Biogenesis and function of microRNA

miR-124

MANISH MUHURI

SCHOOL OF BIOLOGICAL SCIENCES

2013
Biogenesis and function of microRNA

miR-124

MANISH MUHURI

SCHOOL OF BIOLOGICAL SCIENCES

NANYANG TECHNOLOGICAL UNIVERSITY

SINGAPORE

A thesis submitted to Nanyang Technological University

in partial fulfillment of the requirement for the degree of

Doctor of Philosophy

2013
STATEMENT OF ORIGINALITY

I, Manish Muhuri, hereby certify that the work embodied in this thesis is original work done by me and has not been submitted for a higher degree to any other University or Institute.
ACKNOWLEDGEMENTS

"If I have seen further it is by standing on the shoulder of giants" - Sir Isaac Newton.

Working on a PhD might be perceived by some as an endeavor that has to be completed in seclusion, but in practice one cannot survive without the help and support from others, fruitful scientific discussions, collaborative development of tools and papers and valuable pieces of advice. Although it is my name on the front page and whilst this physical document represents a personal journey this work could not have been possible without a whole raft of people. Here, I would like to take the chance to say ‘thank-you’ to everyone.

This thesis grew out of a series of dialogues with my supervisor Dr. Eugene Makeyev. Through his Socratic questioning, he brought me closer to the reality I had initially perceived, eventually enabling me to grasp its rich complexity. His comments on chapter drafts are themselves a course in critical thought upon which I will always draw. His capacity to combine critique with an immediate empathy and commitment towards workers and others engaged in struggle will always inspire me. He continually and convincingly conveyed a spirit of adventure in regard to research and scholarship. Without his guidance and persistent help this thesis would not have been possible. I would like to express my deepest appreciation and gratitude to him.

I must acknowledge the generous support of the Nanyang Technological University throughout this project. Through the support offered by NTU Research Scholarship, I have been able to spend extended periods of time working on my research and have been afforded opportunities to travel to present my research in progress to my peers at conferences – giving me unparalleled exposure and also forming a vital part of my ability to move through this project.

My gratitude is also directed at my Thesis Advisory Committee members Prof. Mark Featherstone, NTU and Dr. Karuna Sampath, Temasek Lifesciences Laboratory for their advice, comments and supervision of my
project. Their examining input gave me vital leads in my research and gave it direction.

I must also thank my friends for being there and supporting me with friendly advice, cups of coffee and random conversations about what is wrong with the world, particularly Piyush who has been a great friend, sounding board and general finder of random ideas since day one. To my friends Archana, Margaret, Priyangee, Siddharth, Ravishankar and Vinay- thank you for your continued friendship, support, hugs, chats, laughs and cups of tea along the way. Lavanya and Sharadha, thank you for taking the time to proof read this thesis and helping me through the technical aspects, your meticulous inputs were invaluable. Thank you Hari, for all the support and encouragement you give me, and the patience and unwaivering faith you have in me. You believe in me even when I don’t believe in myself. For that I am eternally grateful. Love, always. And Narendar for all the conversations, philosophical debates and inspirational talks over coffee and cake. I am truly lucky to count you amongst my friends. To Reelin for getting me out of the house and my eternal mood swings and making sure I ate properly, not forgetting her ability to always lift me up from the deepest emotional dungeons and enabling me to bypass the more persistent obstacles to thesis completion. To Laxmi and the nonstop chatter, laughter, awesome food and the wonderful music sessions which were the ultimate stress busters. To Swarolipi and Chaitali for their continued support and friendship over the years and being an unwaivering source of inspiration. I am very grateful to all those whose names I could not mention but who have given me their friendship, put up with my odd hours, and provided me with practical help.

In that same vein, I want to thank my labmates Prasanna, Piyush, Karen, Weijun and all the other lab people for providing the perfect lab atmosphere and also making me feel at home with their continued support and cooperation.

I have also been saved many times from depression and writer’s blocks by good music. I cannot name all the artists responsible for that, but the most credit goes to Anoushka Shankar, Shiv Kumar Sharma, Zakir Husain and Hans Zimmer.

I would not have contemplated this road if not for my parents, who instilled within me a love of creative pursuits, science and language, all of
which finds a place in this thesis. They blessed my life with an environment in which I could freely pursue my dreams and have the courage to surmount higher mountains. My brother Jaideep and sister-in-law Rituparna have also been the best of friends along this journey. They mean the world to me and this undertaking would not have been possible without their love, support and trust.
CONTENTS

CONTENTS.......................................................................................................................v
LIST OF FIGURES ...........................................................................................................x
LIST OF TABLES.............................................................................................................xvi
LIST OF ABBREVIATIONS...........................................................................................xvii
ABSTRACT .........................................................................................................................xxiii

Chapter 1 ....................................................................................................................... 1

INTRODUCTION ............................................................................................................ 1

1.1. Non-Coding RNAs .............................................................................................. 2
1.2. Long non-coding RNAs ...................................................................................... 7
1.3. Short non-coding RNAs ...................................................................................... 10
1.4. Piwi interacting RNAs (piRNAs) ........................................................................ 11
1.5. Short interfering RNAs (siRNAs) ..................................................................... 12
1.6. MicroRNAs .......................................................................................................... 15
1.7. Structure of microRNA genes .......................................................................... 18
1.8. MicroRNA Biogenesis ......................................................................................... 24
  1.8.1. Nuclear processing by Drosha ................................................................. 25
  1.8.2. Export of pre-miRNA to the cytoplasm from the nucleus by exportin-5 ...... 27
  1.8.3. Cytoplasmic processing by Dicer ............................................................... 28
  1.8.4. RISC ........................................................................................................... 30
  1.8.5. Biogenesis of intronically encoded miRNAs .......................................... 35
1.9. Regulation of miRNA biogenesis .................................................................... 38
  1.9.1. Transcriptional Control ............................................................................ 39
  1.9.2. Post transcriptional Control ..................................................................... 40
  1.9.3. Feedback circuits in miRNA regulation ............................................... 43
1.10. miRNA-mRNA interaction ............................................................................. 44
1.11. Mechanisms of gene silencing mediated by miRNAs .................................... 45
  1.11.1. Gene repression at the level of translation ............................................. 45
  1.11.2. Gene repression at the level of mRNA stability ..................................... 48
  1.11.3. Regulation at the level of transcription ................................................ 49
  1.11.4. miRNA mediated translational activation ............................................ 50
1.12. MicroRNAs and the Nervous system ........................................... 50
  1.12.1. Nervous System .................................................................. 50
  1.12.2. Nervous System specific miRNAs ....................................... 55
  1.12.3. miR-124 in Nervous System development ......................... 57
1.13. Some outstanding problems of miRNA biology ..................... 61
1.14. Recombination systems used for knocking out genes in mouse... 62
  1.14.1. Homologous recombination ............................................. 62
  1.14.2. Site Specific Recombination System .................................. 66

Chapter 2 ......................................................................................... 71
MATERIALS AND METHODS .............................................................. 71
  2.1. Plasmids .................................................................................. 71
  2.2. Cloning Primers Used ............................................................... 79
  2.3. Molecular Cloning Techniques .................................................. 82
    2.3.1. Digestion with restriction enzymes and purification of DNA fragments .......................................................... 82
    2.3.2. Klenow reactions to repair recessed 3’ termini .................. 82
    2.3.3. Ligation of DNA fragments and bacterial transformation .. 82
    2.3.4. Site Directed Mutagenesis .................................................. 83
    2.3.5. Nucleic acid precipitation ................................................... 83
  2.4. Cell lines .................................................................................. 84
  2.5. Purification of RNA from Cell Lines and Tissues ..................... 84
  2.6. Polyacrylamide gel electrophoresis of RNA (Northern Blotting) ... 85
  2.7. 5’ end labeling by T4 Polynucleotide Kinase reaction ............... 86
  2.8. RT qPCR (Reverse Transcriptase quantitative Polymerase Chain Reaction) ................................................................. 86
    2.8.1. DNase treatment: ............................................................... 86
    2.8.2. Reverse Transcription ....................................................... 87
    2.8.3. Quantitative PCR .............................................................. 87
    2.8.4. Primers Used for Real Time PCR ....................................... 88
  2.9. Multiplex PCR .......................................................................... 88
    2.9.1. Primers used for Multiplex PCR ......................................... 89
  2.10. Western Blotting ................................................................. 90
    2.10.1. Protein isolation ............................................................... 90
    2.10.2. Determining Protein Concentration .................................. 90
2.10.3. SDS PAGE (Sodium Dodecyl Sulfate-PolyAcrylamide Gel Electrophoresis) ................................................................. 90
2.10.4. Transfer ................................................................................ 91
2.10.5. Immunostaining .................................................................. 91
2.10.6. Enhanced ChemiLuminescence (ECL) detection ................. 92
2.11. Luciferase Assay ..................................................................... 92
2.12. PrestoBlue Cell Viability Assay .............................................. 93
2.13. Electroporation of constructs into ES cells ......................... 94
2.14. Selection and Colony Picking of ES cells containing desired constructs ............................................................................... 94
2.15. PCR screen for targeted ES cell clones .................................. 95
2.16. Isolation of Genomic DNA from ES cells ............................... 95
2.17. Southern Blotting .................................................................... 96
2.18. Preparation of ES cells for blastocyst injection ..................... 98
2.19. Estimating mRNA half life ....................................................... 98
2.20. Tail Typing ............................................................................. 99
2.21. Non-Radioactive In Situ Hybridization for Embryonic Sections .... 99
  2.21.1. Non-Radioactive Labeling Of LNA Probe ......................... 100
  2.21.2. Day 1: Preparation and sectioning of embryos ................. 100
  2.21.3. Day 2: Preparation and hybridization of sections ............ 101
  2.21.4. Day 3: Post hybridization washes ..................................... 102
  2.21.5. Day 4: Visualization of reaction product ......................... 102
2.22. Immunofluorescence of Frozen Tissue Sections ................. 103
  2.22.1. Preparation and sectioning of embryos ......................... 103
  2.22.2. Permeabilization of Sections ............................................ 104
  2.22.3. Blocking and Incubation with Primary Antibody .......... 104
  2.22.4. Incubation with Secondary Antibody ............................ 104
  2.22.5. Incubation with DAPI and Mounting ......................... 105
2.23. Isolation of Neural Stem Cells from Mouse Brain ............. 105
  2.23.1. Dissociation of Embryonic Day 14 (E14) Mouse CNS Tissue 105
  2.23.2. Chemical Dissociation of Neurospheres ....................... 106

Chapter 3 ......................................................................................... 108
  OBJECTIVES .................................................................................. 108

Chapter 4 ......................................................................................... 109
RESULTS .................................................................................................................................109
4.1. The three miR-124 genes show overlapping but distinct expression patterns in developing mouse embryo .................................................109
4.2. Function of miR-124-1 in vivo ........................................................................................................113
  4.2.1. Generation of knockout mice .................................................................................................113
  4.2.2. Checking the reporter expression by Western Blotting ..............................................................115
  4.2.3. Identification of correctly targeted ES cell clones .......................................................................117
  4.2.4. Breeding and creation of 124-1 conditional knockout mice ......................................................117
  4.2.5. 124-1 homozygous knockout mice exhibit neonatal lethality ..........................................................119
  4.2.6. miR-124-1 expression is reduced in homozygous mutant mice brain ...........................................120
  4.2.7. miR-124-1 knockout leads to an increase in Neural Stem Cell proliferation .........................................................122
  4.2.8. Knockout NSC cultures show higher proliferation rates .............................................................123
4.3. Role of spliceosomal intron in biogenesis of mature miR-124 from miR-124-2 precursor .........................................................................................125
  4.3.1. Functional intron is essential for miR-124-2 biogenesis ..............................................................125
  4.3.2. Swapping of miR-124-2 loop with 124-1 and 124-3 loops ..............................................................129
  4.3.3. Endogenous pri-miR-124-2 sequences adjacent to the pre-miR-124-2 hairpin is essential for its splicing dependence .........................................................132
  4.3.4. Both upstream and downstream sequences are required for the pri-miR-124-2 splicing dependence .................................................................................134
  4.3.5. The composite upstream-downstream element in pri-miR-124-2 reduces the steady-state levels of unspliced pri-miR-124-2 precursor ...........................................137
  4.3.6. The pri-miR-124-2 upstream-downstream element can destabilize RNA in the absence of the pre-miR-124-2 hairpin .................................................................140
  4.3.7. pri-miR-124-2 upstream-downstream element efficiently reduces levels of splicing-deficient but not spliceable precursor transcripts ..............................................143
  4.3.8. pri-miR-124-2 upstream-downstream element retains its activity in exonic context .................................................................147
  4.3.9. The upstream-downstream element leads to faster decay of mRNA in a heterologous context .................................................................149
  4.3.10. Presence of the upstream-downstream element in precursor transcript impedes miRNA biogenesis .................................................................................150
  4.3.11. Biogenesis of mature miR-124 from the natural pri-miR-124-2 intron can be regulated by alternative splicing .................................................................154
4.3.12. Preliminary screen for trans-factors involved in miR-124 biogenesis.................................................................................................................................160

Chapter 5 ........................................................................................................... 163
  CONCLUSIONS ..............................................................................................163
Chapter 6.............................................................................................................165
  DISCUSSION ..............................................................................................165
Chapter 7 ...........................................................................................................170
  FUTURE WORK ........................................................................................170
Chapter 8 ...........................................................................................................173
  REFERENCES ..........................................................................................173
LIST OF FIGURES

Figure 1.1: Emerging classes of non-coding RNAs. A) Long non-coding RNAs. B) Small ncRNAs and IncRNAs in the mitochondrial genome. C) Some additional small ncRNAs KERV, kangaroo endogenous retrovirus; NF90, nuclear factor 90; snoRNA, small nucleolar RNA; sRNAs, small RNAs (arrows indicate direction of transcription) [8].

Figure 1.2: Diagram summarizing the four mechanistic archetypes of IncRNAs [40].

Figure 1.3: LncRNAs and their roles in a variety of biological processes [7].

Figure 1.4: Biogenesis and function of Piwi-interacting RNAs [67].

Figure 1.5: Model for RNA silencing pathways [71].

Figure 1.6: Genomic organization of miRNAs [146].

Figure 1.7: The structure of human pri-miRNAs. (A) Examples of miRNAs with their own transcription units, such as polycistronic miR-17–92-1 cluster. and miR-21. (B) miRNAs that are transcribed in conjunction with other genes [100].

Figure 1.8: Schematic depiction of the three miR-124 gene structures and localization in mouse.

Figure 1.9: microRNAs biogenesis from canonical miRNA genes [164].

Figure 1.10: Structure of human Drosha and DGCR8 [44]. P-rich = Proline rich, RIID = RNaseIII domains, dsRBD = double stranded RNA Binding Domain.

Figure 1.11: Possible mechanism of action of Drosha [84].

Figure 1.12: Domain organization of Exportin-5 [84].

Figure 1.13: Domain organization of Dicer and associated proteins.

Figure 1.14: Possible mechanism of action of Dicer [84].

Figure 1.15: Model for Argonaute action in RISC loading, unwinding and target recognition [222].

Figure 1.16: Domain structure of Argonaute 2 [220].

Figure 1.17: Schematic depiction of the siRNA-guided mRNA cleavage by Argonaute [190].

Figure 1.18: Cytoplasmic miRNA biogenesis steps [231]. a) Formation of the RISC loading complex. b) Dicer dependent (left) and Dicer independent processing of pre-miRNAs.

Figure 1.19: Domain organization of human GW182 [239].

Figure 1.20: Illustration of the canonical intron encoded biogenesis of miRNA [44].

Figure 1.21: Essential characteristics of a mirtron [244].

Figure 1.22: Schematic representation of the ‘mirtron’ miRNA biogenesis [44].

Figure 1.23: Comparison of canonical miRNA and mirtron biogenesis[245].
Figure 1.24: Regulation of miRNA processing. a) The pri-miRNA is cleaved by
the microprocessor complex Drosha–DGCR8 and pre-miRNA is released. b)
Additional specificity factors (e.g. p72 and p68) are required by some miRNAs
for efficient cleavage. c) Drosha mediated cleavage of pri-miR-18a is facilitated
by its interaction with hnRNP. d) TGF-β signalling enhances the efficient
processing of miR-21 precursor by Drosha by inducing SMAD binding to it
[231].

Figure 1.25: Regulation of Argonaute levels in the cell. a) Localization of Ago2
to P-bodies is regulated by serine phosphorylation. b) Stability of human Ago2
is affected by prolyl hydroxylation [231].

Figure 1.26: Feedback regulation loop of Drosha and DGCR8 [231].

Figure 1.27: Feedback mechanism loop between let-7 and lin-28. a) pri-let-7
processing by Drosha is inhibited by lin-28. b) pre-let-7 cleavage by Dicer is
inhibited by lin-28, which also recruits a terminal uridylyl transferase (TUTase)
to pre-let-7 [231].

Figure 1.28: Principles of microRNAs-mRNA interaction [281].

Figure 1.29: Proposed Mechanism of Translational Repression by Ago2.
Binding of the Ago2 MC domain to the mRNA cap (m7G) is mediated by
miRNA-guided deposition of Ago2 to an mRNA target and thus exclusion of
eIF4E and inhibition of translation initiation [283].

Figure 1.30: Mechanism of inhibition of ribosomal subunit joining [289].

Figure 1.31: Mechanisms of post initiation inhibition by miRNAs. A) Inhibition
of translation elongation B) Co-translational protein degradation [289].

Figure 1.32: Mechanisms of post initiation inhibition by miRNAs by ribosome
drop-off [289, 290].

Figure 1.33: miRNA mediated deadenylation of mRNA and decay [239].

Figure 1.34: Parts of the Central Nervous System.

Figure 1.35: Relative spatial and temporal range of adaptive response and
function for different mechanisms that are involved in the progression from
mRNA production to final protein function [315].

Figure 1.36: miRNAs involved in various aspects of synaptic development and
function.

Figure 1.37: Cross talk between SCP1, PTBP1 and miR-124 in non-neuronal
cells [352].

Figure 1.38: Cross talk between SCP1, PTBP1 and miR-124 in neuronal cells
[352].

Figure 1.39: Schematic summary showing modulation of multiple pathways
involved in neurogenesis by miR-124 through translational regulation of its
currently known mRNA targets [316].

Figure 1.40: Illustration showing the DSBR and SDSA pathways [368].

Figure 1.41: Enrichment showing the DSBR and SDSA pathways [368].
Figure 1.42: Overview of bacteriophage λ recombination system used for recombineering [389].

Figure 1.43: Sequence of Lox P.

Figure 1.44: Outcomes of a Cre-lox recombination (The Jackson laboratory).

Figure 1.45: Mechanism of creating knockout mice using Cre-Lox Recombination system.

Figure 1.46: Diagram depicting reactions catalyzed by FLP upon interaction with FRT sites [401].

Figure 4.1: Whole mount in situ analyses of mouse embryos (E8.5-E10.5) using probes against individual pri-miR-124 precursors.

Figure 4.2: RT-PCR analysis of RNA samples from mouse embryonic and adult brain and liver revealing spatio-temporal expression patterns of the three miR-124 genes. Asterisks indicate the positions of correct bands.

Figure 4.3: In situ staining results with mice embryo sections. (A) Thoracic spinal cord segments (E10.5) showing CNS expression of all three pri-miRNAs and exclusive PNS (dorsal root ganglia, DRG) expression of pri-miR-124-1. (B) & (C) Expression of the three miR-124 genes in transverse sections of mice embryo showing differential spatial expression pattern of 124 genes.

Figure 4.4: Illustration of the constructs used to target the three mouse miR-124 genes.

Figure 4.5: Western Blot showing the expression of tags corresponding to each miR-124 gene.

Figure 4.6: Southern blot analyses of genomic DNA samples from selected ES clones. DNAs were cut with indicated restriction enzymes and analyzed using probes specific to the corresponding 3’ homology arms. Wt stands for genomic DNA from the wild type ES cells and the alphanumeric characters are clone numbers of targeted ES cells. For all three miR-124 genes, the longer fragments are derived from the wild-type miR-124 alleles, whereas the shorter fragments correspond to correctly targeted miR-124 alleles. Restriction enzymes used to fragment the genomic DNA before Southern blot analysis are indicated at the bottom.

Figure 4.7: Diagram of the wild-type, conditional (124-1tm1.1) and null (124-1tm1.2) 124 alleles. Note that the conditional allele expresses a “substitute pre-miR-124” from an intronically encoded fragment (~500 bp) of the 124-1 gene. Positions of the recombinant intron (rIntron), the two LoxP sites, and the IRES-EYFP-3xHA cassette are indicated.

Figure 4.8: miR-124 expression seen by LNA staining in E14.5 mice brain sections. The panels on the left correspond to homozygous knockout samples and the panel on the right corresponds to wild type mice brain samples. Sections in each row belong to the same or similar region of the brain to facilitate comparison of miR-124 expression levels.

Figure 4.9: Immunofluorescence analysis showing over-proliferation (arrowheads) of the NSC (neuroepithelial) layer in 124-1-null embryos.
Figure 4.10: Immunofluorescence analysis showing over-proliferation (arrowheads) of the NSC layer in P2 mice brain samples. .........................123
Figure 4.11: Immunofluorescence analysis of adherent NSC cultures obtained from the wild-type, heterozygous (124-1\(^{\text{WT}}/124-1^{\text{tm1.2}}\)) and knockout (124-1^{\text{tm1.2}}/124-1^{\text{tm1.2}}) female E14.5 embryonic cortices. Note that most cells in all three cultures express NSC marker nestin. .............................................123
Figure 4.12: PrestoBlue assay showing significantly faster proliferation of 124-1-null (KO; 124-1^{tm1.2}/124-1^{tm1.2}) NSCs as compared to NSC controls with other genotypes. .............................................................................124
Figure 4.13: Constructs used to test the role of functional intron in miR-124 biogenesis........................................................................125
Figure 4.14: A) Northern blot analysis of splicing dependence of 124 genes in N2a cells. B) Expression levels of the samples used for the Northern indicating the significance of their differences. ...........................................127
Figure 4.15: Splicing Dependence of the three miR-124 genes in N2a cells. ANOVA test was carried out to examine the significance of the difference between the three precursors ........................................................................128
Figure 4.16: Northern blot analysis showing dependence of pri-miR-124-2 but not pri-miR-124-1 and pri-miR-124-3 on a functional intron. RNA samples were analyzed 72 hours post transfection. ........................................128
Figure 4.17: Splicing Dependence of the three miR-124 genes. ANOVA test was carried out to examine the significance of the difference between the three precursors ........................................................................129
Figure 4.18: Diagram showing the stem loop structures of the three miR-124 pri-miRNA sequences. .................................................................................................................................130
Figure 4.19: Illustration of the constructs containing the loop mutants. ....130
Figure 4.20: A) Northern Blot of 124-2 and its loop mutants. B) Splicing dependence of the constructs used. .................................................................131
Figure 4.21: Illustrations for the wild type and mutated mini-miR-124 genes. .................................................................................................................................132
Figure 4.22: A) Northern blot analysis of minimized miR-124 constructs. B) Splicing dependence of the miR-124-2-specific constructs quantified from five transfection experiments. ........................................................133
Figure 4.23: Constructs encoding wild-type pri-miR-124-2 fragment in a spliceable context and its 5' and 3' truncated mutants in spliceable and splicing-deficient context .....................................................................134
Figure 4.24: A) Northern blot analysis of the 124-2 precursor truncation mutants. B) Splicing dependence of 124-2 compared with the end mutants...136
Figure 4.25: Sequence of the upstream-downstream destabilization sequence. The sequence upstream to the 124-2 stem-loop is indicated in red while the sequence downstream is indicated in green. The pre-miR-124-2 sequence is present between these two and indicated in black. ........................................................................137
Figure 4.26: Multiplex PCR primers used in this study. The arrows indicate the primer binding sites. .................................................................138
Figure 4.42: Northern blot analysis of mature miR-153 expression in the absence and presence of upstream-downstream element. The samples analyzed were harvested 48 hours post transfection. .......................................................... 152

Figure 4.43: Quantification of the mature miR-153 expression levels in Figure 4.41. .................................................................................................................. 153

Figure 4.44: Endogenous miR-124-2 status in mice brain. (A) Transcription structure of the mouse 124-2 gene. (B) RT-PCR analysis showing progressive decline in the relative expression of the embryonic (bottom band, 218 bp) and accumulation of the adult (top band, 470 bp) pri-miR-124-2 splice form during mouse brain development. * around 400bp region indicates the presence of a non-specific amplicon. (C) Quantification of the results in (B) averaged from 4 independent RT-PCR analyses ±standard deviation...............................154

Figure 4.45: Constructs encoding the pri-miR-124-2 natural intron 2 and its site-specific mutants lacking indicated splice sites.............................................. 156

Figure 4.46: RT-multiplex PCR analysis of splicing patterns of indicated recombinant pri-miR-124-2 transcripts in N2a cells. This assay resolves transcripts spliced in the “adult” (470 bp) and “embryonic” (218 bp) manner as well as unspliced RNA (159 bp). Note that the 124-2 and SDI-124-2 controls in lanes 1 and 2 lacked natural E5’ss and A5’ss-associated sequences and therefore could only give rise to unspliced RT-PCR products. ......................... 157

Figure 4.47: A) Northern blot analysis of the effect of splicing on miR-124 biogenesis from its “natural” intronic sequence in N2a cells. (B) Quantification of the results in (A) showing relative miR-124 expression levels. ................. 158

Figure 4.48: (A) Northern blot analysis of the effect of splicing on miR-124 biogenesis from its “natural” intronic sequence in HEK293T cells. (B) Quantification of the results in (A) showing relative miR-124 expression levels. ......................................................................................................................... 159

Figure 4.49: Microprocessor component RNAi efficiency examined by RT-qPCR using corresponding primers. The samples were analyzed 72 hours post transfection.......................................................... 161

Figure 4.50: Effect of knocking down specific Microprocessor components on the pri-miR-124-2 splicing dependence.......................................................... 162
LIST OF TABLES

Table 1.1: Diversity and functions of various classes of non-coding RNAs. ...... 5
Table 2.1: Name and description of plasmids used for Experiments. .......... 77
Table 2.2: Primers used for cloning. ......................................................... 79
Table 2.3: Primers used for qPCR. ........................................................... 88
Table 2.4: Multiplex PCR Primers. ......................................................... 89
Table 2.5: Reaction mixture for labeling oligonucleotide probe. .......... 100
Table 2.6: The components of the Hybridization Solution. ..................... 101
Table 2.7: Composition of NTMT Buffer. .............................................. 103
Table 4.1: Number of mice corresponding to different 124-1\textsuperscript{tm1.1} genotypes and sex on breeding. ................................................................. 119
Table 4.2: Number of mice corresponding to different 124-1\textsuperscript{tm1.2} genotypes and sex on breeding. ................................................................. 120
LIST OF ABBREVIATIONS

