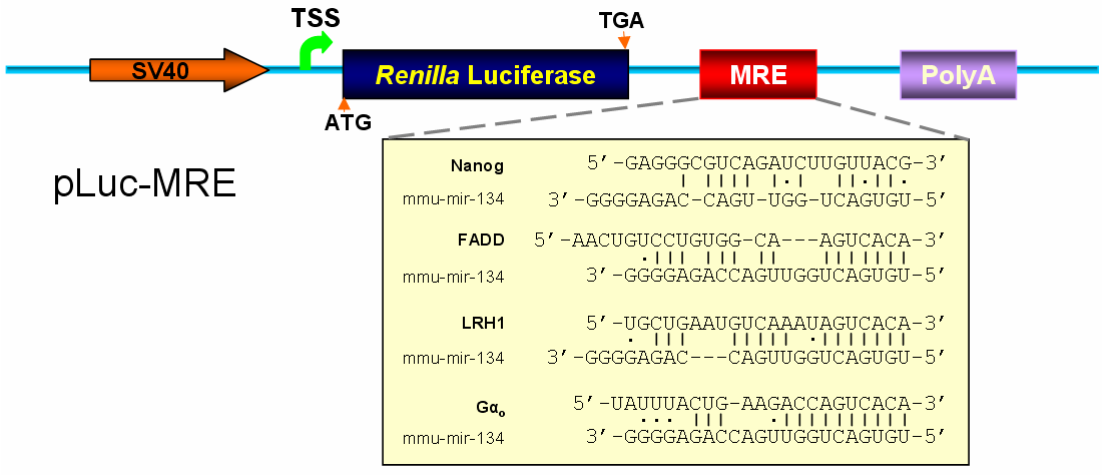


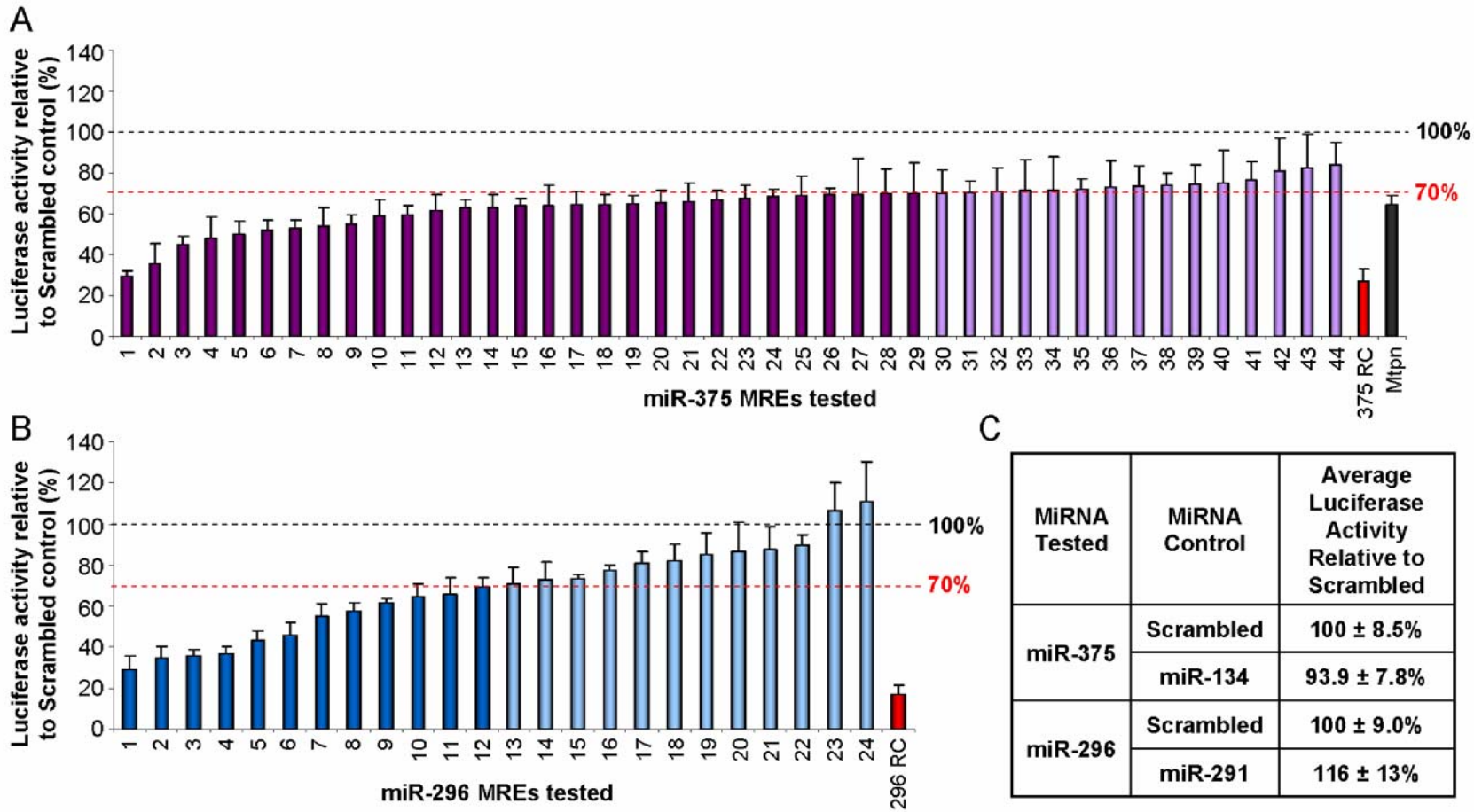

**Figure 4.2. Schematic representation of pLuc-MRE plasmid reporter construct.** Predicted MREs and binding heteroduplexes are also depicted.

In addition to miR-134, two other murine microRNAs were selected for target valdiation: miR-375 and miR-296. MiR-375 was included as its human homolog was

characterized and shown to regulate insulin secretion by binding to myotrophin (Poy *et al*., 2004). MiR-296 was included due to its possible role in ESC differentiation. Requiring a minimum of M = 14 matching base pairs between the microRNA and a target, at most G = 1 unpaired bases in the seed region, and binding energies ( E = -22 Kcal/mol for the microRNA/mRNA complex, 2292, 271 and 2318 3'UTR targets were predicted for miR-375, miR-296 and miR-134 respectively. Of these 3'UTR targets, 44 predicted sites for miR-375, 24 for miR-296, and 158 for miR-134 were randomly subselected for validation *in vitro* with luciferase assays.

As the level of target repression is related to the degree of sequence complementarity between a microRNA and its target mRNA, the reverse complement of each targeting microRNA served as a positive control. None of these three microRNAs were predicted to target the reverse complement of miR-21, so this served as a negative control. Based on the observed standard deviation about the mean (100%) of luciferase activity for negative controls, 30% repression was determined to be the minimum requirement for a positive microRNA target (Figures 4.3C and 4.4A).

Figures 4.3A, B and 4.4B show the results of these assays for miR-375, miR-296 and miR-134 respectively. Luciferase activity was suppressed by at least 30% for 168 out of the 226 tested predictions; in fact, suppression ranged between 40% and 80% for more than half of the predictions tested. Appendices 7.1-3 list the ENSEMBL ids and sequences for the 226 tested predictions. These results suggest that a high proportion of the *rna22* predicted targets may be post-transcriptionally regulated by microRNAs, and gave us confidence in the predictive power of *rna22* for our microRNA of interest, miR-134.




| MiRNA Tested | MiRNA Control | Average Luciferase Activity Relative to Scrambled |
|---|---|---|
| miR-375 | Scrambled | 100 ± 8.5% |
| | miR-134 | 93.9 ± 7.8% |
| miR-296 | Scrambled | 100 ± 9.0% |
| | miR-291 | 116 ± 13% |

**Figure 4.3. Luciferase-based validation of predicted targets for miR-375 and miR-296 (A,B)** Graph showing the effect on luciferase activity in 293T cells of PmiR and Scrambled control transfections against predicted MREs for miR-375 **(A)** and miR-296 **(B)**. Luciferase-MRE constructs containing the perfect reverse complements of each microRNA (375 RC, 296 RC) were used as positive controls (**C)** This table shows the average luciferase activity relative to Scrambled control of an unrelated PmiR against each set of predictions. These served as controls to detect any false positive downregulation in luciferase activity. Luciferase experiments were performed in quadruplicate thrice (n=12), where error bars denote standard error. See Appendix 7.1 and 7.2 for ENSEMBL identifiers and target sequences tested.

A

| MiRNA Control | Average Luciferase Activity Relative to Scrambled |
|---|---|
| Scrambled | 100 ± 7.9% |
| PmiR-21 | 104 ± 9.2% |
| PmiR-124a | 116 ± 9.8% |

B

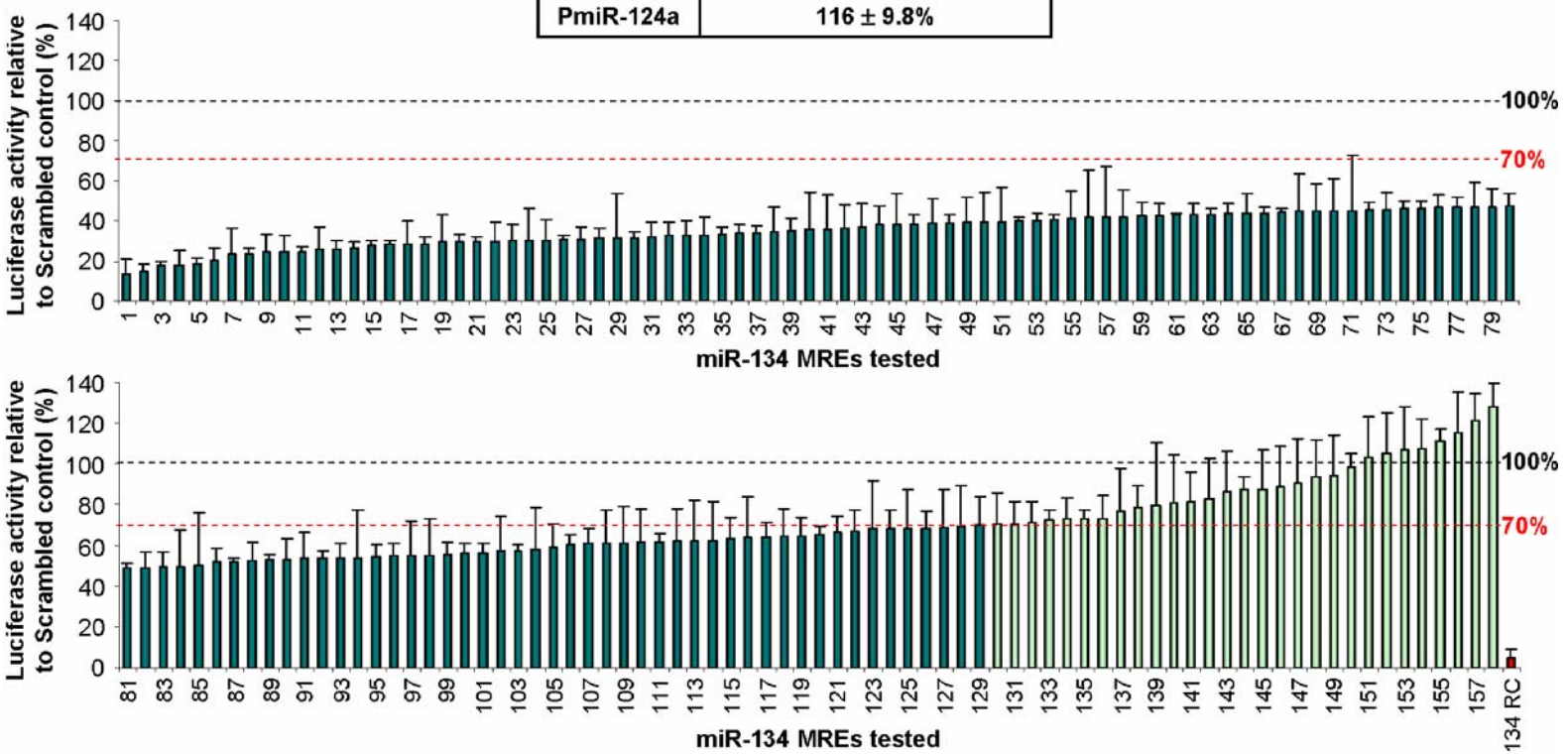



**Figure 4.4. Luciferase-based validation of predicted targets for miR-134 (A)** This table shows the average luciferase activity relative to Scrambled control of two unrelated PmiRs against all 158 predictions. These served as controls to detect any false positive downregulation in luciferase activity **(B)** Graph showing the effect on luciferase activity in 293T cells of PmiR-134 and Scrambled control transfections against predicted MREs for miR-134. A luciferase-MRE construct containing the perfect reverse complement of miR-134 (134 RC) was used as a positive control. Luciferase experiments were performed in quadruplicate thrice (n=12), where error bars denote standard error. See Appendix 7.3 for ENSEMBL identifiers and target sequences tested.

## 4.3 MiR-134 targets *Nanog* and *LRH1*, amongst other genes

Having verified *rna22*'s effectiveness, its target predictions for miR-134 were examined in greater detail. The relationship between miR-134 and the expression profile of all its predicted mRNA targets in mESCs was investigated using gene expression analysis. This revealed that 1,051 of the genes potentially targeted by miR-134 were expressed in mESCs (Figure 4.5A and B). Of these, approximately 50% were upregulated, 20% remained unchanged and 30% were downregulated relative to water only (mock transfection) and Scrambled oligomer control transfections. This suggested that miR-134 had different effects, direct or indirect, on target mRNAs, where it may affect mRNA translation, degradation and splicing.

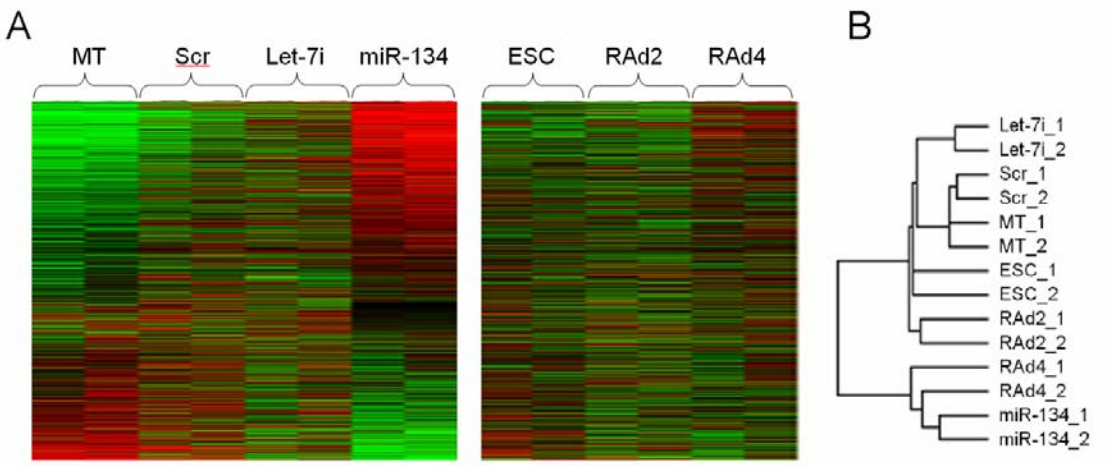



**Figure 4.5. Expression analysis of all *rna22*-predicted miR-134 target genes (A)** Heat map depicting the expression profile of all *rna22*-predicted miR-134 targets in both the PmiR-134 dataset (left panel) and the RA-treated dataset (right panel) **(B)** Dendrogram obtained from hierarchical sample clustering showing the relatedness of all the samples profiled in (A).

Interestingly, the expression profile of predicted target genes after Pre-miR-134 transfection was highly correlated with their corresponding expression profile after 4 days of RA treatment (Figure 4.5A and B). The mean Pearson correlation coefficient between these two expression profiles was 0.51, whereas the mean Pearson

correlation coefficient between predicted target gene expression in Pre-miR-134 transfection or after 4 days of RA treatment and other treatments (ESC/MT/Scr/Pre-Let-7i) was ~ -0.39. These data implied that miR-134's effect on mESCs may be mediated, at least in part, by modulation of the transcript levels of its target mRNAs.

In order to begin to understand the role of miR-134 in the post-transcriptional regulation of its potential direct targets, four predicted mRNA targets were next selected for further characterization. Upon Pre-miR-134 treatment, LRH1, $G\alpha_o$ and Nanog mRNA levels do not change and FADD mRNA levels increase. The predicted MREs in the 3'UTRs of these four genes are depicted in Figure 4.2.

The orphan nuclear receptor liver receptor homologue 1 (LRH1, also known as NR5A2), is a transcription factor that binds to the PE and PP regions of the *Oct4* promoter (Figure X) and regulates its expression (Gu *et al.*, 2005), it is also known to modulate cell proliferation through its interaction with β-catenin (Botrugno *et al.*, 2004). LRH1 is an important modulator of stem cell fate and is required for maintaining Oct4 expression in the epiblast of the embryo; loss of LRH1 leads to early embryonic lethality (Gu *et al.*, 2005). *$G\alpha_o$* (Guanine nucleotide binding protein, alpha o) and *FADD* (Fas-associated via death domain) have been shown to play key roles in neurite extension (Nakata *et al.*, 2005), apoptosis (O'Flaherty *et al.*, 2006) and embryo development (Yeh *et al.*, 1998).

For these four targets, miR-134 elicited a significant decrease in luciferase activity (Figure 4.6A). No effect was observed with Pre-miR-124a, Pre-miR-21, or Scrambled oligomer control that were not predicted to target these genes. Secreted frizzled

receptor protein 2 (Sfrp2) is shown as an example of a predicted MRE that was affected by Pre-miR-134 by a negligible amount (Figure 4.6A). Luciferase activity of miR-134 reverse complement (RC) sequence was suppressed almost completely by Pre-miR-134 (Figure 4.6A), confirming Pre-miR-134 activity in this assay.

Pre-miR-134 could specifically suppress luciferase activity when the entire *Nanog* 3'UTR was cloned within the 3'UTR of the *Renilla* luciferase reporter, verifying that the use of a single copy MRE is indicative of a microRNA's effect on the 3'UTR of a gene (Figure 4.6B). Moreover, the effect of miR-134 on Nanog's 3'UTR was abolished when its predicted MRE was mutated specifically while maintaining the integrity of the rest of the 3'UTR (Figure 4.6B and C). This demonstrates that this site is directly involved in miR-134-mediated repression of Nanog via its 3'UTR.

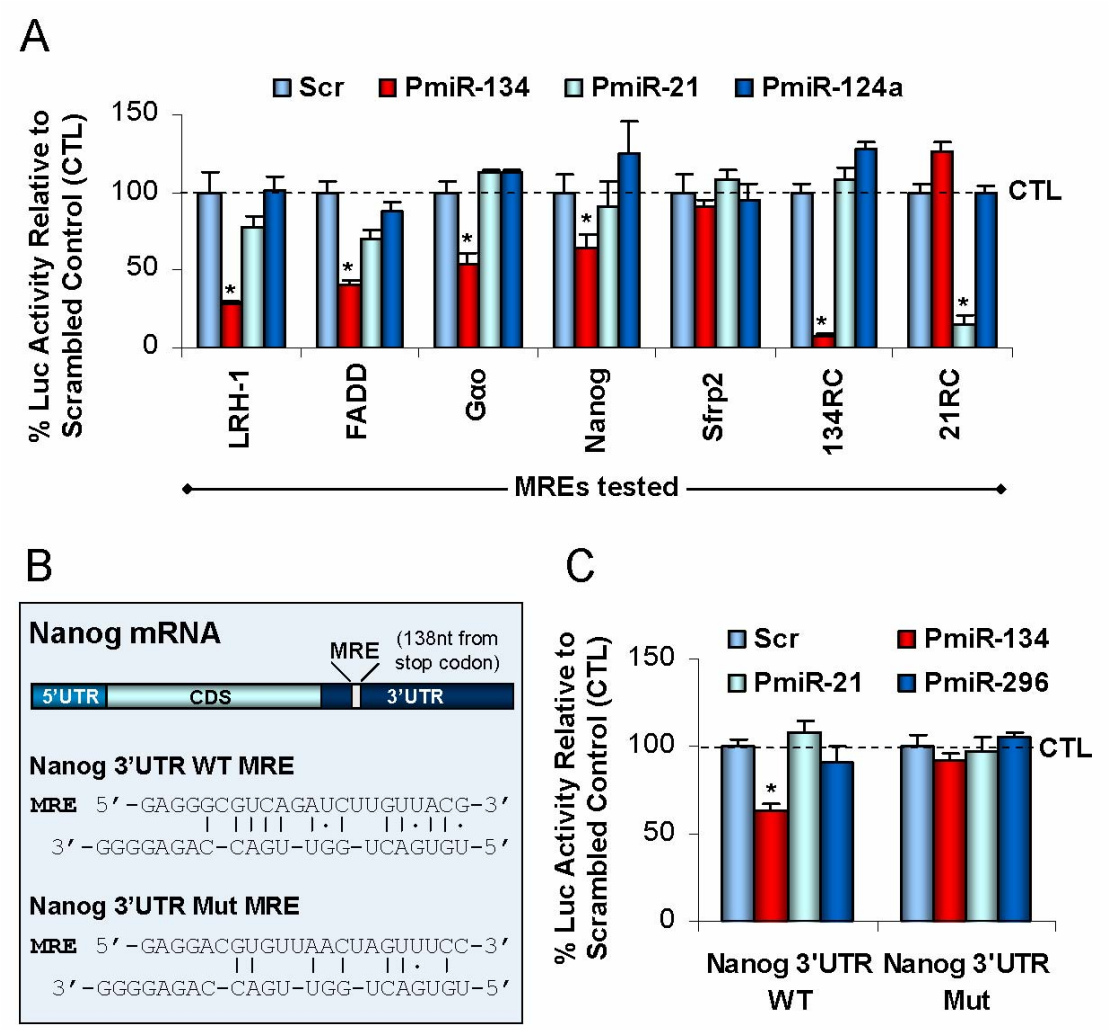



**Figure 4.6. *LRH1*, *FADD*, *Gα*$_o$ and *Nanog* are potential targets of miR-134 (A)** Graph showing the effect on luciferase activity in 293T cells of PmiR and Scr control transfections against five predicted miR-134 MREs. Relative to Scr control, PmiR-134 reduced luciferase-MRE activity of 4 of the 5 constructs, but not PmiR-21 or PmiR-124a. These unrelated PmiRs were used to detect any false positives. The PmiRs were also tested with luciferase-MRE constructs containing the perfect reverse complements of miR-134 (134 RC) and miR-21 (21 RC) as positive and negative controls respectively **(B)** Graph showing the effect on luciferase activity in 293T cells of PmiR and Scr transfections against the entire Nanog 3'UTR. Relative to Scr control, PmiR-134 reduced luciferase-UTR activity, but not PmiR-21 or PmiR-296. All luciferase experiments were performed in quadruplicate thrice (n=12), where error bars denote standard error, and asterisk (*) denotes significance of difference to scrambled RNA oligomer control at $p<0.0001$ **(C)** Sequences and binding heteroduplexes of the predicted miR-134 MRE within Nanog's 3'UTR and its mutant counterpart.

After confirming the effect of miR-134 on Nanog, LRH1, G$\alpha_o$ and FADD at the luciferase level, its effect on the endogenous protein levels of these four selected positive targets in E14 mESCs was explored next (Figure 4.7A, B). Transfection with Pre-miR-134 downregulated the endogenous protein levels of LRH1 (lanes 2 and 4), G$\alpha_o$ (lanes 10 and 12) and Nanog (lanes 14 and 16), both in the presence and absence of RA. Consistent with the induction of miR-134 by RA, RA alone also downregulated the endogenous protein levels of LRH1 (lanes 1 and 3), G$\alpha_o$ (lanes 9 and 11) and Nanog (lanes 13 and 15). Importantly, and for all three tested targets, protein downregulation was not associated with a concomitant down-regulation in mRNA levels (Figure 4.7C), suggesting that miR-134 acted, at least in part, by translationally inhibiting these mRNAs in E14 mESCs.

An intriguing finding was that miR-134 on its own led to an increase (+50%) in *FADD* mRNA levels (Figure 4.7C), whereas FADD protein levels decreased (-80%) (Figure 4.7A, lanes 5 and 6). RA induced a significant increase in *FADD* mRNA (+100%) and enhanced the effect of Pre-miR-134 (+250%) (Figure 4.7C; lanes 7 and 8), with concomitant increase in FADD protein (Figure 4.7A and B; lanes 7 and 8). This highlights that miR-134 on its own can affect mRNA levels and that its effect on protein levels can be context dependent. RA- or N2B27-mediated changes on other components of the post-transcriptional machinery (Harris *et al.*, 2004) may alter the mechanism of action by which miR-134 affects FADD translation.

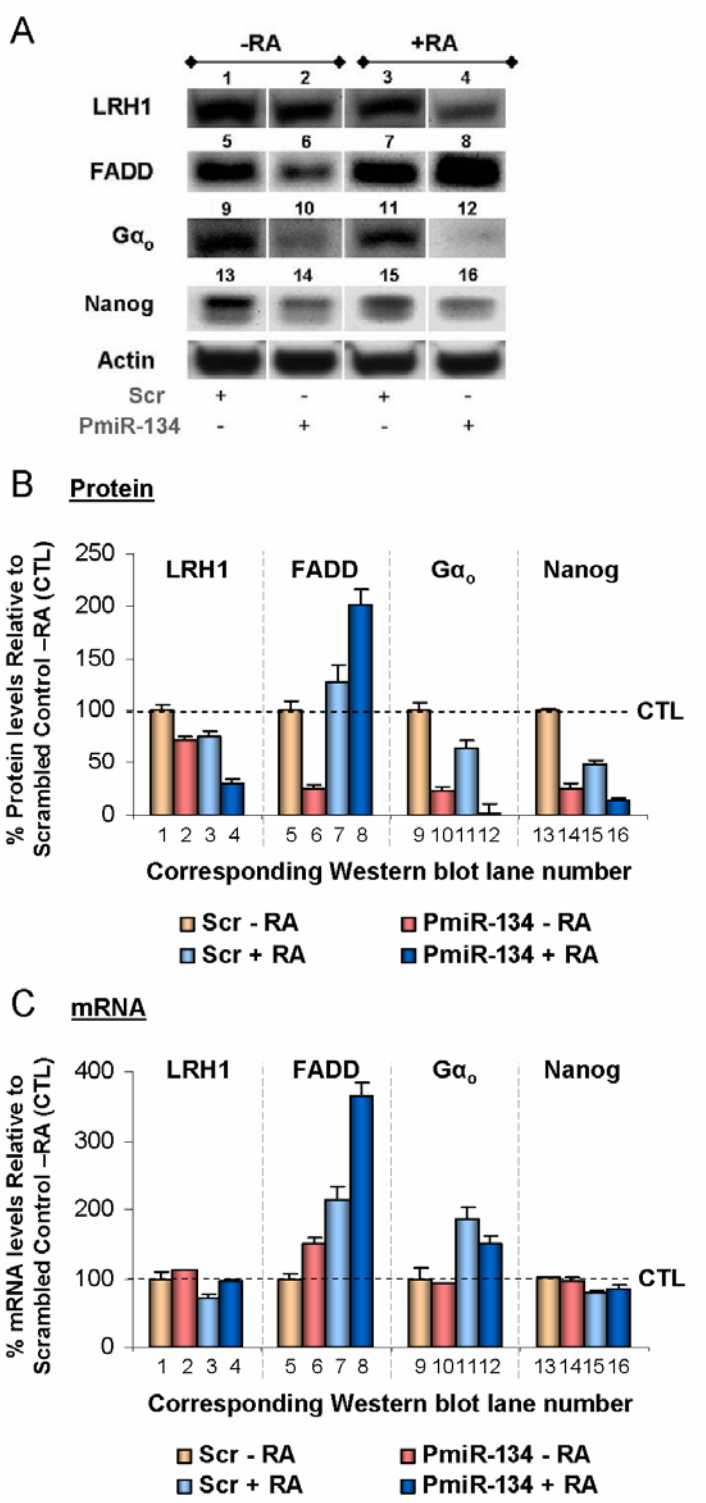



**Figure 4.7. miR-134 reduces the protein levels of predicted pluripotency-associated targets *LRH1*, *Gα$_o$* and *Nanog* without altering their mRNA levels (A)** Western blot analyses showing that PmiR-134 transfection of E14 mESCs, relative to Scr transfection, led to a reduction in the endogenous protein levels of LRH1, FADD, G$\alpha_o$ and Nanog. Experiments were performed in triplicate twice (n=6), and a representative blot is shown (**B)** Graph showing quantification of the western blots in (A) (**C)** Graph showing percentage change in mRNA levels of *LRH1*, *FADD, Gα$_o$* and *Nanog* in PmiR-134 transfected E14 mESCs ± RA relative to Scr transfection without RA-treatment. Experiments were performed in triplicate thrice (n=9), where error bars denote standard error.

## 4.4 Knockdown of miR-134 targets induces mESC differentiation

If these genes were *bona fide* targets of miR-134 that could mediate its effect, downregulation of their expression in mESCs should promote differentiation. To investigate this, *FADD*, *LRH1*, $G\alpha_o$ and *Nanog* levels were perturbed by RNA interference (RNAi) and their effects on mESC differentiation were examined. Quantiative PCR demonstrated that the respective RNAi were effective in significantly reducing transcript levels of *LRH1*, *FADD*, $G\alpha_o$ and *Nanog* (Figure 4.8A).

Knockdown of *Nanog*, *LRH1* and $G\alpha_o$ resulted in differentiation of mESCs as determined by the reduction in *Oct4* promoter activity (Figure 4.8B). To compare the effects of miR-134 and target genes RNAi on mESC differentiation, the levels of transcript markers associated with pluripotency (*Oct4*, *Sox2*) and differentiation (*Nestin*, *Otx2*, *Fgf5*, *Sox17*, *Bmp4*; Figure 4.8C) were measured. RNAi against *LRH1*, $G\alpha_o$ and *Nanog* decreased *Oct4* and *Sox2* with concomitant elevation of ectoderm-associated *Fgf5*, but had variable effects on *Nestin*, Otx2, *Sox17* and *Bmp4* (Figure 4.8C). These data indicated that although *Nanog*, *LRH1* and $G\alpha_o$ were all able to induce some degree of mESC differentiation, none of them were able to completely replicate the effect of miR-134. This suggests that they may be partial mediators of miR-134 activity, and supports the hypothesis that miR-134 acts via modulation of many target genes.

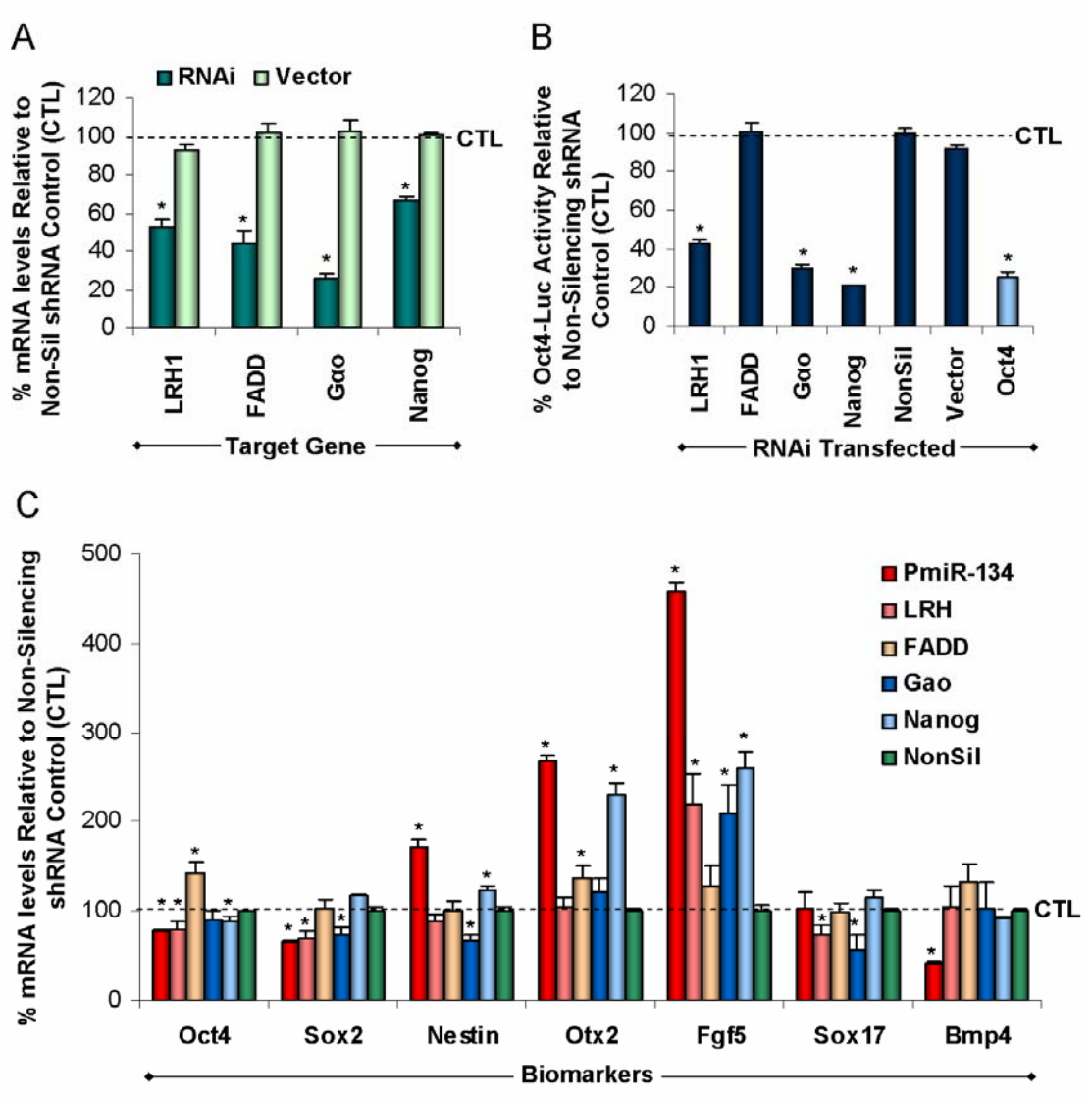



**Figure 4.8. Knockdown of *Nanog*, *LRH1* and $G\alpha_o$ results in differentiation of mESCs.**
**(A)** Graph showing that transfection of gene-specific RNAi in mESCs resulted in a reduction in target mRNAs relative to non-silencing shRNA control (CTL). No reduction was observed in target mRNAs with vector only transfection. The expression of each gene was normalized to internal β-actin, and expressed as the relative percentage change to their respective gene in the CTL. Experiments were performed in triplicate thrice (n=9), where error bars denote standard error, and asterisk (*) denotes significance of difference to non-silencing vector control at $p<0.0001$ **(B)** Graph showing changes in *Oct4* promoter activity in E14 mESCs induced by RNAi to *LRH1*, *FADD,* $G\alpha_o$, *Nanog,* and *Oct4*, relative to the non-silencing shRNA control (CTL). RNAi to *LRH1*, $G\alpha_o$, *Nanog* and *Oct4* elicited reductions in *Oct4* promoter activity. To correct for transfection efficiency, firefly luciferase activity was normalized to *Renilla* luciferase activity, and expressed as relative percentage change to that of the CTL. Experiments were performed in quadruplicate thrice (n=12), where error bars denote standard error, and asterisk (*) denotes significance of difference to non-silencing vector control at $p<0.0001$ **(C)** Graph showing transcript changes in pluripotency and differentiation markers induced by PmiR-134 and RNAi against *LRH1, FADD,* $G\alpha_o$ and *Nanog* relative to non-silencing shRNA control (CTL). Experiments were performed in triplicate thrice (n=9), where error bars denote standard error, and asterisk (*) denotes significance of difference to non-silencing vector control at $p\leq0.001$.

