

Investigation into the Subcellular Localisation and Function of miR-122

Poppy Winlow



**University of
Nottingham**
UK | CHINA | MALAYSIA

**BBSRC DTP
Gene Regulation and RNA Biology
School of Pharmacy**

Supervisor: Catherine Jopling

Declaration

Except where acknowledged in the text, I declare that this dissertation is my own work and it is based on the research work that was undertaken by myself in Gene Regulation & RNA Biology group, School of Pharmacy, Faculty of Science, University of Nottingham, UK.

Abstract

MicroRNAs are key post-transcriptional regulators of gene expression and function either by translational repression or degradation of target mRNAs. To do so, microRNAs must form an RNA-induced silencing complex (RISC) and direct its binding to the 3' UTR of their target mRNAs. Whilst the understanding of microRNA function has been extensively investigated, there has been little research on whether this differs in subcellular locations.

Although the majority of mature microRNA and their targets are concentrated in the cytoplasm, there is growing evidence for microRNA localisation and function in different subcellular compartments. As there has been no direct comparison of microRNA function in different subcellular sites in human cells, this project aims to address this question by applying subcellular fractionation methods and by generating luciferase reporters to compare regulation between subcellular compartments. This project specifically aimed to investigate microRNA regulation at the Endoplasmic Reticulum (ER), as they may have a direct role in silencing transcripts encoding secreted or membrane-localised proteins, and in the nucleus where they may have a range of functions including the regulation of microRNA biogenesis and regulation of nascent transcripts at the chromatin.

The overall aim was to investigate the subcellular localisation and function of miR-122. It is liver-specific and one of the most highly expressed microRNAs, accounting for roughly 70% of the total microRNA pool in Huh7 hepatocellular carcinoma cells, making it an ideal target for study. Furthermore, the role of miR-122 in the positive regulation of Hepatitis C Virus (HCV) RNA in liver cells provides the opportunity to examine if there are any differences in regulation between the ER and cytoplasm in the context of up-regulation of translation by miR-122.

To establish how microRNA regulation occurs at the ER versus cytoplasm, a series of ER-translated luciferase reporters for miR-122 regulation at the 3' UTR were successfully generated and their regulation by miR-122 was compared with that of equivalent reporters translated in the cytoplasm. This approach also enabled the comparison of miR-122 repression of the 3' UTR to miR-122 activation of translation via 5' UTR sites from HCV RNA. In summary, there was evidence for differential regulation via 3' UTR sites at the ER and cytoplasm in some, but not all, tested reporters.

Next, regulation of endogenous miR-122 targets that are known to associate to different subcellular sites were investigated using a membrane fractionation method to isolate ER and cytoplasm-localised mRNAs. The effects of miR-122 inhibition and overexpression on known miR-122 mRNA targets was compared in these fractions and some differences in regulation of these endogenous targets was observed between the ER and cytoplasm.

Finally, to examine the localisation of miR-122 within the nucleus a fractionation method was used to isolate chromatin and nucleoplasmic fractions from the cytoplasm which demonstrated the presence of miR-122 specifically in the chromatin fractions. To investigate the role of miR-122 in the chromatin, CRISPR/Cas9n genome modification was designed to disrupt a potential miR-122 seed match downstream of the pre-miR-122 encoding gene with the aim of investigating whether miR-122 autoregulates in a similar fashion to let-7 in *C.elegans*.

Ultimately these investigations provide new understanding of the subcellular localisation of miR-122 in Huh7 cells, demonstrating differences in miR-122 regulation at the ER and cytoplasm, and generated tools for the further analysis of miR-122 activity at different subcellular sites.

Acknowledgements

Give it a few months and I'm sure I will look back on my PhD with fond memories, but there's one thing that has never wavered and that's my gratitude and love for the people around me who have supported me throughout this journey.

Firstly, I would like to thank my supervisor Dr Catherine Jopling for her constant support, advice and patience. And to Dr Hilary Collins for stepping in during Catherine's maternity leave, although the weekly meeting in cafes with Catherine meant we didn't actually forgo our lab meetings. I am also grateful for all our technicians over the years for maintaining the lab, which is a much harder job than they're given credit for, and to all the supervisors in the Gene Regulation and RNA Biology (GRRB) group for their helpful discussions and feedback in meetings.

To all the PhD members, past and present, of GRRB thank you for not only offering guidance and help but most importantly entertainment. You made coming into the lab everyday a joy. Special mention goes to Aimée who gave me the most incredible start to my PhD and has been a true mentor throughout, and to Angie and Athena for carrying me through to the finish line. To the "Grrb Grrls" (Kat, Hannah, Nicole, Asta and honorary Grrl Mitchell) you have been the most wonderful friends and have given me some fantastic memories, I wouldn't have enjoyed my time in Nottingham half as much without you.

From outside the lab, Zaki, Joanna and James, thank you for the many phone calls listening to me rant and reminding me that I was strong enough to do it. To Callum, my best friend, I haven't stopped smiling since the day we met and I can't wait to conquer the world with you.

And the most special thanks to my parents, Dawn and Julian, who, without a doubt, I would not be where I am without their love and support. Thank you for believing in me, thank you for moulding me into a strong determined person, and thank you for always being there for me. I can't do or say enough to thank you, but I hope I make you proud.

Table of Contents

| | |
|---|------|
| Abstract | iii |
| Acknowledgements | v |
| Table of Contents | vi |
| List of Tables | xii |
| List of Figures | xiii |
| 1 Chapter One: Introduction | 1 |
| 1.1 Overview of Eukaryotic Gene Expression | 2 |
| 1.1.1 Transcription | 2 |
| 1.1.2 Co-transcriptional Events | 3 |
| 1.1.3 Eukaryotic Cytoplasmic Translation | 4 |
| 1.1.4 Eukaryotic Translation at the ER | 8 |
| 1.1.5 Control of Eukaryotic Translation | 10 |
| 1.2 Eukaryotic mRNA turnover | 12 |
| 1.2.1 Deadenylation | 13 |
| 1.2.2 5' – 3' Exonucleolytic decay pathway | 13 |
| 1.2.3 3' -5' Exonucleolytic decay pathway | 14 |
| 1.2.4 Targeted mRNA Decay | 14 |
| 1.2.5 P Bodies | 15 |
| 1.3 RNA Interference: An overview | 16 |
| 1.3.1 Small Regulatory RNAs | 16 |
| 1.3.2 Sequence conservation of microRNA | 18 |
| 1.3.3 Conservation of microRNA targets | 19 |
| 1.3.4 Biological Role of microRNA | 19 |
| 1.3.5 MicroRNA and Cancer | 21 |
| 1.3.6 Canonical Biogenesis Pathway | 22 |
| 1.3.7 Non-canonical Biogenesis | 23 |
| 1.3.8 Regulation of microRNA Biogenesis | 24 |
| 1.4 MicroRNA function | 25 |
| 1.4.1 Target recognition | 25 |
| 1.4.2 RISC composition | 25 |
| 1.4.3 MicroRNA-Mediated Deadenylation and Decay | 26 |
| | vi |

| | | |
|-------|---|----|
| 1.4.4 | MicroRNA-Mediated Regulation of Translation | 28 |
| 1.4.5 | Relative contribution of mRNA decay vs translational repression | 30 |
| 1.4.6 | Translational activation | 31 |
| 1.4.7 | microRNA Turnover | 32 |
| 1.4.8 | MicroRNA Sponges | 33 |
| 1.5 | MicroRNA Subcellular Localisation | 35 |
| 1.5.1 | Interaction between microRNA and P Bodies | 35 |
| 1.5.2 | microRNA and RISC at the Endoplasmic Reticulum (ER) | 36 |
| 1.5.3 | microRNA and the Endosomal Pathway | 37 |
| 1.5.4 | Localisation of microRNAs and RISC to the nucleus | 40 |
| 1.5.5 | Targeting of pri-microRNAs in the Nucleus | 41 |
| 1.5.6 | Other Nuclear Functions | 43 |
| 1.5.7 | Nucleolar microRNA | 44 |
| 1.5.8 | Localisation of microRNA and RISC at the Mitochondria | 44 |
| 1.5.9 | Cytosekeletal association of miRNA | 45 |
| 1.6 | miR-122 and microRNA therapeutics | 47 |
| 1.6.1 | miR-122 gene and expression | 47 |
| 1.6.2 | Physiological Function and Pathology | 48 |
| 1.6.3 | miR-122 regulation of Hepatitis C Virus | 50 |
| 1.6.4 | MicroRNA as therapeutics | 52 |
| 1.7 | Aims and Objectives | 54 |
| 2 | Chapter Two: Materials and Methods | 55 |
| 2.1 | Materials | 56 |
| 2.1.1 | General Cell Culture Reagents | 56 |
| 2.1.2 | Plasmids | 56 |
| 2.1.3 | PCR Primers | 58 |
| 2.1.4 | 2'-O-methylated and RNA oligonucleotides | 59 |
| 2.1.5 | siRNA sequences | 59 |
| 2.1.6 | Antibodies and Antisense Probes | 60 |
| 2.2 | Molecular Cloning | 61 |
| 2.2.1 | Designing guides for Cas9n | 61 |
| 2.2.2 | Phosphorylation and Annealing gDNA Oligonucleotides | 61 |
| 2.2.3 | Polymerase Chain Reaction (PCR) | 61 |
| 2.2.4 | Restriction Digest and Dephosphorylation | 62 |

| | | |
|-------|---|----|
| 2.2.5 | Agarose Gel Electrophoresis | 62 |
| 2.2.6 | Ligation | 63 |
| 2.2.7 | <i>In Vitro</i> Transcription | 63 |
| 2.2.8 | Genotyping | 63 |
| 2.2.9 | Sequencing | 64 |
| 2.3 | Bacterial Preparation and Culture | 65 |
| 2.3.1 | Composition of Reagents and Buffers | 65 |
| 2.3.2 | Preparation of heat-competent <i>Escherichia Coli</i> | 65 |
| 2.3.3 | Bacterial Transformations | 66 |
| 2.3.4 | Plasmid Purification | 66 |
| 2.4 | Cell Culture and Transfections | 67 |
| 2.4.1 | Cell maintenance | 67 |
| 2.4.2 | Cell Thawing | 67 |
| 2.4.3 | Cryopreservation | 67 |
| 2.4.4 | microRNA inhibitor Transfections | 68 |
| 2.4.5 | microRNA mimic/overexpression Transfections | 68 |
| 2.4.6 | Luciferase Reporter Transfections | 68 |
| 2.4.7 | Crispr/Cas9 Nickase (Crispr/Cas9n) Transfection | 70 |
| 2.4.8 | Fluorescence activated cell sorting (FACS) | 70 |
| 2.5 | Luciferase assays | 71 |
| 2.6 | RNA Extraction, Reverse Transcription (RT) and Real-time qPCR (qRT-PCR) | 73 |
| 2.6.1 | RNA Isolation | 73 |
| 2.6.2 | Mature microRNA qRT-PCR | 74 |
| 2.6.3 | mRNA qRT-PCR | 75 |
| 2.7 | Subcellular Fractionation | 76 |
| 2.7.1 | Chromatin-Associated RNA Fractionation | 76 |
| 2.7.2 | Membrane Fractionation | 77 |
| 2.8 | Ago-Immunoprecipitation (Ago-IP) | 78 |
| 2.9 | Northern blotting | 79 |
| 2.9.1 | Buffer Composition | 79 |
| 2.9.2 | microRNA Northern Blotting | 79 |
| 2.9.3 | mRNA northern blotting | 80 |

| | | |
|--------|--|-----|
| 2.10 | Protein Extraction, SDS-PAGE and Western Blotting | 83 |
| 2.10.1 | Buffer Composition | 83 |
| 2.10.2 | Protein Extraction | 83 |
| 2.10.3 | SDS-PAGE | 84 |
| 2.10.4 | Western Blotting | 84 |
| 3 | Chapter Three: Use of Luciferase Reporters to Investigate Regulation by miR-122 at the Cytoplasm and Endoplasmic Reticulum | 86 |
| 3.1 | Introduction | 87 |
| 3.2 | Generation of a Secreted Luciferase Reporter | 89 |
| 3.3 | FlucM4 and GlucM4 reporters are differentially regulated by pre-miR-122p3+4 expression | 92 |
| 3.4 | Effect of eIF4All knockdown on the regulation of FlucM4 and GlucM4 reporters | 96 |
| 3.5 | Generation of NanoLuc Reporters regulated by miR-122 | 99 |
| 3.6 | Regulation of cytoplasmic NlucM4 and secreted NlucSecM4 reporters by pre-miR-122p3+4 transfection | 101 |
| 3.7 | Comparison of secreted and cytoplasmic NanoLuc reporters by endogenous or overexpressed wildtype miR-122 binding to the 3' UTR | 105 |
| 3.8 | Regulation of secreted and cytoplasmic NanoLuc reporters by wildtype miR-122 binding to the HCV 5' UTR | 108 |
| 3.9 | Regulation of miR-122 reporters at the level of RNA stability | 113 |
| 3.10 | Discussion | 115 |
| 4 | Chapter Four: miR-122 Regulation of Endogenous mRNA Targets at the ER and Cytoplasm | 123 |
| 4.1 | Introduction | 124 |
| 4.2 | miR-122 target expression in cytoplasmic and membrane-rich fractions following miR-122 Inhibition with a 2'Ome Oligonucleotide | 127 |

| | | |
|-----|--|-----|
| 4.3 | miR-122 target expression in total cell lysate following miR-122 Inhibition and Overexpression with miRCURY LNA oligonucleotides | 134 |
| 4.4 | miR-122 target expression in cytoplasmic and ER-rich fractions following miR-122 Inhibition and Overexpression | 141 |
| 4.5 | Association of miR-122 with Argonaute Proteins at the Endoplasmic Reticulum and Cytoplasm | 149 |
| 4.6 | Discussion | 151 |
| 5 | Chapter Five: Investigation of the location and function of miR-122 within the nucleus | 157 |
| 5.1 | Introduction | 158 |
| 5.2 | miR-122 is present in the chromatin fraction | 160 |
| 5.3 | Design of a CRISPR/Cas9 strategy to modify a miR-122 seed match in the pri-miR-122 gene | 164 |
| 5.4 | Generation and genotyping of CRISPR/Cas9n cell lines | 168 |
| 5.5 | Discussion | 173 |
| 6 | Chapter Six: Discussion | 179 |
| 6.1 | Summary | 180 |
| 6.2 | Use of Luciferase Reporters for Monitoring miRNA activity at the ER and Cytoplasm | 182 |
| 6.3 | Optimisation of Subcellular Fractionation Methods | 184 |
| 6.4 | Endogenous miRNA-mediated regulation at the ER | 186 |
| 6.5 | Global miRNA-mediated Regulation at Different Subcellular Sites | 188 |
| 6.6 | Investigation into the Role of Chromatin-Localised miR-122 | 189 |
| 6.7 | Microscopy-based techniques for investigating the subcellular localisation of miRNA and RISC components | 191 |
| 6.8 | Concluding Remarks | 193 |
| 7 | References | 194 |

List of Tables

| | |
|--|------------|
| Table 2.1: Suppliers and Composition of Cell Culture Reagents .. | 56 |
| Table 2.2: Luciferase reporter constructs | 57 |
| Table 2.3: PCR primers | 58 |
| Table 2.4: Sequence of pre-microRNA mimics and antisense oligonucleotides..... | 59 |
| Table 2.5: siRNA sequences..... | 59 |
| Table 2.6: List of Antibodies used in Western Blotting..... | 60 |
| Table 2.7: Sequences of oligonucleotide probes for microRNA northern blot DNA | 60 |
| Table 2.8: Composition of Reagents/Buffers for Bacterial Transformation | 65 |
| Table 2.9: Composition of Materials and Buffers for Northern Blotting..... | 79 |
| Table 2.10: Composition of SDS-PAGE gel..... | 84 |
| Table 3.1: Contamination prevents accurate detection of luciferase mRNA by qRT-PCR..... | 114 |
| Table 4.1: miR-122 contamination of Input and IgG control samples. | 150 |

List of Figures

| | |
|--|------------|
| Figure 1.1: Overview of transcription: RNA polymerase binds to DNA..... | 3 |
| Figure 1.2: Eukaryotic Translation Initiation. | 5 |
| Figure 1.3: Overview of Translation at the ER | 9 |
| Figure 1.4: Overview of mRNA turnover pathways. | 12 |
| Figure 1.6: The canonical microRNA biogenesis pathway.. | 23 |
| Figure 1.7: microRNA-mediated mRNA decay..... | 27 |
| Figure 1.8: MicroRNA-Mediated Translational Repression..... | 28 |
| Figure 1.9: Overview of the Endosomal Pathway..... | 38 |
| Figure 1.10: ALG-1 regulation of Let-7 primary transcripts..... | 42 |
| Figure 1.11: Structure of both pri- and pre-miR-122 transcripts... .. | 47 |
| Figure 1.12: Schematic showing the binding of miR-122 to the 5' UTR of HCV..... | 51 |
| Figure 2.1: Luciferase Transfections..... | 69 |
| Figure 2.2: Schematic of Taqman RT/PCR primers..... | 75 |
| Figure 2.3: Schematic of tower transfer. | 81 |
| Figure 3.1: Translation of Gaussia and Firefly Reporters..... | 90 |
| Figure 3.2: Gaussia and Firefly Reporters with mutant miR-122 binding sites. | 91 |
| Figure 3.3: FlucM4 and GlucM4 reporters are differentially regulated over a 48 hr period by pre-miR-122p3+4 expression in Huh7 cells..... | 94 |
| Figure 3.4: Regulation of both FlucM4 and GlucM4 by both pre-miR-122p3+4 and pre-miR-122wt control as technical replicates..... | 95 |
| Figure 3.5: Knockdown of eIF4All with siA potentially has different effects on the regulation of FlucM4 and GlucM4 reporters pre-miR-122p3+4 expression in Huh7 cells. | 97 |
| Figure 3.6: Effect of eIF4All Knockdown with siC on the regulation of FlucM4 and GlucM4 reporters pre-miR-122p3+4 expression in Huh7 cells.. | 98 |
| Figure 3.7: Schematic of NanoLuc Reporters with miR-122 binding sites. | 100 |

| | |
|--|-----|
| Figure 3.8: Cytoplasmic and Secreted NlucM4 reporter expression over a 48 hr time period regulated by pre-miR-122p3+4 expression in Huh7 cells.. | 102 |
| Figure 3.9: Cytoplasmic and Secreted NlucM4 reporters are regulated similarly over a 48 hr time period by pre-miR-122p3+4 expression in Huh7 cells. | 104 |
| Figure 3.10: Cytoplasmic and Secreted Nluc reporter expression over a 48 hour time period following inhibition of endogenous miR-122 in Huh7 cells.. | 106 |
| Figure 3.11: Cytoplasmic and Secreted Nluc reporter expression over a 48 hour time period following overexpression of wildtype miR-122 in Huh7 cells. | 107 |
| Figure 3.12: Schematic of NanoLuc Reporters with the 3' UTR and 5' UTR from HCV..... | 109 |
| Figure 3.13: Cytoplasmic and Secreted Nluc reporter expression under control of the HCV 5' UTR over a 24 hour time period following inhibition of endogenous miR-122 in Huh7 cells..... | 111 |
| Figure 3.14: Cytoplasmic and Secreted Nluc reporter expression under control of the HCV 5' UTR over a 24 hour time period following overexpression of miR-122 in Huh7 cells..... | 112 |
| Figure 4.1: CAT1 is more strongly associated to the ER than ALDOA. | 126 |
| Figure 4.2: Schematic of miR-122 targets Aldolase A (ALDOA), Cationic Amino Acid Transporter 1 (CAT1), and Glucose-6-Phosphatase Catalytic Subunit 3 (G6PC3)..... | 127 |
| Figure 4.3: Protein levels of miR-122 mRNA targets following inhibition of miR-122 with a 2'Ome Antisense Oligonucleotide.. | 129 |
| Figure 4.4: Expression of miR-122 mRNA targets following inhibition of miR-122 with a 2'Ome oligonucleotide..... | 131 |
| Figure 4.5: Ratio of miR-122 target expression in cytoplasmic and membrane-rich fractions following miR-122 Inhibition with a 2'Ome Oligonucleotide. | 132 |
| Figure 4.6: Protein levels of miR-122 mRNA targets following transfection with a miR-122 LNA Inhibitor or Mimic..... | 136 |

| | |
|---|------------|
| Figure 4.7: Expression of miR-122 mRNA targets at 24 and 48 hours following transfection with an LNA miR-122 inhibitor.. | 138 |
| Figure 4.8: Expression of miR-122 mRNA targets at 24 and 48 hours following overexpression of miR-122 with an LNA Mimic.. | 140 |
| Figure 4.9: Expression of miR-122 mRNA targets following transfection with an LNA miR-122 Inhibitor. | 143 |
| Figure 4.10: Ratio of miR-122 target expression in cytoplasmic and membrane-rich fractions following miR-122 inhibition. | 144 |
| Figure 4.11: Expression of miR-122 mRNA targets following transfection with an LNA miR-122 Mimic. | 146 |
| Figure 4.12: Ratio of miR-122 target expression in cytoplasmic and membrane-rich fractions following miR-122 overexpression. q. | 148 |
| Figure 4.13: miR-122 association with AGO1-4 in membrane and cytoplasm fractions. | 149 |
| Figure 5.1: Western blot analysis performed by Dr Aimée Parsons of fractions obtained following chromatin-associated fractionation. | 161 |
| Figure 5.2: miR-122 is present in the chromatin of Huh7 cells.... | 162 |
| Figure 5.3: Northern blot analysis of RNA obtained following chromatin-associated fractionation. | 163 |
| Figure 5.4: Schematic showing relative location and binding potential of downstream miR-122 seed match. | 164 |
| Figure 5.5: Process of Cas9n DNA mutation. | 166 |
| Figure 5.6: Design of guide RNA sequences with respect to the potential miR-122 seed match. | 167 |
| Figure 5.7: FACS analysis of untransfected wild-type Huh7 cells, and Huh7 cells transfected with either single Cas9n guide (control) or paired Cas9n plasmids. | 170 |
| Figure 5.8: Genotyping of CRISPR/Cas9n clonal populations suggests no gene editing occurred. | 171 |
| Figure 5.9: gDNA sequence from CRISPR/Cas9n clonal populations perfectly align to targeted wild type sequence. | 172 |
| Figure 5.10: Conservation of the miR-122 binding site across primates.. | 178 |

List of Abbreviations

| | |
|-------------------|--|
| +ve | positive |
| Ago | Argonaute |
| ATP | Adenosine triphosphate |
| bp | Base pair |
| <i>C.elegans</i> | Caenorhabditis elegans |
| CaCl ₂ | Calcium chloride |
| cDNA | Complementary DNA |
| ChIP | Chromatin Immunoprecipitation |
| CLIP-seq | Crosslinking Immunoprecipitation Sequencing |
| CO ₂ | Carbon Dioxide |
| COPII | Coat Protein II |
| Cox-2 | Cyclooxygenase 2 |
| CRISPR | Clustered regularly interspaced short palindromic repeats |
| CRISPR/Cas9n | CRISPR/Cas9 nickase |
| C-ter | Carboxyl-terminus |
| Cyto | Cytoplasm |
| DGCR8 | DiGeorge critical region 8 |
| DMEM | Dulbecco's Modified Eagle Medium |
| DMSO | Dimethyl sulfoxide |
| DNA | Deoxyribonucleic acid |
| DSB/DSBR | Double strand break/ Double strand break repair |
| DSBs | Double-strand breaks |
| dsRNA | Double-stranded RNA |
| DTT | Dithiothreitol |
| E.coli | Escherichia coli |
| ECL | Electrochemiluminescence |
| EDC | 1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide |
| EDTA | Ethylenediaminetetraacetic acid |
| EGTA | Ethylene glycol-bis(β-aminoethyl ether)-N,N,N',N'-tetraacetic acid |
| EMCV | Encephalomyocarditis virus |
| ER | Endoplasmic Reticulum |

| | |
|-------------------------------------|---|
| FBS | Fetal bovine serum |
| FLuc | Firefly Luciferase |
| FRET | Fluorescent resonance energy transfer |
| GLuc | Gaussia Luciferase |
| GTP | Guanidine triphosphate |
| H ₂ O/diH ₂ O | Deionised water |
| HCl | Hydrogen chloride |
| HCV | Hepatitis C virus |
| HeLa | Henrietta Lacks |
| hnRNP | Heterogeneous nuclear ribonucleoprotein |
| IGF2 | Insulin-like growth factor 2 |
| InDels | Insertions/deletions |
| Inhib | Inhibitor |
| IP | Immunoprecipitation |
| IRES | Internal ribosome entry site |
| LB | Luria-Bertani |
| lncRNA | long non-coding RNA |
| m ⁷ G | 7-Methylguanosine |
| MgCl ₂ | Magnesium chloride |
| microRNA/miR | microRNA |
| mRNA | Messenger RNA |
| mt | mitochondria |
| MVBs | Multivesicular Bodies |
| NaCl | Sodium chloride |
| ND1 | NADH dehydrogenase 1 |
| NEB | New England Biolabs |
| nt | Nucleotides |
| OD | Optical Density |
| PABP | Poly(A)-binding protein |
| PB | P body |
| PBS | Phosphate buffered saline |
| PCR | Polymerase chain reaction |
| piRNA | Piwi-interacting RNA |
| Pre-microRNA | Precursor microRNA |
| Pri-microRNA | Primary microRNA |

| | |
|------------|--|
| qRT-PCR | Quantitative real-time polymerase chain reaction |
| Rand | Random |
| RISC | RNA-induced silencing complex |
| RLC | RISC-loading complex |
| RNA | ribonucleic acid |
| RNA pol | RNA polymerase |
| RNAi | RNA interference |
| RT | Reverse transcription |
| SDS | Sodium dodecyl sulfate |
| sgRNA/gRNA | Single guide RNA/guide RNA |
| siRNAs | Short-interfering RNAs |
| snRNA | Small nuclear RNA |
| TRBP | Trans-activation response RNA binding protein |
| tRNA | Transfer RNA |
| TSAP | Thermosensitive Alkaline Phosphatase |
| UTR | Untranslated region |
| UV | Ultraviolet |

Chapter One: Introduction

1.1 Overview of Eukaryotic Gene Expression

The first step of eukaryotic gene expression requires genes to be transcribed by an RNA polymerase into RNA transcripts in the nucleus, which in the case of protein-encoding mRNA is by RNA Pol II. This pre-mRNA transcript undergoes a series of co-transcriptional events such as the addition of a poly(A) tail and 5' cap and splicing prior to export into the cytoplasm. In eukaryotes, translation occurs in ribosomes found in either the cytoplasm or on the endoplasmic reticulum (ER) membrane where single amino acids are added in a sequence dictated by the genetic code of the mRNA transcript to generate a polypeptide chain that forms the functional protein.

1.1.1 Transcription

The process of copying a DNA sequence into RNA is known as transcription and is carried out by an RNA polymerase enzyme for which there are three types in animals and is outlined in Figure 1.1. Transcription of protein-encoding genes occurs when the RNA polymerase II binds to the promoter region which is facilitated by transcription factors [1, 2]. This region is then “opened” and the polymerase travels along to the transcriptional start site (TSS) on the DNA. The RNA polymerase then moves along the template strand of DNA, from the 3' end towards the 5' end, producing an RNA transcript running in the opposite 5' to 3' direction. Termination of transcription is signalled by the transcription of a poly(A) signal which is recognised by the cleavage and polyadenylation specificity factor (CPSF) and cleavage stimulation factor (CstF) proteins which act as a scaffold for other effector proteins to cleave the nascent mRNA transcript and the RNA polymerase is released from the DNA template.

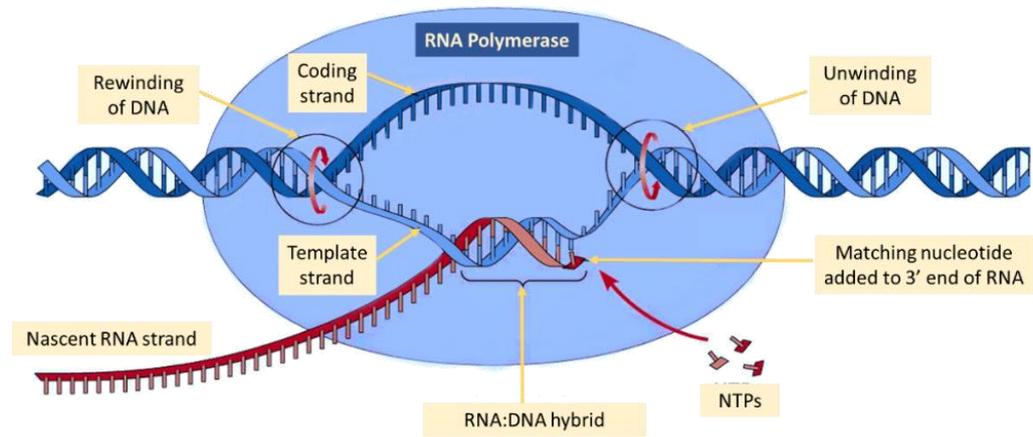


Figure 1.1: Overview of transcription: RNA polymerase binds to DNA. The DNA is unwound which allows RNA polymerase to add nucleotides (NTPs) that match the coding strand of the DNA to form a single-stranded RNA transcript. Image adapted from <https://microbiologynotes.org/transcription-in-prokaryotes-initiation-elongation-and-termination>

1.1.2 Co-transcriptional Events

Cleavage of the nascent RNA transcript and release of the RNA polymerase is coupled to polyadenylation, which is the addition of a series of adenosine bases to the 3' end of the mRNA by PolyA Polymerase (PAP) [3-5]. The nuclear PolyA binding protein (PABPN1), of which several can bind to the tail, increases processivity of PAP to add up to 250 additional adenosines [5-8].

During transcription, a 7-methylguanosine (m^7G) cap is added to the 5' end of the mRNA for protection [9, 10]. This m^7G cap is recognised by the cap-binding complex (CBC) which has been shown to have a wide range of roles in export of the RNA transcripts from the nucleus, splicing efficiency, co-transcriptional processing, and mRNA turnover and translation [11-14].

For some mRNAs, the newly made precursor messenger RNA (pre-mRNA) transcript is then transformed into a mature messenger RNA (mRNA) by the spliceosome. The spliceosome complex contains five small nuclear RNAs (U1, U2, U4, U5 and U6) which associate with additional accessory proteins to form the small nuclear RNA protein complex (snRNP) to recognise and remove introns from pre-mRNA [15].

1.1.3 Eukaryotic Cytoplasmic Translation

The mRNA is exported from the nucleus through the nuclear pore complex (reviewed in [16]) to the cytoplasm where it can undergo translation, which comprises of three main steps: initiation, elongation, and termination. The assembly of the 80S ribosome complex at the translational start site initiates translation and allows the elongation of the peptide chain, see Figure 1.2. Termination involves the detachment of the newly synthesised protein from the ribosomal complex and allows the recycling of ribosomal subunits following their detachment from the mRNA.

Translation Initiation

The translation initiation complex eIF4F assembles on the m⁷G cap at the 5' end. eIF4F is a heterotrimer complex and its formation is the rate-limiting step for initiation [17]. It consists of eIF4E, a 24-kDa 5' cap mRNA-binding protein [18, 19]; eIF4A, an ATP-dependent RNA helicase unwinding the 5' UTR secondary structure of mRNA, in cooperation with single-strand RNA binding protein eIF4B or eIF4H [20, 21]; and eIF4G, a large scaffolding protein that binds to both eIF4E and eIF4A and other proteins [22-24]. eIF4G interacts simultaneously with eIF4E and polyadenylate-binding protein 1 (PABP1) which in essence circularises the mRNA by bringing the two ends of the mRNA in close proximity [25-27] as can be seen in Figure 1.2. This enhances the rate of translation initiation by increasing the affinity of eIF4E for the 5' cap [28], and may facilitate ribosome recycling [26]. To enable the recruitment of the 40S ribosomal subunit, eIF3 interacts with eIF4G, eIF5 and the ternary complex (eIF2–GTP–Met-tRNA_i) [29] and stimulates binding of the 40S subunit. eIF3 also promotes the binding of the resulting 43S complex to the mRNA near the 5' m⁷G cap [30]. The ribosomal subunit then scans the mRNA from the 5' end, which requires the helicase activity of eIF4A for which two models have been suggested. One involves helicase-mediated “ratcheting”, where mRNA is moved through the mRNA binding channel of the ribosome requiring eIF4A, eIF4G and eIF4B with the secondary structures being unwound by the 40S subunit prior to entry.

The other model suggests the unwinding of mRNA occurs before entry into the channel by eIF4A, eIF4G and eIF4B [31-33]. Recognition of the first initiation codon (usually AUG) in the correct context then activates the hydrolysis of GTP bound to eIF2 [34, 35] and results in the recruitment of the 60s ribosomal subunit.

Translation Elongation and Termination

The subsequent release of eIF2-GDP allows the elongation of the peptide chain, facilitated by eIF5A, by the primed 80S ribosome (40S and 60S subunits) at the initiation codon [36]. The correct aminoacyl-tRNAs are recruited for each codon and are transported to the A (acceptor) site of the ribosome by the elongation factor eEF1A. Once the peptide bond is catalysed between the amino acid on the P-site tRNA and that on the A-site tRNA, the nascent peptide is transferred to the A site tRNA. The peptidyl-tRNA is then translocated to the P (peptide) site in the ribosome, facilitated by eEF2, which releases the A-site for a subsequent round of elongation with the now empty tRNAs released from the E site [37].

On recognition of a stop codon at the A-site of the ribosome, the release factors eRF1 and eRF3 form a complex with GTP and bind to the A-site. eRF3 catalyses GTP hydrolysis followed by eRF1 hydrolysis of the peptidyl-tRNA, releasing the completed polypeptide [38]. The 80S ribosome and translation factors must be recycled before subsequent rounds of translation can occur, with the dissociation of the two ribosomal subunits facilitated by the highly conserved ABCE1. ABCE1 hydrolyses ATP into ADP and two phosphates, and utilises the energy generated to separate the ribosome into its 40S and 60S subunits.

Figure 1.2: Eukaryotic Translation Initiation. (Next page) (1) The translation initiation complex eIF4F (eIF4G, eIF4A, and eIF4E) assembles on the m7G cap (shown in red) and interacts with polyA tail binding protein (PABP) at the 3' polyadenylated end of the mRNA. (2) Formation of a 43S pre-initiation complex (40s ribosomal subunit, eIF2-GTP-tRNAMet). (3) 43S complex recruited to the eIF4F complex (4) followed by scanning of the 43S ribosome in 5'-3' direction. (5) Recognition of the AUG start codon and release of eIF2-GDP. (6) 60S ribosomal subunit is then recruited with concomitant displacement of initiation factors. (7) Elongation-competent 80S ribosome is formed.

Cap-independent Translation Initiation

Alternatively, some cellular and viral mRNAs are translated independently of the m⁷G cap and eIF4E via alternative cap-independent mechanisms. Herein, the recruitment of the 40S ribosome is through interaction with an internal ribosome entry site (IRES), circumventing the requirement for the 5' m⁷G cap. Elongation then takes place after the binding of the 60S subunit at the AUG initiation codon.

IRESs were first discovered in picornaviruses in the 1980s [39, 40] and have since been identified in a wide range of viral mRNAs including Poliovirus [41], Hepatitis C virus (HCV) [42], and Encephalomyocarditis virus (EMCV) [43]. However, in eukaryotic cells the total number of mRNAs that may contain IRES elements has been estimated to be only 3% of the total mRNAs in the cell but include growth factors, transcription factors, and oncogenes [44, 45]. There are a number of examples of IRES-mediated translation as a cellular response to pathological and stress conditions, involving growth, nutritional, environmental, and proliferation signals [46, 47].

There are four distinct classes for viral IRES based on their different requirements for initiation factors and the complexity of their secondary and tertiary structures [48]. Class I IRESs have basic secondary structures and require all translational initiation factors, with the exception of eIF4E, whilst class II members are able to tether the translational machinery to the start codon, eliminating the scanning step. Similarly, Class III IRESs do not require scanning and are able to recruit the ribosome directly to the start codon but contain more complex structures than Class II IRESs and require fewer initiation factors. One well-studied Class III member is the HCV IRES which includes binding sites for the microRNA miR-122, whose role in HCV replication is discussed later in 1.6.3. The most sophisticated IRESs in terms of structure belong to Class IV, which usually contain several pseudoknots. Class IV members do not require initiation factors, an AUG start codon, nor Met-tRNA and are the

only class which are located outside of the 5'UTR, typically in intergenic regions.

1.1.4 Eukaryotic Translation at the ER

Proteins that are destined for organelles in the endomembrane system (ER, golgi, lysosomes) or to be secreted out of the cell are targeted to the ER for translation by a signal peptide [49-51]. This signal peptide is a series of hydrophobic amino acids at the N-terminus of the protein which is recognised by the signal-recognition particle (SRP) as it emerges from the ribosome and is responsible for targeting the ribosome (rather than the mRNA directly) to the ER [52-54]. The nascent amino acid chain is then co-translationally fed through to the lumen of the ER through a channel formed by the Sec61p complex [52, 53]. For ER membrane-bound proteins, the signal peptide remains attached to the protein and gets embedded into the membrane, anchoring it, whilst for some proteins the signal peptide is cleaved, releasing the protein into the lumen of the ER [55, 56]. Figure 1.3 outlines the process.

Whilst in the ER the newly made proteins fold into their functional structures and may also undergo post-translational modifications such as glycosylation. Newly synthesised proteins destined for other organelles in the endomembrane pathway are exported from the ER at ER exit sites (ERES) and are incorporated into vesicles coated in coat protein COPII [57-59]. These COPII vesicles are able to fuse to the Golgi complex where they can be further modified or processed and shipped to their final destination [60, 61]. If the proteins do not contain tags targeting them to endomembranes (ER, Golgi, lysosomes), the proteins are shipped from the Golgi apparatus to the plasma membrane or secreted to the cell exterior.

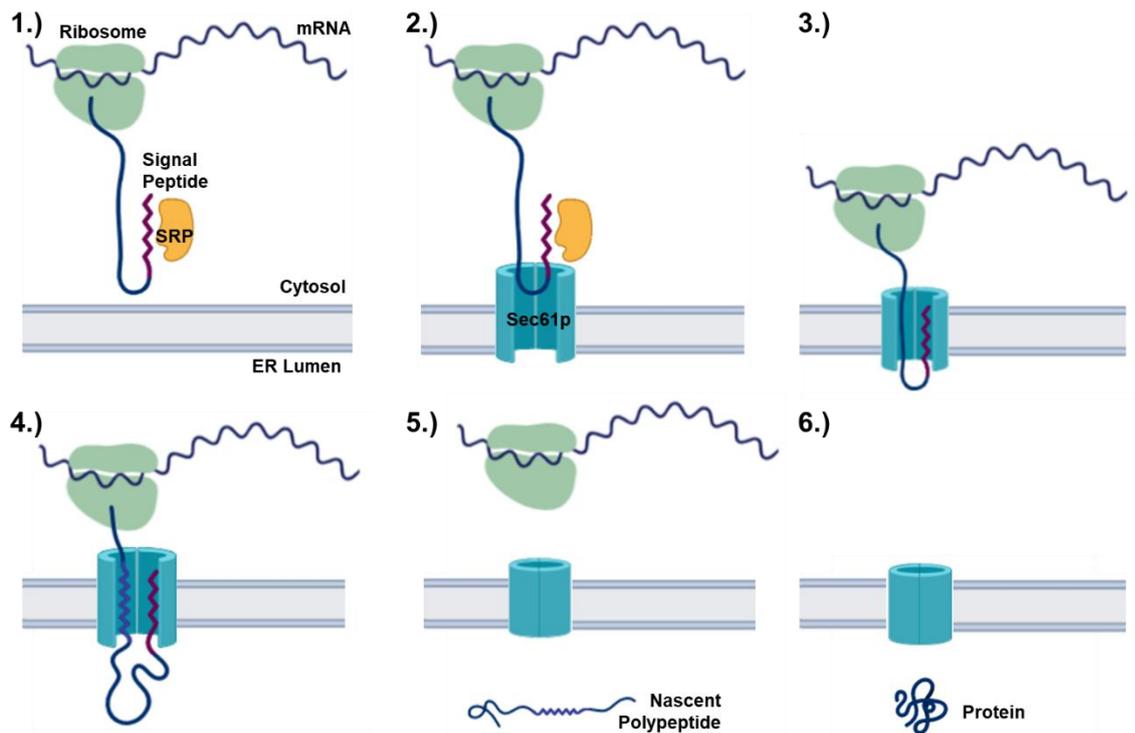


Figure 1.3: Overview of Translation at the Endoplasmic Reticulum (ER). 1) The signal peptide on a newly translated polypeptide is recognised by the signal recognition particle (SRP). 2) The SRP directs the ribosome to the ER membrane whilst translation continues. 3) The polypeptide gets translocated to the lumen of the ER through the Sec61p channel. 4) The ribosome resumes translation of the mRNA, feeding the polypeptide through the pore into the ER lumen. 5) For some proteins, the signal peptide remains attached, anchoring the protein to the ER membrane, whilst it is cleaved for others, releasing the polypeptide into the ER. 6) The polypeptide chain folds into its functional conformation and can undergo post-translational modification before sorting for their final destination in the endomembrane pathway.

1.1.5 Control of Eukaryotic Translation

Global translation initiation can be controlled by mechanisms such as phosphorylation of eIF2 or the eIF4E-binding protein 4E-BP, while translation of specific mRNAs can be controlled by interaction of protein factors with elements within the 3' UTR which can disrupt eIF4F association.

The availability of initiation factors is one such regulation mechanism. For example, phosphorylation of eIF2, as a response to stress, reduces eIF2 availability by inhibiting guanine nucleotide exchange which results in the inhibition of Met-tRNA binding to ribosomes, subsequently reducing global translational activity [62, 63]. Similarly, phosphorylation of 4E-BP by mTORC1 releases the cap-binding protein eIF4E, relieving translational repression. [62, 64]. These downstream effectors of the ERK and mTORC signalling cascades allows eIF4E activity modulation in response to external (insulin receptors and metabotropic receptors) and internal stimuli (cellular energy and oxygen levels via the AMPK pathway, and DNA damage, via p53 signalling) [65, 66].

The binding of PABP to the 3' poly(A) tail is now not thought to be essential for translation but rather stimulates translation [67, 68]. There are also some examples of mRNAs that lack poly(A) tails and are circularised to increase the efficiency of translation. For example, the 3' UTR of histone mRNAs are bound by the stem-loop binding protein (SLBP) which is subsequently bound by the SLBP-interacting protein 1 protein which bridges the 3' UTR to the cap-binding protein complex [69].

Additionally, 3' UTR-protein interactions can regulate initiation through the formation of an inhibitory closed loop that prevents access of eIF4F to the 5' end. In general, one protein recognises and binds a sequence specific motif in the 3' UTR region of the mRNA and an intermediate protein then binds and bridges this first protein to a third protein bind to the 5' m7G cap. This leads to the formation of a closed circuit [70-72].

Although less studied, there is also regulation of the elongation and termination steps of translation. Phosphorylation of eEF2 on threonine 56

at its GTP-binding domain by eEF2 kinase (eEF2K) prevents eEF2 from binding to the ribosome. The MAPK and mTOR pathways negatively regulate eEF2K whilst the AMPK and PKA signalling pathways stimulates eEF2K in response to stress e.g. starvation, hypoxia, oxidative stress [73, 74]. Through proteomic screening and immunoprecipitation studies Hizli et al, also identified a mechanism for eEF2 regulation involving phosphorylation of S595 by cyclinA-CDK2 which in turn recruits eEF2K for T56 phosphorylation [75].

The incidence of rare codons or complex secondary structure in the coding region of an mRNA slows elongation which in rare occasions can trigger a shift in reading frame, thus generating a secondary protein with different sequence and length than original unshifted protein. Frameshifting is most prevalent in RNA viruses, for example, the overlap of *pol* at the 3' end of *gag* in retroviruses, with *pol* in the -1 reading frame, results in a Gag-Pol polyprotein and allows the expression of the viral reverse transcriptase [76, 77]. Examples of frameshifting in eukaryotes include PEG10 and Ma3, however these are thought to be a remnant of retroviral origin [78, 79]. Similarly, the binding of a signal recognition particle to an ER-destined mRNA pauses elongation to allow the nascent protein to anchor onto the ER, where elongation resumes [80]. The rate of elongation is also thought to affect protein folding, with slowing at specific regions allowing proper folding of proteins in complex regions [81, 82]. Frame-shifting or read-through can also result from inhibition of termination. For example, incorporation of a selenocysteine amino acid can occur at a UGA stop codon which is influenced by the sequence context of the codon [83].

1.2 Eukaryotic mRNA turnover

The turnover of mRNA plays an important role in the regulation of gene expression. There are several mechanisms to achieving mRNA decay in eukaryotes, which are outlined in Figure 1.4. The deadenylation pathway is initiated by the shortening of the poly(A) tail followed by the removal of the 5' m⁷G cap which allows exonucleolytic degradation from either the 5' end of the transcript or through the more tightly controlled 3'-5' exosome-mediated decay. Deadenylation-independent decay can also occur where mRNA can be degraded through direct endonucleolytic cleavage of the transcript, or decapping of the 5' end without shortening of the poly(A) tail to permit 5' to 3' degradation.

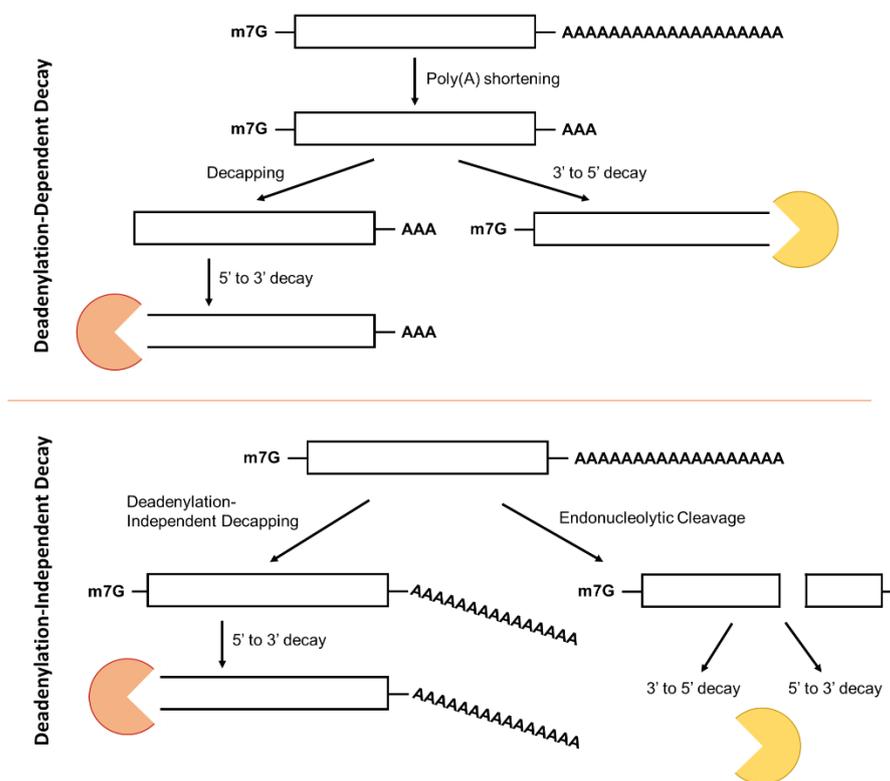


Figure 1.4: Overview of mRNA turnover pathways. The top pathway depicts deadenylation-dependent decay, in which following shortening of the poly(A) tail, either the 5' m⁷G cap of an mRNA is removed resulting to 5' to 3' decay by the XRN1 exonuclease, or 3' to 5' decay by the exosome complex occurs. The bottom pathway portrays deadenylation-independent pathways for mRNA decay, either rapid decapping resulting in 5' to 3' decay, or endonucleolytic cleavage leading to both 3' to 5' and 5' to 3' decay.