µg : Microgram
µl : Microliter
Ago : Argonaute
Amp : Ampicillin
APP : Amyloid Precursor Protein
APS : Ammonium persulfate
ATP : Adenosine-5’-triphosphate
BCA : Bicinchoninic acid
bp : base pairs
BSA : Bovine Serum Albumin
CaCl₂ : Calcium chloride
casiRNA : cis-acting siRNA
cDNA : Complementary DNA
Ci : Curie units
cm : Centimetre
CMV : Cytomegalovirus
Co-IP : Co-Immunoprecipitation
CNS : Central Nervous System
Cre : Cyclization Recombinase
CREB : cAMP response element-binding
DAPI : 4’, 6-diamidino-2-phenylindole
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEPC</td>
<td>Diethyl pyrocarbonate</td>
</tr>
<tr>
<td>DGCR8</td>
<td>DiGeorge syndrome Critical Region gene 8</td>
</tr>
<tr>
<td>DMEM</td>
<td>Dulbecco's Modified Eagle Medium</td>
</tr>
<tr>
<td>DNase</td>
<td>Deoxyribonuclease</td>
</tr>
<tr>
<td>dNTP</td>
<td>Deoxyribonucleotide triphosphate</td>
</tr>
<tr>
<td>Dox</td>
<td>Doxycycline</td>
</tr>
<tr>
<td>dsRBD</td>
<td>double-stranded RNA-binding domain</td>
</tr>
<tr>
<td>dsRNA</td>
<td>Double stranded RNA</td>
</tr>
<tr>
<td>DTT</td>
<td>Dithiothreitol</td>
</tr>
<tr>
<td>E12.5</td>
<td>12.5 days post-conception embryos</td>
</tr>
<tr>
<td>EAE</td>
<td>Experimental Autoimmune Encephalomyelitis</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>EGF</td>
<td>Epidermal Growth Factor</td>
</tr>
<tr>
<td>EGFP</td>
<td>Enhanced Green Fluorescent Protein</td>
</tr>
<tr>
<td>FBS</td>
<td>Fetal Bovine Serum</td>
</tr>
<tr>
<td>FITC</td>
<td>Fluorescein IsoThioCyanate</td>
</tr>
<tr>
<td>Fgf</td>
<td>Fibroblast growth factor</td>
</tr>
<tr>
<td>fmol</td>
<td>Femto Molar</td>
</tr>
<tr>
<td>FRT</td>
<td>Flippase Recognition Target</td>
</tr>
<tr>
<td>GAPDH</td>
<td>Glyceraldehyde 3-phosphate dehydrogenase</td>
</tr>
<tr>
<td>HCl</td>
<td>Hydrogen Chloride</td>
</tr>
<tr>
<td>HDAC</td>
<td>Histone Deactylase</td>
</tr>
<tr>
<td>HEK293T</td>
<td>Human Embryonic Kidney 293T</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>HEPES</td>
<td>4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid</td>
</tr>
<tr>
<td>HILO RMCE</td>
<td>High-efficiency low-background recombination-mediated cassette exchange</td>
</tr>
<tr>
<td>hnRNP K</td>
<td>heterogenous nuclear ribonucleoprotein K</td>
</tr>
<tr>
<td>HPRT</td>
<td>Hypoxanthine-guanine phosphoribosyltransferase</td>
</tr>
<tr>
<td>HRP</td>
<td>Horseradish peroxidase</td>
</tr>
<tr>
<td>hr</td>
<td>Hour</td>
</tr>
<tr>
<td>IRES</td>
<td>Internal Ribosomal Entry Site</td>
</tr>
<tr>
<td>kb</td>
<td>kilobases</td>
</tr>
<tr>
<td>KCl</td>
<td>Potassium chloride</td>
</tr>
<tr>
<td>kDa</td>
<td>kiloDalton</td>
</tr>
<tr>
<td>LB</td>
<td>Luria-Bertani</td>
</tr>
<tr>
<td>lincRNA</td>
<td>Large intergenic ncRNAs</td>
</tr>
<tr>
<td>lncRNA</td>
<td>Long non-coding RNA</td>
</tr>
<tr>
<td>LoxP</td>
<td>locus of X over P1</td>
</tr>
<tr>
<td>lsiRNA</td>
<td>long siRNA</td>
</tr>
<tr>
<td>MgCl₂</td>
<td>Magnesium chloride</td>
</tr>
<tr>
<td>min</td>
<td>Minutes</td>
</tr>
<tr>
<td>miRNA</td>
<td>micro Ribo Nucleic Acid</td>
</tr>
<tr>
<td>miRNP</td>
<td>micro RNA Ribonucleo Protein complex</td>
</tr>
<tr>
<td>mM</td>
<td>Milli Molar</td>
</tr>
<tr>
<td>mRNA</td>
<td>Messenger Ribo Nucleic Acid</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>mtDNA</td>
<td>mitochondrial DNA</td>
</tr>
<tr>
<td>NaCl</td>
<td>Sodium chloride</td>
</tr>
<tr>
<td>natsiRNA</td>
<td>natural antisense transcript-derived siRNA</td>
</tr>
<tr>
<td>ncRNA</td>
<td>Noncoding RNA</td>
</tr>
<tr>
<td>ng</td>
<td>Nanogram</td>
</tr>
<tr>
<td>NS</td>
<td>Nervous System</td>
</tr>
<tr>
<td>NSC</td>
<td>Neural Stem Cells</td>
</tr>
<tr>
<td>nt</td>
<td>Nucleotide</td>
</tr>
<tr>
<td>NEB</td>
<td>New England Biolabs</td>
</tr>
<tr>
<td>°C</td>
<td>Degree centigrade</td>
</tr>
<tr>
<td>OCT</td>
<td>Optimum Cutting Temperature</td>
</tr>
<tr>
<td>PAA</td>
<td>Polyacrylamide</td>
</tr>
<tr>
<td>PABP</td>
<td>Poly(A)-binding protein</td>
</tr>
<tr>
<td>PAGE</td>
<td>Polyacrylamide Gel Electrophoresis</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate-Buffered Saline</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase Chain Reaction</td>
</tr>
<tr>
<td>PFA</td>
<td>Paraformaldehyde</td>
</tr>
<tr>
<td>PGK</td>
<td>Phospho-Glycerate Kinase</td>
</tr>
<tr>
<td>piRNA</td>
<td>Piwi interacting RNA</td>
</tr>
<tr>
<td>PKR</td>
<td>protein kinase R</td>
</tr>
<tr>
<td>pmol</td>
<td>Pico Molar</td>
</tr>
<tr>
<td>PMSF</td>
<td>Phenylmethanesulfonylfluoride</td>
</tr>
<tr>
<td>PNK</td>
<td>Polynucleotide Kinase</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>PNS</td>
<td>Peripheral Nervous System</td>
</tr>
<tr>
<td>PRC</td>
<td>Polycomb Repressive Complex</td>
</tr>
<tr>
<td>pre-miRNA</td>
<td>Precursor micro RNA</td>
</tr>
<tr>
<td>pri-miRNA</td>
<td>Primary micro RNA</td>
</tr>
<tr>
<td>PTBP</td>
<td>Polypurimidine Tract-Binding Protein</td>
</tr>
<tr>
<td>REST</td>
<td>RE1-Silencing Transcription repressor</td>
</tr>
<tr>
<td>RISC</td>
<td>RNA-Induced Silencing Complex</td>
</tr>
<tr>
<td>RNA Pol</td>
<td>RNA Polymerase</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic Acid</td>
</tr>
<tr>
<td>RNase</td>
<td>Ribonuclease</td>
</tr>
<tr>
<td>RNasin</td>
<td>RNase Inhibitor</td>
</tr>
<tr>
<td>RNP</td>
<td>Ribonucleoprotein</td>
</tr>
<tr>
<td>rpm</td>
<td>Revolutions per minute</td>
</tr>
<tr>
<td>rRNA</td>
<td>Ribosomal RNA</td>
</tr>
<tr>
<td>RT qPCR</td>
<td>Reverse Transcription and quantitative Polymerase Chain Reaction</td>
</tr>
<tr>
<td>RT</td>
<td>Reverse transcription</td>
</tr>
<tr>
<td>RT-PCR</td>
<td>Reverse Transcription Polymerase Chain Reaction</td>
</tr>
<tr>
<td>SB</td>
<td>Sample Buffer</td>
</tr>
<tr>
<td>scnRNA</td>
<td>Small scanRNA</td>
</tr>
<tr>
<td>SCP</td>
<td>Small C-terminal domain Phosphatase 1</td>
</tr>
<tr>
<td>SDN1</td>
<td>Small RNA Degrading Nuclease 1</td>
</tr>
<tr>
<td>SDS</td>
<td>Sodium Dodecyl Sulfate</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>SDSA</td>
<td>Synthesis Dependent Strand Annealing</td>
</tr>
<tr>
<td>shRNA</td>
<td>Small hairpin RNA</td>
</tr>
<tr>
<td>siRNA</td>
<td>Small interfering RNA</td>
</tr>
<tr>
<td>snRNA</td>
<td>Short noncoding RNA</td>
</tr>
<tr>
<td>snoRNA</td>
<td>Small nucleolar RNA</td>
</tr>
<tr>
<td>snoRNA</td>
<td>Small nucleolar RNA</td>
</tr>
<tr>
<td>SSC</td>
<td>Saline Sodium Citrate</td>
</tr>
<tr>
<td>ssRNA</td>
<td>Single stranded RNA</td>
</tr>
<tr>
<td>SVZ</td>
<td>Subventricular Zone</td>
</tr>
<tr>
<td>tasiRNA</td>
<td>trans-acting siRNA</td>
</tr>
<tr>
<td>TBE</td>
<td>Tris/Borate/EDTA</td>
</tr>
<tr>
<td>TBST</td>
<td>Tris Buffered Saline with Tween</td>
</tr>
<tr>
<td>TE</td>
<td>Tris EDTA</td>
</tr>
<tr>
<td>TEMED</td>
<td>Tetramethylethylenediamine</td>
</tr>
<tr>
<td>Temp</td>
<td>Temperature</td>
</tr>
<tr>
<td>TNRC</td>
<td>Tri-Nucleotide Repeat Containing</td>
</tr>
<tr>
<td>TRBP</td>
<td>TAR RNA Binding Protein</td>
</tr>
<tr>
<td>tRNA</td>
<td>Transfer RNA</td>
</tr>
<tr>
<td>UBA</td>
<td>Ubiquitin-associated</td>
</tr>
<tr>
<td>UTR</td>
<td>Untranslated Region</td>
</tr>
<tr>
<td>Vegf</td>
<td>Vascular endothelial growth factor</td>
</tr>
</tbody>
</table>
ABSTRACT

MicroRNAs are 19-25 nucleotides long non-coding RNAs that base-pair with cognate mRNA targets and regulate their translation and/or stability. Different miRNAs have been implicated in the regulation of essential biological processes like differentiation, proliferation and apoptosis. miR-124 is a conserved, abundantly expressed neuron-specific microRNA that has been previously shown to contribute to neurogenesis by targeting several important transcripts. miR-124 is expressed in neurons, but not glial cells, and the levels of miR-124 increase over time in the developing nervous system (NS).

Similar to a few other vertebrate microRNAs, mature miR-124 is encoded by three distinct non-allelic genes in human and mouse genome. To understand biological significance of this genetic redundancy, we examined spatio-temporal expression patterns of the three genes in developing mouse embryo. Two miR-124 genes were expressed at detectable levels in both neural stem cells (NSCs) and post-mitotic neurons, whereas the third gene was expressed exclusively in neurons. Notably, knocking out the neuron-specific miR-124 gene in mouse led, in addition to other interesting phenotypes, to over-proliferation of the neuroepithelial layer containing embryonic NSCs which are positive for certain proliferation markers. These data suggest that one of the functions of the multiple miR-124 genes in mouse could be limiting the NSC proliferation potential and ensuring temporal precision of neuronal differentiation.

In a separate line of experiments, we have investigated post-transcriptional mechanisms underlying miR-124 biogenesis. Interestingly, processing of miR-124 from the pri-miR-124-2 precursor requires the presence of a functional intron, consistent with the native genetic structure of this precursor. We show that the unusually strong dependence of 124-2 on the splicing reaction is due to the presence of a tunable cis-regulatory element that induces rapid degradation of the pri-miR-124-2 precursor unless this element is positioned within a spliceable intron. These data uncover a novel post-
transcriptional mechanism regulating miRNA expression which could also possibly modulate the mature miRNA output of the 124-2 gene in vivo.
Chapter 1

INTRODUCTION

Living systems depend on carefully orchestrated gene expression to regulate diverse aspects of their development and physiology. Cells have evolved elaborate mechanisms to generate and utilize transcripts with optimal efficiency, which ensures precise implementation of proliferation, differentiation and programmed cell death programs as well as an adequate response to multiple environmental signals. Since Crick’s proposition of the central dogma of gene expression in the late 1950’s, there has been an explosion in our awareness of the steps regulating the flow of information from DNA to RNA to protein. We now know that the pathway leading from gene activation in the nucleus to mRNA translation and decay at specific cytoplasmic locations is highly regulated and shows an extensive degree of cross-talk between individual steps. Gene expression is tightly regulated to ensure that the spatial and temporal expression pattern of proteins and RNA is maintained in the cell. Considering the diverse repertoire of RNA sequence and structure, it plays crucial roles in biological systems and is regulated by numerous proteins modulating its chemical composition and overall expression patterns. Recent methodological advances, in bioinformatics, genome and transcriptome sequencing, microarray-based analyses and biochemistry are generating new insights into cellular and organismal functions of various mRNA variants and non-coding RNAs, as well as uncovering possible roles of these molecules in evolution.
1.1. Non-Coding RNAs

Noncoding RNAs (ncRNAs) are RNAs that are not translated into protein. ncRNAs include transcripts of different lengths, structures, and functions [1]. The DNA sequence from which an ncRNA is transcribed is often called an RNA gene. Highly abundant and functionally important RNAs are also encoded by these RNA genes. They include transfer RNA (tRNA), ribosomal RNA (rRNA), nuclear snRNAs and snoRNAs. More recently identified groups include short ncRNAs, such as microRNAs (miRNAs), siRNAs, and piRNAs and long ncRNAs (lncRNAs) (Figure 1.1). Recent transcriptomic and bioinformatic studies suggest the existence of thousands of distinct ncRNAs that have been revealed even though the exact number of RNA genes still remains unknown [2-6]. The last decade has unveiled the importance of ncRNAs in gene regulation. However, a lot of work still needs to be done to elucidate the identity of certain functional elements in the primary sequence of non-coding genes that define their roles as RNA molecules [7].

Although ncRNAs still remain a relatively underexplored class, considerable progress has been made over the last decade in understanding molecular mechanisms underlying biogenesis and functions of several ncRNA categories [7]. One important lesson learned from this work has been the realization that the expression of many ncRNAs is regulated by diverse epigenetic, transcriptional and post-transcriptional mechanisms that affect biogenesis of primary transcript, post-transcriptional processing and maturation steps, as well as, changes in its conformation and trafficking of ncRNAs within and between the cell and its milieu. As a result, the non-coding transcriptome is now bestowed with an extraordinary diversity in function and response to environmental changes [8].
Examples of epigenetic and transcriptional regulators of the ncRNA expression include tumor protein 53 (TP53), repressor element 1-silencing transcription factor (REST), zinc finger protein 143 (ZNF143) and cAMP response element-binding (CREB) [9-12]. ncRNAs are transcribed by specific RNA polymerase enzymes [13]. RNA Polymerase II (RNAP II) is responsible for synthesizing many types of non-coding RNAs (ncRNAs) including snRNAs...
and miRNAs. However, a range of essential ncRNAs are generated by Pol III, which curiously includes an RNA regulator of Pol II activity called 7SK [13].

Various classes of ncRNAs can also be subject to extensive post-transcriptional regulation [14, 15] (and see Table 1.1 below). Furthermore, a variety of post-transcriptional processes (such as non-templated modifications, alternative splicing, RNA editing, polyadenylation and capping) have the ability to modify ncRNAs [16-18]. Emerging evidence suggests that these and other post-transcriptional modifications may result in important functional consequences, particularly in the nervous system. For example, adenosine-to-inosine RNA editing can modify miRNA sequence alterations in the interactions between a miRNA and its target mRNA. This modification also occurs in ncRNAs derived from retrotransposons (LINE-1 (L1) and Alu sequences), which form the largest fraction of substrates for primate RNA editing [17]. Intriguingly, these sequences have undergone considerable expansion in primates with the utmost levels of RNA editing occurring in the human brain [19]. Therefore, it has been proposed that RNA editing of retrotransposons may play critical roles in brain functioning and evolution. Moreover, ncRNAs can undergo nuclear-cytoplasmic, nuclear-mitochondrial and axodendritic trafficking facilitated by ribonucleoprotein complexes. This process can be a critical factor for ensuring correct spatio-temporal distribution and ultimately function of mRNAs, ncRNAs and RNA-binding proteins [20]. Additionally, by means of exosomal release and various other transport mechanisms, ncRNAs may play a role in intercellular communication between nearby nerve cells, somatic sites which are at a distance and the germline [21, 22]. These observations imply that understanding of nervous system development and function will require a thorough understanding of ncRNAs life cycles.
### Table 1.1: Diversity and functions of various classes of non-coding RNAs.

<table>
<thead>
<tr>
<th>Class</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Antisense IncRNAs</td>
<td>IncRNAs derived from antisense transcription, present at ~70% of mammalian genomic loci and implicated in the regulation of sense protein-coding genes</td>
</tr>
<tr>
<td>Antisense termini-associated short RNAs</td>
<td>Small ncRNAs with 5’ poly-U tails originating from 3’ termini of genes in antisense orientation, suggesting transcription by RNA-dependent RNA polymerase</td>
</tr>
<tr>
<td>Centromere repeat-associated small interacting RNAs</td>
<td>Small ncRNAs (34–42 nucleotides) derived from centromeric repeats with putative roles in local epigenetic modifications and heterochromatin formation</td>
</tr>
<tr>
<td>Endogenous small interfering RNAs</td>
<td>Small DICER-dependent ncRNAs (21–26 nucleotides) associated with Argonaute proteins (AGO2) and involved in post-transcriptional and epigenetic silencing of protein-coding genes and transposons</td>
</tr>
<tr>
<td>Enhancer RNAs</td>
<td>IncRNAs transcribed from enhancer domains, and expressed coordinately with activity-dependent neuronal genes</td>
</tr>
<tr>
<td>Enhancer-like long IncRNAs</td>
<td>IncRNAs exhibiting enhancer activity, particularly for genes regulating development and differentiation</td>
</tr>
<tr>
<td>Large intergenic RNAs</td>
<td>IncRNAs derived from intergenic regions that act as guides for recruiting PRC2, CoREST and other chromatin-modifying complexes</td>
</tr>
<tr>
<td>IncRNAs</td>
<td>Large family of ncRNAs (~200 nucleotides) with diverse functional roles</td>
</tr>
<tr>
<td>MicroRNA-offset RNAs</td>
<td>Small ncRNAs (~20 nucleotides) produced from microRNA precursors and exhibiting independent expression relative to associated microRNAs, implying distinct functional roles</td>
</tr>
<tr>
<td>MicroRNAs</td>
<td>Small DICER-dependent ncRNAs (20–23 nucleotides) associated with Argonaute proteins (AGO1–4) and involved in post-transcriptional silencing of protein-coding genes and ncRNAs</td>
</tr>
<tr>
<td>Mitochondrial ncRNAs</td>
<td>Small ncRNAs and IncRNAs generated from both strands of the mitochondrial genome, regulated by nuclear-encoded mitochondrial proteins and expressed in cell type- and tissue-specific patterns</td>
</tr>
<tr>
<td>Class</td>
<td>Description</td>
</tr>
<tr>
<td>------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>PIWI-interacting RNAs</td>
<td>Small ncRNAs (26–30 nucleotides) associated with the PIWI subclass of Argonaute proteins and involved in silencing of mRNAs and transposons</td>
</tr>
<tr>
<td>Promoter-associated long RNAs</td>
<td>IncRNAs transcribed from promoter domains of protein-coding genes, particularly cell cycle modulators, and capable of recruiting regulatory factors</td>
</tr>
<tr>
<td>Promoter-associated small RNAs</td>
<td>Small ncRNAs (20–200 nucleotides) possessing 5’ ends that coincide with the transcription start sites of protein-coding genes and ncRNAs</td>
</tr>
<tr>
<td>Small nucleolar RNAs</td>
<td>Small ncRNAs derived from intronic regions with roles in promoting RNA modifications, including pseudouridylation (H/ACA snoRNAs) and methylation (C/D snoRNAs), as well as pre-mRNA processing</td>
</tr>
<tr>
<td>Small RNAs derived from small nucleolar RNAs</td>
<td>Small ncRNAs derived from 3’ ends of H/ACA snoRNAs (20–24 nucleotides) and 5’ ends of C/D snoRNAs (17–19 nucleotides and &gt;27 nucleotides) with microRNA-like functions</td>
</tr>
<tr>
<td>Small nuclear factor 90-associated RNAs</td>
<td>Small ncRNAs (117 nucleotides) associated with the nuclear factor 90 RBP and exhibiting accelerated evolution and expansion in hominids and region-specific expression in human brain</td>
</tr>
<tr>
<td>Splice-site RNAs</td>
<td>Small ncRNAs (17–18 nucleotides) derived from splice sites of highly transcribed genes, and expressed in developmental stage- and region-specific patterns</td>
</tr>
<tr>
<td>Telomere small RNAs</td>
<td>Small DICER-independent telomere-specific ncRNAs (~24 nucleotides)</td>
</tr>
<tr>
<td>Telomere repeat-containing RNAs</td>
<td>IncRNAs transcribed from telomeric repeats and dynamically regulated during the cell cycle with roles in heterochromatin formation and telomere functioning</td>
</tr>
<tr>
<td>Termini-associated short RNAs</td>
<td>Small ncRNAs originating from 3’ termini of protein-coding genes and ncRNAs</td>
</tr>
<tr>
<td>Transcription initiation RNAs</td>
<td>Small ncRNAs (18 nucleotides) originating downstream of transcription start sites of protein-coding genes, implicated in modulating CTCF localization and nucleosome density</td>
</tr>
<tr>
<td>Transcription start site associated RNAs</td>
<td>Small ncRNAs (20–90 nucleotides) originating from the -250 to +50 position relative to the transcription start site of protein-coding genes</td>
</tr>
<tr>
<td>tRNA-derived RNA fragments</td>
<td>Small ncRNAs (19–40 nucleotides) derived from cleavage of mature tRNAs constitutively by DICER and in response to stress by angiogenin (in humans), promoting stress granule assembly and inhibiting miRNA translation</td>
</tr>
<tr>
<td>3’-untranslated region-associated ncRNAs</td>
<td>IncRNAs derived from 3’-untranslated regions of protein-coding transcripts and exhibiting independent developmental stage- and tissue-specific expression profiles</td>
</tr>
<tr>
<td>Vault RNAs</td>
<td>Small ncRNAs (88–100 nucleotides) integral to the vault RNP complex with putative roles in multidrug resistance, apoptosis resistance and innate immunity</td>
</tr>
<tr>
<td>Y RNAs</td>
<td>Small ncRNAs (~100 nucleotides) required for DNA replication, cleaved into microRNAs and implicated in Ro RBP localization and function</td>
</tr>
</tbody>
</table>
1.2. Long non-coding RNAs

The majority of ncRNAs belong to the group of so-called long non-coding RNAs (lncRNA) that provisionally combines a wide range of non-protein coding transcripts with lengths of at least 200 nt [23]. LncRNAs have recently emerged as regulators of several important cellular processes [7, 24]. Although clearly a heterogeneous group, most lncRNAs share some common traits including nuclear localization, low expression levels, low level of interspecies sequence conservation and their occurrence in both poly A+ and poly A- forms [16, 23, 25, 26]. They were first described during the full-length sequencing of cDNA libraries in mouse [27]. LncRNAs can be transcribed from intergenic regions (large intergenic ncRNAs, or lincRNAs) [28, 29] in sense, antisense, bidirectional and overlapping orientations relative to the adjacent protein-coding genes. Other lncRNAs can be produced from the genomic sequences additionally encoding regulatory regions of protein-coding genes [e.g. introns [30] and UTRs [30], promoters [31] and enhancers [32]] as well as from specialized chromosomal regions [e.g., telomeres [33]]. Several lncRNAs have been shown to be derived from the mitochondrial genome [34].

As mentioned above, lncRNAs can be subject to post-transcriptional processing, including capping, polyadenylation, alternative splicing, RNA editing, as well as regulated intracellular trafficking [35, 36]. Some lncRNAs can be involved in the regulation of the genomic loci from which they originate or of their neighboring ones. For example, a subset of lincRNAs have been shown to have enhancer-like activity [32]. However, in comparison other lncRNAs, such as HOX transcript antisense RNA (HOTAIR) [37] and X inactive-specific transcript (XIST) [38], stimulate the formation of repressive chromatin environments across large regions in the genome and even entire chromosomes, respectively. Many lncRNAs show spatio-temporal patterns of expression, indicating that lncRNA expression is often tightly regulated [30, 39]. This might be beneficial given the broad range of lncRNA functions. Besides already mentioned ones, these include lncRNA roles in transcriptional and epigenetic control via transcription factor recruitment and chromatin-
modifying complexes to specific sites in the nucleus and genome; alternative 
splicing and other post-transcriptional RNA modifications by assembling 
nuclear domains containing factors involved in RNA-processing; translational 
control and nuclear-cytoplasmic shuttling [40]. IncRNAs can also act as 
precursors for small ncRNAs, such as small nucleolar RNAs (snoRNAs) and 
miRNAs. Depending on their molecular functions IncRNAs are classified into 4 
archetypes (Figure 1.2):

*Archetype I:* as signals, IncRNA expression acts as a faithful indicator of 
the combinatorial actions of signaling pathways or transcription factors 
(colored ovals) reflecting gene regulation in time and space.

*Archetype II:* as decoys, titration of transcription factors and other 
proteins from chromatin or titration of protein factors into nuclear subdomains 
can be carried out by IncRNAs.

*Archetype III:* as guides, recruitment of chromatin-modifying enzymes 
to target genes that are located either near the site of IncRNA production (i.e., 
in *cis*), or to distant target genes (i.e, in *trans*) is also done by IncRNAs.

*Archetype IV:* as scaffolds, IncRNAs can form ribonucleoprotein 
complexes by bringing together multiple proteins. Histone modifications can 
be affected upon by the IncRNA-RNP by their action on chromatin as 
illustrated. In other instances, the structural properties of IncRNA scaffold 
leads to stabilization of nuclear structures or signaling complexes.
Several lncRNAs have been implicated in human diseases include ANRIL and HOTAIR that bind to chromatin-remodeling complexes PRC1 (Polycomb Repressive Complex1) and PRC2 altering chromatin structure and transcriptional regulation (Figure 1.3). GAS5 lncRNA acts as a decoy for the transcription factor called GR, thus preventing GR from binding to DNA and activating transcription. MALAT1 RNA sequesters SR proteins regulating
mRNA alternative splicing, while BACE-1AS RNA binds to the complementary BACE-1 mRNA and regulates its translation.

![Diagram](image)

**Figure 1.3:** LncRNAs and their roles in a variety of biological processes [7].

### 1.3. Short non-coding RNAs

Non-coding RNAs that are shorter than 200 nt can be collectively referred to as short noncoding RNAs (sncRNAs). These include small nuclear (sn) RNAs, 5S and 5.8S ribosomal RNAs (rRNAs), small nucleolar (sno) RNAs and transfer RNAs (tRNAs) [41]. More recently, a plethora of additional sncRNAs have been identified in eukaryotes including miRNAs, siRNAs, piRNAs, etc [42-45].
Despite their small size, they have emerged as key players in regulating diverse cellular activities. Some of the known functions of sncRNAs include mRNA degradation, gene silencing, translational repression (Fig 1.5), post transcriptional modifications, genome stability and modulation of protein functions [41, 46-49]. Based on their mechanism of biogenesis and protein factors that they associate with, the sncRNAs are classified into:

- Piwi interacting RNAs (piRNAs)
- Small interfering RNAs (siRNAs)
- MicroRNAs (miRNAs)

Another method of classification involves categorizing the sncRNAs based on their Dicer (RNase III family endoribonuclease specific for dsRNA) dependence during biogenesis into Dicer-dependent (miRNAs and siRNAs) and Dicer-independent group (piRNAs).

### 1.4. Piwi interacting RNAs (piRNAs)

piRNAs are Dicer-independent 26-30 nucleotide long ssRNAs (single stranded RNAs), first discovered in the germline of fruit fly *D. melanogaster* using a genetic approach [50, 51]. To date, piRNAs have been additionally identified in vertebrates, nematodes, cnidarians, and poriferans [51-54]. These ncRNAs are produced from intergenic repetitive elements (e.g., retrotransposons) through several incompletely understood biogenesis pathways including an amplification mechanism called the ‘ping-pong’ cycle [55]. piRNAs have been shown to physically associate with the PIWI subfamily of Argonaute proteins, hence their name [42, 56-58]. They possess phosphorylated 5’ ends and a 2’-O-methyl (2’-O-me) modifications at their 3’ ends [56, 59-61]. piRNAs that are transcribed from transposon elements are involved in suppressing transposon activity (function essential for germline integrity), whereas those derived from piRNA clusters are implicated in modulating gene expression [62-65]. Recent
studies have also established the expression of piRNAs in somatic cells including neurons [66].

Figure 1.4: Biogenesis and function of Piwi-interacting RNAs [67].

1.5. Short interfering RNAs (siRNAs)

siRNAs are a class of ~21 nt long sncRNAs, found in a variety of organisms such as fungi (e.g., *Schizosaccharomyces pombe*), protozoa (e.g., Trypanosoma), plants (e.g., Arabidopsis), invertebrates (e.g., *Drosophila* and *Caenorhabditis elegans*) and vertebrates (e.g., human and mouse) [50, 68-71]. They are processed from longer dsRNA precursors though repeated
Chapter 1  

Introduction

fragmentation steps catalyzed by endoribonuclease Dicer releasing ds siRNA intermediates from one of the dsRNA end. These ds siRNA intermediates are composed of two complementary ~21 nt long strands that form a ~19 base pair duplex featuring 2-nt single-stranded overhangs at each of its two 3’ ends [42]. In *Drosophila*, dsRNA fragmentation is catalyzed by one of the two fly Dicer enzymes called Dicer2, whereas in mammals, only the Dicer protein is responsible for this reaction. One of the two strands of the ds siRNA precursor subsequently associates with an appropriate Argonaute (AGO) protein [72, 73]. Based on their origin, siRNAs can be classified into the following categories:

**Exo-siRNAs:** These are produced from exogenous long dsRNAs derived from transgenes, viruses or exogenously supplied dsRNAs.

**Endo-siRNAs:** These are produced from endogenously encoded precursors, which might be long stem-loop structures, sense-antisense transcript pairs and transposon transcripts [72-75].

Endo-siRNAs are divided further into several subclasses depending on their source, length and function [42]:

- **casiRNAs** (cis-acting siRNAs) - 24 nt long, arise from tandem repeats, repetitive elements and transposons and function in chromatin modification via DNA methylation and histone modification [76].

- **natsiRNAs** (natural antisense transcript-derived siRNAs) - generated from dsRNA, which is made by convergently transcribed genes in the size range of 21-24 nt and regulate stress-response genes [42, 77, 78].

- **lsiRNAs** (long siRNAs) - produced from antisense transcript pairs, are 30-40 nt long and induced by changes in growth conditions or pathogen challenge in plants [78].

- **tasiRNAs** (trans-acting siRNAs) - 21nt in length, found in plants and targets genes other than the siRNA-producing gene [79]. It is derived from an mRNA cleaved by a miRNA where the product of a miRNP cleavage is used as
a template to synthesize double stranded RNA using RNA dependent RNA polymerase (RdRP) [80, 81].

scnRNAs (small scanRNAs) - ~28nt in length and mediate gene silencing by induction of histone methylation and ultimately, through elimination of DNA in certain ciliate protozoa. The micronuclear genome is transcribed in both orientations whose product is cleaved by Dicer like protein. This cleavage is followed by diffusion to the old macronucleus and then to the new macronucleus to scan for the region that has to be excluded (Figure 1.5). After incorporation into RITS (RNA-induced transcriptional silencing) complex, histone modification is carried out that recruits proteins which are responsible for DNA elimination.
1.6. MicroRNAs

miRNAs are 19-25 nt long single-stranded RNAs generated by RNase III-type enzymes from local hairpin structures embedded within endogenous transcripts [69, 82, 83]. They are encoded in most eukaryotic organisms and involved in
regulation of essential and a variety of regulatory pathways, which include control of organ development, cell differentiation, proliferation and apoptosis [84-87]. Human and mouse genomes encode hundreds (and possibly thousands) of distinct miRNAs, each potentially targeting an extensive set of mRNA targets [88-91].

MicroRNAs were discovered in 1993 by Victor Ambros, Rhonda Feinbaum, Gary Ruvkun and Rosalind Lee when they studied regulation of a transcription factor gene \textit{lin-14} during larval development of the nematode \textit{C. elegans} [45, 92]. They found that \textit{lin-14} expression was regulated by the \textit{lin-4} gene product known to control the timing of \textit{C. elegans} larval development. The surprise came when the Ambros and Ruvkun teams realized that \textit{lin-4} gives rise to short RNA products, instead of encoding for a protein. It was also noticed that a 61-nt precursor RNA transcribed from the \textit{lin-4} gene was processed to a 22-nt RNA that contained sequences which, on base pairing in an antisense manner, were partially complementary to 7 cognate sites in the 3’ UTR of the \textit{lin-14} mRNA. This complementarity was found to be both necessary and sufficient to inhibit \textit{lin-14} mRNA translation to the LIN-14 protein [93]. Although originally considered a nematode idiosyncrasy, in hindsight, the \textit{lin-4} small RNA was the first known example of an extensive class of microRNAs expressed in most eukaryotic organisms [94-96]. The first glimpse of miRNA evolutionary conservation came in 2000 following the discovery of \textit{let-7}, a gene encoding another heterochronic non-coding RNA, which repressed \textit{daf-12, lin-42, lin-41, lin-28} and \textit{lin-14} expression during \textit{C. elegans} developmental stage transitions from the late-larval to adult cell fates in a manner similar to \textit{lin-4} [97, 98]. \textit{let-7} was soon found to be conserved in other species, including flies and humans [97, 99].

miRNAs function as guide molecules in post-transcriptional gene silencing by either repressing translation or/and destabilizing target mRNAs [84, 100, 101]. In plants, direct cleavage and degradation of the target mRNA occurs by complete or nearly complete complementary base pairing of miRNAs with their messenger RNA targets through a mechanism involving the so-called \textit{Slicing activity} of AGO proteins (RNAi) machinery [102, 103]. In contrast, most animal miRNAs interact in a partially complementary manner with the 3’
untranslated regions (3’ UTRs) of their miRNA targets. This normally leads to inhibition of protein synthesis or/and mRNA destabilization via Slicer-independent mechanisms that preserves the stability of the mRNA target. Some recent studies have even shown that those target mRNAs that are translationally repressed may still remain in association with the stalled ribosomes [104-107]. Other types of regulation, such as translational activation [108-111] and heterochromatin formation [112-116] have also been documented. A few other miRNA functions have also been postulated. For example, due to their distinct property of base pairing with target mRNAs, miRNAs possess the potential to regulate nuclear pre-mRNA processing or modify RNA structure by acting as chaperones or modulate the interactions between mRNA and proteins [108]. Recent studies have indicated that Hepatitis C Virus replication is dependent on miR-122 expression. This happens by binding of miR-122 to two adjacent sites close to the 5’ end of HCV RNA. Once again, this reveals another instance where a miRNA was found to be involved in positive regulation of function [117-119].

For the miRNA/mRNA interaction to occur, 6-8 base pairs (bp) of perfect complementarity between the miRNA 5’ terminus (so called “seed sequence”) and a cognate miRNA target site (MTS) in the mRNA 3’ UTR is normally required [104, 106, 120-122]. miRNAs and their targets seem to form complex regulatory networks. For example, a single miRNA can bind to and regulate many different mRNA targets and conversely, several different miRNAs can bind to and cooperatively control a single mRNA target [123]. Since over one third of all human genes have been predicted to be targeted by miRNAs [106], each cell type must be characterized by the expression of a unique combination of miRNAs that might be responsible for the modification of transcription and utilization of many thousands of distinct mRNA transcripts [124]. Moreover, recent studies have shown that miRNAs often function as a part of larger gene regulation networks containing important transcription factors and RNA-binding proteins contributing to normal organismal development and physiology [125-132].
1.7. Structure of microRNA genes

Even though it was 1993 when the first miRNA, *lin-4*, was discovered [45, 92], the miRNA field gathered pace only after the discovery of the highly conserved let-7 small RNA in 2000 [97]. With >1000 distinct miRNAs encoded in human genome, miRNAs represent one of the largest gene families that have been mapped on all human chromosomes except the Y chromosome [89]. ~55% of *C. elegans* miRNAs are known to have homologues in humans, which indicates the possible phylogenetic conservation amongst many of the bilaterian animal miRNAs and thereby suggesting the importance that miRNAs have had in animal evolution [133].