## 4.5 Discussion

In the last few years, diverse computational methods for predicting microRNA binding sites have been published (Table 1.3). One approach is 'signature-based', and relies on a model in which target MREs posses exact complementarity to the first 6 to 8 bases from the 5' end of the mature microRNAs. This signature is referred to as the 'seed' region (Krek *et al*., 2005; Lewis *et al*., 2003, 2005; Rajewsky and Socci, 2004). Another approach (Enright *et al*., 2003; John *et al*., 2004; Kiriakidou *et al*., 2004) utilizes variants of the dynamic programming solution to the ''local suffix alignment'' problem (Gusfield, 1997). Other algorithms make use of hidden Markov models to find seed matches (Stark *et al*., 2003), or calculate binding interactions for every offset of the microRNAs target sequence and subselect the relative placements that are significant based on an energy-based statistical measure (Rehmsmeier *et al*., 2004). Although these computational methods can be applied to individual genomes in isolation; multiple species conservation of a potential binding site at orthologous positions has typically been used to filter results.

At a time when it was thought that each microRNA only targeted a few genes (Lewis *et al.*, 2003), one of the most significant findings in this study was that a single microRNA could have numerous targets. Luciferase results in 293T cells for miR-134 and miR-375 showed at least 30% repression for 81.7% and 61.4% of the tested MREs respectively. If this ratio of success, obtained by testing randomly selected targets, is representative of the entire datasets then it follows that a large number of the predicted miR-134 and miR-375 targets may be true. In other words, miR-134 may have more than 129/158*2318 = ~1,890 targets, and miR-375 may have more than 27/44*2292 = ~1,400 targets.

If these two microRNAs are not exceptions in this regard, it is probable that many other microRNAs have numerous targets. This hypothesis is in congruence with the average number of transcripts that *rna22* estimates are targeted by a single microRNA: focusing on 3'UTR targets alone, the numbers range from several tens of targets in the worm to about a thousand in the human genome (Miranda *et al.*, 2006).

Furthermore, the experiments here were only performed in a single cell line. Based on the fact that post-trancriptional mechanisms can vary from cell type to cell type (Han *et al.*, 2000; Sammarco et al, 2007), it is a possibility that some microRNAs may only bind to certain targets in specific cell types. If this is true, then experiments in single cell lines may not encompass the entire spectrum of endogenous targets of a particular microRNA. This will be explored further in the next chapter.

With regards to miR-134, there were 2,318 *rna22*-predicted mRNA 3'UTR *cis*-element targets in the mouse genome for miR-134, of which 129 out of 158 were confirmed as positive (>80%). Over one thousand genes expressed in mESCs comprised miR-134 potential targets. This implies that the mechanism by which miR-134 post-transcriptionally regulates mRNA to enhance mESC differentiation is complex.

This hypothesis is supported by the microarray data, which demonstrated that miR-134 is able to induce the global modulation of transcript levels of its potential target mRNAs. Interestingly, RNAi of any single target of miR-134 was not sufficient to replicate the Pre-miR-134 induced changes in mESC transcript levels, again suggestive of a wide-ranging miR-134 target network. Together, these data indicate

that RA-induced elevation of miR-134, or exogenously elevated levels of miR-134 alone may promote mESC differentiation through the co-ordinate post-transcriptional regulation of a potentially large target-gene pool, that includes the pluripotency-associated target genes such as *Nanog*, *LRH1* and $G\alpha_o$. Interestingly, the sequences of the predicted miR-134 target sites in LRH1 and $G\alpha_o$ are conserved in mouse, rat and human; and mouse, rat, human, dog and chicken respectively.

Nanog and LRH1 are known regulators of mESC self-renewal where their interaction with specific promoter regions is integral to the transcriptional regulation of many genes, including Oct4 (Loh *et al.*, 2006). Although *Nanog* mRNA levels remained unchanged in E14 mESCs during Pre-miR-134 treatment over 3 days, but decreased in D3 mESCs, this may reflect subtle differences across cell types, or that D3 mESCs may be more sensitive to elevated levels of Pre-miR-134 than E14 ESCs.

Gene expression analysis demonstrated clearly that amongst the constellation of genes responding to RA treatment in mESCs, a subset of genes responded in the same manner to Pre-miR-134 transfection. Antagonists to miR-134 ablated the effects of miR-134. The data add to growing evidence that miRs function not only by translational inhibition of target mRNAs, but can influence changes in levels of non-target mRNAs. Recent reports show that RNAi can lead to an upregulation of target mRNAs in some cases (Li *et al.*, 2006). Intriguingly, miR-134 on its own led to an increase in *FADD* mRNA levels, whereas FADD protein levels decreased. RA induced a significant increase in *FADD* mRNA and enhanced the effect of PmiR-134, with concomitant increase in FADD protein. As RA induces multiple changes in ESCs, this highlights that the effect of a microRNA may be context dependent.

Thus far, the predicted targets studied have shed some light on the mechanism by which miR-134 modulates mESC differentiation. However, results in chapter 3 also suggested that miR-134 was able to enhance the differentiation of mESCs into neural lineages. As the validated targets thus far could not explain this phenomenon, this led to the hypothesis that miR-134's effect on neural differentiation was mediated by additional targets that had not been identified yet.

# CHAPTER 5. MIR-134 IN NEURAL DEVELOPMENT

## 5.1 Introduction

Aside from this study, the only reported results that describe a functional role for miR-134 were reported by Schratt *et al*. They found that miR-134 was brain-specific and localized to the synaptodendritic compartment of rat hippocampal neurons, where it negatively regulated the size of dendritic spines. This effect was mediated by miR-134 inhibition of the translation of Limk1, a protein kinase that controlled dendritic spine development. Exposure of neurons to extracellular stimuli such as brain-derived neurotrophic factor was able to relieve miR-134 inhibition of Limk1 translation, and may contribute to synaptic development, maturation and/or plasticity (Schratt *et al.*, 2006).

These findings indicated that miR-134 regulated synapse formation in mature, postnatal, rat hippocampal neurons. However, its role in neural cell differentiation at earlier stages had not been investigated. To gain better insight into this phenomenon, a list of genes that have been implicated in neural differentiation and development were selected from the list of 2,800 *rna22*-predicted miR-134 targets for experimental validation. Additionally, as experiments thus far had been performed in mESCs, a model for studying early development *in vitro*, an investigation of miR-134's expression profile during early mouse development *in vivo* was also performed.

## 5.2 MiR-134 may target Chrdl and Dcx, amongst other genes

In order to begin to understand the potential role that miR-134 may play in neural differentiation, 47 out of the 2318 mRNA targets predicted by *rna22* were selected for validation using the luciferase-based reporter assay. These genes were chosen based on their known involvement in neural differentiation and/or development.

Of these 47 MREs, located in a total of 41 genes, luciferase activity was suppressed by at least 30% for 17 of the targets tested (See Appendix 10.4). Figure 5.1A shows the results of these assays in 293T cells for 8 MREs which demonstrated significant repression of luciferase activity when co-transfected with Pre-miR-134. No effect was observed with Pre-miR-296 or Scrambled oligomer that were not predicted to target these genes. Chordin-like-1 MRE 4 (Chrdl1-4) is an example of a predicted MRE that was affected by Pre-miR-134 by a negligible amount (Figure 5.1A). Luciferase activity of miR-134 reverse complement (RC) sequence was suppressed almost completely by Pre-miR-134 (Figure 5.1A), confirming Pre-miR-134 activity in this assay.

To examine the effect of miR-134 on these neural MREs in a more biological context, the luciferase screen was repeated in Neuro2As, a mouse brain neuroblastoma cell line. Appendix 10.4 lists all 47 targets tested and the corresponding luciferase activities respective to Scr control in both cell lines. For the majority of the 47 MREs tested, the luciferase activities after Pre-miR-134 transfection relative to their respective Scr control transfection were strikingly similar across both 293Ts and Neuro2As. This result suggests that 293Ts are indeed a valuable model system in this

context, and that results obtained in this cell line are likely to be representative of *bona fide* microRNA:target interactions.

Figures 5.1A and B depict the 8 MREs which demonstrated significant repression of luciferase activity when co-transfected with Pre-miR-134 in both 293Ts and Neuro2As respectively (Figures 5.1A and B). These 8 MREs are found in the 6 genes *BMP8b* (1 site), *Chordin-like 1* (*Chrdl1,* 3 sites), *Doublecortin* (*Dcx*, 1 site), *Deltex-4* (*Dtx4*, 1 site), *Hoxc10* (1 site) and *Tcf12* (1 site). Dcx is a microtubule-associated protein that is highly expressed by migrating and differentiating neurons (Cohen *et al.*, 2008; Francis *et al.*, 1999; Gleeson *et al.*, 1999). It is a reliable marker of neurogenesis and neuronal precursors, and is necessary for neuronal migration, a pivotal step during the development of brain architeacure and function (Hwang *et al.*; Kawauchi and Hoshino, 2008; Verwer *et al.*, 2007). Dcx may also function in glia-to-neuron communication in the adult human neocortex (Verwer *et al.*, 2007). Dcx knockout mice exhibit severe morphological defects in the rostral migratory stream, delayed neuronal migration and have a fractured pyramidal cell layer in the hippocampus (Corbo *et al.*, 2002; Koizumi *et al.*, 2006).

Bmp8b is a member of the TGFbeta superfamily of growth factors, which promotes primordial germ cell differentiation from mESCs (Kee *et al.*, 2006). A secreted protein with significant homology to Chordin, Chrdl1, is hypothesized to be an important regulator of BMP signaling for the differentiation of mesenchymal cells and the development of specific neurons (Nakayama *et al.*, 2001). Dtx4, a mouse homolog of *Drosophila* Deltex, encodes a modulator of the Notch pathway that may be involved in the regulation of neurogenesis (Kishi *et al.*, 2001). It has been reported to

promote oligodendrogliogenesis from progenitor cells and oligodendrocyte precursor cell differentiation (Cui *et al.*, 2004) and repress neural fate (Ramain *et al.*, 2001). Hoxc10 is expressed in the lumbar spinal cord, and mutations in this gene have been shown to affect motor neuron patterning (Choe *et al.*, 2006). It is thus conceivable that regulation of any or several of these genes may modulate neural differentiation.

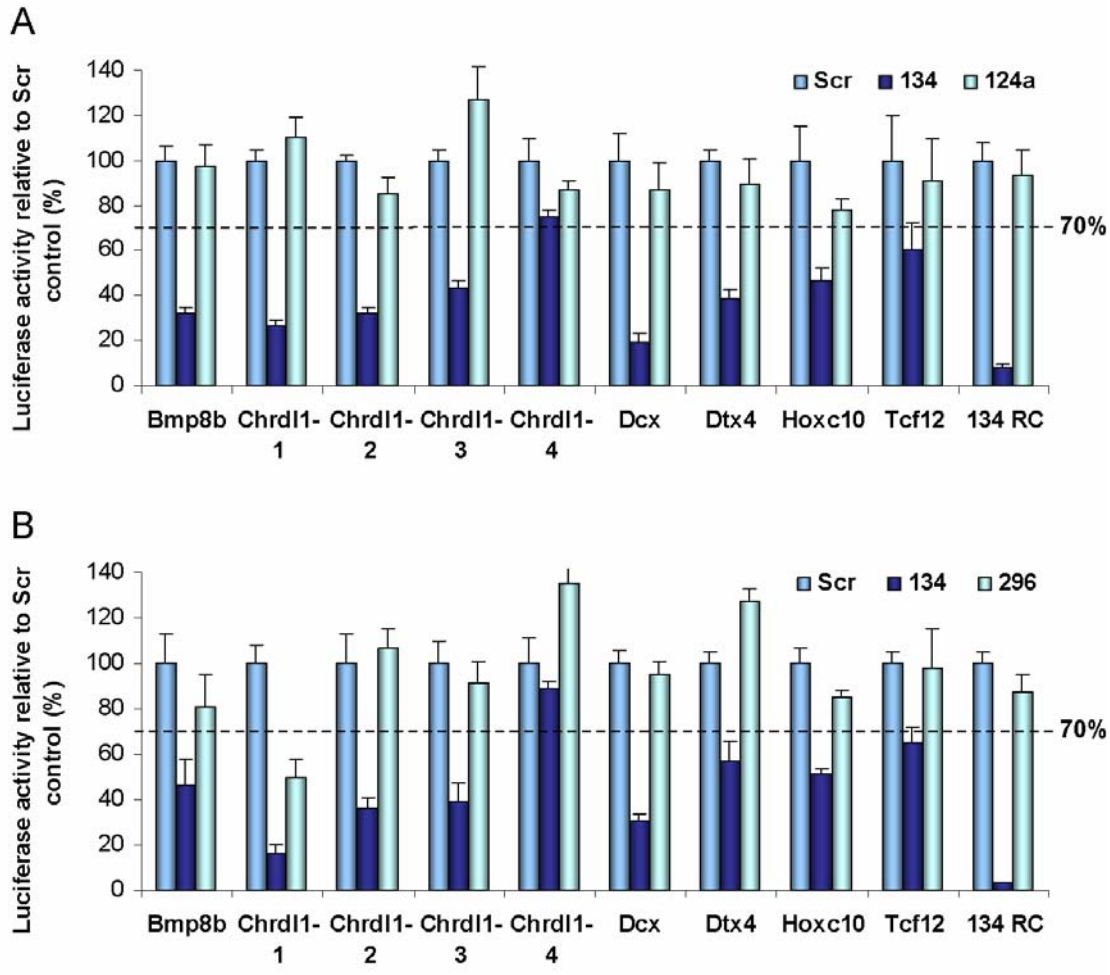



**Figure 5.1.** ***BMP8b, Chrdl1, Dcx, Dtx4 and Hoxc10*** **are potential targets of miR-134** **(A,B)** Graph showing the effect on luciferase activity in **(A)** 293T and **(B)** Neuro2A cells of PmiR and Scr control transfections against nine predicted miR-134 MREs. Relative to Scr control, PmiR-134 reduced luciferase-MRE activity of 8 of the 9 constructs, but not PmiR-296. This unrelated PmiR was used to detect any false positives. The PmiRs were also tested with the luciferase-MRE construct containing the perfect reverse complement of miR-134 (134 RC) as a positive control. All luciferase experiments were performed in quadruplicate thrice (n=12), where error bars denote standard error. See Appendix 10.4 for the entire list of 47 targets tested and the corresponding luciferase activities respective to Scr control in both cell lines.

## 5.3 Expression profiling of miR-134 in embryos and adult tissues

To gain further insight into the possible role of miR-134 in neural differentiation *in vivo*, quantitiative PCR was used to examine the temporal and spatial expression patterns of miR-134 in mouse embryos and adult tissues (Figure 5.2A). During mouse embryogenesis, miR-134 expression was detected in embryonic day (E)7.5 whole embryos at relatively low levels. This increased between E7.5 and E15.5 and decreased towards E17.5 (Figure 5.2A). MiR-134 expression was highest from E11.5 through E15.5, which coincides with major phases of spinal cord neurogenesis, neuronal differentiation, cell migration, gliogenesis, and motor neuron cell death (Choe *et al.*, 2006).

Spatial distribution of miR-134 at E12.5 indicated that miR-134 levels were highest in the midbrain compared to the cortex, hindbrain and hippocampus (Figure 5.2B). Levels of brain-specific miR-124a were lower than miR-134 in the embryonic brain. Interestingly, this pattern was reversed in the adult (postnatal day 60) brain where miR-124a levels were higher in all brain regions relative to miR-134, although the latter's levels remained comparably higher in the midbrain (Figure 5.2C). Although the significance of this switch from miR-134 to miR-124a predominance is not understood, it may represent a change in maturation from fetal to adult tissue. Intriguingly, the overexpression of several microRNAs in adult cardiomyocytes was able to reactivate fetal genes (van Rooij et al, 2006; van Rooij et al, 2007).

Screening of various adult tissues was also performed. This demonstrated that miR-134 expression was notably enriched in brain and spinal cord (as with miR-124a),

although it was also observed in additional tissues including the ovary and stomach (Figure 5.2A).

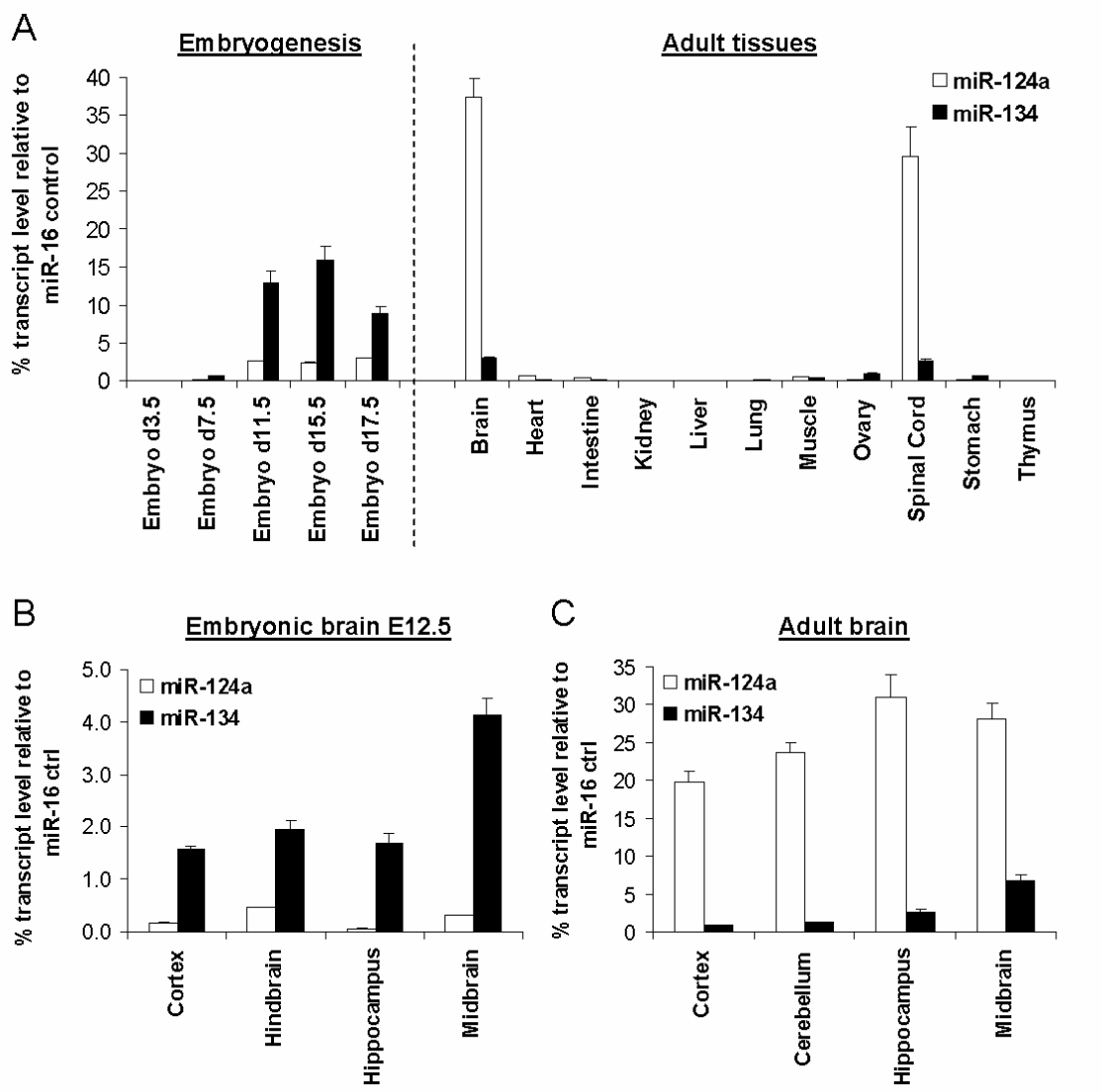



**Figure 5.2. miR-134 expression increases during embryogenesis and is present in several adult tissues (A)** MicroRNA quantitative PCR analysis demonstrating that miR-134 expression increases during mouse embryogenesis and is present in various adult mouse tissues, notably the brain and spinal cord. miR-16 expression was used as a loading control (**B,C)** MicroRNA quantitative PCR analysis showing miR-134 expression in various parts of the **(B)** embryonic and **(C)** adult brain. miR-16 expression was used as a loading control.

MiR-134 and miR-124a distribution were also examined by *in situ* hybridization (Figure 5.3). A diffuse miR-134 expression was observed throughout the E11.5 embryo, although there was significant enrichment in caudal and ventral brain regions. This was consistent with quantitative PCR measurements, showing significant miR-134 expression in hindbrain (Hb) and the ventral telencephalon / midbrain (T) relative to miR-124a at E12.5 (Figure 5.2B)

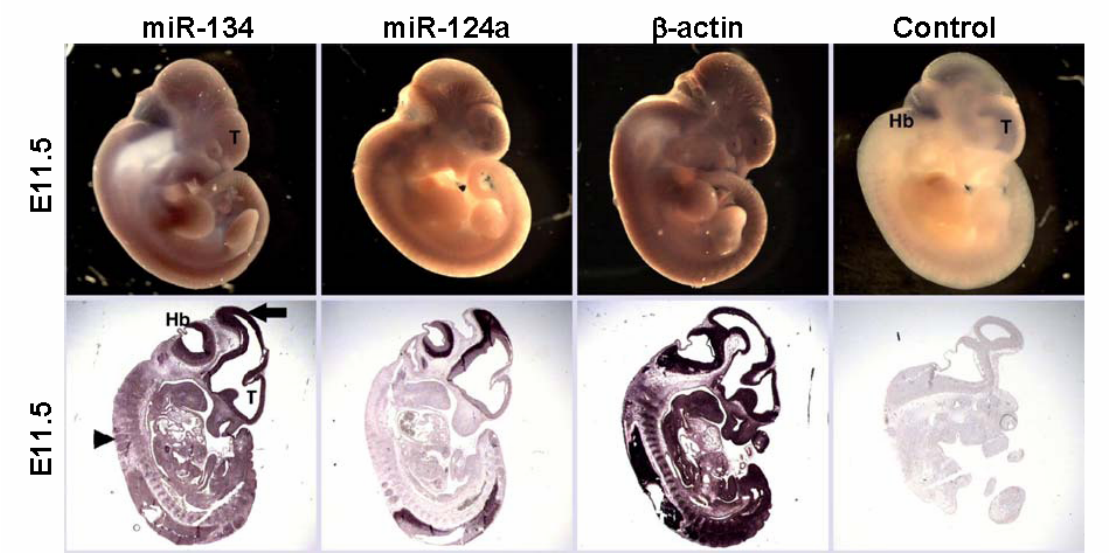



**Figure 5.3. Distribution of miR-134 expression in the E11.5 embryo.** Whole mount and section *in situ* hybridization depicting miR-134, miR-124a and β–actin expression during mouse embryogenesis. Diffuse miR-134 expression was observed throughout the E11.5 embryo, with significant enrichment in the central and peripheral nervous systems (CNS and PNS respectively).

## 5.4 Discussion

The preliminary evidence presented here suggests that beyond its role in enhancing the differentiation *in vitro* of mESCs towards neural lineages, miR-134 may also be involved in modulating neural differentiation *in vivo* or development. This may involve the following *rna22*-predicted targets that have been implicated in these processes and have been validated at the luciferase level: *BMP8b, Chrdl1, Dcx*, *Dtx4*, *Hoxc10* and *Tcf12***.**

The presence of 4 predicted MREs in *Chrdl1*, of which 3 showed very significant luciferase repression across both cell lines tested, was especially fascinating. To date, although multiple target sites in a single mRNA by a specific microRNA have been reported in plants (Arteaga-Vázquez *et al.*, 2006), similar findings have not been described in animals.

Another putative miR-134 target which demonstrated dramatic luciferase repression across both cell lines tested is Dcx. Intriguingly, reduced expression of Dcx in dendrites resulted in reduced branch points, total length and complexity (Cohen *et al.*, 2008). This finding appears to be consistent with Schratt *et al.*'s finding that miR-134 negatively regulated dendritic spine size in hippocampal neurons via *Limk1* (Schratt *et al.*, 2006), and suggests that miR-134 may have more substantial effects on neuronal maturation that may be cell type and/or context dependent.

In addition, *in situ* hybridization and quantitative PCR analysis indicate that the peak of miR-134 expression *in vivo* coincides with major phases of spinal cord neurogenesis, neuronal differentiation, cell migration, gliogenesis, and motor neuron

cell death (Choe *et al.*, 2006). Although at a preliminary stage, it is hoped that this combination of expression analysis and validation of neural target genes will serve as a useful launchpad to extend the developmental role for miR-134 prior to that of spine formation in terminally differentiated neurons.

# CHAPTER 6. MICRORNAS TARGETING OUTSIDE THE 3'UTR

## 6.1 Introduction

It has long been established that RBPs are essential components of translation and mRNA degradation processes, and that these RBPs bind across all regions of mRNAs (George and Tenenbaum, 2006). Therefore, as microRNAs are part of the post-transcriptional regulatory machinery, it would not be surprising that they should probably be active across all regions of a mRNA where they act in conjunction with RBP complexes.

In this final chapter, the computational evidence generated by *rna22* suggesting that microRNAs may target genes outside their 3'UTRs is revisited. *rna22* is not biased in favor of 3'UTRs as its starting point is microRNA sequences, not 3'UTR sequences. Application of *rna22* to the analysis of the worm, fruit fly, mouse and human genomes resulted in the predictions that between 31% and 53% of the known 5'UTR sequences contained one or more targets, and that almost every amino acid coding sequence contained one or more targets (Appendix 10.4).

Although it is unclear whether *rna22*'s low rate of false positives for 3'UTRs predictions can be extrapolated to predictions in 5'UTR and CDS regions, the sheer number of predictions raises the distinct possibility that fairly extensive microRNA regulation may occur via the 5'UTRs and CDSs of gene transcripts in animals, in addition to 3'UTRs. To investigate this, the interactions of microRNAs with the transcription factors Nanog and Sox2 were explored.

## 6.2 MiR-296 targets the coding region of mouse Nanog

To explore the possibility of microRNAs targeting outside the 3′UTR of a transcript, *rna22* was first used to determine which of the 380 murine microRNAs in Release 9.1 of miRBase (Griffiths-Jones, 2006) formed heteroduplexes in Nanog's 5′UTR, coding sequence (CDS) or 3′UTR. The settings for the three parameters of *rna22* were as follows: the number of unmatched bases in the seed region was set to $G$=0; the minimum number of paired-up bases was set to $M$=12; and, the maximum allowed folding energy was set to $E$=-20Kcal/mol. The above choices for $M$ and $E$ were substantially more lenient than those used for 3'UTR predictions – this was an intentional choice meant to allow a broader exploration of the space of *rna22*'s predictions.

MiR-296 was predicted to target the CDS of Nanog at two distinct non-overlapping locations starting at positions 235 and 493 respectively from the translation start site. These were full-length binding sites, i.e. each predicted heteroduplex extended well beyond the span of the microRNA seed. These two sites were in close proximity and flanked the region that coded for Nanog's homeobox domain (Figure 6.1). However, the significance of this arrangement is unclear. *rna22* predicted no miR-296 binding sites in either the 5′UTR or the 3′UTR of Nanog.

Considering the leniency of parameter settings, the fraction of correct predictions in this set will likely be lower than the 75-80% obtained earlier for the 3'UTR predictions. Even considering a reduced rate of correctly predicted binding sites and the fact that microRNA expression exhibits temporal and tissue-specific aspects, the

sheer number of predicted heteroduplexes will likely represent a substantial obstacle to attempts at elucidating the full extent of Nanog's control by microRNAs.

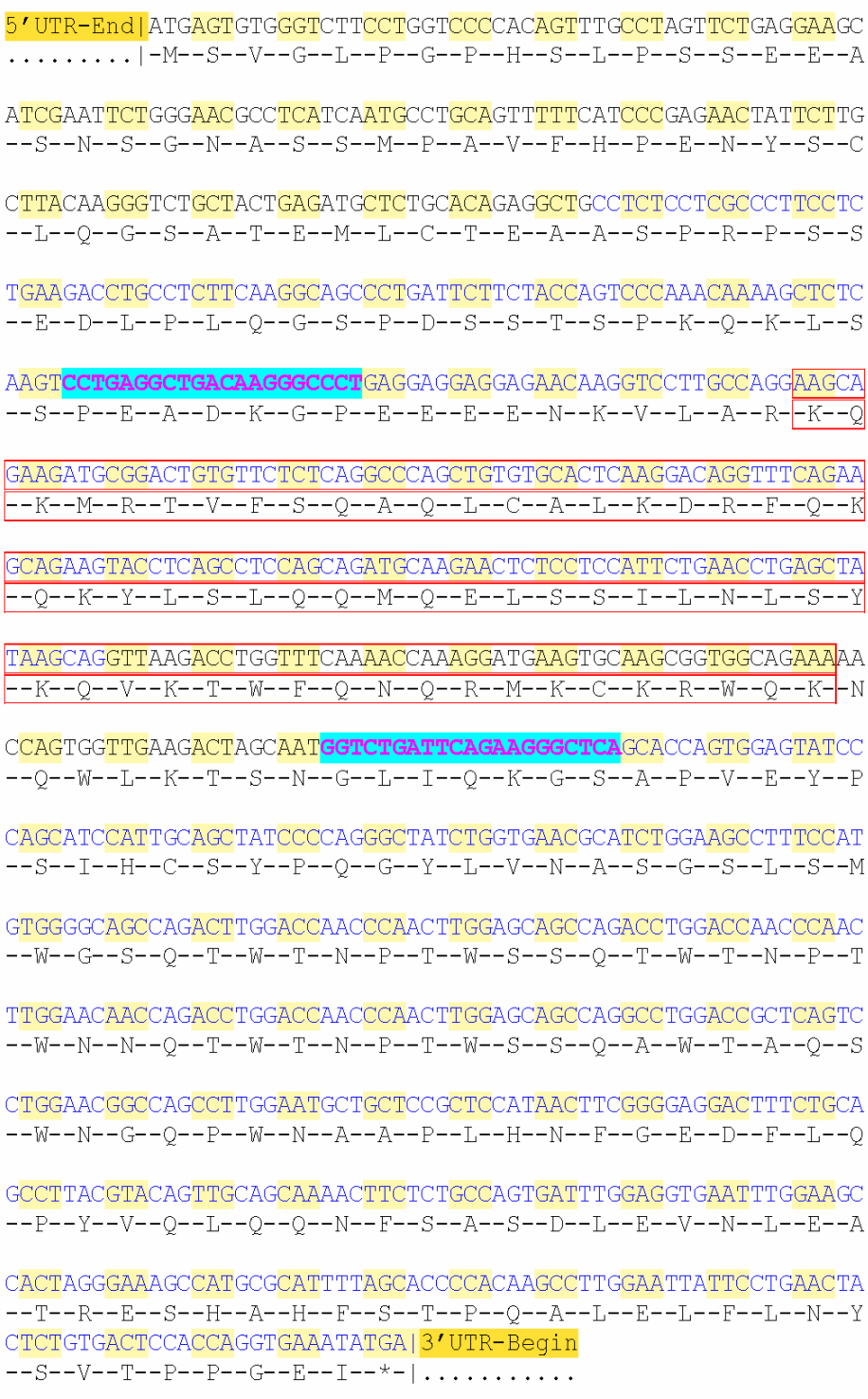



**Figure 6.1. Nucleotide sequence of Nanog's CDS region, codons and the corresponding amino acid translation.** The homeodomain is marked with a red border whereas the two predicted binding sites for miR-296, starting at positions 235 and 493 respectively from the start of translation, are shown in cyan background. Note the symmetric placement of the two binding sites and their proximity to Nanog's homeodomain. Codons are shown in alternating background color.

Luciferase-reporters were used to test the predicted putative MREs. Each MRE was inserted *as a single copy* directly downstream of a *Renilla* luciferase open reading frame (ORF) as described earlier. Each MRE was examined separately in order to ensure that any reduction in luciferase activity could be attributed to a single source. The relative luciferase activity of the control transfection (Scr, scrambled RNA oligomer – represented as 100%) was compared to the activity observed after addition of the cognate microRNA. The reverse complement of miR-296, labeled 296 RC, served as a positive control whereas microRNAs miR-21 and miR-134, which are not predicted to target Nanog-235p or Nanog-493p, served as negative controls.

Figure 6.2A shows that the luciferase activity of both MREs is significantly repressed in the presence of Pre-miR-296 but remains unchanged when Pre-miR-21 or Pre-miR-134 were transfected (negative controls). Notably, the amount of observed repression is different for Nanog-235p and Nanog-493p, suggesting that each of these MREs likely acts as an independent binding site for miR-296 in the CDS of Nanog.

Nanog's full length 5′UTR and 3′UTR were next cloned into the luciferase vector generating two different reporters. As *rna22*'s prediction that miR-134 targets Nanog's 3′UTR had been validated in chaper 4, and it is not predicted to target Nanog's 5′UTR, it was used as a negative control for the 5′UTR assay and as a positive control for the 3′UTR assay. The results in Figure 6.2B show that there is effectively no change in luciferase activity upon co-transfection of Pre-miR-296 with the Nanog 5′UTR and 3′UTR reporters, in agreement with the computational

predictions. Pre-miR-134 functioned as predicted as controls for both assays (Figure 6.2B).