1.2.1 Deadenylation

It is thought that the shortening of the poly(A) tail is required as the first step in the decay of most eukaryotic mRNAs as this prevents PABP1 binding, resulting in both translational repression and release of the 5' cap exposing the mRNA to exonucleolytic degradation. In eukaryotes, it has been proposed that PABP-dependent Pan2/Pan3 complex acts early on in deadenylation, where it shortens the poly(A) tail of target mRNA without causing full mRNA degradation [84-86]. The CCR4–NOT complex is then suggested to continue the removal of the poly(A) tail [87, 88]. However, more recent data suggests the PAN and CCR4–NOT complexes may not act sequentially but may target different groups of RNA for degradation [89, 90].

The CCR4–NOT complex contains a large non-catalytic subunit, CNOT1, which is involved in selective recruitment to mRNA [91, 92]. A nuclease module containing both CAF1 and CCR4 subunits bind to CNOT1 via its MIF4G domain, but it is unknown whether these two nucleases have unique or overlapping/cooperative roles. In humans, biochemical studies have shown that CAF1 and CCR4 have distinct roles but that both are required and act in cooperation for mRNA deadenylation [92].

1.2.2 5' – 3' Exonucleolytic decay pathway

Most commonly, deadenylation is followed by removal of the 7-methylguanosine cap at the 5' end of the mRNA. The decapping complex Dcp2/Dcp1 is recruited to the RNA and has a high affinity for the 5' cap which leads to the displacement of the cap-binding complex eIF4F.

Dcp2 with additional cofactors, leads to the hydrolysis of the m7G cap producing m7GDP and 5' monophosphate RNA [93-95]. Many cofactors enhance the activity of Dcp1/Dcp2 such as Edc1, Edc2, Edc3, and the Lsm1–7 complex while some are able to directly inhibit the initiation of translation such as Scd6 and Stm1 which are reviewed extensively by Li & Kiledjian [96]. Pat1 has been shown to bind both the Dcp1/Dcp2 complex and the CCR4–NOT complex, linking decapping to deadenylation [97, 98].

This removal of the 5' m7G cap exposes the 5' end of the mRNA to degradation by the 5'-3' exonuclease Xrn1 [99, 100]. Xrn1 is highly conserved within eukaryotes at the N-terminal region and catalytic site [101-104]. It is the key enzyme of cytoplasmic mRNA decay, and is also involved in several specific biological pathways such as nonsense-mediated decay and microRNA-mediated decay [105].

1.2.3 3' -5' Exonucleolytic decay pathway

Following deadenylation, the cytoplasmic exosome degrades the 3' end of the mRNA [106, 107]. The exosome is a 3' to 5' exoribonuclease complex containing nine core subunits and a number of additional subunits [108], some of which may also possess endoribonuclease activity [109-111].

The exosome releases cap structures and capped oligonucleotides of less than 10 bases which are capable of being degraded by the DcpS scavenger decapping enzyme [112]. DcpS cleaves between the γ and β phosphates of the cap structure leading to decay [112]. The exosome is not only involved in general mRNA turnover, but also the specialised degradation of AU-rich elements (ARE)-containing RNA transcripts and RNA surveillance discussed briefly below [113-115].

1.2.4 Targeted mRNA Decay

Although most transcripts undergo deadenylation-dependent decay as described above, certain mRNAs are targeted by alternative mechanisms, some of which will recruit the general mRNA turnover machinery.

Several endonucleases capable of internal cleavage have been identified whose activity results in both 5' and 3' mRNA fragments that are vulnerable to degradation by the XRN1 and exosome complex as discussed previously. PMR1 has been shown to possess endonucleolytic cleavage activity but acts on polysomes and specifically targets translating mRNAs [116, 117]. Ire1 on the other hand, also has endonucleolytic cleavage activity but is involved in the ER stress response, specifically targeting ER transcripts during the unfolded-

protein response [118, 119]. There are several mechanisms to degrade aberrant transcripts, such as those containing premature stop codons, to protect the cell from potentially toxic proteins, such as nonsense mediated decay (NMD) and nonstop decay, that involved recruitment of aforementioned decapping and 5' exonuclease complexes [120].

Adenylate-uridylate-rich elements (AREs) are regions of adenine- and uridine-rich bases found in the 3' UTR of mRNAs that target the mRNA for degradation. They are often found in proto-oncogenes, cytokines, and transcription factors. RNA binding proteins such as TTP, AUF1 and HuR have been shown to recognise and bind AREs; AUF1 has been shown to interact with PABP and eIF4G, TTP recruits exosomes, and conversely HuR has been shown to stabilise its mRNA targets in response to genotoxic stress [121]. Additionally, RNA modifications such as N6-methyladenosine (m6A) are recognised by YTH domain proteins and alter the stability of the target transcript. YTHDF2, for example, directly recruits CCR4-NOT and XRN1 to m6A-containing transcripts to promote deadenylation and decay.

1.2.5 P Bodies

Translationally repressed mRNAs form complexes with the decapping machinery that accumulate in cytoplasmic foci termed P bodies to facilitate degradation. In mammalian cells, it has been shown that P bodies contain a high concentration of decapping enzyme DCP1/2, the 5' to 3' exonuclease XRN1, and several decapping activators, alongside the CCR4-NOT complex [95, 122-125]. Some proteins involved in nonsense-mediated decay are observable in P bodies under some conditions, such as stress or are cell-type dependent. In addition, P bodies of metazoa have been shown to contain miRNA and miRNA-associated repression factors, such as AGO and TNRC6 [126-129]. Transcripts that associate with P body components can also exit and re-enter translation under growth/stress conditions, possibly acting as a buffering system to balance the rate of mRNA degradation and transcription [130-132].

1.3 RNA Interference: An overview

Short interfering RNAs (siRNAs), microRNAs (microRNAs), and piwi-interacting RNAs (piRNAs) are all found in eukaryotes and control gene expression through silencing in a phenomenon known as RNA interference (RNAi). RNAi has been identified to be a mechanism of regulation for a range of cellular processes, including cell growth, development, and differentiation.

1.3.1 Small Regulatory RNAs

The three classes of small regulatory RNAs can be distinguished by the subset of Argonaute proteins they associate with; piRNAs bind to the Piwi clade of Argonaute proteins whilst siRNAs and microRNAs associate to the Ago clade. Additionally, piRNAs differ in their biogenesis and conservation, and are reviewed by C.D. Malone and G. J. Hannon [133]. Briefly, they are derived from single-stranded precursors and primarily function in germline cells of animals, whereas siRNAs and microRNAs are derived from double-stranded RNA precursors.

The first microRNA (microRNA) was identified in *C. elegans* in 1993 by Lee *et al* [134]. They isolated lin-4 as a negative regulator of lin-14 at the mRNA level through antisense complementarity in the 3' UTR. Then in 2000, let-7 was similarly identified in *C. elegans* as a regulatory RNA with antisense complementarity to the 3' UTRs of lin-14, lin-28, lin-41, lin-42, and daf-12 [135]. However, it wasn't until 2001 that the term microRNA was coined as a result of several sequencing studies that identified microRNA across a wide range of organisms [136].

Early papers described microRNAs as short, approximately 21 nucleotides (nts) long, single-stranded RNAs that act as regulators of other RNAs in the context of development [137, 138] but have since been identified as post-transcriptional regulators in a wide range of biological pathways and implicated in number of diseases [139]. According to miRbase edition 22, there are currently 48,860 mature microRNA

identified in 227 species, with 2654 mature sequences in humans alone, that are thought to regulate up to a third of our genes [140].

MicroRNAs have been shown to be key post-transcriptional regulators of gene regulation and function either by translational repression of target mRNAs or degradation of mRNAs through deadenylation [141-143]. To do so, microRNAs must associate with an Argonaute protein to form an RNA-induced silencing complex (RISC), discussed in Section 1.4.1.

In animals, the microRNA directs the RISC to target mRNAs by binding with imperfect complementarity to sequences in their 3' UTR [144]. There is some evidence for microRNAs binding outside the 3' UTR, but with a few exceptions such as miR-122 and HCV, discussed later, there is little evidence for a significant regulatory effect in other locations [145-147]. Target mRNAs often have multiple sites for a number of microRNAs that usually act co-operatively to mediate repression [148]. In addition, a single microRNA is thought to be able to regulate over 200 mRNAs and affect a range of cellular pathways.

Short-interfering RNAs (siRNAs) are roughly the same length as microRNAs and similar in structure. Whilst microRNAs are endogenously encoded in the genome, siRNAs can be produced from an externally introduced dsRNA such as foreign nucleic acids from viruses and/or transposons [149, 150]. However, it has been shown that siRNAs are not solely derived from exogenous long dsRNA but can also originate from endogenous genomic sites such as heterochromatin and other repetitive sequences [150, 151]. Figure 1.5 summarises the differences in siRNA versus microRNA biogenesis and mechanism of action. Plant endogenous siRNAs have also been identified which originate from isolated gene transcripts and have specific mRNA targets for silencing. They typically function by perfect complementarity and mRNA cleavage but can also repress translation via imperfect complementarity [152]. Furthermore, siRNA can be chemically synthesised and introduced into cells for research and therapeutic purposes [153].

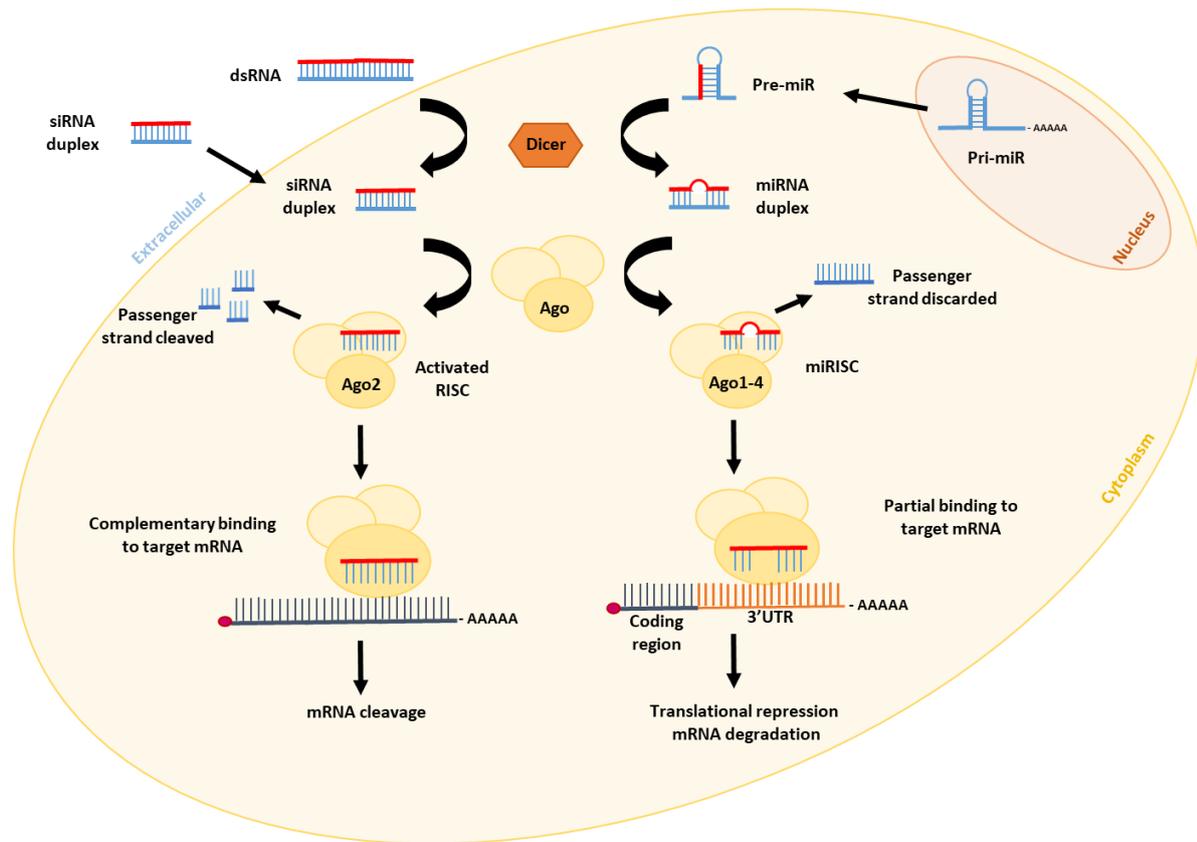


Figure 1.5: siRNA versus microRNA in metazoa. siRNA or the dsRNA they are derived from can come from external sources, whilst the pre-miR transcript is processed from the pri-miR transcript in the nucleus. The Dicer protein is common to both siRNA and microRNA formation, processing dsRNA into a duplex for siRNA or a pre-microRNA stem-loop structure into the microRNA duplex. The active form for both is loaded onto the RISC which can contain Argonaute proteins 1-4 for miRISC and specifically AGO2 for siRNA. The passenger strand is cleaved for siRNA, or simply discarded for loading of microRNA. The activated siRNA-containing RISC forms complementary binding to its mRNA target leading to mRNA cleavage, whilst the miRISC has partial complementarity to its target, leading to either translational repression or mRNA degradation of its target.

1.3.2 Sequence conservation of microRNA

Many microRNA are highly conserved in term of their sequence, function and expression pattern. For example, the sequence of mature miR-122 is identical in all vertebrate species, from humans down to zebrafish [154]. Some microRNAs are expressed in a wide range of species as a family of related molecules, for example, the Let-7 family which is

expressed from mammals to *C.elegans* with slight sequence changes [155].

Although some microRNAs are highly conserved, many are unique to particular species; Bentwich *et al* [156] identified 89 novel microRNAs in humans, 53 of which are not conserved beyond primates. Sempere *et al* [157] used computational approaches to gain insight into the phylogenetic distribution of microRNAs and found a set of 18 microRNAs that emerged concurrently with the emergence of organs such as the brain and heart and theorised that microRNAs played a fundamental role in the evolution of higher organisms.

1.3.3 Conservation of microRNA targets

As microRNA are key post-transcriptional regulators, genes are under selective pressure to retain or avoid conserved miR-binding sites and as such microRNAs have had a profound impact on the evolution of 3' UTRs [158]. It has been shown that the presence and absence of target sites correlates to gene function where increased conservation and number of microRNA target sites are found in genes involved in developmental processes, whilst genes that are broadly expressed and not tissue-specific, such as ribosomal-associated genes and housekeeping genes, tend not to contain microRNA target sites and typically possess short 3' UTRs [159]. Stark *et al* also identified examples of mutually exclusive expression of microRNAs and their target genes; miR-1 is expressed solely in the muscle tissue, and muscle-specific gene sets showed significant avoidance for miR-1 sites [159]. Likewise, they recognised that central nervous system (CNS) genes most significantly avoided miR-124 regulation which is exclusively expressed in the CNS.

1.3.4 Biological Role of microRNA

Most microRNAs display only modest repression of their targets (typically <2-fold) and knock-down of individual microRNA often does not result in a discernible phenotype [160, 161]. MicroRNA are thought to regulate approximately 60% of the human genome, regulating an extensive range of biological processes such as cell proliferation, differentiation, and

survival in addition to having key roles in normal development and homeostasis, with aberration of miR-mediated regulation leading to disease. It is often the case that a single microRNA is able to target several mRNAs and a single mRNA can contain binding sites for several different microRNA. This alongside the strong evolutionary conservation of microRNA sequence and their 3' UTR target sites suggest that microRNAs act to fine-tune molecular pathways conferring robustness in gene regulation rather than acting as individual regulators of individual targets.

Both the spatial and temporal expression of microRNA are linked to their function. For example, several microRNAs have been shown to act as developmental switches [162]. In *C.elegans*, miR-61 was shown to determine vulval cell fate by regulating the expression of LIN-12/Notch and VAV in a feedback loop [163]. Another study shows miR-196 regulating the expression of several Hox genes involved in human limb development [164].

A number of microRNAs have been shown to be expressed only in specific organs such as muscle-specific miR-1 which plays a key role in the development and differentiation of smooth and skeletal muscles, and miR-126 which is expressed in the blood and involved in blood vessel formation [165, 166]. Ason *et al* [167] compared 100 different microRNAs and found variation in microRNA expression even between two closely related species; miR-145, miR-205, and miR-454a exhibit sequence conservation between the related species medaka and zebrafish but show different spatial expression with the most dramatic seen between the gut and head.

Expression of microRNA can also be rapidly regulated, for example, miR-155, whose expression is induced by NF- κ B in response to TLR4 signalling [168]. Targets for miR-155 include the anti-inflammatory molecules SOCS1 and SHIPS1. Conversely, miR-146a has been identified as a negative regulator of TLR signalling with its expression

leading to a reduction of pro-inflammatory cytokine and chemokines helping to control the duration of the inflammatory response [169, 170].

1.3.5 MicroRNA and Cancer

It is unsurprising that dysregulation of microRNA is associated with disease with the range of targets that microRNAs are known to regulate, and which are involved in a myriad of biological processes. Most well-studied is the link between microRNA dysregulation and cancer, with global downregulation of microRNAs associated with cancer [171].

The abnormal expression of microRNA in cancers is often attributed to the amplification or deletion of microRNA genes, the earliest evidence of which comes from B-cell chronic lymphocytic leukaemia where deletions at chromosome 13q14 were frequently observed in patients [172-174]. Aberrant expression of transcription factors can also alter the levels of microRNA expression in cells. For example, loss of function mutations of the transcriptional activator p53 and upregulation of MYC are common in cancers, which have been shown to be responsible for the widespread repression of microRNA [175, 176].

MicroRNA expression can also be altered by the changes to epigenetic regulation that feature in cancers with aberrant DNA methylation and histone acetylation of microRNA genes leading to the repression of microRNA with tumour suppressor roles [177]. Defects in the microRNA biogenesis pathway could also lead to abnormal expression of microRNA [178, 179]. For example, the downregulation of DICER has been observed in various types of tumour and is associated with poor prognosis [180-182].

Conversely, there are some well-studied examples of upregulation of oncogenic microRNAs (oncomiRs) [183, 184]. For example, MYC has been shown to increase expression of the miR-17-92 cluster alongside widespread microRNA downregulation which leads to enhanced cell proliferation and angiogenesis [185, 186]. This is attributed in part to the repression of tumour suppressor PTEN by miR-19 and the transcription factor p21 by miR-17 and miR-20a [187, 188].

1.3.6 Canonical Biogenesis Pathway

MicroRNAs are transcribed, typically by RNA polymerase II, to form primary microRNA (pri-microRNA) transcripts [189-191] that are generally capped and polyadenylated, reminiscent of protein-encoding mRNAs [189]. Whilst the majority of microRNAs are located within protein coding exons, many microRNAs have been observed in non-coding loci of the genome such as introns, lncRNA, and 3' UTR of protein-encoding genes [192], such as the miR-17-92 cluster [193],

The microprocessor, Figure 1.6, is responsible for processing the pri-microRNA transcript into a precursor-microRNA (pre-microRNA) intermediate of approximately 70 nt that contains a stem loop structure that occurs either co-transcriptionally or following transcription [194-196]. The class III RNase Drosha was the first protein of this complex identified through immunoprecipitation and knock down studies [197]. Subsequently, Han *et al* [198] identified DGCR8 as an interacting partner of Drosha using RNAi. Currently, data suggests a mechanism whereby DGCR8 is responsible for substrate recognition and directing Drosha cleavage to a specific site [198, 199].

The nuclear transport protein Exportin-5 is responsible for binding and transporting pre-microRNA to the cytoplasm [200-202]. Here, the pre-microRNA is taken up for processing by Dicer which cleaves the approximately 70 nt long pre-microRNA into a duplex of 21 nt [203-205]. The two microRNA strands generated by Dicer are designated -5p and -3p based on whether they originate from the 5' or 3' arm of the precursor respectively, with one favoured as the mature microRNA strand which is thought to be determined by thermodynamics [206].

In order to achieve mRNA silencing, microRNAs need to form an RNA-induced silencing complex (RISC), composed minimally of microRNA, GW182 and an Argonaute protein. RISC receives the mature microRNA from Dicer that is used as a guide to effect mRNA silencing [207, 208].

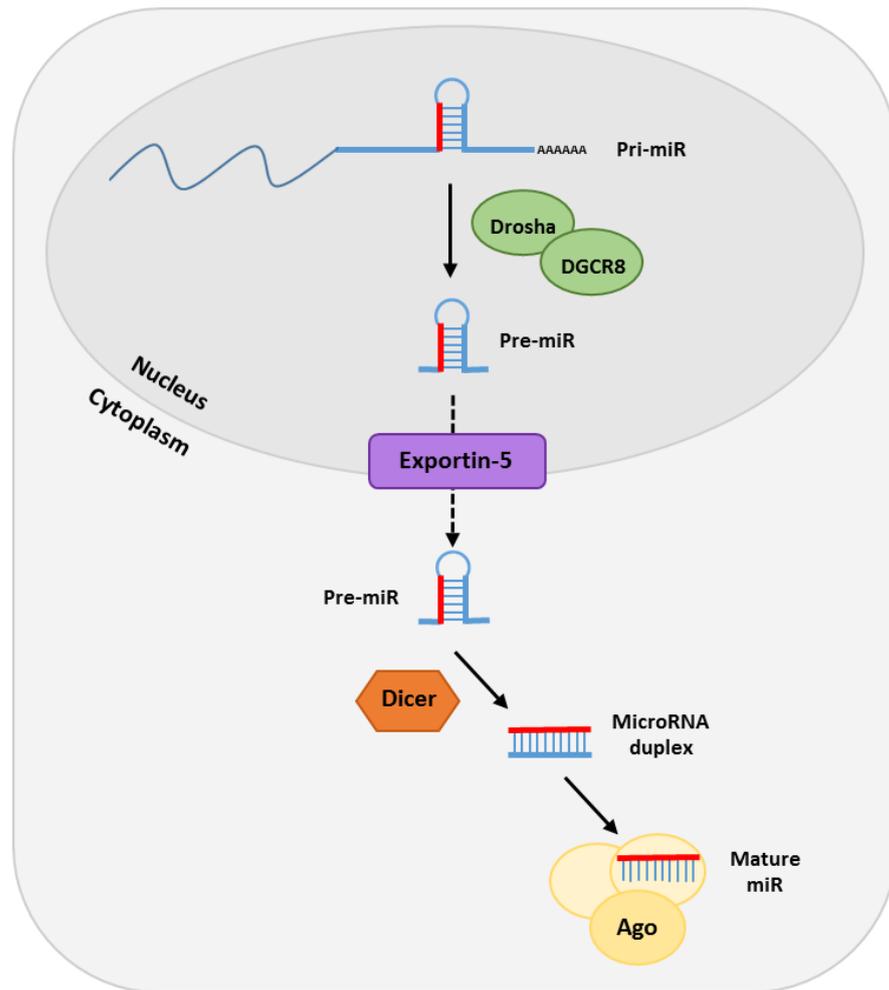


Figure 1.6: The canonical microRNA biogenesis pathway. The pri-microRNA transcript (pri-miR) is cleaved by Drosha/DGCR8 to produce the precursor microRNA (pre-miR) in the nucleus. This pre-miR is then transported to the cytoplasm by Exportin-5 where the enzyme Dicer cleaves it to form a microRNA duplex. Strand selection and loading onto the effector complex RISC is then facilitated by Argonaute proteins.

1.3.7 Non-canonical Biogenesis

There are some microRNAs that have been observed to reach their mature form using alternative mechanisms to the canonical biogenesis pathway. Ruby *et al* [209] identified a group of pri-microRNA transcripts located in introns, termed Mirtrons, which bypassed Drosha processing. They possess a stem loop that when spliced produces a 2 nt overhang at the 3' end that is able to be directly processed by Dicer.

Another microprocessor-independent pathway was identified for pre-miRNA that contain a 7-methylguanylate (m^7G) cap which occurs when the 5' end of the pri-microRNA coincides with the transcriptional start site and transcription termination occurs at the 3' end of the hairpin. This removes the requirement for cleavage by Drosha, and the pre-microRNA is exported via Exportin-1 which typically transports mRNA. The pre-microRNA is still open to processing by Dicer, however the guide strand selection is biased for the 3' strand [210].

So far, only one microRNA has been shown to be processed without the need for Dicer. MiR-451 was shown by Cheloufi *et al* [211] to undergo Drosha processing, but its maturation is Dicer-independent. They found instead that the pre-microRNA is loaded directly onto AGO2 and cleaved to produce a 3' end.

1.3.8 Regulation of microRNA Biogenesis

Due to a single mRNA being regulated by several microRNAs and the ability of a single microRNA to regulate several targets, microRNAs themselves are under tight regulation.

Computational studies have identified promoters in the pri-miRNA sequence similar to promoters found in protein-coding genes [212-214] and several transcription factors have also been shown to have an effect on pri-microRNA transcription [185, 215, 216].

Several accessory factors for the microprocessor complex have been identified that regulate the processing of the pri-microRNA transcript through facilitating Drosha recruitment [196, 217, 218] or targeting Drosha to the nucleus via phosphorylation [219, 220]. The processing of microRNA intermediates is often modulated by factors that bind to the terminal loop with one well-studied example being the processing of pri-let-7 which is inhibited by lin-28 binding to its terminal loop structure [221-226]. Similar to Drosha, a number of proteins have been shown to bind Dicer and affect the processing of the pre-microRNA including PACT, and the key RISC components Ago2 and TAR RNA binding protein (TRBP) [227-230].

1.4 MicroRNA function

MicroRNA are key post-transcriptional regulators and form an active RNA-induced silencing complex (RISC) alongside AGO and other RISC-associated proteins. They act by base-pairing to seed sequences at the 3' UTR of their target mRNA to induce gene silencing either repressing the translation of its target or through mRNA decay. The mechanisms of microRNA-induced silencing are not fully understood with different groups reporting different methods, however, what is currently understood about the process in animals is summarised herein.

1.4.1 Target recognition

Target recognition of mRNA targets is facilitated by Argonaute proteins. The Argonaute protein AGO2 was shown to be able to endonucleolytically cleave RNA [231, 232] whilst all four members of the Argonaute family in mammals (AGO 1-4) can function in microRNA-mediated repression [233-235].

In animals, microRNAs typically only have partial complementarity to sites within the 3' UTR of their mRNA target and partially base pair at nucleotides 2-7 or 2-8 of the microRNA seed sequence, sometimes accompanied by additional pairing to the region surrounding the seed sequence [236]. Lacking complementarity at nt 10-11, the RISC complex is unable to directly cleave the mRNA but instead facilitates gene silencing by recruiting additional proteins such as GW182 that can result in translational repression or mRNA decay. When full target complementarity is achieved by the microRNA, AGO2 can facilitate cleavage of the mRNA target [231, 237, 238]. Both canonical and non-canonical microRNA-target interactions have been extensively reviewed by Seok *et al* [239].

1.4.2 RISC composition

Biochemical and structural analysis show TRBP (trans-activation response RNA binding protein) to associate with both Dicer and AGO2

[230, 240]. More recent studies show that when AGO2, Dicer and TRBP associate and form the RISC-loading complex (RLC), the complex not only aids the assembly of RISCs but also RNA binding and determination of the product length [241-243]. Argonaute proteins of miRISCs that are bound to mRNA targets, have been shown to bind GW182 [244-246]. GW182, known as TNRC6A-C in humans, contains multiple tryptophan (W)-containing motifs which have been shown to bind to AGO2 at W-binding pockets in its PIWI domain [247-249].

Some indirect partners, who bind to the same RNA, can also influence the efficiency of mRNA silencing [126, 250, 251]. For example, the RNA-binding protein Pumilio has been shown to repress the translation of the transcription factor E2F3 by binding to its 3' UTR and stabilising the binding of microRNA to E2F3 mRNA [252]. Similarly, Lim domain-containing proteins (P body components) have been shown to bind both the 5' m7G cap and AGO1/2 concurrently to inhibit translation initiation [253] while Rack1, a 40s subunit component, recruits miRISCs to translating ribosomes to promote post-initiation translational repression [254]. The molecular chaperone Hsp90 has also been shown to interact with Argonaute proteins and has been proposed to promote an "open" conformation to aid in RISC loading [255, 256].

Furthermore, although GW182 binding by Argonaute is important for mRNA silencing, not all miRISC contain GW182. *In vivo*, Argonaute-bound microRNA predominantly do not associate with mRNA and GW182 and may be present as a stable reserve although they have been shown to reform into miRISC associated with GW182 upon physiological stimulation [257, 258].

1.4.3 MicroRNA-Mediated Deadenylation and Decay

There is widespread evidence that microRNAs effect degradation through the cellular 5' to 3' mRNA decay pathway (described in Section 1.2, of which TNRC6 is a key adaptor protein in recruiting deadenylase and decapping proteins.

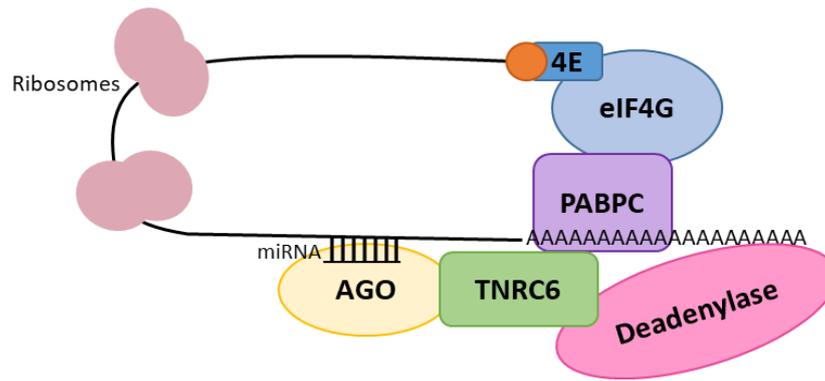


Figure 1.7: microRNA-mediated mRNA decay. The RNA-induced silencing complex is directed by the microRNA to the mRNA target and recruits the adaptor protein TNRC6. Subsequent recruitment of the deadenylase complex Pan2-Pan3 or CCR4-NOT leads to deadenylation at the 5' end of the mRNA target which triggers degradation of the mRNA.

To initiate microRNA-mediated degradation of an mRNA target, the RISC associates with the mRNA, guided by the microRNA, which in turn recruits TNRC6 via the Argonaute protein, see Figure 1.7 [84, 259-262]. Also referred to as AIN1/2 in *Drosophila* and GW182 in *C.elegans*, TNRC6 is present as three paralogs in mammals (A/B/C) and associates with the poly(A)-binding protein (PABP) at the mRNA poly(A) tail [263, 264]. The deadenylase complexes Pan2-Pan3 and CCR4-NOT are recruited to microRNA targets through interaction of TNRC6 to the PABP-dependent poly(A) nuclease (PAN3) and conserved W-containing motifs in the CNOT1/9 subunits of the deadenylase complexes, respectively [262, 265].

The subsequent deadenylation and degradation pathways are summarised in Section 1.2 but ultimately lead to the destabilisation of the mRNA. Whilst it is not agreed whether deadenylation or decapping is the primary trigger of microRNA-mediated mRNA decay in mammals, recent kinetic studies by Chen and Shyu suggest a biphasic deadenylation by both Pan2-Pan3 and CCR4-NOT complexes before Dcp1-Dcp2 complex-directed decapping occurs [85]. However, luciferase reporter assays by Nishihara *et al* suggested that the miRISC is capable of recruiting the decapping factors DCP1, Me31B and HPa to mRNA

targets, independent of ongoing deadenylation, and that when decapping is blocked deadenylated mRNA accumulates in the cell [266]. In agreement, Barisic-Jager *et al* showed that decapping enzymes are directly recruited to the miRISC rather than decapping occurring solely as a consequence of deadenylation [267] which may suggest that the promotion of decapping by the miRISC, independent of deadenylation, accelerates the degradation of the mRNA target.

1.4.4 MicroRNA-Mediated Regulation of Translation

There are also a number of studies to suggest microRNAs are capable of inhibiting translation of their mRNA targets and whilst some early studies concluded that microRNAs repress translation at elongation/termination, most suggest repression at the initiation step. There are three proposed mechanisms outlined in Figure 1.8 by which microRNA achieve inhibition of translation initiation: (1) dissociation of eIF4A, (2) poly(A)-binding protein (PABP) displacement, and (3) recruitment of translational inhibitors.

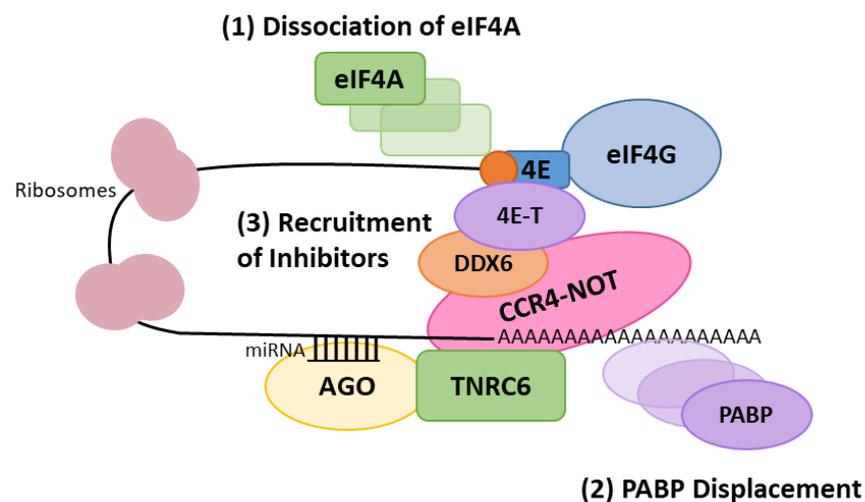


Figure 1.8: MicroRNA-Mediated Translational Repression. MicroRNA can repress translation via (1) displacing the RNA helicase eIF4A from the cap-binding complex, blocking ribosome recruitment and/or (2) displacement of the poly(A) binding protein (PABP). Alternatively, TNRC6 interaction with the CCR4-NOT complex can result in (3) the recruitment of translational inhibitors such as DDC6 and 4E-T.

It has been suggested that microRNA may block the assembly of the eIF4F initiation complex to elicit translational repression [268-272]. Pull-down assays in HEK293 cells, show the miRISC promotes the release of eIF4A1 and eIF4A2 from eIF4F, resulting in microRNA-mediated inhibition of translation initiation [273]. Additionally, treatment with silvestrol (eIF4A inhibitor) relieved microRNA-mediated repression. These data suggest that eIF4A1 and eIF4A2 are potential targets for microRNA in regulation of translation.

Some studies suggest that eIF4A1 and eIF4A2 possess distinct activities with eIF4A1 forming part of the eIF4F complex and eIF4A2 binding to the CCR4-NOT complex as part of microRNA-mediated translational repression [274-276]. Evidence from the Bushell group shows eIF4A2 binds mRNAs enriched in purine motifs and that repression via eIF4A2 is dependent on the CCR4-NOT component CNOT1. Through RIP-seq experiments, they showed mRNAs bound by eIF4A2 had sub-polysomal distribution upon depletion of CCR4-NOT, suggesting repression of translation initiation [277]. It is important to note that there are studies that did not see an association of eIF4A2 with the CCR4-NOT complex however these contradicting experiments used overexpression of either full-length or truncated eIF4A proteins whilst the work supporting the role of eIF4A2 interaction with the CCR4-NOT in miRNA-mediated regulation was primarily focused on the endogenous complexes involved [278-280].

It is also thought that TNRC6 interacts with CCR4-NOT to repress translation of mRNA lacking poly(A) tails by subsequently recruiting translational inhibitors [84, 281-283]. Structural and mutagenesis studies by Mathys *et al* revealed that CNOT1 interacts with the DDX6 (a known translational inhibitor) via its MIF4G domain, and several others saw that this interaction is required for microRNA-mediated translational repression [262, 278, 284]. The eIF4E-binding protein 4E-T has also been proposed to associate with CCR4-NOT, with depletion of 4E-T corresponding to an increase in global translation and is thought to compete with eIF4G for binding to eIF4E to enhance the repression of mRNA targets [272, 285-288].

Furthermore, the interaction between TNRC6 and PABP has also been well documented and shown to be involved in microRNA-mediated translational repression. Crystal structures and mutagenesis studies have demonstrated that TNRC6 binds to the C-ter domain of PABP via a PABP-interacting motif 2 (PAM2) in its C-terminus and that disruption of this interface impedes microRNA-mediated silencing [289-292]. It is proposed that TNRC6 recruits the CCR4–NOT complex to the poly(A) which displaces PABP, disrupting the circularisation and facilitating deadenylation and translational repression [293, 294].

Overall, the mechanism by which microRNAs facilitate translational repression is not fully understood, it may be that these three overarching mechanisms may occur concurrently or overlap, with different groups reporting different findings in different experimental contexts.

1.4.5 Relative contribution of mRNA decay vs translational repression

Whilst the concentration of RISC components, cell type, and the number/position of microRNA-binding sites all need to be considered, it has been suggested that degradation of mRNA mediated by microRNA has been shown to contribute most to overall gene silencing, with translational repression a consequence of mRNA decay.

Through the use of microarray analysis and polysome profiling, Hendrickson *et al* [295] measured the effect of miR-124 on its mRNA targets and the effect on translation, respectively. They found that mRNA levels decreased on average by 35% whilst translation was repressed by 12%, concluding mRNA degradation to account for the largest contribution to post-transcriptional regulation by microRNAs. Using ribosome profiling, Guo *et al* [296] found microRNA-mediated degradation of mRNA was responsible for 84% of silencing and Eichhorn *et al* estimated it was responsible for 66-90% [297]. Ding and Großhans, 2009 [298] found that in *C. elegans* repression of translation initiation, mediated by let-7 often coincides with target degradation and that both require the involvement of GW182. A quantitative mass spectrometry study by Baek *et al* [160], found that whilst a few microRNA targets were

translationally repressed without evident changes in mRNA levels, mRNA degradation provided the greatest contribution to repression for the majority of targets. Yet these studies did not examine the effect of microRNA function at early time points, instead when the cells were in a steady state.

Several kinetic studies established that although mRNA decay may have a greater contribution, the two are closely linked and decay actually follows translational repression [299, 300]. For example, the expression of miR-430 in Zebrafish peaks at 4 hrs post-embryo fertilisation and coincides with the highest reduction of active ribosomes bound to mRNAs, before mRNA deadenylation is concluded [301]. In agreement, ribosome profiling by Eichhorn *et al* [297], also showed that translational repression occurs promptly, but its overall affect is marginal once mRNA decay succeeds.

Whilst it is known that microRNA-mediated mRNA decay can occur independent of deadenylation [264, 302-304] it seems decay of the target is not always the outcome. Relief of microRNA-mediated translational repression has been reported under certain physiological conditions [305, 306] [131]. For example, Bhattacharyya *et al* [131] showed CAT1 mRNA reporters could be de-repressed by miR-122 under stress conditions, in conjunction with the release of CAT1 from P bodies and recruitment to polysomes.

Further studies looking at different organisms, cell types and physiological state are still required to understand the physiological consequences of micro-RNA mediated repression and the kinetics involved and to resolve some of the competing models.

1.4.6 Translational activation

It has been shown that under specific circumstances microRNAs are able to activate translation instead of repress. One such example, by Vasudevan *et al* [307], shows several microRNAs (let-7, miR-369-3 and artificial CXCR4 siRNA) are able to recruit AGO2 to AU-rich elements in the 3' UTR of TNF- α in order to upregulate translation in cells arrested in

the G0/G1 phase, but in proliferating cells are involved in translational repression. Other examples of microRNAs upregulating translation have also been reported when they interact with the 5' UTR of target mRNAs. Orom *et al* [308] showed miR-10a is able to prevent translational repression of ribosomal proteins in response to amino acid deprivation by binding to the 5' UTR of its target mRNAs.

The role of miR-122 in upregulating HCV RNA in liver cells is possibly the most well studied example of an microRNA being involved in positive translational regulation with a mechanism distinct to its repression of endogenous target mRNA. By binding to two sites in the 5'UTR of HCV RNA, miR-122 is thought to play a role in stabilising the IRES structure and recruiting IRES-stabilising molecules in addition to promoting virus replication rather than directly affecting the translation of HCV. As such there is no evidence for a distinct RISC which typically is involved in miRNA-mediated negative regulation, rather miR-122 recruits Ago2 to the target as a simple complex with some evidence for the involvement of TNRC6 and eIF4A2 as well as IRES- trans-activating factors [309-317]. The unique interaction between miR-122 and HCV RNA is discussed in more detail in Section 1.6.3.

1.4.7 microRNA Turnover

When mature microRNAs are incorporated into the RISC, they are thought to generally be stable with a half-life of >24 hours [318-320]. However, microRNA turnover could be modulated as a means to regulate microRNA expression.

Studies have shown turnover of specific microRNAs to be directed by cell cycle. For example, miR-29b is rapidly degraded in cycling HeLa cells and enriched in mitotic cells whilst miR-29a, which differs in sequence only at position 10, is highly stable [321]. Growth factors have also been shown to augment microRNA turnover. In MCF10A and HeLa cells, epidermal growth factor (EGF) stimulated a global decrease in microRNAs, a number of which target genes involved in the EGFR pathway, suggesting the presence of a feedback loop [322, 323]. In

humans, terminal uridylyl transferases have been shown to add oligoU tails to microRNAs to promote degradation, and miR-122 stability is increased by GLD-2-mediated 3'adenylation [324].

Most microRNA degrading enzymes, or microRNAses, are 3'-5' or 5'-3' exonucleases which have been shown to be highly conserved in eukaryotes [323]. One example of a microRNase is polyribonucleotide nucleotidyltransferase1 (PNPase), which in human melanoma cells has been shown to degrade mature microRNA in response to interferon stimulation without affecting the levels of pre- or pri-microRNA [325, 326]. Bail *et al* [221], showed direct regulation of miR-382 by RRP41 (the non-catalytic component of the 3'-5' exoribonuclease complex) and XRN1 mediated by Dicer processing.

1.4.8 MicroRNA Sponges

miRNA sponges are post-transcriptional regulators of microRNA which are RNA transcripts that contain multiple binding sites for miRNA thereby sequestering microRNA from its target mRNA [327-330]. There are several classes of endogenous RNA transcripts that can act as miRNA sponges including endogenous mRNAs, long non-coding RNAs (lncRNAs), pseudogenes, and circRNAs that are reviewed by Alkan & Akgul [331].

One such example of a protein encoding transcript acting as a miRNA sponge is the non-coding transcript containing the CD44 mRNA 3'UTR sequence. This transcript competes for miR-216a, miR-330, and miR-608 which results in the upregulation of CD44 and CDC42 to induce metastasis in the breast cancer cell line MT-1 [332]. LncRNAs are defined as transcripts exceeding 200 nucleotides that are not translated into proteins; One of the first lncRNAs to be described as a miRNA sponge is HULC, a hepatocellular carcinoma upregulated long noncoding RNA, which sequesters miR-372 and results in the de-repression of its target gene PRKACB resulting in tumourigenesis [333]. Pseudogenes on the otherhand, are oligonucleotide sequences that possess a frameshift mutation or premature stop codon but otherwise resemble protein-coding

transcripts. PTENP1 shares 98% homology with the sequence for the tumour suppressor PTEN mRNA and harbours binding sites for a number of miRNAs. Pandolfi et al showed that overexpression of PTENP1 acts as a decoy to increase the expression of PTEN and suppress tumour proliferation [334].

The final group of miRNA sponges are circRNAs which are non-coding RNA transcripts that possess a covalently closed circular structure and therefore lack both a 3' poly (A) tail and 5' cap [335-338]. Memczak et al found that the circRNA CDR1as was densely bound by AGO and sequence analysis identified 63 binding sites for the highly conserved miR-7. Furthermore, they showed that overexpression of CDR1as had a similar effect in zebrafish as knocking down miR-7 resulting in impaired midbrain development [339].

1.5 MicroRNA Subcellular Localisation

Whilst the understanding of microRNA function is progressing, the subcellular sites of microRNA action remain unclear. However, one can argue that it is essential to understand the whereabouts of microRNA and miRISC in order to truly elucidate microRNA function. There is some evidence to suggest different functions in different locations for individual microRNAs and this could influence some of the different observations about microRNA mechanism.

1.5.1 Interaction between microRNA and P Bodies

Using fluorescent microscopy, mammalian Argonaute proteins were shown to predominantly localise to P bodies (PBs) [123] which are also rich in GW182 and mRNA decay factors such as deadenylases. PBs were shown to have a ten times higher concentration of AGO2 than the cytoplasm but accounted for <1% of total cytoplasmic AGO2 [340]. Therefore it has been proposed that P bodies, distinct foci in the cytoplasm, could be the central site of microRNA-mediated mRNA silencing [341]. On the other hand, depletion of PBs was shown to have no effect on both micro-RNA mediated translational repression and mRNA decay [340, 342].

Alternatively, PBs are thought to store mRNA before a return to active translation. For example, Bhattacharyya *et al*/ showed that miR-122 binds to the 3' UTR of CAT1 mRNA sites to target the mRNA to P-bodies and repress the translation [131]. Conversely, they found that under stress conditions the RNA binding protein HuR binds to the 3' UTR of CAT1 at sites adjacent to the miR-122 binding sites, inhibiting the association of miR-122. This prevents miR-122 mediated translational repression of CAT1 mRNA and translation of CAT1 mRNA recommences following relocation from PBs to polysomes.

However, neither poly(A) tail containing mRNA nor PABP has been detected in PBs [123, 343]. Therefore, as poly(A) tails are required for

translation, the microRNA-targeted mRNA would have to be re-adenylated prior to resumed translation, or more likely the observations are a result of non-PB transcripts relocating to the polysomes.

1.5.2 microRNA and RISC at the Endoplasmic Reticulum (ER)

Previously it was assumed that micro-RNA mediated silencing occurs predominantly in the cytosol, however recent evidence identifies increasing links between endomembranes and microRNA in plants and animals.

Mammalian AGO2 was initially identified by Cikaluk *et al* [344] as the ER- and Golgi- associated protein GERP95, and Dicer has also been shown to associate with the ER and Golgi [345, 346]. Suzawa *et al* performed western blots on fractions obtained from HeLa cells using a digitonin-lysis method, and whilst TNRC6A, AGO2, and CNOT1 were preferentially observed in the cytoplasm, small amounts of all three key RISC components were detected in nuclear and ER-containing fractions [347].

In *Arabidopsis*, the ER-bound integral membrane protein AMP1 associates with AGO1 and was shown to regulate microRNA-mediated translational repression [348]. This study shows that although in plants regulation of mRNA levels predominantly occurs through microRNA-mediated cleavage of mRNA transcripts, translational inhibition can also be facilitated by microRNA. In *Drosophila*, Wu *et al* [349] identified an alternative form of RISC, P-miRISC that is formed upon serum starvation and results in up to ten times stronger repression of target mRNA. They demonstrated that this complex is devoid of GW182, but instead AGO1 binds Loqs-PB, and that it associates with the ER.

Using immunohistochemistry, Antoniou *et al*, found that components of the RISC-loading complex (Dicer, TRBP, and PACT) primarily associate at the endoplasmic reticulum in the soma and dendrites of primary neurons [350]. Utilising a digitonin-lysis fractionation method, they showed that short-term stimulation with BDNF (brain-derived neurotrophic factor) is capable of inducing the dissociation of TRBP from Dicer and its redistribution from the ER to the cytoplasm. This led to a

reduction in the processing from pre-to mature of a subset of microRNAs including the miR-16 family, suggesting that these pre-microRNAs are predominantly processed at the ER. Similarly, Stalder *et al* [351], found that microRNA-loaded AGO2 localises to rough ER membranes and associates with RISC factor Dicer, and the human Loqs-PB analogue TRBP. They found that TRBP was required for anchoring the miRISC to the ER membrane and that the ER was the major site for both loading of the microRNA onto AGO2 and cleavage of the target mRNA.

Bhattacharya *et al* showed enrichment of AGO2 and microRNA at the ER of HEK293 cells using a selection of fractionation techniques including digitonin-lysis and sucrose gradients [131]. In addition, they showed preferential localisation at the ER for a miR-122 reporter in Huh7 cells along with association with miR-122 endogenous targets ALDOA and CAT1 and that treatment with thapsigargin (an ER stress inducer) reduced association of microRNA, AGO2 and targeted mRNA with the ER. They performed time-course experiments, using a tet-inducible synthetic miR-122 target, and showed that *de novo* synthesised mRNAs localise to ER-bound polysomes where they associate with AGO2 and microRNAs prior to microRNA-mediated translational repression.