Based on their genomic location, miRNA genes can be: *Intergenic*, *Intronic*, and *Exonic*.

*Intergenic miRNA* are encoded by functional intergenic transcription units containing their own promoters and regulatory elements. According to some studies, majority of intergenic miRNAs are believed to have a primary transcript of 3-4 kb in length, with a distinct transcription start site and a poly (A) signal [134, 135]. Intergenic miRNAs can be either monocistronic or polycistronic (see Figure 1.6) and can even have their own promoters and regulatory elements.

Both protein-coding and non-coding annotated genes located within the introns constitute *Intronic miRNA* genes [94, 95, 136]. About half of all known mammalian miRNAs rely on this coding strategy [137]. Similar to other scenarios, intronic miRNAs can occur as a cluster (polycistronic configuration) or as a single miRNA (monocistronic configuration) [138]. The transcription of intronic miRNAs happens from the same promoter as the host genes followed by processing from the introns of the host gene transcripts. For instance, the intron 8 of the zinc finger protein 5 gene contains both human and mouse miR-186 [139]. Intronic miRNAs are known to usually (albeit not always) share a single primary transcript, that is, they lie in the same orientation as the pre-mRNA in which they reside and are coordinately
expressed with them [137, 140]. In this case, their transcription is often driven by the host pre-mRNA promoter, ensuring co-expression of the miRNA and the spliced mRNA transcript. One example of this coding strategy includes the miR-17-92 cluster found in the intron 3 of C13orf25 (chromosome 13 open reading frame 25) gene [141]. A small number of miRNAs exist in the untranslated regions of protein-coding mRNAs and it is likely that these transcripts generate either the protein or the miRNA (but not both) from a single RNA precursor [142]. Around 50% of the known mammalian miRNAs have been traced within either the introns or exons of non-coding RNAs or within the introns of protein-coding genes, rather than in their own unique transcription units [137]. In certain cases, the location of intronic miRNA genes is indicative of their roles in coordinating physiological processes. For example, miR-7 is situated in the intron of the hnRNP k (heterogenous nuclear ribonucleoprotein K) gene thus suggesting possible co-expression of miR-7 and hnRNP K protein. Of note, there are certain cases where intron-encoded miRNAs “hijack” the intron splicing machinery to ensure their biogenesis. Such introns are called *mirtrons* [143-145].

As their name suggests, *exonic miRNAs* overlap with exonic sequences that are normally found in non-coding genes. These miRNAs are much rarer than the intergenic or intronic ones. These miRNAs typically lack their own promoters and are under the regulation of the host gene promoters. Notably, their maturation is expected to inactivate the host gene transcript due to Drosha cleavage [137] (see Figure 1.6 below).
miRNAs are normally transcribed by RNA polymerase II, regardless of their coding strategy, as capped and polyadenylated primary miRNA precursors (pri-miRNAs) [147, 148] that range from hundreds to thousands of nucleotides in length [147-149]. A large number of Pol-II-specific transcription factors can regulate miRNA gene transcription, thus allowing them to establish elaborate spatio-temporal expression patterns. Of note, RNA Pol III has been shown to transcribe a small set of miRNA genes that are associated with Alu repeats [150].
Some miRNA encoding loci are clustered resulting in similar expression patterns (Figure 1.7), and hence they are transcribed as polycistronic transcripts. However, some others lie at large distances from other miRNAs, suggesting that they form their own transcription units (TUs) [140, 151-153]. Although cases have been reported where individual miRNAs are transcribed from separate gene promoters however, they may only be exceptions.

Figure 1.7: The structure of human pri-miRNAs. (A) Examples of miRNAs with their own transcription units, such as polycistronic miR-17–92-1 cluster, and miR-21. (B) miRNAs that are transcribed in conjunction with other genes [100].

Most mammalian miRNA genes, probably as a consequence of gene duplications, possess multiple isoforms (paralogues). Paralogues are characterized by having identical sequences at 2 – 7 nucleotide positions relative to the 5’-end of the miRNA that have been found to be critical for base pairing with its target. When a whole genome study was done in different
organisms for miRNAs that are encoded in more than 1 locus, it was found that as much as 15% of human miRNAs and 12% of mice miRNAs are encoded in more than locus. This figure is as high as 42% for Zebrafish which is because of the genome duplication that took place in Zebrafishes over the course of evolution.

The transcription of some miRNAs occurs from non-coding transcription units, while the rest from protein-coding TUs. Of the miRNAs being transcribed from the non-coding TUs, about 40% are present in the intronic region, whereas about 10% are placed in the exonic region. In the protein coding transcription units, miRNA encoding loci are generally found in intronic regions that make up about 40% of all miRNA loci. Even when miR-124 is considered, a similar case is seen (Figure 1.8). In human and mouse, the miR-124 isoform of this miRNA is encoded by three distinct non-allelic genes, 124-1, 124-2 and 124-3 [154-156].

![Figure 1.8: Schematic depiction of the three miR-124 gene structures and localization in mouse.](image)

A pertinent example of genetically “redundant” miRNAs is provided by the nervous-system specific miRNA miR-124 that has been examined in this thesis. In human and mouse, the miR-124 isoform of this miRNA is encoded by three distinct non-allelic genes, 124-1, 124-2 and 124-3 [154-156]. Of these, mouse 124-1 is encoded on Chr14; 124-2 on Chr3 and 124-3 on Chr2. Interestingly, 124-2 is encoded intronically, whereas 124-1 is located within an
Chapter 1 | Introduction

exon of a spliced non-coding transcript. Finally, 124-3 is encoded in a non-
spliced transcript (Figure 1.8). Thus, the three miR-124 genes utilize an entire
repertoire of possible miRNA coding strategies.

This type of genetic redundancy is typical for many other mammalian
miRNAs (http://www.mirbase.org/) but its biological significance is poorly
understood. From a practical standpoint, the presence of the three genes has
complicated genetic dissection of mammalian miR-124 functions in vivo. Two
studies have been recently carried out that began addressing the functions of
the individual “redundant” genes. Surprisingly, non-conditional deletion of the
124-1 locus markedly reduced the mature miR-124 levels in several areas of
post-natal mouse brain [157]. Interestingly, this had no apparent effect on
neurogenesis but led to a distinct set of phenotypes characterized by reduced
brain size, increased incidence of apoptosis, aberrant axonogenesis, behavioral
signs of neurodegeneration and highly penetrant postnatal mortality due to yet-
to-be identified reasons [157]. An alternative approach used in another study
involved functional inactivation of miR-124 in mouse brain by a lentivirus-
encoded miR-124 “sponge” RNA [158]. This treatment arrested neuronal
differentiation of adult NSCs [159].

Of note, the human 124-1 gene is located in the 8p23.1 chromosome
band and genetic rearrangements of this region have been linked with various
neurodevelopmental and neuropsychiatric defects including epilepsy,
microcephaly, autism, schizophrenia, bipolar disorders, speech delay and
learning difficulties [160-162]. The small brain phenotype, signs of
neurodegeneration and aberrant hippocampal axonogenesis observed in mice
lacking the 124-1 locus suggest that the loss of 124-1 may account for some of
the symptoms observed in the patients with 8p23.1 aberrations [157].

However, the neuropsychiatric significance of reduced 124-1 expression in
specific brain regions cannot be addressed using the existing non-conditional
mouse model. This prompted us to generate miR-124 knockout mice allowing
conditional miR-124 inactivation using the Cre-LoxP recombination system
(see below).
1.8. MicroRNA Biogenesis

miRNA biogenesis is a complex, multistep process which starts with the miRNA genes transcription in the nucleus and leads to the formation of mature miRNA in the cytoplasm, which can interact with its targets. As mentioned above, RNA polymerase II (Pol II) mediates the transcription of most miRNA genes [147, 148], although Pol III transcription has been reported for a small set of miRNAs that are associated with Alu repeats [150]. miRNA gene transcription is controlled by a range of Pol II-associated transcription factors [163]. Thus, miRNA genes are endowed with an elaborate regulatory mechanism owing to Pol II-dependent transcription depending on specific cell types and conditions. The product of Pol II transcription yields primary miRNAs (pri-miRNAs) which are several kilobases (kb) in length and contain both the 5′ Cap Structures and the 3′ Poly (A) Tail [148], characteristic of Pol II transcripts. Mature miRNA sequences are embedded within local stem loop structures as a part of an imperfectly paired double-stranded stem connected by a short terminal loop.

The existing model for miRNA biogenesis has been devised on the basis of two observations [136]. Firstly, miRNAs are transcribed as long primary transcripts which are initially processed to release stem-loop containing (hairpin) intermediates (pre-miRNAs), which in turn are subsequently cleaved to form mature miRNAs. Secondly, there is a compartmentalization in the nucleus and the cytoplasm (Figure 1.9) of the catalytic activities for the first and the second processing reactions, respectively (Drosha and Dicer; see below). Hence, a dedicated export step is necessary for the nucleus to cytoplasm translocation of the pre-miRNAs.
1.8.1. Nuclear processing by Drosha

Enzymatic cleavage of pri-miRNA at the stem of the hairpin structure commences the biogenesis of miRNA, which releases a ~70-100 nucleotide hairpin called the pre-miRNA or miRNA precursor [136]. Within the nucleus, this reaction is catalyzed by the Microprocessor complex, a multiprotein complex of approximately 500-650 kDa in size, with the core components being Drosha and DGCR8 (Di George Syndrome critical region gene 8) [165-169].

Drosha is a large ~160 kDa protein, which contains a double-stranded RNA-binding domain (dsRBD) and two tandem RNase III domains (RIIIDs) which are central for its endonucleolytic function (Figure 1.10) [170]. Also, the protein’s central region, adjacent to the RIIIDs, has been proved essential for pri-miRNA processing [170]. Drosha interacts with its cofactor DiGeorge syndrome critical region gene 8 (DGCR8), also called Pasha (Partner of Drosha) [166, 167]. DGCR8/Pasha is a large protein (~120 kDa in size) which
contains two dsRBDs along with a WW domain that may function as an interaction module for proline rich polypeptide sequences (Figure 1.10). However, it is presently unknown if this domain may indeed interact with the proline-rich region of Drosha [168].

![Figure 1.10: Structure of human Drosha and DGCR8 [44]. P-rich = Proline rich, RIID = RNaseIII domains, dsRBD = double stranded RNA Binding Domain.](image)

A typical metazoan pri-miRNA consists of an extended stem element made up of about 33 bp, a terminal loop and flanking 5’- and 3’-terminal single stranded RNA segments. DGCR8 interacts with the Y-shaped base of the stem-loop structure and thus guides the positioning of Drosha-catalyzed cleavage at about 11 bp from the stem base (Figure 1.11). This reaction releases the pre-miRNA stem-loop element containing a 2 nucleotide overhang at its 3’ end [166, 167, 169, 171]. Since the cleavage site is situated at ~2 helical turns (~22 nt) from the terminal loop, it was discerned that the Drosha complex possesses the ability to measure the length of the stem [172].
1.8.2. Export of pre-miRNA to the cytoplasm from the nucleus by exportin-5

Exportin 5 (EXP5), a member of the Ran-dependent nuclear transport receptor family, mediates the export of pre-miRNAs to the cytoplasm [173-175]. EXP5 cooperatively binds cargo RNAs and a GTP-bound form of its cofactor Ran in the nucleus. After the hydrolysis of GTP, the cargo is released in the cytoplasm. EXP5 recognizes >14-bp dsRNA stem elements along with a short 3’ overhang (1 – 8 nucleotides), which allows this protein to also participate in nucleocytoplasmic export of tRNAs [173, 176-179]. It is a protein of about 113 kDa containing a Nuclear Transport Receptor (NTR) domain characteristic of many Ran-dependent nuclear transport factors (Figure 1.12). Owing to the compartmentalization of the two processing events, export of pre-miRNAs to the cytoplasm from the nucleus is obviously a vital step in miRNA biogenesis [180, 181].
1.8.3. Cytoplasmic processing by Dicer

Following their export to the cytoplasm from the nucleus, cleavage of pre-miRNAs near the terminal loop is carried out by the cytoplasmic RNase III Dicer that results in the release of ~22 nucleotide miRNA duplexes containing a mature miRNA strand and a partially complementary miRNA* strand [182-186]. While the 3' end of the mature miRNA is determined by Drosha, the 5’ end that is cleaved about 22 nucleotides from the pre-existing terminus of the pre-miRNA is measured and controlled by Dicer. Dicer is a large, highly conserved multi domain protein found in almost all eukaryotic organisms (Figure 1.13). It is ~200 kDa in size and is comprised of two RNase III type domains, a dsRBD and a long N-terminal segment. This N terminal region contains a DEAD-box RNA helicase domain, as well as a DUF283 domain and a PAZ domain and is linked to the RIIIDs by a connector helix [185, 186].

Human Dicer can associate with two closely related proteins PACT (also known as PRKRA) (Figure 1.13) [187] and TRBP (TAR RNA-binding protein; also known as TARBP2) [188, 189]. Although not required for the catalysis these interactions may modulate Dicer, lead to RNA induced silencing complex [133] formation and regulate stability of miRNA [190-192]. Interestingly, PACT is an activator of protein kinase R (PKR), whereas TRBP is an inhibitor. One dsRBD known as Medipal in TRBP is a protein-protein interaction domain that binds Merlin, Dicer, and PACT. The TRBP C-terminus or Medipal interacts with Dicer via its ATPase/Helicase domain [193]. In D. melanogaster, two different Dicer paralogs can also associate with several
partner proteins called Loquacious (LOQS), which contains three dsRNA binding domains, R2D2 and FMRI [194-197].

Figure 1.13: Domain organization of Dicer and associated proteins.

Dicer interacts with pre-miRNA substrates by binding to their 3’ overhangs using its PAZ domain (Figure 1.14). This positions the substrate correctly for cleavage by the two catalytic domains of Dicer, 22 nt away from the PAZ recognition site within the double stranded stem [198, 199]. Recent studies have also reported that Dicer contacts the 5’ end of the pre-miRNA and this interaction contributes to pre-miRNA recognition and correct positioning of the cleavage site [200].

Figure 1.14: Possible mechanism of action of Dicer [84].
1.8.4. RISC

Following Dicer cleavage, one of the two strands of the resulting ~22 nt RNA duplex is loaded into a protein complex, called RISC (RNA Induced Silencing Complex). In its completely assembled form, RISCs can be thought of as a RNP complex containing members of the Ago protein family, characterized by the presence of Piwi and PAZ-domains, as well as a miRNA (or an siRNA) guide strand, and siRNA/miRNA-complementary target. As already mentioned, the mature RNA is that strand of the ~22 nucleotide RNA duplex, which remains in Ago (also called the guide strand) while degradation of the other strand (miRNA* or the passenger strand) takes place [201-203]. In humans, Ago, TRBP (and/or PACT) and Dicer proteins can interact forming the so-called RISC loading complex (RLC), with Ago being the core component [182, 186, 187, 204, 205]. In flies, RLC comprises of Dicer1, LOQS and AGO1 [188, 204, 206]. Interestingly, formation of the human RLC complex does not require ATP hydrolysis, and only after a ternary complex is formed between Ago2, TRBP and Dicer, the exported hairpin joins the RLC [206]. RISC strand loading preference is determined by the thermodynamic stability of the duplex ends of the miRNA (or siRNA). The strand whose 5’ end is less stably base-paired to the complementary strand is normally selected as the miRNA/guide strand and loaded onto RISC [207, 208]. Strand selection is apparently a flexible process, since some pre-miRNAs are known to give rise to miRNAs originating from both strands of the dsRNA stem with comparable efficiencies [188, 209].

Argonaute (Ago) is a key subunit of RISC which is crucial for the miRNA function. In line with this important activity, Ago proteins are known to contribute to diverse facets of embryonic development, cell differentiation and stem cell maintenance [210-213].

Several biological activities of bona fide Ago proteins have been described. Some (but not all) Agos can catalyze mRNA slicing reaction when the recruited guide strand (miRNA or siRNA) is perfectly or nearly perfectly complimentary to its target site. On the other hand, Ago complexes loaded with small RNAs partially complementary to their targets bring about
translational repression or Slicer-independent destabilization of the mRNA [214-217]. Moreover, Ago proteins have been shown to modulate miRNA abundance at a post-transcriptional level. The activity and steady-state levels of mature miRNA are diminished due to the absence of endogenous Ago2 [218, 219]. This particular role of Ago2 is not dependent on its Slicer function and endonuclease activity.

Based on sequence similarities and phylogenetic analysis, the Argonaute protein super-family can be broadly divided into three clades, or subfamilies: Ago subfamily, PIWI subfamily and a recently identified new subfamily the *Caenorhabditis elegans*-specific group 3 Argonautes or Wagos [191, 220, 221].

![Figure 1.15: Model for Argonaute action in RISC loading, unwinding and target recognition [222].](image)
Of the four paralogous Ago proteins in animals (Ago 1-4), only Ago2 has retained demonstrable endonuclease (or Slicer) activity. This activity requires the presence of a conserved catalytic triad containing DDH amino acid residues, which also function as a metal coordination site. Ago 1, Ago 3 and Ago 4 proteins that are encoded in a single 190 kb locus have apparently lost their mRNA Slicing competence during evolution [190, 191, 223]. All four Agos have a molecular weight of ~100 kDa and a well conserved structure comprising the Mid domain, the N-terminal domain, as well as the signature PAZ and Piwi domains (Figure 1.16). The small RNA guide is anchored with the help of the PAZ and Mid domains. PAZ binds the miRNA 3’ end using a stretch of conserved aromatic residues, of which two tyrosines and one histidine interact with the phosphates of the two 3’ terminal nucleotides. The binding pocket for the miRNA 5’ end is provided by the Mid domain, and the Piwi domain forms the catalytic motif of Ago2 as it contains an RNase H motif.

![Figure 1.16: Domain structure of Argonaute 2 [220].](image)

Upon loading of a highly complementary target into an Ago2 complex, the scissile phosphate of the target mRNA is brought opposite to the nucleotides 10 and 11 of the small RNA guide. This allows cleavage of the mRNA with the formation of a 5’ phosphorylated and a 3’ hydroxylated termini [217, 224-227].
Figure 1.17: Schematic depiction of the siRNA-guided mRNA cleavage by Argonaute [190].

Apart from recruiting pre-processed mature miRNA, Argonautes have also been described to be involved in Dicer-independent processing of some pre-miRNAs showing a high degree of complementarity along the hairpin and a small loop [228, 229]. Ago2, owing to its Slicer activity, produces the ac-pre-miRNA or Ago2-cleaved precursor miRNA which is a nicked hairpin generated by the cleavage of the 3’ arm of the prospective passenger strand (Figure 1.18). The ac-pre-miRNA is then converted into mature miRNA using a yet-to-be-understood mechanism [229, 230].
In many cases, mRNA silencing activity of AGO requires an additional partner protein called GW182 [232-235]. Mammals synthesize three GW182 paralogs, termed TNRC (tri-nucleotide repeat containing) proteins 6A, 6B, and 6C. On the other hand, only a single GW182 protein is found in insects. The N-terminal segment of GW182 contains GW, WG or GWG repeats which may interact with AGO and have also been shown to play a role in mRNA repression by recruiting the CCR4-NOT deadenylase complex [233, 235-237]. This region is followed by a glutamine-rich (Q-rich) domain and a putative ubiquitin-associated (UBA) domain, which plays a role in localization of GW182 to cytoplasmic foci called Processing (P) bodies (Figure 1.19). The C-terminal segment is called the ‘silencing domain’ and it is essential for the mRNA repressive activity. It is a bipartite region that is divided into C-terminal and Middle (Mid) subdomains. An RNA-Recognition Motif (RRM) is flanked by these two subdomains. This RRM has apparently lost its RNA-binding function due to the presence of an additional C-terminal alpha helix [238].

Figure 1.18: Cytoplasmic miRNA biogenesis steps [231]. a) Formation of the RISC loading complex. b) Dicer dependent (left) and Dicer independent processing of pre-miRNAs.
1.8.5. Biogenesis of intronically encoded miRNAs

As mentioned above, a large fraction of miRNAs are encoded within introns of protein-encoding or lncRNA genes, where the pre-miRNA excision and splicing form highly coordinated co-transcriptional processes. Interestingly, the Drosha processing step may precede splicing of the host intron without impairing correct joining of the exons [240].

In a handful of cases, miRNAs embedded in short introns (called ‘mirtrons’) do not require Drosha processing for their biogenesis (Figure 1.20). Splicing of short introns possessing the potential to form hairpins, that structure-function and knockdown studies have revealed, generates pre-
miRNA mimics via the mirtron pathway. Mirtrons were originally identified in flies, worms and later in rodents, primates, chicken, cow, and rice [74, 144, 241-243].

Following splicing, the lariat debranching enzyme resolves the branch point of the lariat-shaped intron (where the 3’ branchpoint is ligated to the 5’ end of the intron). The resulting debranched intron forms a hairpin structure which is processed further by exonucleolytic trimming and translocated to the cytoplasm by Exportin-5, cleaved by Dicer and loaded onto RISC for target regulation [143, 241, 244].
In an ideal situation, mirtronic pre-miRNA ends are characterized by the presence of exon-intron boundaries. However, mirtronic pre-miRNA hairpins often reside at one end of the intron. In this case, mirtron biogenesis proceeds through “tailed” intermediates containing extensions at either 5’ or 3’ end of the pre-miRNA hairpin. Prior to eventual dicing, these must undergo additional trimming steps [143]. For example, miR-1017 locus has been reported to have such a 3’ tailed mirtron where the intronic sequence following the hairpin is a ~100 bp tail extending to the splice acceptor site. Post the splicing and debranching events, the RNA exosome (3’-5’ exonuclease complex) trims the 3’ tail succeeding the hairpin, a reaction which depends on a specific component of the nuclear RNA exosome called the Rrp6 exonuclease. The trimming reaction renders the tail short enough to act as a substrate for Exportin-5, which then transports it to the cytoplasm. However, the biogenesis of 5’ tailed mirtrons (e.g. mmu-miR-3103) is understood to a lesser degree. Presumably, it depends on different exonucleases such as XRN1 or XRN2, the two major 5’-3’ exonuclease in the eukaryotes.
1.9. Regulation of miRNA biogenesis

A number of miRNAs have been shown to be controlled by developmental and/or tissue-specific signaling pathways [50, 154, 155, 246-248]. This regulation occurs at both transcriptional and post-transcriptional steps of the biogenesis of miRNA to ensure precise control of miRNA levels.
1.9.1. Transcriptional Control

Transcription is an important step in regulating miRNA biogenesis. Intronic miRNAs are often positioned under the control of the promoters of the host gene thus leading to a strong correlation between the host mRNA and miRNA expression patterns. However, some of the intronic miRNAs are known to utilize distinct miRNA-specific promoters [249]. Intergenic miRNAs also have dedicated promoters and when transcribed by RNA Pol II, they contain the hallmark features of Pol II transcription- 5' cap and a poly (A) tail [249]. Pol II-associated factors may also affect their expression. E.g, myogenic transcription factors such as MyoD1, induces the transcription of miR-1 and miR-133 during myogenesis by binding to their upstream elements [250, 251]. Similarly, the tumor suppressor p53 causes the activation of miR-34 family of miRNAs, which is a part of a miRNA cluster, while MYC downregulates the transcription of a number of miRNAs that have roles in apoptosis and cell cycle [252]. Many other transcription and growth factors also influence miRNA transcription such as MEF2, PU.1, TGF-β and REST [253-255]. Epigenetic mechanisms are also known to play a role in regulating miRNA levels. For instance, DNA methylation of miR-203 locus occurs in T-cell lymphoma. However, it does not occur in normal lymphocytes [256]. Hypermethylation of many more miRNA loci, such as miR-34b/c, -137, -342, -9-1 and -193a is found in multiple human cancers, that are known to be similar to tumor suppressor gene loci [257, 258]. During development miRNA promoters have been shown to be potentially regulated by histone modifications. A set of miRNAs such as miR-1 are known to be upregulated by histone deacetylase (HDAC) inhibitors but certain other miRNAs like miR-27a are repressed [259-261].
1.9.2. Post transcriptional Control

Following the transcription step, the pri-miRNA is subjected to a host of regulatory factors in order to control its processing in the cell at several steps. Apart from the pri-miRNA itself, the factors responsible for its biogenesis also are under a tight check to maintain the levels in the cell.

One of the first points of post-transcriptional regulation of biogenesis of miRNA is provided by Drosha-catalyzed processing. For example, miR-21 levels are up-regulated in the presence of bone morphogenetic factors (BMPs) and transforming growth factor-β (TGF-β). These signaling events lead to the recruitment of SMAD proteins to the pri-miR-21 precursor. This stimulates binding of Drosha and the DEAD-box Microprocessor RNA helicase DDX5 (p68) leading to increased pri-miRNA processing efficiency and elevated steady-state levels of mature miR-21 [262]. Both DDX5 (p68) and the other Microprocessor RNA helicase DDX17 (p72) can themselves contribute to the pri-miRNA regulation processing, possibly by rearranging the pri-miRNA hairpin structure or serving as a framework for deployment of other regulators to the Microprocessor complex [263, 264]. Heterogenous nuclear ribonucleoprotein A1 (hnRNPA1), which specifically recognizes only miR-18a causes Drosha-mediated cleavage of pri-miR-18a. hnRNPA1 creates a more favorable cleavage site for Drosha by binding to the conserved loop of the pri-miRNA and changing the hairpin conformation [265, 266].
Figure 1.24: Regulation of miRNA processing. a) The pri-miRNA is cleaved by the microprocessor complex Drosha–DGCR8 and pre-miRNA is released. b) Additional specificity factors (e.g. p72 and p68) are required by some miRNAs for efficient cleavage. c) Drosha mediated cleavage of pri-miR-18a is facilitated by its interaction with hnRNP. d) TGF-β signalling enhances the efficient processing of miR-21 precursor by Drosha by inducing SMAD binding to it [231].

Dicer-dependent pre-miRNA processing might also serve as an important post-transcriptional control mechanism. This can be realized through nuclear retention of pre-miRNAs preventing their access to cytoplasmically localized Dicer [267]. On the other hand, Dicer reaction might be repressed by competitive inhibitors hindering Dicer-pre-miRNA interaction [268]. Interestingly, let-7 targeting sites have also been identified in Dicer mRNA coding region thus suggesting that let-7 targets Dicer directly and establishes a miRNA/Dicer autoregulatory negative feedback loop [269].

Another possible way to regulate miRNA biogenesis is though RNA editing. One commonly observed modification catalyzed by adenosine deaminases (ADARs) involves conversion of select adenosine (A) bases to inosines (I). This may alter base pairing and structural properties of miRNA
precursors making them poor substrates of the RNase III processing enzymes. The consequences of A to I modifications have been investigated in some detail for miR-142 and miR-151 [270, 271]. On one hand where the Drosha mediated cleavage of pri-miR-142 is inhibited by A-to-I editing, on the other hand the presence of inosine-uracil pairs in double stranded RNAs is observed to stimulate its degradation with the help of ribonuclease Tudor-SN [270, 272].

Several mechanisms that regulate cellular Ago protein levels have been described, which provide yet another possibility for post-transcriptional management of miRNA function and stability. For instance, Ago2 protein is destabilized by the LIN-41 ubiquitin ligase which has been demonstrated to inhibit cellular differentiation [273]. On the other hand, Ago2 stability can be increased by treatment with epidermal growth factor (EGF) [274]. Hydroxylation of Ago2 at the proline 700 by type 1 collagen prolyl-4-hydroxylase (C-P4H(I)), composed of the α-(P4H-α(I)) and β-(P4H-β) subunits has also been shown to have a protein stabilization effect [275]. Finally, phosphorylation of specific serine residues of Ago2 has been linked with its P-body localization.

Figure 1.25: Regulation of Argonaute levels in the cell. a) Localization of Ago2 to P-bodies is regulated by serine phosphorylation. b) Stability of human Ago2 is affected by prolyl hydroxylation [231].
1.9.3. Feedback circuits in miRNA regulation

Several feedback loops have been shown to contribute to regulation of miRNA biogenesis that involve the miRNAs, their targets and the biogenesis factors. Two types of feedback mechanisms are generally observed: *Single- and Double-negative feedback*. An interesting example of feedback regulation involves downregulation of DGCR8 by Drosha by cleaving DGCR8 mRNA and upregulation of Drosha by DGCR8 through protein stabilization [276].

![Figure 1.26: Feedback regulation loop of Drosha and DGCR8 [231].](image)

Another example is the double-negative feedback loop between let-7 and LIN28. LIN28 protein synthesis is suppressed by Let-7, whereas LIN28 blocks let-7 maturation [277-279].
Figure 1.27: Feedback mechanism loop between let-7 and lin-28. a) pri-let-7 processing by Drosha is inhibited by lin-28. b) pre-let-7 cleavage by Dicer is inhibited by lin-28, which also recruits a terminal uridylyl transferase (TUTase) to pre-let-7 [231].

1.10. miRNA-mRNA interaction

Mature miRNAs loaded into RISC use base-pairing mechanism to interact with their mRNA targets. In plants, most miRNAs by virtue of their base pairing to their targets with almost perfect complementarity lead to mRNA degradation using the Ago-dependent slicing mechanism [280]. However, metazoan miRNA normally base pair with their targets imperfectly and in accordance with specific rules. Although the complete set of these rules is still being elucidated, efficient miRNA-mRNA interaction normally requires the miRNA seed region comprising 5’-terminal nucleotides 2 through 8 to form perfect and contiguous base pairs with its cognate target site. Also, the middle of the miRNA-mRNA duplex should contain mismatches or bulges, and finally, there should be reasonable amount of complementarity of the 3’ part of the miRNA to the target, often involving residues 13-16. Some additional factors might also be important such as the presence of an AU-rich sequence context and positioning of miRNA binding sites in proximity to the poly(A) tail or the termination codon for long 3’-UTRs [281].
1.11. Mechanisms of gene silencing mediated by miRNAs

The major physiological function of miRNAs is the regulation of gene expression. Some estimates suggest that miRNAs control up to sixty percent of protein-coding genes involved in a range of cellular processes [282]. Several distinct effects of miRNAs on the gene expression program have been reported including:

1) Gene repression at the level of translation
2) Gene repression at the level of mRNA stability
3) Regulation at the level of transcription
4) Transcriptional and translational activation

1.11.1. Gene repression at the level of translation

This is the most frequently reported mechanism of miRNA action and includes repression of initiation and/or elongation of translation, ribosome drop-off, and nascent polypeptide degradation. Translation might be inhibited at the initiation step by interfering with the function of eIF4E, eukaryotic initiation factor interacting with the mRNA 5’ cap structure. An important structural determinant underlying this activity is the Ago protein “MC” subdomain that
exhibits sequence similarity to eIF4E [283]. eIF4E is known to bind the m7Gppp cap by stacking the methylated base between two tryptophans. The MC subdomain has phenylalanines at similar positions which can engage in similar interaction and sequester the cap away from eIF4E. In line with this model, miRNAs fail to repress translation that is initiated by cap-independent mechanisms [284-286].

Figure 1.29: Proposed Mechanism of Translational Repression by Ago2. Binding of the Ago2 MC domain to the mRNA cap (m7G) is mediated by miRNA-guided deposition of Ago2 to an mRNA target and thus exclusion of eIF4E and inhibition of translation initiation [283].

Another proposed model suggests that binding of the Ago partner GW182 to poly(A)-binding protein (PABP) inhibits mRNA translation by blocking the interaction between PABP and eIF4G, a cap-binding complex component involved in mRNA circularization. Repression at the initiation stage is also thought to be mediated by the inhibition of formation of the mature ribosome. For instance, miRNA miR-2 has been demonstrated to inhibit both 40S ribosomal subunit recruitment and 80S initiation complex formation [287]. One possible mechanism of this activity may depend on Ago2-dependent recruitment of eIF6, a translation initiation factor preventing
premature binding of the large (60S) and the small (40S) ribosomal subunits. Therefore, Ago2-induced increase in local eIF6 concentration is expected to repress initiation of translation [288].