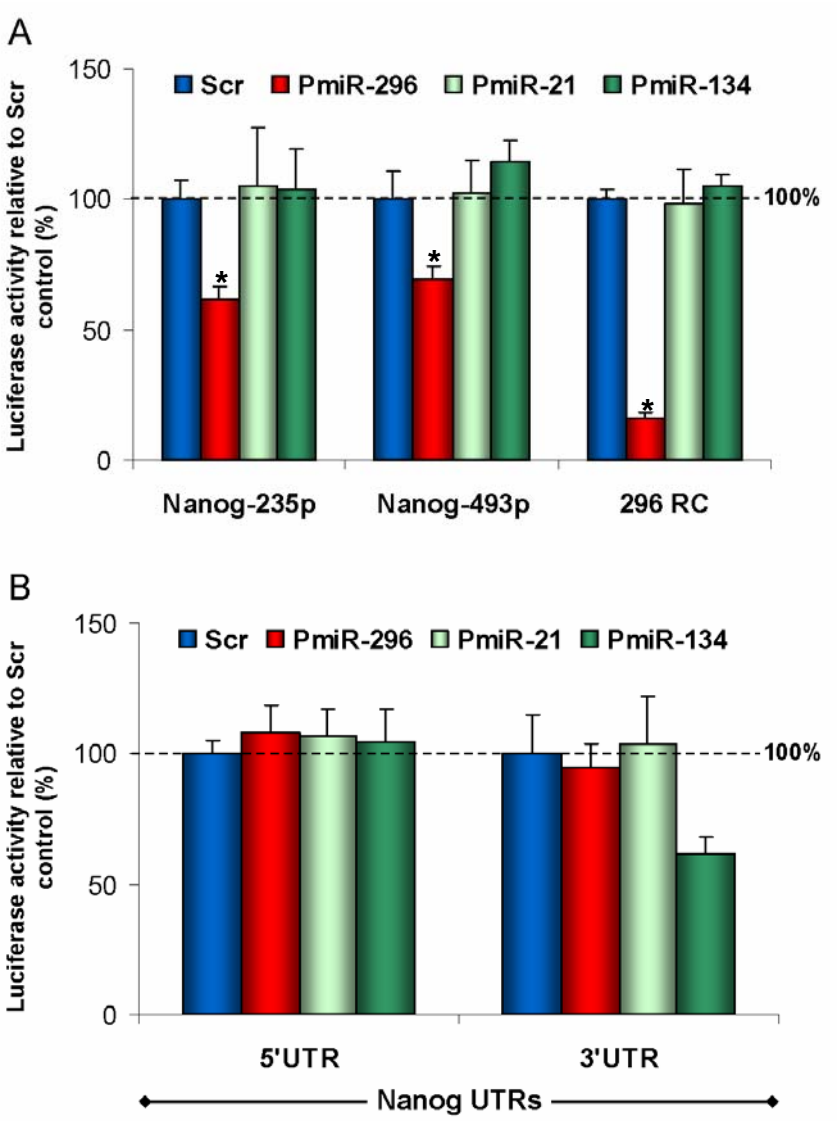



**Figure 6.2. Nanog-235p and Nanog-493p are potential targets of miR-296 (A)** Graph showing the effect on luciferase activity in 293T cells of Pre-miR and Scr control transfections against the Nanog-235p and Nanog-493p MRE constructs. Relative to Scr control, Pre-miR-296 reduced luciferase-MRE activity of both these constructs, but not Pre-miR-21 or Pre-miR-134. These unrelated Pre-miRs were used to detect any false positives. The Pre-miRs were also tested with a luciferase-MRE construct containing the perfect reverse complements of miR-296 (296 RC) as a positive control (**B)** Graph showing the effect on luciferase activity in 293T cells of Pre-miR and Scr transfections against the entire Nanog 5'UTR and 3'UTR constructs. Relative to Scr control, Pre-miR-296 did not reduce luciferase activity, thus suggesting the absence of binding sites for this microRNA in Nanog's 5′UTR and 3′UTR. All luciferase experiments were performed in quadruplicate thrice (n=12), where error bars denote standard error, and asterisk (*) denotes significance of difference to scrambled RNA oligomer control at $p<0.0001$.

To explore the impact of miR-296 on endogenous Nanog in mESCs, Western blotting of mESCs transfected with Pre-miR-296 was performed. This showed a marked decrease in Nanog protein expression (Figure 6.3A). However, no change was observed when Scrambled oligomer or Pre-miR-181a were transfected. Quantitative PCR showed only a slight decrease of Nanog mRNA in the presence of Pre-miR-296, compared to the pronounced reduction in its protein level (Figure 6.3C). These results suggest a mode of action whereby miR-296 induces posttranscriptional down-regulation of endogenous Nanog in mESCs.

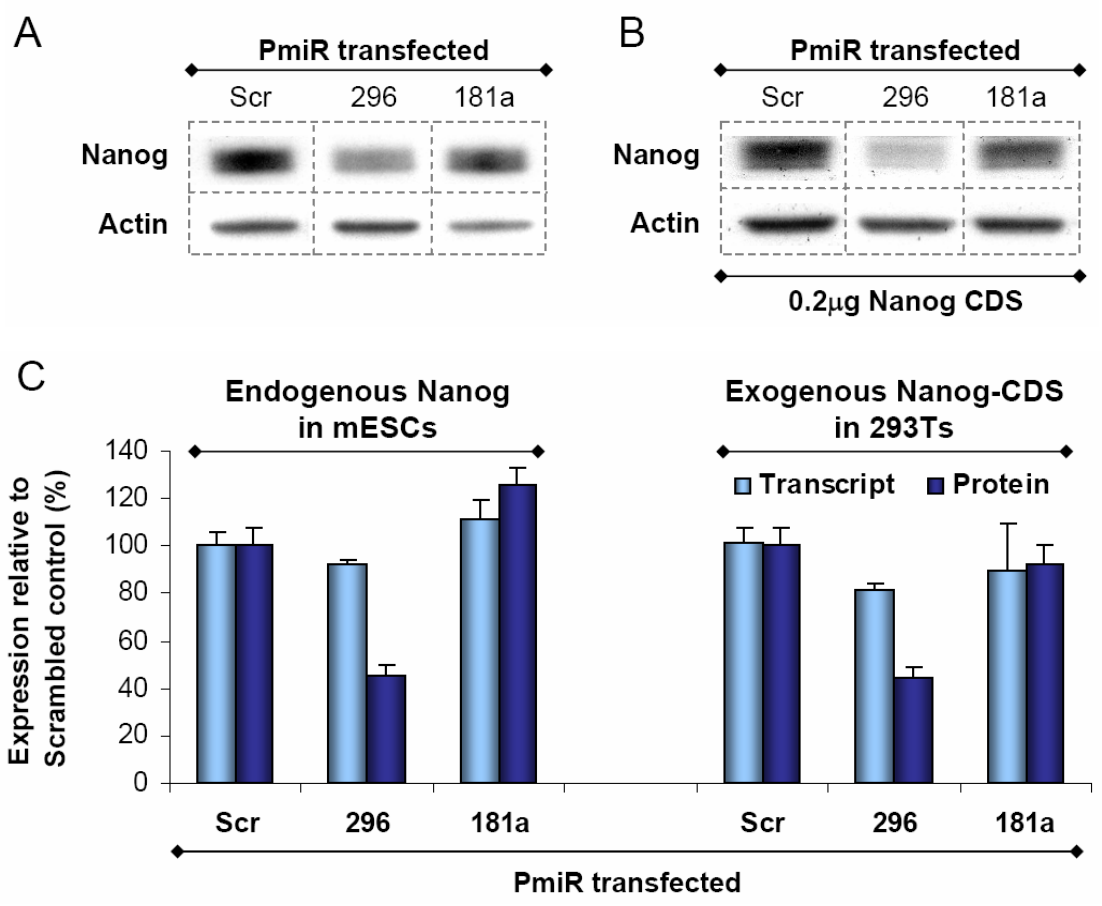



**Figure 6.3. Transfection of Pre-miR-296 reduces the amount of endogenous Nanog protein in mESCs and the amount of exogenous Nanog protein in 293T cells (A)** Western blot analysis of mESCs transfected with control Scrambled RNA oligomer (Scr), Pre-miR-296 or Pre-miR-181a (**B)** Western blot analysis of 293T cells co-transfected with control Scr, Pre-miR-296 or Pre-miR-181a and 0.2μg of Nanog CDS (**C)** Quantitative PCR analysis of mESCs transfected with control Scr, Pre-miR-296 or Pre-miR-181a, and 293T cells co-transfected with control Scr, Pre-miR-296 or Pre-miR-181a and 0.2μg of Nanog CDS, together with the quantitation of the corresponding Western blots in (A) and (B).

Earlier work has suggested the possibility of a complex regulatory interconnection involving miR-296, Nanog, Oct4 and possibly other transcription factors (Loh *et al.*, 2006). To ensure that the above findings were not the result of a complex interplay of the numerous proteins involved in the regulation of mESC pluripotency, Nanog was also studied in an environment shielded from such interactions. Western blot analysis confirmed that Nanog protein is absent from 293T cells (Figure 6.4A) thus making the latter a suitable platform for this purpose. This experiment was further optimized by transfecting various amounts of plasmid and harvesting cells at different time points. The most significant effect was observed with 0.2 µg of plasmid construct, and when cells were harvested 24 hr post-transfection, so these conditions were used for subsequent experiments (Figure 6.4).

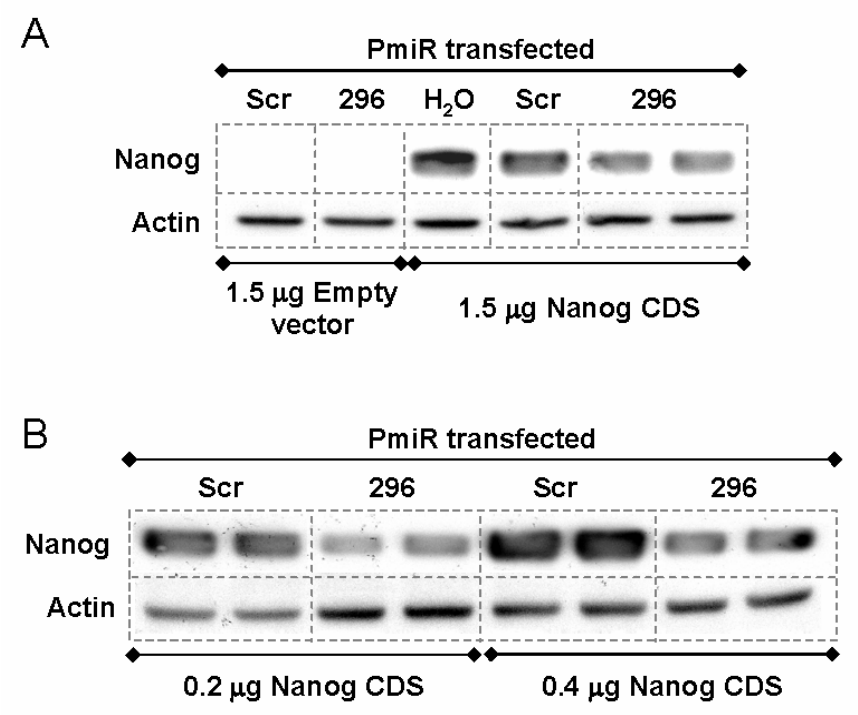



**Figure 6.4. Selection of 293T cells as a suitable cell line to study exogenous Nanog (A)** Western blot analysis of 293T cells co-transfected with water, control Scr, or Pre-miR-296 and 1.5 µg of empty vector or Nanog CDS. Cells were harvested 48 hr post-transfection. **(B)** Western blot analysis of 293T cells co-transfected with control Scr, or Pre-miR-296 and 0.2 or 0,4 µg of Nanog CDS. Cells were harvested 24 hr post-transfection.

Following co-transfection of Pre-miR-296 and the Nanog-CDS construct in 293T cells, quantitative PCR showed a slight decrease of Nanog mRNA in the presence of Pre-miR-296, whereas Western blotting showed a rather pronounced decrease in Nanog protein expression (Figure 6.3B and C). This is consistent with observations in mESCs. When the Nanog-CDS construct was co-transfected with either Scr or Pre-miR-181a (negative controls), quantitative PCR and Western blots revealed no change in Nanog mRNA or protein levels (Figure 6.3B and C).

Given that the construct we employed in 293T cells comprised only the CDS of Nanog, these findings show that miR-296 acts on Nanog's CDS region and induces post-transcriptional repression of exogenous Nanog mRNA in these cells. Even though our data suggests that this repression is predominantly through translation repression, we cannot rule out the possibility that miR-296 somehow induces an increase in Nanog protein turnover; we examine this last point next.

In order to corroborate the hypothesis that the effect of miR-296 is predominantly through translation repression of Nanog mRNA via binding to the Nanog-235p and Nanog-493p sites, silent mutations to each of these sites were introduced and studied . Since these mutations result in no amino acid changes to the Nanog protein, the observed miR-296 effects could viewed purely in the context of either translation repression or mRNA degradation.

Mutational studies involved luciferase reporter assays for the MRE mutants, and quantitative PCR and Western blot analysis of 293T cells that were co-transfected with Pre-miR-296 and mutant versions of Nanog-CDS. Several Nanog-CDS mutants

were designed as follows: a) silent mutations in the Nanog-235p site only and with the Nanog-493p site intact; b) silent mutations in the Nanog-493p site only and with the Nanog-235p site intact; and, c) simultaneous silent mutations at both the Nanog-235p and the Nanog-493p sites.

Figure 6.5A shows the results for the reporter assays and the mutant MREs containing the silent mutations: the mutant MREs have been listed and numbered in order of their increasing ability to disrupt the binding of miR-296. Five mutants of the Nanog-235p MRE and eight mutants of the Nanog-493p MRE were analyzed. With the exception of mutant Nanog-493p/m1 that contained a single-nucleotide silent mutation, all other mutant MREs contained mutations at two or more nucleotide locations. Figure 6.5B shows the sequence for each mutant MRE. As can be seen, in all cases the silent mutations disrupted miR-296's ability to suppress luciferase activity.

Figure 6.6 shows results from the Western blot and quantitative PCR analysis of 293T cells that were co-transfected with Pre-miR-296 and wild-type or mutant constructs of Nanog's CDS. Again, Scr and 296RC served as negative and positive control respectively. We used four Nanog-CDS mutants in these experiments. The mutants' labels are the same as in Figure 6.5, i.e., 235-m4 denotes the Nanog-CDS mutant whose Nanog-235p MRE has been replaced by the mutant-MRE labeled m4 in Figure 6.5 and so on.

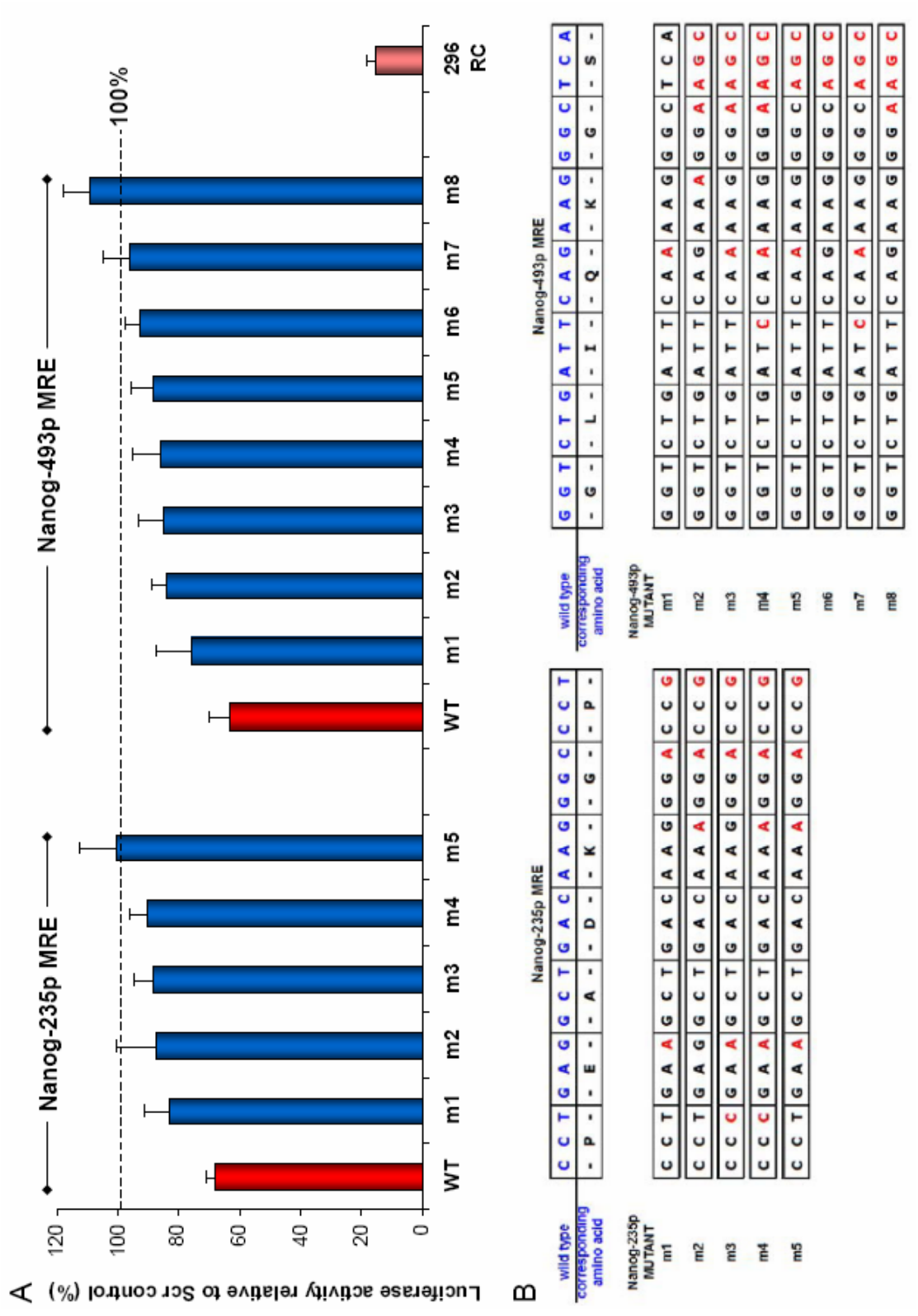



**Figure 6.5. Several MRE mutants are able to rescue the miR-296 induced reduction in luciferase activity of Nanog-235p and Nanog-493p (A)** Luciferase results for co-transfection of Pre-miR-296 with various luciferase-MRE mutant constructs; values for each construct are presented relative to their respective Scr control transfection. Results are shown for five mutants of the Nanog-235p MRE and for eight mutants of the Nanog-493p MRE (**B)** Sequences corresponding to the silent mutations we introduced in each mutant MRE – red letters show the modified nucleotides.

In concordance with the reporter assays, silent mutations to either one of the binding sites disrupted miR-296's ability to repress the translation of the mutant Nanog-CDS in 293T cells. Furthermore, introduction of silent mutations at both the Nanog-235p and Nanog-493p sites *simultaneously* (235-m4 and 493-m2) abolished the downregulation effect of miR-296 on Nanog (Figure 3.23). It is evident that these silent mutations disrupted miR-296's ability to down-regulate Nanog, thus showing that miR-296's effect on Nanog protein levels is not through altered protein turnover states. This observation in conjunction with the lack of changes in mRNA levels demonstrates that the miR-296 effects are predominantly through translation repression of Nanog mRNA via miR-296's interaction with specific CDS elements.

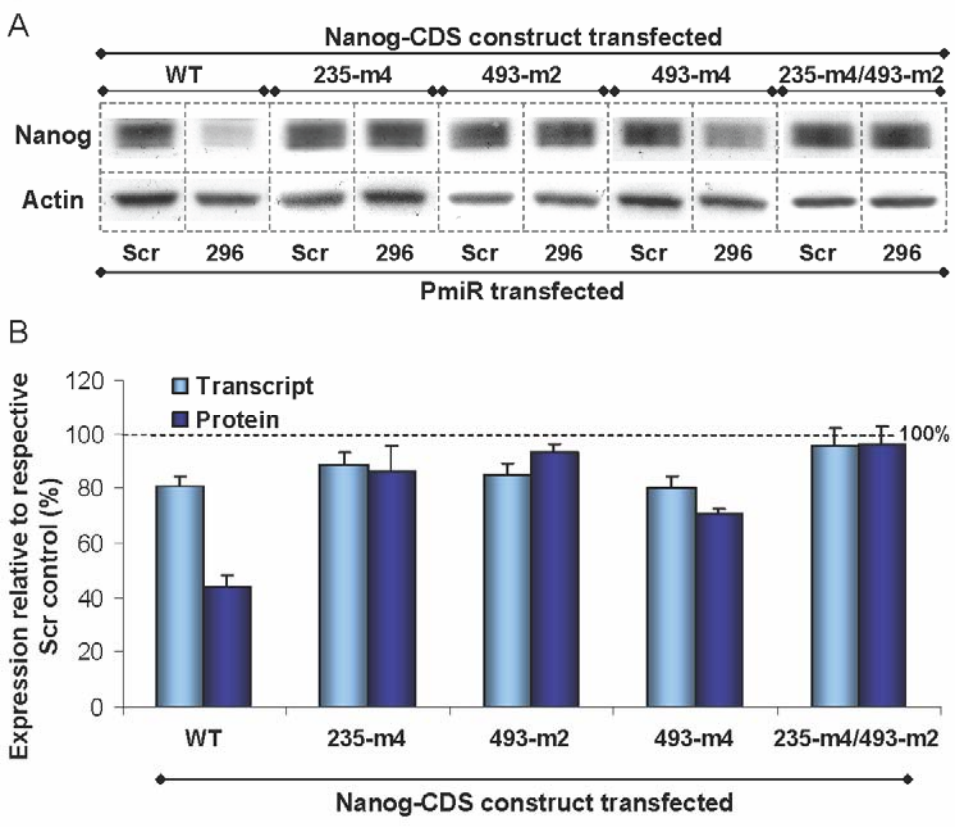



**Figure 6.6. The 235-m4/493-m2 double mutant is able to rescue the miR-296 induced reduction in Nanog protein levels (A)** Western blot analysis of 293T cells co-transfected with control Scrambled RNA oligomer (Scr) or Pre-miR-296, and 0.2µg of Nanog-CDS wild-type and mutant constructs (**B)** Quantitative PCR analysis of 293T cells co-transfected with Pre-miR-296, and 0.2µg of Nanog-CDS wild-type and mutant constructs, together with the quantitation of the corresponding Western blots; values for each construct are presented relative to their respective Scr control transfection.

## 6.3 MiR-296 modulates mESC differentiation

As Nanog is a known regulator of mESC pluripotency (Chambers *et al.*, 2003, Mitsui *et al.*, 2003), these results suggested that miR-296-modulated repression of Nanog protein levels may modulate mESC differentiation. To investigate this, miR-296 expression during mESC differentiation was profiled. Quantitative PCR analyses demonstrated that miR-296 expression increased during the differentiation of E14 mESCs. During RA induction, its expression relative to day 0 increased to about 650% after 4 days, with subsidence by day 6 (100% of day 0 expression, Figure 6.7A). During EB formation, its expression relative to day 0 increased to about 900% after 10 days, with subsidence by day 15 (400% of day 0 expression, Figure 6.7A).

It was next determined whether miR-296 could elicit phenotypic effects. E14 mESCs transfected with scrambled RNA oligomer (Scr) maintained the domed colony structures characteristic of mESCs and expressed high levels of alkaline phosphatase (Figure 6.7B). Alkaline phosphatase is expressed at higher levels in pluripotent ESCs compared to differentiated cells. However, transfection with Pre-miR-296 alone induced visible morphological changes in these cells, with the cells acquiring a flattened epithelial-like morphology typical of differentiation and expressing reduced levels of alkaline phosphatase (Figure 6.7B). Furthermore, colony forming unit assays demonstrated that there was a ~50% decrease in the ability of mESCs treated for 2 days with Pre-miR-296 to form mESC colonies compared to scrambled RNA oligomer treated controls (Figure 6.7C). This provided additional support to the hypothesis that miR-296 alone could modulate mESC differentiation.

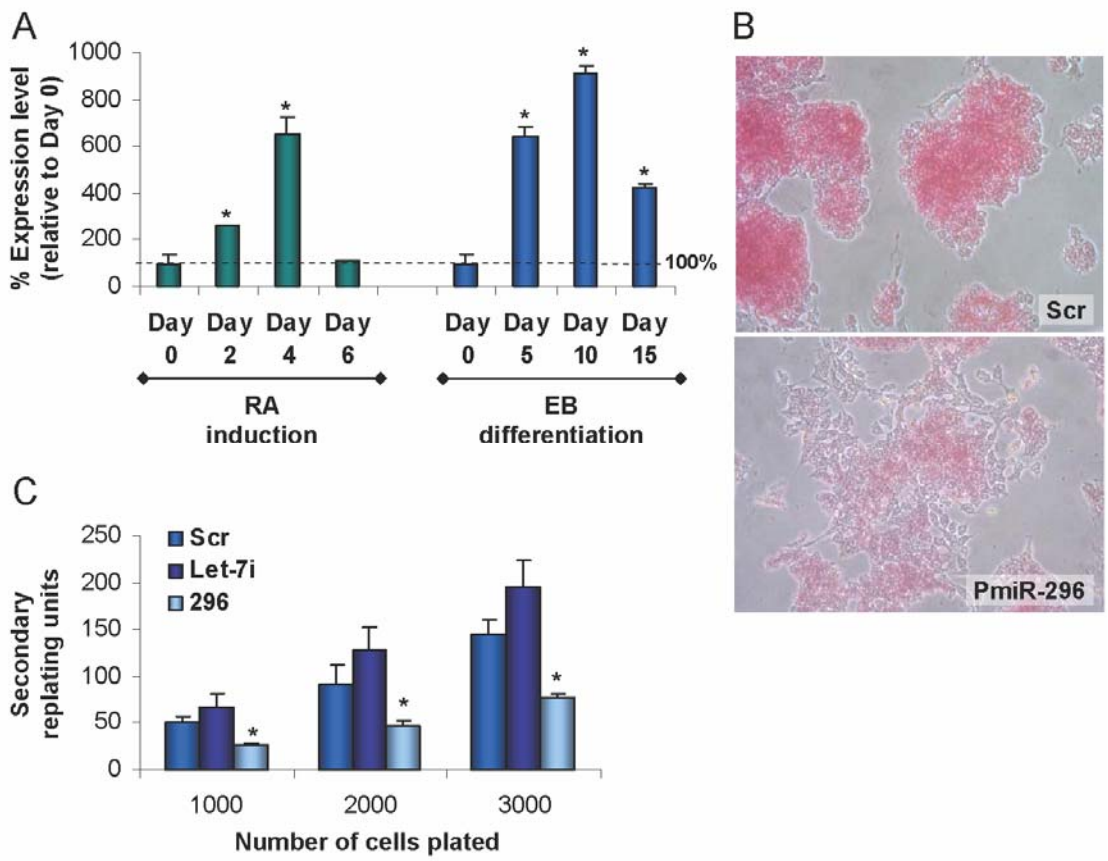



**Figure 6.7. miR-296 is upregulated during mESC differentiation, and reduces the alkaline phosphatase activity and colony forming efficiency of mESCs (A)** MicroRNA quantitative PCR analyses showing that, relative to expression in untreated E14 mESCs, miR-296 expression increases during both RA and EB differentiation of E14 mESCs. Experiments were performed in duplicate thrice (n=6), where error bars denote standard error, and asterisk (*) denotes significance of difference to day 0 at $p<0.0001$ **(B)** Photographs showing alkaline phosphatase (AP) staining of E14 mESCs three days post-transfection with Pre-miR-296 or Scr control; where Pre-miR-296, unlike Scr, induced visible phenotypic changes to a flattened, epithelial-like morphology, and reduced AP activity relative to Scr control **(C)** Colony formation assay results demonstrating that relative to Scr control, transfection of Pre-miR-296 into E14 mESCs significantly reduced the proportion of undifferentiated, colony forming mESCs in the cell population. Let-7i is an example of a microRNA that does not affect ESC differentiation. The data is expressed as mean number of colony forming units ± standard error. Experiments were performed in triplicate thrice (n=9).

The effect of miR-296 alone on cell phenotype as determined by changes in the mRNA and protein levels of key marker genes associated with pluripotency and differentiation was determined next. Quantitative PCR of total RNA samples taken from E14 mESCs three days post-transfection with Pre-miR-296 was performed. Compared to scrambled RNA oligomer treated control, transfection of Pre-miR-296

into undifferentiated E14 mESCs increases expression of the early primitive ectoderm marker, *Fgf5*, neuroectoderm marker, *Sox1*, endoderm marker, *MixlI,* and mesoderm marker *Kdr* (Figure 6.8A).

Transfection of Pre-miR-296 in RA-treated E14 mESCs resulted in accelerated acquisition of the early primitive ectoderm marker, *FgF5*, and neuroectoderm markers *Sox1* and *Otx2* (Figure 6.8A) relative to RA-treated cells transfected with the Scr control oligomer. Pre-miR-296 also enhanced the RA-induced decrease in the pluripotent marker *Nanog*, (Figure 6.8A), but did not elicit any significant changes in *Oct4* or *Sox2* mRNA levels. These results suggest that an elevated level of miR-296 in undifferentiated mESCs promotes a transcriptional expression profile that is indicative of differentiation, and enhances the differentiation of mESCs treated with RA.

As discussed previously, Western blot analyses of Pre-miR-296 transfected E14 mESCs showed that relative to Scr control oligomer control transfection, cells transfected with Pre-miR-296 express significantly decreased amounts of Nanog protein (Figure 6.8B and C). Oct4 and Sox2 protein levels do not change significantly after Pre-miR-296 transfection (Figure 6.8B and C). The quantitative PCR and western blot results suggest that the observed effect of miR-296 on mESC differentiation is not due to direct regulation of Oct4 or Sox2, at either transcript or protein levels.

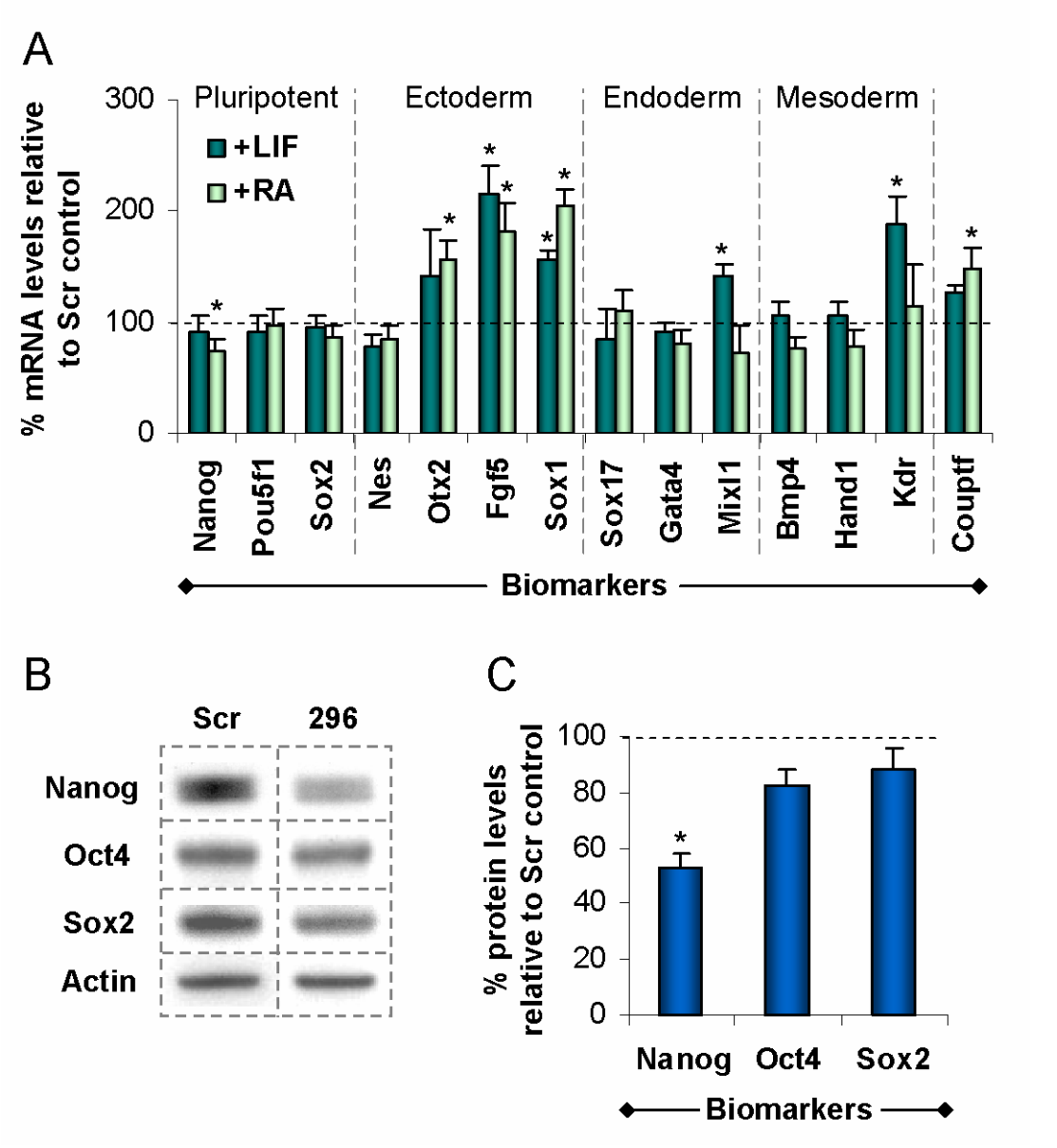



**Figure 6.8. miR-296 modulates the transcript levels of lineage-specific biomarkers in mESCs (A)** Relative to Scr control transfection, transfection of Pre-miR-296 into undifferentiated E14 mESCs led to an appreciable upregulation of the early primitive ectoderm marker, *Fgf5*, neuroectoderm marker, *Sox1*, endoderm marker, *Mix1I,* and mesoderm marker *Kdr*. Relative to RA-treated Scr control transfection, transfection of Pre-miR-296 in RA-treated E14 mESCs resulted in accelerated acquisition of the early primitive ectoderm marker, *FgF5*, and neuroectoderm markers *Sox1* and *Otx2*. Experiments were performed in triplicate thrice (n=9), where error bars denote standard error, and significance of difference to scrambled RNA oligomer control at $p<0.001$ are denoted as asterisk (*).**(B)** Western blot analyses showing that Pre-miR-296 induced significant downregulation of Nanog protein levels, but not Oct4 or Sox2 protein levels, relative to Scr control in E14 mESCs. Experiments were performed in triplicate twice (n=6), and a representative blot is shown (**C**) Quantification of the Western blots in (B).