The enrichment of both miRISCs and ribosomes at the rough ER suggests microRNA-mediated mRNA silencing can be achieved at a greater efficiency at the ER than in the cytoplasm, however the mechanism by which this is achieved and the consequences for regulation of cytoplasmic and ER-associated target mRNAs remain undetermined.

1.5.3 microRNA and the Endosomal Pathway

RISC components AGO2 and GW182 have been shown to co-sediment in endosomes and multivesicular bodies (MVBs) fractions from human cells and depletion of ESCRT (endosomal sorting complex required for transport), a key complex required for MVB formation, inhibits microRNA-mediated silencing [352]. Likewise, in *Drosophila*, GW182 bodies were shown to localise to the cytosolic side of MVBs and lysosomes and

blocking MVB turnover into lysosomes enhances mRNA silencing and promotes microRNA loading onto RISC [353]. Taken together, these data suggest the compartments of the endosomal pathway (Figure 1.9) may act as a central site for RISC loading.

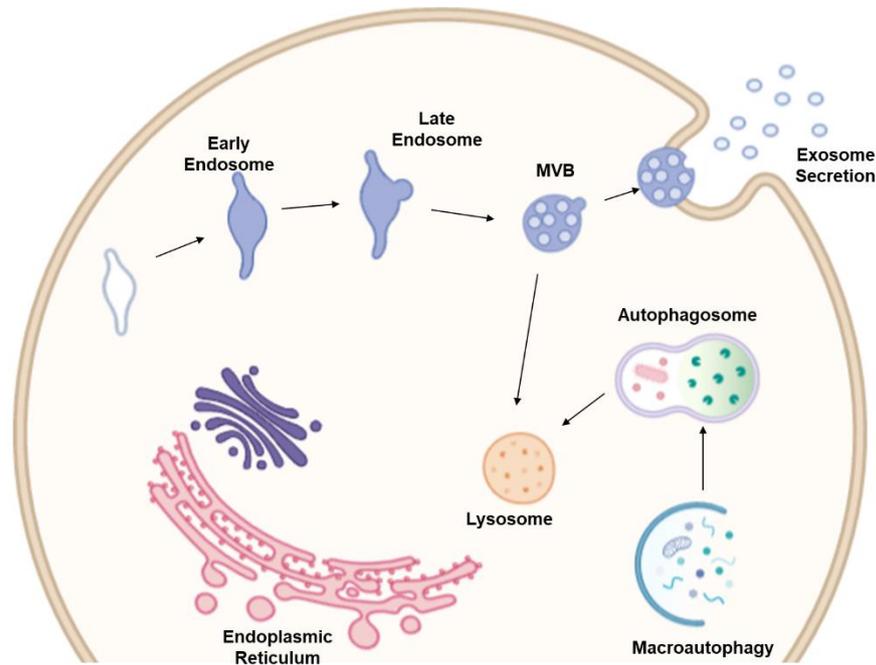


Figure 1.9: Overview of the Endosomal Pathway. Early endosomes receive extracellular molecules and are principally sorting organelles sending material on to either late endosomes for degradation/secretion or recycling back to the cell surface. Molecules designated for secretion form vesicles within multivesicular bodies (MVBs) that fuse to the cell membrane and are secreted as exosomes. Molecules destined for degradation pass through late endosomes and multivesicular bodies (MVBs) and fuse onto lysosomes which are the principle site for intracellular digestion, containing over 40 different hydrolytic enzymes. Lysosomes can also receive cellular waste products from the autophagy pathway for degradation.

The Bhattacharya group showed that proteins involved in mRNA degradation have differential subcellular distribution [354]. AGO2, XRN1 and PABPC1 were shown to be evenly distributed between the cytoplasmic and membrane fractions in HEK293 cells, whilst DCP2 and CNOT4 were predominantly associated within the MVB/endosomal fractions. Using reporter constructs, they showed that microRNA-repressed mRNAs are targeted to late endosomes/MVB, seeing an

almost 10-fold enrichment of the reporter mRNA, which is coupled with dissociation with AGO2 and occurs prior to their degradation. Blocking the maturation of endosome to late endosomes/MVBs via siRNA-mediated knockdown of HRS (hepatocyte growth factor receptor substrate) increased the stability of the target mRNAs which they suggest is through reducing their accumulation at the late endosomes/MVBs and preventing dissociation of AGO2 from the target mRNAs required for degradation.

Extracellular vesicles are lipid-bound particles that are secreted into the extracellular space and include microvesicles, exosomes, and apoptotic bodies. Both exosomes and microvesicles are released by healthy cells and whilst microvesicles bud off directly from the plasma lipid bilayer, exosomes are a part of the endosomal pathway resulting from the fusion of MVBs at the plasma membrane. Exosomes have been shown to be enriched in AGO2 and GW182 and microRNA-rich exosomes have been shown to regulate gene expression in cancer, immune, and neuronal receptor cells [355-357]. Ghosh *et al* [358] found that sequestering miRISCs to polysomes prevents exosomal sorting of microRNAs. This led to an increase in levels of microRNA within growth-restricted mammalian cells but did not increase microRNA activity.

Autophagy is a process used for the degradation, turnover, and renewal of intracellular and cytosolic components and may be responsible for homeostatic regulation of microRNA-mediated mRNA silencing. Macroautophagy is initiated by the formation of an isolation membrane (or phagophore) that envelops cytoplasmic components to form an autophagosome that subsequently fuse with lysosomes, where the cytosolic cargoes are degraded. The origin of the isolation membrane is still unknown with a number of different membrane sources suggested including the ER, Golgi, mitochondria, and endosomes [359, 360]. Gibbins *et al* [361], showed that Dicer and AGO2 co-immunoprecipitate with the autophagy receptor NDP52 and co-localise with autophagosomes in human cells. In addition, they found that depletion of autophagy components ATG5, ATG6, ATG7 and NDP52 lead to an

increase in AGO1, AGO2 and Dicer levels but not GW182, suggesting Dicer and AGO2 are targeted for degradation by NDP52.

1.5.4 Localisation of microRNAs and RISC to the nucleus

A number of studies have detected specific microRNAs in the nucleus; Mature miR-21 was found in both the nuclear and cytoplasmic fractions, where 20% of total miR-21 was localised to the nucleus [231]. Here, Miester *et al* using a fractionation approach described by Dignam *et al* [362] to separate nuclei from the cell pellet, quantified expression levels using radioactive northern blotting. Another microRNA, human miR-29b, was predominantly found in the nucleus and contains a hexanucleotide sequence at its 3' end that was shown to direct microRNA nuclear import of microRNAs or siRNAs that it was attached to [363]. Castanatto *et al*, proposed a model for the stress-induced nuclear transport of microRNA, along with oligonucleotides and siRNA, by a complex they identified as the stress-induced response complex (SIRC) which includes AGO1/AGO2 and several transcription and splicing regulators [364].

Furthermore, screening studies have indicated that most microRNAs are actually found in both the cytoplasm and nucleus along with other RISC factors, namely Dicer, TRBP, and GW182 [365-367]. However, it cannot be determined in these studies whether these early detections were a result of contamination from the cytoplasm as these experiments were performed in dividing cells. Some microRNAs have even been shown to cleave RNA targets in nuclear fractions, indicating an association of the microRNA with AGO2 [231, 368]. Ohrt *et al* [369] identified that AGO2 exists in a 158 kDa RISC in the nucleus, much smaller than the 3 MDa cytoplasmic complex, and is composed of only AGO2 and a short RNA that is loaded in the cytoplasm prior to nuclear import. This was in agreement with Gagnon *et al* [367] who did not detect microRNA loading in nuclear fractions. Immunoprecipitation and knockdown studies indicated that Importin 8 is responsible for the nuclear import of AGO2 [370].

Sarshad *et al*, identified nuclear microRNA-AGO complexes in embryonic stem cells that associate with core RISC components TNRC6 and CCR4-NOT and function to initiate differentiation [371]. They saw that whilst cytoplasmic miRISC act on the 3' UTR of their target mRNA, nuclear AGO is also able to bind to the CDS and introns to stimulate gene silencing. Suzawa *et al* [347] previously showed that TNRC6A is capable of shuttling between the cytoplasm and nucleus of HeLa cells and followed-up with mass spectrometry analysis of TNRC6A interacting proteins identifying proteins involved in the RNA degradation pathway present in both nuclear and cytoplasmic fractions. This suggests that microRNA-mediated RNA silencing may occur in both the nucleus and cytoplasm, and as they categorised some TNRC6A-associated proteins into the spliceosomal pathway, it suggests TNRC6A may be involved in regulating nuclear splicing.

1.5.5 Targeting of pri-microRNAs in the Nucleus

At this time, the function of nuclear microRNAs is not well understood. Chi *et al* [372] generated a genome-wide interaction map for several microRNAs present in mouse brains and identified that while most microRNA-target mRNA interaction sites followed the rules of canonical microRNA function, occurring in the 3' UTR region of target mRNAs, 12% mapped to nuclear-retained intron sequences. Moreover, 4% of such sites mapped to long non-coding RNAs (lncRNAs), suggesting a role of microRNA in regulation of the non-coding transcriptome. Alternatively, these could correspond to the presence of miR sponges in the nucleus which were discussed in detail in Section 1.4.8. Several examples of microRNA directly targeting non-coding transcripts have since been identified [373-375] and also an example of a nuclear microRNA regulating the biogenesis of another where miR-709 is a regulator of pri-miR-15a/miR-16-1 transcript [376]. More recently, pre-miR-122 has been shown to downregulate translation of an mRNA target, by targeting Insig1 pre-mRNA in the nucleus and influencing poly-adenylation site usage [377].

Similarly, in *C.elegans* there has been an example of autoregulation by a microRNA in a positive feedback loop suggested in Figure 1.10. Zisoulis *et al* [378], showed that in nuclear fractions ALG-1, an argonaute homolog, associates with the primary transcript for let-7. They went on to identify a conserved site at the 3' end of the primary transcript which mature let-7 binds in order to promote ALG-1 mediated processing of the primary let-7 transcript.

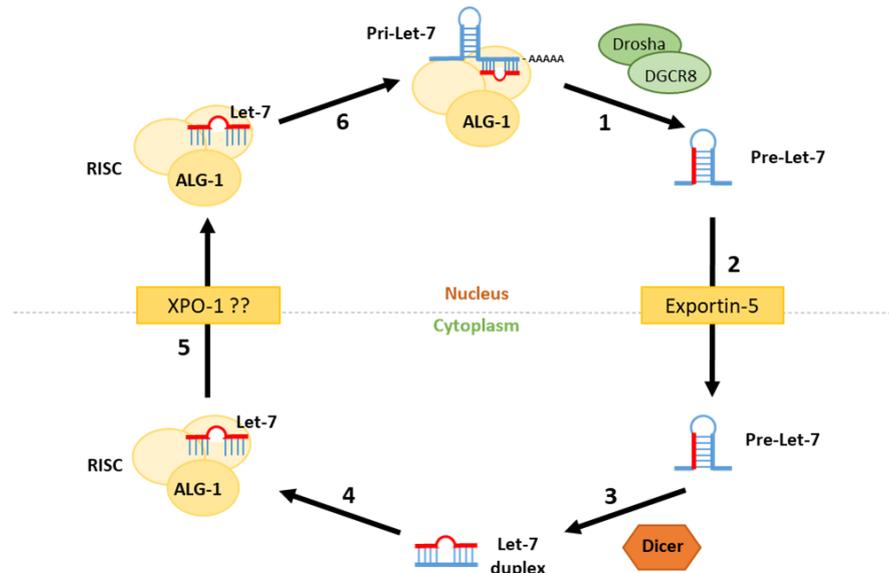


Figure 1.10: ALG-1 regulation of Let-7 primary transcripts. (1) Processing of pri-let-7 transcript to precursor transcript by Drosha/DGCR8. (2) Export to the cytoplasm through Exportin-5. (3) Processing of Pre-let-7 to the mature Let-7 form. (4) Association with ALG-1 to form the effector complex RISC. (5) Import into the nucleus, possibly through XPO-1. (6) Binding of ALG-1 to pri-let-7 mediated through mature let-7 binding sites towards the 3' end.

Using microRNA tracing and *in situ* hybridisation, Wang *et al* [379] detected miR-122 in the nucleus and saw re-entry of the mature microRNA into the nucleus of Huh7 cells. And after seeing an inverse correlation between miR-21 and miR-122 expression, they identified a seed match for miR-122 within the 3' terminal region of the pri-miR-21 transcript. Using reporter constructs, they showed that binding of miR-122 to this region prevented the binding of Drosha-DGCR8 and therefore inhibited the processing of the primary transcript into pre-miR-21 by the

microprocessor complex, providing further evidence of mature microRNA regulating the processing of pri-microRNA.

1.5.6 Other Nuclear Functions

Kim *et al* [380] searched for microRNA target sites in promoter regions within the human genome which led them to identify miR-320 as a cis-regulator of cell cycle gene POLR3D. By recruiting AGO1 and Polycomb group component EZH2 to the promoter for the gene, miR-320 is able to cause the formation of heterochromatin and POLR3D gene silencing. Initial screening studies exposed several microRNA seed matches present in the human genome [381] and since then several other examples of epigenetic silencing at promoters have been described, regulating a wide range of cellular processes [382-385].

It has been proposed that microRNAs may elicit transcriptional activation of target genes by silencing lncRNA transcripts at the promoter that act as negative regulators, thereby “repressing a repressor” [386, 387]. An alternative proposal involves recruitment of positive epigenetic regulators by miRISC to target promoters. One example is the regulation of pro-inflammatory gene cyclooxygenase-2 (COX-2) by endogenous miR-589. Binding of miR-589 to the COX-2 promoter triggers COX-2 transcription through the recruitment of AGO2 and TRNC6A and enrichment of active histone modifications [388].

Several studies support a role of nuclear microRNA in the regulation of alternative splicing. MicroRNA targeting sequences near to an alternative splice site were shown to affect splicing decisions, the effect of which is dependent on the heterochromatin status and availability for RNA polymerase II processivity [389]. Furthermore, AGO1, AGO2, and Dicer were shown to be required for this regulation of alternative splicing and chromatin immunoprecipitation (ChIP) data showed an association between the argonaute proteins and several splicing factors [390]. It may be that the formation of such complexes conceals splicing recognition motifs, thereby preventing the binding of splicing factors.

1.5.7 Nucleolar microRNA

Several microRNAs have been reported in the nucleolus, however their actions there remain unexplained [391-395]. For example, Li *et al* [392] identified eleven microRNAs abundant in the nucleolus of HeLa cells using *in situ* hybridisation, including miR-191 and miR-484. They found that the nucleolar presence was influenced by dsRNA and more interestingly viral infection, which led to the redistribution of nucleolar microRNA to the cytoplasm. Furthermore, treatment with Leptomycin B (known to affect Exportin-1 dependent nuclear export) abrogated the observed redistribution between nucleolus and cytoplasm, suggesting a role for Exportin-1 in regulating the trafficking of nucleolar microRNAs to the cytoplasm. In agreement, Castanotto *et al* [396] found that Exportin-1 co-immunoprecipitates with AGO1 and AGO2 and that Leptomycin B treatment prevented export of miR-16 and miR-29b to the cytoplasm.

Interestingly, Reyes-Gutierrez *et al* [397] reported the presence of spliced transcripts encoding insulin-like growth factor 2 (IGF2) in the nucleolus and through bioinformatics analysis predicted binding sites for five nucleolar microRNAs in the 3' UTR of the IGF2 mRNA. This suggests that the nucleolus could act as a site for mRNA-microRNA interactions before the export of mRNA to the cytoplasm.

1.5.8 Localisation of microRNA and RISC at the Mitochondria

A large number of microRNA have been identified in mitochondria, isolated from a broad range of samples (e.g. rat, mouse, human tissue, human cells) [398-402], and as mitochondria possess their own genome, it has been proposed as a site for microRNA-mediated post-transcriptional regulation.

Bandiera *et al* [403], identified 57 microRNAs highly expressed in the mitochondria of HeLa cells and co-localised with both AGO2 and the mitochondrial transcript COX3 indicating a possible role of miRISC in the regulation of mitochondrial mRNA targets. Zhang *et al* [404], reported that AGO2 but not AGO1 or AGO3 was imported to the mitochondria, where the levels of mitochondrial AGO2 compared to total AGO2 increased from

13% to 33% after differentiation. They showed that miR-1 increases the translation of mitochondrial cytochrome c oxidase subunit 1 (mt-COX1) and mitochondrial-encoded NADH dehydrogenase 1 (ND1) rather than repressing, in a mechanism dependent on AGO2 but not GW182. This translational activation was abrogated upon AGO2 knockdown and treatment with chloramphenicol (a mitochondrial translation inhibitor) and rescued with AGO2 containing a mitochondrial targeting signal.

In agreement, Das *et al* [405], also found AGO2 (but not Dicer nor TRBP) to be present in mitochondria of cardiomyocytes and co-immunoprecipitate with miR-181c and mt-COX1. They also found that miR-181c overexpression did not alter the mRNA level but did decrease the protein level of mt-COX1, showing that mitochondrial miR-181c acts as a post-transcriptional regulator of mitochondrial COX1 mRNA. Further work showed altered levels of mt-COX1 in isolated heart mitochondria upon *in vivo* delivery of miR-181c in rats that resulted in remodelling of complex IV and increased reactive oxygen species production [406].

1.5.9 Cytosekeletal association of miRNA

The intracellular trafficking of mRNA along elements of the cytoskeleton through directed active transport along microtubules, intermediate and/or actin filaments are usually achieved as part of ribonucleotide-protein (RNP) complexes where translation is inhibited. One well-studied example is ZBP1 (Zipcode binding protein 1) which binds to the 3' UTR of β -Actin mRNA to promote translocation of the transcript to sites of active actin polymerisation in order to modulate cell migration and neuronal differentiation in primary fibroblasts and neurons [407]. Hüttelmaier *et al* showed ZBP1 association prevented the formation of the 80S ribosomal complex which is relieved upon Tyrosine 396 phosphorylation of ZBP1 by protein kinase SRC [408]. Therefore, it could be possible for miRNA to be associated with such RNP complexes to repress translation of mRNA as they are trafficked along the cytoskeleton.

The actin cytoskeleton also contributes the structural integrity of neuronal synapses and the development and morphology of dendritic spines. As such, there are several examples of miRNA localised to dendrites that are involved in the regulation of synaptic development, maturation, and/or plasticity [409, 410]. One of the first microRNA shown to be enriched at synapses was miR-134 which is involved in the modulation of synaptic development. In Situ Hybridisation experiments by Schrott et al showed miR-134 to be localised to the synapto-dendritic compartment in hippocampal neurons [411]. They showed that miR-134 represses the translation of Limk1 mRNA whilst the transcripts are transported to the synaptic site in order to regulate the development of dendritic spines. Furthermore, using luciferase constructs they showed this regulation to be relieved by brain-derived neurotrophic factor (BDNF), released upon synaptic stimulation. Therefore, the cytoskeleton has the potential to be an important site of miRNA localisation and regulation, especially in neurons with examples of spatiotemporal regulation by miRNA in dendritic compartments.

1.6 miR-122 and microRNA therapeutics

1.6.1 miR-122 gene and expression

In humans, miR-122 is derived from a single locus on the positive strand of chromosome 18. Pri-miR-122 is initially synthesised as a 7.5 kb transcript that is spliced to yield a transcript roughly 4.5 kb in length and consisting of two exons with the pre-miR-122 hairpin located at the 3' end of the terminal exon [412], as shown in Figure 1.11. Transcription is terminated by microprocessor cleavage and pri-mi-122 is capped but not polyadenylated [412]. The mature sequence for miR-122 was shown to be conserved in twelve different species from mammals to fish [154].

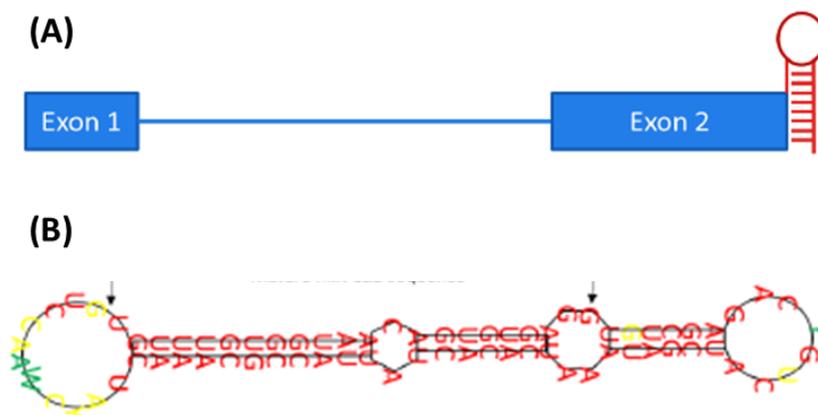


Figure 1.11: Structure of both pri- and pre-miR-122 transcripts. (A) Schematic of the pri-miR-122 transcripts with the pre-miR-122 hairpin in red located adjacent to the terminal hairpin. **(B)** Structure of the pre-miR-122 hairpin with the mature miR-122 sequence located between the two black arrows. Red nucleotides indicate high sequence conservation, yellow indicates moderate, and green indicates bases with the least conservation.

miR-122 is highly liver specific. Lagos-Quintana *et al*/showed it to account for 72% of the microRNA pool in mouse livers but to be absent in all other tissues tested [413]. The liver-enriched transcription factor hepatocyte nuclear factor 4 α (HNF4 α) controls the liver-specific expression of miR-122 by binding directly to a conserved upstream promoter region [312] and expression of miR-122 in the liver has been shown to increase with embryonic development [154].

1.6.2 Physiological Function and Pathology

The high expression level of miR-122 in the liver is associated with its role in the maintenance of normal liver metabolism. It was first identified as a regulator of the cationic amino acid transporter (CAT1) and the 3' UTR of CAT1 mRNA has since been shown to possess several miR-122 binding sites [154]. Bhattacharyya *et al* showed that miR-122 binds to these sites to repress the translation of CAT1 and target the mRNA to P-bodies [131].

Esau *et al* [414] inhibited miR-122 using antagomirs, and consequently identified a role for miR-122 in the regulation of hepatic fatty acid and cholesterol synthesis. Specifically, miR-122 has also been shown to regulate the expression of several genes involved in the regulation of cholesterol metabolism, including Aldolase A (ALDOA), Ndr3 and the 3-hydroxy-3-glutaryl CoA reductase (Hmgcr) [414-416]. Castoldi *et al* also showed direct binding of miR-122 to the 3' UTR of two genes involved in iron homeostasis, Hemochromatosis (Hfe) and Hemojuvelin (Hjv) [417].

Additionally, miR-122 expression is linked to mitochondrial gene expression, with miR-122 shown to interact with PPARGC1A (peroxisome proliferator-activated receptor gamma coactivator 1 alpha), a key regulator of mitochondrial biogenesis [418]. Not only is miR-122 responsible for the regulation of liver metabolism, but it has also been linked to liver maturation. Several targets of miR-122 (Cux1, Rhoa, Iqgap1, Mapre1, Nedd4l, and Slc25a34) are involved in the regulation of cytokinesis, and work with miR-122 knockout livers in mice demonstrated that miR-122 is required for hepatic polyploidisation [419].

Interestingly, the transcription of pri-miR-122 has been shown to be under control of the circadian transcriptional repressor REV-ERB α in mouse livers [319]. Although the mature miR-122 levels were unchanged likely due to its high stability, the levels of pri-miR-122 and pre-miR-122 were shown to fluctuate between four- and ten-fold in abundance throughout the day due to circadian transcription of the miR-122 gene. Gatfield *et al*

went on to show that miR-122 targets PPAR β/δ (peroxisome proliferator-activated receptor β/δ) and SMARCD1/BAF60a (a PPAR β/δ coactivator) responds to fluctuations of pri-miR-122 to regulate metabolism under circadian control.

Repressed levels of miR-122 have also been observed in hepatocellular carcinoma (HCC), with low miR-122 correlating to increased metastasis and an overall poorer prognosis [420]. This, together with its crucial role in liver homeostasis and maturation, has established miR-122 as a key intrinsic tumour suppressor within the liver [421-423]. Recent studies have also identified miR-122 involvement in extra hepatic cancers and inflammatory disease [424, 425]. miR-122 is detectable within extracellular vesicles and therefore has also been proposed as a biomarker not only for HCC, but also non-alcoholic fatty-liver disease as it can be detected in the blood, and for liver injury with miR-122 levels in extracellular vesicles increasing upon injury [426].

Barajas *et al* [427], identified that Glucose-6-Phosphate-Dehydrogenase (G6PD) expression is altered in liver cancer patients, with upregulation linked to higher tumour grade, increased recurrence, and poor prognosis. G6PD is the rate-limiting enzyme of the pentose phosphate pathway (PPP) which is crucial for lipid and nucleotide synthesis [428-430], and is often activated in tumour malignancies. They went on to detect two miR-122 binding sites within the 3' UTR of G6PD, validating these conserved sites using luciferase reporters. They saw that overexpression of miR-122 reduced G6P levels in HCC cells and proposed that miR-122's tumour suppression activity is mediated at least in part through its ability to reduce G6PD activity and inhibiting the PPP, enabling the cancer cells to proliferate rapidly.

There have also been examples of miR-122 involvement in gluconeogenesis with glutaminase (Gls) confirmed as a direct miR-122 target and associated with high tumour grade [431]. Aberrant upregulation of the transcription factor HNF-4 α has been shown to upregulate miR-122 in conjunction with increased expression of key

gluconeogenesis enzymes PEPCK and G6Pase in both Huh7 cells and mouse livers which contributes to the development of lipid metabolism and gluconeogenesis disorders such as Type-2 diabetes [312, 432]. Likewise, miR-122 downregulation is associated with hypoxia diseases and coincides with increased expression of glycometabolism enzymes G6PC3, ALDOA and CS in brain and cardiac tissues which Zeng *et al* subsequently identified as miR-122 targets using target prediction algorithms and proposed the use of miR-122 and its targets as biomarkers for hypoxia [433].

Interestingly, Wang *et al* [379] showed that miR-122 was capable of regulating the processing of pri-miR-21 at the post-transcriptional level as discussed in Section 1.5.4. The authors showed that alongside downregulating miR-21, miR-122 increased PDCD4 activity to promote liver cell apoptosis and suppress HCC tumour growth.

1.6.3 miR-122 regulation of Hepatitis C Virus

Hepatitis C Virus (HCV) is a small positive-sense RNA virus of the Flaviviridae family that chronically infects the liver. It is the cause of hepatitis C which can progress to cirrhosis and hepatocellular carcinomas. HCV has been shown to only replicate strongly in Huh7 human hepatocellular carcinoma cells *in vitro* and utilises intracellular membranes such as the ER to form intracellular lipid membranes, “membranous webs” on which it replicates [434, 435].

By sequestering mature miR-122 using an antisense 2’O-methylated oligonucleotide, Jopling *et al* [436] saw a decrease in HCV RNA and showed that miR-122 is essential for the replication of HCV. Further work by the Sarnow group identified two conserved sites within the 5’ UTR of the HCV RNA (See Figure 1.12) that were shown to be complementary to the miR-122 seed match and miR-122 was shown to bind cooperatively to these two sites to regulate HCV replication [310].

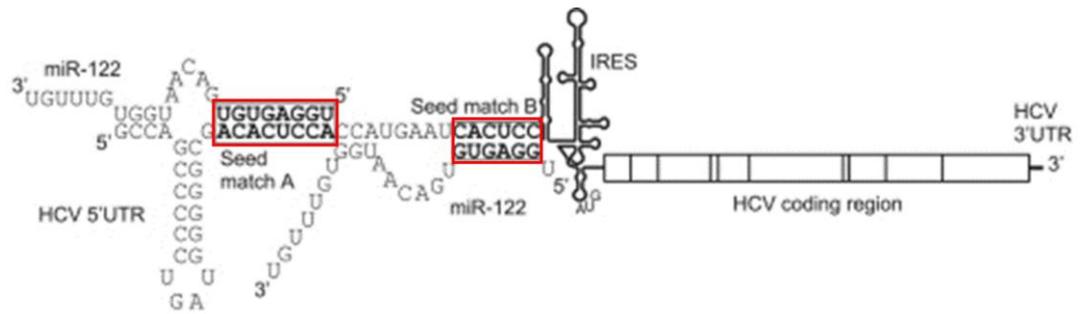


Figure 1.12: Schematic showing the binding of miR-122 to the 5' UTR of HCV. The two conserved seed match regions for miR-122 are outlined in red. Adapted from Jopling *et al* [310].

Enhancement of IRES-dependent translation of HCV by miR-122 was seen in some studies revealing that not only was miR-122 able to enhance viral replication but was responsible for a significant stimulation of translation of viral proteins [314, 437, 438]. Conversely, comparisons between HCV IRES mutants and miR-122 seed match mutants suggested that miR-122 translation stimulation was not sufficient to achieve full regulation of viral replication [309]. Masaki *et al* [439], proposed that miR-122 instead stimulates HCV RNA synthesis by displacing poly(rC)-binding protein 2 (PCBP2), thereby limiting translation. More recent work in the Jopling group shows that eIF4AII is required for efficient HCV replication. It has been shown to interact with HCV RNA directly and thought to be involved in either the retention of or recruitment to viral RNA of miR-122 [440].

Moreover, Randall *et al* [316] found that knockdowns to key proteins in the microRNA biogenesis pathway significantly reduced HCV RNA abundance and that miR-122 was dependent on Drosha, DGCR8, Dicer, and Argonaute proteins. Wilson *et al* also confirm the requirement of AGO proteins for miR-122 mediated regulation of HCV, suggesting an involvement of the RISC [441] however other research shows overexpression of miR-122 can circumvent the requirement for AGO [442]. Some reports show that miR-122 acts cooperatively with AGO to protect HCV RNA from degradation by the Xrn1 or Xrn2 5'-3' endonucleases by binding to the 5' UTR and forming a dsRNA structure

[315, 442, 443]. Indeed, structural analysis showed extensive pairing between the second seed match region of HCV RNA and miR-122 [442, 444]. However, whilst mutant miR-122 was able to provide similar protection to HCV RNA against exonucleases it was unable to efficiently upregulate HCV replication therefore it is unlikely that this mechanism accounts for the full effect of miR-122 mediated regulation of HCV replication [313, 443]. Overall, more work needs to be done in order to truly understand the dynamics between miR-122 and HCV RNA.

1.6.4 MicroRNA as therapeutics

As microRNA have a key role in gene regulation and possess the ability to target multiple mRNAs it makes them prospective candidates for therapeutics. The main issue to overcome in the development of microRNA therapeutics is delivery, with approaches including encapsulation in lipid nanoparticles or conjugation to a targeting moiety such as N-acetylgalactosamine (GalNAc), which has a high affinity for the liver-specific asialoglycoprotein receptor (ASGPR) [445].

Therapeutics that mimic a natural microRNA are typically synthetically derived from oligonucleotide duplexes but can be chemically modified to enhance stability against serum nucleases and encourage delivery to the target site. MicroRNA mimics can either be injected locally, at the site of the tumour for instance, or systemically, with various delivery systems in development.

The use of antisense oligonucleotides (ASOs) to target oncomiRs has also been investigated [446, 447]. These ASOs can be designed with various modifications to the ribose moiety that increase stability and affinity of binding to the microRNA, for example locked nucleic acids (LNA) and 2'-O-methoxyethyl modifications (2'Ome) [448, 449].

Currently, the most advanced microRNA-targeted therapy is miravirsen, a LNA oligonucleotide complementary to the 5' end of miR-122, which has reached phase II clinical trials for hepatitis C. Preclinical studies of miravirsen showed improved liver function, sufficient delivery to the liver and reduced HCV load [415] and the phase IIa trial showed dose-

dependent reductions in HCV RNA levels and successfully concluded the treatment was safe [450]. RG-101, a hepatocyte targeted N-acetylgalactosamine conjugated oligonucleotide that antagonises miR-122, was similarly effective however clinical trials were halted due to adverse effects, and microRNA-targeted approaches are less desirable with the advent of new direct acting antivirals for HCV [451].

It has been suggested that microRNA-expression profiles could present novel diagnostic biomarkers for multiple diseases and be utilised as tools for the progression of diseases and drug response [452]. Whilst the specificity for individual microRNA as biomarkers are weaker, a microRNA signature of a selection of different microRNA may aid the diagnosis of pathology and enable identification of lineage and stage of disease [171, 453, 454]. Polymorphisms in microRNA can also be used to identify disease risk, and can include mutations to the pri-, pre-, and mature-microRNA sequences, microRNA biogenesis machinery, and microRNA *cis*-regulatory elements [455, 456].

Technology such as microarray profiling and next-generation sequencing enabled initial screening of circulating microRNA to generate microRNA-biomarker profiles and microRNA biomarkers have since been identified in samples from blood, plasma, serum, urine, and sputum [457-462]. Extracellular vesicles, also known as exosomes, also provide a source of extracellular microRNAs with exosomal microRNAs detected in blood and media in several cell lines and identified as potential diagnostic biomarkers for cancers and a range of other diseases [463-467].

Despite the potential for microRNA as biomarkers, most studies have not resulted in specific and reproducible profiles. This is likely due to variation in patients (age, sex, prior treatment, confounding illness) or variation resulting from method of detection which need to be addressed in order to improve the prospects of microRNA as clinical biomarkers.

1.7 Aims and Objectives

Whilst the majority of mature microRNA and their targets are concentrated in the cytoplasm, a growing number of microRNA are being identified in a variety of different subcellular compartments. Of interest are microRNA localised at the ER, as they are likely to have a direct role in silencing transcripts encoding secreted or membrane-localised proteins, and in the nucleus where they may have a range of functions including regulation of microRNA biogenesis and regulation of nascent transcripts at the chromatin. The overall aim of my PhD is to look into the subcellular localisation and function of miR-122 to help elucidate its roles in human liver cells. The key questions I aim to address are:

Where is miR-122 located in the cell?

- Isolate microRNA from the ER and nuclear fractions using Membrane Fractionation and Chromatin-associated Fractionation methods.

How does microRNA regulation at the ER compare to in the cytoplasm?

- Look at the effect of miR-122 inhibitors in Huh7 cells on expression of miR-122 target mRNAs in these fractions.
- Establish how microRNA regulation occurs at the ER versus cytoplasm by generating ER-translated luciferase reporters for miR-122 regulation via 3' or 5' UTR sites and comparing their regulation by miR-122 with that of equivalent reporters translated in the cytoplasm.

Does miR-122 function in the nucleus to regulate its own biogenesis?

- Investigate whether miR-122 is able to self-regulate its biogenesis through the potential seed match downstream of the miR-122 encoding gene using CRISPR/Cas9 genome modification.

Chapter Two: Materials and Methods

2.1 Materials

2.1.1 General Cell Culture Reagents

Table 2.1: Suppliers and Composition of Cell Culture Reagents

| | |
|--|--|
| Dulbecco's modified eagle medium (DMEM), L-glutamine, Foetal bovine serum (FBS) | Sigma-Aldrich |
| Non-essential amino acids (NEAA) | Invitrogen |
| Optimem | Gibco |
| Lipofectamine iMax/2000 | Invitrogen |
| Trypsin-EDTA (Trypsin) | Invitrogen |
| Phosphate buffered saline (PBS) | 4.3 mM Na ₂ HPO ₄ , 1.5 mM KH ₂ PO ₄ , 137 mM NaCl, 2.7 mM KCl, pH 7.4 |

2.1.2 Plasmids

The firefly luciferase plasmids pLuc122x2, pLuc122x2M4 and p5'Luc3' has been described previously [314] and are referred in this thesis as FlucWT, FlucM4 and 5'Fluc3' respectively. The 5'Luc3' reporter contains the 5' UTR from pH77 HCV type 1a genome-length cDNA with the first 11 amino acids of the HCV core protein fused in frame to the firefly luciferase start codon, FlucWT contains two copies of the miR-122 binding region from the HCV 5' UTR inserted in its 3' UTR and is repressed by miR-122, and FlucM4 has these sites mutated (p3-4 TC to AG). The plasmid pBi-Gluc-H77C(1a)/JFH was a gift from Charles Rice [468] and contains the 5' UTR from pH77 and has Gaussia Luciferase coding sequence with signal sequence fused in frame. The plasmid pGL3con (Promega Cat E1741) was used as a luciferase reporter control and contains the firefly luciferase coding sequence downstream of an SV-40 promoter. All NanoLuc reporters were generated from pNL2.3 (Promega Cat N1021) which contains the NanoLuc® luciferase coding region fused to an N-terminal secretion signal. Luciferase reporter constructs are detailed in Table 2.2.

Table 2.2: Luciferase reporter constructs

| Reporter | Luciferase | Backbone | Site of Translation | Source |
|------------------|-------------------|--|----------------------------|---------------|
| FlucWT | Firefly | 4x wt miR-122 sites in 3'UTR | Cytoplasm | CLJ Lab |
| FlucM4 | Firefly | 4x mutant miR-122 sites in 3'UTR | Cytoplasm | CLJ Lab |
| 5'Luc3' | Firefly | 5' and 3' UTR from pH77 HCV type 1a genome-length cDNA | Cytoplasm | CLJ Lab |
| GlucM4 | Gaussia | 4x mutant miR-122 sites in 3'UTR | Endoplasmic Reticulum | Rice Lab |
| Nluc | NanoLuc | 4x wt miR-122 sites in 3'UTR | Cytoplasm | PW |
| NlucM4 | NanoLuc | 4x mutant miR-122 sites in 3'UTR | Cytoplasm | PW |
| 5'Nluc3' | NanoLuc | 5' and 3' UTR from pH77 HCV type 1a genome-length cDNA | Cytoplasm | PW |
| NlucSec | NanoLuc | 4x wt miR-122 sites in 3'UTR | Endoplasmic Reticulum | PW |
| NlucSecM4 | NanoLuc | 4x mutant miR-122 sites in 3'UTR | Endoplasmic Reticulum | PW |
| 5'Nluc3' | NanoLuc | 5' and 3' UTR from pH77 HCV type 1a genome-length cDNA | Endoplasmic Reticulum | PW |

2.1.3 PCR Primers

Table 2.3: PCR primers

| Primer | Sequence | Experiment | T _m (°C) |
|-----------------------|----------------------------------|--------------------|------------------------|
| Cat1_qF | GTCTGTCTGTTTCGCGATCCT | qRT-PCR | 64.5 |
| Cat1_qR | CAAGGAAGGTTTCAGAATCCAA | qRT-PCR | 64 |
| G6P_qF | TGGCTCAACCTCATCTTCAA | qRT-PCR | 63.4 |
| G6P_qR | CACAAGAAGAGGGGAACTGG | qRT-PCR | 63.6 |
| AldoA_qF | GTCATCCTCTTCCATGAGAC | qRT-PCR | 67 |
| AldoA_qR | TGGTAGTCTCGCCATTTGTC | qRT-PCR | 67 |
| Gluc_qF | GCACGCCCAAGATGAAGAAG | qRT-PCR | 67 |
| Gluc_qR | GTGCGATGAACTGCTCCATG | qRT-PCR | 66.9 |
| Fluc_qF | TCGCCAGTCAAGTAACAAC | qRT-PCR | 63.4 |
| Fluc_qR | ACTTCGTCCACAAACACAA | qRT-PCR | 63.9 |
| Nluc_qF | CCGTATGAAGGTCTGAGCGG | qRT-PCR/sequencing | 66.9 |
| Nluc_qR | TCTTTTTGCCGTGGAACACG | qRT-PCR/sequencing | 68.2 |
| R_SpeI | GTTGTGGTCTGGACTAGTCTTAG | Cloning | 71.5 |
| F_StuI | CAACCGTCGAGGCCCTGCGTATGGGAGTC | Cloning | 80.8 |
| F_NcoI | GTCGCCCATGGCGTATGGGAGTCAAA | Cloning | 78.5 |
| Nluc_F | TGCCTTCCCTGCCATGGTCTTCACACTCGAAG | Cloning | 82.3 |
| NlucSec_F | TAAAGCCACCATGGACTCCTTCTCCAC | Cloning | 73 |
| Nluc_R | GCCCCGACTAGTGAGTCGCGGCCTACG | Cloning | 81 |
| NcoIMut_F | GCCACGGCGTTAGTATGAGTGTCG | Cloning | 72.1 |
| NcoIMut_R | TAGACGCTTTCTGCGTGAAGACA | Cloning | 68.1 |
| sgDNA_F_Top | CACCGAGGCTGTAGTGAGCTGTG | Cas9 guides | 71.7 |
| sgDNA_F_Bottom | AAACTCACAGCTCACTACAGCCT | Cas9 guides | 64.8 |
| sgDNA_R_Top | CACCGTGTGACAGAGCAAGATCC | Cas9 guides | 70.6 |
| sgDNA_R_Bottom | AAACAGGATCTTGCTCTGTCAACA | Cas9 guides | 65.5 |
| Genotyping_qF | TACTTGAGAGGCTGAGGTGG | Genotyping | 62 |
| Genotyping_qR | AAGAGAAGGACCCACACACC | Genotyping | 63 |

2.1.4 2'O-methylated and RNA oligonucleotides

Table 2.4: Sequence of pre-microRNA mimics and antisense oligonucleotides.

*m indicates 2'O-methylated RNA nucleotide.

| Oligonucleotide | Sequence | Catalogue No |
|--|---|--------------|
| pre-miR-122 | UGGAGUGUGACAAUGGUGUUUGUGUCUAAACUAUCAACGCCAUUAUCACACUAAAUA | |
| pre-p3+4 | UGCUGUGUGACAAUGGUGUUUGUGUCUAAACUAUCAACGCCAUUAUCACACAAAUA | |
| 122-2'OMeB | mAmCmAmAmAmCmAmCmAmUmUmGmUmCmAmCmAmCmUmCmC | |
| Rand-2'OMeB | mGmUmGmUmAmAmCmAmCmGmUmCmUmAmUmAmCmGmCmCmCmAmA | |
| hsa-miR-122-3p miRCURY LNA miRNA Inhibitor | | 339121 |
| Neg Control A miRCURY Inhibitor Control | | 339126 |
| hsa-miR-122-3p miRCURY LNA miRNA Mimic | | 339173 |
| Negative Control miRCURY LNA miRNA Mimic | | 339173 |

2.1.5 siRNA sequences

Table 2.5: siRNA sequences

| siRNA | Sequence | |
|--------------|-----------------------|------------------|
| Control si A | proprietary | Ambion ID AM9611 |
| eIF4AI si A | GGAGAGUGUUUGAUUGUU.TT | Ambion ID s4572 |
| eIF4All si C | proprietary | Ambion ID s4570 |

2.1.6 Antibodies and Antisense Probes

Table 2.6: List of Antibodies used in Western Blotting

| Antibody | Species | Manufacturer | Cat Number | Dilution |
|------------------------|-------------------|---------------------|-------------------|-----------------|
| AldoA | Rabbit Polyclonal | Proteintech | 11217-1-AP | 1:5000 |
| Cat1 | Rabbit Polyclonal | Proteintech | 14195-1-AP | 1:1000 |
| G6P | Rabbit Polyclonal | Abcam | Ab83690 | 1:1000 |
| Tubulin | Mouse Monoclonal | Proteintech | 66240-1-Ig | 1:2500 |
| Histone H2B | Rabbit Polyclonal | Abcam | Ab1790 | 1:500 |
| Anti-Mouse-HRP | Goat Polyclonal | Dako | P0447 | 1:1000 |
| Anti-Rabbit-HRP | Goat Polyclonal | Sigma | A6154 | 1:20000 |

Table 2.7: Sequences of oligonucleotide probes for microRNA northern blot DNA

| Probe | Sequence |
|---------------|-----------------------|
| 122 as | ACAAACACCATTGTCACTCCA |
| U6 as | ATATGGAACGCTTCACGAATT |

2.2 Molecular Cloning

2.2.1 Designing guides for Cas9n

Forward and reverse guide oligonucleotides (gDNA) pairs were designed using ChopChop [469] to introduce double-strand breaks (DSB) at either side of the potential miR-122 seed match.

2.2.2 Phosphorylation and Annealing gDNA Oligonucleotides

Phosphorylation of 100 μ M sgDNA oligonucleotides was achieved using 2 μ L of 10 x T4 DNA ligase buffer (Thermo Fisher), 1 μ L T4 Polynucleotide Kinase (NEB) in a total reaction volume of 22 μ L, incubated at 37°C for one hour, then heat-inactivated at 65°C for 20 minutes.

Once phosphorylated, 5 μ L of each of the corresponding top and bottom oligonucleotides were added to 40 μ L of diH₂O and annealed to each other in a thermocycler (95°C 3 mins; reduce 1°C per min until reaches 26°C; 25°C 5 mins).

2.2.3 Polymerase Chain Reaction (PCR)

Amplification of plasmid DNA was achieved by PCR performed in an Applied Biosystems 2720 thermal cycler with Phusion High-Fidelity DNA polymerase (NEB). Typically, a 50 μ L reaction consisted of 20 ng plasmid DNA as a template, 200 μ M dNTPs, 2 μ M Forward and Reverse Primers, and 1x Phusion High Fidelity Buffer (NEB) and 20 U DNA polymerase. Primer sequence and T_m can be found in Table 2.3.

The PCR was then set up as follows:

| Step | Temperature | Time | Number of Cycles |
|----------------------|--------------------|---------------|-------------------------|
| Initial Denaturation | 98°C | 30 seconds | 1 |
| Denaturation | 98°C | 10 seconds | 35 cycles |
| Annealing | 50-72°C* | 30 seconds | |
| Extension | 72°C | 30 seconds/kb | |
| Final Extension | 72°C | 10 mins | 1 |

*Annealing temperature dependent on optimal temperature for primers used.

2.2.4 Restriction Digest and Dephosphorylation

1 µg of plasmid DNA or total PCR reaction were digested with restriction enzymes (SpeI, StuI, NcoI-HF, BbsI, all NEB) in CutSmart buffer (NEB), performed according to the manufacturer's recommendations, incubating at 37°C for 2 hrs.

If required, the digested vectors were then dephosphorylated using TSAP (Promega) for 20 mins at 37°C before deactivating the enzyme at 75°C for a further 20 mins.

2.2.5 Agarose Gel Electrophoresis

PCR products and restriction digests, unless otherwise stated, were analysed via agarose gel electrophoresis. The agarose gels were made by adding 1-1.5% w/v of agarose with 1 X TBE buffer and heated until a clear liquid forms before adding 4 µL SybrSafe (Invitrogen) per 50 mL of liquid, pouring and allowing to solidify at room temperature. Either 1 kb or 100 bp DNA ladders were loaded (both supplied by NEB) and samples loaded with 6 X DNA loading dye (NEB) added to reach a 1 X concentration. The gels were run at 110 V for 1 hr. A Bio-Rad Universal Hood II UV transilluminator or Cleaver Scientific safe-VIEW-MINI2 blue light box was used to visualise the DNA bands.

If required, gel extraction of the desired band(s) was performed using Monarch Gel Extraction Kit (NEB), following the manufacturer's guidelines, and eluted in 15 μ L RNase-free water. The DNA concentration and purity of each sample was determined using a NanoDrop 1000 Spectrophotometer.

2.2.6 Ligation

For the luciferase reporters, vector backbone and luciferase insert were ligated using 2 μ l of DNA T4 ligase from either NEB (40 U/ μ l) or Promega (3 U/ μ l) as specified in the Results Section, according to the manufacturer's guidelines, and incubated at room temperature for 1 hr or overnight, cycling between 4°C and 16°C every hour for a total of 10 hours.

For generation of CrisprCas9n guides, 100 ng of dephosphorylated digested vector was ligated to the annealed oligonucleotides at a molar ratio of 3:1 (vector:insert) using T4 DNA ligase (Promega) and incubated overnight at 4°C.

2.2.7 *In Vitro* Transcription

For the 5' UTR reporter plasmids, 1 μ g of plasmid was linearised by EcoRI (downstream of the 3' UTR, retaining the 5' UTR uncapped RNA structure) and purified by phenol/chloroform extraction as described in Sections 2.2.4 and 2.6.1, respectively. RNA for transfections was generated from the linearised template by *in vitro* transcription and was carried out using the T7 Megascript kit (Ambion) according to the manufacturer's instructions, then purified by LiCl precipitation overnight at -20°C. The reaction was centrifuged at 16000 g for 5 minutes and the pellet washed with 75% ethanol before resuspending in RNase free water and stored in aliquots of 0.1 μ g/ μ l at -80°C.