Figure 1.30: Mechanism of inhibition of ribosomal subunit joining [289].

miRNA mediated repression of translation can also occur at post-initiation steps including translation elongation due to premature dissociation of elongating ribosomes and co-translational degradation of the nascent polypeptide chains; however molecular mechanisms underlying these effects are poorly understood [290-294].

Figure 1.31: Mechanisms of post initiation inhibition by miRNAs. A) Inhibition of translation elongation  B) Co-translational protein degradation [289].
1.11.2. Gene repression at the level of mRNA stability

Although metazoan miRNAs were originally thought to cause translational repression without altering the mRNA levels, subsequent studies have suggested otherwise, i.e., most miRNAs induce detectable changes in mRNA targets. This has been shown by the fact that the transcript levels of predicted and validated miRNA targets reduce significantly on ectopic expression of the miRNAs and the converse is true as well [238, 288, 295, 296]. As mentioned previously, miRNAs perfectly complementarity to their target mRNAs can directly induce Ago2-dependent endonucleolytic slicing of mRNAs. However, an overwhelming majority of metazoan miRNA are thought to destabilize their target though partial base-pairing followed by mRNA deadenylation and subsequent decapping [238, 297]. miRNA-induced deadenylation requires the CCR4-NOT deadenylation complex and, to a lesser extent, the PAN2-PAN3 deadenylase complex [232, 234, 298, 299]. Several studies have demonstrated that the Argonaute partner GW182 recruits both CCR4-NOT and the PAN2-PAN3 deadenylase complexes via its C-terminal domain [300]. Human GW182 proteins directly bind the CNOT1 subunit of the CCR4-NOT complex whereas the PAN2-PAN3 complex binds GW182 either directly, possibly through the PAN3 kinase-like domain and PABP [234, 236, 301, 302]. Following deadenylation, mRNA is decapped by the DCP1/DCP2 decapping enzyme complex and finally degraded by the cytoplasmic 5'-3' exonuclease XRN1 [303].

Figure 1.32: Mechanisms of post initiation inhibition by miRNAs by ribosome drop-off [289, 290].
Another destabilization pathway may involve miRNA-stimulated sequestration of mRNA targets to cytoplasmic foci known as P bodies. This model came into the picture upon the detection of Argonaute proteins, miRNAs, miRNA targets, as well as deadenylating and decapping factors in P bodies. Within these P bodies, mRNA molecules are localized, where they are shielded from the translational machinery and undergo decay [295, 304]. Even though P bodies are crucial in miRNA mediated silencing, aggregation into microscopic P bodies is not essential for miRNA function. This implies that localization of mRNA into P bodies could be a consequence rather than a cause of silencing [285, 286, 305].

1.11.3. Regulation at the level of transcription

A subset of mature miRNAs has been recently reported to localize to nucleus and induce transcriptional gene silencing through chromatin reorganization [306-308]. This level of regulation apparently depends on complementarity between miRNAs and sequences within the promoter regions of target genes. Upon binding, miRNAs may stimulate methylation of promoter DNA. Several lines of experimental evidence support the existence of such a mechanism. These include direct demonstration of nuclear import of miRNAs, along with
the observation that steady-state levels of target transcripts can be drastically reduced in the absence of mRNA decay [309-311].

1.11.4. miRNA mediated translational activation

Under certain conditions, miRNAs have been reported to induce translational activation [312, 313]. For example, serum starvation prompts the recruitment of AGO-miR-369-3 complex bound to the 3’-UTR of TNF-α mRNA to the fragile X-related protein 1 (FXR1) ultimately leading to stimulation of mRNA translation [314]. Another instance of such an effect is when miR-10a interacts with the 5’-UTR of many mRNAs encoding ribosomal proteins, causing increased translation of these mRNAs in response to stress or nutrient shortage [313].

1.12. MicroRNAs and the Nervous system

1.12.1. Nervous System

The adaptive abilities of biological systems to environmental changes determine their success of survival. According to Conrad Waddington, organismal development and reaction to the environment are governed by an “epigenetic system” that moulds the pathway of embryogenesis; a landscape navigated by the cell depending on external factors and adaptive responses and the topology being defined by a mesh of underlying networks.

This complex network of developmental decisions is perhaps especially relevant in the context of NS, an extremely elaborate structure allowing us to capture and process information, form life-long memories and respond to ever-
changing environment. For these and many of our functions, we rely on a plethora of intricately formed neural circuits. However, this capability is ultimately dependant on a well-conserved toolkit of gene expression mechanisms enabling relevant changes in cellular identity and physiology. Of all the molecular factors that potentially bring about changes in cellular state, our knowledge of posttranscriptional regulation of gene expression is expanding exponentially and it is becoming increasingly clear that post-transcriptional regulation of gene expression plays an important part in these processes. In addition to other post-transcriptional mechanisms, the miRNA system is known to form a critical component of the ‘‘epigenetic landscape’’ shaping the robust regulation and deployment of gene networks underlying brain construction, remodeling and function.

Nervous System (NS) is composed of the central (CNS) and peripheral (PNS) compartments. Of these, CNS is used to process and integrate the information received from all parts of the body and deliver commands adjusting the functions of these parts to changing environmental conditions and current needs of the organism. Both CNS and PNS play fundamental roles in the control of organismal physiology and behavior. The CNS can be subdivided into seven main parts:

1. The **spinal cord**, which is the most posterior part of CNS that receives sensory information from the peripheral organs, processes it and in turn controls their movement and other forms of physiological response.
2. The spinal cord continues rostrally as **medulla oblongata** containing vital centers responsible for several autonomic functions.
3. Over the medulla, lies the **pons**, which conveys information about movement to the cerebellum from the cerebral hemisphere.
4. The **cerebellum** is situated behind the pons and connected to it by fibre tracts called peduncles. It is involved in the learning of motor skills and controls the range and force of movement.
5. The **midbrain**, positioned rostrally to the pons, controls numerous motor and sensory.
6. Midbrain is connected to the **diencephalon** which contains 2 important structures; the **thalamus**, which processes a majority of the incoming
information into the cerebral cortex and the hypothalamus, which regulates autonomic, endocrine and visceral functions.

7. The cerebral hemispheres consist of the cerebral cortex (a heavily wrinkled outer layer) and 3 structures apart from this: the hippocampus (memory storage), the basal ganglia (motor performance control) and the amygdala (coordination of endocrine and autonomic responses of emotional states). The cerebral cortex consists of 4 lobes: frontal, parietal, temporal, and occipital.

![Diagram of the Central Nervous System]

**Figure 1.34: Parts of the Central Nervous System.**

The brain is sometimes also subdivided into three larger regions: the forebrain (diencephalon and cerebral hemispheres), midbrain and the hindbrain (medulla, pons, and cerebellum). The brain stem is comprised of the hindbrain, excluding the cerebellum and midbrain.
In triploblastic animals, the Nervous System is developed from the *ectoderm*, the outermost tissue layer of the embryo. During early development, the *notochord*, a flexible rod-shaped body running along the back of the embryo, prompts the ectoderm to convert to *neuroectoderm*. This gives rise to a strip of neuronal stem cells along the dorsal side of the embryo, called the *neural plate*. The neural plate is responsible for the origin of the nervous system and it is the source of a majority of neurons and glial cells in the mature human nervous system. As the development progresses, the neural plate is further transformed into the neural groove and, ultimately a hollow *neural tube*. This process by which neural tube is formed from the ectoderm is defined as *neurulation*. Owing to rapid cell proliferation, the anterior part of the neural tube expands and ultimately forms the brain. Over time, some of these cells cease to divide which is followed by differentiation into neurons and glial cells, the main building blocks of the nervous system. Migration and self-organization of these newly formed neurons to different portions of the developing brain gives rise to distinct structures and circuitries. On reaching
their regional destination, these neurons then develop axons and dendrites, which permit communication with other neurons and some non-neuronal cells via synapses. Functional neural circuits, that mediate a wide range of sensory, motor, cognitive and behavioral functions, are established by formation of synaptic communication between neurons.

The lower part of the neural tube develops into the spinal cord that in the adult NS consists of the mantle layer (the grey matter) containing somas of neuronal cells. The marginal layer (white matter) is comprised of the nerve fibers which emerge from the mantle layer. The motor areas of the spinal cord correspond to the ventral part of the mantle layer (the basal plates), whilst the sensory areas are localized in the dorsal part (the alar plates). The intermediate layer forms the connection between the basal and alar plates and it is made up of neurons of the autonomic nervous system.

**Neurogenesis** is defined as a process in which neural stem/progenitor cells (NSCs) proliferate and differentiate to form new neurons, that ultimately integrate into neuronal circuitries [316]. During development, radial glia, and neuroepithelial cells, which line the neural tubes, function as Neural Stem Cells, which are responsible for the generation of the entire central nervous system (CNS). Neurons and the two major glial cell types in the CNS arise from Neural Stem Cells in a temporally defined sequence in which the neurons appear first, and are followed by astrocytes (star shaped glial cells with diverse functions), and then oligodendrocytes (glial cells responsible for axon myelination). Once the initial embryonic development is completed, neurogenesis essentially ceases and is restricted to the subventricular zone (SVZ) of the lateral ventricles, the subgranular zone (SGZ) of the hippocampus and possibly the olfactory bulbs. Adult NSCs appear as radial glia cells localized in the ventricles and originating from embryonic neuroepithelial cells. Adult NSCs give rise to cortical neurons through intermediate progenitor cells that display astrocyte-like morphology; at the end of neurogenesis these cells might potentially be converted into bona fide astrocytes [317, 318]. However, in certain specialized niches of the brain, a subset of astrocytes persist as stem cells and keep on producing a large number of neurons that functionally integrate themselves into restricted regions [319]. Many extrinsic stimuli and intrinsic mechanisms can affect adult neurogenesis,
which include: physiological activities and enriched environment [320], hormones [321], growth factors such as fibroblast growth factor (Fgf) and vascular endothelial growth factor (Vegf) [322], transcription factors and diseases including epilepsy, stroke and depression [323].

1.12.2. Nervous System specific miRNAs

It is now acknowledged that an extensive repertoire of adaptive signaling mechanisms appropriate for responses to a wide range of temporal domains and environmental conditions or cellular interactions are possessed by the cells (Fig 1.35). Changes in conformation and catalysis or post-translational modifications of cellular molecules induce only swift, localized and short term responses. However, long term adaptive changes are sustained beyond the lifetime of individual molecules, such as the memories stored in neural networks. Constraints in space can become a major limiting factor in the nervous system because of the complexity in its cellular architecture, which is a pre-requisite for its structure and function. This endows the translational regulation at the RNA level greater significance. Considering that the cellular and transcriptional machinery of the nervous system is highly complex, it is therefore hardly surprising that the NS development and functions are regulated by an extremely wide range of molecular mechanisms including the miRNA pathway. This complex functional architecture inside a single neuron gives rise to many compartments that make it susceptible to potential regulation by miRNAs. Several studies have emphasized the importance of miRNA in neuronal differentiation using knockouts/inactivation of the miRNA biogenesis components. For example, impaired neuronal differentiation, reduction in neuronal size and loss of branching are consequences of Dicer ablation in mice [324-326]. Disruption of the Zebrafish Dicer gene also provided evidence that miRNAs are necessary for neurogenesis as the maternal zygotic mutants show gross morphological defects in brain patterning and morphogenesis [327]. Similarly, a reduction in all types of neurons and glial cells was noticed in Ago1 mutants of Drosophila [328].
However, elimination of Dicer or other components of the miRNA biogenesis pathway does not allow one to resolve function of individual miRNAs. This limitation has been overcome in recent studies using miRNA-specific knockdown and knockout approaches.

For example, simultaneous deletion of miR-132/212 genes in conditional knockout mice caused a tremendous decrease in dendritic length, arborisation, and spine density. Also, simultaneous ablation showed similar effects on dendritic development. This is because the loss of cAMP response element-binding (CREB), which regulates transcription of miR-132, can block its effects on dendritic maturation [329, 330]. Similarly, knockout studies targeting NS-specific miRNAs miR-124 [331], miR-125b [332], miR-132 [333], miR-134 [330, 334], miR-137 [335] and miR-138 [336] have convincingly implicated these miRNAs in different aspects of synaptic development and dendritic arborisation.

Figure 1.36: miRNAs involved in various aspects of synaptic development and function.
1.12.3. miR-124 in Nervous System development

As discussed above, neural development is a highly orchestrated process that necessitates accurate control of gene expression. Although microRNAs (miRNAs) have been associated with the fine-tuning of gene networks, including those in the nervous system (NS) [155, 291, 337-339], not much is known about their functions [330, 339-343]. miR-124, whose mature sequences are conserved from *C. elegans* to humans, is the most abundant miRNA in adult and embryonic CNS [246, 344]. In adult mouse brain, it accounts for 25-48% of all brain expressed miRNAs [246].

miR-124 expression has been seen in neurons, but not astrocytes, and the levels of miR-124 have also been shown to increase over time in the developing NS [155, 156, 338, 345]. Along with another NS-specific miRNA miR-9, miR-124 stimulates neuronal and represses glial differentiation of ES cells *in vitro* [346]. In addition, miR-124 decreases the levels of hundreds of non-neuronal transcripts, so that its introduction into HeLa and other non-neuronal cells promotes a neuronal-like mRNA profile [107, 347, 348]. When mouse neuronal cells were transfected with miR-124, neuron-like differentiation was stimulated by promoting neuron-specific class III β-tubulin (Tuj1) and MAP2 expression [349]. On the other hand, miR-124 inhibition by an antisense 2′-OMe-RNA results in an increase in the levels of non-neuronal transcripts in primary cortical neurons [347]. Also, in the case of primary cortical neurons and differentiating mouse P19 cells, miR-124 has been shown to target the Rho GTPase family, diminishing F-actin density and causing stimulation of tubulin acetylation and regulation of cytoskeletal reorganization thereby, stimulating neuronal outgrowth. This is brought about by cell division cycle 42 (Cdc42) targeting and deactivation of ras-related C3 botulinum toxin substrate 1 (Rac1) of the Rho GTPase family [350]. Neural cell identity also seems to be defined by miR-124 by the downregulation of several target mRNAs *in vivo* [337, 347, 351].

miR-124 can antagonize the function of NRSF/REST (RE1 silencing transcription repressor), which recruits SCP1 (Small Carboxyl terminal
Phosphatase 1) - an anti-neural factor - to RE1 (Repressor Element) containing neuronal genes. Expression of miR-124 results in degradation of non-neuronal transcripts including SCP1 and hence, promotes neurogenesis [351]. Interestingly, transcription of the three genes encoding miR-124 in mammals is in turn repressed by REST, which establishes a double negative feedback regulation circuit [347, 351, 352]. This could explain in part the stimulatory effect of miR-124 on neuronal differentiation [351, 353].

Another miR-124 target is PTBP1 (Polypyrimidine tract binding protein 1, also called PTB/hnRNP1), a global repressor of alternative pre-mRNA splicing, that plays a major role in differentiation of neurons from their progenitors [126, 354]. In addition to other splicing events regulated by this protein, PTBP1 targets a critical exon in the pre-mRNA of PTBP2, a neuronal enriched PTBP1 homolog, leading to the skipping of this critical exon and stimulating nonsense mediated decay (NMD) of PTBP2 mRNA. During neuronal differentiation, miR-124 reduces the levels of PTBP1 leading to the accumulation of spliced PTBP2.

Figure 1.37: Cross talk between SCP1, PTBP1 and miR-124 in non-neuronal cells [352].
miR-124 also targets Sox9 (SRY-box containing gene 9), an important transcription factor expressed in NSCs. When the endogenous miR-124 in mice was knocked down, the subventricular zone (SVZ) stem cells were maintained as dividing precursors, whereas ectopic expression of miR-124 led to increased neuron formation. Blocking miR-124 expression during regeneration also led to hyperplasia, proving that miR-124 is an important regulator of adult neurogenesis. Overexpression of Sox9 has the reverse effect as it abolishes neuronal differentiation and Sox9 knockdown leads to increased neuron formation [331].
Figure 1.39: Schematic summary showing modulation of multiple pathways involved in neurogenesis by miR-124 through translational regulation of its currently known mRNA targets [316].

A growing body of evidence has linked miR-124 with several neurological and neuropsychiatric disorders. For example, the miR-124/PTBP1 circuitry has been recently shown to regulate alternative splicing of the pre-mRNA encoding amyloid precursor protein (APP), a critical component involved in Alzheimer’s disease [355]. miR-124 is also known to target the mRNA of β secretase (BACE1), an enzyme involved in generation of the malignant amyloid peptide. Moreover, over-expression of miR-124 in the mouse hippocampus resulted in long-term potentiation (LTP) defects and disrupted spatial learning and memory, whereas elevated expression of this mRNA in prefrontal cortex led to aberrant social behavior [356]. The critical miR-124 target responsible for these effects is Egr1/Zif268, a zinc finger-containing transcription factor essential for stabilizing synaptic plasticity and spatial and macrophages during experimental autoimmune encephalomyelitis (EAE), whereas over-expression of miR-124 in these cells alleviated the disease
symptoms [357]. Through targeting glucocorticoid receptors, an important component of the stress response pathways, miR-124 may be additionally involved in neurological stress disorders [358, 359]. The high expression levels of miR-124 in the NS also make this miRNA a promising plasma biomarker for cerebral infarction [360].

1.13. Some outstanding problems of miRNA biology

Despite the remarkable recent advances in our understanding of miRNA biogenesis and function, several important questions still await their answers. A number of mammalian miRNAs including miR-124 are encoded in more than one genetic loci but the biological function of this genetic redundancy is poorly understood. Furthermore, ~50% of metazoan miRNAs are encoded in spliceosomal introns [44] but the significance of this bias has not been addressed systematically. In a few relevant examples that have been studied, splicing either had virtually no effect on the Drosha/Microprocessor reaction or modestly stimulated pre-mRNA excision possibly due to a physical interaction between the Microprocessor complex and the spliceosome [240, 361, 362]. Of the three mammalian miR-124 precursors, 124-2 is encoded within an intron while 124-1 and 124-3 are encoded in an exon or an unspliced transcript. To begin addressing partial functions of individual miR-124 genes and understand possible implications of the difference in their coding strategies, I have generated miR-124 conditional knockout mouse models. The following section provides a brief theoretical outline of genetic strategies used in this work.
1.14. Recombination systems used for knocking out genes in mouse

1.14.1. Homologous recombination

Homologous recombination, also called general recombination, is a type of multi-step genetic recombination event which is characterized by nucleotide sequence exchange between two similar or identical strands of DNA [1]. As mentioned, such a process involves a number of steps in which physical breakage is succeeded by the eventual rejoining of DNA strands and it is widely used for repair of potentially lethal double-strand DNA breaks. Also, at the time of meiosis during which germ cells like sperm and ova are made by eukaryotes, homologous recombination is responsible for producing new combinations of DNA sequences. These new genetic combinations of DNA enhance genetic variation in the offspring and help populations adapt to changing environmental and other conditions in the course of evolution. Horizontal gene transfer to exchange genetic material, is another area where homologous recombination is employed between different strains and species of bacteria and viruses [1].

Homologous recombination is found in all domains of life, although specific details of this process might differ. Once a double strand break occurs, in a process called resection, parts of the DNA that are around the break on the 5’-end of the damaged chromosome are removed. This is followed by the strand invasion step, which is characterized by the “invasion” of an undamaged homologous chromosome by an overhanging 3’-end of the damaged chromosome. This results in the creation of an intersection called a Holliday Junction between the two chromosomes, which is a mobile, cross-shaped intersection of four strands of DNA formed after the strand invasion step [363]. Homologous recombination has been most extensively studied and is best characterized in Escherichia coli [364]. The two most accepted models of this pathway are the RecBCD pathway, which aids in the repair of the
double-strand breaks in DNA, and the RecF pathway, which promotes repair of single-strand breaks [365].

Homologous recombination is critical for normal progression of mitosis and meiosis in eukaryotes. In mitosis, it helps in the repair of double strand breaks in DNA which are result of exposure to ionizing radiation or DNA damaging chemicals [366]. In the case of meiosis, it facilitates chromosomal crossover during prophase I [367]. Two models of homologous recombination that are believed to aid in the double-strand break repair of DNA are the SDSA pathway (Synthesis Dependent Strand Annealing pathway) and the DSBR pathway (also called the Double Holliday Junction Model) [368]. This is the method of choice when an allele has to be replaced with an engineering construct without affecting any other locus in the genome.
Mario Capecchi and Oliver Smithies carried out pioneering work showing that homologous recombination can take place between exogenous DNA fragments and the chromosomes in mammalian cells [369-373]. They also found that all endogenous genes could be targeted by this technique irrespective of their activity. Meanwhile, Martin Evans established the use of ES cells for gene targeting into the germ line [374-378]. The two technologies were eventually combined to create knockouts of mouse hprt gene [379, 380] followed by refining the gene targeting strategies and development of positive-negative selection streamlining the generation of gene targeted knockout mice (Figure 1.4) [381]. Since then, gene targeting using homologous recombination has developed into a highly adaptable technology and it has equipped us with the power to introduce mutations, both during development and in the adult
animal that can be activated in specific cells or organs, or at specific time points.

Figure 1.41: Enrichment of ES cells, using positive-negative selection, containing a targeted disruption of a gene in both (A) gene targeting and (B) random integration [382].

Homologous recombination in bacteria (often referred to as recombineering) has recently become a method of choice for engineering gene targeting constructs. Recombineering allows insertion, deletion or alteration of any sequence precisely and is not dependent on restriction sites [383-387]. Two of the popular recombineering methods rely on the Red homologous recombination system from bacteriophage λ or the RecET system from the Rac prophage [383, 388].
1.14.2. Site Specific Recombination System

Cre-Lox system is an example of site-specific recombination, which involves recombination at short $LoxP$ sites in the presence of an enzyme called Cre recombinase [390, 391]. Both the Cre enzyme and the $loxp$ site originate from bacteriophage P1 where they function as a part of the normal viral life cycle [390, 392]. The bacteriophage utilizes Cre-$loxp$ recombination as a means to circularize its genomic DNA and thereby, facilitate its replication. On the molecular level, Cre (Cyclization Recombinase) catalyzes DNA recombination between two $LoxP$ (locus of X over P1) sites each containing specific Cre-binding sequences and a directional core sequence where recombination can occur [393].
Since its discovery, Cre-LoxP system has been transformed into a suitable genetic technology for genome manipulation in mice, plants, yeasts, mammalian cell cultures, and other organisms [394]. One especially popular application of the Cre-Lox system is generation of DNA deletions, insertions, translocations and inversions in conditional knockout animals. The fact that embryonic lethality occurs as a result of knockout of certain genes makes it impossible to study biological functions of these genes in adult animals. In such cases, the use of the Cre-LoxP system allows one to knock out genes at specific developmental stages or within specific tissues.

Cre protein is 343 amino acids long and consists of two domains: the amino (N-terminal) domain, which is the smaller one and the larger carboxyl (C-terminal) domain. The C-terminal domain contains the enzyme’s catalytic site and it is similar in structure to the Integrase family enzyme from bacteriophage lambda. Cre is thought to function as a tetramer [395, 396]. LoxP site consists of 34 bp including an asymmetric 8 bp sequence flanked by two 13 bp-long palindromic sequences.

\[
\text{LoxP} = \text{ATAACTTCGTATA-GCATACAT-TACGAAGTTAT}
\]

Figure 1.43: Sequence of Lox P.

Because the loxP sites and the Cre gene are not innate to the mouse genome, they need to be incorporated by transgenic technology [397]. In cells possessing loxP sites in their genome and also express Cre, a recombination event can occur between these loxP sites. The Cre protein induces a double stranded cut at both loxP sites in the DNA, which is then rapidly rejoined. The result of a recombination event is decided by the orientation in which the loxP sites are present. A direct repeat of loxP sites on the same chromosome will lead to a deletion while inverted loxP sites, when present on the same
Chapter 1 | Introduction

chromosome arm, cause an inversion. When these *loxP* sites are situated on different chromosomes Cre may catalyze corresponding translocation events [391].

![Figure 1.44: Outcomes of a Cre-lox recombination (The Jackson laboratory).](image)

In mouse genetics, Cre and *loxP* lines are developed separately and then crossed to produce a Cre-lox line [397]. The following lines are commonly used in different genetic projects (The Jackson Laboratory):

- **Cre expressing lines:** contain a transgene that expresses Cre under the control of a widespread (general) or tissue-specific (conditional) promoter. They are used to produce general or conditional knockouts respectively.

- **Inducible Cre lines:** the transgene expresses a modified form of Cre recombinase that is non-functional until an inducing agent (such as doxycycline, tetracycline, RU486, or tamoxifen) is administered at a desired time point in embryonic development or adult life.

- **LoxP-flanked (“floxed”) lines:** contain *loxP* sites flanking a critical portion of a target gene or a genomic region of interest.
Cre reporter lines: contain *loxP* sites in combination with visible (fluorescent or lacZ) marker proteins used to trace Cre recombination success and/or alterations in gene expression.

Figure 1.45: Mechanism of creating knockout mice using Cre-Lox Recombination system.

FLP/FRT is another example of a commonly used site-specific recombination system. The FLP (Flippase) enzyme derived from the yeast *Saccharomyces cerevisiae* is a 48 kDa protein interacting in the form of a tetramer with two FRT (*Flippase Recognition Target*) sites and catalyzing the cleavage and ligation of these sites [398]. Similar to Cre/LoxP, the outcome of FLP-mediated recombination depends on the orientation of the FRT sites.

Similar to *loxP*, repeated DNA sequences which are of 13 bp each in length and an 8-bp long spacer region between the repeats make up the FRT site. The spacer region is critical in determining the overall orientation of the FRT recombination site [399-401].
Since, the FLP-FRT recombination can be organ specific, it is used to create genetic mosaics in higher organisms. Using this technology, spatial control by a gene can be studied in an organ of interest, if the organism cannot survive the gene deletion in other organs. Similarly, temporal control by a gene can also be studied by using an inducible promoter to trigger the recombination event at a specific point in development. The FRT sites can be used in conjunction with the lox P sites in the same targeting construct in order to selectively delete a selection marker like Neomycin or Puromycin. Thus, it eliminates any background effects which might be caused due to the expression of these genes.

Figure 1.46: Diagram depicting reactions catalyzed by FLP upon interaction with FRT sites [401].
Chapter 2

MATERIALS AND METHODS

2.1. Plasmids

Plasmids used in this study (including both published and newly generated ones) are listed in Table 2.1. To generate splicing-deficient constructs we amplified a fragment of pEM208 [401] using the 0.1 µM EMO559/EMO560 primers (Table 2.2) containing splice junction mutations. PCR reaction was carried out using KAPA HiFi DNA Polymerase (KAPA Biosystems; 1 unit/µl) in 1× KAPA HiFidelity Buffer. The other components of the reaction included: 1.75 mM MgCl₂, 0.2 mM dNTPs and 20 ng of linearized plasmid DNA per 20 µl reaction. Thermal cycler conditions: Initial denaturation at 95°C for 2 min, followed by 20 cycles of a three-step reaction, denaturation at 98°C for 30 seconds, annealing at 60°C for 30 seconds, extension at 68°C for 20 seconds, and a final extension at 68°C for 5 min. Both the amplicons and the plasmid pEM543 were digested with *NheI* and *XhoI* and run on a 1% agarose gel. The required bands were cut from the gel and purified using Gel Purification Kit (Nucleospin, Macherey & Nagel). It was followed by ligation using T4 DNA Ligase (Fermentas; 40 units/µl).

*124-1* and *124-3* genes were cloned within splicing deficient introns. They were also designed by mutating the splice sites of the intron. This was done by amplifying the *124-1* and *124-3* genes from pEM183 and pEM209, respectively using EMO559 and EMO560 and *NheI/XhoI* digestion followed by ligation using T4 DNA Ligase employing the same reagents and methods as described above.
Pre-miR-124-2 loop mutants were made using EMO1139 (5'CCGTGTTCACAGCGGACCTTGATTTAAATGTCCATACAAATTAAGGCACGCGGTGAATGCCAAGAG3') and EMO1140 (5'CCGTGTTCACAGCGGACCTTGATTTAAATGTCCATACAAATTAAGGCACGCGGTGAATGCCAAGAG3') for replacing with the loop of 124-1 and EMO1141 (5'CCGTGTTCACAGCGGACCTTGATTTAAATGTCCATACAAATTAAGGCACGCGGTGAATGCCAAGAG3') and EMO1142 (5'CTCTTGGCATTCACGCGTGCCTTAATTGTATAGAC(230x589)x(344x589)>AACCGG3') for replacing with the loop of 124-3. pEM560 was used as the template for the PCR reaction and KAPA HiFi (KAPA Biosystems) was used as the amplifying enzyme in a 16-20 cycle PCR reaction. The PCR product was then precipitated using 20 mg/ml Glycogen, 3 M Sodium Acetate and 100% Ethanol. The mixture was incubated at room temperature for 20 mins. It was then followed by centrifugation at 14,000 rpm for 10 min and then washed with 70% ethanol and spun at 14,000 rpm for 5 min. The pellet was suspended in suitable volume of water and the parental strands were digested with DpnI and incubated for 2 hours at 37°C. The digestion product was then transformed into competent TOP 10 cells, the colonies obtained were inoculated into LB broth and plasmid isolation was carried out to obtain the mutated products.

To generate plasmids encoding miR-124 precursor fragments of different size, primers were used to amplify a specific region within the miR-124 genes which encompass ~40 bp upstream and downstream of the stem-loop region. These shortened miR-124 genes were called “mini-124-genes”. EMO1219 (5'TTCCTTCTTCTCCTCAGGAGAAAG3') and EMO1220 (5'GGGCGTTGTCATATGTCGATCCTTCCGT3') were used to amplify mini-124-1 segment, EMO1221 (5'GTCACTTGCTTCTAGATCAAGATCA3') and EMO1222 (5'CATTCTAGAATTGCATATGCCACCCTTCCGT3') were used to amplify mini-124-2 segment, and EMO1223 (5'ACCCGGTCCAGCGCTC3') and EMO1224 (5'CCGCGACGCACTATGCCATCCCGGGAATTT3') were used to amplify mini-124-3 segment using pEM183 (124-1 gene within spliceable intron), pEM208 (124-2 gene within spliceable intron), and pEM209 (124-3 gene within spliceable intron) as the templates and KAPA HiFi Polymerase (KAPA Biosystems) in the PCR
reaction of 16-20 cycles. The vectors pEM558, pEM561, and pEM560 for miR-124-1, 2, and 3, respectively were digested with SacII and NdeI. The amplicons from the PCR reaction were also digested with the same enzymes. The vector and amplicon digested products were run on an agarose gel, purified using the NucleoSpin® Gel and PCR Clean-Up (Macherey-Nagel) and ligated for 16 hours at 4°C. The splicing deficient constructs were created using the spliceable context constructs (generated above) and mutating the splice sites using EMO560 (5'GCGCTAGCCCCTGTGGGGCCGAGTACTCCCTCTCAAAAG3') and EMO559 (5'GTAGTCTCGAGTGAGACACGGGTGGAGAAAGGCAAAGT3') in a KAPA HiFi Polymerase (KAPA Biosystems) catalyzed PCR reaction of 16-20 cycles.

To generate miR-124-expressing constructs truncated at 5' or 3' end, primers EMO1895 (5'TCAACCTCAGGAGACTCATGCACCCGCTACT3') and EMO32 (5'CTGTGGAGAGAAAGGCAAAG3') were used for amplifying the 5' end of miR-124-2 gene and primers EMO1894 (5'AGTTCCTCAGGGTACCTTGCTTCTAGATCAAGATCA3') and EMO32 (5'CTGTGGAGAGAAAGGCAAAG3') were used to amplify the 3' region of miR-124-2. pEM561 was used as the template for the PCR reaction, which used KAPA HiFi Polymerase (KAPA Biosystems) in a PCR reaction of 16-20 cycles. The amplicons were then digested with Bsu36I and NdeI (New England Biolabs® Inc). pEM560, which was used as the vector here for introduction of the amplicons, was also digested with Bsu36I and NdeI. The vector and amplicon digested products were run on an agarose gel, purified using the NucleoSpin® Gel and PCR Clean-Up (Macherey-Nagel) and ligated for 16 hours at 4°C.