As Nanog is an essential regulator of mESC pluripotency, the next question asked was whether the 235-m4/493-m2 double mutant Nanog, which had silent mutations introduced in both miR-296 binding sites but should function in the same way as wild-type Nanog, was able to rescue the observed effects of Pre-miR-296 on mESC differentiation. For these experiments, two independent clones of the 235-m4/493-m2 double mutant Nanog (Ng-DM1 and DM2) were used to increase the reliability of the results obtained.

Earlier results indicated that transfection of Pre-miR-296 into undifferentiated E14 mESCs increased expression of *Fgf5*, *Sox1*, *MixlI,* and *Kdr* (Figure 6.8A). These markers, together with *Couptf*, are consistently upregulated when Pre-miR-296 is co-transfected with the empty vector control (Figure 6.9A). Relative to this empty vector control, transfection of either Ng-DM1 or DM2 into undifferentiated E14 mESCs was able to rescue the Pre-miR-296 induced upregulation of *Fgf5*, *Sox1*, *Mixl1* and *Couptf* (Figure 6.9A). However, both Ng-DM1 and DM2 were unable to rescue the Pre-miR-296 induced upregulation of *Kdr*, suggesting that this may not be modulated by its targeting of Nanog via its CDS.

Furthermore, Western blot analyses showed that transfection of Ng-DM1 or DM2 into mESCs was able to rescue the Pre-miR-296 induced downregulation of Nanog protein (Figure 6.9B). Interestingly, overexpression of the wild-type Nanog CDS construct was able to elevate Nanog protein levels in the presence of Pre-miR-296 to a significantly higher level than that seen when Pre-miR-296 alone is transfected (Figure 6.9B). This may explain the observation that co-transfection of wild-type

Nanog CDS and Pre-miR-296 is able to partially rescue the Pre-miR-296 induced upregulation of *Fgf5*, *Sox1*, *Mixl1* and *Couptf* transcripts (Figure 6.9A).

Ng-DM1 and DM2 were also able to partially rescue the Pre-miR-296-induced visible phenotypic changes and reduced AP activity in mESCs (Figure 6.9C). Taken together, these results suggest that miR-296's effect on ESCs is largely mediated via its repression of Nanog via the two sites in its CDS.

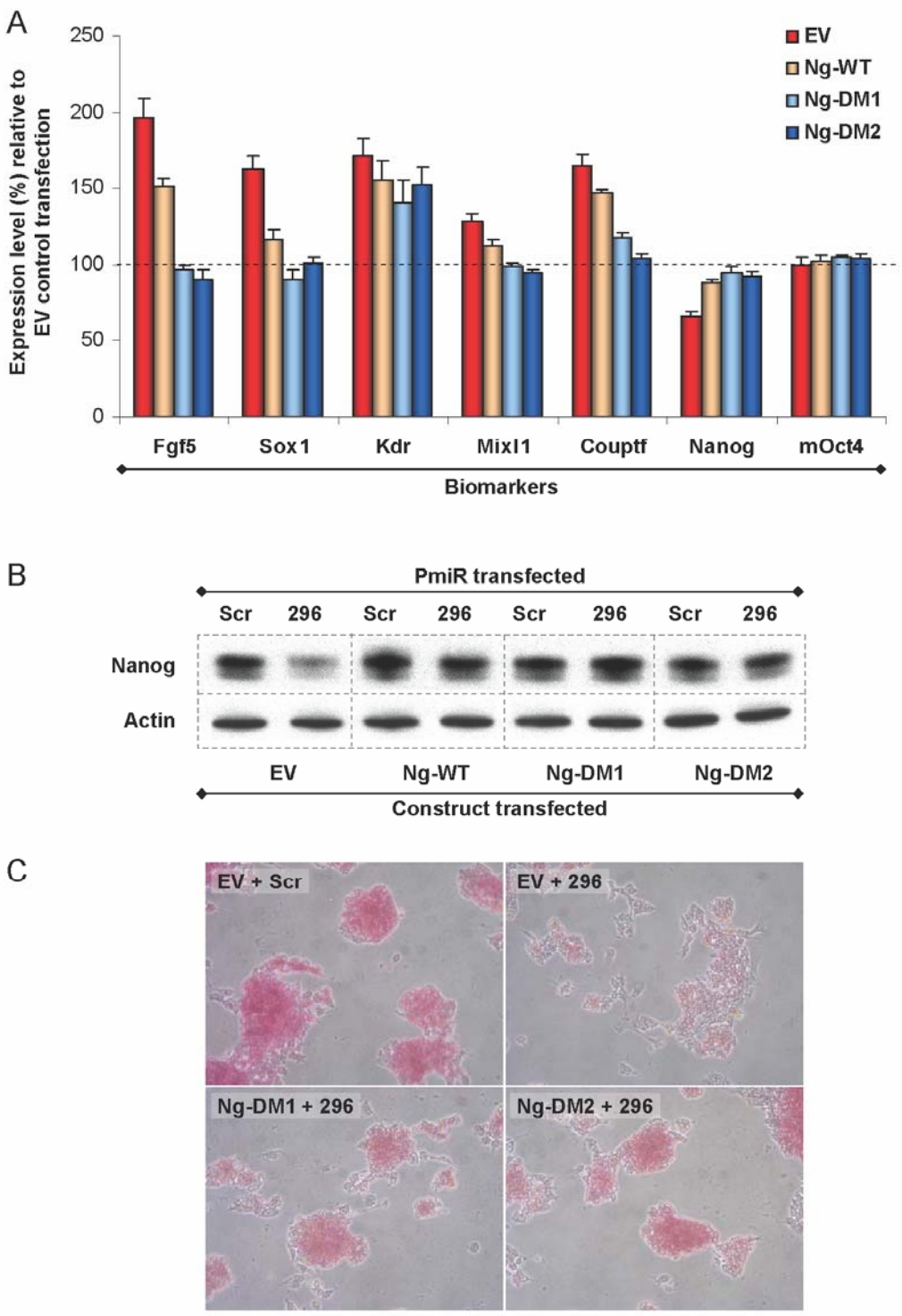



**Figure 6.9. 235-m4/493-m2 double mutants are able to rescue miR-296-modulated mESC differentiation (A)** Relative to the empty vector control, transfection of either of two independent clones of the 235-m4/493-m2 double mutant (Ng-DM1 and DM2) into undifferentiated E14 mESCs was able to rescue the Pre-miR-296 induced upregulation of *Fgf5*, *Sox1*, *Mixl1* and *Couptf*. Experiments were performed in triplicate thrice (n=9), where error bars denote standard error (**B)** Western blot analyses showing that transfection of Ng-DM1 or DM2 into mESCs was able to rescue the Pre-miR-296 induced downregulation of Nanog protein (**C)** Photographs showing alkaline phosphatase (AP) staining of E14 mESCs three days post-transfection; where Ng-DM1 and DM2 were able to partially rescue the Pre-miR-296-induced visible phenotypic changes and reduced AP activity in mESCs.

## 6.4 MiR-134 targets the coding region of mouse Sox2

To demonstrate that the targeting of Nanog's CDS by miR-296 was not an isolated incidence of microRNAs targeting genes in their coding regions, five distinct, full-length *rna-22* predicted targets for miR-134 in the CDS of mouse Sox2 were explored next. Previously, the results in chapters 3 and 4 have shown that miR-134 modulates the differentiation of mESCs, partly via the 3'UTR targeting of Nanog and LRH1.

The 5′UTR and 3′UTR of Sox2 did not contain any *rna*-22 predicted miR-134 binding sites. Luciferase reporter assays show that for the Sox2-637p site, luciferase activity is significantly repressed in the presence of Pre-miR-134 but remains unchanged when transfected with Pre-miR-181a (negative control) (Figure 6.10A). For the other four predicted CDS target sites tested, luciferase activity in Pre-miR-134 transfected cells was not significantly different from the luciferase activity in Scrambled oligomer or Pre-miR-181a transfectants (Figure 6.10A).

As shown previously, Western blot analyses of mESCs transfected with Pre-miR-134 showed a marked decrease in endogenous Sox2 protein expression (Figure 6.10B, left panel). However, no change was observed when the Scrambled oligomer or Pre-miR-181a (negative control) were transfected. Quantitative PCR showed a decrease of Sox2 mRNA in the presence of Pre-miR-134, although this decrease was less pronounced than the decrease in protein expression (Figure 6.10B, left panel). No decrease in the amount of Sox2 mRNA was observed when Scrambled oligomer or Pre-miR-181a were transfected. These results suggest a mode of action whereby miR-134 induces posttranscriptional down-regulation of endogenous Sox2 in mESCs.

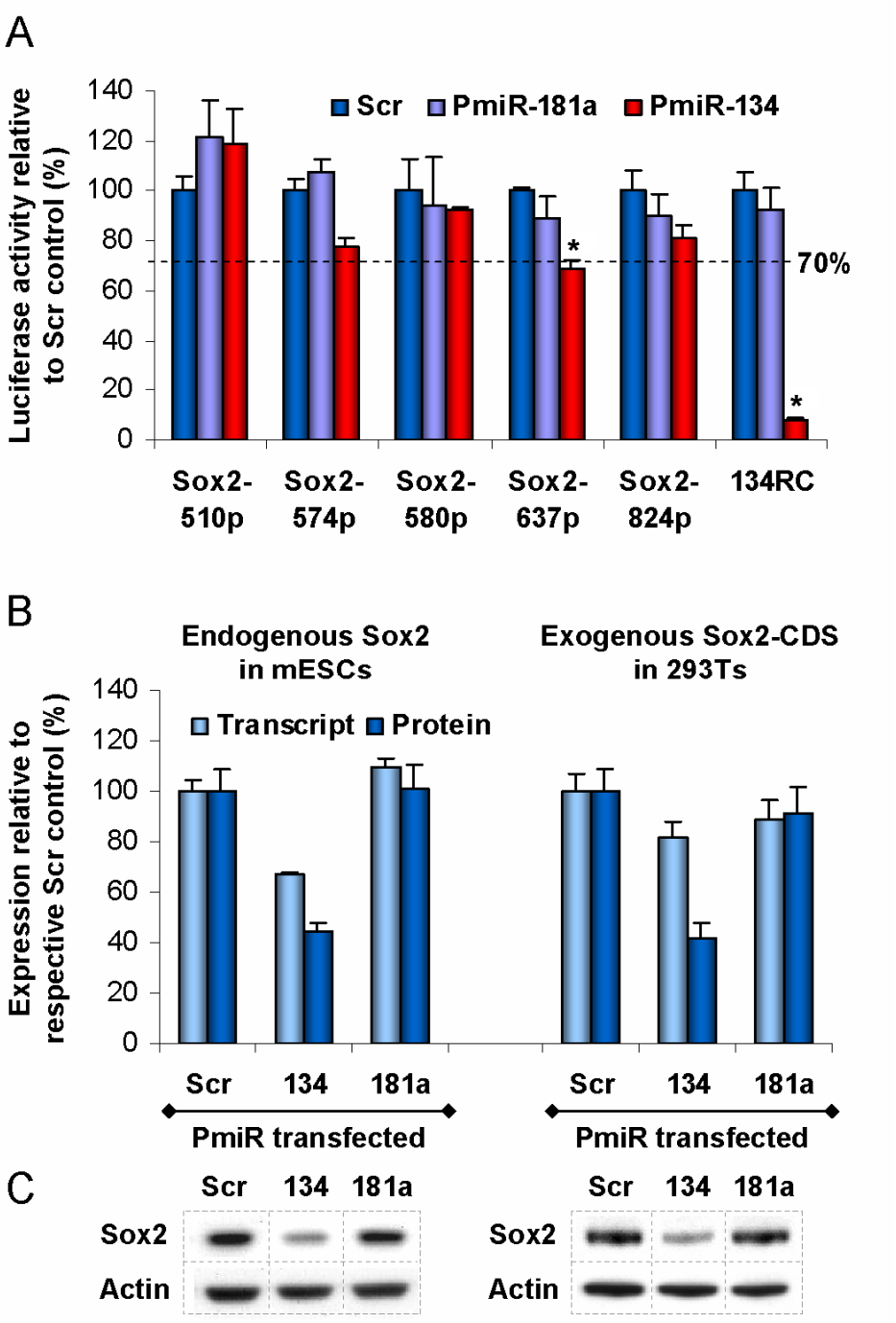



**Figure 6.10. Sox2-637p is a potential target of miR-134 (A)** Graph showing the effect on luciferase activity in 293T cells of Pre-miR and Scrambled oligomer (Scr) control transfections against five predicted Sox2 MRE constructs. Relative to Scr control, Pre-miR-134 reduced luciferase-MRE activity of Sox2-637p, but Pre-miR-181a did not. This unrelated Pre-miR was used to detect any false positives. Luciferase experiments were performed in quadruplicate thrice (n=12), where error bars denote standard error, and asterisk (*) denotes significance of difference to scrambled RNA oligomer control at $p<0.0001$ (**B)** Quantitative PCR analysis of mESCs transfected with control Scr, Pre-miR-134 or Pre-miR-181a, and 293T cells co-transfected with control Scr, Pre-miR-134 or Pre-miR-181a and 0.2µg of Nanog CDS, together with the quantitation of the corresponding Western blots in (C). (**C)** Western blot analysis of mESCs transfected with control Scrambled RNA oligomer (Scr), Pre-miR-134 or Pre-miR-181a (bottom left panel), Western blot analysis of 293T cells co-transfected with control Scr, Pre-miR-134 or Pre-miR-181a and 0.2µg of Sox2 CDS (bottom right panel).

Western blot analysis confirmed that Sox2 protein is absent from 293T cells (data not shown). Following co-transfection of Pre-miR-134 and a Sox2-CDS construct in 293T cells, quantitative PCR showed a slight decrease of Sox2 mRNA in the presence of Pre-miR-134, whereas Western blotting showed a significant decrease in Sox2 protein expression (Figure 6.10B, right panel). When the Sox2-CDS construct was co-transfected with either Scrambled oligomer or Pre-miR-181a (negative controls), quantitative PCR and Western blots revealed no change in Sox2 mRNA or protein (Figure 6.10B, right panel). Given that the construct we employed with 293T cells comprised only the CDS of Sox2, these findings show that miR-134 acts on Sox2's CDS region and induces post-transcriptional repression of exogenous Sox2 mRNA in these cells.

Silent mutations, which result in no amino acid changes to the Sox2 protein, were next introduced to the Sox2-637p site. As discussed previously for miR-296 targeting Nanog CDS, we utilized: a) luciferase reporter assays for the MRE mutants, and b) quantitative PCR and Western blot analysis of 293T cells that were co-transfected with Pre-miR-134 and mutant versions of Sox2-CDS. Figure 6.11A shows the results for the reporter assays and the mutant MREs containing the silent mutations: the mutant MREs have been listed and numbered in order of their increasing ability to disrupt the binding of miR-134. We analyzed four mutants of the Sox2-637p MRE. The table in Figure 6.11A shows the sequence for each mutant MRE. The MRE labeled 134RC was the reverse complementary sequence of miR-134 and served as a positive control. All four silent mutations disrupted miR-134's ability to suppress luciferase activity, with mutant 637-m4 being the most effective (Figure 6.11A).

Figure 6.11B shows results from the Western blot and quantitative PCR analysis of 293T cells that were co-transfected with Pre-miR-134 and wild-type or mutant constructs of Sox2's CDS. The two Sox2-CDS mutants 637-m3 and 637-m4, which demonstrated the most significant rescue of Pre-miR-134 induced repression at the luciferase level were used in these experiments. The mutants' labels are the same as in Figure 6.11A, i.e. 637-m4 denotes the Sox2-CDS mutant whose Sox2-637p MRE has been replaced by the mutant-MRE labeled m4 in Figure 6.11A and so on. In concordance with the reporter assays, the 637-m3 mutant disrupted slightly miR-134's ability to repress the translation of the mutant Sox2-CDS in 293T cells, and the 637-m4 mutant abolished completely the downregulation effect of miR-134 on Sox2 (Figure 6.11B). Furthermore, for the 637-m4 mutant, Sox2 mRNA levels after Pre-miR-134 co-transfection did not change relative to the Scr co-transfection. These observations demonstrate that miR-134's effect on Sox2 protein is predominantly through translation repression of Sox2 mRNA via interaction with a specific CDS element, and provides an additional example of microRNAs targeting genes in their coding regions.

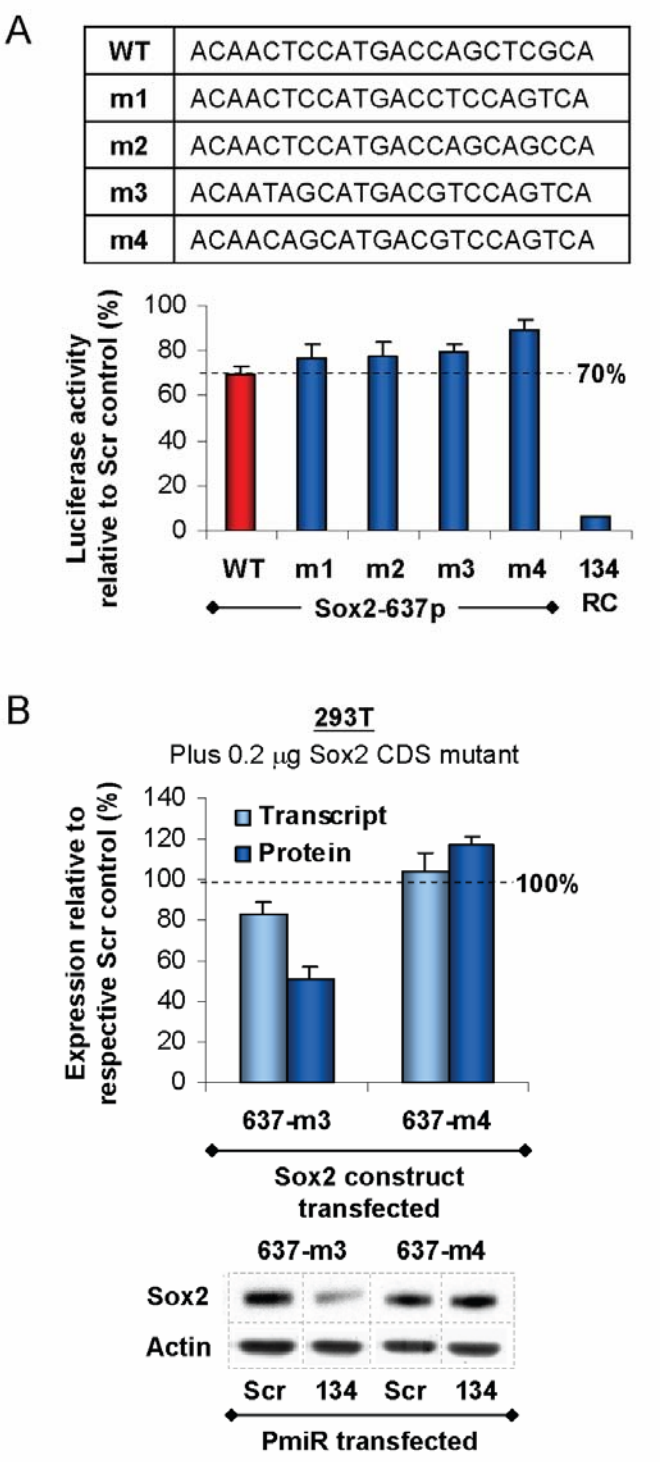



**Figure 6.11. The 637-m4 mutant is able to rescue the miR-134 induced reduction in Sox2 protein levels (A)** Luciferase results for co-transfection of Pre-miR-134 with various luciferase-MRE mutant constructs; values for each construct are presented relative to their respective Scrambled oligomer (Scr) control transfection. The table shows sequences corresponding to the silent mutations introduced in each mutant MRE (**B)** Quantitative PCR analysis of 293T cells co-transfected with Pre-miR-134, and 0.2μg of Sox2-CDS mutant constructs, together with the quantitation of the corresponding Western blots in (C); values for each construct are presented relative to their respective Scr control transfection (**C)** Western blot analysis of 293T cells co-transfected with Scr or Pre-miR-134, and 0.2μg of Sox2-CDS mutant constructs.

## 6.5 Discussion

Luciferase assays, Western blots and quantitative PCR analyses showed that miR-296 post-transcriptionally downregulates the expression of Nanog by binding to two non-overlapping sites that are located in the CDS region of Nanog. The targeting of Nanog's CDS by miR-296 was shown to be true for both endogenous (mESCs) and exogenous (293T cells) Nanog. Separate reporter assays using Nanog's full length 5′UTR and 3′UTR supported the absence of any interaction between miR-296 and Nanog's UTRs. Additionally, any of a number of silent mutations introduced in each of the two CDS binding sites was able to disrupt miR-296's activity. Finally, simultaneous silent mutations at *both* binding sites succeeded in abolishing miR-296's activity. The impact of silent mutations on miR-296's activity was supported by reporter assays, qPCR and Western blots, and suggests that miR-296 acts on Nanog predominantly through translational repression.

These findings lend experimental support to *rna22*'s statistically-significant computational evidence that animal microRNA binding sites must exist along the entire length of messenger RNAs, and show this to be true for the CDS region of a murine mRNA. This suggests the possibility of an expanded sphere of activity for animal microRNAs. It is likely that natural microRNA binding sites exist beyond the confines of a 3′UTR, which would imply a substantially more complex picture of microRNA-based regulation than previously estimated. Notably, in the miR-296/Nanog case that we discussed, the microRNA does not appear to act on the 5′UTR or 3′UTR region of the transcript, which suggests that CDS-only microRNA-targeting may also be a possibility.

The above analysis also showed that the two non-overlapping binding sites act independently of one another and are targeted by the same microRNA. Indeed, in order to abolish Nanog's down-regulation by miR-296, silent mutations had to be introduced simultaneously at both sites to disrupt miR-296's binding.

Even though a definite statement cannot be made based on the available data, it is nonetheless conceivable that the presence of a second binding site for the same microRNA may serve to provide a necessary degree of redundancy, especially considering the importance of Nanog for mESCs. Interestingly, *rna22* has predicted two or more binding sites in Nanog's mRNA for 158 microRNAs, thus suggesting that this redundancy hypothesis may apply to many of the microRNAs that target Nanog. Finally, it is worth noting that a single point mutation at Nanog-493p (Nanog-493p/m1) is sufficient to disrupt miR-296's activity in a very substantial manner. This provides yet another clear demonstration of the remarkable impact that single nucleotide polymorphisms can have on the binding of microRNAs (Abelson *et al.*, 2005; Clop *et al.*, 2006).

As discussed earlier, Nanog is known to be a major regulator of mESC pluripotency (Chambers *et al.*, 2003, Mitsui *et al.*, 2003). MiR-296's ability to modulate post-transcriptional repression of Nanog via its CDS region thus led to an investigation of the effect of miR-296 on mESC differentiation. The results demonstrated that miR-296 was upregulated during RA and EB differentiation, and was able to modulate mESC differentiation as determined by reduced alkaline phosphatase activity, reduced colony forming efficiency and increased expression of differentiation marker transcripts. As transcript and protein levels of Oct4 and Sox2 did not change

significantly with miR-296 overexpression, the data suggests that the effect of miR-296 on mESCs is mainly mediated by post-transcriptional regulation of Nanog. Furthermore, transfection of Nanog constructs which had silent mutations abolishing miR-296 binding was able to significantly rescue the miR-296-modulated differentiation of mESCs. This provides further evidence that miR-296's effect on ESCs is largely mediated via its repression of Nanog via the two sites in its CDS.

Finally, to demonstrate that the targeting of Nanog's CDS by miR-296 is not an isolated incidence of microRNAs targeting genes in their coding regions, the interaction between miR-134 and its five *rna-22* predicted targets in the CDS of mouse Sox2 was explored. The data suggested that miR-134 post-transcriptionally down-regulates the expression of Sox2 by binding to a single site located in the CDS region of Sox2. The targeting of Sox2's CDS by miR-134 was shown to be true for both endogenous (mESCs) and exogenous (293T cells) Sox2. Additionally, a number of silent mutations introduced in the CDS binding site were able to disrupt miR-134's activity. The impact of silent mutations on miR-134's activity was supported by reporter assays, qPCR and Western blots, and suggested that miR-134 acts on Sox2 predominantly through translation repression via interaction with a specific CDS element.

Taken together, these data demonstrating that miR-296 and miR-134 modulate post-transcriptional repression of Nanog and Sox2 respectively provide the first examples of animal microRNAs targeting genes in their coding regions and suggest that the sphere of microRNA influence on gene regulation in animals extends beyond the 3'UTR.

# CHAPTER 7. CONCLUSION

The importance of the complex transcriptional regulatory network including Nanog, Oct4 and Sox2 for the maintenance of ESC pluripotency and self-renewal has been well documented. Mounting evidence over the past few years, from this study and others, suggests that microRNAs may act as an additional layer of gene regulation in these cells. This is likely to be partially mediated via the reciprocal regulation of important ESC factors and microRNAs. Additional support for this hypothesis has emerged from a rather unexpected source.

In a recent publication, Yu *et al.* demonstrate that the four factors Oct4, Sox2, Nanog and Lin28 are sufficient for the reprogramming of human somatic cells to induced pluripotent stem (IPS) cells exhibiting the essential characteristics of ESCs (Yu *et al.*, 2007). These IPS cells are karyotypically normal, express hESC pluripotency markers and possess the ability to differentiate into derivatives of all three primary germ layers (Yu *et al.*, 2007). While Oct4, Sox2 and Nanog have been studied extensively in ESCs, comparatively little is known about Lin28. It is a highly conserved protein that is expressed in mouse embryos, ESCs, EC cells and subsequently downregulated during development *in vivo* and the neuronal differentiation of EC cells (Moss and Tang, 2003; Sempere *et al.*, 2004; Yang and Moss, 2003).

The mammalian orthologs of *C. elegans* Lin28, which is repressed by *lin-4* in worms via 3'UTR targeting, contain conserved predicted MREs in their 3'UTRs for miR-125b (a lin-4 ortholog), Let-7a, and miR-218 (Sempere *et al.*, 2004). The downregulation in Lin28 expression during RA-induced EC cell differentiation parallels the concomitant upregulation of miR-125b, let-7a, and miR-218, suggesting

that these events may be linked (Sempere *et al.*, 2004). The expression of mouse miR-125b is also observed to increase during RA-induced mESC differentiation (Figure 3.2A). Furthermore, human miR-125b and its homolog miR-125a have been shown to target Lin28 via two conserved sites in its 3'UTR (Wu and Belasco, 2005). Taken together, these data suggest a direct role for several microRNAs in the regulation of Lin28, a heterochronic gene recently implicated in reprogramming and pluripotency.

The results of this study present further evidence for microRNA-mediated modulation of ESC pluripotency via the targeting of important regulators of pluripotency and self-renewal. In summary, three significant findings are presented herein.

Firstly, the previously uncharacterized role of miR-134 in mESCs is described. By combining microRNA identification and expression, functional analyses and validation of microRNA-mRNA heteroduplexes, data is presented that suggests a role for miR-134 in modulating differentiation of mESCs, where this promiscuous microRNA exerts its effects through post-transcriptional regulation of multiple mRNAs, including translational attenuation of Nanog and LRH1 via their 3'UTRs. Upregulation of miR-134 also enhances the ectodermal differentiation of mESCs treated with RA or N2B27, where it augments the RA- and N2B27-induced elevation of ectodermal markers (ß-III-tubulin, Nestin, Neurogenin2, Sox1). These results will aid in the understanding of microRNA function in ESCs, and facilitate our ability to inhibit or enhance the differentiation of ectodermal lineages.

Secondly, experimental validation is provided for 168 of 226 *rna22*-predicted microRNA targets at the luciferase level. This is significant as aside from this study,

the number of validated predictions that have been reported thus far is small (Rajewsky, 2006). These experimental validations were also invaluable in enabling the refinement of *rna22*, a microRNA target prediction algorithm. Assuming that the luciferase results for a subset of randomly selected targets are representative of the entire dataset, the analysis suggests that some microRNAs may have more than a thousand targets. This hypothesis challenged the prevailing opinion in the field at the time that each microRNA only targeted a few genes.

Thirdly, 3′UTRs have been the focus of virtually all animal microRNA/mRNA studies to date, both from a computational and from an experimental standpoint. Although Kloosterman et al demonstrated functional targeting of *let-7* sites in the CDS or 5'UTR in zebrafish, this study did not identify any endogenous target sites as it utilized a fragment of the *lin-41* 3'-UTR containing two *let-7* target sites inserted into a *gfp* reporter gene (Kloosterman et al, 2004). Thus, the results of this study demonstrating that miR-296 and miR-134 modulate post-transcriptional repression of Nanog and Sox2 via 2 sites and 1 site in their respective coding regions provide the first examples of animal microRNAs regulating gene expression via endogenous CDS target sites.

This suggestion that the sphere of microRNA influence on gene regulation in animals may be far wider than currently thought is supported by several recent publications. Lytle et al reported that the translational efficiency of luciferase reporters containing let-7 microRNA complementary sites derived from the *C. elegans* lin-41 3' UTR was repressed in human HeLa cells, which express endogenous Let-7a (Lytle et al, 2007).

Although this study did not identify any endogenous 5'UTR microRNA binding sites, it raises the possibility that such targeting may be functional in mammalian cells.

Two landmark papers which further challenged existing beliefs about microRNA function were published by Vasudevan et al and Kedde et al. Vasudevan et al proposed that microRNAs could activate translation. They demonstrate that the ARE in tumor necrosis factor-alpha mRNA is transformed into a translation activation signal upon cell cycle arrest. Its subsequent recruitment of AGO and fragile X mental retardation-related protein 1 (FXR1), factors associated with micro-ribonucleoproteins (microRNPs), is directed by miR-369-3 (Vasudevan et al, 2007). In contrast, translation is repressed by miR369-3 during other cell cycle phases. Furthermore, two other microRNAs, Let-7 and the synthetic microRNA mimic cxcr4, similarly enhance translation of target mRNAs on cell cycle arrest, yet repress translation in proliferating cells (Vasudevan et al, 2007). These results suggest that multiple microRNAs and their associated Ago proteins may enhance or repress translation, in a cell cycle-dependent manner (Buchan and Parker, 2007).

Based on the intriguing observation that the sequences around the microRNA targeting site are highly conserved during evolution, Kedde et al hypothesized that these conserved regions may serve as docking platforms for modulators of microRNA activity. They demonstrated that the expression of dead end 1 (Dnd1), an evolutionarily conserved RBP, inhibited the function of several microRNAs by binding to mRNAs and prohibiting microRNA association with their target sites (Kedde et al, 2007). This effect is mediated through uridine-rich regions in the microRNA-targeted mRNAs, and may occur via physical blockage of microRNA

access to its target site or changing the subcellular localization of the target mRNA (Ketting, 2007). As the RNA-induced silencing complex (RISC) is only one of the numerous mRNA-targeting RBPs, it is perhaps not surprising that the regulation of gene expression by microRNAs may be tuned by RBPs that are not a direct part of the microRNA complex (Ketting, 2007).

In conclusion, the results of this study, combined with other recently published reports, suggest that researchers have barely scratched the surface of understanding microRNA biology and function. MicroRNAs, which were previously thought to associate solely with the 3'UTRs of several target genes to mediate gene repression, have now been shown to have hundreds or possibly thousands of targets. They are also able to target genes in their coding regions, 5'UTRs and perhaps even introns, as well as induce translational upregulation. These tiny regulators of gene expression also appear to be, in turn, regulated by other factors such as RBPs. Shedding greater light on these newly discovered facets of microRNA biology will be just one of the challenges facing researchers in the near future.

The last 5 years has seen an explosion of information on the function and regulation of microRNAs, as well as their mechanism of targeting mRNA *cis*-elements. Building upon the current knowledge base, future research will continue to unravel the mysteries of microRNA biology, including the vital role that microRNA-mediated gene regulation plays in directing cell fate. As this study and others have demonstrated that microRNAs are important modulators of differentiation in ESCs, these new insights will likely facilitate the understanding of ESCs and their use as models of development and disease, and for potential clinical applications.

# CHAPTER 8. MATERIALS AND METHODS

## 8.1 Cell culture, transfection & cell-based assays

### 8.1.1 Routine cell line maintenance

All cell culture reagents and culture plastics were obtained from Invitrogen/Gibco and BD Biosciences, respectively, unless otherwise specified. Cell cultures were maintained at 37°C with 5% $CO_2$. mESC lines, E14 (ATCC: CRL-1821) and D3 (ATCC: CRL-1934), were cultured feeder-free on 0.1% gelatin-coated (Sigma) plates in ESC-medium (Dulbecco's modified Eagle's medium, 15% heat-inactivated ES-standard fetal bovine serum (FBS), 100 μM non-essential amino acids, 2 mM L-glutamine, 55 nM β-mercaptoethanol, 1% (v/v) penicillin/streptomycin and mouse leukemia inhibitory factor (mLIF; $10^3$ U/ml, Chemicon). HEK 293T/17 (ATCC: CRL-11268) cells were maintained in DMEM supplemented with 10% heat-inactivated FBS, 2mM L-glutamine and 1% (v/v) penicillin/streptomycin.

### 8.1.2 Differentiation of mESCs

For RA-induced mESC differentiation, feeder-free mESCs were cultured on 0.1% gelatin-coated plates in LIF-deficient ESC-medium supplemented with 100nM all-*trans* retinoic acid for up to 6 days. For N2B27-mediated differentiation, feeder-free mESCs were cultured on 0.1% gelatin-coated plates in a 1:1 mixture of DMEM/F12 and Neurobasal medium, together with N2 and B27 supplements as described previously (Ying *et al.*, 2003). To form EBs, feeder-free mESCs were trypsinized to a single cell suspension and subsequently cultured on uncoated petri dishes in ESC-medium without mLIF for up to 15 days. Petri dishes were placed on a reciprocal shaker to minimize aggregation of EBs. Media were changed every two days for all mESC differentiation conditions.