2.2.8 Genotyping

Genomic DNA was extracted using GenElute Mammalian Genomic DNA Isolation Kit (Sigma-Aldrich) and colonies genotyped by PCR, using primers in Table 2.3, with products run on an agarose gel (Section 2.2.5). Select

bands were extracted with Monarch Gel extraction kit (NEB), eluting in 15 μ L RNase-free water and sequenced using the genotyping primers following the requirements specified in Section 2.2.9.

2.2.9 Sequencing

Source Bioscience Ltd provided Sanger Sequencing, requiring at least 100 ng/ μ L of purified plasmid DNA and using a stock pCep_R primer (supplied by Source Bioscience) or 3.2 pmol/ μ L of primers if otherwise stated in the Results Section.

2.3 Bacterial Preparation and Culture

2.3.1 Composition of Reagents and Buffers

Table 2.8: Composition of Reagents/Buffers for Bacterial Transformation

| | |
|----------------------------------|--|
| Luria-Bertani (LB) medium | 1% w/v tryptone, 0.5% w/v yeast extract and 1% w/v NaCl in diH ₂ O, pH 7.0. For solid medium, 2% w/v bacteriological agar was added. Medium was autoclaved and stored at room temperature until required. |
| SOC Media | 2% tryptone, 0.5% yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl ₂ , 10 mM MgSO ₄ , and 20 mM glucose. Medium was autoclaved and stored at room temperature until required. |
| Ampicillin | 1000 X stock solution made as 100 mg/mL in sterile diH ₂ O and stored at -20°C. |
| Selective LB Plates | The selective antibiotic was added to melted LB agar medium to make a final concentration of 1X and poured onto petri dishes in a laminar flow hood. Plates were stored at 4°C for up to 1 month. |

2.3.2 Preparation of heat-competent *Escherichia coli*

The DH5α *Escherichia coli* (*E. coli*) strain was originally purchased from Stratagene and has the following genotype: fhuA2 Δ (argF-lacZ) U169 phoA glnV44 Φ80 Δ (lacZ) M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17. DH5α cells were streaked onto a non-selective LB agar plate and incubated at 37°C overnight. A single colony was taken and used to inoculate 10 mL of LB media, grown overnight at 37°C, before this was used to inoculate a 300 mL volume culture of LB media. When the optical density at 600 nm (OD₆₀₀) reached 0.5, the culture was centrifuged in separate 50 mL tubes at 1500 g, 10 minutes at 4°C, removing the supernatant from each. Each pellet was then resuspended in 8.3 mL of 0.1 M MgCl₂, repeating the centrifuging. Each pellet was subsequently resuspended in 8.3 mL of 0.1 M CaCl₂ and incubated on ice for 20 minutes before centrifuging again. The pellets were resuspended together in a total volume of 8.6 mL of 0.1 M CaCl₂ and 1.4 mL of 100% v/v glycerol. The suspension was inverted five times, and

transferred into microcentrifuge tubes as 200 μ L aliquots which were immediately snap frozen in liquid nitrogen and stored at -80°C .

When commercial competent DH5 α cells were used for cloning, these were purchased from NEB (cat no C2987H) and have the phenotype *fhuA2* Δ (*argF-lacZ*) U169 *phoA* *glnV44* Φ 80 Δ (*lacZ*)M15 *gyrA96* *recA1* *relA1* *endA1* *thi-1* *hsdR17*.

2.3.3 Bacterial Transformations

50 ng of intact vector/plasmid, or 5 μ L of ligation products was added to 50 μ L of competent DH5 α *E.Coli* cells and left on ice for ten minutes. The cells were then heat shocked in a water bath at 42°C for one minute and incubated on ice for an additional 5 minutes. 150 μ L of SOC media was added to the mixture and then left to recover for one hour shaking at 37°C . The entire volume was plated onto selective LB-Ampicillin plates and incubated at 37°C overnight.

2.3.4 Plasmid Purification

Single bacterial colonies were used to inoculate 10 mL LB-Amp culture, and plasmid DNA isolated using the Wizard Plus SV Miniprep kit (Promega), eluting in 50 μ L diH₂O. For large-scale purification, Qiagen Maxiprep gravity flow kit was used to isolate the plasmid DNA, following the recommended guidelines, and eluting in 200 μ L of diH₂O. The concentration and purity of plasmid DNA was determined using a Nanodrop 1000 spectrophotometer.

2.4 Cell Culture and Transfections

2.4.1 Cell maintenance

Huh7 cells (hepatoma-derived immortal cell line) were a gift from Stan Lemon (University of North Carolina) and were maintained in Dulbecco's Modified Eagle Medium supplemented with 10% v/v foetal bovine serum (FBS) with an addition of 1% v/v non-essential amino acids and 1% v/v L-Glutamine. Huh7 cells were grown at 37°C with 5% CO₂. Cells were passaged when at 90% confluency using Trypsin-EDTA using a seeding density of 1.4×10^6 cells per T75 flask.

HeLa cells were maintained in DMEM supplemented with 10% v/v FBS and 1% v/v L-Glutamine and grown at 37°C with 5% CO₂. Cells were passaged when at 80 % confluency using Trypsin-EDTA and seeded at a density of 8.4×10^5 cells per T75 flask.

2.4.2 Cell Thawing

Cells were immediately placed in a 37 °C water bath to thaw following removal from liquid nitrogen. The aliquots were then added to cell-line appropriate pre-warmed media, and cells pelleted by centrifugation at 200 g for 5 mins at room temperature. The supernatant was discarded and cells were resuspended in media and cultured as outlined above.

2.4.3 Cryopreservation

Cell lines were cryopreserved in liquid nitrogen. Following trypsinisation, 8.8×10^6 cells were pelleted by centrifugation (200 g, 4°C, 5 mins) and the cell pellet was resuspended in 1 mL of FBS with 10% v/v DMSO. For maximum recovery of cells, vials for cryopreservation were stored at -80°C in an isopropanol freezing container, with a cooling rate of approximately -1°C /minute, before being transferred to liquid nitrogen for long term storage.

2.4.4 microRNA inhibitor Transfections

2'-O-methylated oligonucleotides (Table 2.4) were synthesised by Dharmacon Inc. and the LNA inhibitors (hsa-miR-122 LNA, Cat no: 339121; Neg Ctrl A LNA, Cat no: 339125) were ordered from Qiagen and were complementary to miR-122 or a negative control. For RNA and protein isolation from total cell lysates or for fractionation experiments, cells were grown to 60- 70% confluency on a 10 cm plate and transfected with 20 nM of 2'-O-methylated oligonucleotides or 10 nM LNA inhibitor using Lipofectamine RNAiMax and OptiMem following the manufacturers guidelines. In both cases, a randomised control transfection was carried out in parallel. The cells were then cultured at 37°C before a media change at 6 hrs and then harvested 24 and/or 48 hrs after initial transfection.

2.4.5 microRNA mimic/overexpression Transfections

The wild-type pre-miR-122 and mutant pre-miR-122p3+4 oligos were synthesised by Dharmacon Inc. and the LNA miRNA mimics were ordered from Qiagen (hsa-miR-122-5p miRCURY LNA miRNA, Mimic Cat no: 339173; Negative Control miRCURY LNA miRNA Mimic, Cat no: 339173) and were designed to simulate miR-122 or a negative control. Cells were grown to 60- 70% confluency on a 10 cm plate or 6-well plate and transfected with 20 nM wildtype pre-miR-122/mutant pre-miR-122p3+4 or 2 nM LNA mimic/control using Lipofectamine RNAiMax and OptiMem following the manufacturers guidelines. The cells were then cultured at 37°C before a media change at 6 hrs and then harvested 24 and/or 48 hrs after initial transfection.

2.4.6 Luciferase Reporter Transfections

Huh7 cells were grown to 60-80% confluency in 24-well plates and transfected with lipofectamine 2000 in triplicate, in a method adapted from the manufacturer's guidelines. Different concentrations of plasmid/RNA/mimic/inhibitor were tested, and cells harvested at different timepoints, and these are discussed in detail in the Results Section.

A master mix in Optimem was used to introduce Lipofectamine 2000 to all wells and incubated for 5 minutes. A second set of master mixes was used starting at the smallest concentration of plasmid and then split up and additional plasmid added to second set to generate master mixes for each plasmid concentration, as outlined in Figure 2.1. The Lipofectamine-containing master mix was then added to each plasmid master mix and incubated for 20 minutes. The media on the 24-well plates was removed and replaced with 200 μ l of fresh media before the transfection mix was added directly per well. The cells were then cultured at 37°C and harvested at several time points. Sequences for the oligonucleotides can be found in Table 2.4.

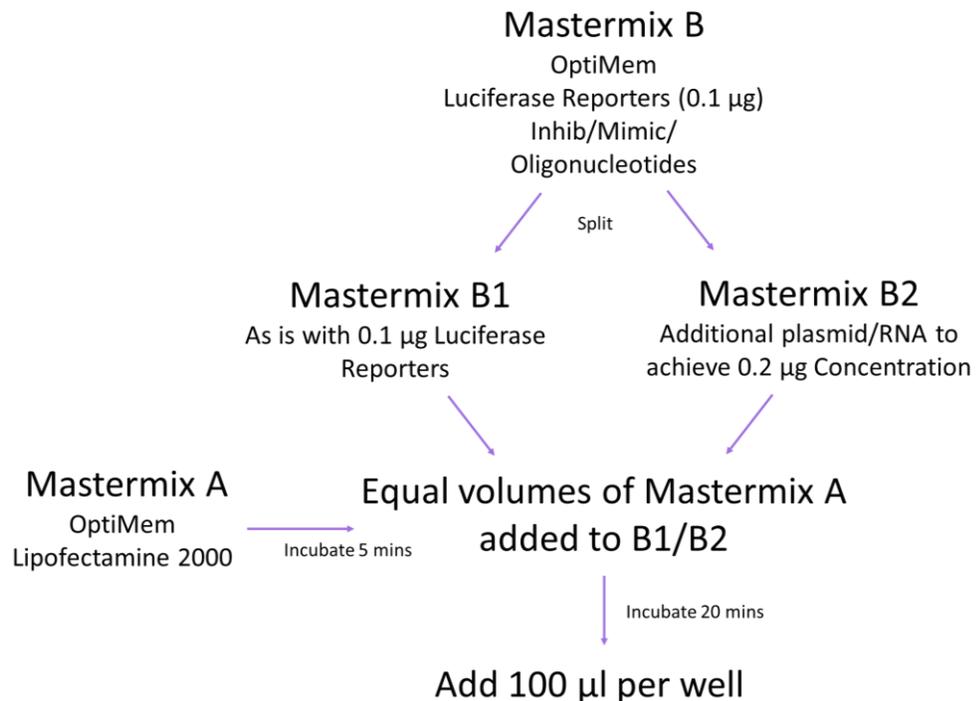


Figure 2.1: Luciferase Transfections. Diagram outlining an example use of mastermixes, including incubation steps, in the co-transfection of luciferase reporters and miR-122 Inhibitors/Mimics or oligonucleotides.

For the Nanoluc reporter assays, each triplicate contained 300 μ l Optimem, 1.2 μ l Lipofectamine2000 and either 10 nM LNA Inhibitor, 2 nM LNA mimic

(both Qiagen, details in Section 2.1.4) or 10 nM pre-miR-122/pre-miR-122p3+4 oligonucleotides (Dharmacon Inc). The cells were co-transfected in triplicate with either 100/200 ng 5'luc3' RNA (as generated in Section 2.2.7) or 50/100/200 ng luciferase reporter plasmids. For the Firefly and Gaussia reporter assays, each triplicate contained 150 µl Optimem, 0.6 µl Lipofectamine2000, 10 nM oligonucleotide, and 50 ng luciferase reporter plasmid.

Ambion silencer select siRNAs were used to knock down eIF4AII (s4572, and s4570) in parallel with Control si1 (AM9611). All were delivered into cells grown to 60-70% confluency in 6 cm dishes at 10 nM final concentration using Lipofectamine RNAiMax. Cells were cultured at 37°C for 6 hrs before splitting into 24-well dishes for subsequent oligonucleotide/reporter transfections at 24 hrs post siRNA transfection.

2.4.7 Crispr/Cas9 Nickase (Crispr/Cas9n) Transfection

Either 4 µg of single or 2 µg of both CrisprCas9 Nickase plasmids was transfected into Huh7 cells (grown to 70% confluency on a 10 cm plate) with 1.5 µL Lipofectamine 2000, following the manufacturer's guidelines.

2.4.8 Fluorescence activated cell sorting (FACS)

Transfected cells were trypsinised and centrifuged at 200 g for 5 min at 4°C. The pellet was resuspended in normal growth media and transferred to a FACS tube. The fluorescence of cell suspension was quantified using the MoFlo XDP (Beckman Coulter) flow cytometer with excitation at 536 nm and emission measured at 617 nm wavelengths. Using un-transfected cells as the background control, single GFP^{+ve} cells were seeded into 96-well plates in 100 µL normal growth media, supplemented with an additional 100 µL before incubating at 37°C. Or when single cell seeding was performed by hand, FACS was used to isolate a bulk population of GFP^{+ve} cells which were allowed to recover in a T75 flask overnight before being seeded into a 96-well plate at 1 cell per well in 200 µL media.

2.5 Luciferase assays

Gaussia reporters were analysed with GAR-2B Assay Kit (Targeting Systems) following the manufacturer's protocol at half volumes; 10 μL of media was transferred to a 96-well optical plate and 25 μL of assay reagent added. For Firefly reporters, cells were washed in 1 x PBS and incubated in 50 μL of 1x Passive Lysis Buffer (Promega, Cat No: E1941) for 5 minutes at room temperature before scraping with a pipette tip. The lysates were then transferred to a 96-well optical plate and analysed using the Single Luciferase Assay Reporter Kit (Promega), following the manufacturer's protocol at half volumes; 25 μL of assay reagent added to 10 μL of lysate.

When assayed without the firefly control, both Nanoluc reporters are assayed using Nano-Glo[®] Luciferase Assay System (Promega). For the secreted reporter, 10 μL of media was diluted in 40 μL of diH₂O before adding to 50 μL luciferase assay buffer in a 96-well plate and incubated for at least 3 minutes. For the lytic method, the media was removed, washed in 1x PBS, and lysed in 50 μL of 1 x Passive Lysis Buffer (Promega). After 5 minutes, the cells were scraped and 10 μL of lysate was added to 10 μL of luciferase assay buffer in a 96-well plate and incubated for at least 3 minutes.

When co-transfected with a firefly control reporter, both Nanoluc reporters are assayed using the Nano-Glo Dual Reporter Assay System (Promega, Cat No: N1620). For the secreted reporter, 20 μL of media was transferred to a 96-well plate before adding 40 μL Stop&Glo Reagent, incubating for at least three minutes. To measure the activity of the firefly control of both the secreted and cytoplasmic reporters, the media was removed, washed in 1x PBS, and lysed in 50 μL of 1 x Passive Lysis Buffer (Promega). After 5 minutes, the cells were scraped and 20 μL of lysate was added to 40 μL of One-Glo EX Reagent in a 96-well plate and incubated for at least 3 minutes. Subsequently for the cytoplasmic reporters, 40 μL of the Stop&Glo Reagent

was added directly to each well and activity measured after a three minute incubation.

All assay reagents were at room temperature at the time of the assay. Luciferase activity was measured using a GloMax luminometer (Promega) in triplicate and all samples were taken as raw fluorescence intensity values. For all reporters the triplicates were then averaged, except where the Dual Reporter Assay system was used where each Nanoluc fluorescence value was normalised to their respective Firefly control prior to calculating an average. The relative difference in activity was then determined by normalising to the average of the transfection control samples.

2.6 RNA Extraction, Reverse Transcription (RT) and Real-time qPCR (qRT-PCR)

2.6.1 RNA Isolation

When RNA was isolated directly from 10 cm dishes, 2 mL Tri Reagent (Sigma) was added to the plate, scraped with a cell lifter (ThermoFisher) and collected, splitting over two microcentrifuge tubes. From 6 cm dishes or 6-well plates, 1 mL of Tri reagent was used and cells collected into a single tube per sample. To extract RNA following IP, 1 mL of Tri reagent was added to all samples regardless of sample volume.

200 μ L chloroform (Sigma) was added to each tube and vigorously shaken for >20 seconds. The samples were incubated at room temperature for 2 minutes before centrifugation at 16000 g for 15 mins at 4°C. The aqueous phase was carefully taken into a new microcentrifuge tube, where necessary combining the duplicates for each sample. 500 μ L of isopropanol (Sigma) and 0.5 μ L Glycoblue (ThermoFisher) was added to each sample and left to precipitate at -80°C for at least 30 mins or overnight. The RNA pellets were then precipitated by centrifugation for 10 mins at 16000 g at 4°C. The supernatant was removed and the pellet washed in 500 μ L 75% ethanol. The tubes were then centrifuged again (16000 g, 4°C, 5 mins) and the RNA pellet resuspended in 20 μ L RNase-free water.

Alternatively, RNA was extracted using the Promega ReliaPrep RNA cell Miniprep system. For samples obtained from membrane fractionation, lysates were precipitated in equal volumes of isopropanol and centrifuged at 16000 g, 4°C for 30 mins. For total lysates of 10 cm dishes, cells were washed twice with PBS and then scraped in 1 mL of PBS on ice and centrifuged at 16000 g for 5 minutes at 4°C. For both fractionated and total cell RNA, the cell pellets were then resuspended in 250 μ L BL + TG buffer from the ReliaPrep RNA cell Miniprep system (Promega), continuing to follow the manufacturer's protocol. Adaptations to the manufacturer's

guidelines included incubating in DNaseI for 30 mins and eluting in 15 μ L RNase-free water.

RNA quantity and quality were determined by UV spectrophotometry using Nanodrop (Thermo Fisher). RNA was stored at -20°C and downstream processing was typically performed within 3 days.

2.6.2 Mature microRNA qRT-PCR

For mature microRNA, cDNA was generated from 100 ng of total RNA using TaqMan MicroRNA Reverse Transcription kit (Applied Biosystems) scaled down to half- reactions except when using serial dilutions to form a standard curve where full volumes were used. For fractionations, 0.5 μ L cel-mir-39 (Norgen) was added to the RT reactions as a control. qRT-PCR was performed in a Qiagen Rotor-Gene Q machine using microRNA Taqman assay kits (Applied Biosystems), scaled down to half-reactions. The RT hairpin linker is designed with a 3' stem-loop structure to eliminate detection of pri-microRNA and therefore increase the specificity for mature microRNA, see Figure 2.2. Both RT primers and qRT-PCR probes were obtained from Applied Biosystems: hsa-miR-122 (Cat 4427975), U6 snRNA (Cat 4427975), cel-miR-39 (Cat 4427975). For each sample, qRT-PCR was performed in triplicate and the variation in expression levels of the microRNA of interest were presented as $2^{-\Delta\text{Ct}}$ for un-normalised fractionation data or corrected by normalising to cel-miR-39 where data were analysed by the $2^{-\Delta\Delta\text{Ct}}$ method.

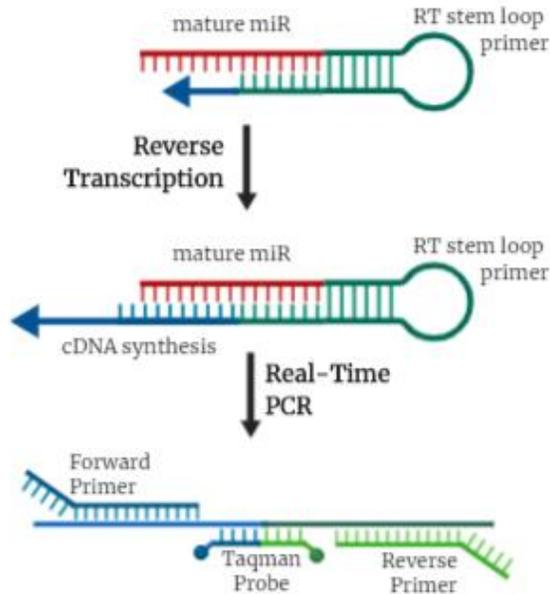


Figure 2.2: Schematic of Taqman RT/PCR primers. Showing the RT primer (dark green) binding to the mature microRNA, and the forward (blue) and reverse (green) primers and taqman probe binding to the cDNA product from the reverse transcription.

2.6.3 mRNA qRT-PCR

For mRNA, GoScript with random primers (Promega) was used to generate cDNA from 100 ng total RNA, according to the manufacturer's guidelines. For fractionation experiments, 0.5 ng 5'Luc3' RNA was added to the RT reactions as a control. Real-time quantitative PCR (qRT-PCR) was performed using GoTaq Green Master Mix (Promega), with primers that were obtained from Sigma-Aldrich (Table 2.3) and carried out using a Qiagen Rotor Gene Q machine. Data were analysed by the $2^{-\Delta\Delta C_t}$ method relative to reference RNA for total RNA experiments or synthetic 5'Luc3' RNA control for fractionation experiments. For immunoprecipitation experiments, data were analysed as $2^{-\Delta C_t}$ relative to the input RNA.

2.7 Subcellular Fractionation

2.7.1 Chromatin-Associated RNA Fractionation

Chromatin and nucleoplasmic fractions were isolated from cytoplasmic fractions following a method adapted from the Proudfoot lab [470]. Huh7 cells were grown to 80% confluency in 10 cm plates and washed twice in ice-cold PBS. 4 mL ice-cold HLP1 (10 mM Tris-HCl pH 7.5, 10 mM NaCl, 0.5% v/v NP-40, 2.5 mM MgCl₂) was added and then incubated on ice for 5 minutes. The cells were scraped into a 50 mL tube and the lysate underlaid with 1 mL HLP2 (10 mM Tris-HCl pH 7.5, 10 mM NaCl, 0.5% v/v NP-40, 2.5 mM MgCl₂ and 10% v/v glycerol). The cells were then centrifuged for 5 mins (150 g, 4°C) and the cytoplasmic fraction was retained. The nuclei pellet was suspended in 125 µL NUN1 (20 mM Tris-HCl pH 8.0, 75 mM NaCl, 0.5 mM RNase-free EDTA, 50% v/v glycerol, supplemented immediately before use with 1X Protease Inhibitor (cOmplete Mini Inhibitor Cocktail, Sigma) and 200 units RNase inhibitors (RNasin® Ribonuclease Inhibitor, Promega) and transferred to a 1.5 mL microcentrifuge tube and incubated on ice for 15 minutes with 1.2 mL of NUN2 (20 mM HEPES, 7.5 mM NaCl, 0.2 mM RNase-free EDTA, 1 M urea, 1% v/v NP-40, supplemented with protease inhibitor). This was vortexed regularly before centrifugation for 10 minutes (16000 g, 4°C). The supernatant containing the nucleoplasmic fraction was retained and the chromatin pellet resuspended in 200 µL HSB (10 mM Tris-HCl pH 7.5, 500 mM NaCl, 10 mM MgCl₂). The chromatin fraction was immediately treated with Turbo DNase (Thermo Fisher) and Proteinase K (NEB) for 10 minutes at 37°C. RNA and protein were extracted by the methods outlined in Sections 2.6.1 and 2.10.2.

2.7.2 Membrane Fractionation

Membrane-associated fractions were isolated using a method adapted from Jagannathan *et al* [471]. Huh7 cells were grown to 80% confluency in a 10 cm plate and washed twice in ice-cold PBS. The cells were incubated for 5 minutes on ice in 1 mL permeabilisation buffer (100 mM potassium acetate, 25 mM K-Hepes pH 7.2, 2.5 mM magnesium acetate, 1 mM EGTA, 0.015% w/v digitonin, 1 mM DTT, supplemented immediately before use with 1X Protease Inhibitor (cOmplete Mini Inhibitor Cocktail, Sigma) and 200 units RNase inhibitors (RNasin® Ribonuclease Inhibitor, Promega). The soluble material (cytosol fraction) was drained and transferred to microcentrifuge tubes using a P1000 pipette. The cells were then washed with 5 mL of wash buffer (100 mM potassium acetate, 25 mM K-Hepes pH 7.2, 2.5 mM magnesium acetate, 1 mM EGTA, 0.004% w/v digitonin, 1 mM DTT) before 1 mL lysis buffer (400 mM potassium acetate, 25 mM K-Hepes pH 7.2, 15 mM magnesium acetate, 1% NP-40, 0.5% w/v sodium deoxycholate, 1 mM DTT, supplemented with protease and RNase inhibitors) was added. The cells were incubated on ice for 5 minutes and the soluble material (membrane fraction) collected in microcentrifuge tubes. Both the cytosolic and membrane fractions were centrifuged for 10 minutes (7500 g, 4°C) to remove cell debris and the supernatants transferred to clean microcentrifuge tubes. RNA and protein were extracted by the methods outlined in Sections 2.6.1 and 2.10.2.

2.8 Ago-Immunoprecipitation (Ago-IP)

20 μ L of magnetic protein A/G beads (Thermo Scientific) were resuspended in 500 μ L of ER-lysis buffer from Section 2.7.2 and blocked with 1 mg/mL BSA (Alpha Diagnostics), 1 mg/mL Glycogen, and 1 mg/mL yeast tRNA (Sigma) for 30 mins at 4°C on a rotator. Beads were washed with 500 μ L ER-lysis buffer (from Section 2.7.2) before being incubated with either 2 μ g AGO1-4 antibody (mAb 2A8, Sigma) or normal mouse-IgG control antibody (mAb sc-2025, Santa Cruz) for 2 hours, rotating at 4°C. Huh7 cells were grown to 80-90% confluency in 10 cm dishes. Where used, crosslinking was achieved by incubating the 10 cm dishes with 1% v/v formaldehyde for 2 mins at room temperature before quenching with 250 nM glycine, pH7 at room temperature for 5 mins. Both crosslinked and un-crosslinked plates were washed twice with ice-cold PBS before undergoing membrane-associated fractionation as described in Section 2.7.2. Alternatively, for total cell lysates, after washing in PBS the 10 cm plates were scraped using a cell lifter to collect the cells in 1 mL of PBS and collected by centrifugation at 300g for 5 minutes at 4°C. The cell pellet was then resuspended in 500 μ L NET-2 buffer (150 mM NaCl, 0.05% v/v NP-40, 50 mM Tris-HCl, pH 7.4, supplemented with fresh protease and RNase inhibitor) prior to 3 rounds of sonication in 20 second intervals. The cells were centrifuged at 16,000g for 10 mins at 4°C to collect cell debris.

In all cases, 10% of the lysate volume used for each IP sample were saved as “Input” and the remaining lysate was incubated with 20 μ L of beads for 1 hour at 4°C on a rotator. Pre-cleared lysates were removed from the beads and 400 μ L was added to both the AGO- and IgG-conjugated beads. The samples were incubated overnight at 4°C on a rotator. Beads were then washed 4x in 1 mL of ice-cold NET-2 buffer rotating at 4°C for 5 minutes between washes. To reverse crosslinking, samples were heated to 75°C for 10 mins after the washes. RNA was isolated from the input cell lysate, the AGO1-4 immunoprecipitate, and the control IgG immunoprecipitate using Tri

reagent (Sigma) (Section 2.6.1) and analysed by RT-qPCR as described in Section 2.6.

2.9 Northern blotting

2.9.1 Buffer Composition

Table 2.9: Composition of Materials and Buffers for Northern Blotting

| Reagent | Composition |
|--------------------------------|--|
| 10 X MOPS | 0.2 M MOPS, 0.05 M NaAc, 0.01 M EDTA, pH 7.0 |
| mRNA loading dye | 95% v/v formamide, 0.04% w/v bromophenol blue |
| 2 X SSC | 0.3 M sodium chloride, 0.03 M trisodium citrate, pH 7 |
| 0.1x SSPE | 0.02 M EDTA and 2.98 M NaCl in 0.2 M phosphate buffer, pH 7.4 |
| Blotting Buffer | 0.05 M Tris-HCl, 0.192 M glycine, 20% v/v methanol |
| Denaturation Buffer | 1xMOPS, 17.5% v/v formaldehyde, 50% v/v deionised formamide |
| miRNA loading dye | 50% glycerol, 0.001 M EDTA, 0.4% bromophenol blue, 0.4% xylene cyanol |
| 20x SSC | 3 M NaCl, 0.3 M tri-sodium citrate |
| Methylene Blue Solution | 0.1% methylene blue dye, 0.5M sodium acetate |
| Church-Gilbert Solution | 0.18 M Na ₂ HPO ₄ , 0.07 M NaH ₂ PO ₄ , 7% SDS |
| EDC Solution | 0.16 M EDC, 0.13 M 1-methylimidazole, pH 8 with HCl. |

2.9.2 microRNA Northern Blotting

Northern blotting for microRNA was performed using a minimum of 10 µg of RNA, denatured at 65°C for 10 minutes in RNA loading dye. The denatured RNA was separated on a 15% bis-acrylamide gel containing 7M urea, and 1X MOPS and run in 1 X MOPS buffer at 18 W for 1.5 hours. The RNA was then transferred to a Hybond NX membrane (Amersham), using a standard semi-dry transfer method (10 V, 4°C, 1.5 hours) in blotting buffer. The membrane was then crosslinked either by heating to 60°C with EDC (Sigma)

solution for 1 hour, or UV crosslinking using a Stratalinker 2400 (Stratagene) at 1200 mJ/min.

Radioactive probes were prepared by incubating 1 μL of anti-sense oligo (100 μM), 1 μL PNK buffer (Promega), 4.5 μL RNase-free water, 1 μL T4 PNK (Promega) and 2.5 μL ^{32}P γ -ATP, for at least 30 mins at 37°C. 40 μL RNase-free water was added and then spun through a G50 Sephadex Column (GE Healthcare) to eliminate unincorporated γ -ATP. Membranes were pre-hybridised in 7 mL ULTRAhyb Hybridization Buffer (Thermo Fisher) at 37°C in a hybridisation oven with rotation, prior to incubating overnight with 7 μL of 5' ^{32}P -labeled anti-sense oligo probes, sequences can be found in Table 2.7.

The membranes were washed twice in 2 x SSC/0.5% w/v SDS for 30 minutes at 37°C, blotted dry and placed on a phosphor screen overnight, before imaging on a Storm phosphorimager (GE healthcare). Before re-probing, membranes were stripped by washing three times in 0.1 x SSPE/0.1% w/v SDS at 65°C.

2.9.3 mRNA northern blotting

A minimum of 1 μg RNA was heated at 55°C for 15 min in denaturation buffer. 2 μl mRNA loading dye was added to the samples before loading onto a 1% agarose gel containing 1x MOPS with 2.2 M formaldehyde, alongside 6 μl RNA ladder (RNA millennium, Ambion). The gel was run in 1x MOPS containing 10% formaldehyde at 100 V for 2 hrs. The gel was then soaked in fresh 0.05 M NaOH for 20 mins, rinsed in RNase free H₂O and then soaked in 20X SSC for 30 min. The RNA was transferred to a Zetaprobe membrane (Bio-rad) overnight by a capillary action, outlined in Figure 2.3. To create the transfer system, a wick was placed over a glass plate with its ends submerged in 20X SSC. The gel was then placed on top of the wick, followed by the zetaprobe membrane and three pieces of blotting paper, all pre-soaked in SSC buffer. A stack of paper towels was placed on top, followed by a glass plate and finally weights to compress the tower.

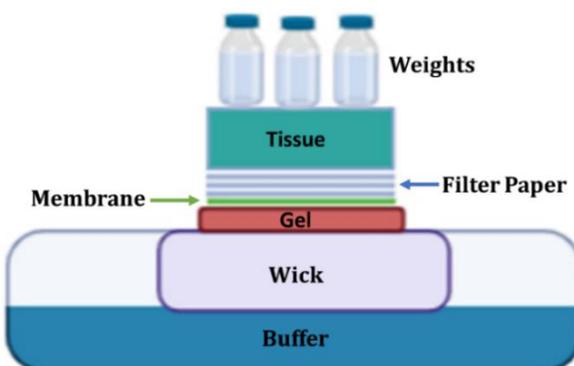


Figure 2.3: Schematic of tower transfer. Used for capillary transfer of RNA from formaldehyde gel to Zetaprobe membrane in mRNA northern blotting

The membrane was crosslinked using a Stratalinker 2400 (Stratagene) at 1200 mJ/min and stained with methylene blue solution. The membrane was rinsed with H₂O and the positions of rRNA bands and ladder marked.

Approximately 10 µg of the plasmid of interest was digested to produce a 200-600 bp fragment (Fluc and Gluc: Spel and Stul, Nluc/NlucSec: Spel and NcoI, Act: EcoRI) and run on a 1% agarose gel. The fragment was extracted using the Monarch gel extraction kit (Section 2.2.5) and eluted in 100 µL RNase-free water. 5 µL of the probe template was boiled at 95°C for 5 min in water (15 µL total volume). ³²P α-dCTP was incorporated to denatured probes by incubating at room temperature for 1 hr with 5 µL labelling buffer (Promega), 0.5 µl 20 mM dATP/dGTP/dTTP, 1 µL 400µg/mL BSA, 2.5 µL ³²P α-dCTP, 1 µL Klenow exo- (NEB). The unincorporated ³²P α-dCTP was removed using a G50 Sephadex column (Cyiva), following the manufacturer's recommendations, and then stored at -20°C and used within a week. Prior to use, the probe was boiled for 5 min at 95°C.

Membranes were pre-hybridised in either Church-Gilbert or Ultrahyb (Ambion) for 1 hr at 65°C or 42°C, respectively in a hybridisation tube in a hybridisation oven with continuous rotation, prior to incubating overnight with

³²P-labeled oligo probes at the same temperature. For Church-Gilbert: membranes washed for 10 mins each at room temperature in (i) 2X SSC/0.1% SDS (ii) 0.5X SSC/0.1% SDS (iii) 0.1X SSC/0.1% SDS. For Ultrahyb: membranes washed at 42°C twice each in (i) 2X SSC/0.1% SDS for 5 mins and (ii) 0.1X SSC/0.1% SDS for 15 mins. Membrane was then blotted dry, wrapped in Saran wrap, and placed on a phosphor screen overnight before imaging on a Storm phosphorimager (GE healthcare). To re-probe, the membrane was boiled in 200 mL H₂O with 1 g SDS and left on a rocker for 30 mins, repeated twice, before being rinsed with H₂O and pre-hybridised again.

2.10 Protein Extraction, SDS-PAGE and Western Blotting

2.10.1 Buffer Composition

| Component | Composition |
|-----------------------------------|---|
| RIPA Buffer | 50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 1.0% NP-40, 0.5% w/v sodium deoxycholate, 0.1% w/v SDS |
| SDS Loading Dye | 10% SDS, 500mM DTT, 50% Glycerol, 500mM Tris-HCL 0.05% bromophenol blue dye |
| TGS Buffer | 25 mM Tris-HCl pH 8.3, 192 mM glycine, 0.1% w/v SDS |
| Transfer Buffer | 50 mM Tris-HCl pH 8.3, 192 mM glycine, 20% v/v methanol |
| 1 X TBST | 10 mM Tris-HCl pH 8, 30 mM NaCl, 0.1% v/v Tween |
| 5 % Milk Blocking Solution | 5% w/v Marvel 0% fat milk powder in 1 X TBST |

2.10.2 Protein Extraction

Protein was extracted from cells by washing 2x ice-cold PBS and incubating in RIPA buffer for 5 mins on ice. 1.5 mL and 500 μ L of RIPA buffer was used directly on 10 cm dishes and 6-well plates, respectively, for total lysates. For fractionation experiments, cells were collected in 1 mL of PBS, where 80% of the volume was spun down and the cell pellet resuspended in 1 mL of RIPA buffer. The samples were then centrifuged at 16000 g for 5 mins at 4°C to pellet the cell debris.

Protein concentration was determined using a Bradford assay (BioRad), standardised to bovine serum albumin (BSA, Alpha Diagnostics). Protein samples were stored at -80°C.

2.10.3 SDS-PAGE

Table 2.10: Composition of SDS-PAGE gel

| Component | 4% Stacking gel | 12.5% Resolving gel |
|----------------------|-----------------|---------------------|
| diH ₂ O | 1.36 mL | 1.57 mL |
| 30% w/v Acrylamide | 340 µL | 2.08 mL |
| 1.5 M Tris pH 8.8 | 250 µL | 1.25 mL |
| 10% w/v SDS solution | 20 µL | 50 µL |
| 10% w/v APS solution | 20 µL | 50 µL |
| TEMED | 2 µL | 5 µL |

4 X SDS loading dye was added to the samples which were then boiled at 95°C for 10 minutes. SDS-PAGE gels (4% stacking, 12.5% resolving) was prepared as shown in Table 2.10 above. Samples were loaded onto the gel and run at 120 V for 90 minutes in 1 X TGS buffer.

2.10.4 Western Blotting

Following separation by electrophoresis, protein samples were transferred from SDS-PAGE gels onto a methanol-soaked PVDF membrane (Amersham) either using a semi-dry transfer machine (70 mA, 90 mins, room temperature; Trans-Blot® SD Semi-Dry Transfer Cel, BioRad) where blotting paper is soaked in transfer buffer, or a wet-transfer machine (80 mA, 90 mins, 4°C; Criterion™ Blotter, BioRad). Non-specific antibody binding was then blocked by adding a 5% milk solution to the membrane for one hour, rotating at room temperature, prior to incubation at 4°C overnight, rotating in 5 mL 5% milk containing the primary antibody at dilutions specified in 2.1.6. The membrane was then washed in TBST for five minutes, three times, at room temperature on a rocker to remove excess antibody and then incubated for one hour at room temperature in the secondary antibody

on a rocker (dilutions specified in 2.1.6). The membrane was again washed in TBST three times to remove excess antibody.

ECL Prime Western Blotting Detection Reagent (Amersham) was added to the membrane and developed following the manufacturer's instructions. The excess developing solution was blotted off and the membrane sealed in clear plastic. A Fujifilm LAS-4000 was used to detect and image the chemiluminescent signal. If stripping was required, the membrane was soaked in 1X Re-Blot Plus Strong Solution (Merck) for 10 mins at room temperature and rinsed in TBST before re-probing.

Chapter Three:

**Use of Luciferase
Reporters to
Investigate Regulation
by miR-122 at the
Cytoplasm and
Endoplasmic
Reticulum**

3.1 Introduction

It has been presumed that microRNA-mediated silencing occurs predominantly in the cytosol, however more recent studies have found evidence for miRNA and RISC components at endomembranes such as the endoplasmic reticulum (ER) [131, 344-347, 349]. In fact, mammalian AGO2 was initially identified as an ER and Golgi-associated protein before its identification as an argonaute protein [344]. Dicer, TRBP, and PACT have been found to associate at the ER in primary neurons and similarly a small subpopulation of the key RISC components TNRC6A, AGO2, and CNOT1 were identified in ER-containing fractions of HeLa cells [350]. Interestingly, Stalder *et al* found that not only does AGO2 localise to the ER and associate with Dicer and TRBP, but TRBP anchors the miRISC to the ER where it is a major site for mRNA silencing [351].

Proteins that are destined for organelles in the endomembrane system (ER, Golgi, lysosomes) or to be secreted out of the cell are targeted to the ER for translation by a signal peptide [49-51]. There are multiple ribosomes translating an individual mRNA at any one time, and as new signal peptides will emerge the mRNA will remain tethered to the ER over multiple rounds of translation [52-54]. Increasing evidence suggests that the ER is a site for general protein synthesis, with ER-bound ribosomes translating mRNAs encoding cytosolic proteins, not just those encoding membrane/secretory proteins (comprehensive review by Reid & Nicchitta, 2015 [472]). For example, in HEK293 cells 50% of all ribosomes are associated to the ER and 75% of the translational activity at the ER is directed to the synthesis of cytosolic proteins [473, 474]. As the rough ER contains high levels of both miRISCs and ribosomes, this suggests that miRNA-mediated mRNA silencing could potentially be achieved at greater efficiency at the ER than in the cytoplasm. Whilst there is a plethora of studies looking at the role of microRNA in the ER stress response, and a number looking at the extracellular export of microRNA from the ER to vesicles and the disruption

of microRNA at the ER in disease conditions, there has not been a direct comparison of endogenous miRNA regulation at the ER and the cytoplasm in human cells.

To establish how microRNA regulation occurs at the ER versus cytoplasm, ER-translated luciferase reporters for miR-122 regulation were generated and their regulation by miR-122 was compared with that of equivalent reporters translated in the cytoplasm. Luciferase reporter assays are able to determine a direct functional connection between the microRNA of interest and the luciferase reporter as the amount of light produced by the luciferase protein provides a quantitative measure of the protein synthesis from the luciferase reporter gene. Co-transfecting cells with reporter plasmids and miRNA mimics or inhibitors also allows the implementation of gain-/loss-of-function experiments. The liver-specific miR-122 was chosen to study as it has been extensively used in the Jopling lab with high expression in Huh7 hepatocellular carcinoma cells, accounting for up to 70% of their total microRNA content [413]. Huh7 cells are also of interest in that being hepatocyte cells, they are highly active in the synthesis of proteins and lipids for secretion and as such, contain high amounts of rough and smooth endoplasmic reticulum. As Huh7 cells allow the replication of HCV, the investigation of miR-122 in Huh7 cells allows the comparison of miR-122 repression of the 3' UTR to the unique mechanism of miR-122 up-regulation of translation via 5' UTR sites from Hepatitis C Virus (HCV) RNA. Furthermore, HCV uses intracellular membranes such as the ER to form intracellular lipid membranes on which it replicates, which adds interest to seeing if there are any differences in regulation between the ER and cytoplasm in the context of up-regulation of translation by miR-122.

3.2 Generation of a Secreted Luciferase Reporter

The Jopling lab had previously generated Firefly luciferase reporters that are regulated by miR-122 binding to the 3' UTR (plasmids Section 2.1.2, Table 2.2) for which plasmid maps can be found in Supplementary Figures 1-3 [314]. FlucWT contains two copies of the miR-122 binding region from the HCV 5' UTR inserted in its 3' UTR and which can be repressed by miR-122. As there are two miR-122 binding sites in the HCV 5' UTR binding region, this reporter has four miR-122 binding sites in total. FlucM4 has these sites mutated at positions 3-4 (TC to AG) and is no longer regulated by endogenous miR-122 but repressed by ectopically introduced miR-122p3+4 with compensatory mutations that restore binding to the mutant [310]. Upon transfection into Huh7 cells, the firefly luciferase is translated in the cytoplasm and the protein is expressed intracellularly.

The plasmid pBi-Gluc-H77C(1a)/JFH contains the 5' UTR from pH77 and has the Gaussia Luciferase coding sequence with signal sequence fused in frame, Supplementary Figure 4 [468]. As this luciferase contains a secretion signal upstream of the luciferase coding region, it is translated at the ER and then secreted into the media. Therefore, generating a reporter construct using Gaussia luciferase regulated by the same miR-122 binding sites as the Firefly luciferase reporters would allow a comparison between miR-122 regulation at the ER and cytoplasm. Figure 3.1 shows the site of translation of the Firefly and Gaussia reporters.

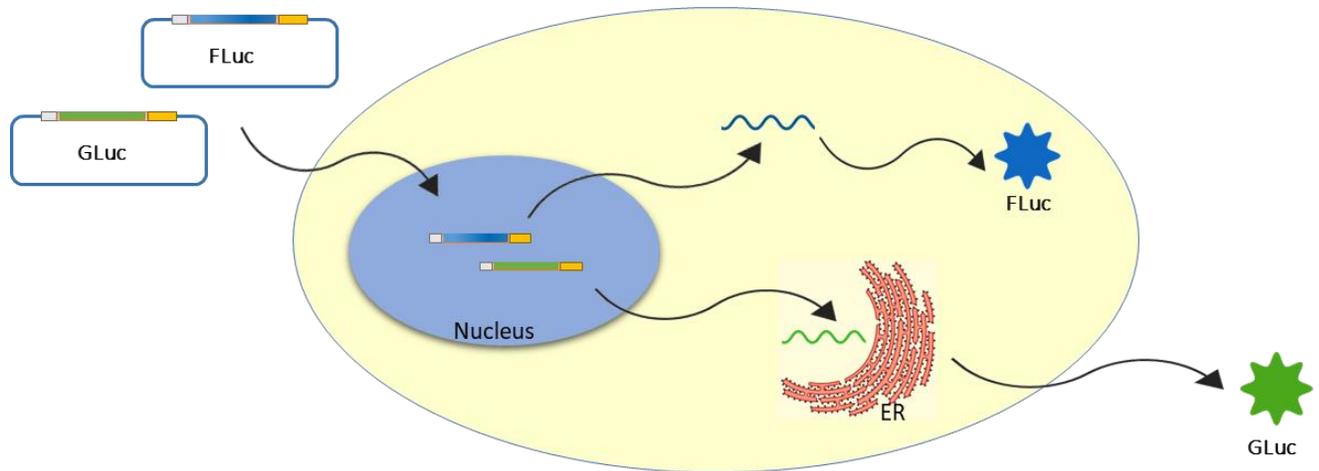


Figure 3.1: Translation of Gaussia and Firefly Reporters. Upon transfection into Huh7 cells, the Firefly reporter (blue) is translated in the cytoplasm and the protein is expressed intracellularly whilst the Gaussia luciferase reporter (green) is translated at the endoplasmic reticulum (ER) and then secreted into the media.

The Gaussia luciferase coding region from pBi-Gluc-H77C(1a)/JFH was amplified using PCR primers *StuI*_F and *SpeI*_R and inserted in place of the firefly luciferase coding region in the FlucM4 plasmid using *StuI* and *SpeI*. The resulting plasmid was designated GlucM4 and confirmed by sequencing with Source Bioscience stock pCep_R primer. Therefore, the GlucM4 and original FlucM4 luciferase plasmids possess the same vector backbone, promoter and Poly(A) signal as each other. Figure 3.2 provides a schematic showing the location of the miR-122 binding sites, and secretion signal the case of GlucM4, in relation to the luciferase coding region.

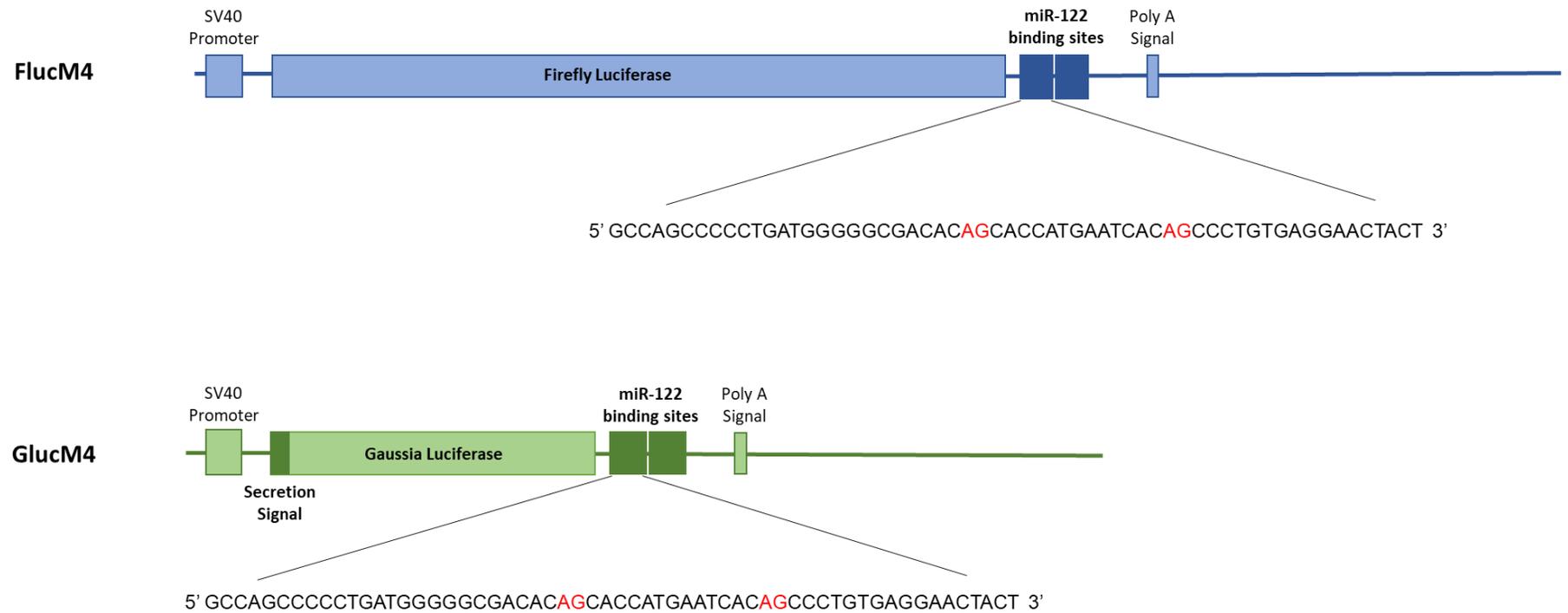


Figure 3.2: Gaussia and Firefly Reporters with mutant miR-122 binding sites. Schematic of the luciferase reporters with FlucM4 in blue and GlucM4 in green, showing the location of the SV40 promoter, poly(A) signal and two copies of the mutant miR-122 binding sites from the HCV 5' UTR in relation to the luciferase coding region, with the location of the secretion signal for Gaussia also indicated. The sequence of the mutant miR-122 binding sites are shown, with the mutations at positions 3 +4 highlighted in red.