For creating versions of the above constructs with splicing deficient introns, primers EMO560 (5'GCGCTAGCCCCTGTGGGGCCGAGTACTCCCTCTCAAAAG3') and EMO559 (5'GTAGTCTCGAGTGAGACACGGGTGGAGAAAGGCAAAGT3') were used in a PCR reaction of 16-20 cycles catalyzed by KAPA HiFi Polymerase (KAPA Biosystems) using the constructs generated above as template. The amplicons and the vectors were then digested with NheI and XhoI (New England Biolabs® Inc), run on an agarose gel and purified using NucleoSpin® Gel and PCR Clean-Up (Macherey-Nagel) and ligated for 16 hours at 4°C.
Pri-miR-124 constructs lacking the pre-miR-124 stem-loop structure were generated by using the primers EMO2253 (5’CAGGAGAAAGGCCTCTCTCTGAGGGCTGTCTGAGCACCTTGGGTCC3’) and EMO2254 (5’GGACCCAAGGTGCTCAGACAGCCCCATTCGAGAGAGGCCCTTTCTCCTG3’) for miR-124-1 gene, EMO2255 (5’CTAGATCAAGATCGACGTCTGGGACTCTACGGCTGACTTTG3’) and EMO2256 (5’CAAGTGCAGCCGTAGGCTCCGCTCGAGAGCAGAGTCTCTGATCTTGGATCTAG3’) for miR-124-2 gene, and EMO2257 (5’CAGGCCCTGAGGGCCCTCTGGAGAGGCCTCCGCCGCTCCTTTC3’) and EMO2258 (5’GAAAGGAGCGGCGGAGGCCCTCTCCAGAGGGGCCCTCAGGGCGTG3’) for miR-124-3 gene. These primers were used to amplify the gene segments from the templates pEM558, pEM561, and pEM560 for miR-124-1, 2 and 3, respectively in a PCR reaction of 16-20 cycles catalyzed by KAPA HiFi Polymerase (KAPA Biosystems). This was followed by the rest of the Site Directed Mutagenesis protocol described previously. For creating the splicing deficient introns, the constructs made above (i.e., the ones lacking the stem-loop structure) were used as templates along with EMO560 (5’GCGCTAGCCCTCTGGTGGCCGAGTACTCCCTCAAAAG3’) and EMO559 (5’GTAGTCTCGAGTGGACACCGGTTGGAGAGAGAAGCGCAAAGGT3’) in a PCR reaction of 16-20 cycles catalyzed by KAPA HiFi Polymerase (KAPA Biosystems). pEM1046, pEM1047, and pEM1048 are constructs with miR-124 genes without stem-loop within spliceable introns and pEM1049, pEM1050, and pEM1051 are constructs containing miR-124 genes without stem-loop within splicing deficient introns.

To generate Renilla luciferase constructs containing the pri-miR-124-2 upstream-downstream destabilization element, the destabilization sequence was amplified from the plasmid lacking the miR-124-2 stem-loop using primers EMO2934 (5’GCTGGTACCGGTACTAGTAGACTCATGCACCCGCTACT3’) and EMO533 (5’GGACTAGTATTTGGAAAAGGCCAGCACT3’) in a reaction catalyzed by KAPA HiFi (KAPA Biosystems) in a PCR reaction of 16-20 cycles. The PCR product was then cut with Acc65I and SpeI (New England Biolabs® Inc). pEM831 (plasmid containing the Renilla luciferase gene) was cut with Acc65I and XbaI and the digested products were run on a gel and
purified using NucleoSpin® Gel and PCR Clean-Up (Macherey-Nagel). The products were then ligated at 4°C for 16 hours. pEM1158 is the construct which contains Renilla luciferase with miR-124-2 "destabilization context" sequence in the 3'UTR.

To generate constructs encoding endogenous pri-miR-124-2 intron, primers EMO1563 (5’AAATACTAGTAACCTGCTTTTGAGCCCTTT3’) and EMO1564 (5’ATCTCTCAGAGCAAAGATGTGTCTGTCTTTCA3’) were used for amplifying intron flanked by exon 2 and exon 3 using a Bacterial Artificial Chromosome as the template. The PCR reaction was carried out using KAPA HiFi Polymerase (KAPA Biosystems) in a reaction of 16-20 cycles. The amplicons were then digested with NheI and XhoI (New England Biolabs® Inc). pEM561, which was used as the vector for insertion of the amplicon, was also digested with NheI and XhoI. The vector and amplicon digested products were run on an agarose gel, purified using the NucleoSpin® Gel and PCR Clean-Up (Macherey-Nagel) and ligated for 16 hours at 4°C.

Constructs encoding endogenous pri-miR-124 intron with mutated splice sites were generated by using primers EMO1940 (5’CTCCGGCGGTA GAGCGGAACCTCCTCATT3’) and EMO1941 (5’GAATGAGGGGAG TTGGGCTCTAGGCCGAGGAG3’) for embryonic exon 5’ splice site mutation, primers EMO1938 (5’CTCTCCGAGCTCAGGAGGGCTGTTGAGTTA ATTTTTCT3’) and EMO1939 (5’GAAAAATATTACATCAACACGGCCTG CCGAGTCCCAGGAG3’) for adult exon 5’ splice site mutation. For mutating the 3’ splice site, primers EMO1936 (5’GAAAAAGTTAAACGT TTTGTATTTATGTATTCATCCGGTTATATGGAGTTAATAATGTAG AATGC3’) and EMO1937 (5’GCATTCTACATTATTATCTCCATATA AGCGGATGAAACATAAAATTCAAAAAACGTTTAATTCTCTTTCT3’) were used. The PCR reaction was carried out using KAPA HiFi Polymerase (KAPA Biosystems) in a reaction of 16-20 cycles followed by the rest of the Site Directed Mutagenesis protocol described above. The process of creating adult and embryonic exon mutants in the same construct is a three step process where first the adult splice site was mutated followed by mutation of the embryonic splice sites and finally mutation of the 3’ splice sites. pEM926 (endogenous miR-124 2 gene), pEM1006 and pEM1007 (endogenous miR-124-2 gene with
respective splice sites mutated - adult and embryonic - along with the 3’ splice sites) and pEM1031 (endogenous miR-124-2 gene with both adult and embryonic splice sites mutated) were the constructs generated for this experiment.

To generate constructs containing endogenous pri-miR-124-2 intron with inactivated adult or embryonic 5’ splice site, pEM926 was used as a template and EMO1938 (5’CTCTCCCCGGGACTCGACCCGGCCTGTTGAGTTAATTTTTC3’) and EMO1939 (5’GAAAAATTAACTCAACAGCGGGCTGCCAGTCCCGGGAGAGGAG3’) were used as primers for mutating the adult form splice site and EMO1940 (5’CTCCGCCGGTAGACCCCAACTCCCTCATTCC3’) and EMO1941 (5’GAATGAGGGAGTTGGGCCTCTACGGGCGGAG3’) were used as primers for mutating the embryonic form splice site. KAPA HiFi (KAPA Biosystems) was used as the enzyme in a PCR reaction of 16-20 cycles. The PCR products were then processed using the site directed mutagenesis protocol mentioned above. pEM1255 and pEM1256 (endogenous miR-124-2 gene with the 5’-splice site of adult form and embryonic form mutated, respectively) were the constructs made for this study.

To generate Renilla luciferase constructs containing the stem-loops (pre-miRNA sequences) of mouse miR-1 and mouse miR-153 into the construct containing upstream-downstream destabilization element of pre-miR-124-2 as well as the construct containing Renilla luciferase 3’ UTR only, the pre-miRNA sequences of miR-1 and miR-153 were amplified from plasmids pEM1003 and pEM1002, respectively. Primers EMO3530 (5’AACCGGTACCTCACTGTCTTTGGTGAGTGA3’) and EMO3565 (5’GACACATCTAGAACCCTTACGGCCT3’) were used for amplifying pre-miR-1 and EMO3532 (5’AACCGGTACCAGCCACAAGATGTAA3’) and EMO3566 (5’CCACTCTAGATACCCAGAGCGGAGTTGCCT3’) were used for amplifying pre-miR-153 in a reaction catalyzed by KAPA HiFi (KAPA Biosystems) in a PCR reaction of 16-20 cycles. The PCR products and the vectors (pEM831 and pEM1158) were then cut with KpnI and XbaI (New England Biolabs® Inc) and the digested products were run on a gel and purified using NucleoSpin® Gel and PCR Clean-Up (Macherey-Nagel). The products were then ligated at 4°C for 16 hours. pEM1247 and pEM1249 were the constructs containing miR-1 and miR-
153 in the 3’ UTR of Renilla luciferase and pEM1250 and pEM1252 were the constructs with miR-1 and miR-153 in the upstream-downstream element which is present in the 3’UTR of Renilla luciferase.

Table 2.1: Name and description of plasmids used for Experiments.

<table>
<thead>
<tr>
<th>Plasmid Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>pEM222</td>
<td>Parent plasmid containing intron module cloned under mPGK promoter</td>
</tr>
<tr>
<td>pEM183</td>
<td>Mouse miR-124-1 expressed under CMV promoter; used for amplification of 124-1 gene to be cloned under mPGK</td>
</tr>
<tr>
<td>pEM208</td>
<td>Mouse miR-124-2 expressed under CMV promoter; used for amplification of 124-2 gene to be cloned under mPGK</td>
</tr>
<tr>
<td>pEM209</td>
<td>Mouse miR-124-3 expressed under CMV promoter; used for amplification of 124-3 gene to be cloned under mPGK</td>
</tr>
<tr>
<td>pEM579</td>
<td>eYFP, 3× HA and rabbit globin 3’ UTR clones downstream to miR-124-1</td>
</tr>
<tr>
<td>pEM581 and 580</td>
<td>mCerulean, 3× Myc and rabbit globin 3’ UTR clones downstream to miR-124-2 within spliceable and splicing defective intron, respectively</td>
</tr>
<tr>
<td>pEM582</td>
<td>dsRed, 3× Flag and rabbit globin 3’ UTR clones downstream to miR-124-3</td>
</tr>
<tr>
<td>pEM613</td>
<td>Final construct containing 124-1 for injection into ES cells</td>
</tr>
<tr>
<td>pEM614</td>
<td>Final construct containing 124-2 for injection into ES cells</td>
</tr>
<tr>
<td>pEM615</td>
<td>Final construct containing 124-3 for injection into ES cells</td>
</tr>
<tr>
<td>pEM558</td>
<td>124-1 within spliceable intron</td>
</tr>
<tr>
<td>pEM560</td>
<td>124-2 within spliceable intron</td>
</tr>
<tr>
<td>pEM561</td>
<td>124-3 within spliceable intron</td>
</tr>
<tr>
<td>pEM647</td>
<td>124-1 within splicing deficient intron</td>
</tr>
<tr>
<td>pEM559</td>
<td>124-2 within splicing deficient intron</td>
</tr>
<tr>
<td>pEM596</td>
<td>124-3 within splicing deficient intron</td>
</tr>
<tr>
<td>pEM744</td>
<td>124-2 with the loop of 124-1</td>
</tr>
<tr>
<td>pEM745</td>
<td>124-2 with the loop of 124-3</td>
</tr>
<tr>
<td>---------</td>
<td>-----------------------------</td>
</tr>
<tr>
<td>pEM543</td>
<td>Vector control containing no miR-124 gene</td>
</tr>
<tr>
<td>pEM781</td>
<td>mini-124-1 gene within spliceable intron</td>
</tr>
<tr>
<td>pEM782</td>
<td>mini-124-2 gene within spliceable intron</td>
</tr>
<tr>
<td>pEM783</td>
<td>mini-124-3 gene within spliceable intron</td>
</tr>
<tr>
<td>pEM931</td>
<td>mini-124-1 gene within splicing deficient intron</td>
</tr>
<tr>
<td>pEM932</td>
<td>mini-124-2 gene within splicing deficient intron</td>
</tr>
<tr>
<td>pEM933</td>
<td>mini-124-3 gene within splicing deficient intron</td>
</tr>
<tr>
<td>pEM953</td>
<td>5’ region of miR-124-2</td>
</tr>
<tr>
<td>pEM954</td>
<td>3’ region of miR-124-2</td>
</tr>
<tr>
<td>pEM958</td>
<td>5’ region of miR-124-2 in splicing deficient intron</td>
</tr>
<tr>
<td>pEM959</td>
<td>3’ region of miR-124-2 in splicing deficient intron</td>
</tr>
<tr>
<td>pEM1046</td>
<td>miR-124-1 gene without stem-loop within spliceable intron</td>
</tr>
<tr>
<td>pEM1047</td>
<td>miR-124-2 gene without stem-loop within spliceable intron</td>
</tr>
<tr>
<td>pEM1048</td>
<td>miR-124-3 gene without stem-loop within spliceable intron</td>
</tr>
<tr>
<td>pEM1049</td>
<td>miR-124-1 gene without stem-loop within splicing deficient intron</td>
</tr>
<tr>
<td>pEM1050</td>
<td>miR-124-2 gene without stem-loop within splicing deficient intron</td>
</tr>
<tr>
<td>pEM1051</td>
<td>miR-124-3 gene without stem-loop within splicing deficient intron</td>
</tr>
<tr>
<td>pEM831</td>
<td>vector for the construct containing destabilization sequence</td>
</tr>
<tr>
<td>pEM1158</td>
<td><em>Renilla</em> luciferase with miR-124-2 &quot;destabilization context&quot; sequence in the 3'UTR</td>
</tr>
<tr>
<td>pEM926</td>
<td>Endogenous miR-124 2 gene (between exon 2 and exon 3) within spliceable intron</td>
</tr>
<tr>
<td>pEM1006</td>
<td>Endogenous miR-124-2 gene with the adult splice sites mutated along with the 3’ splice sites</td>
</tr>
<tr>
<td>pEM1007</td>
<td>Endogenous miR-124-2 gene with the embryonic splice sites mutated along with the 3’ splice sites</td>
</tr>
<tr>
<td>pEM1031</td>
<td>Endogenous miR-124-2 gene with both adult and embryonic splice sites mutated</td>
</tr>
<tr>
<td>pEM1255</td>
<td>Endogenous miR-124-2 gene with the 5’-splice site of adult form mutated</td>
</tr>
</tbody>
</table>
| pEM1256 | Endogenous miR-124-2 gene with the 5’-splice site of
2.2. Cloning Primers Used

Table 2.2: Primers used for cloning.

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Sequence; 5’ to 3’</th>
</tr>
</thead>
<tbody>
<tr>
<td>EMO578 (LoxP_F1)</td>
<td>GTCGACAATATAAATCTTCGTATAATGTATGTA TGCTATACGAAGTTATTAG</td>
</tr>
<tr>
<td>EMO579 (LoxP_R1)</td>
<td>CTAGCTAATAACCTTCGTATAGCATACA TTATACGAAGTTATTATATTGTCGACTGCA</td>
</tr>
<tr>
<td>EMO580 (intron_F1)</td>
<td>GCGCTAGCCCTGTTGGGTAGTACTGAGTACT CCCTCCTCAA</td>
</tr>
<tr>
<td>EMO581 (intron_R1)</td>
<td>GTAGTCTCGAGTGACAGCTGTGGAGA GGAAAGGCAAGG</td>
</tr>
<tr>
<td>EMO582 (Fluor_F1)</td>
<td>GCCACCATGCTGTGAGCAAGGCGA</td>
</tr>
<tr>
<td>EMO583 (Fluor_R1)</td>
<td>AAGGGGATCCGAATTCAGGTGACGGGA TCTCA</td>
</tr>
<tr>
<td>EMO560 (intron_up9)</td>
<td>GCGCTAGCCCTGTTGGCCGGAGTACT CCCTCTCAAAAG</td>
</tr>
<tr>
<td>EMO559 (intron_down9)</td>
<td>GTAGTCTCGAGTGACAGCGGTGGAG AGAAAGGCAAGG</td>
</tr>
<tr>
<td>EMO560 (intron_up9)</td>
<td>GCGCTAGCCCTGTTGGCCGGAGTACT CCCTCTCAAAAG</td>
</tr>
<tr>
<td>EMO559 (intron_down9)</td>
<td>GTAGTCTCGAGTGACAGCGGTGGAG AGAAAGGCAAGG</td>
</tr>
<tr>
<td>EMO580 (intron_F1)</td>
<td>TAAGGCACGCGGTTGAGTGA</td>
</tr>
<tr>
<td>EMO581 (intron_R1)</td>
<td>TCTGACACGGCAAGGGAAGGCA</td>
</tr>
<tr>
<td>EMO1139 (m_miR-124-2_loop1_F1)</td>
<td>CCGGTGTCAACAGCGGACCTGTGATTAA ATGTCATACAAATAGGCACGCAGGTGA ATGCAAGAG</td>
</tr>
<tr>
<td>EMO1140 (m_miR-124-2_loop1_F1)</td>
<td>CCGGTGTCAACAGCGGACCTTGATTTAA ATGTCATACAAATAGGCACGCAGGTGA ATGCAAGAG</td>
</tr>
<tr>
<td>Primer Name</td>
<td>Sequence</td>
</tr>
<tr>
<td>-------------</td>
<td>----------</td>
</tr>
<tr>
<td>2_loop1_R1</td>
<td>ATGTCCATAAATTAAGGCACGC CGTGTTGCTACAGGACCTTTGATTAA TGTCTATACAAATTAAGGCACGC CGTGTTGCTACAGGACCTTTGATTAA ATGCAAAGAG</td>
</tr>
<tr>
<td>EMO1141 (m_miR-124-2_loop3_F1)</td>
<td>CC GTGCATCCAGC CGACCG TTACTTTCGCTTCAATGAATTCGGCAGGTG</td>
</tr>
<tr>
<td>EMO1142 (m_miR-124-2_loop3_R1)</td>
<td>CTCTTGGCATTCACCGCGTGCCTTAAT TGTATAGACATTAAATCAAGGTCCGCT GTGAACACCG</td>
</tr>
<tr>
<td>EMO983 (m_miR-124-2_mut1_F1)</td>
<td>CAC TGTCTCTAGATGACATGAGATCAGAAA AAGACTCTGCTTCCGTTGACAGGC</td>
</tr>
<tr>
<td>EMO984 (m_miR-124-2_mut1_R1)</td>
<td>GCT GTGGAACACCGGAGACAGAGTCTTT TTCTGATCTTCTCTAGAAGCAAGTG</td>
</tr>
<tr>
<td>EMO985 (m_miR-124-2_mut2_F1)</td>
<td>GGTGAATGCCAAGAGCGGAGCCTCTG CACTTG AAGGACATCCG</td>
</tr>
<tr>
<td>EMO986 (m_miR-124-2_mut2_R1)</td>
<td>CGGATGTCTTCTCAAGTCAAGATCAGGGCTT GCTCTTGGCATTCACCGCGTGCCTTAAT TGTATAGACATTAAATCAAGGTCCGCT GTGAACACCG</td>
</tr>
<tr>
<td>EMO1495 (m_miR124-2_mut3_F1)</td>
<td>CAG ACTGTCTCTAGATGACATGAGATCAGAAA AAGACTCTGCTTCCGTTGACAGGC</td>
</tr>
<tr>
<td>EMO1496 (m_miR124-2_mut3_R1)</td>
<td>GCT GTGGAACACCGGAGACAGAGTCTTT TTCTGATCTTCTCTAGAAGCAAGTG</td>
</tr>
<tr>
<td>EMO2020 (miR-124-1_kink_F1)</td>
<td>CGCGGTGAATGCCAAGAATGGGGCTT ACGGTTCTTCTAGATGACATGAGATCAGAAA AAGACTCTGCTTCCGTTGACAGGC</td>
</tr>
<tr>
<td>EMO2021 (miR-124-1_kink_R1)</td>
<td>GCAGGCCCTCCTGGAACCAAGGAGGTGCTC AGAAGAAACGTGAAGCCATTTCTTGGCC TATCAGCCAG</td>
</tr>
<tr>
<td>EMO1219 (mini-124-1_F1)</td>
<td>TTCCTTTCTTCTCTTTCGAGGACAAAG</td>
</tr>
<tr>
<td>EMO1220 (mini-124-1_R1)</td>
<td>GGGCGTTGTCATATGTCGATCCTTCCG T</td>
</tr>
<tr>
<td>EMO1221 (mini-124-2_F1)</td>
<td>GTCACTTTGCTTCTAGATCAGCATAGATCA</td>
</tr>
<tr>
<td>EMO1222 (mini-124-2_R1)</td>
<td>CATTCTAGAATTGCAATGCCCACCT TCCTAA</td>
</tr>
<tr>
<td>EMO1223 (mini-124-2_F1)</td>
<td>ACCCGGTCCAGGCGCTC</td>
</tr>
<tr>
<td>EMO1224 (mini-124-3_R1)</td>
<td>CCGCGACGCATATGCTCCGCGGCGACTT</td>
</tr>
<tr>
<td>EMO1895 (mmu_miR-124-2_L5')</td>
<td>TCAACCTCAGGAGACTTCAATGACCCCGC</td>
</tr>
<tr>
<td>EMO1894 (mmu_miR-124-2_S5')</td>
<td>AGTTCCACTAGGTCACCTTCTCATAGATCA</td>
</tr>
<tr>
<td>EMO32 (intron_down1)</td>
<td>CGTGGAGAGAAGGCAAAGAG</td>
</tr>
<tr>
<td>Name</td>
<td>Sequence</td>
</tr>
<tr>
<td>------</td>
<td>----------</td>
</tr>
<tr>
<td>EMO2253 (mmu-miR-124-1_del_F1)</td>
<td>CAGGAGAAAGGCTCTCTCTCTCGAATGGGGTCTGTCGACACCTTTGGGTCC</td>
</tr>
<tr>
<td>EMO2254 (mmu-miR-124-1_del_R1)</td>
<td>GGACCCCAAGGTGCTCAGACAGCCCACTCTGCTTTCTCTTGGGTCAG</td>
</tr>
<tr>
<td>EMO2255 (mmu-miR-124-2_del_F1)</td>
<td>CTAGATCAAGATCAGAGACTCTGCTCTCGAGAGCGGAGCCTACGGCTGCACTTG</td>
</tr>
<tr>
<td>EMO2256 (mmu-miR-124-2_del_R1)</td>
<td>CAAGTGCAGCCGTAGGCTCCGCTCGAGAGCAGAGTCTCTGATCTTGATCTAG</td>
</tr>
<tr>
<td>EMO2257 (mmu-miR-124-3_del_F1)</td>
<td>CAGCCCTGAGGGCCCCTCTGGAGAGGCGCCTCCGCCGCTCCTTTC</td>
</tr>
<tr>
<td>EMO2258 (mmu-miR-124-3_del_R1)</td>
<td>GAAAGGAGCGGCGGAGGCGCCTCTCCAGAGGGGCCCTCAGGGCTG</td>
</tr>
<tr>
<td>EMO2934 (miR124-2_Kpn1_F1)</td>
<td>GCTGGTACCGGTACTAGTAGACTCATGCAAGTACCTGACAGCAGCA</td>
</tr>
<tr>
<td>EMO533 (m_miR124a_loc_down9)</td>
<td>GGACTAGTATTTTGAAAGGACCAGCACT</td>
</tr>
<tr>
<td>EMO1563 (miR124-2_endointron_F1)</td>
<td>AAATACTAGTAAACCTGGCTTTTGAGCCCCT</td>
</tr>
<tr>
<td>EMO1564 (miR124-2_endointron_R1)</td>
<td>ATCTCTCGAGAGCAAGATGTGTCCTGC TTCA</td>
</tr>
<tr>
<td>EMO1940 (124-2_endo_5'mutint_F1)</td>
<td>CTCCGCCGTAGAGCCCAACTCCCTCTCATT</td>
</tr>
<tr>
<td>EMO1941 (124-2_endo_5'mutint_R1)</td>
<td>GAATGAGGGGAGTTGGGCTCTACGGGAGG</td>
</tr>
<tr>
<td>EMO1938 (124-2_endo5'mutintA_F1)</td>
<td>CTCTCCCGGGACTCGGCAGCCGCTGTTGAGTTAATT</td>
</tr>
<tr>
<td>EMO1939 (124-2_endo5'mutintA_R1)</td>
<td>GAAAAATTAACTCAAGCGGCTGCCGGAATGAAACATAAAAATCAAAAAC</td>
</tr>
<tr>
<td>EMO1936 (124-2_endo_3'mutint_F1)</td>
<td>GAAAAGTTAAAACGTTTTTGATTTATGTTTCATTCCGCTTATGAGAAGAAAATATGTAAGATG</td>
</tr>
<tr>
<td>EMO1937 (124-2_endo_3'mutint_R1)</td>
<td>GCATTCTACATTATTTTACTCCATATAAGCGGAATGAACATAAAAATCAAAAAAGTTTAACTTTTTC</td>
</tr>
<tr>
<td>EMO3530 (m_miR1_Kpn1_F1)</td>
<td>AACCAGTACCTCAGTCTCGGTGTG</td>
</tr>
<tr>
<td>EMO3565 (m_miR1_Xba_R1)</td>
<td>GACACATCTAGAACCCTTACAGTCGGCT</td>
</tr>
<tr>
<td>EMO3532 (m_miR153_Kpn1_F1)</td>
<td>AACCAGTACCAGGCACCCAAGATGTGTA</td>
</tr>
<tr>
<td>EMO3566 (m_miR153_Xba_R1)</td>
<td>CCACTCTAGATACCAGCAAGCGAGGTGCT</td>
</tr>
</tbody>
</table>

81
Chapter 2  Materials and Methods

2.3. Molecular Cloning Techniques

2.3.1. Digestion with restriction enzymes and purification of DNA fragments

The plasmids or PCR products which have to be used for cloning were digested with restriction enzymes (New England Biolabs® Inc) using corresponding 1× buffer systems, BSA. Five units of restriction enzyme was typically used to cut 1 µg of DNA and the reactions were incubated at 37°C (or as appropriate) for 2-16 hours. The digestion products were purified using agarose gel electrophoresis and the desired bands were excised and cleaned up with a Nucleospin Gel Purification Kit (Macherey & Nagel).

2.3.2. Klenow reactions to repair recessed 3' termini

Reactions were carried out using 1× of any NEB buffer, 0.2 mM each of the four dNTPs and 5 units of DNA polymerase I Klenow fragment enzyme (New England Biolabs® Inc, 5 units/µl) per 1 µg of DNA at room temperature for 30 min. The reactions were stopped by adding 1 µl of 0.5 M EDTA, pH 8.0.

2.3.3. Ligation of DNA fragments and bacterial transformation

The purified vector and insert were mixed in a ratio of 1:3 in the presence of 1×NEB T4 DNA Ligase Buffer (New England Biolabs® Inc), 400 units of T4 DNA Ligase (New England Biolabs® Inc, 400 units/µl) in an appropriate volume made up with Millipore water. This reaction mixture is incubated at 4°C for 16 hours. Reactions were then heat inactivated at 70°C for 10 mins and chilled on ice for 10 min. 1-2 µl of the ligation mixture was then added to 25 µl of competent E. coli cells (typically TOP10 cells; Invitrogen) and incubated for
30 min on ice. The cells were then heat-shocked for 30 seconds at 42°C and chilled on ice for 1 min. Following the addition of 300 µl of LB medium the cells were then incubated at 37°C for 1 hour. 150 µl of the cell suspension was then plated onto an LB agar plate containing appropriate antibiotic and incubated at 37°C overnight to allow recombinant bacterial colonies to form.

2.3.4. Site Directed Mutagenesis

QuikChange mutagenesis protocol from Stratagene was used with several modifications. Parental DNA plasmid was “amplified” using mutagenic primers and a High Fidelity Thermostable DNA Polymerase (KAPA HiFi, KAPA Biosystems) for 16-20 PCR-like cycles. The reaction products were then ethanol-precipitated, washed with 70% ethanol and the pellet was suspended in suitable volume of water. If necessary, the parental DNA was digested with DpnI restriction enzyme for 2 hours at 37°C. The mutagenized DNA was then used to transform competent TOP 10 cells (Invitrogen).

2.3.5. Nucleic acid precipitation

An equal volume of 1:1 mixture of phenol-chloroform was added to the samples, mixed vigorously and allowed to stand at room temperature for 2 min. Acid phenol was used for RNA and Tris-HCl (pH 8.0) saturated phenol was used for DNA. The mixtures were then centrifuged at 2 to 8°C for 3 to 5 min and the aqueous phase was transferred to a new tube. 1/10<sup>th</sup> volume of 3 M sodium acetate was added to one volume of nucleic acid sample and mixed well. Three volumes of 100% ethanol was then added followed by incubation of samples at -80°C for 30 min and centrifugation at 14,000 rpm at 4°C for 10 min. Nucleic acid pellets were washed with 70-75% ethanol, air dried and rehydrated in required amount of water or TE buffer and stored at -80° or -20°C.
2.4. Cell lines

HEK293T and N2a (Neuro2a) cells were obtained from ATCC and cultured in DMEM (Hyclone) supplemented with 10% Foetal bovine serum (FBS; Hyclone), 1% Penicillin/Streptomycin (Gibco) and 1% Sodium Pyruvate (Gibco) at 37°C and 5% CO₂. The same medium without Penicillin/Streptomycin was used in transfection experiments. Cell stocks were made in the proliferation medium that was supplemented with 5% DMSO, slowly frozen at -80°C and stored in liquid nitrogen.

Embryonic stem cells were grown on mitotically inactive feeder cell layers (PMEF) on a plate coated with 0.1% gelatin (Chemicon International) in Knockout DMEM medium (GIBCO) supplemented with 15% Fetal Bovine Serum for ES cells (FBS; Hyclone), 0.1 mM MEM non-essential amino acids (GIBCO), 2 mM L-glutamine (GIBCO), 0.1 mM β-mercaptoethanol (Sigma), 1% Penicillin/Streptomycin (GIBCO) and 1000 units/ml leukemia inhibitory factor (LIF; ESGRO). Cell stocks were made in 2X freezing medium (20% DMSO, 60% complete ES cell medium and 20% FBS) by resuspending the cells to a single cell suspension.

2.5. Purification of RNA from Cell Lines and Tissues

RNA was extracted from cells and tissues using TRIzol® (Invitrogen). 1 ml of TRIzol was added to confluent cell culture in a 60 mm dish or 50-100 mg of tissue sample. The lysates were supplemented with 200 μl of chloroform and shaken vigorously at room temperature. The samples were then spun at maximum speed for 10 min at 4°C after which it separated into three phases: lower organic TRIzol-chloroform phase, interphase and upper aqueous phase. The upper aqueous phase (~500 μl) was carefully transferred to another tube containing 500 μl of 100% isopropanol. This was vigorously shaken and
incubated at -20°C for ~20 min and spun at a maximal speed at 4°C. The RNA pellets were then washed with 70% ethanol, air-dried and rehydrated in an appropriate volume of RNase-free water.