### 8.1.3 Preparation of mouse tissues

(*Note: Mouse tissues were isolated by Phil Gaughwin*)

Whole mouse embryos were removed from pregnant C57BL/6J mice euthanized by exposure to $CO_2$ at embryonic days (E) 7.5, 11.5, 15.5 and 17.5 (observation of vaginal plug recorded as E0.5). To examine tissue-specific expression patterns of microRNAs, euthanized adult (at least postnatal day (P) 60) and embryonic (E12.5) tissue was obtained by dissection under RNAse-free conditions. Whole embryos and embryonic / adult organ derived tissue samples were homogenized in Trizol within one hour of tissue collection. E12.5 whole brain tissue was dissected in ice-cold Dulbecco's Phospate-Buffered Saline (DPBS) to remove meninges and separate the hippocampal anlage, cortical sheets, remaining neural tissue rostral to the midbrain-hindbrain isthmus / boundary (MHB), and tissue caudal to the MHB including cervical spinal cord (defined as hippocampus, cortex, midbrain and hindbrain, respectively in Fig. 7B).

### 8.1.4 Transfection

Pre-miR microRNA precursors (Pre-miRs), Anti-miR microRNA inhibitors (Anti-miRs) and Scrambled oligomer (Scr, AGACUAGCGGUAUCUUUAUCCC) were purchased from Ambion. Block-iT fluorescent RNA oligomer (Invitrogen), a FITC-tagged 25-mer, was used as a surrogate for Pre-miRs and Anti-miRs to optimize transfection efficiency. These were transfected into mESCs at a final concentration of ~100 nM using Lipofectamine 2000 (Invitrogen) as per manufacturer's instructions. For example, for a 12-well transfection, mESCs were seeded at a density of 120,000 cells per well. After 24 h, 7 μl of Pre-miRs was added to 100 μl of Opti-Mem and 3.5

μl of Lipofectamine 2000 was added to 100 μl of Opti-Mem and incubated separately for 10 min. These two mixtures were then combined and incubated for a further 20 min. This final transfection mix was subsequently added to cells in 1 ml of medium. 1.5 μg of gene-specific shRNA plasmids, non-silencing shRNA plasmid or empty vector were transfected into mESCs in 12-well plates and drug selection was performed with 1 μg/ml of puromycin (Sigma) 24 h post-transfection for a period of 2 days.

### 8.1.5 Immunostaining & Cellomics High Throughput Screening

Cells were fixed in 4% paraformaldehyde, washed twice with DPBS and then incubated for 5 min at -20$^{o}$C in 95% ethanol (v/v in DPBS). Cells were then washed thrice with DPBS, blocked for 1 h in 5% normal goat serum in DPBS with 0.1x Triton-X100, and incubated overnight with anti-Class III beta-tubulin (Ab18660; Abcam, Cambridge, UK) primary antibody at 4ºC. Cells were next washed twice with DPBS and incubated for 1 h with the corresponding secondary antibody (Goat anti-rabbit IgG 568; Molecular Probes). Cells were washed and mounted. Immunofluorescence was observed using a Leica microscope, and image analysis was conducted using IM50 software. The Alkaline Phosphatase Detection Kit (Chemicon) was used to determine Alkaline Phosphatase activity, according to manufacturer's instruction.

Image acquisition and subsequent neurite outgrowth measurements were performed using the Cellomics ArrayScan II High Content Screening (HCS) platform (Swallowfield, UK). Two separate fluorescent filters at 20× objective magnification were used; Channel 1 illuminated Hoechst 33342-labelled nuclei, allowing automated

focusing upon the cells, and Channel 2 excited objects stained with Alexa Fluor 568-labelled anti-Class III beta-tubulin. Both images were acquired through a 535 nm × 35 nm bandwidth dicroic emission filter. Exposure times for each wavelength were determined empirically. All images were then analyzed using Cellomics neurite outgrowth software, and data analysis for the number of neurons per total nucleus count in each field, average intensity of neuronal cell bodies and neurites are presented. Data shown were an average of values from 16 random fields from 3 replicate wells in two independent experiments.

### 8.1.6 *In situ* hybridization

(*Note:ISH was performed by Wang Weijia and Liu Rubing*)

For sectioned RNA *in situ* hybridization, sagittal and transverse sections of outbred CD-1 (Charles River) wildtype embryos at stages E3.5 to E14.5 were prepared essentially as previously described (Wang *et al.*, 2001). Noon of the vaginal plug date was designated as E0.5. Embryo tissues were dissected free, fixed in 4% paraformaldehyde overnight, washed in DPBS and dehydrated through graded ethanols, followed by two changes of Americlear (Fisher) and embedded in Paraplast (Fisher) overnight. Sections of 7 μm were cut and floated onto Plus+ slides (Fisher), air dried, and stored at 4°C.

Whole-mount *in situ* hybridization was performed essentially as previously described (Wang *et al.*, 2001; Kloosterman *et al.*, 2006). Embryos were fixed in 4% paraformaldehyde overnight, washed in DPBS, and dehydrated through graded methanols and stored at –20˚C for up to several months. Before use, embryos were rehydrated stepwise from methanol to DPBS+0.1% Tween-20 (PTW) and briefly

digested with 10 μg/ml proteinase K. Digoxigenin-labeled RNA probes were purchased from Exiqon (www.exiqon.com) and used for both sectioned and wholemount RNA in situ and hybridized at 65°C overnight. Following washing in PTW, embryos or slides were incubated in a 1:2000 dilution of AP-anti-DIG antibody (Boehringer-Mannheim) followed by extensive washing and then staining in Purple Precipitating Reagent or NBT/BCIP (Boehringer-Mannheim) at 4°C-30˚C. Stained wholemount embryos were photographed in PTW and then stored in 50% ethanol at 4˚C. Sections were treated with aqueous mounting solution, coverslipped and stored at 4˚C prior to photography.

### 8.1.7 Colony formation assay

E14 mESCs were transfected with Pre-miRs (100 nM) and/or Anti-miRs (100 nM), or scrambled RNA oligomer (100 nM) or lipofectamine only (Mock Transfection) in 12-well culture plates at a density of 5 x $10^5$ cells per well. These cells were subsequently trypsinized at 48 h and re-suspended in ESC-medium. Various cell numbers (200, 400) were plated onto mouse feeder layers in 6-well culture plates to form secondary ESC-colonies. After seven days, emerging colonies were stained using Wright-Giemsa (Sigma) stain and counted. Colony morphology and number provide an indication of the number of colony-forming undifferentiated mESCs present in a population of cells (Qi *et al.*, 2004).

### 8.1.8 pOct4/pNanog-Luciferase reporter assays

(*Note: pOct4/pNanog-Luc reporter assays were performed by YT and TWL*)

To perform the pOct4/pNanog-Luciferase assays, feeder free mESC lines, E14 and D3, were seeded 24 h before transfection at a density of 2.5 x $10^4$ cells per well in 96-well culture plates (Costar). Pre-miRs (100 nM), Anti-miRs (100 nM), scrambled RNA oligomer (100 nM) or shRNA constructs (1 μg), plus pOct4-Luciferase (75 ng) or pNanog-Luciferase (75 ng), and pRL-SV40 (1 ng) (Promega) were co-transfected using Lipofectamine 2000 according to manufacturer's instructions. The pRL-SV40 plasmid served as an internal control for the normalization of transfection efficiency. For the transfection of shRNA constructs, selection with 1 μg/ml of puromycin (Sigma) was performed 24 h post-transfection for a period of two days. Firefly and *Renilla* luciferase activities were measured with the Dual-Luciferase® Reporter system (Promega) 48 h post-transfection by Centro LB960 Luminometer (Berthold Technologies Gmbh & Co.). Data generated were expressed relative to control transfection (scrambled oligo for Pre-miRs, non-silencing shRNA for shRNA constructs) after internal normalization to *Renilla* luciferase readings.

### 8.1.9 microRNA target validation assay

(*Note:Target validation assays were performed by YT and AYS)*

HEK 293T cells were seeded 24 h before transfection at a density of 5 x $10^4$ cells/well in black-walled 96-well culture plates. 120 ng of miR-375 or miR-296 overexpression vectors or empty vector; or 12.5 nM of PmiR-134 or scrambled RNA oligomer; were cotransfected with 2 ng of the *R*Luc-MRE vector with Lipofectamine 2000 as per manufacturer's instructions. Concurrently, additional controls were also performed using unpredicted *R*Luc-MREs (eg. reverse complement of miR-21) vs cognate microRNA or predicted *R*Luc-MRE vs non-cognate microRNAs (eg. miR-21). Non-cognate microRNA transfections against each MRE also served to determine the false

positive rate of Pre-miR transfections against that particular MRE. Forty eight hours post-transfection, Firefly and *Renilla* luciferase activities in cell lysates were measured consecutively with the Dual-Luciferase® Reporter system (Promega) by a luminometer (Centro LB960; Berthold Technologies Gmbh & Co. KG, Germany, http://www.bertholdtech.com). Constitutively expressed Firefly luciferase gene activity in psiCHECK-2 served as a normalisation control for transfection efficiency. After correction for transfection efficiency, the normalised *Renilla* luciferase readouts were expressed as relative percentage change to that of the scrambled RNA oligomer transfection. All luciferase assays were repeated a minimum of three times with 4 culture replicates each.

## 8.2 DNA manipulation

### 8.2.1 General DNA manipulation techniques

PCR was performed using the BD Advantage 2 PCR kit (BD Biosciences). The protocol used to amplify DNA fragments from genomic DNA was as follows:

| | |
|---|---|
| 10x buffer with Mg | 5.0 µl |
| dNTP (10mM) | 1.0 µl |
| Genomic DNA (100ng/µl) | 1.5 µl |
| H2O | 39.5 µl |
| Polymerase | 1.0 µl |
| Forward primer (10 µM) | 1.0 µl |
| Reverse primer (10 µM) | 1.0 µl |
| Total | 50.0 µl |

Annealing was done by mixing 1.5 µl of forward oligo (100 µM), 1.5 µl of reverse oligo (100 µM) and 47 µl of annealing buffer. These were vortexed well and annealed under the following conditions: 95$^{o}$C for 5 min, 70$^{o}$C for 5 min, ramp to 4$^{o}$C, 4$^{o}$C for 10 min.

Restriction digest was carried out for 5-16 h at 37°C, at a ratio of 5 enzyme unit to 1 μg of DNA. The protocol used to digest psiCHECK-2 vector was as follows:

| | |
|---|---|
| Water | 28.6 μl |
| Vector (1 μg/μl) | 4.0 μl |
| NEB buffer 3 | 4.0 μl |
| BSA | 0.4 μl |
| Xho1 (20,000U/ml) | 1.0 μl |
| Not1 (10,000U/ml) | 2.0 μl |
| Total | 40.0 μl |

PCR purification and gel extraction were performed as per manufacturer's instructions, using the QIAquick PCR Purification kit and QIAquick Gel Extraction kit respectively (Qiagen).

Ligation was carried out at room temperature for 5 min using the NEB Quick Ligation kit (NEB). The typical vector to insert ratio used was 1 molecule of vector to 5 molecules of insert. The protocol used was as follows:

| | |
|---|---|
| Vector (200ng/μl) | 0.5 μl |
| Insert (25ng/μl) | 2.0 μl |
| Water | 6.5 μl |
| 2x Ligase buffer | 10.0 μl |
| T4 DNA Ligase | 1.0 μl |

Transformation was performed using MAX Efficiency DH5α (Invitrogen) as follows: 1 μl of ligation mix was added to 25 μl of bacteria, tapped gently and incubated on ice for 30 min. The mixture was heat shocked at 42°C for 1 min, then allowed to recover on ice for 5 min. 400 μl of SOC medium was added and the transformation mix was incubated at 37°C on a shaker for 1 h. 150 μl of this transformation mix was spread on LB plates with the appropriate antibiotic resistance.

Clones carrying the DNA of interest were identified by gene specific PCR. Bacterial colonies were first picked into a 96-well plate with 50 μl of LB media containing the appropriate antibiotic in each well. This was incubated at 37 $^{o}$C for 1 h and PCR was then performed with the following conditions:

| | |
|---|---|
| PCR Supermix | 17.5 μl |
| Forward primer (10 μM) | 1.0 μl |
| Reverse primer (10 μM) | 1.0 μl |
| Bacteria | 0.5 μl |

Minipreps and maxipreps were performed as per manufacturer's instructions, using the QIAprep Spin Miniprep kit and HiSpeed Plasmid Maxi kit respectively (Qiagen).

### 8.2.2 pOct4/pNanog-Luciferase reporters

pOct4-Luciferase and pNanog-Luciferase reporters were a kind donation from the lab of Ng Huck Hui (Chew *et al.*, 2005).

### 8.2.3 Construction of microRNA overexpression plasmids

To generate the overexpression vector for mmu-miR-375, a 500 bp fragment was amplified by PCR from mouse genomic DNA and inserted into a modified pIRES2-EGFP vector (*EcoRI* and *BamHI* sites; Clontech).

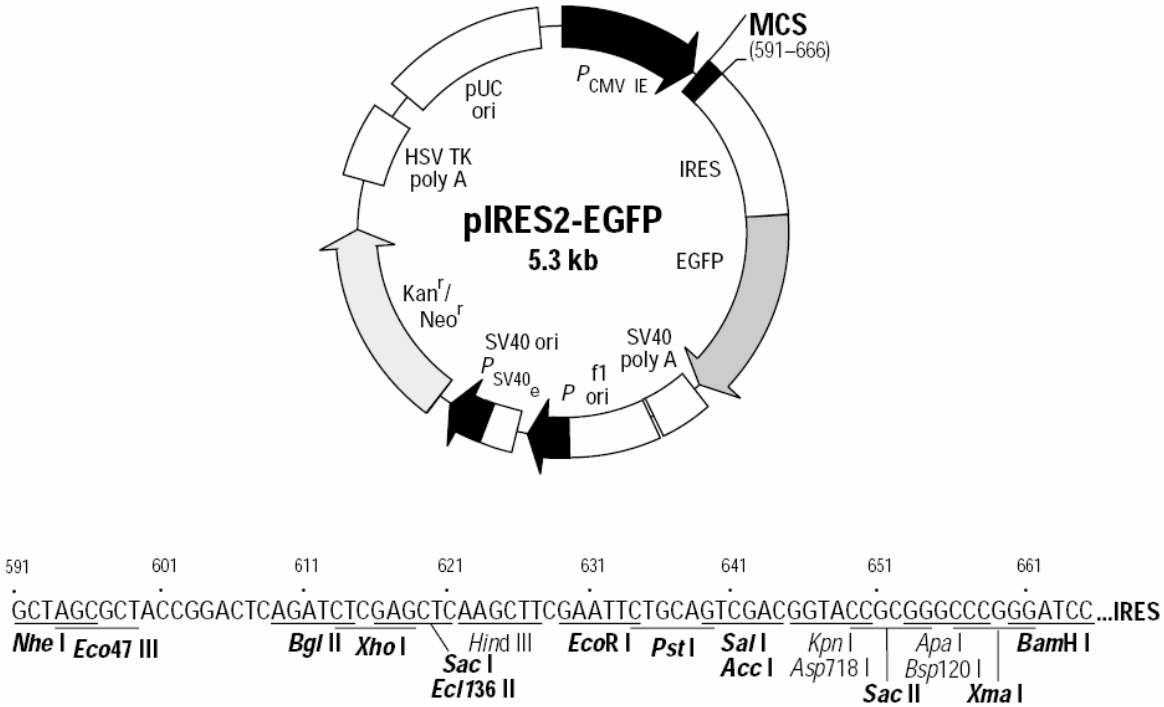



**Figure 8.1. Vector map of pIRES2-EGFP.** From www.clontech.com.

To generate the mmu-miR-296 and mmu-miR-21 overexpression vectors, 500 bp fragments were amplified by PCR from mouse genomic DNA and inserted into the pLL3.7 lentiviral vector (*XhoI* and *HpaI* sites; a kind gift from the Center for Cancer Research, MIT).

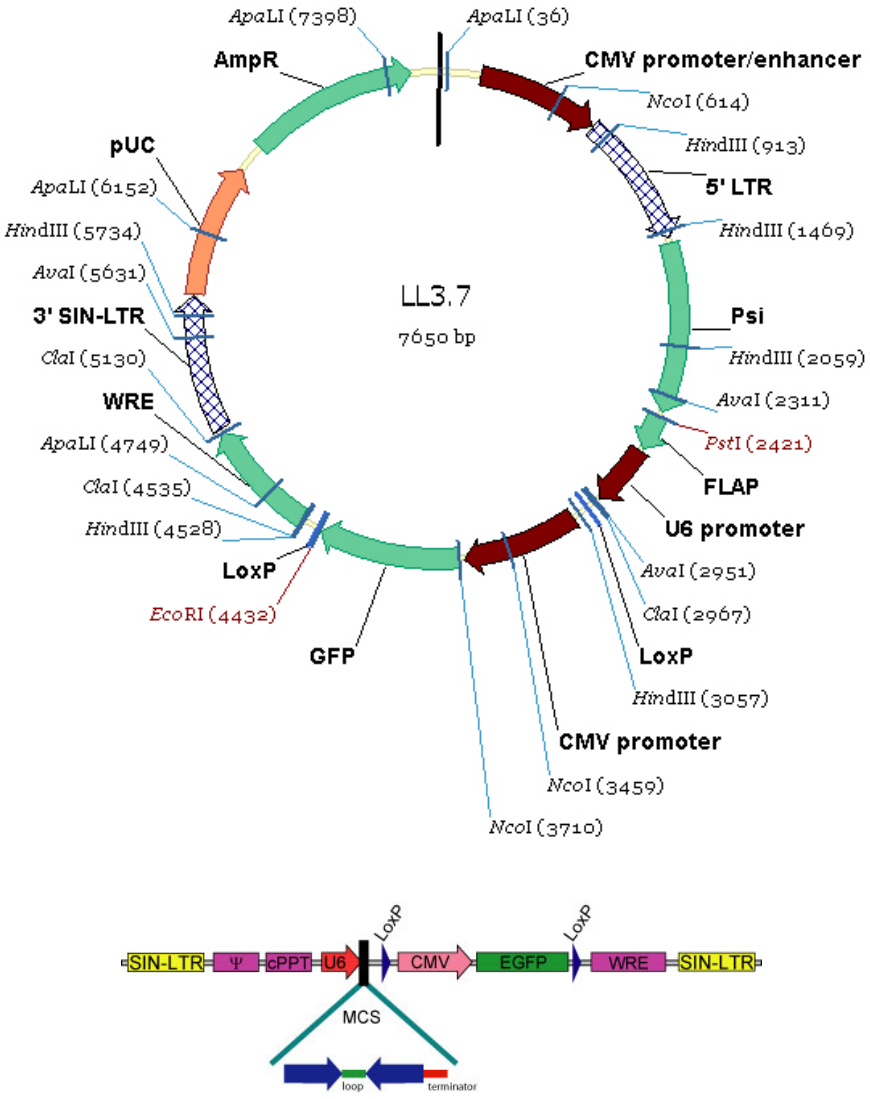



**Figure 8.2. Vector map of pLL3.7.** From www.sciencegateway.org.

### 8.2.4 Construction of pLuc-MRE plasmids

(*Note:pLuc-MRE plasmids were constructed by YT and AYS*)

For pLuc-MRE constructs, the *rna22*-predicted microRNA response elements (MRE, 20-30 nt) were synthesized as sense and antisense oligomers, annealed and cloned into psiCHECK-2 (Promega) at restriction sites *Xho*I & *Not*I, directly 3' downstream of *Renilla* luciferase.

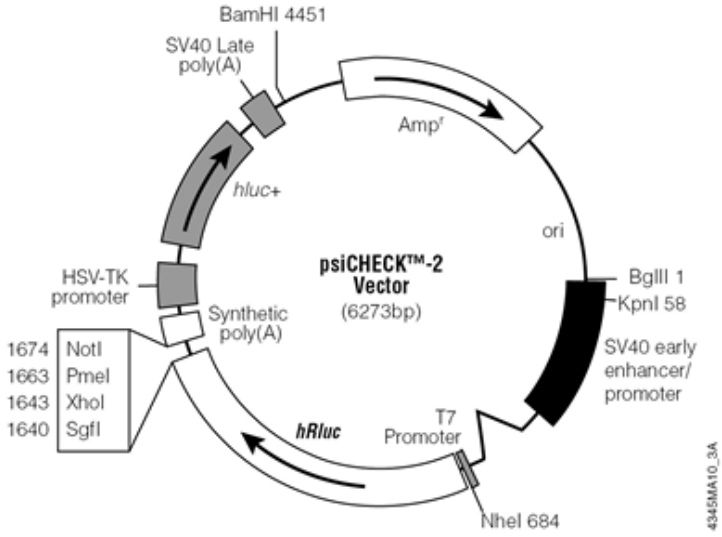



**Figure 8.3. Vector map of psiCHECK-2.** From www.promega.com.

To generate the pLuc-Nanog5′UTR construct, the 5′UTR of Nanog was synthesized as sense and antisense oligomers, annealed and cloned into psiCHECK-2 at the *NheI* restriction site, directly 5′ upstream of *Renilla* luciferase. The pLuc-Nanog3'UTR construct was generated by cloning the entire Nanog 3'UTR into psiCHECK-2 (Promega) at the *Not*I restriction site. To generate the Nanog 3'UTR mutant construct, PCR directed mutagenesis was performed with a pair of primers containing the mutant MRE sequence (GAGGACGTGTTAACTAGTTTCC). The PCR product was subsequently used as a template for full-length PCR. The Nanog 3'UTR mutant was also cloned into psiCHECK-2 at the *Not*I restriction site.

### 8.2.5 Construction of gene-specific RNAi plasmids

(*Note:RNAi plasmids were constructed by TWL and AYS*)

Nineteen-base-pair gene-specific regions for RNA interference (RNAi) were designed (Reynolds *et al.l*, 2004, Ui-Tei *et al.*, 2004). Oligonucleotides were cloned into pSUPER.puro (*Bgl*II & *Hin*dIII sites; Oligoengine), which expresses 19-nucleotide hairpin-type short hairpin RNAs (shRNAs) with a 9-nucleotide loop, as described previously (Geijsen *et al.*, 2004).

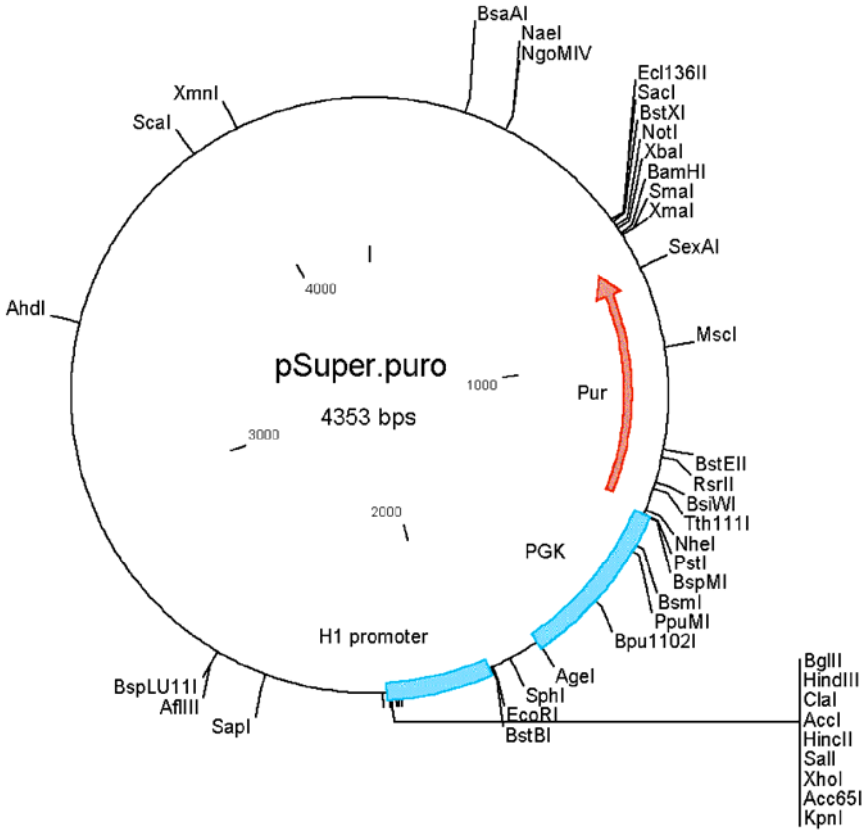



**Figure 8.4. Vector map of pSuper.puro**. From www.oligoengine.com.

To ensure specificity, all sequences were analyzed by BLAST to ensure that they did not have significant (>15 nucleotide matches) sequence similarity with any other gene. For each target gene, RNAi against four regions of the mRNA were designed

and screened, and the construct with the highest mRNA suppression was selected for subsequent pOct4-Luciferase studies. The oligonucleotides used were:

| | |
|---|---|
| ***LRH1*** | 5'-GATCCCCGTGTGTGGCGATAAAGTGTttcaagagaACACTTTATCGCCACACACTTTTTA-3'<br>5'-AGCTTAAAAAGGACCTACACTTATGAGAAtctcttgaaTTCTCATAAGTGTAGGTCCGGG-3' |
| ***FADD*** | 5'-GATCCCCGAATATGTCCCCAGTACTAttcaagagaTAGTACTGGGGACATATTCTTTTTA-3'<br>5'-AGCTTAAAAAGAATATGTCCCCAGTACTAtctcttgaaTAGTACTGGGGACATATTCGGG-3' |
| ***$G\alpha_o$*** | 5'-GATCCCCGGAAGACGGACTCCAAGATttcaagagaATCTTGGAGTCCGTCTTCCTTTTTA-3'<br>5'-AGCTTAAAAAGGAAGACGGACTCCAAGATtctcttgaaATCTTGGAGTCCGTCTTCCGGG-3' |
| ***Oct4*** | 5'-GATCCCCGAAGGATGTGGTTCGAGTATTCAAGAGATACTCGAACCACATCCTTCTTTTTA-3'<br>5'-AGCTTAAAAAGAAGGATGTGGTTCGAGTATCTCTTGAATACTCGAACCACATCCTTCGG-3' |

### 8.2.6 Construction of Nanog-CDS & Sox2-CDS plasmids

(*Note:CDS plasmids were constructed by Zhang Jinqiu*)

The wild-type Nanog-CDS and Sox2-CDS were cloned into pcDNA3.1 (Invitrogen) at restriction sites *BamHI* and *XhoI* using the following primers respectively:

**NanogBamHIF** 5′-GCGGGATCCACCATGAGTGTGGGTCTTCCTGG-3′
**NanogXhoIR** 5′-CGGCTCGAGTCATATTTCACCTGGTGGAG-3′

**Sox2BamHIF** 5′- GCGGGATCCCACCATGTATAACATGATGGAGACGGA-3′
**Sox2XhoIR** 5′- CGGCTCGAGTCACATGTGCGACAGGGGCAGT-3′

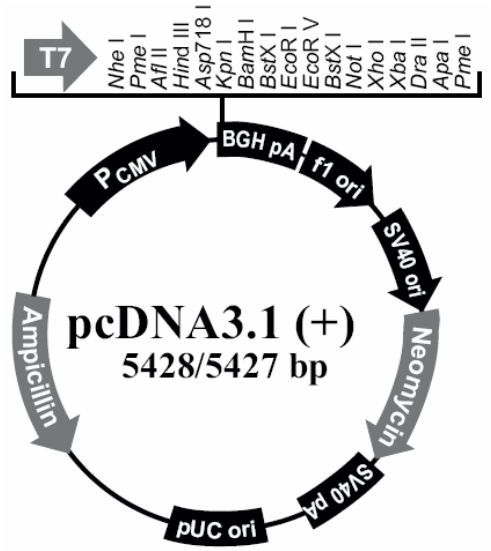



**Figure 8.5. Vector map of pcDNA3.1.** Adapted from www.invitrogen.com.

To generate each CDS mutant construct, PCR directed mutagenesis was performed with a pair of primers containing the mutant MRE sequence. The PCR products were subsequently used as a template for full-length PCR. The CDS mutants were also cloned into pcDNA3.1 (Invitrogen) at restriction sites *BamHI* and *XhoI*. The primer sequences used for mutagenesis were as follows:

**235-1mutmNanogF1:**
5'-AGCTCTCAAGTCCCGAAGCTGACAAAGGACCGGAGGAGGAGGA-3'
**235-1mutmNanogR1:**
5'-TCCTCCTCCTCCGGTCCTTTGTCAGCTTCGGGACTTGAGAGCT-3'
**493-1mutmNanogF1:**
5'-ACTAGCAATGGTCTGATTCAGAAAGGAAGCGCACCAGTGGA-3'
**493-1mutmNanogR1:**
5'-TCCACTGGTGCGCTTCCTTTCTGAATCAGACCATTGCTAGT-3'
**493-2mutmNanogF2:**
5'-ACTAGCAATGGTCTGATCCAAAAGGGAAGCGCACCAGTGGA-3'
**493-2mutmNanogR2:**
5'-TCCACTGGTGCGCTTCCCTTTTGGATCAGACCATTGCTAGT-3'

**637-5mutSox2F:**
5'-GTCAGCGCCCTGCAGTACAACAGCATGACGTCCAGTCAGACCTACATGAACGGCT-3'
**637-5mutSox2R:**
5'-AGCCGTTCATGTAGGTCTGACTGGACGTCATGCTGTTGTACTGCAGGGCGCTGAC-3'
**637-6mutSox2F:**
5'-GTCAGCGCCCTGCAGTACAATAGCATGACGTCCAGTCAGACCTACATGAACGGCT-3'
**637-6mutSox2R:**
5'- AGCCGTTCATGTAGGTCTGACTGGACGTCATGCTATTGTACTGCAGGGCGCTGAC-3'

## 8.3 RNA and protein work

### 8.3.1 RNA extraction & quantitative PCR

For Northern blot, quantitative PCR and microarray analyses, total RNA was extracted from cells using Trizol reagent (Invitrogen) as per manufacturer's instructions and subsequently column-purified with RNeasy kits (Qiagen). Briefly, cell pellets were resuspended in Trizol before addition of chloroform (200 µl of chloroform per 1ml of Trizol). Samples were shaken vigorously for 15 s, left to settle for 2 min, and centrifuged for 20 min at 13,000 x g. After centrifugation, the aqueous layer of each sample was removed and added to an equal volume of 70% ethanol. This mixture was then applied to an RNeasy column, and washed with RW1 and RPE buffers (Qiagen). The column was centrifuged for an additional 2 min at 13,000 x g to remove any remaining ethanol carryover and RNA was eluted in 30-100 µl of RNase-free water.

For microRNAs, cDNA synthesis was performed with 50 ng of total RNA according to the manufacturer's instructions (Applied Biosystems). Endogenous microRNA levels were measured with inventoried Taqman probes and PCR Master Mix (Applied Biosystems). miR-16 was used as an internal control. All amplicons were analyzed using ABI Prism 7900HT Sequence Detection System (SDS) 2.2 software.

For coding genes, cDNA synthesis was performed with 1 µg of total RNA at 37$^{o}$C for 2 h using the High Capacity cDNA Archive Kit (Applied Biosystems), and subsequently diluted 10 times. Endogenous mRNA levels of pluripotency and differentiation markers were measured with inventoried Taqman probes and PCR Master Mix (Applied Biosystems). Endogenous mRNA levels of LRH1, FADD, and

G$\alpha_o$ were measured using SYBR Green PCR Master Mix (Applied Biosystems). β-actin was used as an internal control. All amplicons were analyzed using ABI Prism 7900HT Sequence Detection System (SDS) 2.2 software.

Primer sequences for quantitative PCR detection of miR-134 target genes:

| | **Forward** | **Reverse** |
|---|---|---|
| **LRH1** | 5' - ctcccccaaacgtgaggaaca - 3' | 5' - tttcgcacgtgaggagaccg - 3' |
| **FADD** | 5' - ggcagatctgcaggtggcatt - 3' | 5' - tccgcagcgccttaaccagt - 3' |
| **Gα$_o$** | 5' - tgctgggggctggagaatca - 3' | 5' - cccagagtcgcatcatggca - 3' |
| **β-actin** | 5' - accaactgggacgacatggagaag – 3' | 5' - tacgaccagaggcatacagggaca - 3' |

### 8.3.2 Northern blot

(*Note: Northern blots were performed by AYS and YT*)

All microRNAs were probed using $\alpha^{32}$P-UTP-labeled RNA probes which were transcribed *in vitro* from DNA oligomer templates (Thomson *et al.*, 1999). 5S rRNA was probed using $\gamma$ $^{32}$P-end-labeled DNA probe. Both types of probes were synthesized using the *mir*Vana$^{TM}$ microRNA Probe Construction kit (Ambion) and subsequently column-purified with *mir*Vana$^{TM}$ Probe & Marker kit (Ambion) as per manufacturer's recommendations. The specific activity of each probe was routinely $10^7$ cpm/μg RNA/DNA. 10 μg of total RNA were size-fractionated by 15% TBE-Urea-PAGE (Invitrogen) and electroblotted onto MagnaProbe nylon membrane (GE Osmonics Labstore). The membrane was probed overnight at 42ºC in 6 x SSC, 0.2% SDS, 5 x Denhardt's solution with $^{32}$P-labelled RNA probes ($10^6$ cpm/ml), washed thrice at 30ºC in 6 x SSC, 0.2% SDS, and complexes detected by Phosphorimage (GE-Healthcare).

### 8.3.3 Microarray

MicroRNA microarray was performed using a service provider (LC Sciences). The assay started from 2 to 5 μg total RNA sample, which was size fractionated using a YM-100 Microcon centrifugal filter (from Millipore) and the small RNAs (< 300 nt) isolated were 3'-extended with a poly(A) tail using poly(A) polymerase. An oligonucleotide tag was then ligated to the poly(A) tail for later fluorescent dye staining; two different tags were used for the two RNA samples in dual-sample experiments. Hybridization was performed overnight on a μParaflo microfluidic chip using a micro-circulation pump (Atactic Technologies). On the microfluidic chip, each detection probe consisted of a chemically modified nucleotide coding segment complementary to target microRNA (from miRBase, http://microrna.sanger.ac.uk/sequences/) or other RNA (control or customer defined sequences) and a spacer segment of polyethylene glycol to extend the coding segment away from the substrate. The detection probes were made by synthesis *in situ* using photogenerated reagent (PGR) chemistry. The hybridization melting temperatures were balanced by chemical modifications of the detection probes. Hybridization used 100 μL 6xSSPE buffer (0.90 M NaCl, 60 mM $Na_2HPO_4$, 6 mM EDTA, pH 6.8) containing 25% formamide at 34 °C. After hybridization, detection used fluorescence labeling using tag-specific Cy3 and Cy5 dyes. Hybridization images were collected using a laser scanner (GenePix 4000B, Molecular Device) and digitized using Array-Pro image analysis software (Media Cybernetics). Data were analyzed by first subtracting the background and then normalizing the signals using a LOWESS filter (Locally-weighted Regression). For two color experiments, the ratio of the two sets of detected signals ($log_2$ transformed, balanced) and p-values of the t-test were calculated; differentially detected signals were those with less than 0.01 p-values.