3.3 FlucM4 and GlucM4 reporters are differentially regulated by pre-miR-122p3+4 expression

Reporter plasmids were co-transfected into Huh7 cells with either pre-miR-122p3+4, a synthetic pre-miR-122 oligonucleotide mutated with complementary nucleotides at positions 3 and 4, or pre-miR-122wt, a synthetic pre-miR-122 wildtype oligonucleotide. Both oligonucleotides are digested by endogenous Dicer to generate mature WT or p3+4 miR-122. The sequence of both can be found in Figure 3.3. As Huh7 cells contain a lot of endogenous miR-122, the mutant reporters allowed the effects of the mutant miRNA regulation without any background miR-122 regulation.

Previous work with the Firefly luciferase reporters in the Jopling lab confirmed FlucM4 regulation by miR-122p3+4 over a 48 hour timecourse [317]. Therefore, following similar conditions, transfections were performed in triplicate, each three wells containing 10 nM oligonucleotide, and 50 ng luciferase reporter plasmid, and harvested at 18-, 24-, and 48-hours post-transfection. For the Gaussia reporters, media was taken from the same wells at each timepoint and analysed directly with the GAR-2B Assay kit whilst the FlucM4 transfections were performed in parallel for each timepoint as the cells were lysed prior to analysis with the Single Luciferase Assay Reporter Kit as described in Section 2.5. The raw fluorescence intensity values were then measured, and the relative fluorescence intensity was calculated as an average of the technical triplicates for pre-miR-122p3+4 transfection expressed as a ratio of the pre-miR-122wt control condition.

As shown in Figure 3.3, at 18 hours post-transfections both FlucM4 and GlucM4 reporters show repression from pre-miR-122p3+4 at 0.55 and 0.31 the level of fluorescence for the wild-type control conditions, respectively. At 24 hours both reporters show stronger repression than at 18 hours, at 0.22 and 0.36 relative to the level of the control condition for FlucM4 and GlucM4,

respectively. However, at 48 hours a significant difference between the two reporters was observed, with GlucM4 remaining strongly repressed (0.13) whilst the FlucM4 reporter is beginning to recover and return to the expression of the wild-type conditions (0.72): GlucM4 is 87% repressed whilst FlucM4 is 3-fold less repressed at 28%. Although regulation of the firefly luciferase reporter is more variable than the Gaussia reporter, this suggests differential regulation occurring between the ER and cytoplasm, with stronger and more long-lasting miRNA-mediated repression of translation at the ER.

Renilla luciferase is a common transfection control for Firefly but unfortunately, as Gaussia uses the same substrate as Renilla, we were unable to easily co-transfect with this control reporter and therefore unable to take into account transfection efficiencies. In contrast to Figure 3.3, where only averages of the technical triplicates were considered, in Figure 3.4 the fluorescence for each individual technical triplicate (both pre-miR-122wt and pre-miR-122p3+4 transfections) at 24 and 48 hours was calculated relative to the average for the control condition (pre-miR-122wt) and presented as individual values for each technical replicate. Although there is some variability in the control conditions, it confirmed that in all cases pre-miR-122p3+4 significantly represses both FlucM4 and GlucM4 reporters and that the two reporters are differentially regulated at the 48 hour timepoint. At 48 hours, the FlucM4 fluorescence was more variable than GlucM4 with the latter showing a much stronger repression by pre-miR-122p3+4.

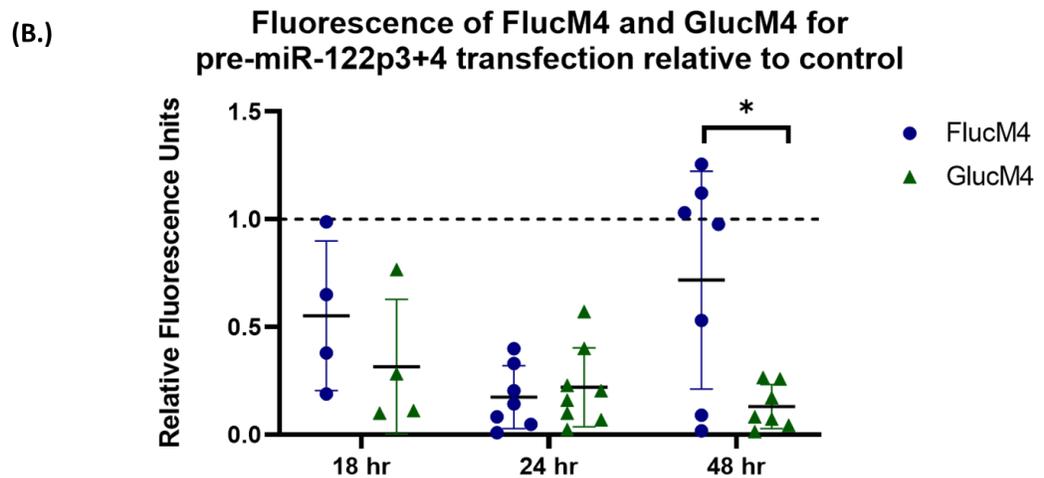
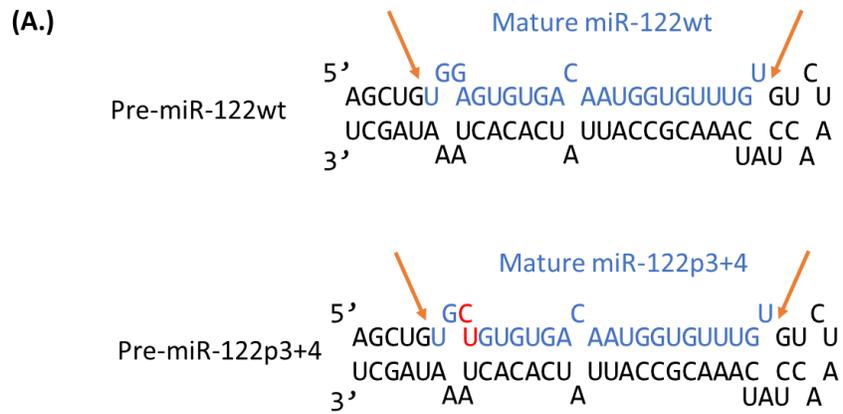


Figure 3.3: FlucM4 and GlucM4 reporters are differentially regulated over a 48 hr period by pre-miR-122p3+4 expression in Huh7 cells. (A) Reporter plasmids co-transfected with either pre-miR-122p3+4, a synthetic pre-miR-122 oligonucleotide mutated with complementary nucleotides at positions 3 and 4 (red), or pre-miR-122wt, a synthetic pre-miR-122 wildtype oligonucleotide. Mature miR-122 sequence, shown in blue, produced after cleavage with Dicer (cleavage sites indicated with an orange arrow). (B) Transfection with pre-miR-122p3+4 represses both FlucM4 (Blue) and GlucM4 (Green) reporters similarly at 18 hours and 24 hours, but at 48 hours the FlucM4 expression shows recovery towards the control conditions in some experiments whilst the GLucM4 expression remains strongly repressed. Luciferase values are an average of technical triplicates for pre-miR-122p3+4 transfection expressed as a ratio of the pre-miR-122wt control condition. Data represents at least four independent experiments, and error bars represent standard deviation. Statistical analysis by Student's T-test where $*=p<0.05$.

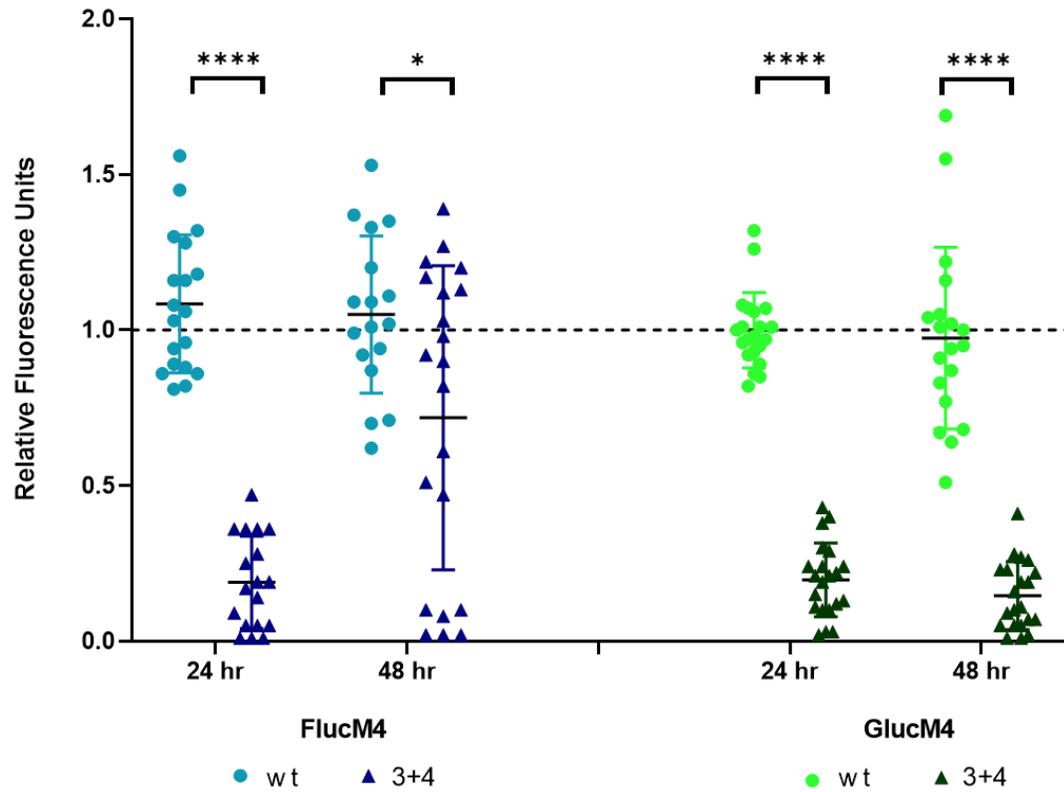


Figure 3.4: Regulation of both FlucM4 and GlucM4 by both pre-miR-122p3+4 and pre-miR-122wt control as technical replicates. Fluorescence data expressed as individual technical replicates relative to the average of the control pre-miR-122wt conditions. Whilst there is some variability in control conditions (wt, circles), transfection with pre-miR-122p3+4 (3+4, triangles) significantly represses both FlucM4 (blue) and GlucM4 (green) reporters as expected, and the two reporters show differential regulation at 48 hours confirming what was seen in Figure 3.3. Triplicate technical replicates from four independent experiments, error bars represent standard deviation. Statistical analysis by Students t-test; **** $p < 0.0001$, * $p < 0.05$.

3.4 Effect of eIF4All knockdown on the regulation of FlucM4 and GlucM4 reporters

As stronger repression of ER-localised reporters had been observed at 48 hours with pre-miR-122p3+4 expression, the FlucM4 and GlucM4 transfection experiment was repeated following knockdown of eIF4All to investigate if eIF4All has a differential role in miRNA function at the ER. As previously stated, eIF4All, but not eIF4AI, is critical for miR-mediated gene regulation and is thought to interact with the CCR4-NOT complex to aid translational repression [275]. Additionally, eIF4All has been shown to contribute to miR-122 activation of HCV which is known to replicate in ER-derived membranes [314, 475].

Huh7 cells were treated with eIF4AllsiA or a control non-targeting siRNA (detailed in Section 2.4.6), before seeding them into 24-well plates and transfections with the FlucM4 and GlucM4 reporters as in Section 3.3. The reporters showed the same pattern as in Figure 3.4 in cells transfected with the control siRNA, with GlucM4 remaining strongly repressed by pre-miR-122p3+4 at 48 hours whilst the FlucM4 expression begins to increase and return to that of wildtype conditions (Figure 3.5). The knockdown of eIF4All with siA shows no effect on FlucM4 repression by pre-miR-122p3+4 but was shown to relieve the repression seen for the GlucM4 reporter (at 0.04 and 0.07 with the control siRNA) with fluorescence at 0.36 and 0.47 relative to the control following knockdown of eIF4All at 24 and 48 hours, although due to a limited number of biological replicates the difference was not statistically significant.

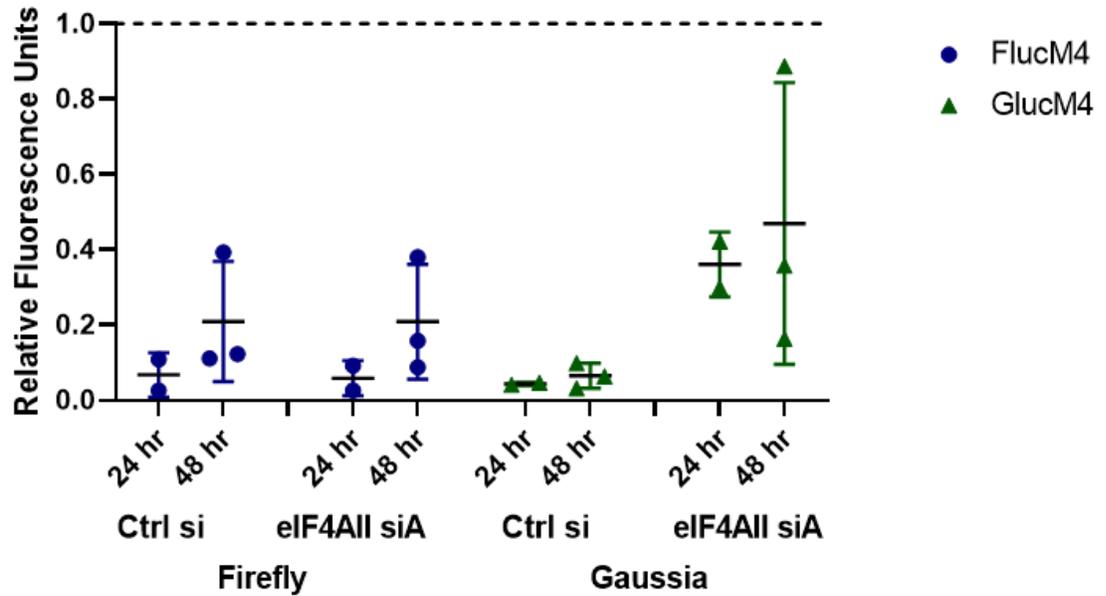


Figure 3.5: Knockdown of eIF4All with siA potentially has different effects on the regulation of FlucM4 and GlucM4 reporters pre-miR-122p3+4 expression in Huh7 cells. Following knockdown of eIF4All using siA, reporter plasmids were co-transfected with either pre-miR-122p3+4 or pre-miR-122wt oligonucleotides. Luciferase values are an average of technical triplicates expressed as a ratio of the pre-miR-122wt control condition. Data represents at least 2 independent experiments and error bars represent standard deviation.

To determine if this was a genuine biological effect of eIF4All, the transfections were repeated with an alternative eIF4All siRNA, siC. However, as seen in Figure 3.6, eIF4AllsiC transfection did not replicate that of siA, having little effect on either FlucM4 or GlucM4. Overall, the regulation by pre-miR-122p3+4 following eIF4All knockdown was more variable than previous experiments and the controls did not exhibit the same pattern of repression as seen without siRNA in Figure 3.4 and in the eIF4All siA experiments in Figure 3.6, with both FlucM4 and GlucM4 remaining around 0.25.

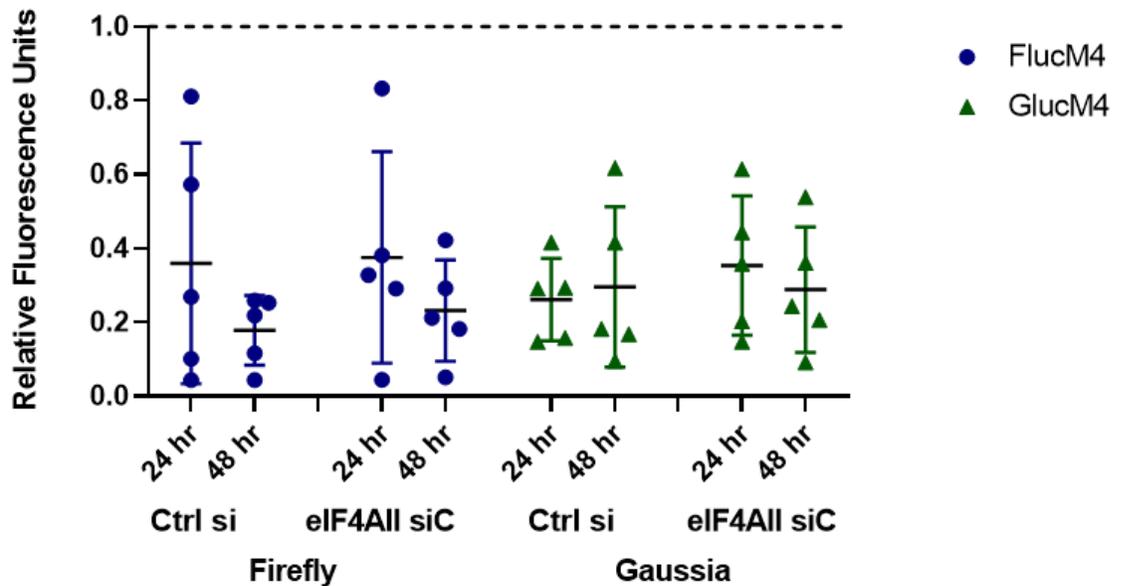


Figure 3.6: Effect of eIF4All Knockdown with siC on the regulation of FlucM4 and GlucM4 reporters pre-miR-122p3+4 expression in Huh7 cells. Following knockdown of eIF4All using siC, reporter plasmids were co-transfected with either pre-miR-122p3+4 or pre-miR-122wt oligonucleotides. Luciferase values are an average of technical triplicates expressed as a ratio of the pre-miR-122wt control condition. Data represents five independent experiments, and error bars represent standard deviation.

Although knockdown of eIF4All with siA was shown to relieve the repression seen for the GlucM4 reporter but not FlucM4 reporter, there was no observable difference in regulation of the reporters with eIF4All siC. While published work from the Jopling lab showed successful knockdown of eIF4All with siA and siC [317], given the variability of both the control and siRNA data in this chapter and lack of western blot to confirm successful knockdown, it is difficult to conclude from this experiment whether eIF4All plays a role in miRNA-mediated regulation of these reporter.

3.5 Generation of NanoLuc Reporters regulated by miR-122

To further investigate the different regulation seen at the cytoplasm and ER from Firefly and Gaussia reporters, additional luciferase reporters were constructed to incorporate the NanoLuc luciferase into vectors containing wildtype and mutant miR-122 binding sites. NanoLuc is a small 19.1 kDa reporter enzyme for which plasmids are available encoding a cytoplasmic form or a secreted version that is fused to an N-terminal signal peptide [476]

All NanoLuc reporters were generated from pNL2.3 (Supplementary Figure 5), encoding the secreted form, with the Nanoluc coding region amplified with or without the signal peptide; Nluc_R was used to incorporate an SpeI restriction site at the 3' end of the luciferase, Nluc_F was used to incorporate a start codon (ATG) after the secretion signal and create an NcoI site at the 5' end of the luciferase for the cytoplasmic reporters, and NlucSec_F was designed to mutate the valine residue to aspartate at the start of signal peptide to incorporate an NcoI site for the secreted reporters. The amplified NanoLuc coding regions +/- secretion signal were inserted in place of the firefly coding region in the FlucWT and FlucM4 reporters and sequencing confirmed generation of Nluc, NlucSec, NlucM4, and NlucSecM4. The resulting Nluc reporters (plasmids Section 2.1.2, Table 2.2) possess the same vector backbone, promoter and Poly(A) signal as well as the same miR-122 binding region as the Fluc and Gluc reporters located in the 3' UTR of either cytoplasmic or secreted NanoLuc plasmids as shown in Figure 3.7.

The wildtype reporters allow the study of regulation by endogenous and overexpressed miR-122 whilst the reporters with mutant binding sites allows the study of regulation by ectopically expressed mutant miR-122. Generating ER- and cytoplasmically translated reporters containing the same luciferase eliminates the possibility that differences in regulation at the two sites are due to inherent differences between luciferases such as coding sequence or ORF length but also using NanoLuc reporters allows the use of Firefly luciferase plasmids as a transfection control.

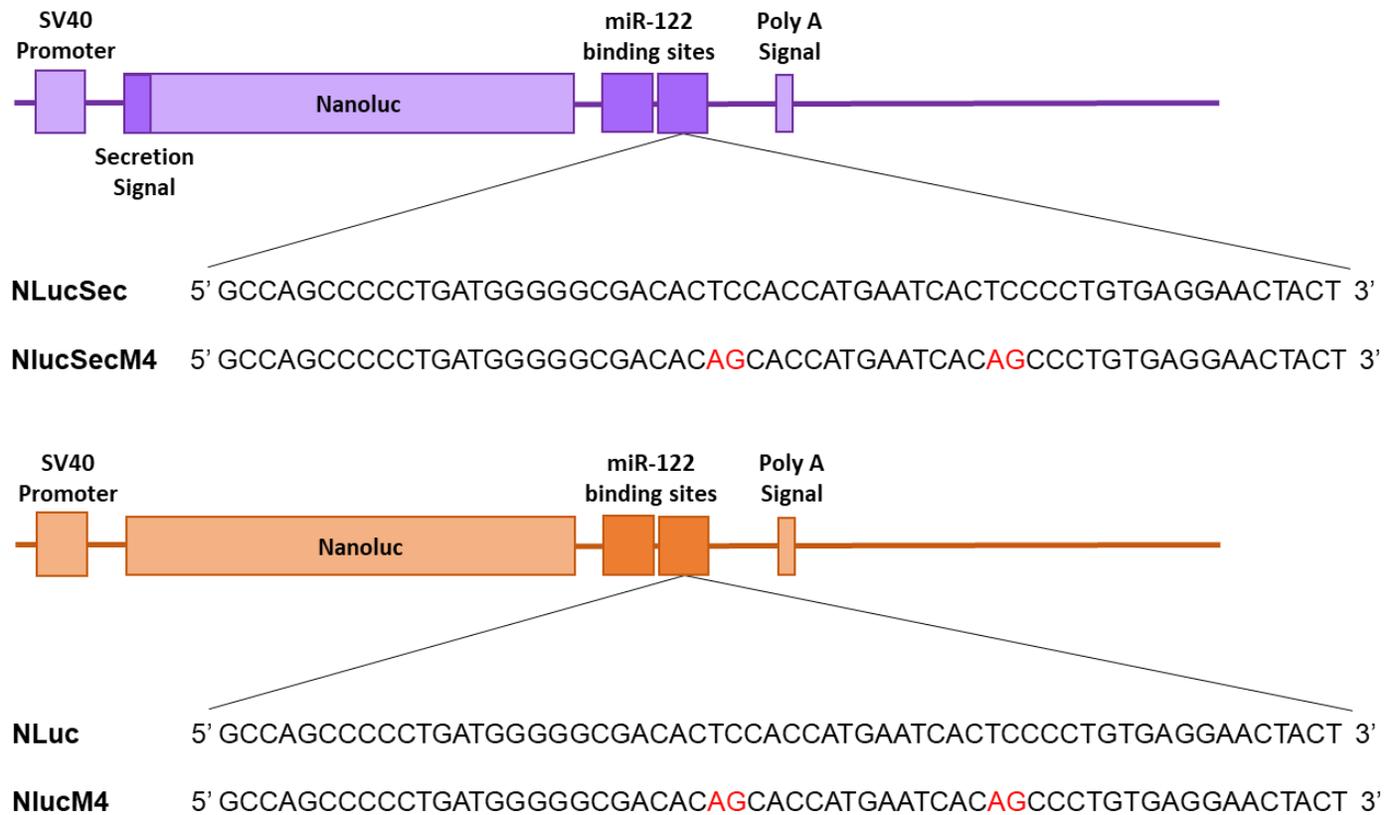


Figure 3.7: Schematic of NanoLuc Reporters with miR-122 binding sites. Showing the NanoLuc coding region in relation to the SV40 promoter, polyA signal, and miR-122 binding sites which contains either 2 copies of the HCV 5' UTR miR-122 binding sites with the wildtype sequence or mutated at two regions in red (M4). Cytoplasmic reporters (NLuc/NLucM4) are shown in orange and the ER-translated reporters (NLucSec/NLucSecM4) are shown in purple with the secretion signal at the 5' end of the luciferase coding region.

3.6 Regulation of cytoplasmic NlucM4 and secreted NlucSecM4 reporters by pre-miR-122p3+4 transfection

NanoLuc reporter plasmids containing the mutant miR-122 binding sites (Figure 3.7) were co-transfected into Huh7 cells with either pre-miR-122p3+4 or pre-miR-122wt as shown in Figure 3.3. The transfections were performed in triplicate, with each three wells containing 10 nM oligonucleotides and 50, 100 or 200 ng luciferase reporter plasmids. As assaying at 18 hours post-transfection did not result in strong repression of Firefly and Gaussia reporters in Section 3.3, both NlucM4 and NlucSecM4 expression was analysed at 24- and 48-hour timepoints using 10 μ L of media for the secreted reporter and 10 μ L of lysate for the cytoplasmic reporter.

In Figure 3.8 the relative fluorescence values were not normalised to a transfection control. Repression mediated by pre-miR-122p3+4 of the NlucM4 and NlucSecM4 reporters is seen at all timepoints but is weaker than the repression of the reporters seen in Section 3.3 where FlucM4 was 78% repressed and GlucM4 64% repressed at 24 hours. The level of repression for the ER-localised NlucSecM4 is similar at both timepoints and slightly but not significantly reduced at the 200 ng concentration (at 59% and 39% repressed relative to the control at 24 and 48 hours respectively) whilst the cytoplasmic NlucM4 reporter is unchanged with different concentrations or times, staying approximately 78% repressed. In contrast to the difference seen with the FlucM4 and GlucM4 reporters in Figure 3.3, the cytoplasmic and secreted NanoLuc reporters show no difference in regulation at any of the tested timepoints and plasmid concentrations.

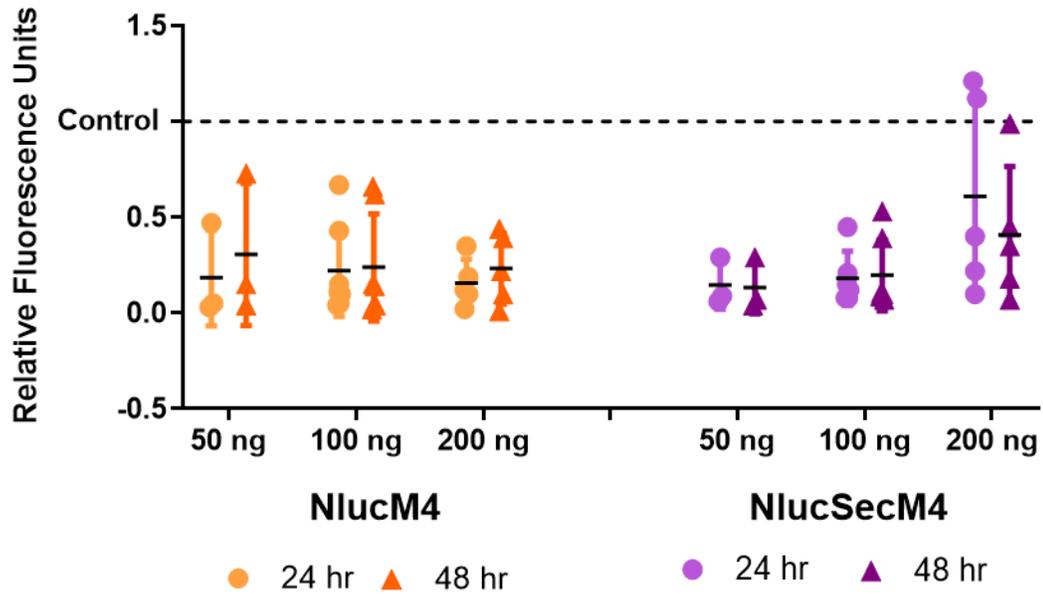


Figure 3.8: Cytoplasmic and Secreted NlucM4 reporter expression over a 48 hr time period regulated by pre-miR-122p3+4 expression in Huh7 cells. Reporter plasmids with mutant miR-122 binding sites were co-transfected with either pre-miR-122p3+4, a synthetic pre-miR-122 oligonucleotide mutated with complementary nucleotides at positions 3 and 4, or pre-miR-122wt, a synthetic pre-miR-122 wildtype oligonucleotide. Luciferase values are an average of technical triplicates expressed as a ratio of the pre-miR-122wt control condition with 50, 100, and 200 ng referring to plasmid concentration. Data represents at least three independent experiments, and error bars represent standard deviation.

Unlike when comparing Firefly and Gaussia reporters in Section 3.3, the use of the same luciferase for both ER and cytoplasmic reporters is well suited to the use of a common transfection control and dual luciferase kits are available for NanoLuc to include a Firefly luciferase transfection control. Therefore, the NanoLuc Reporters were co-transfected with 50 ng pGL3con (Firefly luciferase control plasmid) and assayed using the Nano-Glo Dual Reporter Assay System. This allowed the NanoLuc fluorescence value to be normalised to the respective firefly control prior to calculating an average, and the relative difference in luciferase activity for the average of each triplicate was then calculated by normalising to the average of the transfection control samples.

As seen in Figure 3.9, this confirms the pattern of repression we saw with the unnormalised NanoLuc reporters but with much stronger repression, with both cytoplasmic NlucM4 and secreted NlucSecM4 showing very strong repression by pre-miR-122p3+4 (all more than 80% repressed) with no difference between time of harvesting and concentration of plasmid. The inclusion of a firefly control reporter resulted in much less variation in data and stronger repression by pre-miR-122p3+4 overall.

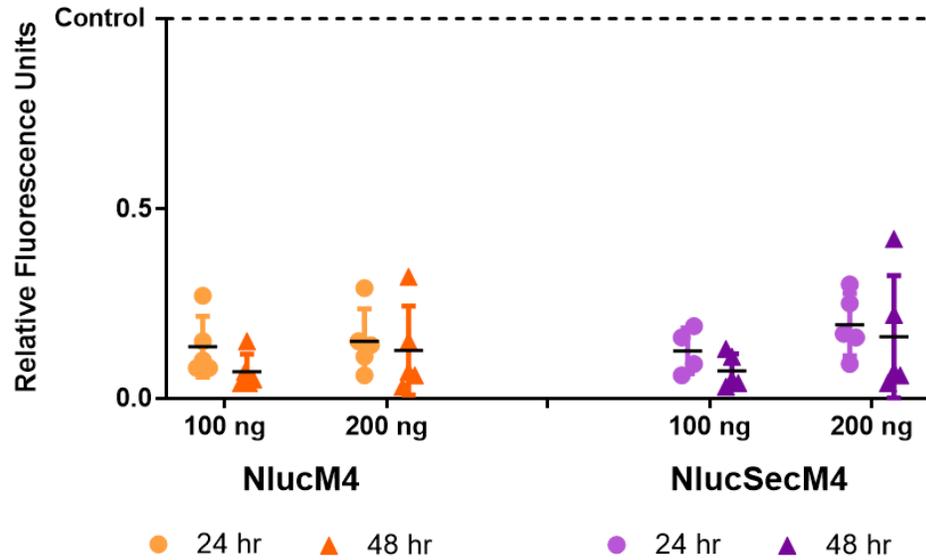


Figure 3.9: Cytoplasmic and Secreted NlucM4 reporters are regulated similarly over a 48 hr time period by pre-miR-122p3+4 expression in Huh7 cells. Reporter plasmids with mutant miR-122 binding sites co-transfected with a Firefly luciferase control reporter and either pre-miR-122p3+4, a synthetic pre-miR-122 oligonucleotide mutated with complementary nucleotides at positions 3 and 4, or pre-miR-122wt, a synthetic pre-miR-122 wildtype control oligonucleotide and harvested at 24 hrs and 48 hrs. Fluorescence data expressed as the ratio of Nluc to Fluc for pre-miR-122p3+4 transfection relative to the pre-miR-122wt control with 100/200ng referring to plasmid concentration. Data represents at least four independent experiments, and error bars represent standard deviation.

3.7 Comparison of secreted and cytoplasmic NanoLuc reporters by endogenous or overexpressed wildtype miR-122 binding to the 3' UTR

The transfection of mutant miR-122 may not be representative of endogenous miRNA, and the very strong repression observed might make it difficult to detect changes between the cytoplasm and ER. As a result, regulation of cytoplasmic and ER-localised reporters by endogenous miR-122 was investigated.

Reporter plasmids containing wildtype miR-122 binding sites downstream of the NanoLuc coding region (Figure 3.7) were co-transfected with a firefly luciferase control reporter (pGL3con). This allowed the study of wildtype miR-122 regulation following co-transfection with a miR-122 LNA inhibitor, detailed in Section 2.4.5. Huh7 cells were again transfected in triplicate, each three wells containing 100 or 200 ng luciferase reporter plasmids, 50 ng pGL3con, and 10 nM miR-122 inhibitor, with the concentration of reporters chosen based on other work in the Jopling research group. Both cytoplasmic (Nluc) and secreted (NlucSec) reporters were assayed using 20 μ L of lysate or media, respectively, harvesting at 24 and 48 hours. Fluorescence data is expressed as the ratio of Nluc to Fluc for LNA Inhibitor transfection relative to the respective LNA negative control.

Following miR-122 inhibition, both cytoplasmic Nluc and secreted NlucSec show an increased expression at both 24 and 48 hours when comparing the ratio of NanoLuc expression to Firefly expression relative to the control inhibitor, with minimal difference between 100 or 200 ng reporter plasmid transfected (Figure 3.10). However, at 48 hours NlucSec expression increased at both 24 and 48 hours at 2.77- and 2.53-fold that of the control condition with a greater increase than that observed with cytoplasmic Nluc

(2.12 and 1.71) when co-transfected at 200 ng with a 2-way Anova determining the difference to be statistically significant ($p=0.04$).

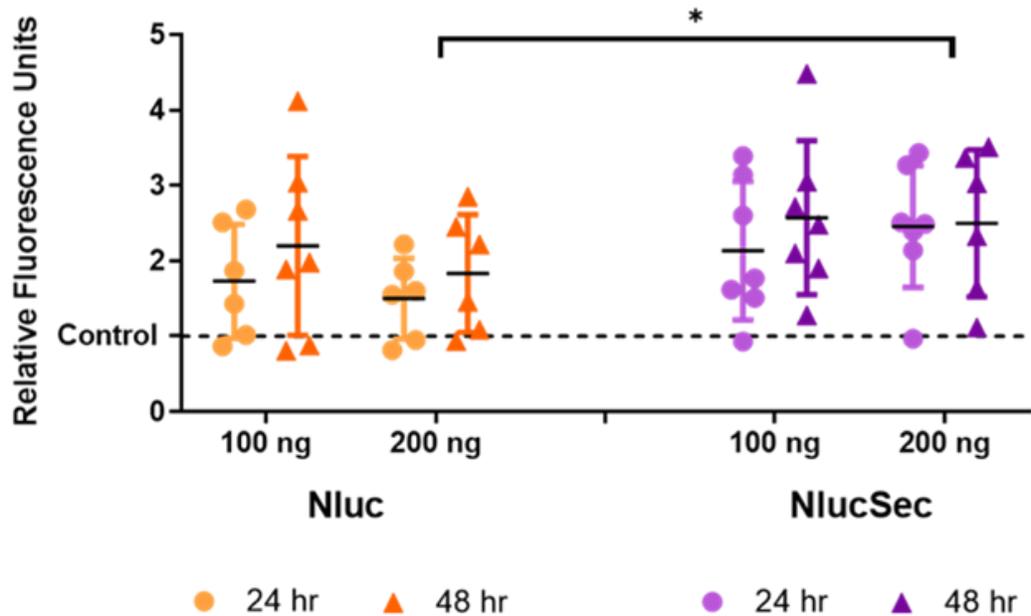


Figure 3.10: Cytoplasmic and Secreted Nluc reporter expression over a 48 hour time period following inhibition of endogenous miR-122 in Huh7 cells. Reporter plasmids containing wildtype miR-122 binding sites downstream of their luciferase coding region were co-transfected with a firefly luciferase control reporter and either a miR-122 LNA inhibitor or negative control.. Fluorescence data expressed as the ratio of Nluc to Fluc for LNA Inhibitor transfection relative to a LNA Negative control with 100/200 ng referring to plasmid concentration. Data represents at least six independent experiments, and error bars represent standard deviation. Statistical analysis by 2-way Anova where $*=p<0.05$.

It was then tested whether there would be a difference in regulation by overexpressed wildtype miR-122. The same conditions were followed as with the LNA inhibitor but co-transfecting with a miR-122 mimic at 2 nM as other work in the Jopling research group identified the requirement of less concentrated LNA mimics. The fluorescence data in Figure 3.11 is expressed as the ratio of Nluc to Fluc for LNA Mimic transfection relative to the respective LNA negative control. Both reporters are similarly repressed

across the 48 hour timecourse and for both reporters, transfection at 200 ng produces a slightly greater repression although not statistically significant. As with inhibition of miR-122 and the secreted GlucM4 experiments, NlucSec shows a greater response to miR-122 overexpression than the cytoplasmic reporter with expression at 0.49 and 0.37 that of the control condition at 24 and 48 hours versus 0.67 and 0.65 for the Nluc reporter, although in this case the difference was not statistically significant.

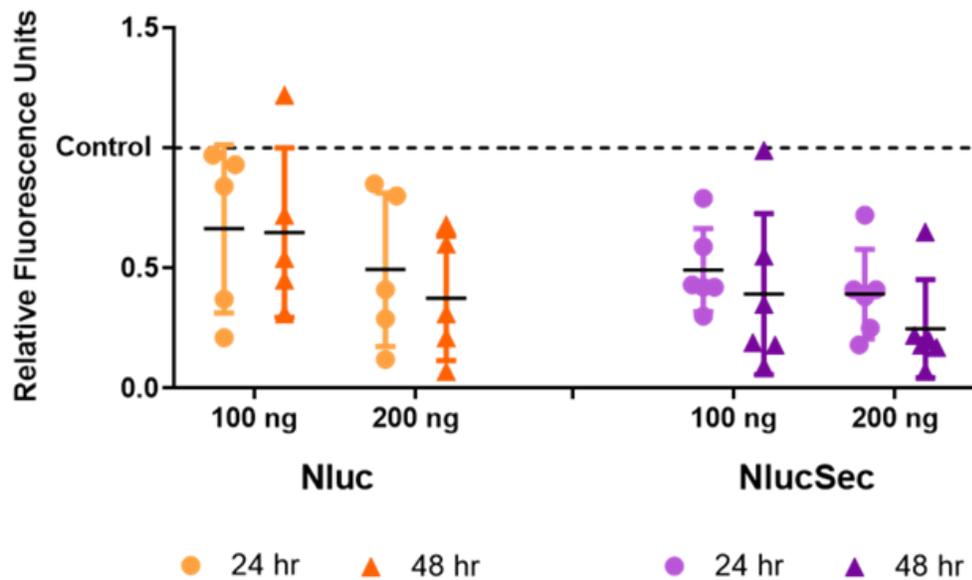


Figure 3.11: Cytoplasmic and Secreted Nluc reporter expression over a 48 hour time period following overexpression of wildtype miR-122 in Huh7 cells. Reporter plasmids containing wildtype miR-122 binding sites downstream of their luciferase coding region were co-transfected with a firefly luciferase control reporter and either a miR-122 LNA mimic or negative control. Fluorescence data expressed as the ratio of Nluc to Fluc for LNA Mimic transfection relative to a LNA mimic control with 100/200 ng referring to plasmid concentration. Data represents at least six independent experiments, and error bars represent standard deviation.

3.8 Regulation of secreted and cytoplasmic NanoLuc reporters by wildtype miR-122 binding to the HCV 5' UTR

Given the unique interaction between miR-122 and HCV and the fact that HCV is known to replicate in ER-derived membranes, cloning NanoLuc into reporter plasmids containing the 5' UTR and 3' UTR from HCV provided the opportunity to examine if there are any differences in regulation between the ER and cytoplasm in the context of up-regulation of translation by miR-122.

The Jopling lab has previously designed firefly reporter plasmids that contain the T7 promoter and can be used as a template for generation of mRNA lacking a cap and poly(A) tail with the firefly luciferase coding region flanked by the complete HCV (genotype 1a) 5' and 3' UTR which confer stability and allow efficient translation [443]. Translation of this RNA is stimulated by miR-122 binding to two sites in the 5' UTR. The amplified NanoLuc inserts with and without the secretion signal (generated in Section 3.5) were inserted into the 5'Fluc3' plasmid following digestion with Spel and NcoI. Sequencing using the corresponding Nluc_R/Nluc_F/NlucSec_F primers confirms the incorporation of the NanoLuc +/- secretion signal in the 5'Luc3' reporter backbone. Figure 3.12 shows a schematic of the 5'Nluc3' and 5'NlucSec3' reporters.

Uncapped, unpolyadenylated 5'Nluc3' and 5'NlucSec3' reporter RNA was produced by *in vitro* transcription. The same method was used to produce the Firefly luciferase reporter 5' M4Fluc3' which acted as the transfection control. This 5'M4Fluc3' control RNA contains the miR-122 seed matches mutated at positions 3 and 4 from TC to AG therefore not regulated by wildtype miR-122 but translation is regulated by the HCV IRES as with the 5'Nluc3' and 5'NlucSec3' reporters.

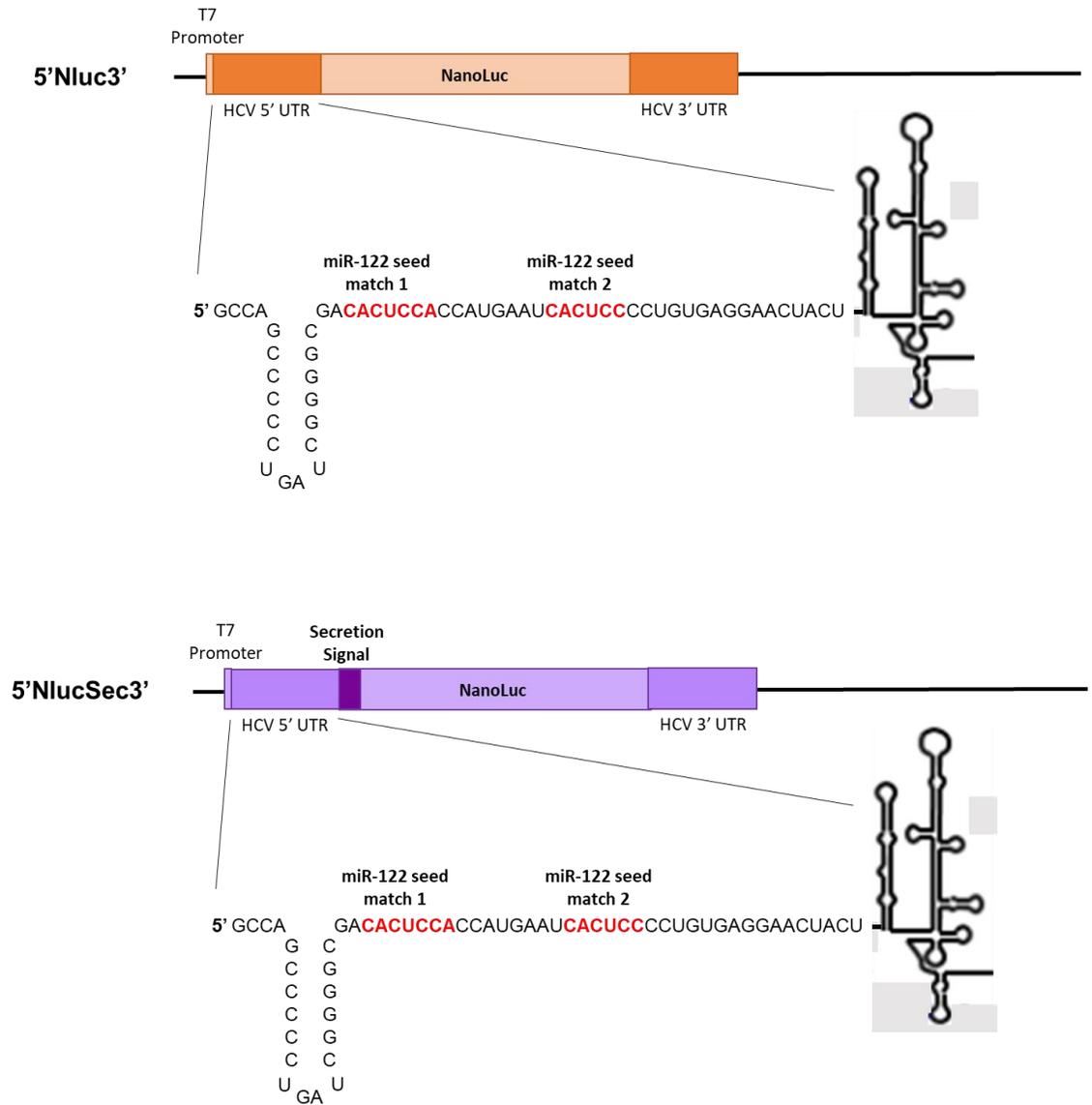


Figure 3.12: Schematic of NanoLuc Reporters with the 3' UTR and 5' UTR from HCV. Showing the luciferase coding region between the 3' UTR and 5' UTR from HCV, which includes two binding sites for miR-122 in red, under control of a T7 promoter. Cytoplasmic 5'Nluc3' is shown in orange and the ER-translated 5'NlucSec3' reporter is shown in purple with the secretion signal at the 5' end of the luciferase coding region.

Huh7 cells were co-transfected in triplicate with 50 ng control reporter RNA, either 100 or 200 ng reporter RNA with either 10 nM LNA miR-122 Inhibitor or negative control across three wells. Both cytoplasmic 5'Nluc3' and secreted 5'NlucSec3' were assayed using 20 μ L of lysate or media, respectively, harvesting at 4, 6, 8 and 24 hours to account for the quicker expression of RNA reporters which do not need to be transcribed following transfection, in contrast to plasmid transfection. Fluorescence data was expressed as the ratio of Nluc to Fluc for LNA Inhibitor transfection relative to the respective LNA negative control.

Following inhibition of miR-122, both cytoplasmic 5'Nluc3' and secreted 5'NlucSec3' are repressed as expected, confirming that miR-122 activates translation via the HCV 5' UTR as previously observed with 5'Fluc3'. There is a greater effect seen with 100 ng reporter for both secreted and cytoplasmic reporters (Figure 3.13). At 100 ng plasmid concentration, the cytoplasmic 5'Nluc3' expression is 0.44, 0.55, 0.50 and 0.33 relative to the control inhibitor at 4, 6, 8, and 24 hours respectively, with the secreted 5'NlucSec3' expression similarly at 0.39, 0.45, 0.49, and 0.38 relative to the control inhibitor across the 24 hours timecourse. At 200 ng concentration, both cytoplasmic and secreted NanoLuc reporters are repressed by an average of 35.3% and 35.4% respectively across the 24 hour timecourse, at 0.73, 0.75, 0.56, and 0.54 expression relative to the control inhibitor for the cytoplasmic reporters, and 0.56, 0.57, 0.67, and 0.78 for the secreted 5'NlucSec3' reporter. Ultimately, there is no difference over the 24-hour timecourse for either reporter, nor does there seem to be any difference in regulation between the cytoplasmic and ER-translated reporters.