2.6. Polyacrylamide gel electrophoresis of RNA (Northern Blotting)

Denaturing polyacrylamide gels were typically prepared using the following recipe: 15% 19:1 acrylamide: bisacrylamide, 7-8 M urea, 1×TBE, 0.87 µl/ml TEMED and 0.03% ammonium persulphate made up with RNase-free water. RNA samples were prepared by mixing them with equal or larger volumes of loading buffer containing 98% formamide, 10 mM EDTA, 0.025% bromophenol blue and 0.025% xylene cyanol. Samples were electrophoresed at ~10-20 V/cm until bromophenol blue reached the bottom of the gel. RNA was then transferred to Hybond N+ membrane (Amersham) using semi-dry transfer apparatus (BioRad Transblot SD) for ~30 min at 2.5 mA/cm² in 0.5×TBE. RNA transferred to membrane was cross-linked with 254 nm UV at 150,000 µJ/cm² using the UV cross-linker unit of HL-2000 Hybri Linker (UVP Laboratories). Blocking of the membrane was carried out in pre-warmed 6×SSC, 7% SDS at 65°C overnight. Oligonucleotide probes were labeled with γ³²-ATP in a T4 Polynucleotide Kinase reaction (see below) and hybridized overnight at 42°C. Blot was washed thrice with 3×SSC, 0.1% SDS, ~10 min each time, at 42°C. Membranes were washed with several changes of 3×SSC, 0.1% SDS and exposed to a Phosphorimage screen for 2 hours to overnight and scanned using a Typhoon Trio Variable Mode Imager (GE Healthcare). This protocol was followed for detecting both miRNAs and U6 snRNA.
2.7. 5’ end labeling by T4 Polynucleotide Kinase reaction

Reactions containing 1×T4 PNK buffer, 1.5 µM of DNA or RNA oligonucleotide, 3 µl of ([γ]-32P)ATP (Perkin Elmer, 6,000 Ci/mmol), 11 µl of Millipore water, and 1 µl of T4 PNK enzyme (10,000 units/ml, NEB) (~20 µl total volume) were incubated at 37°C for 30 to 60 min followed by dilution to 50 µl with Millipore water. Labeled oligonucleotides were purified using G-25 spin columns (Geneaid) according to the manufacturer’s instructions.

2.8. RT qPCR (Reverse Transcriptase quantitative Polymerase Chain Reaction)

2.8.1. DNase treatment:

RNA samples were treated with DNase RQ (Promega) prior to subsequent RT-qPCR analyses to remove possible genomic DNA contamination. The reaction contained: up to 20 µg of RNA, 1×DNase buffer, 5 units of RNase-free DNase (1 unit/µl; Promega), 40 units of RNasin (40 units/µl; Promega) made up to 100 µl with DEPC water. Samples were incubated at 37°C for 1 hour. The samples were then extracted with acid phenol-chloroform and precipitated as described above. DNA-free RNA pellets were normally dissolved in ~30 µl of DEPC water and the RNA concentrations were determined using a Nanodrop spectrophotometer (Thermo Scientific).
2.8.2. Reverse Transcription

2.5 µg of DNase-treated RNA was mixed with 1 µl of a 100 µM random decamer (N10) primer and made up to 5.5 µl with DEPC water followed by incubation at 70°C for 10 min. Then 4.5 µl of the reaction mix containing 1×RT buffer, 1 mM dNTPs, RNasin (Promega), Superscript III RT enzyme (Invitrogen), 10 mM DTT was added to make up the final volume of 10 µl and incubated at 50°C for 1.5 hours followed by incubation at 70°C for 10 min. RT-negative controls were prepared in a similar manner, except that no Superscript enzyme was added in this case. Both RT-positive and negative reaction products were diluted to 100 µl with DEPC water and placed at -80°C for prolonged storage.

2.8.3. Quantitative PCR

Quantitative analysis of gene expression by real-time PCR was carried out using a Step One Plus Real Time System (Applied Biosystems). Each Real-Time PCR reaction (20 µl) was made up of 5 µl of diluted cDNA, 10 µl of 2×Fast SYBR Green Master Mix (Applied Biosystems; or the similar product from KAPA Biosystems) and 0.15 µM of each forward and reverse primer. All reactions were performed using the following thermal cycler conditions: 95°C for 20 min, followed by 45 cycles of a three-step reaction; denaturation at 95°C for 5 seconds, annealing and extension at 60°C for 45 seconds. The reaction was followed by a melting curve analysis to ensure amplification specificity. Transcript levels were normalized to HPRT or GAPDH mRNA levels. Relative fold change in expression was calculated using the “ΔΔCt” method as recommended by Applied Biosystems.
2.8.4. Primers Used for Real Time PCR

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>EMO999 (hDrosha_F1)</td>
<td>TTTGAAGAAGCAATTGGAGTAATT TT</td>
</tr>
<tr>
<td>EMO1000 (hDrosha_R1)</td>
<td>CTTTCATGATGATCTGGGAAT</td>
</tr>
<tr>
<td>EMO1001 (hDDX5_F1)</td>
<td>GAAACCCCTGGGGAGAAATTAGTTA</td>
</tr>
<tr>
<td>EMO1002 (hDDX5_R1)</td>
<td>TTCGGGCAATTTGACCT</td>
</tr>
<tr>
<td>EMO1005 (hDDX17_F1)</td>
<td>AAAACTGATGCAGCTTTGTGGA</td>
</tr>
<tr>
<td>EMO1006 (hDDX17_R1)</td>
<td>ACTTCCATAGCCACTGCCATT</td>
</tr>
<tr>
<td>EMO927 (hDGCR8_F1)</td>
<td>GAGCCTGACTCTATGGGTGCT</td>
</tr>
<tr>
<td>EMO928 (hDGCR8_R1)</td>
<td>CTCAGCCAAAGTCTGAATTT</td>
</tr>
<tr>
<td>EMO249  (hHPRT1-Primerbank _F)</td>
<td>TGGACAGGACTGAACGTCTTGT</td>
</tr>
<tr>
<td>EMO250  (hHPRT1-Primerbank _R)</td>
<td>CCAGCAGGTCAGAAAGAATTTA</td>
</tr>
<tr>
<td>EMO2810 (IRES_F2)</td>
<td>AATCCCGGTTATTTTCCACCA</td>
</tr>
<tr>
<td>EMO2811 (IRES_R1)</td>
<td>AGAGGAACTGTCTTCCACG</td>
</tr>
<tr>
<td>EMO2912 (miR-124_intron_F1)</td>
<td>CTCAAAAAGCGGGCATGACT</td>
</tr>
<tr>
<td>EMO2914 (miR-124_intron_R1)</td>
<td>TTGTCTTCTTGACAGATGGAC</td>
</tr>
<tr>
<td>EMO1262 (REL-F1)</td>
<td>ACAAGTACCTCCACCGCTTGG</td>
</tr>
<tr>
<td>EMO1263 (REL-R1)</td>
<td>GACACTCTCAGCATGGACGA</td>
</tr>
</tbody>
</table>

2.9. Multiplex PCR

Whenever required, cDNA samples prepared as described above were also analyzed by multiplex PCR. We typically used 4 µl of diluted cDNA for 20 µl PCR mixtures. Other reaction components included 1×KapaTaq Buffer, 200
µM dNTPs, and appropriate primers added to a final concentration of 0.4 µM, and 0.4 units of KapaTaq DNA Polymerase (KAPA Biosystems). All reactions were performed using the following thermal cycler conditions. Initial denaturation at 95°C for 3 min, followed by 35 cycles of a 3-step reaction: denaturation at 95°C for 30 seconds, annealing at 60°C for 1 min followed by extension at 72°C for 1 min, and final extension is carried out at 72°C for 3 min. The PCR products are then run on a 2% agarose gel and photographed using a Bio-Rad Gel Doc system.

2.9.1. Primers used for Multiplex PCR

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>EMO1208 (multiplex_F1)</td>
<td>CGACCTGCAGTCGACAATATAA</td>
</tr>
<tr>
<td>EMO1209 (multiplex_R1)</td>
<td>AAGACGGCAATATGGTGGAA</td>
</tr>
<tr>
<td>EMO1210 (multiplex_R2)</td>
<td>AACAGCTTGACACAACAAAAAGATTGT</td>
</tr>
<tr>
<td>EMO3016 (multiplex_endo124_F1)</td>
<td>TGAGCCCTTTGGCTTTCTT</td>
</tr>
<tr>
<td>EMO3017 (multiplex_endo124_R1)</td>
<td>TGTGTCTGCTTCACAACATGC</td>
</tr>
<tr>
<td>EMO3259 (multiplex_endo124_F2)</td>
<td>CGGAAAAATCTATATGAAATGAAAAA</td>
</tr>
</tbody>
</table>
2.10. Western Blotting

2.10.1. Protein isolation

Cell cultures were washed twice with ice-cold 1× Dulbecco PBS (PAA). Cellular proteins were extracted using NP40 lysis buffer containing 20 mM Tris-HCl, pH 7.4, 150 mM NaCl, 5 mM EDTA, 1% NP40, 10% glycerol, 1 mM PMSF and 1×Complete™ Protease Inhibitor Cocktail (Roche). The lysates were cleared by centrifugation at 14,000 rpm at 4°C in a table-top microcentrifuge and the protein-containing supernatants were collected for further analyses.

2.10.2. Determining Protein Concentration

To determine protein concentrations 25 µl diluted protein samples were mixed with 200 µl BCA assay reagent (Thermo Scientific) containing freshly mixed Reagent A and Reagent B (50:1 v/v ratio). BSA standards (from 0 to 2 µg/µl) were also assayed alongside with the experimental protein samples to obtain a standard curve. The samples were mixed well with the BCA reagent, incubated for 30 min at 37°C and quantified using a TECAN plate reader. Standard curve was plotted for the absorbance values of BSA standards. Unknown protein concentrations in experimental samples were determined by comparing their absorbance values with the standard curve.

2.10.3. SDS PAGE (Sodium Dodecyl Sulfate- PolyAcrylamide Gel Electrophoresis)

4-20% gradient precast SDS-PAGE gels (Thermo Scientific) were used. Equal amounts of protein samples were made up to equal volumes with 5×Sample Buffer (0.312 M Tris-HCl, 2% SDS, 25% β-mercaptoethanol and 0.05%
Chapter 2 | Materials and Methods

bromophenol blue) and 1×Sample Buffer followed by heating for 5 min at 98°C in a dry bath. The denatured samples were loaded onto the gel and electrophoresed in 1×gel running buffer (25mM Tris, 192mM glycine, and 0.1% SDS at pH~8.8) at 100 mV for 60 min until the dye was seen coming out of the gel. Bio-Rad Precision Plus protein marker was used as a reference ladder for size estimation.

2.10.4. Transfer

Following SDS-PAGE separation, proteins were transferred from the gel to a membrane (typically nitrocellulose membrane from Whatman) at 100 to 125mV or 0.65mA/cm² for 1.5 to 2 hours in 1×Transfer buffer (25mM Tris, 192mM glycine, and 10% methanol) using a “wet transfer” apparatus from Bio-Rad.

2.10.5. Immunostaining

After the transfer, membrane was washed once with water, stained with Ponceau S (0.5% Ponceau S in 5% acetic acid) for 3 min and washed in water to visualize the red colour protein banding pattern. Membrane was destained by washing it with 1×TBST (10 mM Tris-HCl, 100 mM NaCl and 0.1% Tween 20) for 2 to 3 times until red banding pattern of Ponceau S staining disappeared. The membrane was then blocked with 5% non-fat milk in 1×TBST for 1 hour at room temperature. Membrane was then washed with 1×TBST thrice for 5 min each to remove traces of milk. Overnight probing with a primary antibody (dilution 1:2000) in TBST with 1% BSA and 0.05% sodium azide was carried out at 4°C. The membrane was then washed with 1×TBST thrice, each for 5 min followed by incubation with of an appropriate HRP-conjugated secondary antibody in TBST with 5% milk (no sodium azide) for 1 hour at room
temperature. Membrane was finally washed thrice with 1×TBST, each for 5 min.

2.10.6. Enhanced ChemiLuminescence (ECL) detection

To detect HRP-labeled protein bands, equal volumes of Detection Reagent 1 (peroxide solution; Thermo Scientific) and Detection Reagent 2 (luminol enhancer solution; Thermo Scientific) were mixed just before use. 2 to 3 ml of the above mixture was spread on each membrane and incubated for 3-5 min. The membrane was drained dry and wrapped in Saran Wrap, followed by exposure with X-ray films for 10 seconds to 30 min.

2.11. Luciferase Assay

The Dual-Glo™ Luciferase Assay System (Promega) was used to assay luciferase activity in cells grown in 96-well plates. This system includes two reagents that allow step-wise detection of firefly and Renilla luciferase expression levels. Of these, the first one called Dual-Glo™ Luciferase Reagent can be added directly to cells in growth medium without washing or preconditioning. This reagent induces cell lysis and acts as a substrate for firefly luciferase, producing a stable luminescent signal that can be read over a period of two hours, with greater than 60% of the luminescent signal retained. Addition of the second reagent called Dual-Glo™ Stop & Glo® quenches the luminescence from the firefly reaction by at least 10,000-fold and provides the substrate for Renilla luciferase in a reaction that can also be read within 2 hours (with a similar retention in signal). In practice, 2-3 hours before transfection ~30,000 cells (HEK293T or N2a) were plated per well of a 96-well plate in antibiotic free media using a multichannel pipette in a total volume of 100 µl. After transfection, the cells were incubated at 37°C for 24 hours. 50 µl media
was removed from each well and 80 µl of Dual-Glo™ Luciferase Reagent was added to the cells for measuring firefly luciferase activity. The plate was incubated for about 10 min and the firefly luminescence was measured using a plate luminometer. To measure Renilla luciferase activity, 80 µl of Dual-Glo™ Stop & Glo® Reagent was then added into each well and mixed. The plate was again incubated for at least 10 min and Renilla luminescence was measured with luminometer. The ratio of the luminescence from the experimental reporter to luminescence from the control reporter was calculated and normalized to appropriate control transfection wells.

### 2.12. PrestoBlue Cell Viability Assay

PrestoBlue™ Cell Viability Reagent is a ready-to-use reagent for rapidly evaluating the viability and proliferation of a wide range of cell types. PrestoBlue™ reagent is quickly reduced by metabolically active cells, providing a quantitative measure of viability and cytotoxicity in as little as 10 min. Cells to be tested for proliferation were seeded using a multichannel pipette in a 96-well plate with each well containing ~30,000 cells in a volume of 100 µl. The plate was incubated for 24 hours and PrestoBlue™ Reagent was added directly to the cells in culture medium. The PrestoBlue™ Reagent is supplied as a 10× solution and has to be added according to the volume of medium in the wells. The plate was incubated for at least 10 min at 37°C and readings were taken in a fluorescence spectrophotometer with an excitation filter of 540-570 nm and emission filter of the range 580-610 nm. Higher fluorescence values in this assay were interpreted as greater total metabolic activity and therefore higher number of viable cells.
2.13. Electroporation of constructs into ES cells

ES cells were grown to an optimal density and washed once with PBS and a single cell suspension was generated using 0.25% Trypsin/EDTA followed by the addition of ES medium and pipetting up and down vigorously. Cells were spun down at 1,000 rpm for 3 min and the pellet was washed twice with PBS. 100 µg of double-stranded DNA in PBS was taken (clean and free from organic compounds and salt) and added to cell suspension in PBS. This solution was transferred to an electroporation cuvette and pulsed in Gene Pulser II (Bio-Rad) at 250 V, 500 µF, 6-7 seconds, at room temperature. After electroporation, the cells were incubated at room temperature for 5 mins and then transferred to ES medium followed by plating onto 0.1% gelatin coated plates containing PMEF monolayer. The electroporated cells were grown initially for 24 hours without selection agent before using an antibiotic containing medium.

2.14. Selection and Colony Picking of ES cells containing desired constructs

Twenty four hours after electroporation, G418 was added to the ES medium to 200 µg/ml and the electroporated cells were allowed to form colonies. After 5-6 days of selection, small round ES colonies were visible on the dish. Ideal colonies were lifted from the dish using a pipette with a 200 µl tip. Each colony was placed into one well of a 96-well plate with no feeder cells. Once all colonies were picked, the colonies were trypsinized using Trypsin/EDTA followed by the addition of ES medium to stop the reaction. Pipetting was done to mechanically dissociate colonies, while monitoring the process under a microscope, to obtain a single cell suspension. All the colonies were transferred to corresponding wells of another 96-well plate containing feeder cells.
2.15. PCR screen for targeted ES cell clones

The newly picked ES cell colonies were trypsinized in 96-well plates for 15 min at 37°C. Growth media was added to stop trypsinization followed by pipetting up and down several times and then aliquoting into 2 halves one to 96-well feeder plates and the other half to PCR plates. The PCR plates were spun down at 2,000 rpm for 5 min. Medium was decanted and lysis buffer (10% Tween 20, 10 µl/ml PK and 10× buffer from Expand PCR Kit (Roche Biosystems) was added to each well. The plate was incubated at 55°C for 6 hours followed by 95°C for 15 min to denature PK. Water was added to dilute the cell lysates in each well. A 2×PCR reaction mix was made containing: 10×Expand Long Template PCR System Buffer 3, dNTPs to a final concentration of 0.2 mM, primers to a final concentration of 0.1 µM, and Expand Long Template Enzyme Mix (5 units/µl). The thermal cycler conditions used were: initial denaturation at 94°C, denaturation at 95°C, annealing at 55°C and extension at 68°C, followed by final extension at 68°C. Equal volume of each diluted cell lysate was added to each well followed by addition of PCR mix. Correct clones were then cryopreserved using the abovementioned method for freezing ES cells.

2.16. Isolation of Genomic DNA from ES cells

The ES cells were washed with 1X PBS and trypsinized with 0.25% Trypsin/EDTA. ES-DMEM (without LIF and G418) was added and mixed properly to obtain a single cell suspension. The cell suspension was then added to a 0.1% gelatin coated plate and incubated for 30 min at 37°C. The supernatant containing the ES cells (since the MEFs stick faster to plate) was taken and spun down at 500 rpm for 4 min. The cells were resuspended with 1X gDNA Lysis Buffer (100 mM Tris-HCl pH 8.0, 200 mM NaCl, 5 mM EDTA,
0.2% SDS, dH₂O) mixed with 200 µg/ml of Proteinase K and incubated overnight at 55°C. Equal volume of Tris saturated Phenol was added and shaken vigorously till an emulsion was formed and spun at maximum speed for 10 min at RT. The aqueous layer was taken out and equal volume of Phenol:Chloroform (1:1) was added followed by shaking and spinning at maximum speed for 10 min at RT. The aqueous layer was taken and added to an equal volume of chloroform and then shaken and spun at maximum speed for 5 min. The aqueous layer was taken out and added to 2.5 volumes of 100% ethanol and 1/10th volume of 3 M sodium acetate followed by gentle mixing and incubation at -20°C for 1 hr. It was then spun at maximum speed for 15 min at 4°C. The pellet was then washed with 70% ethanol, air dried and resuspended in TE Buffer.

2.17. Southern Blotting

20 µg of genomic DNA was digested with 200 units of NcoI restriction enzyme overnight. The digested fragments were concentrated using phenol:chloroform (1:1) extraction and ethanol precipitation and finally rehydrated in TE Buffer (pH 8.0). The DNA was run on 0.7% Agarose gel prepared in 1×TAE at 5 V/cm. After the run, the gel was stained with EtBr for 10 min and photographed. DNA was denatured by soaking it for 45 min at room temperature in 10 gel volumes of denaturation solution (0.5 M NaOH and 1.5 M NaCl) with gentle rocking followed by brief rinsing in dH₂O and neutralization by soaking for 20 min at room temperature in 10 gel volumes of neutralization buffer (0.5 M Tris-Cl, pH 7.4, 1.5 M NaCl, 1 mM EDTA) with gentle rocking. The neutralization buffer was changed and the gel was soaked for a further 20 min. The gel was then washed with 10×SSC for 10 min. The cut Hybond N+ membrane was floated on dH₂O until it was completely wet from beneath and then immersed in 10×SSC (150 mM sodium citrate and 1.5 M NaCl, pH 7.6) for 5 min. Two pieces of wet blotting paper, longer and wider
than the gel, were layered in an agarose gel tank to form a support for the gel and
smoothened to get rid of air bubbles. The tank was filled with 10×SSC and the
gel was placed on the wet blotting paper while make sure there are no air
bubbles. The gel was surrounded with parafilm and the wet membrane was
placed on top of the gel. The blotting papers were immersed in 10×SSC and
placed on top of the wet membrane. A paper towel stack was placed on top and
pressed down by a weight. The DNA was transferred overnight. Transferred
DNA was crosslinked to the membrane with 254 nm UV light at 1500 µJ/cm².
The membrane was dried thoroughly and then soaked in 10×SSC followed by
incubation at 42°C in Pre-hybridization buffer (ExpressHyb™ hybridization
buffer; Clonetech, USA), supplemented with 500 µl of 20% SDS, 500 µl of
FISH Sperm DNA (5 mg/ml in TE denatured for 10 min at 100°C followed by
incubation in ice for 2 min) for 4 hours. To prepare the probe, 50 ng of DNA in
dH₂O was denatured by boiling for 5 min followed by incubation in ice for 2
min. Reaction mix was prepared using Amersham Megaprime Kit containing:
denatured DNA, 10×Reaction Buffer, 0.1 µM random primer (final
concentration), γ-P³²-dCTP (Perkin Elmer, USA) and 5 units of Klenow
enzyme (New England Biolabs® Inc; 5 units/µl). Incubation was done for 1
hour at 37°C and the DNA probe was purified using a G-50 column by spinning
it for 2 min at 3,000 rpm. The probe was then boiled, chilled on ice and then
added to the Hybridization buffer, supplemented with 500 µl 20% SDS and 500
µl FISH sperm DNA (5 mg/ml in TE denatured for 10 min at 100°C followed
by incubation on ice for 2 min) and incubated overnight at 42°C. The membrane
was then washed three times with Wash buffer (0.5×SSC, 1% SDS) for 10 min
each at 42°C and then phosphorimaged using Typhoon Trio variable Mode
Imager (GE Healthcare).
2.18. Preparation of ES cells for blastocyst injection

Subconfluent ES cells (~50% confluency) were grown on tissue culture dish with MEF feeder layer. The cells were washed once with 1×PBS and trypsinized using 0.25% Trypsin/EDTA. The cells were then resuspended by pipetting them gently up and down for 10 times in ES medium to ensure a single cell suspension. The suspension was placed in a pre-gelatinized plate (0.1% gelatin) and incubated for 10-15 min at 37°C. Floating ES cells were collected leaving the MEFs attached to the dish (MEFs attach faster to the plate). The cell suspension was spun at 1,000 rpm for 5 min and resuspended in ES complete medium without G418. The cells were placed on ice and sent to Biopolis for injection into blastocysts.

2.19. Estimating mRNA half life

N2a cells were transfected with pEM831 or pEM1158 plasmids overnight to allow the expression of the corresponding RLuc mRNAs. On the next day, the RNA polymerase II inhibitor actinomycin D (Sigma) was added to cultures to the final concentration of 5 µg/ml and the total cellular RNA was extracted with TRIzol® at 0-24 hour time points. RT–qPCR with RLuc-specific primers EMO1262 (5’ACAAGTACCTCACCGCTTGG3’)/EMO1263 (5’GACACTCTCAGCATGGACGA3’) was used to follow the degradation time course of RLuc mRNA. The data was normalized to the mGAPDH mRNA RT–qPCR signals using primers EMO1134 (5’AAATGGGGTGAGGCCGGTGC3’)/EMO1135 (5’ATCGGCAGAGGGGCGGAGA3’) and was plotted as a function of the actinomycin D treatment time and fitted to a biphasic exponential decay function to determine the apparent mRNA half-lives.
2.20. Tail Typing

A short fragment (~0.5 cm) of the tip of the tail was cut using sterilized scissors and placed into a 1.5 ml eppendorf tube. 200 µl of lysis buffer (containing Proteinase K; 20.2 mg/ml) was added to it and incubated overnight at 55°C with shaking. The samples were then heated at 95°C for 10 min to inactivate Proteinase K and chilled at ice for at least 5 min. The tubes were then spun at maximum speed for 5 min to pellet down hair and debris. The supernatant was transferred to a new eppendorf tube containing 200 µl isopropanol. The tubes were mixed well and incubated at room temperature for 5-10 min until the DNA precipitated. They were spun at maximum speed for 10 min. The pellets were washed with 70% ethanol, air-dried and rehydrated in 50 µl of TE Buffer, pH 8.0.

2.21. Non-Radioactive *In Situ* Hybridization for Embryonic Sections

The non-radioactive in-situ hybridization of embryonic sections with DIG-ddUTP labelled probe is a 4 day long sequential process to detect miRNA expression in the sections.
2.21.1. Non-Radioactive Labeling Of LNA Probe

The LNA oligos used for *in-situ* hybridization of frozen mice brain sections were labeled with Digoxigenin (DIG)-ddUTP using the DIG Oligonucleotide 3’-End Labeling Kit, 2nd Generation (Roche Applied Science). 100 pmol of LNA Detection probe (Exiqon) was taken from the stock concentration of 25 µM (or 25 pmol/µl), and water was added to adjust the final volume to 10 µl. This mixture was further supplemented with the following reagents:

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Volume</th>
<th>Final Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reaction buffer (5X)</td>
<td>4 µl</td>
<td>1X</td>
</tr>
<tr>
<td>CoCl₂-solution</td>
<td>4 µl</td>
<td>5mM</td>
</tr>
<tr>
<td>DIG-ddUTP solution</td>
<td>1 µl</td>
<td>0.05 mM</td>
</tr>
<tr>
<td>400 U terminal transferase</td>
<td>1 µl</td>
<td>20 U/µl</td>
</tr>
</tbody>
</table>

The above mixture was incubated at 37°C for 1 hour, after which it was stopped by adding 2 µl of 0.2 M EDTA. The probe was then purified using a Sephadex G25 column by spinning the column at 1,000g for 4 min and then collecting the flow-through containing the labeled probe. The approximate concentration of probe that was obtained was 30 ng/µl. The labeled probe was then aliquoted and stored at -80°C for future use.

2.21.2. Day 1: Preparation and sectioning of embryos

The embryos were dissected in PBS and it was ensured that all the membranes were carefully removed. The embryos were transferred to embedding molds (Thermo Scientific) and all the extra PBS was removed using an automatic pipette. The mold was then flooded with excess of OCT and the embryo gently swirled to ensure that any remaining water/PBS was moved away from the embryo. This was then dipped into a container containing 100% ethanol and the
entire setup was frozen in a crushed dry ice/ethanol bath. Frozen blocks were stored at -80°C or sectioned immediately using a cryostat into 10-15 µm thick slices.

2.21.3. Day 2: Preparation and hybridization of sections

The boundary of the sections was marked using the Liquid Blocker Super PAP Pen (Electron Microscopy Sciences). The sections were post-fixed in 4% paraformaldehyde (PFA) in PBS for 10 min followed by two PBS washes for 5 min each. Starting with the PBS wash, all the steps henceforth were carried out in 50 ml Coplin Jars. The excess PBS was drained off and then sections were incubated with 2 µg/ml of Proteinase K in PBS for 5 min. The proteinase solution was drained off and the slides were rinsed with PBS once for 5 min and then fixed again in 4% PFA for 5 min followed by three PBS washes of 5 min each. They were then dehydrated in 70% ethanol for 2 min and then in 95% ethanol for 2 min and then air dried. 2 ml of the labelled probe was taken and added to 1 ml of filtered hybridization solution and heated at 80°C for 2 min. A humidified box was prepared containing tissues soaked in DEPC-water. The slides were placed horizontally in this box, the sections were covered with 200 ml of the hybridization solution (containing the probe) and parafilm coverslips were lowered onto the surface of the sections carefully to avoid trapping air bubbles. The control sample had no probe added to the hybridization solution. The humidified box was sealed and hybridization was carried out overnight at 55°C.

Table 2.6: The components of the Hybridization Solution.

<table>
<thead>
<tr>
<th>Reagents</th>
<th>Initial Concentration</th>
<th>Final Concentration</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>Formamide (deionized)</td>
<td>100%</td>
<td>50%</td>
<td>50 ml</td>
</tr>
<tr>
<td>Dextran Sulphate</td>
<td>50%</td>
<td>10%</td>
<td>20 ml</td>
</tr>
<tr>
<td>Denhardt's Solution</td>
<td>100×</td>
<td>1×</td>
<td>1 ml</td>
</tr>
</tbody>
</table>
### Chapter 2 | Materials and Methods

<table>
<thead>
<tr>
<th>Yeast tRNA</th>
<th>10 mg/ml</th>
<th>250 mg/ml</th>
<th>2.5 ml</th>
</tr>
</thead>
<tbody>
<tr>
<td>NaCl</td>
<td>5 M</td>
<td>0.3 M</td>
<td>6 ml</td>
</tr>
<tr>
<td>Tris-HCl, pH8</td>
<td>1 M</td>
<td>20 mM</td>
<td>2 ml</td>
</tr>
<tr>
<td>EDTA</td>
<td>0.5 M</td>
<td>5 mM</td>
<td>1 ml</td>
</tr>
<tr>
<td>NaPO₄</td>
<td>1 M</td>
<td>10 mM</td>
<td>1 ml</td>
</tr>
<tr>
<td>Sarcosyl</td>
<td>20%</td>
<td>1%</td>
<td>5 ml</td>
</tr>
<tr>
<td>DEPC-H₂O</td>
<td></td>
<td></td>
<td>11.5 ml</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
<td><strong>100 ml</strong></td>
</tr>
</tbody>
</table>

#### 2.21.4. Day 3: Post hybridization washes

The coverslips were floated off by incubating the slides horizontally in 5×SSC. This step had to be carried out carefully to avoid tissue tearing. The sections were transferred to a Coplin jar and subjected to a high-stringency wash in a solution containing 50% formamide and 2×SSC at 65°C for 30 min. The slides were then subjected to milder washes in 2×SSC followed by 0.1×SSC for 15 min each at 37°C. The washes were succeeded by another wash with PBT (PBS + 0.1% Tween 20) for 15 min at RT. The slides were then placed horizontally in the humidified box and blocked for 1 hour using 10% heat-inactivated goat serum in PBT. About 200-300 µl per slide was used for blocking. The blocking solution was then removed and the sections were incubated with alkaline phosphatase-coupled anti-digoxigenin antibody (Roche; 1:5000 dilution) in PBT with 1% goat serum at 4°C overnight. Approximately 320 µl of the antibody solution was used per slide.

#### 2.21.5. Day 4: Visualization of reaction product

The antibody was removed the next day, the slides placed in a Coplin jar and washed four times in PBT at room temperature for approximately 8 hours in total. The slides were then washed twice for 10 min each in freshly prepared
NTMT buffer. The slides were again placed in the humidified box and about 300 µl of BM-purple AP substrate (Roche) containing 0.5 mg/ml levamisole was added to the sections. The slides were then incubated within a darkened humidified box overnight or as required. When color developed sufficiently, the slides were taken out and washed for 2-4 min in PBS followed by dipping in DEPC water. They were then left to air dry for atleast 2 hours and then sealed with coverslip using DPX and left overnight at room temperature for proper sealing.

Table 2.7: Composition of NTMT Buffer.

<table>
<thead>
<tr>
<th>Reagents</th>
<th>Initial Concentration</th>
<th>Final Concentration</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>NaCl</td>
<td>5 M</td>
<td>100 mM</td>
<td>2 ml</td>
</tr>
<tr>
<td>Tris-HCl pH9.5</td>
<td>1 M</td>
<td>100 mM</td>
<td>10 ml</td>
</tr>
<tr>
<td>MgCl₂</td>
<td>1 M</td>
<td>50 mM</td>
<td>5 ml</td>
</tr>
<tr>
<td>Tween-20</td>
<td>100%</td>
<td>0.1%</td>
<td>0.1 ml</td>
</tr>
<tr>
<td>Distilled Water</td>
<td>82.9 ml</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
<td>100 ml</td>
</tr>
</tbody>
</table>

2.22. Immunofluorescence of Frozen Tissue Sections

2.22.1. Preparation and sectioning of embryos

The embryos were dissected in PBS and it was ensured that all the membranes were carefully removed. The embryos were transferred to embedding molds and fresh-frozen in OCT as described above. Cryostat sections (10-15 µm thick) were transferred to Superfrost/Plus Slides, which were stored at -80°C or (preferably) used immediately for immunostaining.
2.22.2. Permeabilization of Sections

Frozen slides were prewarmed to room temperature for ~30 min. Liquid Blocker Super PAP Pen (Electron Microscopy Sciences) was used to mark the boundary of the sections. The sections were fixed with 4% paraformaldehyde at room temperature for about 15 min followed by three washes for 5 min each in PBS. The tissues are then permeabilized with 0.5% Saponin or 0.1% Triton-X for 20 min at RT. The slides were then washed thrice in PBS, 5 min each wash. Beyond this point, it was ensured that the tissue sections do not dry up.