For the mRNA gene expression microarray, two sample repeats were profiled for all treatments. Total RNA was Trizol extracted; then first and second strand cDNA synthesis, cRNA Biotin-labelling and cRNA purification were performed subsequently using the Illumina TotalPrepRNA Amplification kit (Ambion;#IL1791). 850 ng labeled cRNA were mixed with Hyb Mix (Hyb E1+Formamide) as according to instructions in BeadStation 500X Gene Expression System (Illumina) and used for microarray (Illumina Mouse Refseq-8 Array) staining. Hybridization and wash procedures were performed using Illumina BeadChip protocols where 34 μl of the hybridization mix with labeled cRNA sample was dispensed onto the center of each array and allowed to hybridize with rotations in a pre-heated oven for 16-20 h at 55ºC. Washes were performed with Wash E1BC solution, Block E1 solution was used for blocking, and detection performed by Block E1 solution+Streptavidin-Cy3. Samples were scanned using standard Illumina procedures. Total chip data is deposited for public assess with GEO repository accession number: GSE4522. The microarray data were extracted using the software provided by the manufacturer (BeadStudio).

### 8.3.4 Protein extraction & Western blot

Cell pellets were washed in chilled DPBS, and incubated for 20 min in ice cold lysis buffer (50 mM Tris (pH7.0), 150 mM sodium chloride, 10 mM sodium pyrophosphate, 50 mM sodium fluoride, 50 mM sodium β-glycerophosphate, 1 mM sodium orthovanadate, 2 mM benzamidine, 1 mM dithiothreitol, 0.5% sodium deoxycholate, 0.2% Nonidet P-40) containing freshly added protease inhibitors (0.5 mM phenylmethylsulfonyl fluoride, 10 μg/ml leupeptin, 2 μg/ml aprotonin). Lysates were cleared by centrifugation at 4ºC for 20 min at 12,100 x *g*, and protein

concentrations were determined using Bradford Dye (Bio-Rad). Bovine serum albumin (BSA) was used as protein standards.

For Western blot analysis, 10 µg of total protein was size fractionated by SDS-PAGE on 10% *bis*-Tris acrylamide NuPAge gels in 1 x MOPS-SDS running buffer, pH 7.3 (Invitrogen), at 200 V for 60 min at 4ºC, and transferred to Hybond-P® PVDF membrane (GE Healthcare) in 1 x NuPage transfer buffer (Invitrogen) with 10% methanol at 40 V for 60 min at 4ºC. The membrane was probed with specific primary antibodies (Oct4, sc8628; Sox2, sc17320; BMP4, sc6896; Gata4, sc1237; Nestin, sc21248; $G\alpha_o$, sc387; FADD, sc6036; Santa Cruz Biotechnology; LRH-1, F4100-44; USBiological; Nanog, AB5731; β-actin, MAB1501; Chemicon) and secondary horseradish peroxidase conjugated antibodies (anti-goat HRP, sc2768; anti-rabbit HRP, sc2030; anti-mouse, sc2005; Santa Cruz). Antibody-protein complexes were identified by ECL-Plus (Amersham) and film.

### 8.4 Statistical analysis

Unless otherwise stated, the unpaired Student's *t*-test was used to determine statistical significance.

(*Note: All analysis of Illumina microarray data was performed by Henry Yang from the Bioinformatics Institute*)

For microarray data, the Pearson's correlation was used for the similarity test. Dendrograms were drawn using Eisenplots (clustering and treeview tools), with arm lengths proportional to the respective correlation coefficients.

## CHAPTER 9. BIBLIOGRAPHY

Abbott A.L., Alvarez-Saavedra E., Miska E.A., Lau N.C., Bartel D.P., Horvitz H.R., Ambros V. (2005) The let-7 MicroRNA family members mir-48, mir-84, and mir-241 function together to regulate developmental timing in *Caenorhabditis elegans. Dev Cell* **9**, 403-414.

Abelson J.F., Kwan K.Y., O'Roak B.J., Baek D.Y., Stillman A.A., Morgan T.M., Mathews C.A., Pauls D.L., Rasin M.R., Gunel M*., et al.* (2005) Sequence variants in SLITRK1 are associated with Tourette's syndrome. *Science* **310**, 317-320.

Aboobaker A.A., Tomancak P., Patel N., Rubin G.M., Lai E.C. (2005) *Drosophila* microRNAs exhibit diverse spatial expression patterns during embryonic development. *Proc Natl Acad Sci USA* **102**, 18017-18022.

Aiba K., Sharov A.A., Carter M.G., Foroni C., Vescovi A.L., Ko M.S.H. (2006) Defining a developmental path to neural fate by global expression profiling of mouse embryonic stem cells and adult neural stem/progenitor cells. *Stem Cells* **24(4)**, 889-895.

Aliotta J.M., Passero M., Meharg J., Klinger J., Dooner M.S., Pimentel J., Quesenberry P.J. (2005) Stem cells and pulmonary metamorphosis: New concepts in repair and regeneration. *J Cell Physiol* **204**, 725-741.

Altuvia Y., Landgraf P., Lithwick G., Elefant N., Pfeffer S., Aravin A., Brownstein M.J., Tuschl T., Margalit H. (2005) Clustering and conservation patterns of human microRNAs. *Nucleic Acids Res* **33**, 2697-2706.

Ambros V., Bartel B., Bartel D.P., Burge C.B., Carrington J.C., Chen X., Dreyfuss G., Eddy S.R, Griffiths-Jones S., Marshall M. *et al*. (2003) A uniform system for microRNA annotation. *RNA* **9**, 277-279.

Arteaga-Vázquez M., Caballero-Pérez J., Vielle-Calzada J.P. (2006) A family of microRNAs present in plants and animals. *Plant Cell* **18(12)**, 3355-3369.

Ashraf S.I., McLoon A.L., Sclarsic S.M., Kunes S. (2006) Synaptic protein synthesis associated with memory is regulated by the RISC pathway in Drosophila. *Cell* **124(1)**, 191-205.

Ason B., Darnell D.K., Wittbrodt B., Berezikov E., Kloosterman W.P., Wittbrodt J., Antin P.B., Plasterk R.H.A. (2006) Differences in vertebrate microRNA expression. *Proc Natl Acad Sci USA* **103**, 14385-14389.

Avilion A.A., Nicolis S.K., Pevny L.H., Vivian N., Lovell-Badge R. (2003) Multipotent cell lineages in early mouse development depend on sox2 function. *Genes Dev* **17**, 126-140.

Bain G., Kitchens D., Yao M., Huettner J.E., Gottlieb D.I. (1995) Embryonic stem cells express neuronal properties *in vitro*. *Dev Biol* **168(2)**, 342-357.

Barberi T., Klivenyi P., Calingasan N.Y., Lee H., Kawamata H., Loonam K., Perrier A.L., Bruses J., Rubio M.E., Topf N. *et al*. (2003) Neural subtype specification of fertilization and nuclear transfer embryonic stem cells and application in parkinsonian mice. *Nat Biotech* **21**, 1200-1207.

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

Bentwich I. (2005) Prediction and validation of microRNAs and their targets. *FEBS Lett* **579**, 5901-5910.

Berezikov E. and Plasterk R.H.A. (2005) Camels and zebrafish, viruses and cancer: a microRNA update. *Hum Mol Genet* **14**, R183-R190.

Berezikov E., Cuppen E., Plasterk R.H.A. (2006) Approaches to microRNA discovery. *Nat Gen* **38**, S2-S7.

Bernstein E., Caudy A.A., Hammond S.M., Hannon G.J. (2001) Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* **409**, 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., Hannon G.J. (2003) Dicer is essential for mouse development. *Nat Genet* **35**, 215-217.

Bhattacharyya S.N., Habermacher R., Martine U., Closs E.I., Filipowicz W. (2006) Relief of microRNA-mediated translational repression in human cells subjected to stress. *Cell* **125(6)**, 1111-1124.

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

Botquin V., Hess H., Fuhrmann G., Anastassiadis C., Gross M.K., Vriend G., Schöler H.R. (1998) New POU dimer configuration mediates antagonistic control of an osteopontin preimplantation enhancer by Oct-4 and Sox-2. *Genes Dev* **12**, 2073-2090.

Botrugno O.A., Fayard E., Annicotte J.S., Haby C., Brennan T., Wendling O., Tanaka T., Kodama T., Thomas W., Auwerx J. *et al*. (2004) Synergy between LRH-1 and β-catenin induces $G_1$ cyclin-mediated cell proliferation. *Mol Cell* **15**, 499-509.

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. *et al*. (2005) Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* **122**, 947-956.

Boyer L.A., Mathur D., Jaenisch R. (2006) Molecular control of pluripotency. *Curr Opin Genet Dev* **16**, 455-462.

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

Buchan J.R. and Parker R. (2007) Molecular biology. The two faces of microRNA. *Science* **318(5858)**, 1877-1878.

Burdon T., Chambers I., Stracey C., Niwa H., Smith A. (1999) Signalling mechanisms regulating self-renewal and differentiation of pluripotent embryonic stem cells. *Cells Tissues Organs* **165**, 131-143.

Burdon T., Stracey C., Chambers I., Nichols J., Smith A. (1999) Suppression of SHP-2 and ERK signalling promotes selfrenewal of mouse embryonic stem cells. *Dev Biol* **210**, 30-43.

Burrow C.R. (2000) Retinoids and renal development. *Exp Nephrol* 8(4-5), 219-225.

Cai X., Hagedorn C.H., Cullen B.R. (2004) Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. *RNA* **10**, 1957-1966.

Calin G.A., Dumitru C.D., Shimizu M., Bichi R., Zupo S., Noch E., Aldler H., Rattan S., Keating M., Rai K. *et al*. (2002) Frequent deletions and down-regulation of micro-RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. *Proc Natl Acad Sci USA* **99**, 15524-15529.

Carmell M.A., Xuan Z., Zhang M.Q., Hannon G.J. (2002) The Argonaute family: tentacles that reach into RNAi, developmental control, stem cell maintenance, and tumorigenesis. *Genes Dev* **16**, 2733-2742.

Caudy A.A., Myers M., Hannon G.J., Hammond S.M. (2002) Fragile X-related protein and VIG associate with the RNA interference machinery. *Genes Dev* **16**, 2491-2496.

Chambers I., Colby D., Robertson M. Nichols J., Lee S., Tweedie S., 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., Bartel D.P. (2004) MicroRNAs modulate hematopoietic lineage differentiation. *Science* **303**, 83-86.

Chen K. and Rajewsky N. (2007) The evolution of gene regulation by transcription factors and microRNAs. *Nat Rev Genet* **8(2)**, 93-103.

Cheng A.M., Saxton T.M., Sakai R., Kulkarni S., Mbamalu G., Vogel W., Tortorice C.G., Cardiff R.D., Cross J.C., Muller W.J. *et al*. (1998) Mammalian Grb2 regulates multiple steps in embryonic development and malignant transformation. *Cell* **95**, 793-803.

Chew J.L., Loh Y.H., Zhang W., Chen X., Tam W.L., Yeap L.S., Li P., Ang Y.S., Lim B., Robson P. *et al*. (2005) Reciprocal transcriptional regulation of Pou5f1 and Sox2 via the Oct4/Sox2 complex in embryonic stem cells. *Mol Cell Biol* **25**, 6031-6046.

Choe A., Phun H.Q., Tieu D.D., Hu Y.H., Carpenter E.M. (2006) Expression patterns of Hox10 paralogous genes during lumbar spinal cord development. *Gene Expr Patterns* **6(7)**, 730-737.

Choi D., Lee H.J., Jee S., Jin S., Koo S.K., Paik S.S., Jung S.C., Hwang S.Y., Lee K.S., Oh B. (2005) *In vitro* differentiation of mouse embryonic stem cells: Enrichment of endodermal cells in the embryoid body. *Stem Cells* **23**, 817-827.

Cohen D., Segal M., Reiner O. (2008) Doublecortin supports the development of dendritic arbors in primary hippocampal neurons. *Dev Neurosci* **30(1-3)**, 187-199.

Cui X.Y., Hu Q.D., Tekaya M., Shimoda Y., Ang B.T., Nie D.Y., Sun L., Hu W.P., Karsak M., Duka T., *et al*. (2004) NB-3/Notch1 pathway via Deltex1 promotes neural progenitor cell differentiation into oligodendrocytes. *J Biol Chem* **279(24)**, 25858-25865.

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

Doetschman T.C., Eistetter H., Katz M., Schmidt W., Kemler R. (1985) The *in vitro* development of blastocyst-derived embryonic stem cell lines: formation of visceral yolk sac, blood islands and myocardium. *J Embryol Exp Morph* **87**, 27-45.

Downing G.J. and Battey J.F. Jr. (2004) Technical assessment of the first 20 years of research using mouse embryonic stem cell lines. *Stem Cells* **22**, 1168-1180.

Elbashir S.M., Lendeckel W., Tuschl T. (2001) RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes Dev* **15**, 188-200.

Enright A.J., John B., Gaul U., Tuschl T., Sander C., Marks, D.S. (2003) MicroRNA targets in *Drosophila*. *Genome Biol* **5**, R1.

Erlandsson A. and Morshead C.M. (2006) Exploiting the properties of adult stem cells for the treatment of disease. *Curr Opin Mol Ther* **8(4)**, 331-337.

Esau C., Kang X., Peralta E., Hanson E., Marcusson E.G., Ravichandran L.V., Sun Y., Koo S., Perera R.J., Jain R. *et al*. (2004). MicroRNA-143 regulates adipocyte differentiation. *J Biol Chem* **279**, 52361-52365.

Evans M.J. and Kaufman M. (1981) Establishment in culture of pluripotent cells from mouse embryos. *Nature* **292**, 154-156.

Evans M.J. and Kaufman M. (1983) Pluripotential cells grown directly from normal mouse embryos. *Cancer Surv* **2**, 185-208.

Clop A., Marcq F., Takeda H., Pirottin D., Tordoir X., Bibé B., Bouix J., Caiment F., Elsen J.M., Eychenne F*., et al.* (2006) A mutation creating a potential illegitimate microRNA target site in the myostatin gene affects muscularity in sheep. *Nat Genet* **38**, 813-818.

Filippov V., Solovyev V., Filippova M., Gill, S. S. (2000) A novel type of RNase III family proteins in eukaryotes. *Gene* **245**, 213-221.

Finley M.F., Devata S., Huettner J.E. (1999) BMP-4 inhibits neural differentiation of murine embryonic stem cells. *J Neurobiol* **40**, 271-287.

Flynt A.S., Li N., Thatcher E.J., Solnica-Krezel L., Patton J.G. (2007) Zebrafish miR-214 modulates Hedgehog signaling to specify muscle cell fate. *Nat Genet* **39(2)**, 259-263.

Fortin K.R., Nicholson R.H., Nicholson A.W. (2002) Mouse ribonuclease III. cDNA structure, expression analysis, and chromosomal location. *BMC Genomics* **3**, 26.

Friedrich G. and Soriano P. (1991) Promoter traps in embryonic stem cells: a genetic screen to identify and mutate developmental genes in mice. *Genes Dev* **5**, 1513-1523.

Friel R., van der Sar S., Mee P.J. (2005) Embryonic stem cells: Understanding their history, cell biology and signaling. *Adv Drug Deliv Rev* **57**, 1894-1903.

Friocourt G., Liu J.S., Antypa M., Rakic S., Walsh C.A., Parnavelas J.G. (2007) Both doublecortin and doublecortin-like kinase play a role in cortical interneuron migration. *J Neurosci* **27(14)**, 3875-3883.

Fujikura J., Yamato E., Yonemura S., Hosoda K., Masui S., Nakao K., Miyazaki J.-i. Niwa H. (2002) Differentiation of embryonic stem cells is induced by GATA factors. *Genes Dev* **16**, 784-789.

Gassmann M., Donoho G., Berg P. (1995) Maintenance of an extra-chromosomal plasmid vector in mouse embryonic stem cells. *Proc Natl Acad Sci USA* **92**, 1292-1296.

Geijsen N., Horoschak M., Kim K., Gribnau J., Eggan K., Daley G.Q. (2004) Derivation of embryonic germ cells and male gametes from embryonic stem cells. *Nature* **427**, 148-154.

George A.D. and Tenenbaum S.A. (2006) MicroRNA modulation of RNA-binding protein regulatory elements. *RNA Biol* **3(2)**, 57-59.

Giraldez A.J., Cinalli R.M., Glasner M.E., Enright A.J., Thomson J.M., Baskerville S., Hammond S.M., Bartel D.P., Schier A.F. (2005) MicroRNAs regulate brain morphogenesis in zebrafish. *Science* **308(5723)**, 833-838.

Graham V., Khudyakov J., Ellis P., Pevny L. (2003) SOX2 functions to maintain neural progenitor identity. *Neuron* **39**,749-765.

Gregory R.I., Yan K.P., Amuthan G., Chendrimada T., Doratotaj B., Cooch N., Shiekhattar R. (2004) The Microprocessor complex mediates the genesis of microRNAs. *Nature* **432**, 235-240.

Griffiths-Jones, S. (2006) miRBase: the microRNA sequence database. *Methods Mol Bio*l **342**, 129-138.

Grishok A., Pasquinelli A.E., Conte D., Li N., Parrish S., Ha I., Baillie D.L., Fire A., Ruvkun G., 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**, 23-34.

Gu P., Goodwin B., Chung A.C., Xu X., Wheeler D.A., Price R.R., Galardi C., Peng L., Latour A.M., Koller B.H. *et al*. (2005) Orphan nuclear receptor LRH-1 is required to maintain Oct4 expression at the epiblast stage of embryonic development. *Mol Cell Biol* **25(9)**, 3492-3505.

Guo Y., Costa R., Ramsey H., Starnes T., Vance G., Robertson K., Kelley M., Reinbold R., Scholer H., Hromas R. (2002) The embryonic stem cell transcription factors Oct4 and FoxD3 interact to regulate endodermal-specific promoter expression. *Proc Natl Acad Sci USA* **99**, 3663-3667.

Gusfield D. (1997) Algorithms on strings, trees, and sequences: computer science and computational biology. (Cambridge University Press).

Hammond S.M., Bernstein E., Beach D., Hannon G.J. (2000) An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells. *Nature* **404**, 293-296.

Hammond S.M., Boettcher S., Caudy A.A., Kobayashi, R., Hannon, G. J. (2001) Argonaute2, a link between genetic and biochemical analyses of RNAi. *Science* **293**, 1146-1150.

Han J., Day J.R., Thomson K., Connor J.R., Beard J.L. (2000) Iron deficiency alters H- and L-ferritin expression in rat brain. *Cell Mol Biol* **46(3)**, 517-528.

Han J., Lee Y., Yeom K.H., Kim Y.K., Jin H., Kim V.N. (2004) The Drosha–DGCR8 complex in primary microRNA processing. *Genes Dev* **18**, 3016-3027.

Hanna L.A., Foreman R.K., Tarasenko I.A., Kessler D.S., Labosky P.A. (2002) Requirement for Foxd3 in maintaining pluripotent cells of the early mouse embryo. *Genes Dev* **16**, 2650-2661.

Harris M.N., Ozpolat B., Abdi F., Gu S., Legler A., Mawuenyega K.G., Tirado-Gomez M., Lopez-Berestein G., Chen X. (2004) Comparative proteomic analysis of all-trans-retinoic acid treatment reveals systematic posttranscriptional control mechanisms in acute promyelocytic leukemia. *Blood* **104(5)**, 1314-1323.

Hart A.H., Hartley L., Ibrahim M., Robb L. (2004) Identification, cloning and expression analysis of the pluripotency promoting Nanog genes in mouse and human. *Dev Dyn* **230**, 187-198.

Hay D.C., Sutherland L., Clark J., Burdon T. (2004) Oct-4 knockdown induces similar patterns of endoderm and trophoblast differentiation markers in human and mouse embryonic stem cells. *Stem Cells* **22**, 225-235.

Houbaviy H.B., Murray M.F., Sharp P.A. (2003) Embryonic stem cell-specific microRNAs. *Dev Cell* **5(2)**, 351-358.

Hromas R., Ye H., Spinella M., Dmitrovsky E., Xu D., Costa R.H. (1999) Genesis, a winged helix transcriptional repressor, has embryonic expression limited to the neural crest, and stimulates proliferation *in vitro* in a neural development model. *Cell Tiss Res* **297**, 371-382.

Hutvágner G., McLachlan J., Pasquinelli A.E., Bálint E., Tuschl T., Zamore P.D. 8. (2001) A cellular function for the RNA-interference enzyme Dicer in the maturation of the *let-7* small temporal RNA. *Science* **293**, 834-83.

Hwang I.K., Yoo K.Y., Li H., Choi J.H., Kwon Y.G., Ahn Y., Lee I.S., Won M.H. (2007) Differences in doublecortin immunoreactivity and protein levels in the hippocampal dentate gyrus between adult and aged dogs. *Neurochem Res* **32(9)**, 1604-1609.

Ishizuka A., Siomi M. C., Siomi H. (2002) A *Drosophila* fragile X protein interacts with components of RNAi and ribosomal proteins. *Genes Dev* **16**, 2497-2508.

Ivanova N., Dobrin R., Lu R., Kotenko I., Levorse J., DeCoste C., Schafer X., Lun Y., Lemischka I.R. (2006) Dissecting self-renewal in stem cells with RNA interference. *Nature* **442**, 533-538.

Jackson M., Baird J.W., Cambray N., Ansell J.D., Forrester L.M., Graham G.J. (2002) Cloning and characterization of Ehox, a novel homeobox gene essential for embryonic stem cell differentiation. *J Biol Chem* **277(41)**, 38683-38692.

Jaenisch R. and Bird A. (2003) Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat Genet* 33, S245-S254.

Jin P., Zarnescu D.C., Ceman S., Nakamoto M., Mowrey J., Jongens T.A., Nelson D.L., Moses K., Warren S.T. (2004) Biochemical and genetic interaction between the fragile X mental retardation protein and the microRNA pathway. *Nat Neurosci* **7**, 113-117.

John B., Enright A.J., Aravin A., Tuschl T., Sander C., Marks D.S. (2004) Human microRNA targets. *PLoS Biol* **3(7)**, e363.

Johnston R.J. and Hobert O. (2003) A microRNA controlling left/right neuronal asymmetry in *Caenorhabditis elegans*. *Nature* **426(6968)**, 845-849.

Kanai-Azuma M., Kanai Y., Gad J.M., Tajima Y., Taya C., Kurohmaru M., Sanai Y., Yonekawa H., Yazaki K., Tam P.P., Hayashi Y. (2002) Depletion of definitive gut endoderm in Sox17-null mutant mice. *Development* **129**, 2367-2379.

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**, 489-501.

Kawauchi T. and Hoshino M. (2008) Molecular pathways regulating cytoskeletal organization and morphological changes in migrating neurons. *Dev Neurosci* **30(1-3)**, 36-46.

Keays D.A. (2007) Neuronal migration: unraveling the molecular pathway with humans, mice, and a fungus. *Mamm Genome* **18(6-7)**, 425-430.

Kee K., Gonsalves J.M., Clark A.T., Pera R.A. (2006) Bone morphogenetic proteins induce germ cell differentiation from human embryonic stem cells. *Stem Cells Dev* **15**, 831-837.

Kerr C.L., Shamblott M.J., Gearhart J.D. (2006) Pluripotent stem cells from germ cells. *Methods Enzymol* **419**, 400-426.

Ketting R.F. (2007) A Dead End for MicroRNAs. *Cell* **131(7)**, 1226-1227.

Ketting R.F., Fischer S.E., Bernstein E., Sijen T., Hannon G.J., 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**, 2654-2659.

Khvorova A., Reynolds A., Jayasena S.D. (2003) Functional siRNAs and microRNAs exhibit strand bias. *Cell* **115**, 209–216.

Kim K. (2005) Personal communication from *Keystone Symposia on Stem Cells, Senescence and Cancer*, Singapore.

Kim V.N. (2005) MicroRNA biogenesis: coordinated cropping and dicing. *Nat Rev Mol Cell Biol* **6(5)**, 376-385.

Kiriakidou M., Nelson P.T., Kouranov A., Fitziev P., Bouyioukos C., Mourelatos Z., Hatzigeorgiou, A. (2004) A combined computational-experimental approach predicts human microRNA targets. *Genes Dev* **18**, 1165-1178.

Kiriakidou M., Tan G.S., Lamprinaki S., De Planell-Saguer M., Nelson P.T., Mourelatos Z. (2007) An mRNA m7G cap binding-like motif within human Ago2 represses translation. *Cell* **129**, 1141-1151.

Kleinsmith L.J. and Pierce G.B. Jr. (1964) Multipotentiality of single embryonal carcinoma cells. *Cancer Res* **24**, 1544-1551.

Kloosterman W.P. and Plasterk R.H.A. (2006) The diverse functions of microRNAs in animal development and disease. *Dev Cell* **11**, 441-450.

Kloosterman W.P., Wienholds E., de Bruijn E., Kauppinen S., Plasterk R.H. (2006) *In situ* detection of microRNAs in animal embryos using LNA-modified oligonucleotide probes. *Nat Methods* **3**, 27-29.

Kloosterman W.P., Wienholds E., Ketting R.F., Plasterk R.H. (2004) Substrate requirements for let-7 function in the developing zebrafish embryo. *Nucleic Acids Res* **32(21)**, 6284-6291.

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**, 2269-2271.

Koizumi H., Higginbotham H., Poon T., Tanaka T., Brinkman B.C., Gleeson J.G. (2006) Doublecortin maintains bipolar shape and nuclear translocation during migration in the adult forebrain. *Nat Neurosci* **9(6)**, 779-786.

Krek A., Grun D., Poy M.N., Wolf R., Rosenberg L., Epstein E.J., MacMenamin P., da Piedade I., Gunsalus K.C., Stoffel M., *et al*. (2005) Combinatorial microRNA target predictions. *Nat Genet* **37**, 495-500.

Krichevsky A.M., King K.S., Donahue C.P., Khrapko K., Kosik K.S. (2004) A microRNA array reveals extensive regulation of microRNAs during brain development. *RNA* **9(10)**, 1274-1281.

Krichevsky A.M., Sonntag K.C., Isacson O., Kosik K.S. (2006) Specific microRNAs modulate embryonic stem cell-derived neurogenesis. *Stem Cells* **24(4)**, 857-864.

Kyba M., Perlingeiro R.C., Daley G.Q. (2002) HoxB4 confers definitive lymphoid-myeloid engraftment potential on embryonic stem cell and yolk sac hematopoietic progenitors. *Cell* **109**, 29-37.

Kuroda T., Tada M., Kubota H., Kimura H., Hatano S.Y., Suemori H., Nakatsuji N., Tada T. (2005) Octamer and Sox elements are required for transcriptional cis regulation of Nanog gene expression. *Mol Cell Biol* **25**, 2475-2485.

Lagos-Quintana M., Rauhut R., Lendeckel W., Tuschl T. (2001) Identification of novel genes coding for small expressed RNAs. *Science* **294**, 853-858.

Lagos-Quintana M., Rauhut R., Yalcin A., Meyer J., Lendeckel W., Tuschl T. (2002) Identification of tissue-specific microRNAs from mouse. *Curr Biol* **12(9)**, 735-739.

Lau N.C., Lim L.P., Weinstein E.G., Bartel D.P. (2001) An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans. Science* **294**, 858-862.

Lai E.C., Tam B., Rubin G.M. (2005) Pervasive regulation of *Drosophila* Notch target genes by GY-box-, Brd-box-, and K-boxclass microRNAs. *Genes Dev* **19**, 1067-1080.

Landthaler M., Yalcin A., Tuschl, T. (2004) The human DiGeorge syndrome critical region gene 8 and its *D. melanogaster* homolog are required for microRNA biogenesis. *Curr Biol* **14**, 2162-2167.

Leahy A., Xiong J.-W., Kuhnert F., Stuhlmann H. (1999) Use of developmental marker genes to define temporal and spatial patterns of differentiation during embryoid body formation. *J Exp Zool* 284, 67-81.

Lee R.C., Feinbaum R.L., Ambros V. (1993) The *C. elegans* heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell* **75**, 843-854.

Lee Y., Kim M., Han J., Yeom K.H., Lee S., Baek S.H., Kim V.N. (2004) MicroRNA genes are transcribed by RNA polymerase II. *EMBO J* **23**, 4051-4060.

Lee Y., Ahn C., Han J., Choi H., Kim J., Yim J., Lee J., Provost P., Rådmark O., Kim S., *et al*. (2003) The nuclear RNase III Drosha initiates microRNA processing. *Nature* **425**, 415-419.

Lee Y.S., Kim H.K., Chung S., Kim K.S., Dutta A. (2005) Depletion of human micro-RNA miR-125b reveals that it is critical for the proliferation of differentiated cells but not for the downregulation of putative targets during differentiation. *J Biol Chem* **280**, 16635-16641.

Lensch M.W., Daheron L., Schlaeger T.M. (2006) Pluripotent stem cells and their niches. *Stem Cell Rev* **2(3)**, 185-201.

Lewis B.P., Shih I.H., Jones-Rhoades M.W., Bartel D.P., Burge C.B. (2003) Prediction of mammalian microRNA targets. *Cell* **115**, 787-798.

Lewis B.P., Burge C.B., Bartel, D.P. (2005) Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* **120**, 15-20.

Li L.C., Okino S.T., Zhao H., Pookot D., Place R.F., Urakami S., Enokida H., Dahiya R. (2006) Small dsRNAs induce transcriptional activation in human cells. *Proc Natl Acad Sci USA* **103(46)**, 17337-17342.

Li M., Pevny L., Lovell-Badge R., Smith A. (1998) Generation of purified neural precursors from embryonic stem cells by lineage selection. *Curr Biol* **8**, 971-974.

Li Y., Wang F., Lee J.A., Gao F.B. (2006) MicroRNA-9a ensures the precise specification of sensory organ precursors in *Drosophila*. *Genes Dev* **20(20)**, 2793-2805.

Lickert H., Kutsch S., Kanzler B., Tamai Y., Taketo M.M., Kemler R. (2002) Formation of multiple hearts in mice following deletion of b-catenin in the embryonic endoderm. *Dev Cell* **3**, 171-181.

Lingel A., Simon B., Izaurralde E., Sattler, M. (2004) Nucleic acid 3'-end recognition by the Argonaute2 PAZ domain. *Nat Struct Mol Biol* **11**, 576-577.

Liu J., Carmell M.A., Rivas F.V., Marsden C.G., Thomson J.M., Song J.J., Hammond S.M., Joshua-Tor L., Hannon G.J. (2004) Argonaute2 is the catalytic engine of mammalian RNAi. *Science* **305**, 1437-1441.

Liu N., Lu M., Tian X., Han Z. (2007) Molecular mechanisms involved in self-renewal and pluripotency of embryonic stem cells. *J Cell Physiol* **211**, 279-286.

Liu Q., Rand T.A., Kalidas S., Du F., Kim H.E., Smith D.P., Wang X. (2003) R2D2, a bridge between the initiation and effector steps of the *Drosophila* RNAi pathway. *Science* **301**, 1921-1925.

Loh Y.H., Wu Q., Chew J.L., Vega V.B., Zhang W., Chen X., Bourque G., George J., Leong B., Liu J., *et al*. (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., *et al*. (2005). MicroRNA expression profiles classify human cancers. *Nature* **435**, 834-838.

Luciano D.J., Mirsky H., Vendetti N.J., Maas, S. (2004) RNA editing of a microRNA precursor. *RNA* **10**, 1174-1177.

Lund E., Guttinger S., Calado A., Dahlberg J. E., Kutay U. (2004) Nuclear export of microRNA precursors. *Science* **303**, 95-98.

Luo Y., Schwartz C., Shin S., Zeng X., Chen N., Wang Y., Yu X., Rao M.S. (2006) A focused microarray to assess dopaminergic and glial cell differentiation from fetal tissue or embryonic stem cells. *Stem Cells* **24**, 865-875.

Lytle J.R., Yario T.A., 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 USA* **104(23)**, 9667-9672.

Ma J.B., Ye K., Patel D.J. (2004) Structural basis for overhang-specific small interfering RNA recognition by the PAZ domain. *Nature* **429**, 318-322.

Maden M. (2001) Role and distribution of retinoic acid during CNS development. *Int Rev Cytol* **209**, 1-77.

Mansfield J.H., Harfe B.D., Nissen R., Obenauer J., Srineel J., Chaudhuri A., Farzan-Kashani R., Zuker M., Pasquinelli A.E., Ruvkun G., *et al*. (2004) MicroRNA-

responsive ‘sensor’ transgenes uncover Hox-like and other developmentally regulated patterns of vertebrate microRNA expression. *Nat Genent* **36**, 1079-1083.

Maroney P.A., Yu Y., Fisher J., Nilsen T.W. (2006) Evidence that microRNAs are associated with translating messenger RNAs in human cells. *Nat Struct Mol Biol* **13**, 1102-1107.

Martin G. (1981) Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc Natl Acad Sci USA* **78**, 7634-7638.

Martin G. (1980) Teratocarcinomas and mammalian embryogenesis. *Science* **209**, 768-776.

Martin G.R. and Evans M.J. D(1975) Differentiation of clonal lines of teratocarcinoma cells: formation of embryoid bodies *in vitro*. *Proc Natl Acad Sci USA* **72**, 1441-1445.

Martin G.R., Wiley L.M., Damjanov I. (1977) The development of cystic embryoid bodies *in vitro* from clonal teratocarcinoma stem cells. *Dev Biol* **61**, 230-244.

Mathonnet G., Fabian M.R., Svitkin Y.V., Parsyan A., Huck L., Murata T., Biffo S., Merrick W.C., Darzynkiewicz E., Pillai R.S., *et al*. (2007) MicroRNA inhibition of translation initiation *in vitro* by targeting the cap-binding complex eIF4F. *Science* **317(5845)**, 1764-1767.

Matsuda T., Nakamura T., Nakao K., Arai T., Katsuki M., Heike T., Yokota T. (1999) STAT3 activation is sufficient to maintain an undifferentiated state of mouse embryonic stem cells. *EMBO J* **18**, 4261-4269.