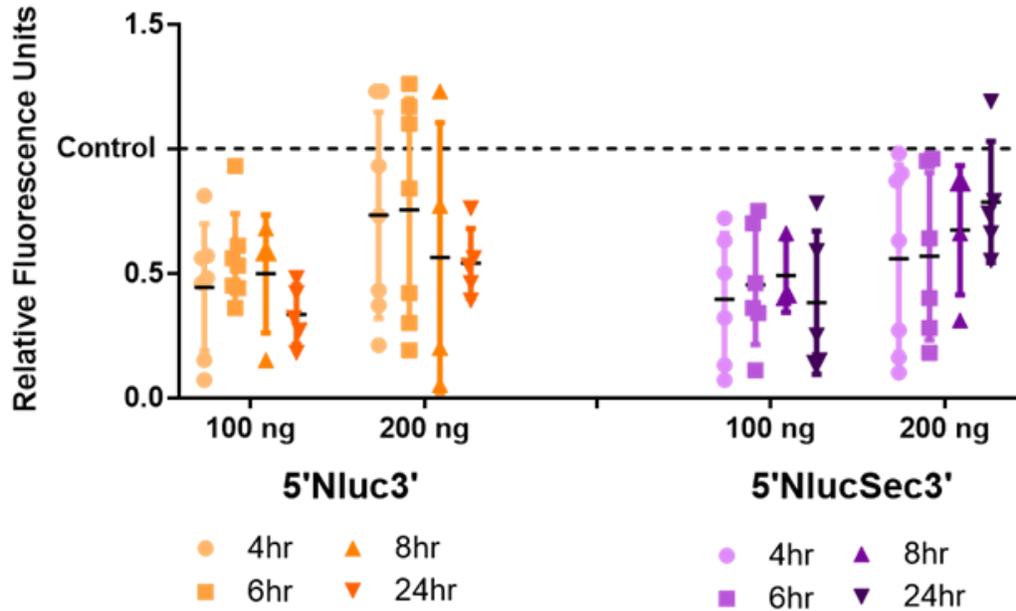


Figure 3.13: Cytoplasmic and Secreted Nluc reporter expression under control of the HCV 5' UTR over a 24 hour time period following inhibition of endogenous miR-122 in Huh7 cells. Reporter plasmids containing the 5' UTR from HCV with two miR-122 binding sites were co-transfected with a firefly luciferase control reporter and either a miR-122 LNA inhibitor or negative control. Fluorescence data is expressed as the ratio of Nluc to Fluc for LNA Inhibitor transfection relative to a LNA Negative control with 100/200 ng referring to plasmid concentration. Data represents at least four independent experiments, and error bars represent standard deviation.

The 5'Nluc3' and 5'NlucSec3' reporters were also used to compare regulation by overexpressed wildtype miR-122 binding to the 5' UTR, using 2 nM LNA miR-122 mimic in the same transfection conditions as with the LNA Inhibitor. In Figure 3.14 the fluorescence data was expressed as the ratio of Nluc to Fluc for LNA Mimic transfection relative to the respective LNA negative control and both reporters show increased expression following overexpression of miR-122 with an LNA mimic as expected.

For the secreted 5'NlucSec3' reporter, there is no difference in expression between 100 ng and 200 ng concentration of plasmid, remaining around 1.5-fold increase upon miR-122 overexpression across 4 to 8 hours, but

increasing to 2.58- and 2.31- fold increase at 24 hours for 100 and 200 ng respectively. The pattern of fluorescence intensity suggests a greater activation of translation from the cytoplasmic reporter 5'Nluc3' than the ER reporter following miR-122 overexpression, increasing from 2.25-fold to 3.07-fold at 100 ng and 1.82- to 2.61-fold at 200 ng over the 24-hour timecourse. However, as it is much more variable than the secreted reporter it is not a statistically significant difference.

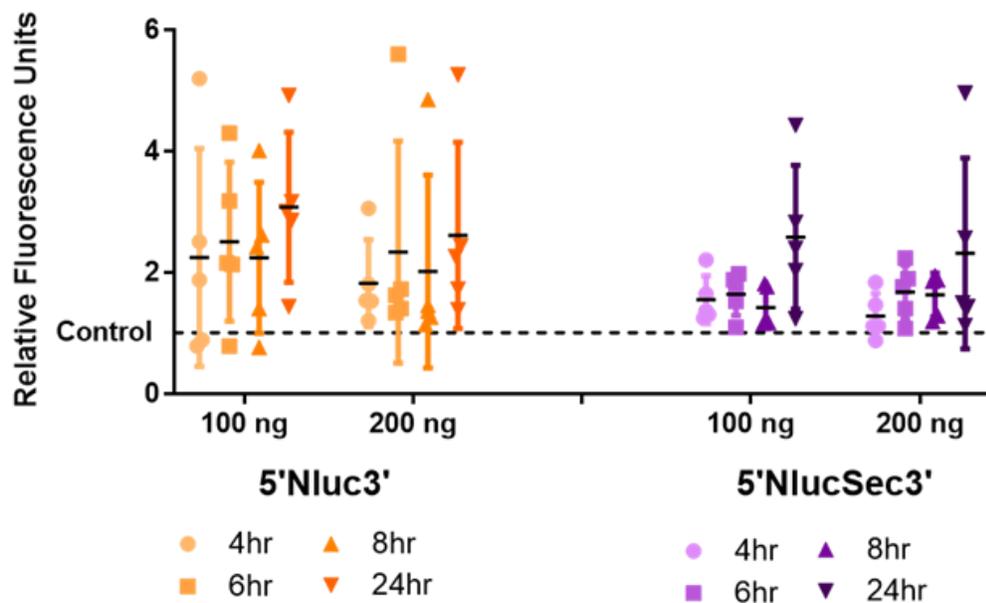


Figure 3.14: Cytoplasmic and Secreted Nluc reporter expression under control of the HCV 5' UTR over a 24 hour time period following overexpression of miR-122 in Huh7 cells. Reporter plasmids containing the 5' UTR from HCV with two miR-122 binding sites were co-transfected with a firefly luciferase control reporter and either a miR-122 LNA mimic or negative control. Fluorescence data is expressed as the ratio of Nluc to Fluc for LNA Mimic transfection relative to a LNA mimic control with 100/200 ng referring to plasmid concentration. Data represents at least four independent experiments, and error bars represent standard deviation.

3.9 Regulation of miR-122 reporters at the level of RNA stability

miRNAs binding to 3' UTRs have been shown to repress translation and destabilise mRNAs so in an attempt to test directly whether there were differences in miR-122 regulation of target mRNA stability in the cytoplasm versus the ER, the luciferase reporters (FlucM4, GlucM4, Nluc, NlucSec) were transfected in scaled up reactions and the total RNA was extracted for analysis by RT-qPCR and/or Northern Blotting.

The FlucM4 and GlucM4 were transfected as in previous experiments and total RNA was harvested at 48 hours using Tri reagent. qRT-PCR was performed with primers specific for firefly and gaussian luciferase mRNA, using 1 μ L of undiluted RNA that has not been subjected to reverse transcription in the qPCR reaction for the “no RT” control. The Nluc reporters were transfected into Huh7 cells with 100 ng reporter and either 10 nM LNA Inhibitor or 2 nM LNA Mimic and harvested at 48 hours. They then underwent an additional DNAase treatment and qRT-PCR was performed for NanoLuc. RNA was diluted to 100 ng for the “no RT” control to match the concentration of RNA used in the RT reaction. Despite extensive efforts to eliminate this, there were continued problems with DNA or PCR product contamination with representative data shown in Table 3.1.

To circumvent the qRT-PCR contamination, northern blots were attempted with the same RNA from the GlucM4 and FlucM4 assays, loading 100 ng total RNA and probing with radioactive probes generated from Gluc PCR products and an existing Fluc probe used in the Jopling Lab. Despite the gel running well, methylene blue stain confirming good transfer, and good scintillation counts for the probes (Fluc 406k, Gluc 260k), probing for Fluc and Gluc overnight was unable to produce bands. An example image can be found in the appendix, Supplementary Figure 6.

| Primer Set | Condition | cDNA | no RT |
|----------------|--------------|-------|-------|
| FlucCLJ | p3+4 | 23.98 | 24.13 |
| | Ctrl | 16.13 | 12.49 |
| FlucPW | p3+4 | 20.05 | 16.85 |
| | Ctrl | 17.58 | 13.35 |
| Gluc | p3+4 | 25.87 | 23.75 |
| | Ctrl | 16.56 | 13.86 |
| Nluc | Inhib | 24.22 | 24.99 |
| | Ctrl | 24.91 | 24.56 |
| | Mimic | 25.08 | 24.15 |
| | Ctrl | 23.54 | 24.24 |
| NlucSec | Inhib | 23.87 | 24.17 |
| | Ctrl | 25.24 | 24.25 |
| | Mimic | 25.60 | 25.07 |
| | Ctrl | 25.43 | 24.54 |

Table 3.1: Contamination prevents accurate detection of luciferase mRNA by qRT-PCR. Ct values obtained from qRT-PCR using respective luciferase primers for both the cDNA generated from luciferase assays or the RNA directly as a no RT control. Total RNA extracted from Huh7 cells 48 hours after transfection of luciferase reporters with RNA for the “No RT” control from Nluc and NlucSec samples diluted to 100 ng as used in the RT reaction but was undiluted in the Fluc and Gluc samples.

3.10 Discussion

We chose to investigate miR-122 regulation at the ER as it is known to contain high levels of both miRISCs and ribosomes, with some studies suggesting the possibility of alternative roles for miRNA residing at the ER [344-347, 349-351]. Despite this, a direct comparison of miRNA-mediated regulation of mRNAs translated at the ER and cytoplasm in human cells has not been done before. To do so, luciferase reporter constructs were successfully generated to study miR-122 regulation via 3' UTR and 5' UTR binding sites using both endogenous miR-122 and synthetic pre-miR-122 oligonucleotides. Luciferase reporter assays are useful for determining a direct functional connection between the microRNA of interest and the luciferase reporter and the low stability of luciferase reporters allows the real-time monitoring of miRNA activity in cells. Reporter assays are able to achieve a high sensitivity with low background noise, but are substrate dependent and, depending on the luciferase, dependent on cellular lysis.

Excitingly, the comparison between the cytoplasmically translated firefly luciferase reporter (FlucM4) and secreted Gaussia luciferase (GlucM4) showed that the two reporters are differentially repressed by a synthetic pre-miR-122 oligonucleotide at the 48 hour timepoint (Figure 3.3). This suggests differential regulation occurring between the ER and cytoplasm, with stronger and more long-lasting miRNA-mediated repression of translation at the ER. This difference might be explained by concentration of miRNA at the ER, for example, and although we used ectopically expressed rather than endogenous miRNA it still undergoes processing by Dicer. Therefore, it was reasonable to expect that it would also be enriched at the ER although this has not been investigated and would be difficult to do so by qRT-PCR as it would pick up both the ectopically expressed and endogenous miR-122.

Subsequent experiments with the NlucM4 and NlucSecM4 reporters co-transfected with a control luciferase reporter showed less variability between biological replicates but similar regulation with pre-miR-122p3+4, with both

cytoplasmic and secreted reporters strongly repressed at 24 and 48 hours, without the relief in the cytoplasmic NanoLuc that was seen with the equivalent Firefly Reporter. Nevertheless, additional experiments using NanoLuc reporters containing wild-type miR-122 binding sites at the 3' UTR were performed. This provided the opportunity to study endogenous miR-122 regulation, as the synthetic pre-miR-122 oligonucleotides are an artificial system and might not fully replicate miR-122 regulation at either site. Following inhibition of miR-122 with an LNA inhibitor, both Nluc and secreted NlucSec showed an increased expression but at 48 hours, NlucSec is more greatly expressed than Nluc when transfected at 200 ng. Similarly, following overexpression of miR-122 with an LNA mimic NlucSec potentially shows a greater response at 24 and 48 hours, but in this case it is not statistically significant. Despite other work in the Jopling lab influencing the transfection conditions, it may be that transfection of greater than 2 nM and 10 nM for the LNA mimic and inhibitor, respectively, might allow the difference between cytoplasmic and ER regulation to be more easily detectable. Overall, there does seem to be a trend for greater regulation of secreted than cytoplasmic reporters, but experimental variability made it difficult to draw firm conclusions.

Using NanoLuc afforded the opportunity to incorporate a control reporter as potential sources of variability can come from differences in cell number, lysis efficiency, cell viability, transfection efficiency, temperature, and measurement time. For the initial experiments with Firefly and Gaussia, there was not a suitable control as both NanoLuc and Renilla luciferases are capable of catalysing the same substrate as Gaussia (coelenterazine). Data from the unnormalised NanoLuc experiments (Figure 3.8) was very variable and comparison with the same reporters normalised to a firefly reporter (Figure 3.9) confirms the benefit of including a transfection control. An alternative transfection control to further investigate the differences we saw between Gaussia and Firefly reporters would be co-transfecting with a vector expressing β -galactosidase. Cell lysis could be achieved in the same buffer

as required for analysis of firefly luciferase, and β -galactosidase activity of each lysate could be measured using an assay system that contains the substrate ONPG (o-nitrophenyl- β -d-galactopyranoside) and measuring the absorbance at 420nm by spectrophotometry. Alternatively, normalisation to the total protein using a Bradford assay could account for variability in the total cell number but our attempts at using protein concentration to normalise FlucM4 and GlucM4 expression introduced more variability than analysis of raw luminescence values, likely due to variability in protein extraction.

As we were able to conclusively determine differences in regulation of the cytoplasmic Firefly luciferase and secreted Gaussia luciferase but not detect as strong an effect with the NanoLuc reporters it may be that the difference is due to inherent differences in the reporters, such as mRNA stability, codon-optimisation, and protein turnover. Whilst the expression of transiently transfected intracellular reporters is known to reach maximum accumulation around 48 hours and then decline beyond this [477], as Gaussia requires both protein expression and then secretion into the media the time-course may be different. In addition, secreted reporter proteins may be able to avoid intracellular turnover and continue to accumulate beyond 48 hours [478]. Whilst a systematic comparison of luciferase reporters has not been performed, Wider and Picard (2017) showed that Gaussia protein accumulated beyond 48 hours and up to 66 hours and it is known that Firefly luciferase has a short protein half-life and more susceptible to enzyme inhibition compared to Renilla Luciferase [478]. To investigate our reporters in Huh7 cells, measuring the reporter expression with qRT-PCR at several timepoints after treatment with the transcription inhibitor ActinomycinD would determine if there is a difference in mRNA stability between Gaussia and Firefly. Likewise, if Firefly luciferase is turned over quicker than Gaussia, its fluorescence would decrease quicker after treatment with the translation inhibitor cycloheximide. Unfortunately, attempts to detect differences in regulation at the mRNA level were unsuccessful. Despite extensive efforts, there were persistent problems with DNA or PCR product contamination for

qRT-PCR and northern blot methodology and/or radioactive probes require optimisation to achieve quantification of the luciferase RNA.

For Gaussia expression in Section 3.3, media was harvested from the same well over the timecourse, whereas for secreted Nanoluc in Sections 3.6-3.8, media was harvested from different wells as lysates were also harvested to measure the control reporter. Likewise, both the cytoplasmic Firefly reporter and Nanoluc reporter were harvested from different wells at each time point therefore the difference in regulation seen with the Gaussia reporter may be due to a difference in harvesting as it isn't replicated with the NanoLuc system. Changing the media at each timepoint would mean that instead of seeing a build-up of Gaussia protein, we see a snapshot of expression over the timecourse to replicate the harvesting techniques more closely between secreted and cytoplasmic reporters. Alternatively, if the NanoLuc reporter could accommodate, rather than a secretion signal that results in the reporter protein being harvested from the media, a KDEL membrane retention signal could be inserted into the reporter plasmid. The reporter would still be translated at the ER but would remain there, thus requiring lysis prior to assay, and would enable a more direct comparison of regulation at the ER and Cytoplasm.

It is also worth considering the codon usage for the different luciferases, as optimisation for the human bias can increase the translational efficiency of the luciferase proteins. According to the suppliers all three luciferases (Gaussia, Firefly, and NanoLuc) are codon optimised for humans. To confirm this, the sequences of our reporter constructs were analysed by Chris Roberts (Heery lab) using a script that calculates the difference between the actual codon used and the optimal codon based on the Kazusa database for human codon usage [479, 480] and scores the total divergence from the optimal codon usage, correcting for the differences in the number of codons for each luciferase (Supplementary Table 1-3). The closer this score is to zero, the closer that gene is to optimal human codon usage, with human

gene P15K1A, for example, scoring 0.082. Gaussia was determined to have a score of 0.051, Firefly 0.098, and NanoLuc 0.067 which confirms all three are codon optimised and that this is unlikely to account for the difference in regulation at 48 hours for Gaussia and Firefly. However, this is just a rough calculation and does not consider G/C content or secondary structures for example.

Calculations have determined Gaussia luciferase to have 64.1% GC content whilst Firefly has a relatively low 45% [481, 482] which could account for the difference as G-C rich mRNA are more likely to be structured and could potentially impact the scanning ribosome [483]. In addition, the open reading frame (ORF) for Gaussia luciferase is much shorter than that of Firefly, at 558 bp and 1659 bp respectively, which alongside the GC content could affect the rate of translation of the luciferases. It is important to note that the 3' UTR sequences of both reporters are identical therefore there should be no difference in the efficiency of miRNA binding. In addition, the raw fluorescence for the control pre-miR-122wt transfections were plotted, Supplementary Figure 7, to consider if the differences seen were due to different expression between reporters under wild-type conditions. This confirmed there is no difference in the fluorescence between FlucM4 and GlucM4 at both 24 and 48 hours and similarly between NlucM4 and NlucSecM4, with all four mutant constructs roughly doubling in fluorescence intensity between 24 and 48 hours. However, the inherent differences between the Gaussia and Firefly reporters were the main drive to generate a series of NanoLuc reporters and could account for the stronger and more long-lasting miRNA-mediated repression of GlucM4 at the ER.

Cloning NanoLuc into reporter plasmids containing the 5' UTR and 3' UTR from HCV allowed the unique opportunity to investigate if there are any differences in regulation between the ER and cytoplasm in the context of up-regulation of translation mediated by miR-122. Whilst there is no difference over the 24 hour time course in the context of miR-122 inhibition, the pattern

of regulation suggests a greater response with the cytoplasmic reporter following miR-122 overexpression. However, this data is not statistically significant, likely due to the variability in relative fluorescence for the cytoplasmic 5'Nluc3' reporter. This variability could possibly come from the additional lysis step compared to the secreted reporter, which is assayed directly from the media, although the incorporation of the control luciferase plasmid should account for this as it also analysed using the lysate. As luciferase data following transfection of both RNA reporters are more variable than the 3' UTR plasmid reporters it may be that RNA transfection, or generation of the RNA constructs, introduces more variability, enabling small changes to go undetected. For the 5'Nluc3' reporters, it might be that there is a difference in regulation in the context of an actual HCV infection, which would disrupt the endomembranes, but not in a healthy Huh7 cell context. This would be interesting to investigate in the future, but data from the NanoLuc reporters does seem to show a contrast between 3' and 5' UTR regulation.

It is also important to consider the effect of transfecting both mutant pre-miR-122 oligonucleotides and overexpression of miR-122 with the LNA mimic on translation at the ER and total cell in general. Overexpression of wild-type miR-122 may dysregulate global regulation and if ER translation is affected, then there may be an indirect effect on the mRNA/protein levels measured of the targets rather than a direct effect of miR-122 regulation. Alternatively, high levels of the mutant miR-122 may compete with RISC and other regulatory machinery, therefore the changes observed may be due to decreased total miRNA or decreased wild-type miR-122 rather than specifically an increase in the mutant miR-122 regulation on the mutant reporter. Likewise, excess mutant versions of miR-122 may trigger the stress response which could lead to a pause in global translation until the mutant miRNA are eliminated. To investigate the possibility that global translation is affected, overexpression of a non-targetting miRNA alongside the existing reporter conditions could be used as a control.

Additionally, for future work it would be recommended to include a control for the signal peptide on the NanoLuc reporters to both confirm the mutated signal peptide (Val – Asp) still directs the reporter to the ER and that the changes in expression seen are a direct effect of regulation at the ER. This could be a mutation to the signal peptide that abrogates its signal retention signal so it is no longer recognised by the SRP and therefore the mRNA remains in the cytoplasm. Alternatively, substitution with a different signal peptide from another secreted protein could confirm the targeting of the reporter to the ER which should produce the same regulation pattern as seen with the existing signal peptide. As discussed earlier, an inclusion of a KDEL membrane retention signal rather than a secretion signal would enable the assay of ER-translated protein from the lysate eliminating the variability between harvesting of Nluc and NlucSec reporters.

In summary, luciferase reporter assays comparing cytoplasmic Firefly and secreted Gaussia showed a difference in regulation at 48 hours, suggesting differential regulation occurring between the ER and cytoplasm. This was followed up by generating ER- and cytoplasmically translated reporters containing the same luciferase (NanoLuc) which eliminates the possibility that differences in regulation at the two sites are due to inherent differences between luciferases such as coding sequence or ORF length. The NanoLuc reporter system also allowed the use of Firefly luciferase plasmids as a control reporter as potential sources of variability can come from differences in cell number, lysis efficiency, cell viability, transfection efficiency, temperature, and measurement time. Although we did not replicate the regulation we saw with the Firefly and Gaussia reporters with the NanoLuc reporters, the changes that we see are worth following up, possibly with more concentrated LNA inhibitors and mimics. Additionally, although eIF4All knockdown with the different siRNA showed conflicting results, it would be of interest to perform the luciferase reporter assays following knockdown of RISC components such as TNRC6A/B/C, AGO1-4, DDX6, and eIF4All with the full set of NanoLuc reporters and include controls to confirm knockdowns

are successful. Finally, as the luciferase reporters are an artificial system, they might not fully replicate the endogenous regulation by miR-122 in Huh7 cells, and so investigation into miR-122 targets that are known to associate to different subcellular sites would be interesting and is addressed in the next chapter.

Chapter Four:

**miR-122 Regulation of
Endogenous mRNA
Targets at the ER and
Cytoplasm**

4.1 Introduction

Recent studies have found miRNA and RISC components at alternative subcellular sites to the cytoplasm, such as the endoplasmic reticulum (ER), and as the ER enables microRNA-mediated mRNA silencing of secreted transcripts, this suggests different subpopulations of microRNA could act on a different subset of mRNA targets with the possibility of differential regulation at the different sites [49-51, 344-351]. Additionally, preliminary data from the Jopling lab, discussed in Section 4.5, showed miR-122 interaction with RISC in the membrane fraction and suggested this association may be stronger than in the cytoplasm.

Whilst the luciferase reporters developed in Chapter 3 allowed the study of direct regulation by miR-122, through both 3' UTR binding site repression and 5' UTR binding site activation, it is an artificial system and therefore unable to consider the complexities of endogenous miRNA-mediated regulation in a cell. Therefore, a membrane fractionation method was used to isolate ER and cytoplasm-localised mRNAs and the effects of miR-122 manipulation on known miR-122 targets were compared in these fractions, providing an insight into the difference in microRNA regulation at the ER and the cytoplasm.

Membrane-associated fractions were isolated using a sequential detergent method adapted from Jagannathan *et al* [471] which makes use of the difference between plasma membrane and ER membrane in terms of lipid composition. The detergent digitonin solubilises the cholesterol-rich plasma membrane but leaves ER and nuclear membranes intact, thereby releasing cytosolic elements. Then subsequent permeabilisation with Nonidet P-40 (NP40) and sodium deoxycholate (DOC) lyses the remaining membranes, allowing the isolation of ER-rich fractions. Although this is a crude fractionation approach, separating intracellular membrane-bound organelles into one fraction, it is compatible with a range of other techniques such as

AGO-IP and polysome fractionation and therefore useful for interrogation of miRNA function in the ER versus cytoplasm.

miR-122 has a number of well-documented mRNA targets but to allow comparison between regulation at the ER and the cytoplasm, targets that differentially associate with the ER and cytoplasm needed to be identified. Aldolase A (ALDOA) is a cytoplasmically localised glycolytic enzyme that is highly expressed in adult muscle tissues, but repressed in adult liver, kidney and intestines and is one of the top three upregulated microRNA in hepatocellular carcinoma cells [415, 420, 484-486].

The cationic amino acid transporter (CAT1, also known as SLC7A1) is expressed in all adult mammalian tissue except the liver [154]. During liver development, CAT1 mRNA decreases with increased levels of miR-122 and has been shown to contain three miR-122 binding sites in its 3' UTR [131, 154]. CAT1 has been reported to predominantly localise in caveolae (plasma membrane domains) but Bhattacharyya *et al* showed that miR-122 repression of CAT1 targets the mRNA to P-bodies [131, 487]. CAT1 transcription has been reported to be induced during ER stress through phosphorylation of eIF2A and increased levels of ATF4 and XBP transcription factors [488, 489]. In addition, unpublished data from the Jopling lab (Figure 4.1) shows CAT1 to be six-fold more enriched in membrane-associated fractions versus cytoplasmic fractions when compared to actin, whereas ALDOA is two-fold enriched, although with some experimental variability this difference is not statistically significant ($p=0.056$).

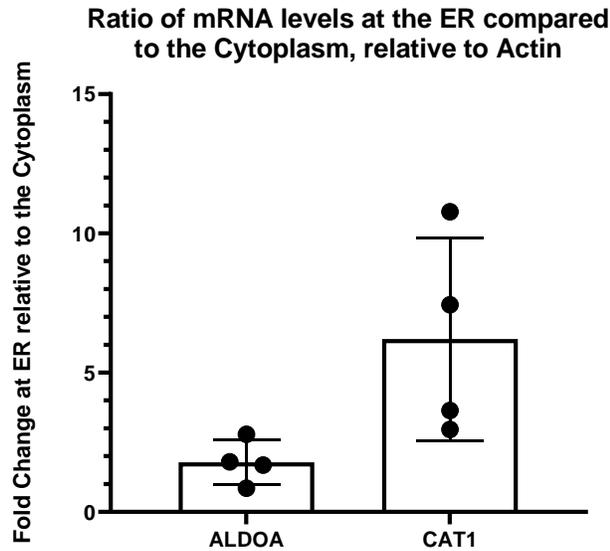


Figure 4.1: CAT1 is more strongly associated to the ER than ALDOA. Preliminary work performed by Dr Catherine Jopling, where Huh7 cells were subjected to membrane fractionation to isolate cytoplasmic and endoplasmic reticulum (ER)-rich fractions. Tri Reagent was used to extract RNA and qRT-PCR was performed to determine ALDOA and CAT1 expression levels. Data normalised to the expression of Actin and presented as mRNA expression in membrane fractionations relative to its relative expression in the cytoplasmic fractions. Bar height represents the average of four independent experiments, and error bars represent standard deviation.

In addition to ALDOA, G6PC3 (subunit of G6Pase) expression increases with miR-122 downregulation in hypoxic conditions which led to its subsequent identification as a miR-122 target [432, 433]. Glucose-6-phosphatase (G6Pase) is a multicomponent system located at the ER that catalyses the final step of glycogenolysis [490, 491], therefore the mRNA encoding the transmembrane G6PC3 is expected to be localised at the ER.

In this chapter, miR-122 regulation of targets ALDOA, CAT1, and G6PC3 was compared in total cell lysates and in ER-rich and cytoplasmic fractions. Levels of miR-122 were modulated utilising both 2'Ome and LNA antisense oligonucleotides for miR-122 inhibition and an LNA mimic for miR-122 overexpression.

4.2 miR-122 target expression in cytoplasmic and membrane-rich fractions following miR-122 Inhibition with a 2'Ome Oligonucleotide

Review of the literature resulted in the selection of ALDOA, CAT1, and G6PC3 as endogenous miR-122 targets to compare regulation at the ER and cytoplasm. miRDB is a database of miRNA-mRNA target interactions which incorporates predicted targets/binding sites from the bioinformatic tool MirTarget and functional annotations from literature mining [492, 493] and was used to confirm miR-122 targeting of the chosen mRNA. Figure 4.2 provides a schematic of the mRNA targets (ALDOA, CAT1, G6PC3) with their coding regions, 5' UTRs and 3' UTRs identified, along with the approximate location of miR-122 binding sites in their 3' UTRs.

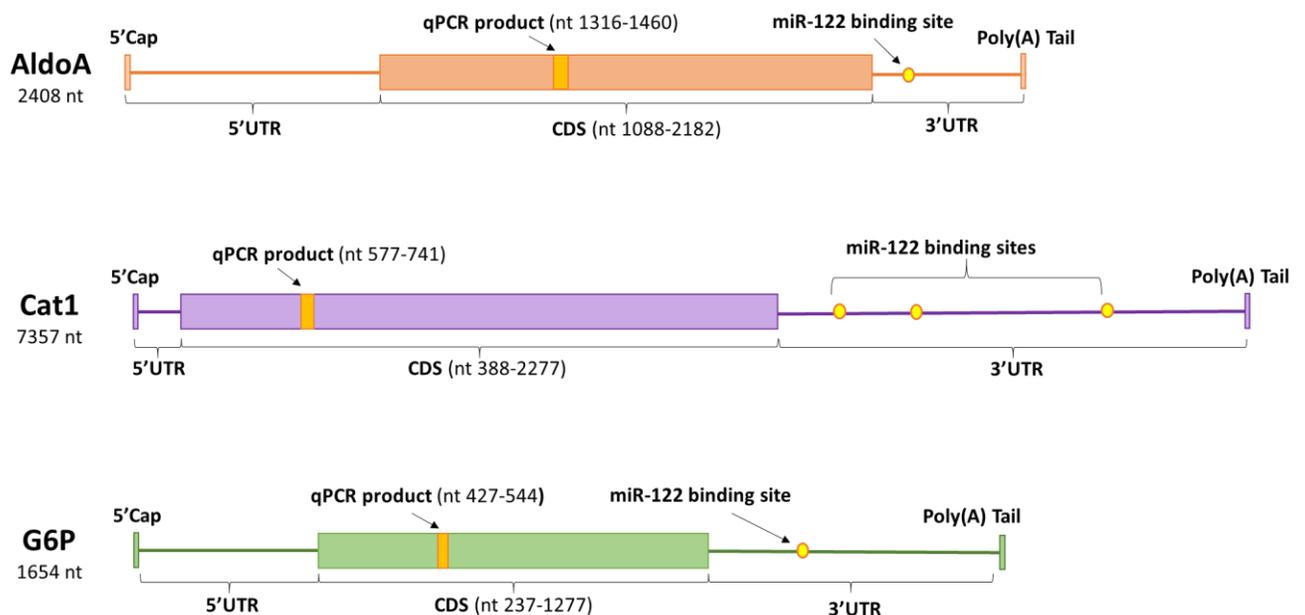
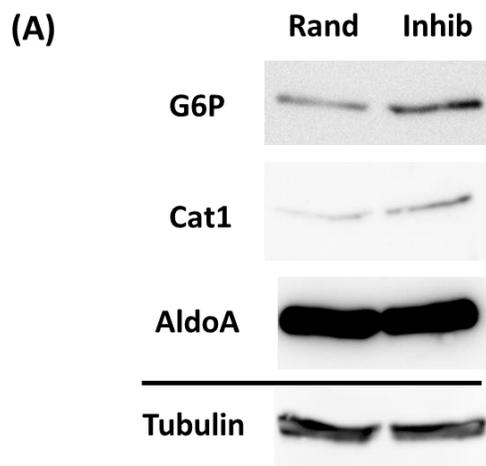


Figure 4.2: Schematic of miR-122 targets Aldolase A (ALDOA), Cationic Amino Acid Transporter 1 (CAT1), and Glucose-6-Phosphatase Catalytic Subunit 3 (G6PC3). Three mRNA known to be regulated by miR-122, with their coding regions, 5' UTRs and 3' UTRs identified, along with the approximate location of miR-122 binding sites in their 3' UTRs (yellow circle), and the location and size of the qRT-PCR product using primers found in Table 2.3 (yellow box).

To allow investigation of miR-mediated repression following miR-122 inhibition, Huh7 cells were transfected with either an antisense 2' O-methylated oligonucleotide (2'Ome) which sequesters miR-122 and has previously been shown to successfully inhibit miR-122 activity, or a randomised control 2'-O-methylated oligomer, thus allowing the levels of active endogenous miR-122 to be controlled [314].

First, western blots were performed to determine whether the inhibition of miR-122 had the expected effect of increasing protein expression from the chosen mRNA targets in Huh7 cells. Total protein was extracted 48 hours post transfection and analysed by western blot following the protocol described in Section 2.10.4 with antibodies against ALDOA, CAT1 and G6PC3, with tubulin as a protein loading control (Antibodies found in Table 2.6). The full gel images can be found in Supplementary Figure 10, with the bands corresponding to each protein identified in Figure 4.3A.

Densitometric analysis of the bands was performed and upon normalisation to the loading control tubulin, confirms an increase of G6PC3 and CAT1 protein at 2.53- and 3.48-fold respectively upon inhibition of miR-122 with a 2'Ome antisense oligonucleotide to miR-122 relative to the random 2'Ome control, Figure 4.3B. ALDOA protein levels only showed a slight increase on miR-122 inhibition, however the signal was strong and densitometric analysis is a very crude method of quantification so it may be that any change upon miR-122 inhibition is not detectable. Nevertheless, ALDOA is a well-documented target of miR-122 and the western blot confirms both G6PC3 and CAT1 as miR-122 targets in Huh7 cells so further analysis included all three target mRNAs.



(B)

| mRNA Target | Relative Density |
|-------------|------------------|
| G6P | 2.53 |
| Cat1 | 3.48 |
| AldoA | 1.10 |

Figure 4.3: Protein levels of miR-122 mRNA targets following inhibition of miR-122 with a 2'Ome Antisense Oligonucleotide. Total protein was extracted at 48 hours and **(A)** analysed by western blot with antibodies against ALDOA, CAT1 and G6PC3, with tubulin as a protein loading control. **(B)** Densitometry analysis of western blot signal was performed presenting the density of each protein, normalised to tubulin, following inhibition relative to their random control.

Next, to determine whether miR-122 inhibition had different effects on the level of the three mRNA targets in the ER and cytoplasm, Huh7 cells were transfected with either a miR-122 2'Ome or a random control and subjected to membrane fractionation. qRT-PCR was used to detect the expression levels of miR-122 targets ALDOA, CAT1 and G6PC3 in 100 ng of RNA extracted from both the cytoplasmic and membrane-rich fractions.

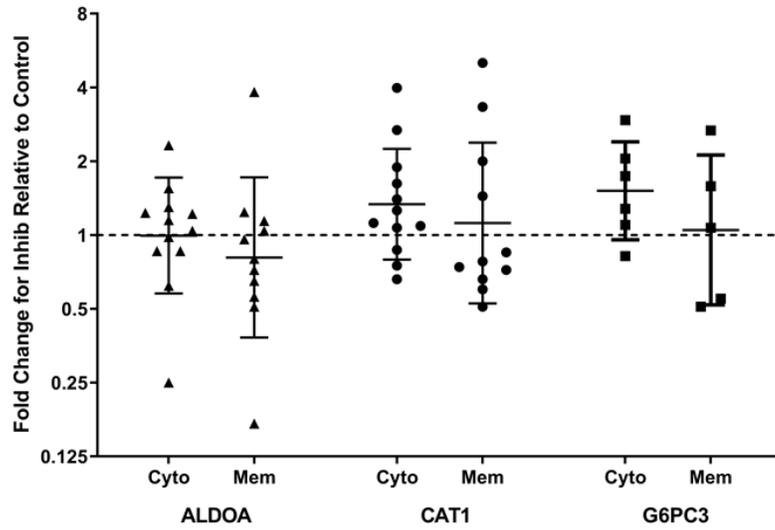
qRT-PCR primers were designed to amplify in the coding region of each mRNA and can be found in Table 2.3. Figure 4.2 provides a schematic of the mRNA targets (ALDOA, CAT1, G6PC3) with the approximate location the qRT-PCR product. For qRT-PCR, whilst there are several housekeeping

genes commonly used for normalisation that are unaffected by treatment/condition there have not been any identified that are consistently expressed across subcellular sites. As a result, the Ct values were normalised to either the housekeeping mRNA actin or a synthetic *in vitro* transcribed (IVT) RNA of known concentration that was spiked-in at the RT step.

The preliminary analysis of ALDOA and CAT1 expression at the ER and cytoplasm used Actin as the control as both are more enriched in the ER-fractions versus the cytoplasmic fractions than Actin. When normalised to the expression of actin in Figure 4.4A, there is no change upon 2'Ome treatment in ALDOA, CAT1, or G6PC3 expression with no statistical difference in response to miR-122 inhibition between the two fractions.

As normalisation to Actin resulted in variable data and could introduce a bias between the cytoplasmic and membrane fractions, IVT RNA was also used as a normaliser, with a known quantity added to each sample prior to the RT step. When normalised to the IVT spike-in in Figure 4.4B, all three miR-122 targets show a slightly increased mRNA level in the cytoplasm but a decrease in mRNA level in the membrane fraction following inhibition of miR-122 with a 2'Ome relative to the transfection control. T-tests show this is a significant difference in regulation of all three mRNAs at the membrane versus cytoplasm ($p < 0.05$). Additionally, 2-way ANOVA of the qRT-PCR data normalised to IVT RNA shows a statistical difference between cytoplasmic and ER-rich fractions ($p = 0.0032$) but no difference in regulation between the three mRNA. Overall, normalisation to IVT RNA suggests that inhibition of miR-122 results in a decrease in the level of each mRNA in the ER but does not have an effect on or slightly increases the mRNA levels at the cytoplasm.

(A) Expression of miR-122 mRNA targets for 2'Ome miR-122 Inhibitor transfection relative to the random control, normalised to Actin



(B) Expression of miR-122 mRNA targets for 2'Ome miR-122 Inhibitor transfection relative to the random control, normalised an IVT RNA spike-in

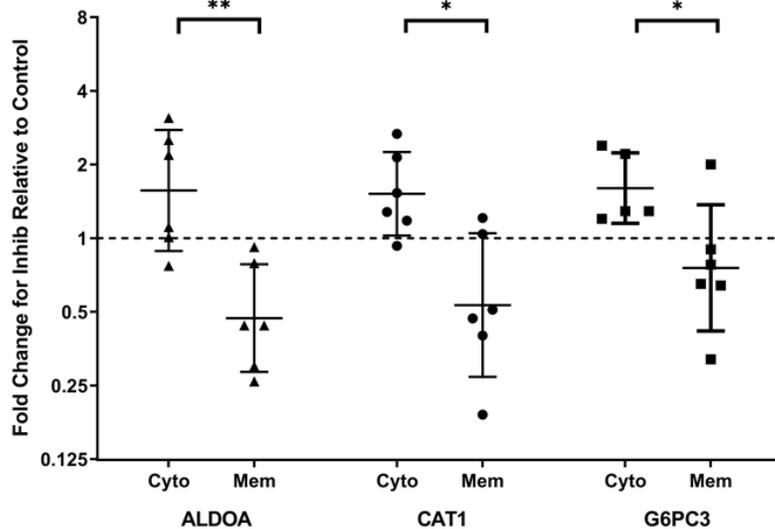


Figure 4.4: Expression of miR-122 mRNA targets following inhibition of miR-122 with a 2'Ome oligonucleotide. Huh7 cells subjected to membrane fractionation to isolate cytoplasmic (Cyto) and endoplasmic reticulum (ER)-rich (Mem) fractions following transfection with either a 2'Ome Oligonucleotide targeting miR-122 or a random control. qRT-PCR for these samples was used to detect the expression levels of miR-122 mRNA targets ALDOA, CAT1, and G6PC3, presented as fold change for 2'Ome relative to the control, and were normalised to either **(A)** Actin, or **(B)** an IVT RNA spike-in. Student's T-test of mRNA levels normalised to the IVT RNA **(B)** determines a significant difference between cytoplasmic and ER rich fractions (* $p < 0.05$, ** $p < 0.01$). Graph represents data from at least five independent experiments plotted on a logarithmic scale, with error bars representing the geometric SD.

Since the digitonin-lysis fractionation method used is fairly crude, to address possible differences in the cleanness of fractionation between biological repeats the ratio of the mRNA target expression in the membrane relative to the respective cytoplasmic fraction was also determined.

The fold change following miR-122 inhibition relative to the control 2'Ome presented in Figure 4.4 is presented in Figure 4.5 as the ratio of expression in the membrane versus the respective cytoplasmic fraction for each mRNA. When normalised to the IVT RNA all three mRNA (ALDOA, CAT1, G6PC3) were reduced following miR-122 inhibition in the membrane relative to cytoplasmic fraction in the membrane fractions averaging 0.30, 0.35, and 0.33 respectively, confirming the conclusions gained from Figure 4.4.

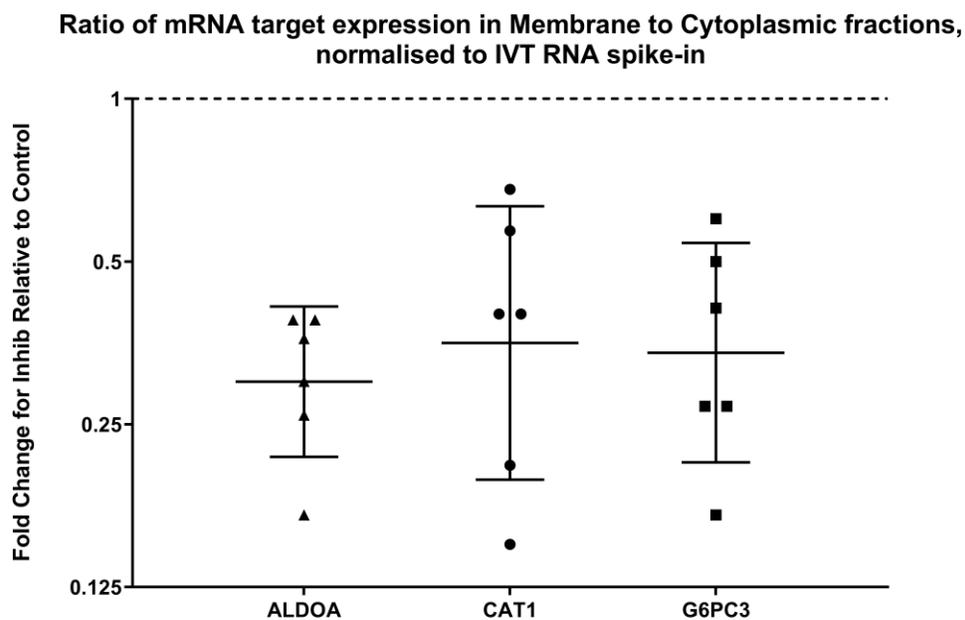


Figure 4.5: Ratio of miR-122 target expression in cytoplasmic and membrane-rich fractions following miR-122 inhibition with a 2'Ome Oligonucleotide. qRT-PCR analysis of miR-122 targets, normalised to (A) IVT RNA spike-in, calculated as a fold change for 2'Ome relative to the control and presented as the ratio of expression in the membrane to cytoplasmic fractions. Graph represents data from at 6 independent experiments plotted on a logarithmic scale, with error bars representing the geometric SD.

In conclusion, actin normalisation introduces more variability between experiments either due to bias from the differential expression of Actin, inconsistent fractionation, or that actin mRNA levels are affected, directly or indirectly, by miR-122 inhibition. Normalisation to an *in vitro* transcribed RNA spiked-in at the RT stage shows all three mRNA (ALDOA, CAT1, G6PC3) are reduced following miR-122 inhibition in the membrane relative to cytoplasmic fraction. When the ratio of fold change in membrane to cytoplasmic fractions was calculated to take into account variability in fractionation, again all three mRNA showed decreased expression in the membrane fractions versus the cytoplasmic fractions. Overall, normalisation to IVT RNA confirms that some miR-122 regulation of these targets is occurring at the mRNA level with differences seen at the membrane compared to cytoplasm.

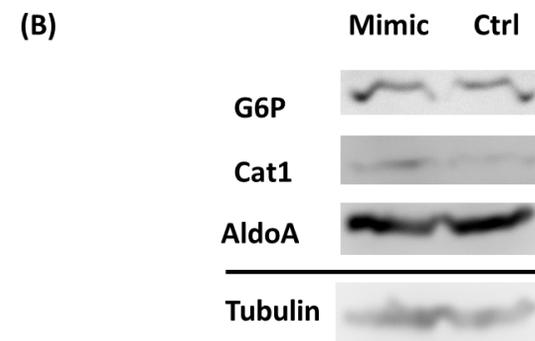
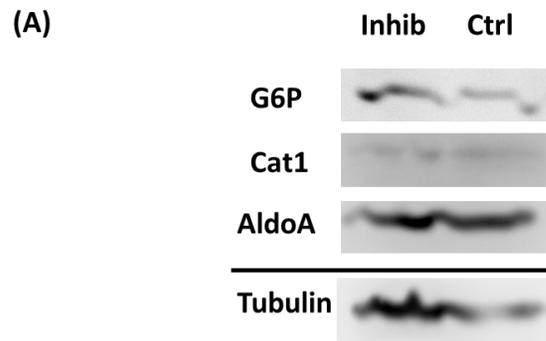
4.3 miR-122 target expression in total cell lysate following miR-122 Inhibition and Overexpression with miRCURY LNA oligonucleotides

Whilst there were significant differences between the ER and cytoplasm in expression of ALDOA, CAT1, and G6PC3 following inhibition with miR-122 2'Ome oligonucleotides, to determine if this was a genuine biological effect and hoping to reduce variability, the transfections were repeated with miR-122 inhibition achieved using a miRCURY LNA Inhibitor specific to miR-122. MiRCURY LNA Inhibitors boast a more potent response than traditional 2'Ome inhibitors by sequestering the target miRNA in highly stable heteroduplexes to achieve greater effectiveness and specificity. The corresponding Negative Control A is designed to avoid matches of >70% homology to any sequence in any organism in the NCBI and miRBase databases. In addition, to investigate the effect of miR-122 overexpression, a miRCURY LNA mimic was used with a corresponding Negative Control with no known homology in human. The LNA mimic is devised as an unmodified guide strand matching the miR-122-5p sequence, and two LNA-modified RNA strands complementary to miR-122-3p sequence but does not elicit any activity. To confirm an effect of miR-122 inhibition and overexpression with LNA oligonucleotides on the levels of ALDOA, CAT1 and G6PC3, qRT-PCR and western blots were performed first on total cell lysates.

Assisted by Dr Angela Downie, western blots were performed with antibodies for ALDOA, CAT1, and G6PC3 on total protein extracted from Huh7 cells transfected with either 10 nM miR-122 LNA inhibitor, 2 nM LNA mimic or their respective control. The concentration of reporters was chosen based on other work in the Jopling research group and the use of the same LNA oligonucleotides in Chapter 3. Full gel images can be found in the appendix, Supplementary Figures 11 and 12.

The probing for tubulin showed that the samples were unevenly loaded for the inhibition experiments, however this was corrected for in the densitometric analysis, Figure 4.6. It was determined that the levels of ALDOA, CAT1 and G6PC3 were increased 48 hours after transfection with the miR-122 LNA inhibitor at 2.35-, 3.27- and 5.32-fold higher density relative to the control, respectively. This reinforces the regulation of CAT1 and G6PC3 that was seen with the western blot in Section 4.2 but also confirms ALDOA is regulated by miR-122 as previously the inhibition of miR-122 with a 2'Ome inhibitor did not show a relief in repression of ALDOA. Furthermore, although the images were of poor quality and of a singular experiment, together with the mRNA and published data it is reasonable to conclude that 10 nM LNA Inhibitor relieved miR-122 repression of these targets.