2.22.3. Blocking and Incubation with Primary Antibody

The slides were then removed from PBS; excess PBS was tapped off onto paper towels and the slides were then laid horizontally in a humidified box. Blocking Buffer (containing 10% horse serum, 10% goat serum and 1% BSA in PBS with 0.1% Triton-X) was added onto the slides to cover the tissues and blocking was carried out in the closed humidified box for 1 hour at room temperature or overnight at 4°C.

2.22.4. Incubation with Secondary Antibody

Before adding the secondary antibody, the tubes were spun down at max speed to precipitate any debris or impurity coming from the serum. The slides were then washed thrice in PBS for 5 min each. The excess PBS is removed and the slides are laid horizontally in the humidified box. Secondary antibody at a dilution of 1:500 was added to the sections and incubated for 1 hour at room temperature in dark. Similar to the primary antibody, the secondary antibody dilution should be spun down to remove impurities. After the incubation time, the slides were washed in PBS thrice for 5 min each.
2.22.5. Incubation with DAPI and Mounting

Excess PBS was removed and the slides transferred back to the humidified box. Around 300 µl of DAPI (4',6-Diamidino-2-Phenylindole) at a concentration of 0.5 µg/ml was added to the sections. The slides were incubated in dark for around 5 min at RT after which they were washed briefly in PBS once. One slide at a time, excess PBS was removed and after placing them horizontally on a flat surface, the remaining PBS was wiped off from the sides. 3-5 drops of mounting medium was added and a coverslip was put over it slowly and carefully to avoid trapping air bubbles. The slides were incubated in dark overnight at RT and observed under fluorescence microscope on the next day. Alternatively, the slides were stored at -20°C in the dark for longer periods.

2.23. Isolation of Neural Stem Cells from Mouse Brain

2.23.1. Dissociation of Embryonic Day 14 (E14) Mouse CNS Tissue

At gestational day E14 (where E0 is the day when a postcoital plug is detected), the pregnant mouse was sacrificed using an IACUC-approved protocol and the embryos were dissected. Brain was removed from all embryos and transferred to a dish containing cold PBS. Thereafter, the desired brain regions were dissected out and placed in cold PBS on ice. Once all the brains were dissected, all tissues were transferred to a 15 ml polypropylene tube. The tissues were then allowed to settle down and the supernatant was pipetted off carefully. The brain tissues were resuspended in 1 ml Complete Embryonic NeuroCult™ Proliferation Medium (StemCell Technologies). By using a 1 ml micropipette tip set at 0.9 ml, the tissue was triturated 5 times to create a single cell
Chapter 2 | Materials and Methods

suspension. If undissociated tissue was still observed, the clumps were allowed to settle for 1-2 mins and then the supernatant containing the single cells was transferred to a new tube. Complete medium was added to resuspend the cells to a final volume of 10 ml and it was then spun down for 5 mins at 150×g. The cell pellet was again resuspended by brief trituration in 1 ml medium and then the final volume was made up to 4 ml. The cell suspension was filtered through a 40 µm cell strainer (BD Bioscience). The cells were then transferred to cell culture flasks for growing neurospheres.

2.23.2. Chemical Dissociation of Neurospheres

4-7 day old neurosphere cultures were examined under a microscope to determine if the neurospheres were ready for passaging. If neurospheres are attached to the bottom of the culture flask, the flask was tapped briefly to detach them. Viable neurospheres should be semi-transparent and phase contrast bright with many of the cells on the outer surface displaying microspikes. Ideally, cultures should contain neurospheres measuring approximately 100 µm or less in diameter. The procedure works less effectively if neurospheres are allowed to grow too large (>100 µm). A disposable plastic pipette was pre-wet with Neurocult Proliferation Medium (Mouse) lacking cytokines. Pre-wetting the pipette helps prevent cell loss due to sticking to the pipette. The culture medium containing suspended neurospheres was removed and transferred to a 50 ml polypropylene tube, which was then spun at 90 g for 5 mins. The supernatant was discarded carefully and 500 µl of NeuroCult® Chemical Dissociation Solution A (StemCell Thechnologies) per pellet was added. Using a 1 ml pipette set at approximately 450 µl, the neurospheres were gently resuspended. The lab timer was now set for 8 mins and 125 ml of NeuroCult® Chemical Dissociation Solution B (StemCell Thechnologies) was added to the cell suspension. The tube was gently tapped to mix the cells with the solution and then immediately the timer was started. At the 3-minute time point, the cell suspension was mixed by gently pipetting up and down 8 times without triturating using a 1 ml micropipette set at 450 µl. The cell suspension
was incubated till the 7-minute time point after which the cell suspension was
gently pipetted again up and down 8 times with a 1 ml micropipette set at 450
µl. Exactly at the end of 8 min, 40 µl of NeuroCult® Chemical Dissociation
Solution C was added and the cell suspension was mixed with a 1 ml
micropipette up and down 8 times. Then, 350 ml of Mouse NeuroCult®
Proliferation Medium (StemCell Technologies) was added to the final volume
to 1 ml and the cell suspension was mixed well again by pipetting it up and
down. This cell suspension should ideally contain single cells and no
aggregates. To wash the cells for subculturing, 10 ml of Mouse NeuroCult®
Proliferation Medium was added and then centrifugation was carried out for 5
min at 150×g. The cells were gently resuspended in medium again and
transferred to a fresh flask for further propagation.
• miR-124 is encoded by three different genes which reside in three distinct locations in mammalian genomes. Mouse precursor miR-124-1 is encoded on Chr14; miR-124-2 on Chr3 and miR-124-3 on Chr2. Are all three genes expressed in the same manner or do they show detectable differences in their spatio-temporal expression patterns? Answering this question constituted the first specific aim of this work.

• Moreover generally, it is unclear why miR-124, as well as many other miRNAs, is encoded in mammalian genomes by more than one non-allelic gene. Therefore, the second specific aim of this thesis was to begin understanding the biological reason(s) for this genetic “redundancy”.

• Finally, as mentioned in the Introduction and depicted in Figure 1.8, the three miR-124 precursors are encoded in different manners: miR-124-1 is encoded in an exon, miR-124-2 in a spliceosomal intron and miR-124-3 in a non-spliceable transcript. We therefore set out to identify possible functional reason(s) for these distinct coding strategies.
Chapter 4

RESULTS

This chapter describes three interrelated lines of experiments carried out to address the questions outlined above. Sections 4.1 and 4.2 deal with questions regarding redundancy in expression and in vivo functions of the three miR-124 genes, whereas Section 4.3 describes fundamental differences in processing of the three miR-124 precursors.

4.1. The three miR-124 genes show overlapping but distinct expression patterns in developing mouse embryo

Previous studies showed that the expression of the three miR-124 genes is regulated at the transcriptional level by REST, a repressor complex active in non-neuronal cells but inactivated in neurons [404]. This common regulation mechanism was proposed to result in largely overlapping expression patterns of the three miR-124 genes [405]. However, the pronounced neurological defects in mice lacking the 124-1 locus uncovered in a recent knockout study [157] indicated that the three miR-124 genes may have distinct biological functions. This prompted us to re-examine the expression of the 124-1, 124-2 and 124-3 genes using in situ hybridization, RT-(q)PCR and Northern blot analyses using gene-specific probes.

To study the detailed expression patterns of the individual miR-124 genes at earlier developmental stages, whole-mount in situ hybridization experiments were carried out in developing mouse embryos using gene-specific
miR-124 probes. Whole mouse embryos at E8.5, E9.5, and E10.5 were hybridized *in situ* with RNA probes specific to 124-1, 124-2, or 124-3 primary transcripts (pri-miRNAs). In each case, we used a mixture of two ~0.5 kb RNA probes complementary to pri-miRNA segments located immediately upstream and downstream of the pre-miRNA hairpin.

![Figure 4.1](image_url): Whole mount *in situ* analyses of mouse embryos (E8.5-E10.5) using probes against individual pri-miR-124 precursors.

These experiments highlighted a distinct pattern of spatio-temporal expression of the three miR-124 genes. Specifically, genes 124-2 and 124-3 were turned on in the neuroepithelial (NE) layer immediately after neurulation (Fig 4.1; E8.5 embryos), whereas 124-1 was not expressed until the onset of neuronal differentiation (Fig 4.1; E10.5 embryos).
The in situ hybridization data were corroborated by the RT-PCR results (Figure 4.2). Indeed, the expression of 124-1 was detectable only beginning from E10.5, whereas 124-2 and 124-3 RT-PCR signals could be detected as early as E9.0. As expected, all three miR-124 genes were expressed in adult brain samples but not in liver samples (Figure 4.2).

To gain further insights into the expression patterns of the three miR-124 genes, we carried out in situ hybridization experiments using cryosections prepared from E12.0 mouse embryos. These analyses showed that all three miR-124 genes are co-expressed at this stage in the mantle layer of the neural tube, an area containing differentiating and mature neurons (Figure 4.3B-C). However, only 124-2 and 124-3 are expressed in the neuroepithelial (or ependymal) layer containing neural stem cells (NSC). Interestingly, 124-3 is expressed at comparable levels in the NSCs and differentiating neurons, whereas 124-2 is expressed in differentiating neurons at a much higher level compared to the NSCs. 124-1, but not 124-2 and 124-3, is expressed throughout peripheral NS including dorsal root ganglia and cranial ganglia (Figure 4.3B).
Finally, \textit{124-1} and \textit{124-2}, but not \textit{124-3}, were expressed in the retina (not shown). The combined activity of the three miR-124 genes appears to generate a medial-to-lateral gradient of mature miR-124 throughout the CNS, with low expression in the NSCs and increasingly high expression in neurons undergoing differentiation (LNA staining in Figure 4.3C). Overall, these experiments led us to conclude that the spatio-temporal expression patterns of the three miR-124 genes are overlapping but clearly distinct.
4.2. Function of miR-124-1 in vivo

4.2.1. Generation of knockout mice

This distinct spatio-temporal expression pattern of the three miR-124 genes and the severe phenotypes manifested by mice lacking the 124-1 locus [157] encouraged us to attempt generation of conditional knockout alleles for all the three miR-124 genes. We reasoned that this effort will allow us to elucidate partial functions of the individual miR-124 genes. Figure 3.4 outlines three targeting constructs designed for the corresponding mouse miR-124 genes.

Figure 4.4: Illustration of the constructs used to target the three mouse miR-124 genes.

In order to generate conditional knockout mice, endogenous mouse miR-124 genes were replaced by cassettes encoding “substitute” miR-124 gene fragments with three distinct fluorescent protein reporters (Figure 4.4). The constructs contained the following critical elements:

- **Homology arms:** the homology arms served to insert the entire cassette into correct mouse loci using homologous recombination. The length of
the 3’ homology arm was considerably shorter (~3 Kb) than of the 5’ homology arm (~10 Kb) to facilitate PCR detection of correct recombination events.

- **Intron-encoded miR-124 “substitutes”:** Corresponding miR-124 gene fragments were inserted within an intron. The part of the intron upstream to the miR-124 genes is taken from the adenoviral genome whereas the downstream region has been constructed from the sequence corresponding to a mouse IgG gene.

- **IRES:** IRES (*Internal Ribosomal Entry Sequence*) enabled internal initiation of translation of the downstream fluorescent protein genes by recruiting initiating ribosomes to the mRNA.

- **Fluorescent Protein Reporter Genes:** Distinct fluorescent reporter genes were used to mark cells expressing targeting constructs and facilitate advanced miR-124 spatio-temporal expression studies.
  - miR-124-1 construct was marked with a eYFP gene.
  - miR-124-2 construct was marked with a mCerulean gene.
  - miR-124-3 construct was marked with a dTomato gene.

- **Tags:** Each of the three fluorescent protein genes was additionally marked with a distinct tag for detection purposes. All these tags contained the specific epitopes repeated three times to increase the sensitivity of immunodetection.
  - miR-124-1 construct contained 3×HA tag.
  - miR-124-2 construct contained 3×Flag tag.
  - miR-124-3 construct contained 3×Myc tag.

- **Neomycin/G418 resistance gene:** It served as a selection marker so that only correctly recombined cells could grow in the presence of antibiotic G418. Eukaryotic expression of this gene was achieved from a PGK (Phosphoglycerate Kinase) promoter.
• **FRT sites:** The Flippase Recognition Target (FRT) sites were added to the constructs to enable deletion of the PGK promoter-Neomycin/G418 resistance cassette in mouse line expressing FLP recombinase.

• **LoxP sites:** The LoxP sites were used to achieve conditional knockout of miR-124 loci in mice expressing Cre recombinase.

An important difference of the above miR-124-1 construct from the earlier published one [157] is the conditional nature of our allele and the ability to perform live imaging and isolation of miR-124-1 expressing cells using the IRES-EYFP-3×HA cassette. Another advantage of our allele is that following the Cre-mediated excision step it retains the overall 124-1 gene structure simply substituting the pre-miR-124-1 hairpin with a LoxP site (Fig. 4.4). The allele published by [157] is a large deletion eliminating the entire pri-miR-124-1 sequence.

### 4.2.2. Checking the reporter expression by Western Blotting

As an intermediate step in preparing the targeting constructs we confirmed the expression of the tagged fluorescent protein reporters from corresponding expression plasmids using Western Blotting with antibodies against corresponding tags (Figure. 4.5).
Figure 4.5: Western Blot showing the expression of tags corresponding to each miR-124 gene.

$pEM494$ plasmid contained the EYFP reporter tagged with $3\times HA$. The primary antibody used for the detection of the tag was rabbit anti-HA followed by the use of anti-rabbit HRP as the secondary antibody.

$pEM558$ contained mCerulean reporter tagged with $3\times$Flag. Mouse anti-Flag was used as the primary antibody followed by anti-mouse HRP as the secondary antibody.

$pEM539$ contained dTomato reporter tagged with $3\times$Myc. Rabbit anti-Myc was used as the primary antibody and anti-rabbit HRP was used as the secondary antibody. As expected, expression plasmids containing reporter genes without corresponding tags as well as mock-transfected cells did not show any staining.

All the plasmids containing the reporter genes without the tag as well as mock transfected cells did not give any band, which indicates no background/non-specific expression.
4.2.3. Identification of correctly targeted ES cell clones

Targeted ES cell colonies (see Materials and Methods for experimental details) were analyzed for proper insertion of the cassette by using genomic PCR with primers specific to the 3’ recombination site and Southern blotting using probes specific to the 5’ and 3’ homology arms (Figure 4.6 and not shown). These analyses allowed us to identify correct ES clones for all three targeting experiments (Figure 4.6). The desired ES clones thus obtained were further used for injection into blastocysts.

![Figure 4.6: Southern blot analyses of genomic DNA samples from selected ES clones. DNAs were cut with indicated restriction enzymes and analyzed using probes specific to the corresponding 3’ homology arms. Wt stands for genomic DNA from the wild type ES cells and the alphanumeric characters are clone numbers of targeted ES cells. For all three miR-124 genes, the longer fragments are derived from the wild-type miR-124 alleles, whereas the shorter fragments correspond to correctly targeted miR-124 alleles. Restriction enzymes used to fragment the genomic DNA before Southern blot analysis are indicated at the bottom.]

4.2.4. Breeding and creation of 124-1 conditional knockout mice

The positive ES cell clones were sent to a mouse knockout service facility (Biopolis, Singapore) for injection into blastocysts and creation of chimeric
mice. Only 124-1 chimeric males obtained through this effort transferred the recombinant allele to their progeny and no germ line transmission could unfortunately be achieved for the 124-2 and 124-3 chimera even following repeated rounds of blastocyst injection.

Given the possible links of 124-1 with cancerogenesis and neurological/neuropsychiatric disorders [160-162] and its unique spatio-temporal expression patterns [406], we decided to focus our further studies on this gene. For the sake of convenience, the wild-type 124-1 allele was referred to as 124-1<sup>wt</sup>, the targeted conditional allele containing the substitute intronic pre-miR-124-1 and the IRES-EYFP-3×HA cassette was called 124-1<sup-tm1.1</sup>, and the knockout allele was referred to as 124-1<sup-tm1.2</sup> (Figure 4.7).

![Available 124-1 alleles](image)

**Figure 4.7:** Diagram of the wild-type, conditional (124-1<sup-tm1.1</sup>) and null (124-1<sup-tm1.2</sup>) 124 alleles. Note that the conditional allele expresses a “substitute pre-miR-124” from an intronically encoded fragment (~500 bp) of the 124-1 gene. Positions of the recombinant intron (rIntron), the two LoxP sites, and the IRES-EYFP-3xHA cassette are indicated.

The chimeric agouti male mice were bred with wild type C57BL/6 mice in order to transmit the targeted allele to the progeny and establish the 124-1<sup-tm1.1</sup> containing line. Mice homozygous for 124-1<sup-tm1.1</sup> were phenotypically normal and expressed EYFP in embryonic brain in a pattern similar to that of the wild-type 124-1 gene.
When these animals were interbred, their progeny was born with Mendelian ratios. With ~25% homozygous $I24-1^{tm1.1}/I24-1^{tm1.1}$ mice surviving past weaning ($n=285$). Thus, the $I24-1^{tm1.1}$ allele had no detectable effect on mouse survival even in its homozygous form (Table 4.1).

Table 4.1: Number of mice corresponding to different $I24-1^{tm1.1}$ genotypes and sex on breeding.

<table>
<thead>
<tr>
<th>Genotype</th>
<th>$I24-1^{+/+}$</th>
<th>$I24-1^{tm1.1/+}$</th>
<th>$I24-1^{tm1.1/tm1.1}$</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Males</td>
<td>47</td>
<td>68</td>
<td>38</td>
<td>153</td>
</tr>
<tr>
<td>Females</td>
<td>31</td>
<td>68</td>
<td>33</td>
<td>132</td>
</tr>
<tr>
<td>Total</td>
<td>78</td>
<td>136</td>
<td>71</td>
<td>285</td>
</tr>
</tbody>
</table>

### 4.2.5. $I24-1$ homozygous knockout mice exhibit neonatal lethality

To begin characterizing the biological role of $I24-1$, the $I24-1^{tm1.1}$ mice were crossed with a sperm cell-specific Cre driver line (PRM-Cre) and obtained heterozygous pre-miR-124-1 knockout progeny ($I24-1^{wt}/I24-1^{tm1.2}$). However, when $I24-1^{wt}/I24-1^{tm1.2}$ animals were interbred, only 3.8% of the homozygous $I24-1$ null animals ($I24-1^{tm1.2}/I24-1^{tm1.2}$) survived past weaning (21 days) vs. the expected 25% ($n=261$; $p<0.0001$; $\chi^2$ test). Strikingly, 100% of the female $I24-1$–null pups died prior to weaning and all the $I24-1$–null animals surviving to adulthood were males (Table 4.2).
Table 4.2: Number of mice corresponding to different \(124-1^{\text{tm}1.2}\) genotypes and sex on breeding.

<table>
<thead>
<tr>
<th></th>
<th>124-1(^{+/+})</th>
<th>124-1(^{\text{tm}1.2/+})</th>
<th>124-1(^{\text{tm}1.2/1.2})</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Males</td>
<td>46</td>
<td>72</td>
<td>10</td>
<td>128</td>
</tr>
<tr>
<td>Females</td>
<td>51</td>
<td>82</td>
<td>0</td>
<td>133</td>
</tr>
<tr>
<td>Total</td>
<td>97</td>
<td>154</td>
<td>10</td>
<td>261</td>
</tr>
</tbody>
</table>

Of note, the previously published non-conditional knockout mice lacking the entire \(124-1\) locus also displayed neonatal mortality but this was apparently a sex-independent effect [157]. This indiscriminate mortality might be due to the loss of miR-124-independent functions encoded in the deleted genomic fragment. Moreover, the PGK-neo cassette retained in these animals [see Fig 1a of [157]] may dramatically modify the \textit{bona fide} miR-124-null phenotypes.

4.2.6. miR-124-1 expression is reduced in homozygous mutant mice brain

To check the expression of miR-124 in mice brain samples, \textit{in situ} LNA staining was done using the protocol described in Materials and Methods. Brain samples from E14.5 mice and P2 mice were frozen in OCT, sectioned using the cryostats at a thickness of 14 µm. The LNA sequence used is complementary to the mature miR-124 sequence.
Figure 4.8: miR-124 expression seen by LNA staining in E14.5 mice brain sections. The panels on the left correspond to homozygous knockout samples and the panel on the right corresponds to wild type mice brain samples. Sections in each row belong to the same or similar region of the brain to facilitate comparison of miR-124 expression levels.

As seen from the above staining result, the miR-124 expression is noticeably reduced in the homozygous knockout mutants as compared to wild type mice brain (Figure 4.8). Similar difference in mature miR-124 expression levels was detected between wild-type and \textit{J24-1}-null animals at postnatal day 2 (P2; not shown). Thus, we concluded that knockout of the \textit{J24-1} gene in mice effectively leads to a reduction in the overall levels of mature miR-124 expression. This reduction, however, is incomplete due to residual expression of the \textit{J24-2} and \textit{J24-3} genes.
4.2.7. miR-124-1 knockout leads to an increase in Neural Stem Cell proliferation

Given the known association of 124-1 with cancer, possible contribution of this gene to Neural Stem Cell (NSC) proliferation was further examined. For this, 124-1tm1.2/tm1.2 (homozygous knockout) embryos and their wild-type littermates were sectioned at E14.5 using a cryostat and the sections were immunostained with antibodies for NSC-specific markers (Ptbp1 and Sox2) and neuronal markers (e.g., Map2). The stained sections were examined using a confocal microscope (Zeiss LSM510 Meta Laser Scanning Confocal Microscope). This analysis revealed that the NSCs were apparently over-proliferated in the 124-1tm1.2/124-1tm1.2 knockout embryos (see white arrowheads in Figure 4.9).

Figure 4.9: Immunofluorescence analysis showing over-proliferation (arrowheads) of the NSC (neuroepithelial) layer in 124-1-null embryos.
This conclusion was confirmed when we stained newborn P0 stage mice brain with the NSC marker Sox2 along with the neuronal marker Map2. Compared to the wild type, ventricular cells of 124-1 null brain expressed Sox2 at a considerably elevated level in the NSC layer (white arrows in Figure 4.10).

### 4.2.8. Knockout NSC cultures show higher proliferation rates

To directly test possible differences in NSC proliferation rates, we established adherent *in vitro* NSC cultures from E14.5 embryos as described in Materials and Methods (Figure 4.11).
Figure 4.12: PrestoBlue assay showing significantly faster proliferation of $I24-1$-null (KO; $I24-1^{tm1.2}/I24-1^{tm1.2}$) NSCs as compared to NSC controls with other genotypes.

These cell cultures were prepared from female embryos belonging to each genotype: wild-type ($I24-1^w/I24-1^w$), heterozygous ($I24-1^w/I24-1^{tm1.2}$) and knockout ($I24-1^{tm1.2}/I24-1^{tm1.2}$). Cell proliferation assays were then carried out with adherent NSC cultures using PrestoBlue cell viability assay (Invitrogen). Satisfyingly, $I24-1$ null cells proliferated significantly faster than the wild-type ones, whereas the heterozygotes showed an intermediate proliferation rate (Figure 4.12). Thus, we concluded that miR-124 produced from the $I24-1$ locus functions as a negative regulator of NSC proliferation. Further experiments will be required to identify mechanisms underlying this effect as well as molecular reasons for the neonatal lethality in $I24-1$ null pups.
4.3. Role of spliceosomal intron in biogenesis of mature miR-124 from miR-124-2 precursor

4.3.1. Functional intron is essential for miR-124-2 biogenesis

To understand potential role of intron in miR-124 biogenesis from pri-miR-124-2, corresponding fragments of the three miR-124 precursors were sub-cloned into two expression vectors containing either spliceable intron or intron inactivated by mutations of its 5’ and 3’ splice sites (splicing-deficient intron). In both cases, the cassettes were under control of the mouse PGK promoter (Figure 4.13).

![Figure 4.13: Constructs used to test the role of functional intron in miR-124 biogenesis.](image)

This effort generated the following 6 constructs (also see Table 2.1):

1) pEM558 - 124-1 within spliceable intron;

2) pEM560 - 124-2 within spliceable intron;
3) pEM561 - 124-3 within spliceable intron;

4) pEM647 - 124-1 within splicing-deficient intron;

5) pEM559 - 124-2 within splicing-deficient intron;

6) pEM596 - 124-3 within splicing-deficient intron.

Mouse N2a cells were transfected with the above six constructs using Lipofectamine 2000 (Invitrogen) and the corresponding RNA samples were prepared using Trizol (Invitrogen) at 48 and 72 hours post transfection. The miR-124 expression was analyzed in these samples using Northern blotting with a miR-124-specific oligonucleotide probe and a U6 snRNA specific probe was used to control lane loading (see Materials and Methods for additional experimental details; Figure 4.14 and not shown). Interestingly, both spliceable and splicing-deficient constructs containing fragments of the 124-1 and 124-3 genes generated comparable amounts of mature miR-124 (Figure 4.14A). However, only spliceable but not splicing-deficient 124-2 construct produced detectable amounts of miR-124 (Figure 4.14A). This severely reduced production of mature miR-124 from the SDI-124-2 construct was a reproducible and statistically significant effect (Figure 4.14B). Another way to present these results is illustrated in Figure 4.15 where “splicing dependence” of miRNA production is calculated as a ratio of mature miR-124 expression levels generated from spliceable vs. splicing-deficient contexts. The splicing dependence value for miR-124-2 precursor was significantly higher than that for miR-124-1 and miR-124-3 precursors.
Figure 4.14: A) Northern blot analysis of splicing dependence of 124 genes in N2a cells. B) Expression levels of the samples used for the Northern indicating the significance of their differences.
Similar results were obtained when we repeated the above experiments using human embryonic kidney cells HEK293T (Figures 4.16 and 4.17). We concluded that, unlike the other two miR-124 precursors, pri-miR-124-2 shows a striking dependence on the presence of functional spliceosomal intron.
4.3.2. Swapping of miR-124-2 loop with 124-1 and 124-3 loops

To gain insights into molecular mechanisms underlying the distinct intron dependence of the three miR-124 precursors, we first compared their pre-miRNA stem loop structures. Although the stem elements were identical across the three precursors, the corresponding loop structures differed in the three pre-mRNAs: pre-miR-124-2 had an 11 nt long loop, pre-miR-124-3, 12 nt loop and pre-miR-124-1, 13 nt loop (Figure 4.18). We therefore substituted the pre-miR-124-2 loop with those from pre-miR-124-1 or pre-miR-124-3 as outlined in Figure 4.18 (see Figure 4.19) and examined the effect of this modification on the splicing dependence of these two modified 124-2 precursors. The corresponding constructs were introduced into N2a cells, followed by harvesting RNA samples 72 hours post transfection and analyzing the miR-124 expression by Northern blotting. This analysis suggested that the loop element has no effect on 124-2 biogenesis (Figure 4.20).
Figure 4.18: Diagram showing the stem loop structures of the three miR-124 pri-miRNA sequences.

Figure 4.19: Illustration of the constructs containing the loop mutants.
Figure 4.20: A) Northern Blot of 124-2 and its loop mutants. B) Splicing dependence of the constructs used.
4.3.3. Endogenous pri-miR-124-2 sequences adjacent to the pre-miR-124-2 hairpin is essential for its splicing dependence

Since the loop was not responsible for the pri-miR-124-2 splicing dependence, we examined possible effect of sequences upstream and downstream of the pre-miR-124-2 stem-loop element. For this purpose, we generated spliceable and splicing-deficient constructs encoding pre-miR-124-1, -2 and -3 devoid of their natural contexts (Figure 4.21). These minimized constructs were then introduced into N2a cells to check the miR-124 expression Northern blot analysis was carried out as above.

Figure 4.21: Illustrations for the wild type and mutated mini-miR-124 genes.
Strikingly, the splicing dependence of minimized miR-124-2 construct was significantly lower than that of full-length miR-124-2 construct (Figure 4.22). This result strongly suggested that the endogenous 5’ or/and 3’ sequences

Figure 4.22: A) Northern blot analysis of minimized miR-124 constructs. B) Splicing dependence of the miR-124-2-specific constructs quantified from five transfection experiments.
adjacent to the pre-miR-124-2 stem-loop structure were essential for its splicing dependence.

4.3.4. Both upstream and downstream sequences are required for the pri-miR-124-2 splicing dependence

To localize the sequence element endowing the pri-miR-124-2 precursor with unusually strong splicing dependence, we generated additional spliceable and splicing-deficient constructs lacking endogenous sequences either upstream or down-stream of the stem-loop element (Figure 4.23).

These mutant constructs were used to transfect N2a cells and the miR-124 expression was analyzed by Northern blotting (Figure 4.24A). Surprisingly, both 5’ and 3’ truncated 124-2 cassettes generated readily detectable amounts of mature miR-124 from both spliceable and splicing-deficient contexts. As a result, splicing dependence of these cassettes was dramatically lower than that
of the full-length 124-2 cassette (Figure 4.24B). These results suggested that the element required for pri-miR-124-2 splicing dependence was composed of sequences both upstream and downstream of the pre-miR-124-2 hairpin (Figure 4.25).
Figure 4.24: A) Northern blot analysis of the 124-2 precursor truncation mutants. B) Splicing dependence of 124-2 compared with the end mutants.
4.3.5. The composite upstream-downstream element in pri-miR-124-2 reduces the steady-state levels of unspliced pri-miR-124-2 precursor

To gain insights into molecular mechanisms allowing the composite upstream-downstream element to regulate pri-miR-124-2 splicing dependence, we carried out multiplex PCR analyses of the recombinant miR-124 precursor transcripts. The RNA samples from corresponding N2a transfection experiments were used to synthesize cDNA by RT-PCR using the SuperScript® III Reverse Transcriptase. The samples were then amplified by multiplex PCR using three universal primers EMO1208, EMO1209 and EMO1210. This assay was designed to afford simultaneous detection of precursor transcripts retaining the recombinant intron (unspliced products) or having this intron spliced out (spliced products) (see Figure 4.26).
The PCR products were then analyzed in a 2% agarose gel (Figure 4.27). As expected, we detected both spliced (bottom PCR bands) and unspliced (top PCR bands) products for constructs encoding long pri-miR-124-1 and pri-miR-124-3 fragments in spliceable context, whereas only unspliced products were detected for their splicing-deficient counterparts (see lanes 1, 3, 4 and 6 in Figure 4.27A). However, the situation was clearly different for the two constructs encoding long pri-miR-124-2 fragment. The spliceable construct generated readily detectable amounts of spliced products and extremely small but detectable amounts of unspliced products (lane 2 in Figure 4.27A). No products could be detected when the long pri-miR-124-2 fragment was expressed off splicing-deficient context (lane 5 in Figure 4.27A; also see lanes 1-2 in Figure 4.27B). Strikingly, unspliced products were expressed at readily detectable levels from constructs where the pri-miR-124-2 upstream-downstream element was either partially (lanes 7-10 in Figure 4.27A) or completely (lanes 3-8 in Figure 4.27B) deleted. Since all of the constructs used in this experiment contained identical mPGK promoters, these data suggested that the pri-miR-124-2 upstream-downstream element could destabilize RNA, an effect especially pronounced in the splicing-deficient context.
Figure 4.27: RT-multiplex PCR analysis of indicated recombinant miR-124 expression constructs.
4.3.6. The pri-miR-124-2 upstream-downstream element can destabilize RNA in the absence of the pre-miR-124-2 hairpin

To test if the upstream-downstream element had autonomous RNA destabilization function, we generated spliceable and splicing-deficient constructs containing the three pri-miR-124 fragments devoid of the pre-miR-124 hairpin structures (Figure 4.28).