McManus M.T. (2003) MicroRNAs and cancer. *Semin Cancer Biol* **13**, 253-258.

Meister G., Landthaler M., Patkaniowska A., Dorsett Y., Teng G., Tuschl T. (2004) Human Argonaute2 mediates RNA cleavage targeted by microRNAs and siRNAs. *Mol Cell* **15**, 185-197.

Meshorer E. and Misteli T. (2006) Chromatin in pluripotent embryonic stem cells and differentiation. *Nat Rev Mol Cell Biol* **7**, 540-546.

Metzler M., Wilda M., Busch K., Viehmann S., Borkhardt A. (2004) High expression of precursor microRNA-155/BIC RNA in children with Burkitt lymphoma. *Genes, Chromosomes Cancer* **39**, 167-169.

Meynert A. and Birney E. (2006) Picking pyknons out of the human genome. *Cell* **125(5)**, 836-838.

Mimeault M. and Batra SK. (2006) Concise review: Recent advances on the significance of stem cells in tissue regeneration and cancer therapies. *Stem Cells* **24(11)**, 2319-2345.

Miranda K.C., Huynh T., Tay Y., Ang Y.S., Tam W.L., Thomson A.M., Lim B., Rigoutsos I. (2006) A pattern-based method for the identification of microRNA binding sites and their corresponding heteroduplexes. *Cell* **126(6)**, 1203-1217.

Miska E.A., Alvarez-Saavedra E., Townsend M., Yoshii A., Sestan N., Rakic P., Constantine-Paton M., Horvitz H.R. (2004) Microarray analysis of microRNA expression in the developing mammalian brain. *Genome Biol* **5(9)**, R68.

Miska E.A. (2005) How microRNAs control cell division, differentiation and death. *Curr Opin Genet Dev* **15**, 563-568.

Mitsui K., Tokuzawa Y., Itoh H., Segawa K., Murakami M., Takahashi K., Maruyama M., Maeda M., Yamanaka S. (2003) The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ESCs. *Cell* **113(5)**, 631-642.

Moores C.A., Perderiset M., Kappeler C., Kain S., Drummond D., Perkins S.J., Chelly J., Cross R., Houdusse A., Francis F. (2006) Distinct roles of doublecortin modulating the microtubule cytoskeleton. *EMBO J* **25(19)**, 4448-4457.

Moss E.G. and Tang L. (2003) Conservation of the heterochronic regulator Lin-28, its developmental expression and microRNA complementary sites. *Dev Biol* **258(2)**, 432-442.

Mourelatos Z., Dostie J., Paushkin S., Sharma A., Charroux B., Abel L., Rappsilber J., Mann M., Dreyfuss G. (2002) miRNPs: a novel class of ribonucleoproteins containing numerous microRNAs. *Genes Dev* **16**, 720-728.

Muckenthaler M., Gray N.K., Hentze M.W. (1998) IRP-1 binding to ferritin mRNA prevents the recruitment of the small ribosomal subunit by the cap-binding complex eIF4F. *Mol Cell* **2**, 383-388.

Nakata H. and Kozasa T. (2005) Functional characterization of Galphao signaling through G protein-regulated inducer of neurite outgrowth 1. *Mol Pharmacol* **67(3)**, 695-702.

Nakayama N., Han C.E., Scully S., Nishinakamura R., He C., Zeni L., Yamane H., Chang D., Yu D., Yokota T., *et al*. (2001) A novel chordin-like protein inhibitor for bone morphogenetic proteins expressed preferentially in mesenchymal cell lineages. *Dev Biol* **232(2)**, 372-87.

Nakielny S. and Dreyfuss G. (1999) Transport of proteins and RNAs in and out of the nucleus. *Cell* **99**, 677-690.

Nichols J., Zevnik B., Anastassiadis K., Niwa H., Klewe-Nebenius D., Chambers I., Scholer H., Smith, A. (1998) Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. *Cell* **95**, 379-391.

Nichols J., Chambers I., Taga T., Smith A. (2001) Physiological rationale for responsiveness of mouse embryonic stem cells to gp130 cytokines. *Development* **128**, 2333-2339.

Nishimoto M., Fukushima A., Okuda A., Muramatsu M.,. (1999) The gene for the embryonic stem cell coactivator UTF1 carries a regulatory element which selectively interacts with a complex composed of Oct-3/4 and Sox-2. *Mol Cell Biol* **19**, 5453-5465.

Niwa H. (2001) Molecular mechanism to maintain stem cell renewal of ESCs. *Cell Struct Funct* **26**, 137-148.

Niwa H., Burdon T., Chambers I., Smith A. (1998) Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3. *Genes Dev* **12**, 2048-2060.

Niwa H., Miyazaki J., Smith A.G. (2000) Quantitative expression of Oct3/4 defines differentiation, dedifferentiation or self-renewal of ESCs. *Nat Genet* **24**, 372-376.

Niwa H., Toyooka Y., Shimosato D., Strumpf D., Takahashi K., Yagi R., Rossant, J. (2005) Interaction between Oct3/4 and Cdx2 determines trophectoderm differentiation. *Cell* **123**, 917-929.

Nottrott S., Simard M.J., Richter J.D. (2006) Human let-7a microRNA blocks protein production on actively translating polyribosomes. *Nat Struct Mol Biol* **13**, 1108-1114.

O'Flaherty J., Mei Y., Freer M., Weyman C.M. (2006) Signaling through the TRAIL receptor DR5/FADD pathway plays a role in the apoptosis associated with skeletal myoblast differentiation. *Apoptosis* **11(12)**, 2103-2113.

Ocbina P.J., Dizon M.L., Shin L., Szele F.G. (2006) Doublecortin is necessary for the migration of adult subventricular zone cells from neurospheres. *Mol Cell Neurosci* **33(2)**, 126-135.

Okada Y., Shimazaki T., Sobue G., Okano H. (2004) Retinoic-acid-concentration-dependent acquisition of neural cell identity during *in vitro* differentiation of mouse embryonic stem cells. *Dev Biol* **275(1)**, 124-142.

Okumura-Nakanishi S., Saito M., Niwa H., Ishikawa F. (2005) Oct-3/4 and Sox2 regulate Oct-3/4 gene in embryonic stem cells. *J Biol Chem* **280**, 5307-5317.

Olsen P.H. and Ambros V. (1999) The lin-4 regulatory RNA controls developmental timing in *Caenorhabditis elegans* by blocking LIN-14 protein synthesis after the initiation of translation. *Dev Biol* **216**, 671-680.

Osborn S.L., Sohn S.J., Winoto A. (2007) Constitutive phosphorylation mutation in Fas-associated Death Domain. (FADD) results in early cell cycle defects. *J Biol Chem* **282(31)**, 22786-22792.

Palmqvist L., Glover C.H., Hsu L., Lu M., Bossen B., Piret J.M., Humphries R.K., Helgason C.D. (2005) Correlation of mouse embryonic stem cell gene expression profiles with functional measures of pluripotency. *Stem Cells* **23**, 663-680.

Pan G. and Thomson J.A. (2007) Nanog and transcriptional networks in embryonic stem cell pluripotency. *Cell Res* **17**, 42-49.

Pan G., Li J., Zhou Y., Zheng H., Pei D. (2006) A negative feedback loop of transcription factors that controls stem cell pluripotency and self-renewal. *FASEB J* **20**, 1730-1732.

Pan G.J., Chang Z.Y., Scholer H.R., Pei D. (2002) Stem cell pluripotency and transcription factor Oct4. *Cell Res* **12**, 321-329.

Pasquinelli A.E., Reinhart B.J., Slack F., Martindale M.Q., Kuroda M.I., Maller B., Hayward D.C., Ball E.E., Degnan B., Müller P., *et al*. (2000) Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. *Nature* **408**, 86-89.

Pesce M., Gross M.K., Scholer H. (1998) In line with our ancestors: Oct-4 and the mammalian germ. *BioEssays* **20**, 722-732.

Peters L. and Meister G. (2007) Argonaute proteins: mediators of RNA silencing. *Mol Cell* **26**, 611-623.

Petersen C.P., Bordeleau M.E., Pelletier J., Sharp P.A. (2006) Short RNAs repress translation after initiation in mammalian cells. *Mol Cell* **21**, 533-542.

Pfeffer S., Lagos-Quintana M., Tuschl T. (2003) Cloning of small RNA molecules. pp. 26.4.1-26.4.18. In *Current Protocols in Molecular Biology Vol. 4*, Ausubel F. *et al*., eds. (John Wiley & Sons).

Pillai R.S., Bhattacharyya S.N., Filipowicz W. (2007) Repression of protein synthesis by microRNAs: how many mechanisms? *Trends Cell Biol* **17**, 118-126.

Poy M.N., Eliasson L., Krutzfeldt J., Kuwajima S., Ma X., Macdonald P.E., Pfeffer S., Tuschl T., Rajewsky N., Rorsman P., *et al*. (2004) A pancreatic islet-specific microRNA regulates insulin secretion. *Nature* **432(7014)**, 226-230.

Qi X., Li T.G., Hao J., Hu J., Wang J., Simmons H., Miura S., Mishina Y., Zhao G.Q. (2004) BMP4 supports self-renewal of embryonic stem cells by inhibiting mitogen-activated protein kinase pathways. *Proc Natl Acad Sci USA* **101(16)**, 6027-6032.

Qian J., Jiang Z., Li M., Heaphy P., Liu Y.H., Shackleford G.M. (2003) Mouse *Wnt9b* transforming activity, tissue-specific expression, and evolution. *Genomics* **81**, 34-46.

Rajewsky N. (2006) MicroRNA target predictions in animals. *Nat Genet* **38 Supp 1**, S8-S13.

Rajewsky N., and Socci N.D. (2004) Computational identification of microRNA targets. *Dev Biol* **267**, 529-535.

Ramain P., Khechumian K., Seugnet L., Arbogast N., Ackermann C., Heitzler P. . (2001) Novel Notch alleles reveal a Deltex-dependent pathway repressing neural fate. *Curr Biol* **11(22)**, 1729-1738.

Rao M. (2004) Conserved and divergent paths that regulate self-renewal in mouse and human embryonic stem cells. *Dec Biol* **275**, 269-286.

Rao S, and Orkin S.H. (2006) Unraveling the transcriptional network controlling ESC pluripotency. *Genome Biol* **7**, 230.

Rehmsmeier M., Steffen P., Hochsmann M., Giegerich R. (2004) Fast and effective prediction of microRNA/target duplexes. *RNA* **10**, 1507-1517.

Reinhart B.J., Slack F.J., Basson M., Pasquinelli A.E., Bettinger J.C., Rougvie A.E., Horvitz H.R., Ruvkun G. (2000) The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans. Nature* **403**, 901-906.

Reynolds A., Leake D., Boese Q., Scaringe S., Marshall W.S., Khvorova A. (2004) Rational siRNA design for RNA interference. *Nat Biotechnol* **22**, 326-330.

Rhoades M.W., Reinhart B.J., Lim L.P., Burge C.B., Bartel B., Bartel D.P. (2002) Prediction of plant microRNA targets. *Cell* **110**, 513-520.

Rigoutsos I., Huynh T., Miranda K., Tsirigos A., McHardy A., Platt D. (2006) Short blocks from the noncoding parts of the human genome have instances within nearly all known genes and relate to biological processes. *Proc Natl Acad Sci USA* **103(17)**, 6605-6610.

Robertson E.J. (1987) Embryo-derived cell lines. pp. 71-112. In *Teratocarcinomas and Embryonic stel cells: A practical approach,* Robertson E.J. ed. (IRL Press).

Rodda D.J., Chew J.L., Lim L.H., Loh Y.H., Wang B., Ng H.H., Robson P. (2005) Transcriptional regulation of nanog by OCT4 and SOX2. *J Biol Chem* **280**, 24731-24737.

Rodriguez A., Griffiths-Jones S., Ashurst J. L., Bradley A. (2004) Identification of mammalian microRNA host genes and transcription units. *Genome Res* **14**, 1902-1910.

Rouault T.A. (2005) The intestinal heme transporter revealed. *Cell* **122**, 649-651.

Sammarco M.C., Ditch S., Banerjee A., Grabczyk E. (2007) Ferritin L and H subunits are differentially regulated on a post-transcriptional level. *J Biol Chem* Dec 26 [Epub ahead of print]

Sasai Y. (2002) Generation of dopaminergic neurons from embryonic stem cells. *J Neurol* **249(suppl 2)**, II41-II44.

Sato N., Meijer L., Skaltsounis L., Greengard P., Brivanlou A.H.,. (2004) Maintenance of pluripotency in human and mouse embryonic stem cells through activation of Wnt signaling by a pharmacological GSK-3- specific inhibitor. *Nat Med* **10**, 55-63.

Schratt G.M., Tuebing F., Nigh E.A., Kane C.G., Sabatini M.E., Kiebler M., Greenberg M.E. (2006) A brain-specific microRNA regulates dendritic spine development. *Nature* **439**, 283-289.

Schüller U., Lamp E.C., Schilling K. (2001) Developmental expression of heterotrimeric G-proteins in the murine cerebellar cortex. *Histochem Cell Biol* **116(2)**, 149-159.

Schwab R., Palatnik J.F., Riester M., Schommer C., Schmid M., Weigel D. (2005) Specific effects of microRNAs on the plant transcriptome. *Dev Cell* **8**, 517-527.

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

Seggerson K., Tang L., Moss E.G. (2002) Two genetic circuits repress the Caenorhabditis elegans heterochronic gene lin-28 after translation initiation. *Dev Biol* **243**, 215-225.

Segupathy P., Megraw M., Hatzigeorgiou A.G. (2006) A guide through present computational approaches for the identification of mammalian microRNA targets. *Nat Meth* **3**, 881-886.

Seitz H., Royo H., Bortolin M.L., Lin S.P., Ferguson-Smith A.C., Cavaillé J. (2004) A large imprinted microRNA gene cluster at the mouse Dlk1-Gtl2 domain. *Genome Res* **14(9)**, 1741-1748.

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

Serakinci N. and Keith W.N. (2006) Therapeutic potential of adult stem cells. *Euro J Cancer* **42(9)**, 1243-1246.

Slack F.J., Basson M., Liu Z., Ambros V., Horvitz H.R., Ruvkun G. (2000) The lin-41 RBCC gene acts in the *C. elegans* heterochronic pathway between the let-7 regulatory RNA and the LIN-29 transcription factor. *Mol Cell* **5**, 659-669.

Smirnova L., Gräfe A., Seiler A., Schumacher S., Nitsch R., Wulczyn F.G. (2005) Regulation of microRNA expression during neural cell specification. *Eur J Neurosci* **21(6)**, 1469-1477.

Smith A.G., Heath J.K., Donaldson D.D., Wong G.G., Moreau J., Stahl M., Rogers D. (!998) Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides. *Nature* **336**, 688-690.

Smith A.G. (2001) Embryo-derived stem cells: Of mice and men. *Annu Rev Cell Dev Biol* **17**, 435-462.

Song J.J., Liu J., Tolia N.H., Schneiderman J., Smith S.K., Martienssen R.A., Hannon G.J., 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**, 1026-1032.

Song J.J., Smith S.K., Hannon G.J., Joshua-Tor L. (2004) Crystal structure of Argonaute and its implications for RISC slicer activity. *Science* **305**, 1434-1437.

Stark A., Brennecke J., Russell R.B., Cohen S.M. (2003) Identification of *Drosophila* MicroRNA targets. *PLoS Biol* **1(3)**, E60.

Stefanovic S. and Puceat M. (2007) Oct-3/4: not just a gatekeeper of pluripotency for embryonic stem cell, a cell fate instructor through a gene dosage effect. *Cell Cycle* **6(1)**, 8-10.

Stevens L.C. (1970) The development of transplantable teratocarcinomas from intratesticular grafts of pre- and postimplantation mouse embryos. *Dev Biol* 21, 364-382.

Stevens L.C., Varnum D.S., Eicher E.M. (1977) Viable chimeras produced from normal and parthenogenetic mouse embryos. *Nature* **269**, 515-517.

Stevens L.C. (1978) Totipotent cells of parthenogenetic origin in a chimeric mouse. *Nature* **276**, 266-267.

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., *et al*. (2004) Human embryonic stem cells express a unique set of microRNAs. *Dev Biol* **270(2)**, 488-498.

Tabara H., Yigit E., Siomi H., Mello C.C. (2002) The dsRNA binding protein RDE-4 interacts with RDE-1, DCR-1, and a DExH-box helicase to direct RNAi in *C. elegans*. *Cell* **109**, 861-871.

Takamizawa J., Konishi H., Yanagisawa K., Tomida S., Osada H., Endoh H., Harano T., Yatabe Y., Nagino M., Nimura Y., *et al*. (2004) Reduced expression of the let-6 microRNAs in human lung cancers in association with shortened postoperative survival. *Cancer Res* **64**, 3753-3756.

Tamura Y., Kataoka Y., Cui Y., Takamori Y., Watanabe Y., Yamada H. (2007) Multi-directional differentiation of doublecortin- and NG2-immunopositive progenitor cells in the adult rat neocortex *in vivo*. *Eur J Neurosci* **25(12)**, 3489-3498.

Teleman A.A. and Cohen S.M. (2006) *Drosophila* lacking microRNA miR-278 are defective in energy homeostasis. *Genes Dev* **20**, 417-422.

Thermann R., and Hentze M.W. (2007) Drosophila miR2 induces pseudo-polysomes and inhibits translation initiation. *Nature* **447**, 875-878.

Thomson A.M., Rogers J.T., Walker C.E., Staton J.M., Leedman P.J. (1999) Optimized RNA gel-shift and UV cross-linking assays for characterization of cytoplasmic RNA-protein interactions. *Biotechniques* **27(5)**, 1032-1039.

Thomson A.M., Cahill C.M., Cho H., Kassachau K.D., Epis M.R., Bridges K.R., Leedman P.J., Rogers J.T. (2005) The acute box cis-element in human heavy ferritin mRNA 5'-untranslated region is a unique translation enhancer that binds poly(c)-binding proteins. *J Biol Chem* **280**, 30032-30045.

Thomson J.M., Parker J., Perou C.M., Hammond S.M. (2004) A custom microarray platform for analysis of microRNA gene expression. *Nat Methods* **1(1)**, 47-53.

Tokuzawa Y., Kaiho E., Maruyama M., Takahashi K., Mitsui K., Maeda M., Niwa H. Yamanaka, S. (2003) Fbx 15 is a novel target of Oct3/4 but dispensable for ESC self-renewal and mouse development. *Mol Cell Biol* **23**, 2699-2708.

Tolkunova E., Cavaleri F., Eckardt S., Reinbold R., Christenson L.K, Schöler H.R., Tomilin A. (2006) The caudal-related protein Cdx2 promotes trophoblast differentiation of mouse embryonic stem cells. *Stem Cells* **24**, 139-144.

Tomioka M., Nishimoto M., Miyagi S., Katayanagi T., Fukui N., Niwa H., Muramatsu M., Okuda A. (2002) Identification of Sox-2 regulatory region which is under the control of Oct-3/4-Sox-2 complex. *Nucleic Acids Res* **30**, 3202-3213.

Ui-Tei K., Naito Y., Takahashi F., Haraguchi T., Ohki-Hamazaki H., Juni A., Ueda R., Saigo K. (2004) Guidelines for the selection of highly effective siRNA sequences for mammalian and chick RNA interference. *Nucleic Acids Res* **32**, 936-948.

van Rooij E., Sutherland L.B., Liu N., Williams A.H., McAnally J., Gerard R.D., Richardson J.A., Olson E.N. (2006) A signature pattern of stress-responsive microRNAs that can evoke cardiac hypertrophy and heart failure. *Proc Natl Acad Sci U S A* **103(48)**, 18255-18260.

van Rooij E., Sutherland L.B., Qi X., Richardson J.A., Hill J., Olson E.N. (2007) Control of stress-dependent cardiac growth and gene expression by a microRNA. *Science* **316(5824)**, 575-579.

Vasudevan S., Tong Y., Steitz J.A. (2007) Switching from repression to activation: microRNAs can up-regulate translation. *Science* **318(5858)**, 1931-1934.

Verwer R.W., Sluiter A.A., Balesar R.A., Baayen J.C., Noske D.P., Dirven C.M., Wouda J., van Dam A.M., Lucassen P.J., Swaab D.F. (2007) Mature astrocytes in the adult human neocortex express the early neuronal marker doublecortin. *Brain* **130(Pt 12)**, 3321-3335.

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., *et al*. (2006) A genetic screen implicates microRNA-372 and microRNA-373 as oncogenes in testicular germ cell tumours. *Cell* **124(6)**, 1169-1181.

Wakiyama M., Takimoto K., Ohara O., Yokoyama, S. (2007) Let-7 microRNA-mediated mRNA deadenylation and translational repression in a mammalian cell-free system. *Genes Dev* **21**, 1857-1862.

Wang B., Love T.M., Call M.E., Doench J.G., Novina, C.D. (2006) Recapitulation of short RNA-directed translational gene silencing *in vitro*. *Mol Cell* **22**, 553-560.

Wang W., Chan E.K., Baron S., Van de Water T., Lufkin T. (2001) Hmx2 homeobox gene control of murine vestibular morphogenesis. *Development* **128**, 5017-5029.

Wang Y., Medvid R., Melton C., Jaenisch R., Blelloch R. (2007) DGCR8 is essential for microRNA biogenesis and silencing of embryonic stem cell self-renewal. *Nat Genet* **39(3)**, 380-385.

Wichterle H., Lieberam I., Porter J.A., Jessell T.M. (2002) Directed differentiation of embryonic stem cells into motor neurons. *Cell* **110**, 385-397.

Wienholds E., Koudijs M.J., van Eeden F.J., Cuppen E., Plasterk R.H. (2003) The microRNA-producing enzyme Dicer1 is essential for zebrafish development. *Nat Genet* **35**, 217-218.

Wienholds E., Kloosterman W.P., Miska E., Alvarez-Saavedra E., Berezikov E., de Bruijn E., Horvitz H.R., Kauppinen S., Plasterk R.H. (2005) MicroRNA expression in zebrafish embryonic development. *Science* **309**, 310-311.

Wightman B., Ha I. Ruvkun G. (1993) Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in *C. elegans*. *Cell* **75**, 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., Gough N.M. (1988) Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells. *Nature* **336**, 684-687.

Wilson P.A. and Hemmati-Brivanlou A. (1995) Induction of epidermis and inhibition of neural fate by Bmp-4. *Nature* **376**, 331- 333.

Wobus A.M., Kaomei G., Shan J., Wellner M.C., Rohwedel J., Ji G., Fleischmann B., Katus H.A., Hescheler J., Franz W.M. (1997) Retinoic acid accelerates embryonic stem cell-derived cardiac differentiation and enhances development of ventricular cardiomyocytes. *J Mol Cell Cardiol* **29(6)**, 1525-1539.

Wu H., Xu H., Miraglia L.J., Crooke S.T. (2000) Human RNase III is a 160-kDa protein involved in preribosomal RNA processing. *J Biol Chem* **275**, 36957-36965.

Wu L. and Belasco J.G. (2005) MicroRNA regulation of the mammalian lin-28 gene during neuronal differentiation of embryonal carcinoma cells. *Mol Cell Biol* **25(21),** 9198-9208.

Xu P., Vernooy S.Y., Guo M., Hay B.A. (2003) The Drosophia microRNA miR-14 suppressESC death and is required for normal fat metabolism. *Curr Biol* **13(9)**, 790-705.

Yamada T., Yoshikawa M., Takaki M., Torihashi S., Kato Y., Nakajima Y., Ishizaka S., Tsunoda Y. (2002) *In vitro* functional gut-like organ formation from mouse embryonic stem cells. *Stem Cells* **20**, 41-49.

Yan K.S., Yan S., Farooq A., Han A., Zeng .L, Zhou M.M. (2003) Structure and conserved RNA binding of the PAZ domain. *Nature* **426**, 468-474.

Yang D.H. and Moss E.G. (2003) Temporally regulated expression of Lin-28 in diverse tissues of the developing mouse. *Gene Expr Patterns* **3(6)**, 719-726.

Yang W., Chendrimada T.P., Wang Q., Higuchi M., Seeburg P.H., Shiekhattar R., Nishikura K. (2006) Modulation of microRNA processing and expression through RNA editing by ADAR deaminases. *Nat Struct Mol Biol* **13**, 13-21.

Yang Z., Ebright Y.W., Yu B., Chen X. (2006) HEN1 recognizes 21–24 nt small RNA duplexes and deposits a methyl group onto the 2′ OH of the 3′ terminal nucleotide. *Nucleic Acids Res* **34**, 667-675.

Yashiro K., Zhao X., Uehara M., Yamashita K., Nishijima M., Nishino J., Saijoh Y., Sakai Y., Hamada H. (2004) Regulation of retinoic acid distribution is required for proximodistal patterning and outgrowth of the developing mouse limb. *Dev Cell* **6(3)**, 411-422.

Yeh W.C., Pompa J.L., McCurrach M.E., Shu H.B., Elia A.J., Shahinian A., Ng M., Wakeham A., Khoo W., Mitchell K., *et al*. (1998) FADD: essential for embryo development and signaling from some, but not all, inducers of apoptosis. *Science* **279(5358)**, 1954-1958.

Yekta S., Shih I.H., Bartel D.P. (2004) MicroRNA-directed cleavage of HOXB8 mRNA. *Science* **304**, 594-596.

Yeom Y.I., Fuhrmann G., Ovitt C.E., Brehm A., Ohbo K., Gross M., Hubner K., Scholer H.R. (1996) Germline regulatory element ofOct-4 specific for the totipotent cycle of embryonal cells. *Development* **122**, 881-894.

Yi R., Qin Y., Macara I.G., Cullen, B.R. (2003) Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes Dev* **17**, 3011-3016.

Ying Q.L., Nichols J., Chambers I., Smith A. (2003) BMP induction of Id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3. *Cell* **115**, 281-292.

Ying Q.L., Stavridis M., Griffiths D., Li M., Smith A. . (2003) Conversion of embryonic stem cells into neuroectodermal precursors in adherent monoculture. *Nat Biotechnol* **21(2)**, 183-186.

Yu J., Vodyanik M.A., Smuga-Otto K., Antosiewicz-Bourget J., Frane J.L., Tian S., Nie J., Jonsdottir G.A., Ruotti V., Stewart R. et al. (2007) Induced pluripotent stem cell lines derived from human somatic cells. *Science* **318(5858)**,1917-1920.

Yuan H., Corbi N., Basilico C., Dailey, L. (1995) Developmental-specific activity of the FGF-4 enhancer requires the synergistic action of Sox2 and Oct-3. *Genes Dev* **9**, 2635-2645.

Zeng X., Miura T., Luo Y., Bhaskar B., Condie B., Lyons I., Puri R.K., Rao M.S., Freed W. (2004) Properties of pluripotent human embryonic stem cells BG01 and BG02. *Stem Cells* **22(3)**, 292-312.

Zeng Y. and Cullen B.R. (2004) Structural requirements for premicroRNA binding and nuclear export by Exportin 5. *Nucleic Acids Res* **32**, 4776-4785.

Zhang H., Kolb F. A., Brondani V., Billy E., Filipowicz W. (2002) Human Dicer preferentially cleaves dsRNAs at their termini without a requirement for ATP. *EMBO J* **21**, 5875-5885.

Zhang H., Kolb F. A., Jaskiewicz L., Westhof E., Filipowicz W. (2004) Single processing center models for human Dicer and bacterial RNase III. *Cell* **118**, 57-68.

Zhao G.Q., Deng K., Labosky P.A., Liaw L., Hogan B.L. (1996) The gene encoding bone morphogenetic protein 8B is required for the initiation and maintenance of spermatogenesis in the mouse. *Genes Dev* **10(13)**, 1657-1669.

# CHAPTER 10. APPENDICES

## 10.1 Gene names & sequences of miR-375 target predictions tested

| Rank | ENSEMBL GENE-ID | TESTED TARGET SITE | ENSEMBL DESCRIPTION |
|---|---|---|---|
| 1 | ENSMUSG00000007617 | GAGAAGAGCCAGAACGAACAAG | homer homolog 1 |
| 2 | ENSMUSG00000047298 | TGAATGTGAGCCAGGGAACAAG | potassium channel, subfamily V, member 2 |
| 3 | ENSMUSG00000014504 | ATTTAACTGAGCTGTGAACAAG | signal recognition particle 19 |
| 4 | ENSMUSG00000028560 | TGATGGATGAGCTGGGAACAAA | ubiquitin specific protease 1 |
| 5 | ENSMUSG00000040265 | CCATACTGAGCTGCAGAACAAG | dynamin 3 |
| 6 | ENSMUSG00000001507 | AAACGCAGCCCGAAGGAACAAA | integrin alpha 3 |
| 7 | ENSMUSG00000011256 | ATACAATGTGAGCAAGAACAAG | disintegrin and metalloproteinase domain 19 |
| 8 | ENSMUSG00000040247 | AAGAGAGCCAGAGAAGAACAAA | TBC1 domain family, member 10c |
| 9 | ENSMUSG00000026872 | TATGCAGGGCTATAAGAACAAA | zinc finger homeobox 1b |
| 10 | ENSMUSG00000057093 | TCCCCTGCAGCTGTTGAACAAA | RIKEN cDNA C030039L03 |
| 11 | ENSMUSG00000023088 | CATGGGAACCAAAATGAACAAA | ATP-binding cassette, sub-family C (CFTR/MRP), member 1 |
| 12 | ENSMUSG00000035051 | CCATTGTTAGCTAATGAACAAA | expressed sequence AW494914 |
| 13 | ENSMUSG00000053870 | TGGTAGTGATGCTGTGAACAAA | fucose-1-phosphate guanylyltransferase |
| 14 | ENSMUSG00000049928 | AGAGTGAGCCATGATGAACAAG | glucagon-like peptide 2 receptor |
| 15 | ENSMUSG00000021870 | CCACTGAGCTATTGGGAACAAG | sarcolemma associated protein |
| 16 | ENSMUSG00000044081 | CATGTTAAGCGCGAAGAACAAG | RIKEN cDNA 4930441O14 gene |
| 17 | ENSMUSG00000031958 | CAGAAAAAGTTGAAAGAACAAA | lactate dehydrogenase D |
| 18 | ENSMUSG00000042156 | GTAACACAGTGCCGAGAACAAA | Zinc finger protein DZIP1 |
| 19 | ENSMUSG00000025956 | TTAGAGCTGCTCAAGGAACAAA | RIKEN cDNA 2310038H17 gene |

| | | | |
|---|---|---|---|
| 20 | ENSMUSG00000050295 | TCTGAGTTGGCACATGAACAAA | forkhead box C1 |
| 21 | ENSMUSG00000038022 | CGAGCTGGAGGCAAAGAACAAG | NM_177883, NM_177883.2 |
| 22 | ENSMUSG00000042074 | AGCACGCAGAAGAAAGAACAAA | Mus musculus 12 days embryo eyeball cDNA |
| 23 | ENSMUSG00000057522 | GTGTCTGACTGAGCAGAACAAA | speckle-type POZ protein |
| 24 | ENSMUSG00000030722 | CACCTTTGTGCCGGAGAACAAG | nuclear factor of activated T-cells, calcineurin-dependent 2 interacting protein |
| 25 | ENSMUSG00000030319 | CACTCAGGTCGTAGTGAACAAA | cullin-associated and neddylation-dissociated 2 |
| 26 | ENSMUSG00000021820 | AGCCAGCTGAGAAGAGAACAAA | calcium/calmodulin-dependent protein kinase II gamma |
| 27 | ENSMUSG00000022865 | CCTCTAAGAGTCAGTGAACAAA | coxsackievirus and adenovirus receptor |
| 28 | ENSMUSG00000033542 | GTCTTCTGAGCTCATGAACAAG | Rho guanine nucleotide exchange factor (GEF) 5 |
| 29 | ENSMUSG00000030678 | CTGGAGCCAGAGGCAGAACAAG | MYC-associated zinc finger protein |
| 30 | ENSMUSG00000025907 | TTAAAAGCTGGCACTGAACAAA | RB1-inducible coiled-coil 1 |
| 31 | ENSMUSG00000018470 | AGTAGCGATTCGCATGAACAAA | potassium voltage-gated channel, shaker-related subfamily, beta member 3 |
| 32 | ENSMUSG00000033342 | ACACGATGATAAAACGAACAAG | RIKEN cDNA 4833424O15 gene |
| 33 | ENSMUSG00000052917 | CCAGCCGAAGGTGATGAACAAA | SUMO1/sentrin specific protease 7 |
| 34 | ENSMUSG00000047261 | TCAATGTGATGGAATGAACAAA | growth associated protein 43 |
| 35 | ENSMUSG00000001829 | GAAACATGCAGCGGAGAACAAG | suppressor of K+ transport defect 3 |
| 36 | ENSMUSG00000045817 | GTTGTGATGCGGAAAGAACAAA | zinc finger protein 36, C3H type-like 2 |
| 37 | ENSMUSG00000031021 | GATGTGGGCAGCTTTGAACAAG | RIKEN cDNA 2310004K06 gene |
| 38 | ENSMUSG00000033420 | AAGTCGTGATGGGCAGAACAAA | anthrax toxin receptor 1 |
| 39 | ENSMUSG00000026782 | CTGGTGATGCTAGAAGAACAAG | abl-interactor 2 |
| 40 | ENSMUSG00000020057 | CACGCATAAATGGATGAACAAG | RIKEN cDNA 1200002N14 gene |
| 41 | ENSMUSG00000055044 | ATGGCGGAGTGAAGGAACAAAC | PDZ and LIM domain 1 |

| | | | |
|---|---|---|---|
| 42 | ENSMUSG00000023828 | TCCAGGAAGCAGGATGAACAAA | solute carrier family 22 |
| 43 | ENSMUSG00000031174 | CCTGAGTTTGTGAATGAACAAG | retinitis pigmentosa GTPase regulator |
| 44 | ENSMUSG00000034998 | CTAAGGCAGCTGGAGGAACAAG | forkhead box N2 |