On the other hand, overexpression of miR-122 with an LNA mimic was shown to have no effect on the protein levels with the relative density of the three targets remaining close to the control condition, Figure 4.6 B and D. As endogenous miR-122 is highly expressed in Huh7 cells, it may be that there is no detectable difference resulting from the LNA mimic as the miR-122 binding sites of these mRNAs is saturated by endogenous miR-122. However, whilst densitometric analysis was applied to account for the uneven loading of samples from the inhibition experiment this is just a crude method of quantification of western blots, and quantification may not be accurate based on one single experiment.



(C)

| mRNA Target | Relative Density: Inhib/Control |
|-------------|---------------------------------|
| G6P | 5.32 |
| Cat1 | 3.27 |
| AldoA | 2.35 |

(D)

| mRNA Target | Relative Density: Mimic/Control |
|-------------|---------------------------------|
| G6P | 1.37 |
| Cat1 | 1.25 |
| AldoA | 1.08 |

Figure 4.6: Protein levels of miR-122 mRNA targets following transfection with a miR-122 LNA Inhibitor or Mimic. Total protein was extracted at 48 hours following **(A)** miR-122 inhibition (Inhib), **(B)** overexpression (Mimic) or respective control (Ctrl) and analysed by western blot with antibodies against ALDOA, CAT1 and G6PC3, with tubulin as a protein loading control. Densitometry analysis of western blot signal was performed presenting the density of each protein, normalised to tubulin, for **(B)** miR-122 inhibition relative to their control and **(C)** miR-122 mimic relative to their control.

In order to select appropriate post-transfection conditions and data normalisation strategy, the effect of miR-122 inhibition and overexpression by LNA oligonucleotides on the level of each mRNA was initially studied in total cell lysates. Huh7 cells were transfected with 10 nM LNA inhibitor, and total RNA was harvested at 24- and 48-hour timepoints. As total RNA was extracted from the cells rather than fractionated samples, there was no need to include the IVT RNA spike-in as traditional housekeeping mRNA GAPDH, Actin and 18S rRNA are expected to be expressed at consistent levels across the treatment conditions and suitable for normalisation.

Figure 4.7 shows the level of each mRNA (ALDOA, CAT1, GCPC3) following miR-122 inhibition relative to the control condition, normalised separately to GAPDH, Actin and 18S rRNA. When normalised to GAPDH, no consistent increase in mRNA levels was observed upon miR-122 inhibition of any of the three mRNA targets at 24 hours. Some experiments resulted in an increase in mRNA levels at 48 hours above a 1.5-fold change whilst some experiments resulted in mRNA levels close to that of the control condition. In a similar pattern to GAPDH, when normalised to Actin, ALDOA shows a modest increase in mRNA levels at 48 hours and of the four independent experiments only one repeat for both CAT1 and G6PC3 increased in expression at 48 hours. Only normalisation to 18S rRNA at 48 hours shows the relief of expression expected when inhibiting miR-122 but the fold change expression is a lot more variable than the data normalised to Actin or GAPDH. All three methods of normalisation suggest that 24 hours is not sufficient to see an effect on mRNA level of the three targets. Overall, ALDOA and G6PC3 consistently increase in expression at 48 hours versus 24 hours regardless of normalisation strategy whereas CAT1 shows a similar trend but a more modest response. It may be possible that the overall effect of miR-122 is mainly at the level of translation rather than mRNA stability as there were strong effects seen on protein levels with both the 2'Ome and LNA inhibitor.

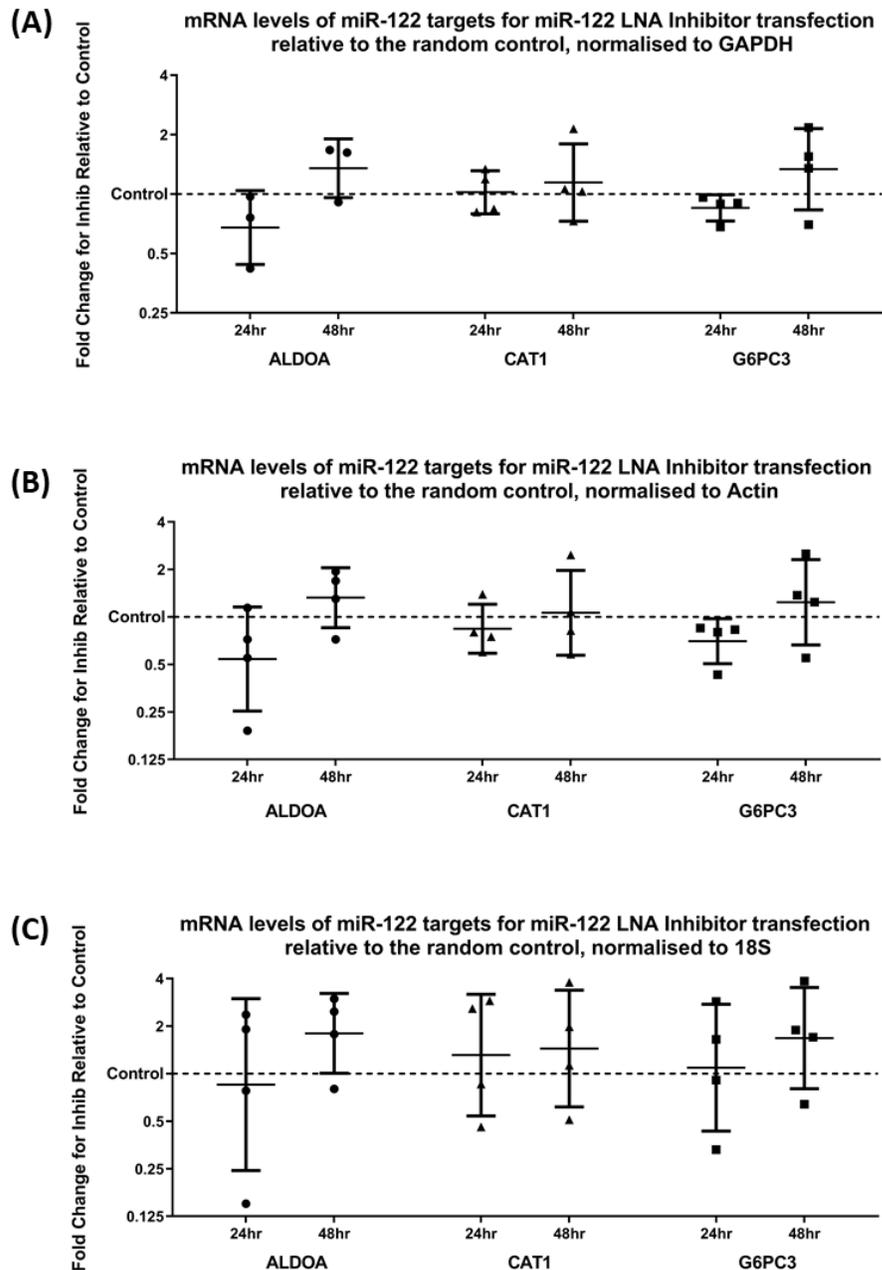


Figure 4.7: Expression of miR-122 mRNA targets at 24 and 48 hours following transfection with an LNA miR-122 inhibitor. Total RNA extracted from Huh7 cells following transfection with either a LNA miR-122 inhibitor or negative control. qRT-PCR was used to detect the expression levels of miR-122 mRNA targets ALDOA, CAT1, and G6PC3, presented as fold change for inhibition relative to the control, and were normalised to either **(A)** and GAPDH **(B)** Actin, or **(C)** 18S. 2-way ANOVA of expression determines no statistically significant effect for time nor target when normalising to Actin or 18S, but that there is a difference between 24 and 48 hours for the expression of the mRNA targets when normalised to GAPDH, $p=0.037$. Data represents four independent experiments plotted on a logarithmic scale, with error bars representing the geometric SD.

To confirm an effect of miR-122 overexpression on the levels of ALDOA, CAT1 and G6PC3, the same conditions were followed as with the LNA inhibitor, but a 2 nM miR-122 mimic was co-transfected instead. Again, normalisation to 18S rRNA introduces the most variability but does show a decrease in mRNA levels of CAT1 and G6PC3 at 24 hours, but with ALDOA levels less affected by miR-122 overexpression. Normalisation to Actin seems to show a reduction in ALDOA at 48 hours and G6PC3 at both 24 and 48 hours but no consistent effect on CAT1. As with normalisation to 18S, when normalised to the expression of GAPDH, CAT1 and G6PC3 are repressed at 24 hours but ALDOA is less affected. Overall, G6PC3 expression shows the most consistent repression with miR-122 overexpression for all three normalisation methods but with the most variability. Across all three graphs in Figure 4.8, the mRNA levels of CAT1 and ALDOA at 24 hours suggest some repression, but that there may be some relief of this at 48 hours.

In this section, western blots of protein extracted at 48 hours (Figure 4.6) show a strong increase in expression of the miR-122 targets ALDOA and G6PC3 following miR-122 inhibition but did not show an effect of miR-122 overexpression with the protein levels of the three miR-122 targets remaining close to the control condition. However, the effects of miR-122 inhibition and overexpression on mRNA levels were more variable. Although qRT-PCR analysis of the three targets shows some increase in mRNA levels at 48 hours when normalised to GAPDH, whether inhibition of miR-122 with LNA oligonucleotides affects total mRNA levels remains inconclusive. Following overexpression of miR-122 with the LNA mimic at 2 nM, there was some repression seen for G6PC3 and CAT1, but remained inconclusive for ALDOA due to variability between experiments.

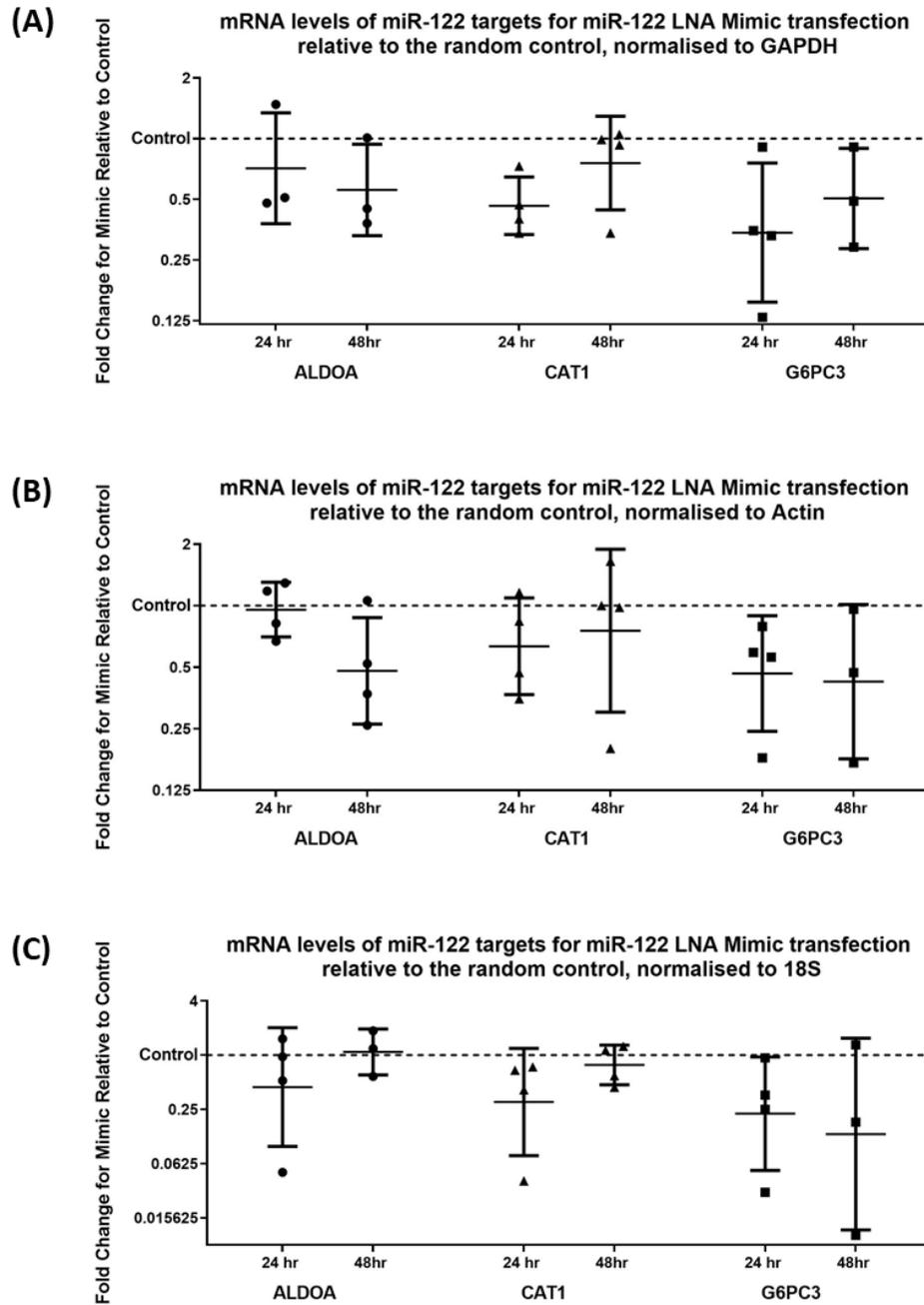


Figure 4.8: Expression of miR-122 mRNA targets at 24 and 48 hours following overexpression of miR-122 with an LNA Mimic. Total RNA extracted from Huh7 cells following transfection with either a LNA miR-122 mimic or negative control. qRT-PCR was used to detect the expression levels of miR-122 mRNA targets ALDOA, CAT1, and G6PC3, presented as fold change for mimic relative to the control, and were normalised to either **(A)** and GAPDH **(B)** Actin, or **(C)** 18S. 2-way ANOVA of expression determines no statistically significant effect for time nor target for all three methods of normalised. Data represents four independent experiments plotted on a logarithmic scale, and error bars represent geometric SD.

4.4 miR-122 target expression in cytoplasmic and ER-rich fractions following miR-122 Inhibition and Overexpression

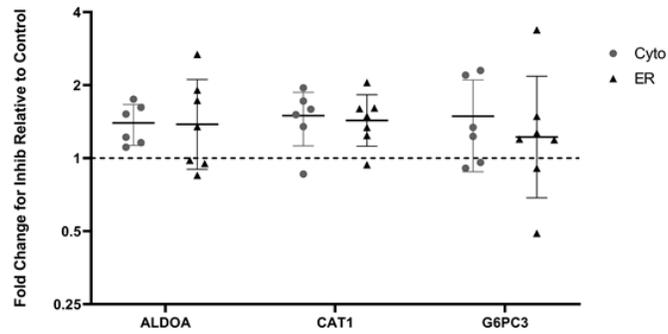
To determine whether inhibition of miR-122 by a different molecule showed the same effects as the 2'Ome inhibitor, modulation of endogenous miR-122 levels with LNA oligonucleotides was performed which also enabled the use of more normalisers. Based on the qRT-PCR and western blot data from total cell lysates, the same LNA inhibitor and mimic were transfected into Huh7 cells at 10 nM and 2 nM which were then subjected to membrane fractionation to isolate cytoplasmic and ER-rich fractions. Harvesting of RNA took place 48 hours post transfection as there was not much effect on the expression of the miR-122 targets at 24 hours in the total cell lysates in Section 4.3. Unfortunately, the Reliaprep kit was not compatible with the buffers used in the membrane fractionation and so Tri Reagent was used for RNA extraction instead.

Building on difficulties in normalising the data in earlier data from the 2'Ome treatment, more normalisation strategies were introduced in these experiments. For the qRT-PCR, Actin was discounted a normaliser due to its variability in Section 4.2, and although GAPDH is also predominantly localised in the cytoplasm it provided the most consistent results in Section 4.3 with the qRT-PCR data from total cell lysates. 18S rRNA was chosen as an additional normaliser as it is a core ribosomal component, present in both the cytoplasm and ER. However, the amount present in each fraction may be variable, therefore IVT RNA was also used as a known quantity was added to each sample. Unfortunately, this only takes into account variability in the RT reaction and cannot account for variability introduced from the fractionation or RNA extraction method. As a result, the qRT-PCR data is normalised to multiple reference RNA.

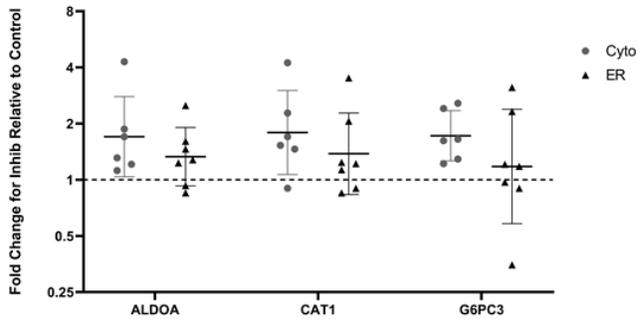
Figure 4.9 shows the expression of ALDOA, CAT1, and G6PC3 after miR-122 inhibition relative to the control in both cytoplasmic and ER-associated fractions. When normalised to GAPDH, all three miR-122 targets show an increased expression in both the cytoplasm and ER averaging to roughly the same level, but with more variability in the membrane fractions. Normalisation to 18S rRNA results in more variable data but again all three targets show an increased expression upon miR-122 inhibition. The increase in expression is greater in the cytoplasmic fractions but due to the variability between experiments, the difference is not statistically significant when T-tests were performed. Normalisation to the IVT spike-in resulted in a lot of variability between experiments making it difficult to draw a firm conclusion about the effect of miR-122 overexpression on any of the targets. Statistical analysis of each normalisation method determined no significant difference in regulation between ER and cytoplasmic fractions.

In contrast to Figure 4.9 where the expression in both the ER and cytoplasmic fractions were presented separately, in Figure 4.10 the ratio of the mRNA target expression in the membrane relative to the respective cytoplasmic fraction was also determined to account for variability introduced by the fractionation method. When normalised to GAPDH, most experiments resulted in less than 2-fold increase in expression of all three targets, similar to what was seen in Section 4.2. Normalisation to 18S rRNA shows a very similar trend to GAPDH but with less variability, with all three showing potentially a lower expression in the ER-associated fraction than cytoplasm. Finally, when normalised to the IVT spike-in, the ratio of mRNA levels at the membrane to cytoplasm were very variable between experiments, reiterating what was seen in Figure 4.9.

(A) mRNA levels of miR-122 targets at the Cytoplasm and ER for miR-122 LNA Inhibitor transfection relative to control, normalised to Gapdh



(B) mRNA levels of miR-122 targets at the Cytoplasm and ER for miR-122 LNA Inhibitor transfection relative to control, normalised to 18S



(C) mRNA levels of miR-122 targets at the Cytoplasm and ER for miR-122 LNA Inhibitor transfection relative to control, normalised to IVT RNA

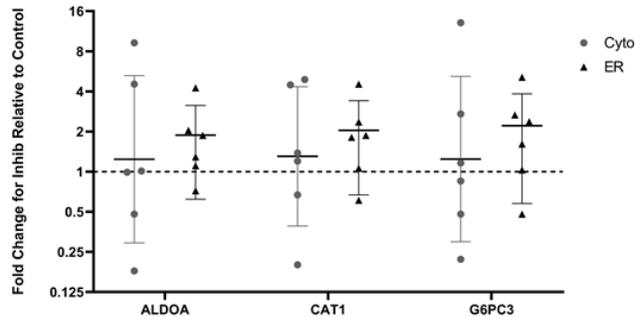


Figure 4.9: Expression of miR-122 mRNA targets following transfection with an LNA miR-122 Inhibitor. Huh7 cells subjected to membrane fractionation to isolate cytoplasmic (Cyto) and endoplasmic reticulum (ER)-rich fractions 48 hours after transfection with either a LNA Inhibitor targeting miR-122 or a random control. qRT-PCR for these samples was used to detect the expression levels of miR-122 mRNA targets ALDOA, CAT1, and G6PC3, presented as fold change for inhibition relative to the control, and were normalised to either **(A)** GAPDH **(B)** 18S or **(C)** an IVT spike-in. 2-way ANOVA of expression determines no statistically significant effect for target nor fraction for all three methods of normalised. Data represents at least five independent experiments plotted on a logarithmic scale, with error bars representing the geometric SD.

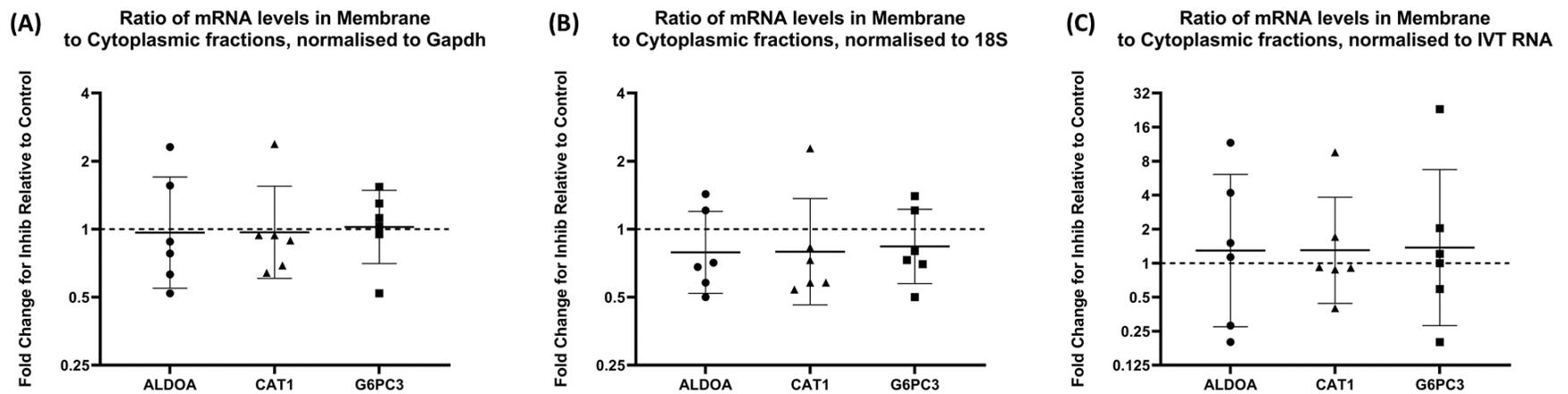
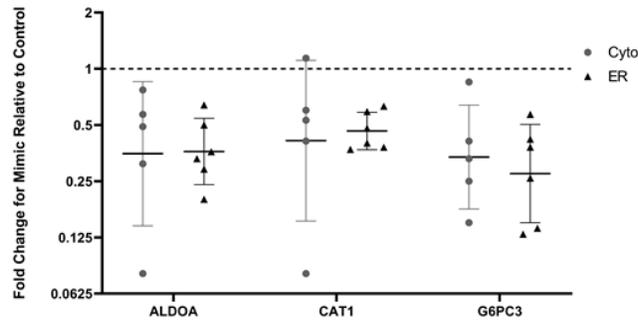


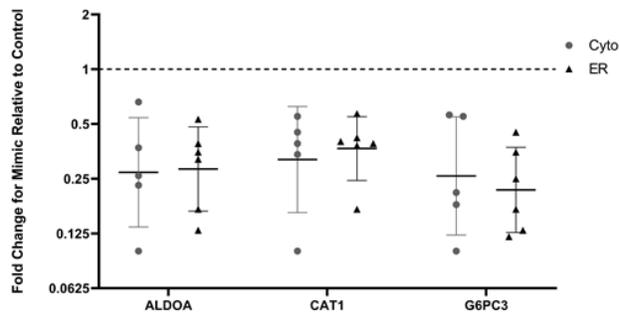
Figure 4.10: Ratio of miR-122 target expression in cytoplasmic and membrane-rich fractions following miR-122 inhibition. qRT-PCR analysis of miR-122 targets, normalised to **(A)** GAPDH **(B)** 18S or **(C)** a spike-in, calculated as a fold change for miR-122 inhibition relative to the control and presented as the ratio of expression in the membrane to cytoplasmic fractions. Data represents at least five independent experiments plotted on a logarithmic scale, with error bars representing the geometric SD..

To further investigate the effect of miR-122 overexpression on the levels of ALDOA, CAT1 and G6PC3 in membrane and cytoplasmic fractions, the same conditions were followed as with the LNA inhibitor but co-transfecting with a miR-122 mimic at 2 nM, again harvesting the RNA at 48 hours. In Figure 4.11, following miR-122 overexpression, when normalised to the level of GAPDH, there is an average repression of all three mRNA at more than 2-fold that of the control condition in both the ER and cytoplasm but with more variability in the cytoplasmic fractions. Normalisation to 18S rRNA resulted in a similar trend, with strong repression in both the cytoplasmic and ER-associated fractions but with stronger repression overall than normalisation to GAPDH. Despite normalisation to IVT RNA levels yielding more consistent results in Section 4.3, this method of normalisation resulted in the most variable data which averages to a more modest repression in the cytoplasmic fractions and much more variation in the membrane-rich fractions. Overall, there is no significant difference between the expression of the miR-122 targets following miR-122 overexpression with any normalisation method, with the differences seen likely to be due to variability in results.

(A) mRNA levels of miR-122 targets at the Cytoplasm and ER for miR-122 LNA Mimic transfection relative to control, normalised to Gapdh



(B) mRNA levels of miR-122 targets at the Cytoplasm and ER for miR-122 LNA Mimic transfection relative to control, normalised to 18S



(C) mRNA levels of miR-122 targets at the Cytoplasm and ER for miR-122 LNA Mimic transfection relative to control, normalised to IVT RNA

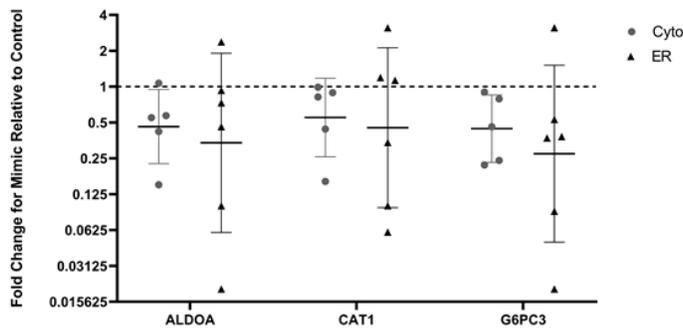


Figure 4.11: Expression of miR-122 mRNA targets following transfection with an LNA miR-122 Mimic. Huh7 cells subjected to membrane fractionation to isolate cytoplasmic (Cyto) and endoplasmic reticulum (ER)-rich fractions 48 hours after transfection with either a LNA Mimic of miR-122 or a random control. qRT-PCR for these samples was used to detect the expression levels of miR-122 mRNA targets ALDOA, CAT1, and G6PC3, presented as fold change for mimic relative to the control, and were normalised to either **(A)** GAPDH **(B)** 18S or **(C)** ICT spike-in. 2-way ANOVA of expression determines no statistically significant effect for target nor fraction for all three methods of normalised. Data represents at least five independent experiments plotted on a logarithmic scale, and error bars represent geometric SD.

To account for variability introduced by the fractionation method, the ratio of the mRNA target expression in the membrane relative to the respective cytoplasmic fraction was also determined. As can be seen in Figure 4.12, excluding the one clear outlier for each mRNA, miR-122 overexpression with an LNA mimic has little effect on the difference in expression between the two fractions for all three targets when normalised to GAPDH. Normalisation to the levels of 18S rRNA is a bit more variable but again showing very similar levels of expression between the two fractions. Finally, normalisation to IVT RNA once again resulted in the most variable data with a slight increase in expression at the ER versus the cytoplasm for the miR-122 targets. This data confirms that miR-122 overexpression has similar effects in both membrane and cytoplasmic fractions on the level mRNA for each target.

In this section, normalisation to GAPDH and 18S rRNA levels resulted in a variable increase in all three miR-122 targets following inhibition whilst the mRNA levels remained unaffected by inhibition when normalised to IVT RNA. When the ratios of expression at the membrane to cytoplasmic fractions were considered, there was no significant difference between the expression of any of the miR-122 targets at the cytoplasm or ER following miR-122 inhibition. Similarly, there was a strong reduction in all mRNAs following miR-122 overexpression at similar levels in the ER-associated and cytoplasmic fractions when normalised to the levels of GAPDH and 18S, with normalisation to IVT resulting in the most variable data. When the ratio of mRNA expression in the membrane to cytoplasmic fractions were calculated, there was no significant difference with overexpression of miR-122 having similar effect at both the ER and cytoplasm. Overall, there was no significant difference between the expression of the miR-122 targets following miR-122 overexpression with any normalisation method, with any differences likely to be due to variability in results.

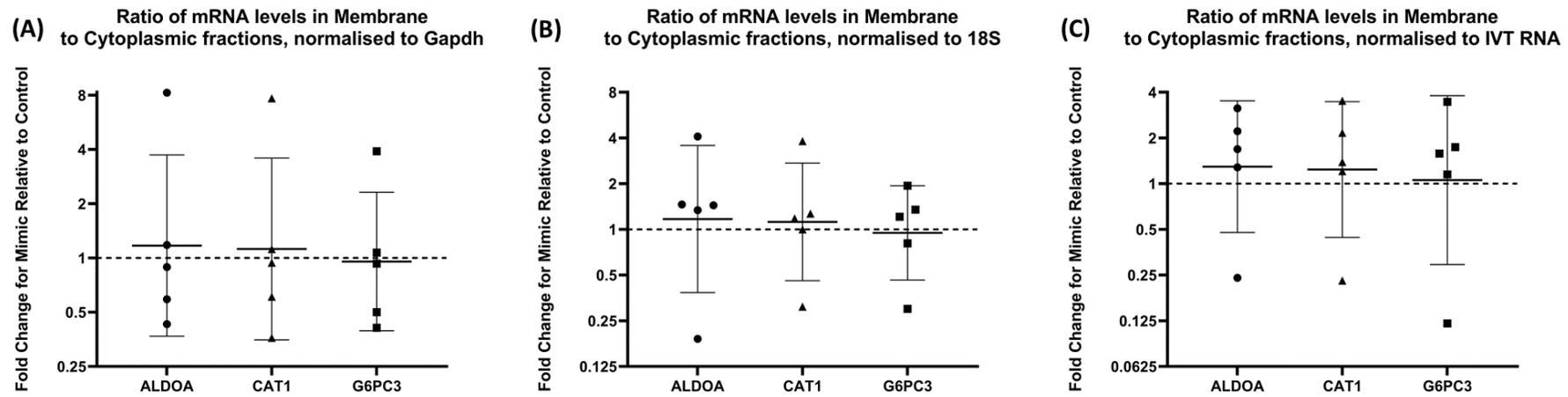


Figure 4.12: Ratio of miR-122 target expression in cytoplasmic and membrane-rich fractions following miR-122 overexpression. qRT-PCR analysis of miR-122 targets, normalised to **(A)** GAPDH **(B)** Actin or **(C)** a spike-in, calculated as a fold change for miR-122 mimic transfection relative to the control and presented as the ratio of expression in the membrane to cytoplasmic fractions. Data represents at least five independent experiments plotted on a logarithmic scale, and error bars represent geometric SD.

4.5 Association of miR-122 with Argonaute Proteins at the Endoplasmic Reticulum and Cytoplasm

Preliminary work performed by Dr Catherine Jopling involved immunoprecipitation with an AGO(1-4) antibody of the membrane and cytoplasmic fractions (Figure 4.13). It was found that miR-122 expression in the AGO-IP samples was higher in the ER compared to the cytoplasm, suggesting that miR-122 is more strongly associated with AGO at the ER. The same approach was used to look at RISC association with miR-122 and its target mRNAs in both the cytoplasmic and ER-rich fractions.

miR-122 expression in IgG- and AGO-IP samples

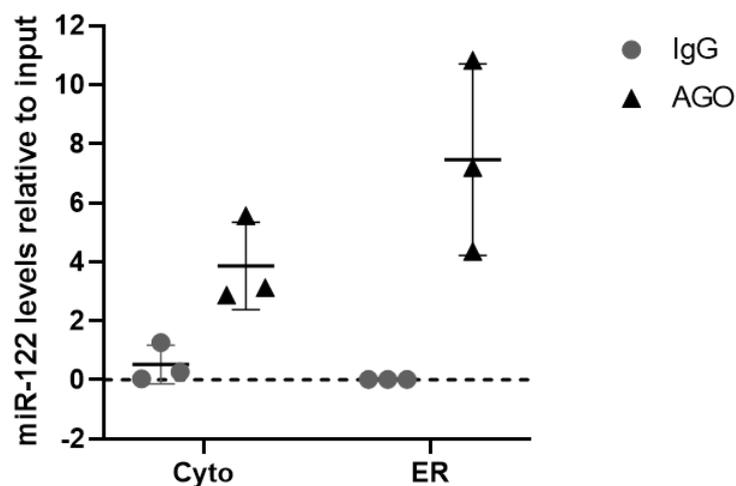


Figure 4.13: miR-122 association with AGO1-4 in membrane and cytoplasm fractions. Preliminary work performed by Dr Catherine Jopling, where Huh7 cells were subjected to membrane fractionation to isolate cytoplasmic and endoplasmic reticulum (ER)-rich fractions. 10% of these samples were used as “input” controls and the remainder of the lysates were split between immunoprecipitation using a control IgG antibody and AGO(1-4) antibody. Tri Reagent was then used to extract RNA and qRT-PCR was performed using the miR-122 TaqMan assay kit. Data presented as miR-122 expression in the IgG-IP (gray) and AGO-IP samples (black) relative to the total expression in the input samples. Data represents three independent experiments, and error bars represent standard deviation.

Huh7 cells were subjected to membrane fractionation to isolate cytoplasmic and endoplasmic reticulum (ER)-rich fractions with 10% of these samples used as “input” controls. The remaining lysates were used for immunoprecipitation (IP) using a control IgG antibody and AGO(1-4) antibody. Tri Reagent was used to extract RNA and qRT-PCR was performed using the miR-122 TaqMan assay kit as described in Section 2.6.2. Unfortunately, despite several attempts at optimisation of the immunoprecipitation protocol used in the preliminary work, miR-122 was persistently detected in the control IgG-IP. Table 4.1 shows the raw Ct values for miR-122 in the Input, IgG-IP, and AGO-IP samples from one representative experiment. Crosslinking with UV and formaldehyde was introduced without fractionating the cells in order to overcome the contamination of IPs with miR-122 but resulted in poor yields.

Table 4.1: miR-122 contamination of Input and IgG control samples. Tri Reagent was then used to extract RNA and qRT-PCR was performed using the miR-122 Taqman assay kit, presented is one example of such experiment.

| | | Average Ct |
|-------------|--------------|-------------------|
| Cyto | Input | 19.72 |
| | IgG | 18.62 |
| | Ago | 18.82 |
| ER | Input | 18.23 |
| | IgG | 19.08 |
| | Ago | 18.75 |

4.6 Discussion

In this chapter, a digitonin-lysis membrane fractionation protocol was used to isolate ER-rich fractions from cytosolic fractions allowing the comparison of miR-122 regulation at the ER and cytoplasm, following modulation of endogenous miR-122 levels using both 2'Ome and LNA antisense oligonucleotides for miR-122 inhibition and an LNA mimic for miR-122 overexpression.

Densitometric analysis of a western blot confirmed an increase of G6PC3 and CAT1 protein at 2.53- and 3.48-fold respectively upon inhibition of miR-122 with a 2'Ome antisense oligonucleotide to miR-122 relative to the random 2'Ome control but only a slight increase of ALDOA protein levels upon miR-122 inhibition. This difference may be due to stability and turnover of the protein, with ALDOA protein possibly conferring a longer half-life and as protein was harvested at a single time-point, a detectable change in protein level may occur at a later point. This could be investigated by pSILAC (pulsed stable isotope labelling by amino acid in cell culture) which labels newly synthesised proteins with heavy nitrogen, carbon, or hydrogen isotopes and analysis by mass spectrometry can quantify the protein synthesis and decay of the protein of interest. However, there was a strong increase in expression at the protein levels of all three miR-122 targets following miR-122 inhibition with an LNA inhibitor, confirming regulation by miR-122 of the targets at least in whole cell lysates. Although the western blots are of poor quality and single repeats, the similar result with both inhibitors supports the conclusion that miR-122 does repress all three mRNAs in Huh7.

In this chapter, inhibition of miR-122 with 2'Ome oligonucleotides resulted in an increase in ALDOA, CAT1 and G6PC3 mRNA levels in cytoplasmic fractions that coincided with a decrease in level in the membrane fractions. It may be that the changes seen in mRNA levels are not necessarily due to stabilisation/destabilisation by miR-122, but rather that miR-122 inhibition could lead to dissociation from the membrane resulting in the decreased expression in membrane fractions.

However, when miR-122 inhibition was achieved using an LNA inhibitor, normalisation to GAPDH and 18S rRNA levels resulted in a variable increase in all three miR-122 targets following inhibition with no difference seen in regulation of the targets between the membrane and cytoplasmic fractions. Although normalisation to IVT RNA provided the most consistent results in Section 4.3, it resulted in the most variability between experiments with the miR-122 LNA inhibitor in Section 4.4.

Despite different normalisation strategies, when compared with miR-122 inhibition with 2'Ome oligonucleotides there is a similar increase in expression of all three mRNAs in the cytoplasm with the miR-122 LNA inhibitor. However, where there was a difference in regulation between the cytoplasmic and ER-associated fractions with a decrease in expression seen at the ER following miR-122 inhibition with 2'Ome oligonucleotides, the use of LNA Inhibitors did not show any difference in regulation between the two sites. This difference in effect between inhibition achieved with 2'Ome and LNA oligonucleotides could be due to differences in normalisation methods and the general difficulties of normalising in fractionation experiments, or that the 2'Ome Inhibitor was more effective.

To investigate this difference, the effect of the 2'Ome Inhibitor on the total mRNA levels for the targets could be assessed. This was not performed during this project as there was a great deal of variability between experiments, and therefore the decision was made to move onto the more robust system of LNA oligonucleotides which should provide a more potent inhibition of miR-122 and allowed the investigation into the effects of overexpression in a parallel method. Furthermore, to determine if the decreased mRNA levels at the ER are due to dissociation, techniques such as *in situ* hybridisation could be used to visualise the localisation of mRNA. Alternatively, polysome analysis could be used to investigate whether the amount of mRNA associated with the translational machinery in each fraction is changing.

In addition, overexpression of miR-122 was achieved with an LNA mimic. Whilst there was no effect seen at the protein level with western blots, there was strong repression seen at 48 hours of the three mRNAs from total cell lysates. It may be that the LNA mimic affects the degradation of mRNA but that the proteins are very stable which would result in no effect of the mimic on the protein level of the targets, but this was only one repeat and so difficult to draw a conclusion. Following fractionation, miR-122 overexpression with the LNA mimic resulted in a strong reduction in all mRNAs to a similar extent in both the ER and cytoplasm, consistent with the similar regulation seen with miR-122 LNA inhibitor. This contrasts with the lack of effect with the LNA inhibitor and the difference between expression in the cytoplasmic and membrane fractionations seen with the 2'Ome inhibitor. Whilst 2 nM of the LNA mimic is sufficient to achieve strong repression of all three miR-122 targets it may not have same subcellular distribution as endogenous miR-122 and therefore may not replicate endogenous regulation at the ER and cytoplasm.

Poor reproducibility was a consistent problem which made it difficult to interpret the results and may account for some of the differences seen between miR-122 inhibition with 2'Ome and LNA oligonucleotides. There are several possible reasons for this including variability in efficiency of fractionation, variability in RNA extraction, and inability to find a suitable normalisation technique.

Whilst the digitonin-lysis fractionation method is compatible with a range of other techniques such as AGO-IP and polysome fractionation it remains a crude fractionation approach, separating intracellular membrane-bound organelles into one fraction. The mRNA levels in both the 2'Ome and LNA control transfection groups were compared for the three targets in both the cytoplasmic and ER-associated fractions (Supplementary Figures 13 and 14). Despite the preliminary work with ALDOA and CAT1 (Figure 4.1), in the 2'Ome experiments the Actin normalised mRNA levels did not show differential expression of the targets ALDOA and G6P and none of the three targets showed differential expression when normalised to IVT RNA across repeated

experiments. When looking at the Δ CT levels for the LNA control transfection conditions, only G6PC3 showed differential expression when normalised to both GAPDH and 18S, with a lower CT in the ER-associated fractions correlating to an increased expression in these fractions. However, when normalised to IVT RNA, it appears all three miR-122 targets show significantly higher expression in the ER-associated fractions compared to the cytoplasm which is unexpected for ALDOA. It may be that the membrane fractionation protocol used is not very efficient at isolating the ER from the cytoplasm, therefore Western blots would be used in future work to quantify ER and cytoplasmic marker proteins such as Calnexin and Tubulin, respectively, to verify clean fractionation. Attempts at alternative methods, such as the use of differential centrifugation, were made whilst performing the inhibition and overexpression with LNA oligonucleotides in total cell lysates but resulted in poor yield. With more time, the alternative methods for membrane fractionation could be optimised [494, 495].

In addition, the cytoplasmic fractions repeatedly yielded less RNA than the membrane fractions. Initially, Tri Reagent was used following fractionation in Section 4.2 as it typically results in a greater yield than kit-based approaches, combating the lower yields that result from fractionation compared to whole cell lysates. However, this method introduces experimenter variability in taking the aqueous phase therefore a Reliaprep kit was used in Section 4.3 when isolating RNA from total cell lysates which resulted in much more consistent extraction. Attempts were made to take the ER and cytoplasmic fractions directly to the Reliaprep kit, but the lysis buffers were not compatible. Despite several optimisation attempts, the resulting RNA yields were too poor to use in qRT-PCR and so Tri Reagent was used to extract RNA following membrane fractionation in Section 4.4.

Furthermore, despite normalising input amounts to 100 ng and input volume to 1 μ L for the RT step any difference in fractionation and RNA extraction will result in variability across experiments. The Nanodrop used to determine RNA concentration is affected by RNA quality and

primarily measures rRNA and does not control for variation inherent in the RT or PCR reactions [496]. Reference RNA are most commonly used to normalise qRT-PCR data as it controls for different RNA amounts input into the RT step, and in theory are consistent between experimental conditions/cell/tissue type, but none have been validated as consistent between subcellular compartments. Alternatively, an artificial RNA, such as the *in vitro* transcribed (IVT) RNA used in our experiments, incorporated into the RT step can account for variability in the RT and PCR reactions but cannot control for differences in fractionation or RNA extraction that normalising to endogenous reference RNA such as 18S rRNA or GAPDH can. As a result, the qRT-PCR data was normalised to multiple reference RNA but there was no singular reference RNA that resulted in consistent expression levels across experiments. For future work, it would be of interest to run a reference gene panel using ER-rich and cytoplasmic samples with the aim of identifying a selection of reference RNA that are consistently expressed between the ER and cytoplasm and are not affected by the inhibition or overexpression of miR-122.

Western blots quantifying ER protein markers would also determine if overall ER integrity of the cells is affected by the changes in miR-122 levels, resulting from the transfection of 2'Ome and LNA oligonucleotides, which may have an indirect effect on overall mRNA expression that are reflected in the qRT-PCR. Alternatively, a control ER mRNA could be included that does not possess miRNA binding sites and therefore would reflect changes in overall ER translation to confirm the changes in the target mRNA expression levels are a direct result of miR-122 regulation.

With more time, optimisation of the AGO-IP protocol would be of interest to investigate the association of miR-122 and the chosen targets with the argonaute proteins in the ER-associated and cytoplasmic fractions. Preliminary work found a stronger association of AGO(1-4) with miR-122 at the ER but attempts to follow up were hampered by persistent contamination of miR-122 in the input and control IgG-IP samples despite

extensive efforts to eliminate this; new reagents were ordered and seals broken on fresh nuclease-free water, DNAaway™ (ThermoFisher) was used on pipettes, pipettes/tubes/tips were UV-irradiated and qRT-PCR was performed with filter tips in the airflow hood.

In this chapter, the regulation of endogenous miR-122 targets that are known to associate to different subcellular sites were investigated using a membrane fractionation method to isolate ER and cytoplasm-localised mRNAs. The effects of miR-122 inhibition and overexpression on known miR-122 targets was compared in these fractions and an increase was observed in the mRNA levels in the cytoplasmic fractions coinciding with a decrease in membrane fractions upon inhibition with a 2'Ome oligonucleotide. Optimisation of the membrane fractionation protocol and normalisation strategy is required to make firm conclusions about the regulation of miR-122 targets at the ER. Future work could include additional targets as only three miR-122 targets were investigated in this project or use of techniques such as proximity-specific ribosome profiling and APEX-seq, which would allow the investigation into global regulation of miR-122 targets at the ER than focusing on individual miR-122 targets.

Chapter Five:

Investigation of the location and function of miR-122 within the nucleus

5.1 Introduction

Previous work by Dr Andrew Lewis in the Jopling lab applied microRNA *in situ* Hybridisation (ISH) to investigate the subcellular localisation of miR-122 within Huh7 cells [311]. This showed cytoplasmic and perinuclear staining, accompanied by some small nuclear foci, that corresponded to miR-122 suggesting a subset of miR-122 may be present in the nucleus within Huh7 cells. Although microRNAs are primarily thought to function in the cytoplasm, the presence of miR-122 within the nucleus is intriguing as a few other microRNA and RISC factors have previously been identified in the nucleus [231, 321, 347, 365-368]. Of note, mature miR-21 was found in both the nuclear and cytoplasmic fractions of HeLa cells, where 20% of total miR-21 was localised to the nucleus [231]. Likewise, during the course of this work, Wang *et al* found that miR-122 is highly expressed in the nucleus of Huh7 cells and directly silences miR-21 by binding to the pri-miR-21 transcript and preventing processing by the Drosha-DGCR8 microprocessor [379].

Understanding whether nuclear miR-122 associates with chromatin gives an insight into whether it regulates nascent RNA. It is possible to biochemically separate chromatin-associated RNA from nucleoplasmic RNA; Wuarin and Schibler [497] were able to isolate RNA and proteins from the chromatin using urea and a non-ionic detergent, and Pandya-Jones and Black [498] optimised this by first separating the nuclei from the cytoplasmic fraction to obtain distinct chromatin-associated and nucleoplasmic fractions. Determining whether mature miR-122 is chromatin-associated would give insight into whether the microRNA can bind to nascent transcripts and therefore regulate the many RNA processing events that occur co-transcriptionally, including microRNA biogenesis. Other studies have suggested a wide range of roles for nuclear-localised microRNA including regulation of the non-coding transcriptome, directing alternative splicing, regulating the biogenesis of another microRNA, influencing poly-adenylation site usage, and as epigenetic regulators [373-376].

In addition, there has been one example of auto-regulation by a microRNA; in *C.elegans*, ALG-1, a Argonaute homolog, promotes downstream processing of the let-7 microRNA primary transcripts, in a mechanism mediated by mature let-7 binding to a conserved complementary site in its own primary transcript [378]. Initial analysis of the pri-miR-122 gene identified a seed match for the mature miR-122 downstream of the pre-miR-122 hairpin, suggesting a similar mechanism could also act at the pri-miR-122 gene.

In this chapter, we further examined the localisation of miR-122 within the nucleus using a fractionation method adapted by the Proudfoot lab from the Wuarin and Schibler method [311, 470], and found the presence of miR-122 specifically in the chromatin fractions. Furthermore, to investigate the role of miR-122 within the chromatin, CRISPR/Cas9n genome modification was designed to disrupt the potential miR-122 seed match that we identified downstream of the pre-miR-122 encoding gene, with the aim of investigating whether miR-122 autoregulates in a similar fashion to let-7 in *C.elegans*.

5.2 miR-122 is present in the chromatin fraction

To understand the differential expression of miR-122 within the nucleus, chromatin fractionation was used to isolate chromatin and nucleoplasmic fractions from the cytoplasmic fractions in Huh7 cells. In a method adapted from the Proudfoot lab [470] and original chromatin RNA/protein isolation method from Wuarin and Schibler [497], a series of lysis buffers were used to first separate the nuclei from the cell pellet and then isolate the chromatin-associated fraction from the nucleoplasmic fraction before isolating the RNA and proteins from these three fractions.

Western blots performed by Dr Aimée Parsons on protein samples extracted from the three fractions, confirm that the chromatin-associated fractionation protocol efficiently separates the nucleoplasm from the cytoplasm and chromatin from the nucleoplasm with no cross-contamination. In Figure 5.1, an antibody binding to histone protein H2B showed H2B only to be present in the chromatin fractions as expected. Likewise, an antibody binding to tubulin was used to detect the cytoplasm as this protein forms a major component of the cytoskeleton and was present only in the cytoplasm fractions. The western blot confirms that the nucleoplasm and the chromatin are efficiently separated, as the chromatin and cytoplasmic markers only picked up in their respective fractions and nothing bound in the nucleoplasm fraction.