N2a cells were transfected with these constructs (124-1D, 124-2D, 124-3D, SDI-124-1D, SDI-124-2D and SDI-124-3D, respectively) and their RNA expression products were analyzed by Northern blot and RT- multiplex PCR
As expected, none of these pre-miR-124 deletion constructs generated mature miR-124 (Figure 4.29).

**Figure 4.29: Northern Blot of the miR-124-2 constructs and the stem-loop deletion mutants.**

Strikingly, the distribution of the primary transcripts lacking the pre-miR-124 element was similar to that of the corresponding wild-type constructs.
containing intact pre-miR-124 stem-loop structures (compare lanes 1-6 with lanes 7-12 in Figure 4.30). Indeed, the steady-state levels of the unspliced RNA product generated from 124-2D were visibly lower than in the case of 124-1D and 124-3D (compare relative intensities of the top and bottom bands in lanes 7-9 of Figure 4.30). Furthermore, no unspliced transcript was detected for the SDI-124-2D, similar to the situation with SDI-124-2 (lanes 11 and 5 in Figure 4.30).

Comparable results were obtained when we repeated this experiment in HEK293T cells (Figure 4.31). We concluded that the pri-miR-124-2 upstream-downstream element can reduce steady-state levels of unspliced transcripts in the absence of pre-miR-124-2 precursor.

Figure 4.31: Multiplex PCR of miR-124 wild type genes and stem-loop deletion mutant constructs in HEK293T cells.
4.3.7. pri-miR-124-2 upstream-downstream element efficiently reduces levels of splicing-deficient but not spliceable precursor transcripts

To analyze the effect of the pri-miR-124-2 upstream-downstream element on transcripts levels in a quantitative manner, we analyzed the N2a RNA samples described in the previous chapter using RT-qPCR with exon-specific primers [EMO2810 and EMO2811] detecting both spliced and unspliced transcripts or intron-specific primers [EMO2912 and EMO2913] detecting unspliced transcripts only (Figure 4.32).

One effect detected by this assay was generally higher expression level of transcripts containing spliceable introns as compared to their splicing-deficient counterparts (Figure 4.33). This was in line with previously published data suggesting that the presence of spliceosomal introns can stimulate gene

![Figure 4.32: RT-qPCR assays used to examine the effect of the upstream-downstream element. A) Assay using exonic primers (B) Assay using intronic primers. Arrows indicate the position of the primers.](image-url)
expression [407-410]. More relevant to our line of inquiry was the observation that the steady-state levels of the pri-miR-124-2 precursor RNA expressed from splicing-deficient construct were considerably lower than the corresponding values for the pri-miR-124-1 and pri-miR-124-3 precursors (Figure 4.33). On the other hand, the total levels of pre-miR-124-2 pre-miRNA precursor expressed from spliceable construct were comparable to or even slightly higher than those of the spliceable pri-miR-124-1 and pri-miR-124-3 precursors (Figure 4.33). These results suggested that the pri-miR-124-2 upstream-downstream element efficiently reduced levels of splicing-deficient but not spliceable precursor transcripts.
Figure 4.33: RT-qPCR analysis of RNA samples. A) Relative expression levels of exonic mRNA corresponding to different stem-loop deletion constructs. B) RNA expression ratio between spliceable and splicing-deficient contexts.
Figure 4.34: Results of RT-qPCR analysis using intronic primers. A) Relative expression levels of intronic mRNA corresponding to different stem-loop deletion constructs. B) RNA expression ratio between spliceable and splicing-deficient contexts.
Interestingly, although the amount of unspliced transcripts generated from spliceable pri-miR-124-2 construct was lower than that for pri-miR-124-1 or pri-miR-124-3 (Figure 4.34), this difference was dramatically more pronounced when we compared corresponding values for splicing-deficient constructs (Figure 4.34). Taken together, these results suggested that the pri-miR-124-2 upstream-downstream element efficiently reduced expression levels of transcripts produced from splicing-deficient context whereas this effect was clearly attenuated for transcripts produced from spliceable constructs.

### 4.3.8. pri-miR-124-2 upstream-downstream element retains its activity in exonic context

Our data so far were consistent with the model that the pri-miR-124-2 upstream-downstream element destabilized transcript lacking functional introns; however, the destabilization effect was at least partially alleviated when the destabilization element was encoded within a spliceable intron (e.g., compare the levels of unspliced precursors between lanes 8 and 11 of Figure 4.30 or Figure 4.31). To better understand the role of splicing in this process, we generated a construct where the pri-miR-124-2 upstream-downstream element was encoded exonically, in the 3’UTR of *Renilla* luciferase (RLuc) gene containing a functional intron in its 5’UTR (pEM1158; Figure 4.35). We then transfected N2a cells with this construct, along with the appropriate control lacking the pri-miR-124-2 upstream-downstream element (pEM831) and examined the RLuc expression levels using luciferase assay (Figure 4.36). RLuc expression from pEM1158 was significantly lower than that from the control plasmid pEM831 (Figure 4.36; p<0.0001; Student’s t-test). Similar results were obtained when we repeated this experiment in HEK293T cells (Figure 4.37). We concluded that the pri-miR-124-2 upstream-downstream element may retain its RNA destabilization activity in heterologous exonic context despite the presence of a functional intron elsewhere in the transcript.
Figure 4.35: Designing Renilla luciferase (RLuc) constructs.

Figure 4.36: Graph showing relative RLuc expression levels in the presence and absence of the pri-miR-124-2 upstream-downstream degradation element in N2a cells.
4.3.9. The upstream-downstream element leads to faster decay of mRNA in a heterologous context

To directly test whether the pri-miR-124-2 upstream-downstream element had an RNA destabilization activity, *Renilla* luciferase constructs described in the previous chapter (pEM831 and pEM1158; see Figure 4.35) were transiently expressed in N2a cells for 12 hours. The transcription was then arrested with 5 µg/ml Actinomycin D and the transcript decay time courses recorded (Figure 4.38). pEM1158 transcripts containing the upstream-downstream element were degraded noticeably faster than the pEM831 control transcripts. Fitting the curves using a biphasic exponential decay model showed that the apparent half-life of the pEM831 transcripts was ~180 min, whereas the corresponding value for pEM1158 was >5 fold smaller (~32 min). This data strongly suggested that the upstream-downstream sequence functions as a potent RNA destabilization element.
Figure 4.38: Estimation of the apparent half-life of Renilla luciferase in the presence and absence of the pri-miR-124-2 upstream-downstream degradation element. N2a cells were transfected with constructs containing hRenilla luciferase with (pEM1158) or without (pEM831) the degradation element in the 3’UTR and the RNA decay time course was recorded after arresting transcription with Actinomycin D. Data were fitted using a biphasic exponential decay model (green and red lines).

4.3.10. Presence of the upstream-downstream element in precursor transcript impedes miRNA biogenesis

Our data so far supported the model that the unusually strong splicing dependence of miR-124 biogenesis from the pri-miR-124-2 precursor was due to destabilization of unspliced precursor transcripts by the upstream-downstream element. To directly test if this element could alter miRNA production levels, we inserted fragments of two heterologous miRNA genes, miR-1 (a muscle-specific miRNA) and miR-153 (a brain-specific miRNA), into the RLuc 3’UTR either containing (pEM1158) or lacking (pEM831) the pri-miR-124-2 upstream-downstream sequence (Figure 4.39). We then transfected N2a cells with these constructs along with the appropriate vector-only controls and analyzed the miRNA expression using Northern blots (Figures 4.40-4.43). miR-1 and miR-153 were expressed at readily detectable levels from the vectors lacking the upstream-downstream element. However, the presence of this element reduced the miRNA production dramatically (Figures 4.40-4.43).
Thus, the presence of the upstream-downstream element in precursor transcript is sufficient to impede miRNA biogenesis.

Figure 4.39: Designing miRNA expression constructs containing (right) or lacking (left) the pri-miR-124-2 upstream-downstream RNA destabilization element.

<table>
<thead>
<tr>
<th></th>
<th>48h</th>
<th>72h</th>
<th>48h</th>
<th>72h</th>
<th>48h</th>
</tr>
</thead>
<tbody>
<tr>
<td>miR-1 stem-loop</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Upstream-downstream element</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
</tbody>
</table>

Figure 4.40: Northern blot analysis of mature miR-1 expression in the absence and presence of upstream-downstream element. The samples analyzed were harvested 48 and 72 hours post transfection, as indicated.
Figure 4.41: Quantification of the mature miR-1 expression levels in Figure 4.39.

Figure 4.42: Northern blot analysis of mature miR-153 expression in the absence and presence of upstream-downstream element. The samples analyzed were harvested 48 hours post transfection.
Figure 4.43: Quantification of the mature miR-153 expression levels in Figure 4.41.
4.3.11. Biogenesis of mature miR-124 from the natural pri-miR-124-2 intron can be regulated by alternative splicing

Since the above experiments were carried out using recombinant introns, we examined the effect of splicing on miR-124 biogenesis from its natural pri-miR-124-2 context. Within pri-miR-124-2, miR-124 is encoded in
intron 2 flanked by exons 2 and 3 (Figure 4.44A). Analysis of the available pri-
miR-124-2-specific EST sequences from the UCSC Genome database
uncovered two putative alternative 5’ splice sites (5’ss) marking the boundary
between exon 2 and intron 2 (Figure 4.44A). Of these, the upstream 5’ss (that
we tentatively referred to as “embryonic” or E5’ss) was frequently used in EST
clones derived from embryonic NS, whereas the downstream 5’ss (“adult”;
A5’ss) was often present in adult brain ESTs (Figure 4.44A). RT-PCR analysis
of RNA samples from embryonic (E9.0, E10.5, E12.5, E15.5, and E17.5) and
postnatal (P45, P60 and P180) mouse brain using KAPA Taq and exon 2- and
3-specific primers [EMO3016 and EMO3017] indeed confirmed that the 5’ss
preference changed as a function of development from the predominant use of
E5’ss in embryonic brain to A5’ss in adult brain (Figure 4.44B-C).
With this in mind, we generated a series of “natural” pri-miR-124-2 expression constructs encoding either wild-type intron 2 flanked by exon 2 and exon 3 sequences (N-124-2) or its mutants that lacked functional A5’ss [N-124-2(mutA5’)], E5’ss [N-124-2(mutE5’)], both A5’ss and 3’ss [N-124-2(mutA5’/3’ss)], both E5’ss and 3’ss [N-124-2(mutE5’/3’ss)], and finally construct lacking E5’ss, A5’ss and 3’ss [N-124-2(SDI)] (Figure 4.45). These six constructs were transiently expressed in N2a cells and the splicing patterns of the corresponding transcripts were examined using an RT-multiplex PCR assay using KAPA Taq and primers EMO3016, EMO3017 and EMO3259 designed to detect both “embryonic” and “adult” spliced products, as well as unspliced RNA (Figure 4.46). All three RNA products were detected in the wild-type N-124-2 sample (see lane 3 in Figure 4.46). As expected, the N-124-2(mutA5’) transcripts were spliced exclusively in the “embryonic” manner.
(lane 7 in Figure 4.46), whereas the N-124-2(mutE5’) transcripts were spliced in the “adult” manner (lane 8 in Figure 4.46). Both “adult” and “embryonic” splicing was virtually abolished in transcripts lacking functional 3’ss [N-124-2(mutA5’/3’ss), N-124-2(mutE5’/3’ss) and N-124-2(SDI)] and only the unspliced forms were readily detectable in these cases (lanes 4-6 in Figure 4.46).

Figure 4.46: RT-multiplex PCR analysis of splicing patterns of indicated recombinant pri-miR-124-2 transcripts in N2a cells. This assay resolves transcripts spliced in the “adult” (470 bp) and “embryonic” (218 bp) manner as well as unspliced RNA (159 bp). Note that the 124-2 and SDI-124-2 controls in lanes 1 and 2 lacked natural E5’ss and A5’ss-associated sequences and therefore could only give rise to unspliced RT-PCR products.

To examine the effect of intron 2 splicing on miR-124 biogenesis, we repeated the N2a transfection experiment and analyzed the samples using Northern blotting with a miR-124-specific probe (Figure 4.47A).
Figure 4.47: A) Northern blot analysis of the effect of splicing on miR-124 biogenesis from its “natural” intronic sequence in N2a cells. (B) Quantification of the results in (A) showing relative miR-124 expression levels.

N-124-2 construct generated readily detectable amounts of mature miR-124 (lane 4 in Figure 4.47A and Figure 4.47B). Mutation of the A5’ss either alone or together with the 3’ss led to a minor decrease in the miR-124 production output (lanes 5 and 8 in Figure 4.47A and Figure 4.47B). A
Figure 4.48: (A) Northern blot analysis of the effect of splicing on miR-124 biogenesis from its “natural” intronic sequence in HEK293T cells. (B) Quantification of the results in (A) showing relative miR-124 expression levels.
substantially more pronounced reduction in miR-124 yield was detected for constructs lacking either E5’ss or both E5’ss and 3’ss (lanes 6 and 9 in Figure 4.47A and Figure 4.47B). Strikingly, elimination of all three splice sites [N-124-2(SDI)] led to nearly complete loss of mature miR-124 (lane 7 in Figure 4.47A and Figure 4.47B).

Similar results were obtained when we repeated this experiment in HEK293T cells; although in this case some residual expression of mature miR-124 was still detectable for the N-124-2(SDI) construct (Figure 4.48 and not shown). We concluded that miR-124 biogenesis from its natural intronic context required efficient splicing, similar to the situation with recombinant introns. Moreover, “embryonic” splicing pattern appeared to stimulate miR-124 production more efficiently than the “adult” one.

4.3.12. Preliminary screen for trans-factors involved in miR-124 biogenesis

To begin understanding possible control of the pri-miR-124-2 splicing dependence by trans-acting factors, we carried out a candidate RNAi screen targeting known components of the Microprocessor complex including DROSHA, DGCR8/PASHA, p68/DDX5, p72/DDX17. For this purpose, HEK293T cells were transfected with the corresponding ON-TARGETplus siRNA pools (Dharmacon; 100 pmol per well of a 6 well) using Lipofectamine 2000. 48 hours later, cells were post-transfected with 124-2 and 124-2-SDI constructs and allowed to express miR-124 for another 24 hours. RNA samples were then analyzed for RNAi efficiency (Figure 4.49) using the following RT-qPCR primers: EMO999 and EMO1000 (for hDrosha), EMO1001 and EMO1002 (for hDDX5), EMO1003 and EMO1004 (for hDDX17) and EMO927 and EMO928 (for hDGCR8). The sequences for these primers have been mentioned in Table 2.3.
Figure 4.49: Microprocessor component RNAi efficiency examined by RT-qPCR using corresponding primers. The samples were analyzed 72 hours post transfection.

The qPCR results confirm that there is an effective knockdown of all the microprocessor components after 72 hours, if not 24 hours. Hence, these siRNAs can further be used to study miR-124 biogenesis by transfecting them along with miR-124 constructs and looking for changes in efficiency of miR-124 biogenesis. This analysis confirmed efficient knockdown of all four Microprocessor components (Figure 4.49).
Figure 4.50: Effect of knocking down specific Microprocessor components on the pri-miR-124-2 splicing dependence.

Interestingly, when we analyzed the miR-124 expression using Northern blot, splicing dependence of pri-miR-124-2 was visibly reduced in all four RNAi samples (not shown). This effect was especially pronounced for the DDX17/p72 knockdown (Figure 4.50). We concluded that the biogenesis of miR-124 from all three primary transcripts occurs in a microprocessor dependent manner rather than through the mirtron pathway. Additional work will be required to identify trans factors rendering miR-124 biogenesis splicing-dependent.
Chapter 5

CONCLUSIONS

• miR-124-1 is turned on later in development than 124-2 and 124-3.

• miR-124-1 is expressed only in differentiated neurons while 124-2 and 124-3 are expressed in both neurons and NSCs.

• Mice containing cassette encoding EYFP and a substitute 124-1 sequence inserted at the endogenous pre-miR-124-1 sequence are born with Mendelian ratio, phenotypically normal and also show a wild type-like miR-124 expression pattern.

• 124-1 homozygous knockout mice show highly penetrant neonatal lethality. Of the 3.8% null mutants surviving in heterozygous crosses, all animals are males. No female 124-1 null mutants survived past weaning this far.

• In situ hybridization analyses showed decreased miR-124 expression in E14.5 and P2 brains of 124-1 null animals.

• Increased neural stem cell proliferation was detected in 124-1 knockout embryos. This phenotype was also confirmed using in vitro NSC cultures.

• Functional intron is essential for miR-124 biogenesis from pri-miR-124-2 but not for pri-miR-124-1 and pri-miR-124-3 precursors.
Chapter 5  |  Conclusions

- Splicing dependence of pri-miR-124-2 processing does not depend on the loop element of the stem-loop structure.

- Element composed by sequences immediately upstream and downstream of the pre-miR-124 stem-loop is responsible for the pri-miR-124-2 splicing dependence.

- When located within an unspliceable part of a transcript, the upstream-downstream element brings about efficient RNA degradation. However, this destabilization effect is reduced or abolished when the upstream-downstream element is located within a spliceable intron.

- The upstream-downstream element destabilizes mRNA precursors and reduces miRNA biogenesis efficiency even when located activity in a heterologous contexts (e.g. 3’UTR).

- miR-124-encoding intron 2 of the pri-miR-124-2 precursor is alternatively spliced with the upstream 5’ splice site being preferentially used in embryonic and the downstream 5’ splice site in adult NS.

- In its natural context, miR-124 processing requires efficient splicing of intron 2. Stimulation of miR-124 biogenesis is especially potent when intron 2 is removed following the “embryonic” rather than “adult” splicing pattern.

- Knockdown of the Microprocessor components appears to reduce pri-miR-124-2 splicing dependence. This effect is especially pronounced in DDX17/p72 knockdown samples.
Chapter 6

DISCUSSION

MicroRNAs (miRNAs) are small non-coding RNAs that base-pair with their target mRNAs and function as important post-transcriptional regulators of various normal and disease-related biological processes [410, 411]. Metazoan miRNAs act largely as repressors of gene expression by either destabilizing their targets or inhibiting their translation [410, 412]. miRNAs constitute one of the largest gene families, potentially accounting for ~1% of the coding capacity of the human genome [410]. Most mammalian miRNA genes have multiple isoforms (paralogues) with as much as 15% of human miRNAs being encoded in more than one genomic locus. miRNAs are often expressed in tissue-specific manner. An important example of this trend is miR-124, the most abundant miRNA in adult and embryonic CNS [245, 343]. It is expressed in neurons and stimulates neuronal and represses glial differentiation of ES cells in vitro [345] as well as decreases the levels of non-neuronal transcripts [346, 413].

An important finding of this study is that the three distinct non-allelic genes encoding miR-124 are expressed in an overlapping but not identical spatio-temporal fashion in developing mouse NS. Indeed, our in situ hybridization analyses show a delayed onset of the 124-1 gene expression as compared to the 124-2 and 124-3 genes. 124-2 and 124-3 expression starts as early as E8.5, whereas 124-1 is turned on only at E10.5 (Figure 4.1). The staining results were confirmed by RT-PCRs showing the same pattern of expression of these genes (Figure 4.2). On the other hand, in situ hybridization using embryonic cryosections revealed a distinct spatial expression pattern for the three miR-124 genes. 124-1 was expressed in the mantle layer (characterized by the presence of differentiated neurons), whereas the expression of 124-2 and 124-3 was more diffused and found in both the NSC
and mantle layers (Figure 4.3). These and other observations suggest that the three genes encoding identical mature miR-124 molecules may not be truly redundant.

A direct support for specialized functions of individual miR-124 genes comes from our miR-124 targeting work in mouse. Although we are still preparing \textit{124-2-} and \textit{124-3-} targeted mouse lines, we have successfully generated two \textit{124-1} lines. One of these (\textit{124-1}^{tm1.1}) expresses an EYFP cassette and a substitute \textit{124-1} fragment and it can be used for conditional knockout work and miR-124 expression studies (Figure 4.7). Mice homozygous for \textit{124-1}^{tm1.1} were found to be phenotypically normal and immunostaining for EYFP revealed a pattern similar to the \textit{124-1 in situ} staining results (not shown). Genotyping of the \textit{124-1}^{tm1.1} litters shows that this allele is inherited in a Mendelian manner with no detectable sex bias (Table 4.1). This result suggests that interrupting the pri-miR-124-1 precursor with a large heterologous insertion is phenotypically benign and, on a more technical level, that the \textit{124-1} targeting procedure did not induce inadvertent deleterious rearrangements elsewhere in the mouse genome.

Non-conditional removal of the substitute \textit{124-1} sequence and the EGFP cassette using sperm cell-specific \textit{Cre} driver line (PRM-Cre) allowed us to generate a \textit{124-1-null} allele (\textit{124-1}^{tm1.2}). Surprisingly, unlike their \textit{124-1}^{tm1.1} counterparts, only 3.8\% of the homozygous \textit{124-1}^{tm1.2} animals survived past weaning instead of the expected 25\% (p<0.0001). Another striking observation that was also made was that all the female pups which were \textit{124-1}^{tm1.2} homozygous died prior to weaning and thus, all the surviving ones were males (Table 4.2). The existence of such an effect is interesting since previously published deletion of the entire pri-miR-124-1 sequence in mouse led to neonatal mortality in apparently sex-independent manner. Given the pan-neuronal expression of the \textit{124-1} gene, possible causes for this phenotype include respiratory failure, abnormal feeding and defects in homeostasis [414]. However, defective respiration can be ruled out since most of the \textit{124-1-null} pups survive past 24 hours after birth.
One obvious biochemical effect of the 124-1 knockout detected by our in situ hybridization experiments is a noticeably reduced expression of mature miR-124 in homozygous knockout brains especially pronounced during later stages of development (Figure 4.8). This suggests that 124-1 contributes a large fraction of mature miR-124 to its overall expression pool. Moreover, in absence of 124-1, there is apparently no compensatory mechanism to synthesize extra miR-124 by the remaining 124 genes – 124-2 and 124-3.

At the cellular level, our immunostaining experiments using appropriate cell lineage markers suggest that the loss of 124-1 leads to detectable NSC overproliferation (Figure 4.9 and Figure 4.10). This phenotype was also obvious in our in vitro experiments using PrestoBlue NSC proliferation assays (Figure 4.12). The highest proliferation rates were seen for homozygous 124-1\textsuperscript{tm1.2} (124-1-null) cells with the heterozygous ones showing an intermediate rate of proliferation and the lowest rates seen for the wild type cells. These data directly implicate the 124-1 gene as a critical role in regulation of NSC proliferation rate. However, further experiments will clearly be required to understand if this effect is related to the neonatal lethality phenotype (see Future Work chapter).

miRNA biogenesis is a complex, multistep and compartmentalized process whose details are still poorly understood. In this study, we examined differences in miR-124 biogenesis from its three precursor genes. Strikingly, we show here that generation of mature miR-124 from the pri-miR-124-2 but not the pri-miR-124-1 or pri-miR-124-3 precursor shows an exceptionally strong dependence on the presence of a functional spliceosomal intron (Figure 4.14). The pri-miR-124-2 splicing dependence did not depend on the loop structure, the only feature distinct between the three pre-miR-124 hairpins (Figure 4.20). Rather, our deletion experiments allowed us to identify a composite element consisting of sequences both upstream and downstream of the pre-miR-124-2 stem-loop. Deletion of either 5’ or 3’ parts or both parts together of this upstream-downstream element abolished the splicing dependence (Figure 4.22 and Figure 4.24). Although additional work will clearly be required to understand the precise structure of the upstream-downstream element, its bipartite nature suggest possibility of long-range secondary structure.
interactions between the upstream and the downstream moieties. Both upstream and downstream parts contain noticeable stretches of interspecies conservation (not shown), which may instruct future molecular studies.

Several lines of evidence suggest that the upstream-downstream element destabilizes RNA in a manner that is tunable by its proximity to functional spliceosomal introns. (1) Our RT-PCR and RT-qPCR analyses (see e.g. Figures 4.26, Figure 4.29, Figure 4.30, Figure 4.32 and Figure 4.33) show that the presence of the upstream-downstream element in a splicing-deficient intron dramatically reduces the transcript steady state levels. On the other hand, positioning the destabilization element within a spliceable intron rescues the overall RNA stability including detectable accumulation of unspliced products (see e.g Figures 4.26, Figure 4.29, Figure 4.30, Figure 4.32 and Figure 4.33). (2) Insertion of the upstream-downstream element in exonic 3'UTR part of an otherwise spliced RLuc transcript reduces the luciferase production (Figures 4.35-4.36). (3) Finally, our RNA stability assays using an Actinomycin D-based approach directly show that the upstream-downstream element endows RNA with a substantially reduced half-life (Figure 4.37).

Although originally uncovered using recombinant introns, miR-124 retains its splicing dependence in its natural pri-miR-124-2 intron 2 (Figures 4.46-4.47). An interesting additional finding in this regard is that intron 2 is alternatively spliced during development with the upstream (“embryonic”) 5’ splice site being preferentially used in embryonic and the downstream (“adult”) 5’ splice site in adult NS. Moreover, miR-124 processing is shown to be especially efficient when intron 2 undergoes “embryonic” rather than “adult” type of splicing (Figures 4.46-4.47). It is therefore tempting to speculate that the developmental splicing program of intron 2 modulates the production output of the mature miR-124 from the 124-2 gene, which in turn fine-tunes the miR-124 spatio-temporal expression.

Finally, our preliminary trans-factor studies using RNAi-mediated knockdowns suggest that cellular levels of known Microprocessor components may play a role in enforcing pri-miR-124-2 splicing dependence. Interestingly, knockdown of DDX17/p72, a DEAD box helicase protein appears to cause the
most potent reduction in the splicing dependence (Figure 4.49). We will continue this line of experiments to further understand the role of DDX17 and possibly other protein candidate miR-124 biogenesis in general and bridging Microprocessor with the spliceosome complex in particular.

In summary, our studies identified interesting differences in developmental expression patterns of the three mammalian miR-124 genes, began addressing biological function of one of these genes in vivo and uncovered a novel post-transcriptional mechanism potentially tuning miR-124 expression during NS development. Our ongoing and future work in these directions is outlined in the next chapter.
Our initial analyses of the $124-I^{tm1.2}/124-I^{tm1.2}$ ($124-I$-null) mice have revealed a significantly elevated incidence of neonatal mortality in these animals. To detect abnormal feeding behavior, newborn mice can be monitored using a video surveillance setup along with regular measurement of their body weight. A common problem associated with inadequate feeding or/and lethal metabolic defects is hypoglycemia [415]. This can be addressed by monitoring the blood glucose levels.

In addition, the morphology of the $124-I$-null brain and other tissues at stages preceding weaning can be examined. The brain structure can be studied using Nissl and Golgi staining techniques as well as immunostaining for neuronal (tubulin βIII, Tau, MAP2), astroglial (GFAP), oligodendrocyte (O4, MBP) and microglial (CD11b, IBA1) markers. Incidence of apoptosis can also be examined using antibodies against cleaved caspase 3.

To understand the molecular mechanisms underlying the neonatal death phenotype, the gene expression profiles of the wild-type, heterozygous ($124-I^{wt}/124-I^{tm1.2}$) and 124-null ($124-I^{tm1.2}/124-I^{tm1.2}$) pre-weaning brain samples can be compared using RNA-seq. Our initial RNA seq data revealed a set of seven genes that were abnormally upregulated in $124-I$ null mice and were located in Chr 14 (same Chr as miR-124-1 gene). Out of these seven genes, six are located within 5 Mb of the $124-I$ locus and one is located within 10 Mb. On qPCR analyses of these targets in mouse NSC cultures for all the genotypes, it was found that TRIM13 was upregulated the most in the homozygous knockout mice.
To explore the formal possibility that the 124-1-null pups die due to the loss of miR-124 expression in cells other than neurons, the pre-miR-124-1 hairpin can be deleted exclusively in the NS by crossing mice carrying the conditional 124-1tm1.1 allele with a Nestin-Cre driver line (maintained in our lab). If the Nestin-Cre conditional KO (cKO) pups fail to develop the neonatal mortality phenotype, non-neural cells/tissues that have been previously reported to express miR-124 at detectable levels can be investigated. Also, monitoring the basic behavior and physiology of the cKO pups can give us leads into the role of 124-1 in mice.

As mentioned previously, according to our EST analyses, the exon preceding the pre-miR-124-2 intron contains two alternate 5’ss elements that appear to be used with different frequencies. Analyses of the underlying molecular mechanisms need to be carried out using 124-2-based transgenes that can be integrated into mouse P19 embryonic carcinoma cells using our recently developed HILO-RMCE procedure [416]. P19 cells can be efficiently differentiated into neurons in vitro [401], which will allow us to examine possible differences in the miR-124 production efficiencies as a function of neuronal differentiation stage. To distinguish the transgenic mature miRNA from the endogenous miR-124 up-regulated in P19 cells upon neuronal differentiation [417], we will substitute pre-miR-124-2 with a non-neuronal pre-miRNA or a miRNA-like shRNA designed against luciferase or LacZ mRNAs [416]. This approach will also allow us to identify critical cis-regulatory elements.

The tunable element adjacent to the pre-miR-124-2 hairpin may recruit presently unknown trans-factors accelerating degradation of unspliced pri-miR-124-2 and, on the other hand, interacting with the splicing machinery to avert degradation. To identify these factors and investigate the role of Ddx17/p72, we will first knock down the expression of known components involved in pri-miRNA processing including non-catalytic subunits of the Microprocessor complex (e.g., Dgcr8/Pasha, Ddx5/p68 and Ddx17/p72) and Microprocessor modulators including Brca1, Smad3, p53, and Dhx9 [418-422]. The RNA interference (RNAi) experiments will be carried out in Neuro2a cells using...
Dox-inducible shRNA transgenes [416] or/and gene-specific synthetic siRNA pools (Dharmacon).

We can also carry out affinity purification or/and pull-down experiments followed by label-free mass-spec analyses of proteins interacting with the pri-miR-124-2 destabilization element bait. Using previously reported strategies, we can also establish in vitro pri-miRNA processing assays and then use this platform to identify relevant degradation activities. The in vitro assays will also allow us to address the formal possibility that the tunable destabilization element functions as a self-cleavable ribozyme.

To systematically identify other NS-specific intron-encoded miRNAs whose biogenesis is significantly modulated by the splicing reaction we will treat Neuro2A cells, NSC cultures and primary cultures of embryonic cortical neurons with commercially available spliceosome inhibitors (spliceostatin A and pladienolide-B) and will examine the miRNA expression using NextGen sequencing (miRNA-seq). miRNA showing significant changes in response to drug treatment and distinct from mirtrons [423] will be shortlisted for further analyses.

Another line of experimentation can involve subcloning the intronic pre-miRNA fragments previously known highly expressed NS-specific genes into spliceable and splicing-deficient contexts and compare the miRNA expression levels using northern blotting or/and commercially available miRNA-specific RT-qPCR assays. miRNA genes showing tight splicing dependence will be compared with the 124-2 gene to deduce common sequence or/and secondary structure motifs that might facilitate bioinformatic identification of similar elements on a genome-wide scale. This specific direction will likely lead to identification of fundamental mechanisms regulating mammalian miRNA biogenesis.
Chapter 8

REFERENCES


31. Tiffany Hung, Y.W., Michael F Lin,Ashley K Koegel,Yojiro Kotake, Gavin D Grant, Hugo M Horlings, Nilay Shah, Christopher Umbricht, Pei Wang, Yu Wang, Benjamim Kong, Anita Langerod, Anne-Lise Børresen-Dale, Seung K Kim, Marc van de Vijver, Saraswati Sukumar,


34. Rackham, O., et al., Long noncoding RNAs are generated from the mitochondrial genome and regulated by nuclear-encoded proteins. RNA, 2011.


338. Sempere, L., et al., Expression profiling of mammalian microRNAs uncovers a subset of brain-expressed microRNAs with possible roles in