## 10.2 Gene names & sequences of miR-296 target predictions tested

| Rank | ENSEMBL GENE-ID | TESTED TARGET SITE | ENSEMBL DESCRIPTION |
|---|---|---|---|
| 1 | ENSMUSG00000037904 | CCTTCAGGATAAAGGGGCCCT | Ankyrin repeat domain 9. |
| 2 | ENSMUSG00000033751 | AGGGAGGAGATGCTGGGGCCC | papillomavirus L2 interacting nuclear protein 1. |
| 3 | ENSMUSG00000059810 | GATTGCAGTTAGGAGGGGCCC | regulator of G-protein signaling 3; C2 membrane binding, PDZ protein/protein interaction, ATP/GTP binding domains. |
| 4 | ENSMUSG00000047415 | CGGAGAGATGTCAGGGGCCCT | Sphingosylphosphorylcholine receptor (G protein-coupled receptor 68). |
| 5 | ENSMUSG00000024146 | GACTGAGAGCCATGGGGCCCA | postsynaptic protein Cript; CRIPT protein. |
| 6 | ENSMUSG00000026754 | TGTGTTGTGGCTGTGGGGCCC | Golgi autoantigen, golgin subfamily A member 1 (Golgin-97). |
| 7 | ENSMUSG00000063931 | CACTCAGTGAGACTGGGCCCT | Xaa-Pro dipeptidase (X-Pro dipeptidase) (Proline dipeptidase) (Prolidase) (Imidodipeptidase) (Peptidase 4). |
| 8 | ENSMUSG00000020657 | TCCCTGGGAATGAGGGGCCCG | Rab-related GTP-binding protein. |
| 9 | ENSMUSG00000030032 | CAGTAGGGCTGAGTGGGCCCT | D3Mm3e protein. |
| 10 | ENSMUSG00000031393 | AGGGAGTGACACCAGGGCCCT | Methyl-CpG-binding protein 2 (MeCP-2 protein) (MeCP2). |
| 11 | ENSMUSG00000037531 | AAAGGCATGAGGAAGGGTCCT | 39s ribosomal protein L47, mitochondrial precursor (L47mt) (MRP-L47) (Fragment). |

| | | | |
|---|---|---|---|
| 12 | ENSMUSG00000032099 | TGGATTGTCAGGGTGGGTCCT | Seminal vesicle protein 7 precursor (SVS VII) (Caltrin) (Calcium transport inhibitor). |
| 13 | ENSMUSG00000026208 | ATGGAAGAGTGGGGGGCCTGA | Desmin. |
| 14 | ENSMUSG00000029767 | AGGATGTACAGAGGGGGCTCT | Calumenin precursor (Crocalbin). |
| 15 | ENSMUSG00000028101 | ACTGAGGGATGAGGGGGTCCT | Protein inhibitor of activated STAT protein 3. |
| 16 | ENSMUSG00000021483 | TGTGCTTGAGGGCTGGGCTCT | cell cycle related kinase; CDK-related protein kinase PNQLARE. |
| 17 | ENSMUSG00000046997 | ATCTTTAGGAAGCGGGGCCTT | SPRY domain-containing SOCS box 4; SPRY domain-containing SOCS box protein SSB-4. |
| 18 | ENSMUSG00000030825 | AGGATTCTAGGCAGGGGCTCT | dehydrogenase/reductase (SDR family) member 10. |
| 19 | ENSMUSG00000023017 | GGGGGCAGCAGGAGGGGCTCT | amiloride-sensitive cation channel 2, neuronal; acid sensing ion channel; degenerin 2. |
| 20 | ENSMUSG00000033735 | GGGGTGAGGAGGGTGGGCTCT | Sepiapterin reductase (SPR). |
| 21 | ENSMUSG00000026764 | GGAGTCTGATGGAGGGGTCCT | Kinesin heavy chain isoform 5C (Kinesin heavy chain neuron-specific 2). |
| 22 | ENSMUSG00000031958 | GGCCTGCTGGGACTGGGCCCT | D-lactate dehydrogenase. |
| 23 | ENSMUSG00000039000 | GAGGTGGAGGAGAAGGGTCCT | ubiquitin protein ligase E3C. |
| 24 | ENSMUSG00000035606 | AGAGACAGAGGCCAGGGCCCT | kyphoscoliosis. |

## 10.3 Gene names & sequences of miR-134 target predictions tested

| Rank | ENSEMBL GENE-ID | TESTED TARGET SITE | ENSEMBL DESCRIPTION |
|---|---|---|---|
| 1 | ENSMUSG00000063613 | CTTTTTGTAAGAAACAGTCACA | RECORD HAS BEEN REMOVED |
| 2 | ENSMUSG00000041098 | TGCCTCTGGCACAGCAGTCACA | RECORD HAS BEEN REMOVED |
| 3 | ENSMUSG00000051314 | ATCCTCTGTGTCACCAGTCACT | G protein-coupled receptor 43. |
| 4 | ENSMUSG00000056069 | AGGTTTCAGTTAGCCAGTCACA | cDNA sequence BC052328 |
| 5 | ENSMUSG00000041977 | CTGGCATGTTCAACCAGTCACA | Rho guanine nucleotide exchange factor (GEF) 11 |
| 6 | ENSMUSG00000027180 | ATTACCTGGCAACACAGTCACC | F-box only protein 3 |
| 7 | ENSMUSG00000028137 | GCCTCTCTGCCAGTCAGTCACA | trinucleotide repeat containing 4 |
| 8 | ENSMUSG00000056131 | AACTTTCTGATTAGTAGTCACA | phosphoglucomutase 3 |
| 9 | ENSMUSG00000050379 | CTCAAAGGTGTGATCAGTCACT | septin 6 |
| 10 | ENSMUSG00000036985 | TTTTCTTTGGTCTTTAGTCACC | zinc finger, DHHC domain containing 9 |
| 11 | ENSMUSG00000018486 | TTGGTCAAAGCCAGTCACT | Wnt-9b protein precursor (Wnt-15) (Wnt-14b). |
| 12 | ENSMUSG00000030869 | CCACACATCTGGCCTAGTCACA | NADH dehydrogenase (ubiquinone) 1, alpha/beta subcomplex, 1 |
| 13 | ENSMUSG00000054266 | ATCCTGAGGTAACCCAGTCACA | serine (or cysteine) proteinase inhibitor, clade B, member 9d; serine protease inhibitor 9. |
| 14 | ENSMUSG00000028980 | AGGGCTTTACCAATCAGTCACA | hexose-6-phosphate dehydrogenase (glucose 1-dehydrogenase). |
| 15 | ENSMUSG00000026398 | TGCTGAATGTCAAATAGTCACA | Orphan nuclear receptor NR5A2 (Liver receptor homolog LRH-1) |
| 16 | ENSMUSG00000059857 | ACTTTATCATTATCCAGTCACA | Netrin G1 precursor (Laminet-1). |
| 17 | ENSMUSG00000050244 | TGTGCTCTGGGAAGAAGTCACA | HEAT repeat containing 1 |
| 18 | ENSMUSG00000039981 | CACCCCTTTATTACTAGTCACA | zinc finger CCCH-type containing 12D |
| 19 | ENSMUSG00000039358 | TCTTTTTTCTTAGTCAGTCACA | D(1B) dopamine receptor (D(5) dopamine receptor). |
| 20 | ENSMUSG00000037411 | AATCGTTTGTGTTCCAGTCACA | serine (or cysteine) peptidase inhibitor, clade E, member 1 |

| | | | |
|---|---|---|---|
| 21 | ENSMUSG00000025326 | TTTGCTTCTAACACCAGTCACA | Ubiquitin-protein ligase E3A (Oncogenic protein- associated protein E6-AP). |
| 22 | ENSMUSG00000026227 | CCCTCTGAGTCTGTGAGTCACA | PREDICTED: hypothetical protein LOC72792 |
| 23 | ENSMUSG00000063672 | ATCAGCTGGTTAGGGAGTCACA | NKX6-3 (Fragment). |
| 24 | ENSMUSG00000052591 | CCCCCTGCCTCACACAGTCACA | RIKEN cDNA 3110070M22 gene (3110070M22Rik), mRNA |
| 25 | ENSMUSG00000039108 | AGCACAGGGTGAGCCAGTCACA | cDNA sequence BC040823 |
| 26 | ENSMUSG00000003227 | GGCGTGTGTCTGTCCAGTCACA | Tumor necrosis factor receptor superfamily member EDAR precursor (Anhidrotic ectodysplasin receptor 1) |
| 27 | ENSMUSG00000042448 | CCTTGGGTCTACAGAAGTCACA | Homeobox protein Hox-D1 (Hox-4.9). |
| 28 | ENSMUSG00000024143 | TTAAATGTTCATACCAGTCACA | ras homolog gene family, member Q; small GTPase Tc10; ras-like protein. |
| 29 | ENSMUSG00000030556) | GTGTGATCTCTTTGCAGTCACA | leucine rich repeat containing 28 |
| 30 | ENSMUSG00000031392 | TGTGTTTGAGTCATGAGTCACA | Interleukin-1 receptor-associated kinase 1 (IRAK-1) (IRAK) (Pelle-like protein kinase) (mPLK). |
| 31 | ENSMUSG00000000782 | ACTCTGTGGTCAACCAGCCACA | transcription factor 7, T-cell specific |
| 32 | ENSMUSG00000027746 | AGAAGCTGTTAATATAGTCACA | ubiquitin-fold modifier 1 |
| 33 | ENSMUSG00000020707 | AGCCTGTGGTCAACAAGTCACA | RIKEN cDNA 2410006N06; U 2-3-0. |
| 34 | ENSMUSG00000032059 | TCAAAGTCTAATCAAAGTCACA | disrupted in bipolar disorder 1 homolog (human) |
| 35 | ENSMUSG00000024683) | TTTGTGTTCAAGGCCAGTCACA | mitochondrial ribosomal protein L16 |
| 36 | ENSMUSG00000032440 | TAGCACTTGACAATCAGTCACA | TGF-beta receptor type II precursor (TGFR-2) (TGF-beta type II receptor). |
| 37 | ENSMUSG00000036523 | AACCTTGGGTTATCATGTCACG | gene regulated by estrogen in breast cancer protein |
| 38 | ENSMUSG00000030232 | CGACTTATTTTAACCGGTTACA | AE binding protein 2 |
| 39 | ENSMUSG00000024526 | ATACTGTAATAAACCAGTCACA | Cell death activator CIDE-A (Cell death-inducing DFFA-like effector A). |
| 40 | ENSMUSG00000026069 | CCTTATTTGCTTACCAGTCACA | interleukin 1 receptor-like 1 |

| | | | |
|---|---|---|---|
| 41 | ENSMUSG00000035342 | AACCCTCTGTGTCACCGTCACC | leucine zipper, putative tumor suppressor 2 |
| 42 | ENSMUSG00000036405 | GTTGTCTGTGTTAGCTGTCACA | RECORD_REMOVED |
| 43 | ENSMUSG00000031847 | ACTAGAAGGCTTATTAGTCACA | RIKEN cDNA 1700030J22 gene (1700030J22Rik), mRNA |
| 44 | ENSMUSG00000023960 | AGCTTCTCACAAGCTAGTCACA | ectonucleotide pyrophosphatase/phosphodiesterase 5 |
| 45 | ENSMUSG00000021481 | TCCTTTATTTGGATCAGTCACA | zinc finger protein 346 |
| 46 | ENSMUSG00000049624 | GCGTGGTACCCAGTCAGTCACA | Sialin (Solute carrier family 17 member 5) (Sodium/sialic acid cotransporter). |
| 47 | ENSMUSG00000050700 | TACTAACCAGGGACCAGTCACA | EMILIN 5 precursor (Elastin microfibril interface located protein 5) (Elastin microfibril interfacer 5). |
| 48 | ENSMUSG00000046807 | TTCCTTTGGTTTTAGGGTCACA | expressed sequence AI646023 |
| 49 | ENSMUSG00000020393 | AACCTGTAGGCAGCCAGTCACA | Kremen protein 1 precursor (Kringle-containing protein marking the eye and the nose) (Dickkopf receptor). |
| 50 | ENSMUSG00000030609 | TTTGTGTATAATGTTAGTCACA | interferon stimulated exonuclease gene 20-like 1 |
| 51 | ENSMUSG00000001089 | GACTTTTGGATAACCAGTTTCA | leucine zipper protein 1 |
| 52 | ENSMUSG00000026788 | AGACTCATGTTCCCCAGTCACA | Zinc finger protein 297B. |
| 53 | ENSMUSG00000031077 | AACTGTCCTGTGGCAAGTCACA | FADD protein (FAS-associating death domain-containing protein) (Mediator of receptor induced toxicity). |
| 54 | ENSMUSG00000000127 | TAGGACTCTGTCCTCAGTCACC | fer (fms/fps related) protein kinase, testis specific 2. |
| 55 | ENSMUSG00000024983 | GGGCCGTGATGACTCAGTCACA | vesicle transport through interaction with t-SNAREs homolog 1A (yeast) |
| 56 | ENSMUSG00000025203 | CAAGAGTTGAACTTGAGTCACA | stearoyl-Coenzyme A desaturase 2 |
| 57 | ENSMUSG00000042365 | GTGGTGCTAGCCTGCAGTCACA | RIKEN cDNA A530023G15 gene |
| 58 | ENSMUSG00000021932 | GTGCATGGGTCACCGAGTCACA | RIKEN cDNA 2610207P08 gene (2610207P08Rik), mRNA |
| 59 | ENSMUSG00000054667 | GTCACAGGTACAAATAGTCACA | insulin receptor substrate 4. |
| 60 | ENSMUSG00000024535 | GCCAGCCTCTGCTTCAGTCACA | sorting nexing 24. |

| | | | |
|---|---|---|---|
| 61 | ENSMUSG00000031924 | ATGCTGGGATAGCGAAGTCACA | cytochrome b5 outer mitochondrial membrane precursor. |
| 62 | ENSMUSG00000021266 | CACCCTGTCAAATCAAGTCACA | Tryptophanyl-tRNA synthetase (Tryptophan--tRNA ligase) (TrpRS). |
| 63 | ENSMUSG00000025036 | AGTAACGTTCAGTCCAGTCACA | Sideroflexin 2. |
| 64 | ENSMUSG00000034341 | GCAGACTGTCAGTCCAGTCACA | WW domain binding protein 2 (WBP-2). |
| 65 | ENSMUSG00000028057 | TTCTTTTCTGGCAACTGTCACA | Ras-like without CAAX 1 |
| 66 | ENSMUSG00000027887 | TCCATCTCCTCAGACAGTCACA | Mitsugumin 29. |
| 67 | ENSMUSG00000000276 | AAATCTGCTGGTCACAGTCACA | Diacylglycerol kinase, epsilon (Diglyceride kinase) (DGK-epsilon) (DAG kinase epsilon). |
| 68 | ENSMUSG00000055436) | AATCCCTTTGTCACCAGTCACT | splicing factor, arginine/serine-rich 11 |
| 69 | ENSMUSG00000027430 | CCATTTCCATCATTGAGTCACA | histidyl tRNA synthetase 2 |
| 70 | ENSMUSG00000031443 | GGGGCTGGGCAGTGCAGTCACA | coagulation factor VII |
| 71 | ENSMUSG00000021756 | TGCCTCTAAGGACAAAGTCACA | interleukin 6 signal transducer |
| 72 | ENSMUSG00000048773 | TTAATTCTGGTGACAAGTCACG | Peptidyl-tRNA hydrolase 2, mitochondrial precursor (PTH 2). |
| 73 | ENSMUSG00000003154 | GCATCTGGTGGGCCAGTGTATG | forkhead box J2 |
| 74 | ENSMUSG00000005686 | CCAAGCTCCTGATCCAGTCACA | AMP deaminase 3 (AMP deaminase isoform E) (AMP deaminase H-type) (Heart-type AMPD). |
| 75 | ENSMUSG00000028657 | TGGTCTTGAACTCACAGTCACA | Palmitoyl-protein thioesterase 1 precursor (Palmitoyl-protein hydrolase 1). |
| 76 | ENSMUSG00000042201 | TCCATCTGGTGCATGGGTCACA | RIKEN cDNA C530024P05 gene (C530024P05Rik), mRNA |
| 77 | ENSMUSG00000047910 | TGTGGGCACTCAATGAGTCACA | protocadherin beta 16. |
| 78 | ENSMUSG00000047091 | GTAACTGAGTCAATTAGTTAAG | PDZ domain containing 6 |
| 79 | ENSMUSG00000042129 | GGCTGCAGTTAACCAGTATATA | Ras association (RalGDS/AF-6) domain family 4 |
| 80 | ENSMUSG00000002881 | CCACAGGCTCTAAGCAGTCACA | NGFI-A binding protein 1 (EGR-1 binding protein 1). |

| | | | |
|---|---|---|---|
| 81 | ENSMUSG00000046598 | ACCTTCTCCTGAGCCAGTCACA | 3-hydroxybutyrate dehydrogenase (heart, mitochondrial); D-beta-hydroxybutyrate dehydrogenase. |
| 82 | ENSMUSG00000005649 | AGAGAGTGTCCCACTAGTCACA | Calcium-binding protein CaBP5. |
| 83 | ENSMUSG00000060733 | AGCTTTCTGTAGCCAGTCACCA | inositol polyphosphate multikinase |
| 84 | ENSMUSG00000024044 | CTACACTGTCAACTGTGTCACA | erythrocyte protein band 4.1-like 3 |
| 85 | ENSMUSG00000054843 | CCCTATCTGCAACACAGTCACA | attractin like 1 |
| 86 | ENSMUSG00000032030 | CGACTCTATTTTGCCAGTCACA | Cullin homolog 5 (CUL-5). |
| 87 | ENSMUSG00000041774 | ATCTCCTGGAACTAGAGTCACA | Ubiquitin-conjugating enzyme E2L 3. |
| 88 | ENSMUSG00000042642 | CCTTCACCAAGGATGGGTCACA | RFad1, flavin adenine dinucleotide synthetase, homolog (yeast) |
| 89 | ENSMUSG00000019803 | GCCATAGGGTCAGCGAGTCACA | Orphan nuclear receptor NR2E1 (Nuclear receptor TLX) (Tailless homolog) (Tll) (mTll). |
| 90 | ENSMUSG00000029670 | GAGCTCTGTCAGCTCAGTCGCA | inhibitor of growth family, member 3 |
| 91 | ENSMUSG00000009418 | AAACACCTCTGAGACAGTCACA | neuron navigator 1. |
| 92 | ENSMUSG00000038446 | CTTGACTCTCTTCCTAGTCACG | Pre-mRNA splicing factor PRP17 (Cell division cycle 40 homolog). |
| 93 | ENSMUSG00000031748 | TATTTACTGAAGACCAGTCACA | Guanine nucleotide-binding protein G(o), alpha subunit 1. |
| 94 | ENSMUSG00000035311 | ATATTGGTCATGCTGTGTTATG | v-maf musculoaponeurotic fibrosarcoma oncogene family, protein B (avian) |
| 95 | ENSMUSG00000020719 | GGGGTCTTGTAGGCTAGTCACA | Probable RNA-dependent helicase p68 (DEAD-box protein p68) (DEAD-box protein 5) (DEAD-box RNA helicase DEAD1) (mDEAD1). |
| 96 | ENSMUSG00000040441 | CAGCTCTGGCACATCAGCCATG | expressed sequence C78409 (C78409), mRNA |
| 97 | ENSMUSG00000022987 | CTTTTTGGTGACTGTGTTAATA | zinc finger protein 641 |
| 98 | ENSMUSG00000033475 | GTATGTTCTGTCAGCTAGTCAC | RIKEN cDNA 1110002E23 gene (1110002E23Rik), mRNA |
| 99 | ENSMUSG00000003119 | CTCTAATCAATTCTGAGTCACA | CDC2-related protein kinase 7; protein kinase for splicing component. |
| 100 | ENSMUSG00000048982 | ACAGTGAGGTGACCAGTTACAA | glycoprotein hormone beta 5 |

| | | | |
|---|---|---|---|
| 101 | ENSMUSG00000001444 | CTCTGCCCGAACTACAGTCACG | T-box 21; T-box transcription factor. |
| 102 | ENSMUSG00000041598 | TGCCTCTGGCCAGCAGTACACA | CDC42 effector protein (Rho GTPase binding) 4 |
| 103 | ENSMUSG00000041936 | CGGCACTTGTGACCCAGTCACA | agrin. |
| 104 | ENSMUSG00000058900 | CTAATTTGGGTAAACAGTTATA | regulator of sex limited protein-Slp 1 |
| 105 | ENSMUSG00000008730 | CACTCTCATGGCAACCAGTACA | homeodomain interacting protein kinase 1 |
| 106 | ENSMUSG00000051147 | GGGCCCAAGAAAACCAGTCACA | Arylamine N-acetyltransferase 2 (Arylamide acetylase 2) (N-acetyltransferase type 2) (NAT-2). |
| 107 | ENSMUSG00000025577 | TTTACTGGTTACGTCAGTCCCA | chromobox homolog 2 (Drosophila Pc class) |
| 108 | ENSMUSG00000018209 | TCCTGTGGGGCTGGTAGTCACA | serine/threonine kinase 4; STE20-like kinase MST1; Yeast Sps1/Ste20-related kinase 3 (S. cerevisiae). |
| 109 | ENSMUSG00000043468 | TCTCTGCTGGTACAGCAGTCAA | a disintegrin and metallopeptidase domain 30 |
| 110 | ENSMUSG00000053870 | TACACTGGTAGCTATGTACACA | fucose-1-phosphate guanylyltransferase |
| 111 | ENSMUSG00000022772 | GTGCTGTGGAAAGCCAGTCACA | SUMO/sentrin specific protease 5. |
| 112 | ENSMUSG00000050890 | GGGACTGGTAATCACTGTCACA | PDLIM1 interacting kinase 1 like |
| 113 | ENSMUSG00000026078 | CCAGTCTGGTTTGCATAGTTAC | phosducin-like 3 |
| 114 | ENSMUSG00000053838) | AGCAGAATCTGTCGTAGTCACA | NudC domain containing 3 |
| 115 | ENSMUSG00000017652 | GCACGGGTGTCAGCCTGTCACA | CD40 antigen |
| 116 | ENSMUSG00000048458 | CTTGGCTTGTGGACCAGTCATA | RIKEN cDNA 6530418L21 gene (6530418L21Rik), mRNA |
| 117 | ENSMUSG00000026383 | TCCTACTGAGTCACTAGTTACA | erythrocyte protein band 4.1-like 5 |
| 118 | ENSMUSG00000049699 | CTAGCTGAGTCTGGCAGTCACA | Urocortin II precursor (Ucn II). |
| 119 | ENSMUSG00000027164 | GCCTTGCCCTCTTCCAGTCACA | TNF receptor associated factor 6. |
| 120 | ENSMUSG00000053004 | GACTGTGTGGCTTCTAGTCACA | Histamine H1 receptor. |
| 121 | ENSMUSG00000038827 | CTTTCAAGTTGTATTAGTTACT | cDNA sequence BC026590 |
| 122 | ENSMUSG00000028654 | TTCTCTGCCGGCCTGAGTCACA | L-myc proto-oncogene protein. |
| 123 | ENSMUSG00000051412 | GCCAACTATTAGCCTAGTCATG | synaptobrevin like 1 |

| | | | |
|---|---|---|---|
| 124 | ENSMUSG00000045854 | AGAGCTGTATTAAAGGGTCACA | RIKEN cDNA 2610208E05 gene (2610208E05Rik), mRNA |
| 125 | ENSMUSG00000038879 | TGAGAGGCTCAGCTATGTTATA | NIPA-like domain containing 2 |
| 126 | ENSMUSG00000055538 | CCTGGCATGTCACATGGTCGCA | RIKEN cDNA 2310047A01 gene (2310047A01Rik), mRNA |
| 127 | ENSMUSG00000059994 | TACTTCTGAGTCACTGTCATCA | Fc receptor-like 1 |
| 128 | ENSMUSG00000021867 | AGACCCTCTGGTTACTATTATT | Adult male small intestine cDNA, RIKEN full-length enriched library, clone:2010100M18 |
| 129 | ENSMUSG00000026163 | CTCTATCCCAGCTTAGGTCACA | sphingosine kinase type 1-interacting protein |
| 130 | ENSMUSG00000032202 | TGTCTGTATTTAAGCAGTCACA | Ras-related protein Rab-27A. |
| 131 | ENSMUSG00000026585 | GTCCTTTGGACAATACAGTCCA | kinesin-associated protein 3 |
| 132 | ENSMUSG00000051803 | CGTGACTGTCAGCCAGTCAGCA | No description</p> |
| 133 | ENSMUSG00000028266 | TGTCTGCAGTAGACCAGTCACC | LIM domain transcription factor LMO4 (LIM-only protein 4) (LMO-4) (Breast tumor autoantigen). |
| 134 | ENSMUSG00000044313 | CCAGGCCGTCAGCTAGCTCACC | cDNA sequence BC037703 |
| 135 | ENSMUSG00000046275 | TTCTTCTCATCTCCTAGTCACA | tumor suppressor candidate 5 (Tusc5), mRNA |
| 136 | ENSMUSG00000057784 | TACTATTTTCAACCAGTTACTA | RECORD_REMOVED |
| 137 | ENSMUSG00000030168 | GACTCTGGTTCAACATAGCACA | adiponectin receptor 2 |
| 138 | ENSMUSG00000028572 | AGTTCATGTGCAGCGAGTCACA | Hook homolog 1. |
| 139 | ENSMUSG00000025781 | AAACTTGGTCCACTGAGTTACA | ATP synthase, H+ transporting, mitochondrial F1 complex, gamma polypeptide 1 |
| 140 | ENSMUSG00000001021 | CATCTCTGCTCAGCCATGTGCA | S100 calcium binding protein A3 |
| 141 | ENSMUSG00000063888 | AGACATTCTGGCAGCTAGTATA | ribosomal protein L7-like 1 |
| 142 | ENSMUSG00000028868 | AAGTGACTCTCTCAAGGTCACA | WAS protein family, member 2 |
| 143 | ENSMUSG00000021712 | ATGCTTATGGTTAACAGTTACT | tripartite motif protein 23 |
| 144 | ENSMUSG00000030560 | CTCACAGAGTGATTTAGTCACA | cathepsin C |
| 145 | ENSMUSG00000034101 | ACTGTCTGTTCAGCAGTCAGCA | catenin (cadherin associated protein), delta 1 |
| 146 | ENSMUSG00000031285 | GTTCTACCCCTGTTCAGTCACA | Doublecortin (Lissencephalin-X) (Lis-X) (Doublin). |

| | | | |
|---|---|---|---|
| 147 | ENSMUSG00000047694 | GAAAACGTGTTAACTAGTTTCA | Yip1 domain family, member 6 |
| 148 | ENSMUSG00000047139 | ATACTTTGGTTATCCATTCATA | CD24a antigen |
| 149 | ENSMUSG00000049076 | ATTCAGAGTTAGCTACGTCACT | centaurin, beta 2 |
| 150 | ENSMUSG00000046541 | AACCAGGAATGAACCAGTCACA | zinc finger protein 526. |
| 151 | ENSMUSG00000026489 | AAGTCTGGTTAATTAGGTGATA | chaperone, ABC1 activity of bc1 complex like (S. pombe) |
| 152 | ENSMUSG00000047462 | AATTCTCTGGTGATTAGTAGCA | RIKEN cDNA A530099J19 gene (A530099J19Rik), mRNA |
| 153 | ENSMUSG00000024844 | TTTTTTGGCAAAAACAGTCACT | Barrier-to-autointegration factor (Breakpoint cluster region protein 1) (LAP2-binding protein 1). |
| 154 | ENSMUSG00000021303 | AGTGTCACAAGCCTAAGTTATG | guanine nucleotide binding protein (G protein), gamma 4 subunit |
| 155 | ENSMUSG00000042579 | TCTCCTGGCTCAGCTGTCACCA | RIKEN cDNA 4632404H12 gene (4632404H12Rik), mRNA |
| 156 | ENSMUSG00000038736 | AATCTTTCTTTCTCTAGTCACA | chronic myelogenous leukemia tumor antigen 66. |
| 157 | ENSMUSG00000045885 | TGGGTCTGGAGCTCGGCTCACA | No description</p> |
| 158 | ENSMUSG00000058706 | TAGTCTCCAGTTAATCAGTCCA | PREDICTED: hypothetical protein LOC68364+D151 |

## 10.4 Luciferase results for predicted neural MREs

| No | Gene | Relative Luciferase Activity (134 vs Scr) | |
|---|---|---|---|
| | | 293T | Neuro2As |
| 1 | Acvr1c | 1.32 | 1.06 |
| 2 | Akt3 | 0.68 | 1.09 |
| 3 | Aph1b | 0.75 | 0.89 |
| 4 | Bmp15 | 0.73 | 1.13 |
| 5 | Bmp2k | 0.75 | 0.79 |
| 6 | Bmp8b | 0.33 | 0.49 |
| 7 | Bmper | 0.68 | 1.08 |
| 8 | Bmpr1b | 0.73 | 1.04 |
| 9 | Chrdl1 | 0.27 | 0.16 |
| 10 | Chrdl1 | 0.32 | 0.36 |
| 11 | Chrdl1 | 0.44 | 0.39 |
| 12 | Chrdl1 | 0.75 | 0.89 |
| 13 | Ctnnb1 | 0.72 | 1.01 |
| 14 | Dpp8 | 0.67 | 1.01 |
| 15 | Dpp8 | 0.72 | 1.04 |
| 16 | Dpp8 | 0.71 | 0.92 |
| 17 | Dcx | 0.19 | 0.30 |
| 18 | Dpp8 | 0.98 | 1.10 |
| 19 | Dtx4 | 0.39 | 0.57 |
| 20 | E2f1 | 0.77 | 1.04 |
| 21 | E2f3 | 0.73 | 1.01 |
| 22 | Efnb2 | 0.59 | 1.04 |
| 23 | Egfr | 1.23 | 0.88 |
| 24 | Fzd6 | 0.65 | 0.74 |
| 25 | Gab1 | 0.93 | 0.84 |
| 26 | Hhip | 0.53 | 1.17 |
| 27 | Hoxa1 | 1.04 | 1.09 |
| 28 | Hoxb13 | 0.79 | 0.91 |
| 29 | Hoxc10 | 0.47 | 0.51 |
| 30 | Zfp42 | 0.79 | 1.03 |
| 31 | Myb | 0.66 | 1.34 |
| 32 | Mybl2 | 0.83 | 0.90 |
| 33 | Notch1 | 0.75 | 1.09 |
| 34 | Ntn2l | 1.13 | 1.35 |
| 35 | Otx2 | 0.76 | 1.16 |
| 36 | Sema3c | 0.97 | 1.17 |
| 37 | Sfrp2 | 0.91 | 0.85 |
| 38 | Shh | 0.87 | 1.28 |
| 39 | Smad3 | 0.92 | 1.07 |
| 40 | Smad5 | 1.03 | 0.96 |
| 41 | Tcf12 | 0.61 | 0.65 |
| 42 | Tcf4 | 0.97 | 1.17 |
| 43 | Tcf7 | 1.12 | 0.91 |
| 44 | Tgfb2 | 0.85 | 1.07 |
| 45 | Tgfbr1 | 1.10 | 0.86 |
| 46 | Tgfbr3 | 0.59 | 1.21 |
| 47 | Tgfbr3 | 0.68 | 0.92 |
| 48 | 134 RC | 0.07 | 0.03 |

## 10.5 Summary of *rna22*'s predictions for four model genomes

A

| Genome | Number of Processed 3′UTRs | Number of 3′UTRs Containing One or More "Target Islands" (% Processed 3′UTRs) | Number of Nucleotides in Processed 3′UTRs | Number of "Target Islands" in Processed 3′UTRs |
|---|---|---|---|---|
| *C. elegans* | 13,186 | 9752 (73.9%) | 3,048,704 | 27,700 |
| *D. melanogaster* | 14,965 | 13,104 (87.6%) | 6,671,035 | 63,918 |
| *M. musculus* | 20,257 | 18,597 (91.8%) | 18,058,224 | 180,157 |
| *H. sapiens* | 25,589 | 23,616 (92.3%) | 25,597,040 | 243,211 |

B

| Genome | Number of Processed 5′UTRs | Number of 5′UTRs Containing One or More "Target Islands" (% Processed 5′UTRs) | Number of Nucleotides in Processed 5′UTRs | Number of "Target Islands" in Processed 5′UTRs |
|---|---|---|---|---|
| *C. elegans* | 11,713 | 3654 (31.2%) | 797,941 | 7085 |
| *D. melanogaster* | 15,461 | 12,139 (32.7%) | 4,129,409 | 37,078 |
| *M. musculus* | 19,978 | 10,298 (51.5%) | 4,398,970 | 31,967 |
| *H. sapiens* | 25,042 | 13,350 (53.3%) | 6,947,437 | 46,007 |

C

| Genome | Number of Processed CDSs | Number of CDSs Containing One or More "Target Islands" (% Processed CDSs) | Number of Nucleotides in Processed CDSs | Number of "Target Islands" in Processed CDSs |
|---|---|---|---|---|
| *C. elegans* | 25,811 | 23,515 (91.1%) | 34,476,529 | 362,110 |
| *D. melanogaster* | 19,177 | 19,059 (99.4%) | 32,199,294 | 270,617 |
| *M. musculus* | 31,535 | 31,345 (99.4%) | 42,926,064 | 420,238 |
| *H. sapiens* | 33,869 | 33,545 (99.0%) | 50,737,171 | 476,677 |

(A) Results from the analysis of 3′UTRs.
(B) Results from the analysis of 5′UTRs.
(C) Results from the analysis of CDSs.

(*adapted from Table 1 of Miranda et al., 2006*)

## 10.6 Related publications


1. Miranda K.C.*, Huynh T.*, **Tay Y.***, Ang Y.S.*, Tam W.L., Thomson A.M., Lim B., Rigoutsos I. (2006) A pattern-based method for the identification of microRNA binding sites and their corresponding heteroduplexes. *Cell* **126**, 1203-1217.

2. **Tay Y.**, Thomson A.M., Lim B. (2007) MicroRNAs in ESC differentiation and prediction of their targets. Chap 35, p476-489. In MicroRNAs From Basic Science to Disease Biology, K. Appasani, ed (Cambridge University Press)

*3.* **Tay Y.***, Tam W.L.*, Ang Y.S.*, Gaughwin P.M., Yang H.H., Wang W., Liu R., George J., Ng H.H., Perera R.J., Lufkin T., Rigoutsos I., Thomson A.M., Lim B. (2007) MicroRNA-134 modulates the differentiation of mouse embryonic stem cells where it causes post-transcriptional attenuation of Nanog and LRH1. *Stem Cells Oct 18 Epub ahead of print.*

*4.* **Tay Y.**, Thomson A.M., Huynh T., Zhang J., Lim B., Rigoutsos I. (2007) microRNA-296 post-transcriptionally regulates the murine homeobox protein Nanog by targeting its amino acid coding region. *In preparation.*


* Co-first author