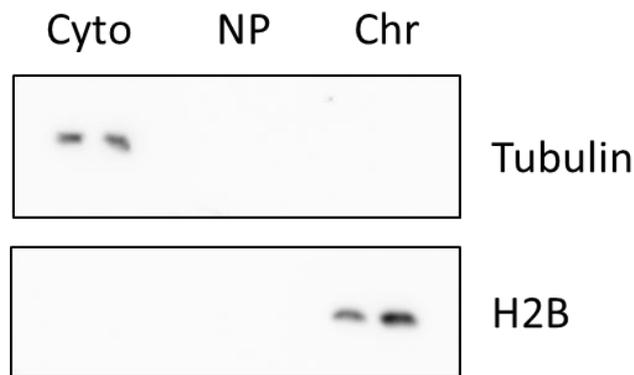
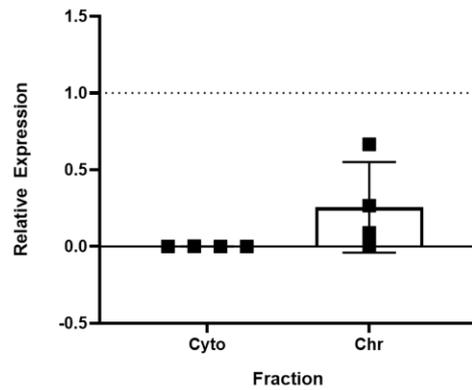


Figure 5.1: Western blot analysis performed by Dr Aimée Parsons of fractions obtained following chromatin-associated fractionation. Equivalent volumes of material were loaded from samples taken from Huh7 cells and represent the cytoplasm (Cyto), nucleoplasm (NP) and chromatin (Chr).

After confirmation of clean fractionation, the expression level of miR-122 within these fractions was analysed using qRT-PCR. The samples were first normalised by RNA concentration, where 100 ng total RNA was used as input for the RT reaction, and in Figure 5.2 the expression of miR-122 in chromatin and cytoplasmic fractions is presented either (A) relative to the expression levels in the nucleoplasm ($2^{-\Delta Ct}$) or (B) as raw Ct values.

In Figure 5.2(A), there is slightly less miR-122 in the cytoplasm and chromatin compared to the nucleoplasm. The low levels in the cytoplasm were unexpected and may suggest that the cytoplasmic material was poorly fractionated or extracted or degraded, for which it would be desirable to check RNA integrity by gel electrophoresis. However, these results do confirm previous observations that mature miR-122 is present in the nucleus [311, 379]. Most interestingly, whilst the expression level of miR-122 in the cytoplasm is likely to be inaccurate, when considering the raw expression levels in Figure 5.2B miR-122 is detected in the chromatin fractions at a comparable level to the nucleoplasm.

A.) Expression of miR-122 in Chromatin and Cytoplasm Relative to the Nucleoplasm



B.) Raw Expression levels of miR-122 following Chromatin-Associated Fractionation

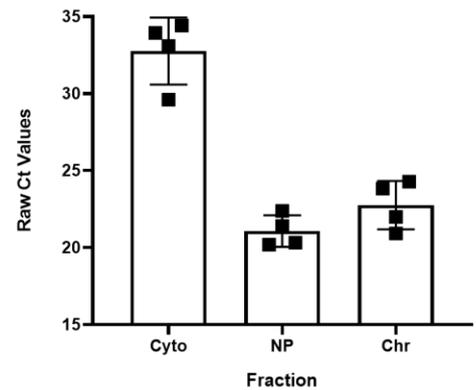


Figure 5.2: miR-122 is present in the chromatin of Huh7 cells. qRT-PCR data showing the **A.)** expression of miR-122 relative to the nucleoplasmic fraction by 2^{-dCt} , and **B.)** raw Ct values of miR-122 in Huh7 cells following Chromatin-Associated Fractionation. (Cyto= cytoplasm, NP= nucleoplasm, Chr= chromatin). Bars show the average of n=4 independent experiments with values from individual experiments shown as black squares and error bars showing standard deviation.

To corroborate the results obtained by qRT-PCR, northern blots were performed. 10 μ g of RNA for each sample was loaded and probed using radioactive anti-sense oligonucleotides for miR-122 and U6 snRNA. Unfortunately, the probe for miR-122 was unable to detect miR-122 in any fraction and the signal for U6 was much weaker than expected, see Figure 5.3. Whilst radioactive probes are thought to provide the strongest signal for northern blots, as a whole, northern blots are less sensitive than qRT-PCR and are more prone to degradation of RNA samples. It may be that the antisense probe design required optimisation or that the samples had been degraded before/during the blotting procedure.

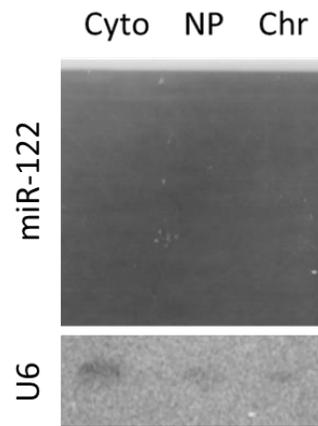


Figure 5.3: Northern blot analysis of RNA obtained following chromatin-associated fractionation. Radioactive probes for miR-122 and U6 snRNA bound to samples from Huh7 cells following chromatin-associated fractionation. Cytoplasm=Cyto, Nucleoplasm=NP, Chromatin=Chr.

5.3 Design of a CRISPR/Cas9 strategy to modify a miR-122 seed match in the pri-miR-122 gene

As miR-122 was found in association with chromatin, the next step was to investigate if it is regulating the processing of nascent transcripts. We chose pri-miR-122 as a candidate transcript due to the presence of a seed match downstream of the miR-122 encoding gene, similar to the mechanism for autoregulation of pri-let-7 in *C.elegans*. Such sites were thought to be conserved to some degree as seed matches were also identified in both mouse and rat with the potential for supplementary pairing downstream, Figure 5.4.

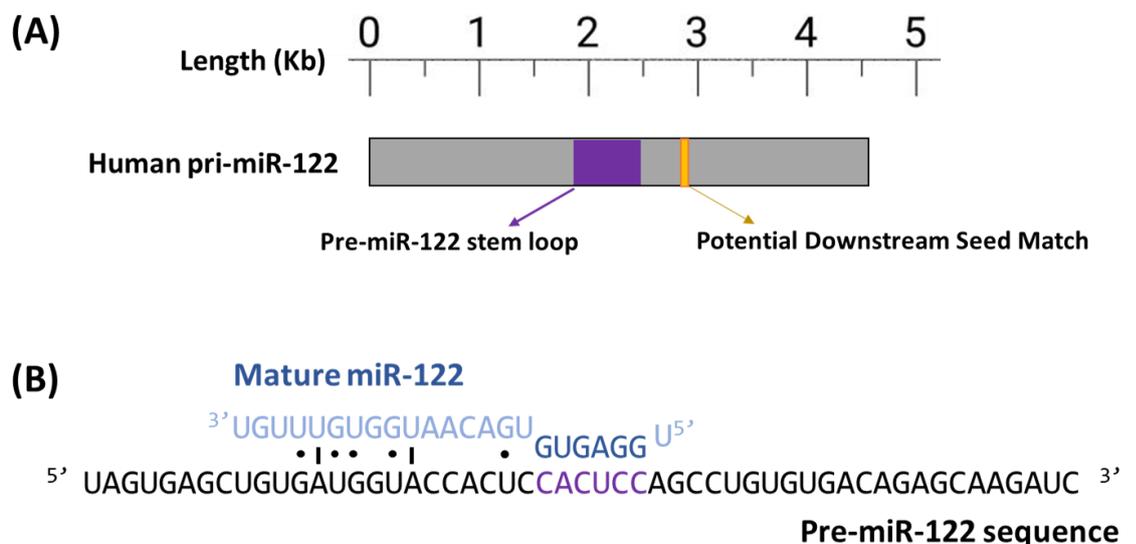


Figure 5.4: Schematic showing relative location and binding potential of downstream miR-122 seed match. Showing (A) the location of the potential downstream seed match (yellow) in relation to the pre-miR-122 stem loop (purple) in the pri-miR-122 sequence for humans. (B) mature miR-122 (blue) interacting with the potential seed match (purple) in the human pri-miR-122 sequence with possible wobble base U-G pairings indicated by a dot.

In order to investigate the role of this site, a CRISPR/Cas9 approach was used with the aim of disrupting/altering the potential seed match which is crucial to microRNA target binding. To target this genomic region, the Cas9n requires targeting by short (20 bp) guide RNAs (sgRNAs) to proximal sequences on opposing strands of the target site that contain an adjacent protospacer adjacent motif (PAM) [499].

In this strategy, a Cas9 nuclease variant (SpCas9n) with a D10A mutation in the RuvC domain, was used to reduce the frequency of off-target effects. This enzyme is capable of cleaving only the target strand, generating a single-strand break rather than a double-strand break at the target site [499, 500]. Single strand breaks are repaired by the high-fidelity base excision repair (BER) typically without mutation. Therefore, to achieve gene editing, two pairs of guides are designed to target opposing strands of the target locus, introducing two single strand breaks within close proximity. This is recognised as a double strand break and results in removal of surrounding nucleotides and DNA repair through non-homologous end joining (NHEJ), a cellular mechanism of double strand break repair. The NHEJ pathway is inherently error-prone and so often during repair will produce a sequence that contains either insertions or deletions (InDels). Requiring the targeting from a pair of sgRNA reduces the chance of off-target editing occurring as the pairs target distinct sequences that are unlikely to be present in close proximity at an alternative locus [499, 500]. In this case, we hoped to use the CRISPR/Cas9 Nickase (CRISPR/Cas9n) system to disrupt the potential miR-122 seed match by introducing InDels, which is outlined in Figure 5.5.

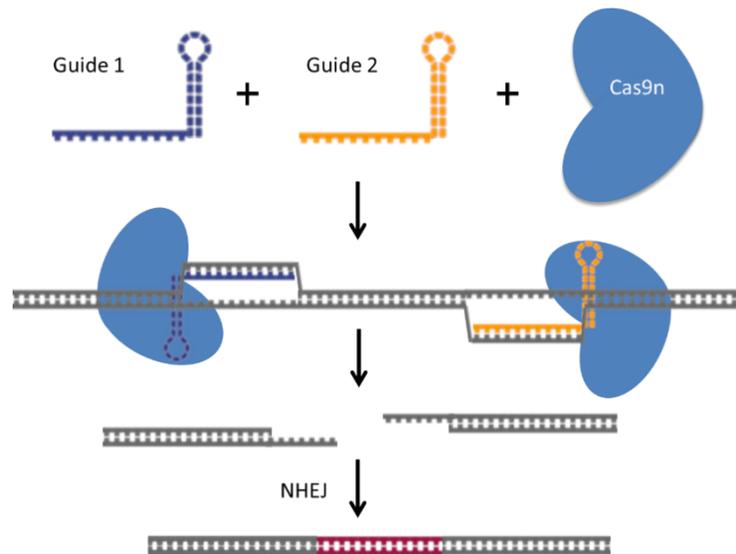


Figure 5.5: Process of Cas9n DNA mutation. The guide sequences target the SpCas9n enzyme to a specific locus where two single strand nicks are made in the DNA flanking the target site. This is recognised as a double strand break which is then repaired by error prone NHEJ.

A pair of 20 nt sgRNAs, seen in Figure 5.6, were designed using the bioinformatics tool ChopChop [469, 501] to introduce double-strand breaks (DSB) at either side of the potential miR-122 seed match. It identifies pairs of specific guide sequences ending in the protospacer adjacent motif (PAM) NGG, one on the forward and one on the reverse strand. Briefly, the software filters sgRNAs with predicted self-complementarity and avoids regions of high-structure that limits accessibility for the Cas9. It then scans the genome to identify potential off-target sites, allowing a few mismatches, whilst also calculating the efficiency of DSBs to result in frameshift mutations to score for appropriate guides.

The sgRNA guide expression vectors were generated by annealing and phosphorylating oligonucleotides comprising the guide sequences (sgDNA_F_Top and sgDNA_F_Bottom; sgDNA_R_Top and sgDNA_R_Bottom) and ligating the resulting oligonucleotide pairs into Bbs1-digested Cas9n-containing px461 vector.

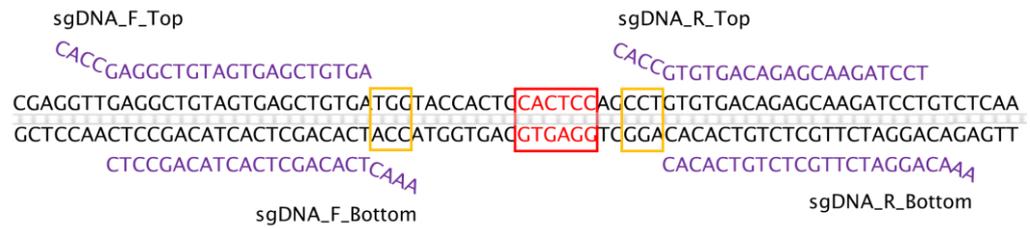


Figure 5.6: Design of guide RNA sequences with respect to the potential miR-122 seed match. Sequence approximately 500 bp downstream of miR-122 encoding DNA, showing sgDNA sequences in purple used to create the gRNA-containing plasmid, and endogenous PAM sequences boxed in yellow, in respect to the miR-122 seed match in red.

Clones containing the correct guide inserts were confirmed by sequencing with the human U6 forward primer (Supplementary Figure 15). Following sequence verification, the px461-sgDNAF/R Cas9n plasmids were cultured and purified by maxiprep to obtain the plasmids for transfection. A plasmid map for such can be found in Supplementary Figure 16) and outlines the location of the guide RNA insertion in terms of the Cas9n encoding gene.

5.4 Generation and genotyping of CRISPR/Cas9n cell lines

The px461-sgDNA Cas9n plasmids were transfected into Huh7 cells in an attempt to carry out genome modification at the potential miR-122 seed match. To allow identification of successfully transfected cells, the px461 vector includes the sequence coding for eGFP as a bicistronic reporter alongside the Cas9n encoding gene, shown in Supplementary Figure 16. Therefore, if the cells express the mRNA for the fluorescent reporter GFP, the cells will also be producing the mRNA for the Cas9n protein of interest and should also be expressing the sgRNA.

Initially, the calcium phosphate method was used but yielded a very low transfection efficiency of less than 5%, estimated by counting cells glowing green when excited at 488 nm, indicating GFP^{+ve}. As a result, Lipofectamine 2000 was used to transfect Huh7 cells with either a single px461-sgDNA Cas9n expression vector (to act as a control) or co-transfecting with both forward and reverse px461-sgDNA Cas9n guides, achieving approximately 15% and 20% transfection efficiency respectively. 24 hours post-transfection fluorescence-activated cell sorting (FACS) was used to isolate successfully transfected cells and subsequently seed single cells into 96-well plates. Unfortunately, no growth from single cells was seen. This was likely a result of the cells not being in ideal conditions (37°C, 5% CO₂) as they were outside an incubator for a large proportion of the day. Attempts were made to mitigate this, including being transported on ice to the FACS facility, and being immediately placed in pre-warmed fresh media after sorting. However, the harshness of sorting first into a GFP^{+ve} bulk population and then secondly into single cells may have caused the cell death.

A series of tests was performed on Huh7 cells, seeded by hand as single cells into 96-well plates, to investigate whether media containing extra FBS or conditioned media (50% old media removed and re-added along with 50 % fresh media) increased the chance of cell proliferation. Ultimately from visual inspection the choice of media did not make a difference, however splitting the cells by hand rather than using FACS increased cell survival. Therefore, a direct repeat of the transfections was performed but FACS was only used to sort cells into a bulk population of GFP⁺ cells. Figure 5.7 shows the FACS analysis from untransfected wild-type control cells, and Huh7 cells transfected with either single-guide Cas9n or both px461-sgDNA plasmids, successfully identifying GFP⁺ cells from the transfected samples. The bulk sort of GFP⁺ cells through FACS was able to isolate 19.42k and 16.55k cells for co-transfected and single guide control conditions, respectively. Seeding of single cells was then performed manually into 96-well plates after allowing the bulk populations to recover in media at 37°C overnight.

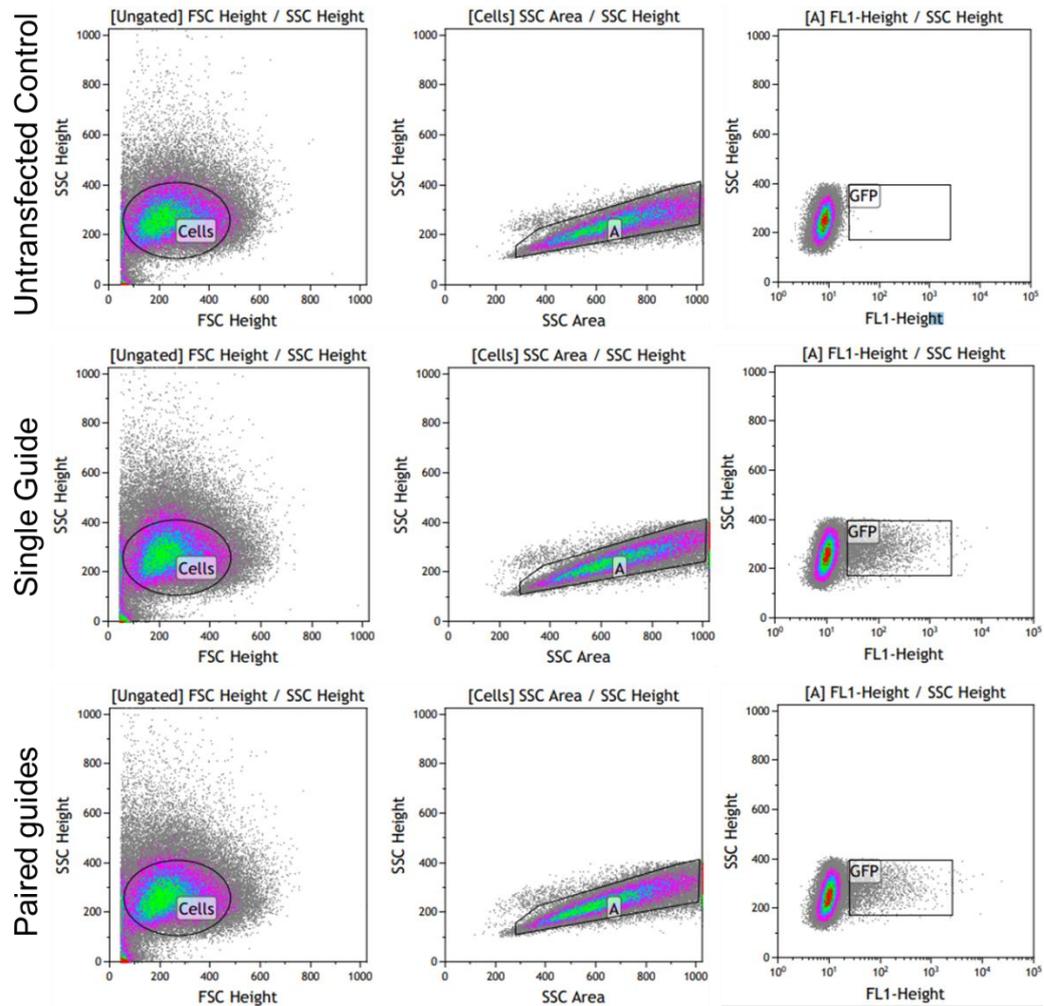


Figure 5.7: FACS analysis of untransfected wild-type Huh7 cells, and Huh7 cells transfected with either single Cas9n guide (control) or paired Cas9n plasmids. Ungated FSC Height/SSC Height separates healthy cells (circled) from apoptotic cells/debris. SSC Area/SSC Height isolates single cells (circled) from doublets passing through the capillary. FL-1 Height/SSC Height identifies those cells fluorescing at 510 nm (boxed) due to incorporation of GFP-tag and therefore isolating cells transfected with the px461-sgDNA Cas9n plasmids.

After several weeks, 15 clonal populations of Cas9n cells and 3 populations of single guide controls were achieved. Once a sufficient number of cells were available (4–8 weeks), genomic DNA was extracted from these populations. The targeted region of genomic DNA was amplified by PCR using the Genotyping_F/R primers. Due to the error prone nature of NHEJ a variety of insertions and deletions are expected and therefore changes in product size are expected for successful mutation. Unfortunately, all Cas9n transfected cells (clones A-H) produced bands of equivalent size to the single guide controls sgA and sgB (See Figure 5.8) suggesting no InDels were achieved. One single guide clone, and two paired px461-sgDNA clones were chosen, highlighted in yellow in Figure 5.8, and genomic DNA sent for sequencing, using the Genotyping_F primer. Figure 5.9 shows alignment of these sequences with the wild-type regions which confirmed all three were wild-type sequences for the region of the pri-miR-122 gene spanning the miR-122 seed match.

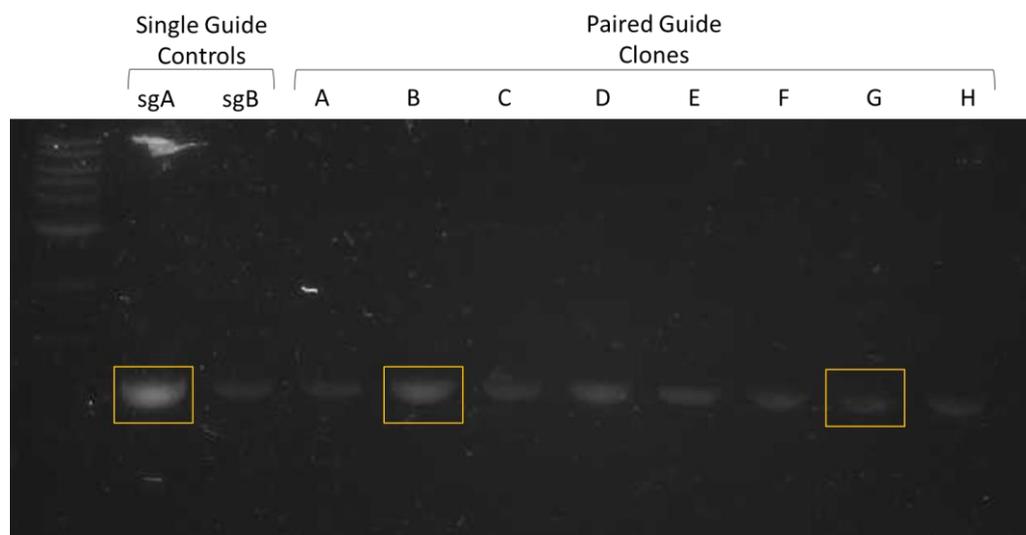


Figure 5.8: Genotyping of CRISPR/Cas9n clonal populations suggests no gene editing occurred. Agarose gel electrophoresis of PCR product using genotyping F/R primers of genomic DNA from single guide controls and paired px461-sgDNA Cas9n clones. Excised bands highlighted in yellow box.

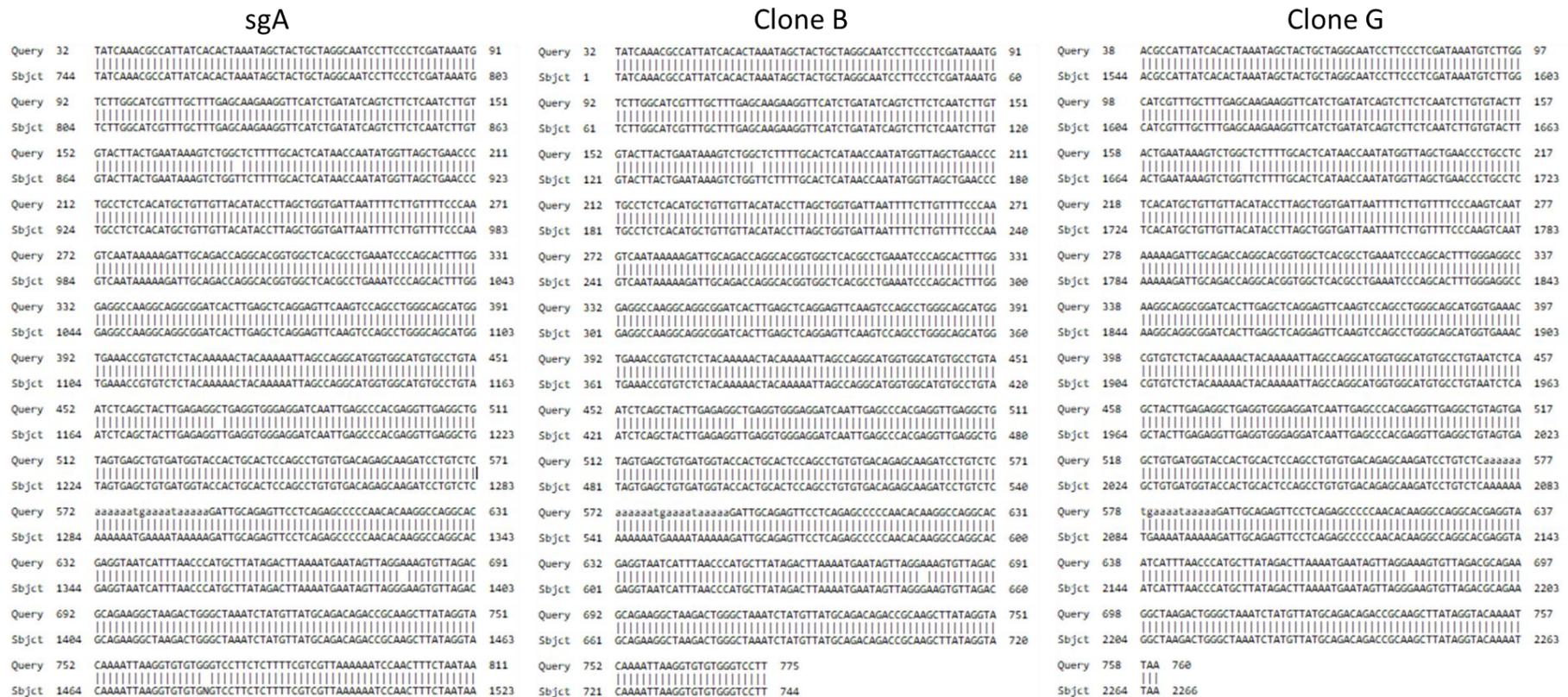


Figure 5.9: gDNA sequence from CRISPR/Cas9n clonal populations perfectly align to targeted wild type sequence. Nucleotide blast alignment (NCBI) of wild type sequence (Subject) for the pri-miR-122 encoding gene with sequences obtained (Query) for the single-guide control (sgA) and paired px461-sgDNA Cas9n clones B and G.

5.5 Discussion

Using a chromatin-associated fractionation technique, miR-122 has been shown to be present in the nucleoplasm and chromatin of Huh7 cells as well as the cytoplasm. This substantiates the preliminary data from Dr Andrew Lewis, which identified a subset of miR-122 in the nucleus of Huh7 cells using *in situ* hybridisation.

For qRT-PCR analysis of the fractions, normalisation to standard housekeeping controls was deemed unsuitable as U6 snRNA for example is predominantly localised to the nucleus and would likely bias the expression levels. As a result, the fractionation samples were normalised to 100 ng total RNA concentration for input into the RT reaction. However, this method of normalising could account for the unexpectedly low levels of miR-122 seen in the cytoplasmic fractions as the high concentration of rRNA in the cytoplasm compared to the nuclear fractions will have affected the total RNA concentration. For future investigation into microRNA expression levels, it would be possible to spike in non-endogenous microRNA, such as cel-miR-39 to control for differences in RT reaction conditions. In addition, it could also be possible to compare the abundance of nascent (unspliced) and mature (spliced) mRNA to account for differences in fractionation technique as nuclear fractions should be enriched in unspliced transcripts and relative ratio of fractions should be similar between experiments. However, it is unlikely that the expression of mature microRNA in the nuclear fractions is due to contamination with nascent pri-microRNA as the RT and qRT-PCR probes used are designed in such a fashion that they are specific to the mature form.

Northern blot analysis of miR-122 expression in these fractions was attempted as a means to validate the qRT-PCR results but unfortunately no signal for miR-122 was produced over several attempts. Northern blots are much less sensitive than qRT-PCR despite being a more direct measure, and although radioactively labelled probes were used, quantification of miR-122 was not achieved. Optimisation of the antisense oligonucleotide probes could be achieved however as the detection of U6 was poorer than expected, despite the U6 probe being

used previously in the CLJ lab without issue, it is more likely that the samples were either not concentrated enough before loading and/or were subjected to RNA degradation during the process or there were technical problems with the blotting, crosslinking or probing techniques. Follow up work would involve optimising the northern blot methodology in order to corroborate the expression of miR-122 in the chromatin-associated fraction. But as many RNA processing events such as splicing and pri-miR-122 processing are co-transcriptional, and so occur in association with the chromatin, finding miR-122 in this fraction using qRT-PCR suggests miR-122 may have the potential ability to bind nascent transcripts and regulate transcription/co-transcriptional processing.

This is consistent with the work done by Wang *et al* [379] published during the course of this study, who not only identified miR-122 in the nucleus of Huh7 cells but also showed that miR-122 is capable of regulating the biogenesis of another microRNA by binding to a cognate site on the pri-microRNA transcript. They used ISH and qRT-PCR to confirm miR-122 presence in the nucleus following purification of the nuclear and cytoplasmic compartments and found that transfection of a miR-122 mimic decreased the levels of pre-miR-21 whilst transfection with an anti-miR-122 increased pre-miR-21 levels. Their work involved the construction of luciferase reporters and pri-miR expression plasmids to investigate a putative seed match region for miR-122 upstream of the pre-miR-21 encoding region of the primary transcript and found that miR-122 is capable of directly binding to this site to regulate the biogenesis of miR-21.

It is difficult to determine if the amount detected is a biologically relevant amount of miR-122 in the chromatin as this is dependent on the stoichiometry between chromatin and miR-122 where the number of targets and binding sites present for miR-122 in the chromatin is unknown. However, as miR-122 was found in association with chromatin and studies have shown miR-122 capable of regulating the biogenesis of miR-21 in the nucleus, we chose to investigate whether miR-122 was able to regulate the processing of nascent pri-miR-122 transcripts. This candidate was chosen for investigation as we identified the

presence of a potential seed match on the pri-miR-122 transcript that could lead to autoregulation by miR-122 in a fashion similar to that seen of Let-7 in *C.elegans*. To do so, we designed a genome editing strategy utilising CRISPR/Cas9n to disrupt this potential seed match which is vital for miR-122 target binding.

Disappointingly, the CRISPR/Cas9n modification of the potential seed match was unsuccessful, with the screening of the genomic DNA showing no incorporation of InDels. Although not peer-reviewed, Schubert and Yan performed a systematic assessment of how gRNA design affected the efficiency of nicking and InDel formation [502]. They found that D10A Cas9 mutants, such as the px461 plasmid used here, have been shown to have more efficient editing when the two nick sites are 40-70 bp apart and that efficiency of targeting drops dramatically if nick distances are too distant >85bp or too close <25bp like our approach targeting an 18 nt region. Equally, the orientation of the PAM sequences in respect to the guide RNA and target locus can alter the efficiency of editing, with PAM sequences facing outside the target region (PAM-out) showing higher efficiency. Despite using a range of bioinformatic tools to design the sgRNA guides, the only guides generated that were capable of targeting such a small region (the 6 nt binding site in

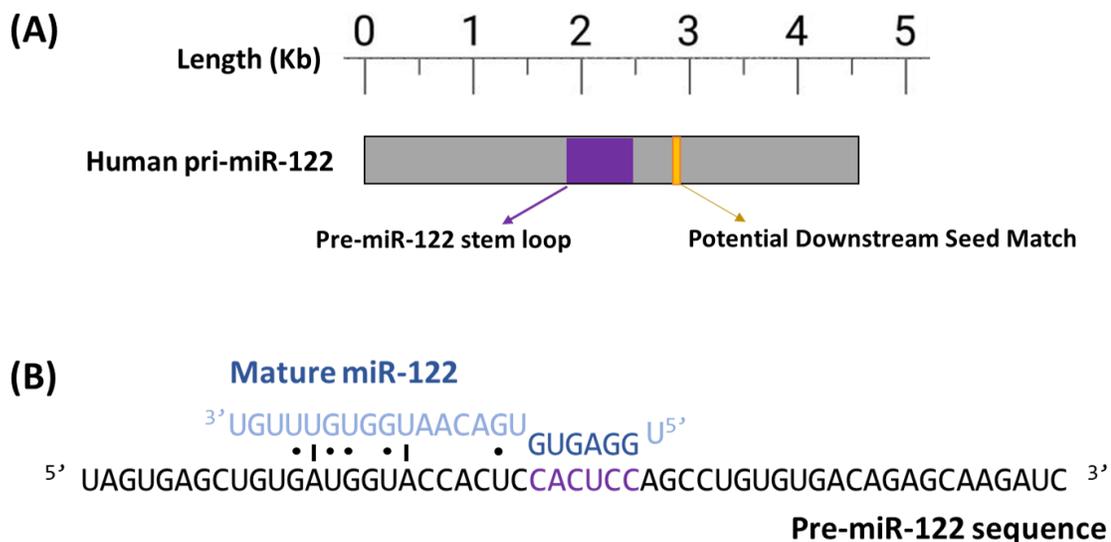


Figure 5.4) and contained endogenous PAM sequences without having predicted off-target effects were all within close distance of each other.

It would be possible to change to an alternative CRISPR/Cas9 method as the techniques available and updates to the approach have come a long way, with dramatic progress made even by the time our attempt of gene editing concluded. Firstly, a wild-type version of the Cas9 enzyme could have been used as this would require only a single Cas9n to cut at the target site and increase the likelihood of NHEJ occurring. Whereas with the nickase approach, the requirement for two nick sites and therefore two separate sgRNA-Cas9n plasmids has a lower probability of both nicks occurring and therefore reduces the chance of a double-strand break being recognised and stimulating NHEJ. The increased risk of off-target effects with the wild-type Cas9 approach may be acceptable for the increased probability of achieving successful editing, with screenings of additional edited clones performed to reduce the chances of observed effects due to off-target effects. An alternative Cas9, xCas9, was generated through Phage-assisted continuous evolution that has additional mutations in the PAM interacting domains which allows for expanded PAM recognition (including NG, GAA and GAT) [503]. This ability to recognise a broad range of PAM could have increased the selection of guide RNAs that were able to target the seed match region, some of which may have two nick sites the optimal 40-70 bp apart.

CRISPR/Cas9 systems capable of editing single bases have also been engineered. Cytosine base editors (CBEs) have a cytidine deaminase fused to Cas9 to target the enzyme to a specific locus [504, 505]. The CBE creates a C to T base change by converting cytidine to uridine within the target site, which subsequently, through base excision repair, gets converted to thymidine. Similarly, adenosine base editors (ABEs) have been created to create an A to G base change within the target region [506, 507]. The ABEs have a Cas9 protein fused to an engineered adenine deaminase (evolved from *E.Coli* TadA) that converts adenosine to inosine which gets recognised by the cell as a guanosine. These base editors are designed to target a small editing window proximal to the

PAM sequence and as a single base change in the seed match would be enough to disrupt microRNA binding such a CRISPR approach would be suited to target the small 6 nt seed match of interest in the pri-miR-122 gene.

In addition, prime editing uses a Cas9 nickase fused to an engineered reverse transcriptase and a prime editing guide RNA (pegRNA). The pegRNA both specifies the target site and encodes the desired edit into the target genomic locus allowing insertions/deletions and single nucleotide changes without introducing DSBs. Therefore, achieving similar efficiency but fewer off-target effects than homology-directed repair-based methods. However, due to availability of reagents at my disposal and ease of cloning, the CRISPR/Cas9n approach was chosen as the gene editing method.

During this time, we also received back alignments from Dr Vladimir Ovchinnikov, a post-doctoral fellow in Mary O'Connell's group, School of Life Sciences, University of Nottingham, which suggested this sequence downstream of the pre-miR-122 coding sequence was less conserved than previously thought. He assessed the conservation of this binding site across a 3058 nt region beginning at the pre-miR-122 sequence from 17 genomes from a range of vertebrate sequences. These alignments can be found in Supplementary Figure 17 (as 800 nt alignments for size constraints), but unfortunately resulted in conservation only across primates, see Figure 5.10 below. Whilst we did find potential binding sites in rat and mouse, the exact seed match sequence and location of these differed and the lack of conservation of the region as a whole suggests it may not have been selected for in evolution.

```

Cja      CTGCAGTGAGCTGTGATGGCACC GCGGCACTCCA-3p
Ptr      CTGTAGTGAGCTGTGATGGTACCACTGCACTCCA-3p
Hsa      CTGTAGTGAGCTGTGATGGTACCACTGCACTCCA-3p
Mm1      CTGTAGTGAGCTGTGATGGCACC ACTGTACTCCA-3p
*** ***** ** * * *****
          :|:: ;|   ;:;| |||
3' -guuugugguaacagugaggu-5' miR-122

```

Figure 5.10: Conservation of the miR-122 binding site across primates. Sequence downstream of the pre-miR-122 coding region performed by Dr Vladimir Ovchinnikov.

For future investigations, in order to look directly at whether miR-122 does interact with candidate sites in pri-microRNAs, a method for RNA immunoprecipitation (RIP) would need to be developed. By pulling down Ago1-4, it would be possible to isolate the RISC and analyse through qRT-PCR which, if any, pri-microRNA transcripts are bound by miR-122.

Chromatin-fractionation of Huh7 and Huh7 Δ miR-122 cells could be used to determine whether the processing of pri-microRNA transcripts is affected by miR-122. Attempts were made to look at the effect on processing of pri-miR-122 in chromatin-associated RNA by mature miR-122 in Huh7.5 and Huh7.5 Δ miR-122 from the Rice lab but were unsuccessful. Largely this was due to persistent problems with DNA contamination of primer sets, despite comprehensive efforts to eliminate the source, which lead to a high amount of background in the no template controls and false detection of the deleted hairpin structure in the Δ miR-122 cell line. Likewise, the Δ miR-122 cells could be used to investigate miR-122 regulation of other microRNA primary transcripts with initial transcripts chosen by comparing the mature microRNA content of WT and Δ miR-122 cells.

Chapter Six: Discussion

6.1 Summary

The overall aim of this thesis was to perform a direct comparison of miRNA function in different subcellular sites in human cells, specifically at the Endoplasmic Reticulum (ER), as miRNA here may have a direct role in silencing transcripts encoding secreted or membrane-localised proteins, and in the nucleus where they may have a range of functions including the regulation of microRNA biogenesis and regulation of nascent transcripts at the chromatin.

A series of ER-translated luciferase reporters for miR-122 regulation at the 3' UTR were successfully generated and their regulation by miR-122 was compared with that of equivalent reporters translated in the cytoplasm. In addition, secreted and cytoplasmic luciferase reporters were constructed which enabled miR-122 activation of translation via 5' UTR sites from HCV RNA to be compared at the ER versus cytoplasm. In summary, there was evidence for differential regulation via 3' UTR sites at the ER and cytoplasm in some, but not all, tested reporters. Although the regulation seen with the Firefly and Gaussia reporters was not replicated with the NanoLuc reporters, the differences in regulation are worth following up, possibly with more concentrated LNA inhibitors and mimics.

In conjunction, miRNA-mediated regulation at the ER and cytoplasm was compared by applying a membrane fractionation method to isolate ER and cytoplasm-localised mRNAs and the effects of miR-122 inhibition and overexpression on endogenous miR-122 targets was monitored in these fractions. An increase was observed in the mRNA levels in the cytoplasmic fractions that coincided with a decrease in membrane fractions upon inhibition with a 2' Ome oligonucleotide, although optimisation of the membrane fractionation protocol and normalisation strategy is required to make firm conclusions about the regulation of miR-122 targets at the ER.

Additionally, a fractionation method was used to isolate chromatin and nucleoplasmic fractions from the cytoplasm which detected the presence of miR-122 specifically in the chromatin fractions. As a potential miR-122 seed match

was identified downstream of the pri-miR-122 encoding gene, CRISPR/Cas9n genome modification was designed to disrupt this sequence with the aim of investigating whether miR-122 autoregulates in a similar fashion to let-7 in *C.elegans* but was unsuccessful, with the screening of the genomic DNA showing no incorporation of InDels.

6.2 Use of Luciferase Reporters for Monitoring miRNA activity at the ER and Cytoplasm

Luciferase reporter assays are a useful tool for monitoring miRNA activity in cells, as the amount of light produced by the luciferase protein provides a quantitative measure of the protein synthesis from the luciferase reporter gene thereby determining a direct functional connection between the microRNA of interest and the luciferase reporter. In this thesis, luciferase reporters were designed to enable the monitoring of repression by miR-122 via the 3' UTR and up-regulation in translation via the 5' UTR. In addition, the incorporation of a signal peptide enables the comparison of regulation of secreted reporters translated at the ER with that of equivalent reporters translated in the cytoplasm, for example the comparison of cytoplasmic Firefly and secreted Gaussia luciferase reporters in this thesis. Furthermore, a series of NanoLuc reporters were designed both with and without the inclusion of a signal peptide which enabled the comparison of regulation at the ER and cytoplasm with reporters of the same open reading frame (ORF) to eliminate any inherent differences in the reporter proteins (discussed in detail in Section 3.10).

The cytoplasmic Firefly and secreted Gaussia luciferase reporters showed a difference in regulation at 48 hours which suggests the possibility of differential regulation occurring between the ER and cytoplasm. However, there was not as strong an effect with the NanoLuc reporters, with the possibility that the difference seen with the Firefly and Gaussia reporters is due to inherent differences in the reporters, such as mRNA stability, codon-optimisation, and GC-content as discussed in Section 3.10. Although the regulation we saw with the Firefly and Gaussia reporters was not replicated with the NanoLuc reporters, the changes that were observed are worth following up, possibly with more concentrated LNA inhibitors and mimics.

In future, it would be of interest to perform the NanoLuc reporter assays following knockdown of RISC components such as TNRC6A/B/C, AGO(1-4), DDX6, and

eIF4All, as these are proteins known to mediate miRNA-mediated repression and have been shown to localise to different subcellular sites such as the ER [350]. In particular, an alternative form of RISC, P-miRISC, has been identified in *Drosophila* that forms upon serum starvation and associates with the ER where it was shown to result in up to 10x stronger repression of target mRNA [ref 324]. The knockdown of RISC components could be applied to the experiments investigating the endogenous miR-122 targets rather than just to the luciferase reporters and would be able to determine if these proteins play a different role in regulation at the ER compared to the cytoplasm.

In addition, for the 5'Nluc3' RNA reporters, it might be that there is a difference in regulation in the context of an actual HCV infection, which would disrupt the endomembranes and so a comparison of miR-122 regulation of these RNA reporters at the ER could be compared after infection with an attenuated HCV virus or replicon.

6.3 Optimisation of Subcellular Fractionation

Methods

In this thesis, membrane-associated fractions were isolated using a sequential detergent method adapted from Jagannathan *et al* [446] which allowed the comparison of miR-122 regulation at the ER and cytoplasm. In addition, the presence of miR-122 in the nucleoplasm and chromatin of Huh7 cells was detected using a chromatin-associated fractionation technique adapted by the Proudfoot lab.

Poor reproducibility across experiments was a consistent problem that made it difficult to interpret the results, with inconsistency in the efficiency of fractionation being a possible source of variability. Fractionation methods such as these are useful in isolating subcellular sites due to their ease, sufficient yield for downstream biochemical analysis, and compatibility with a range of other techniques such as AGO-IP. However, as completely clean separation of subcellular compartments is not possible and mRNAs may move during lysis, results are obtained solely for the fractions and cannot be applied in their entirety to the specific organelle of interest.

Alternative methods for membrane fractionation are available and attempts were made at optimising a differential centrifugation method but were unsuccessful [495]. The different sedimentation rates of the subcellular compartments make it possible to separate pellets that correspond to the heavy membranes (Endomembranes, MVBs), light membranes (ER) and the cytosol. Once again, this method of fractionation provides an enrichment rather than purification of the membranous compartments but would allow more specificity for the ER and may reduce variability between experiments.

Iodixanol (Optimem) gradient fractionation could be used to separate out the ER-enriched fractions from other subcellular membranous compartments such as the endomembranes as it separates cellular components based on their different densities [352]. This would provide a more ER-enriched fraction than the current

membrane fractionation method which simply separates membranous compartments (ER, golgi, and endosomes) from the cytoplasm, but would still require proper characterisation of the fractions and as it is a continuous iodixanol gradient is still prone to cross-contamination of fractions.

These membrane fractionation approaches could also be used to isolate other subcellular sites such as the mitochondria and components of the endosomal pathway. A large number of miRNA and RISC components have been identified in the mitochondria, isolated from a broad range of samples (e.g. rat, mouse, human tissue, human cells) [398-402], and as mitochondria possess their own genome and mitochondrial translation machinery they are likely to be a site of miRNA-mediated regulation [508]. Likewise, RISC components (AGO2, TNRC6) and CNOT4 have been found to associate with endosomes and multivesicular bodies (MVBs) in human cells, with ESCRT (required for MVB formation) knockdown inhibiting miRNA-mediated silencing, and HRS knockdown (required for endosome maturation) increasing the stability of mRNAs [327, 329]. Therefore, fractionation techniques similar to those used in this thesis would allow the investigation of microRNA-mediated post-transcriptional regulation of the endosomal pathway and at the mitochondria.

6.4 Endogenous miRNA-mediated regulation at the ER

In this thesis, membrane fractionation enabled the comparison of miR-122 regulation of endogenous targets ALDOA, CAT1, and G6PC3 between ER-rich and cytoplasmic fractions. The effects of miR-122 inhibition and overexpression on these known miR-122 mRNA targets was compared in these fractions and different regulation of these endogenous targets was observed between the ER and cytoplasm. However, this was just an initial investigation monitoring the mRNA and protein levels of three known targets. It would be interesting to include a much bigger panel of miR-122 targets, including more membrane/secretory protein encoding mRNAs such as the molecular chaperone ankyrin repeat domain 13C (ANKRD13C), the collagen synthesis enzyme prolyl 4-hydroxylase subunit α 1 (P4HA1) and the vesicle trafficking protein SEC22 homolog C (SEC22C). Moreover, it would be interesting to investigate the regulation by miRNAs other than miR-122 as there may be differences in where they regulate their targets.

Additionally, it could be possible to clone a known-cytoplasmic miR-122 target, such as ALDOA, to incorporate an ER-localisation signal such as SNAP-KDEL thereby directing the mRNA target to be translated at the ER. On the other hand, it has been shown that overexpression of GFP-Sec61 β saturates the mRNA binding sites at the ER and could therefore be used to displace endogenous mRNAs from the ER, such as G6PC3 [509]. These techniques would allow the direct comparison of regulation at the ER and cytoplasm by endogenous miR-122 of the same mRNA target.

In this thesis, efforts were made to investigate the regulation of miR-122 at the ER under physiological conditions, but it would be interesting to apply the luciferase reporters and membrane fractionation experiments to ER stress and hypoxia conditions. The unfolded protein response (UPR) is activated by the accumulation of unfolded or misfolded proteins in the ER and results in the attenuation of mRNA translation and increase in the production of molecular

chaperones. If ER stress remains unresolved, the UPR activates apoptosis pathways. Several microRNAs have been shown to fine-tune the cell's response to ER stress by promoting either cell death or cell survival [510, 511]. In addition, while AGO is predominantly localised to the cytoplasm it has been shown to relocate to stress granules upon ER stress [512]. Furthermore, miR-122 overexpression as a result of hepatocellular cancers results in the repression of the UPR pathway and so it would be interesting to examine whether ER stress, inducible with thapsigargin, influences miR-122 regulation at the ER which could be investigated using both the luciferase reporters and the membrane fractionation experiments developed in this thesis.

6.5 Global miRNA-mediated Regulation at Different Subcellular Sites

Rather than focusing on individual miR-122 mRNA targets, there are methods that could investigate differences in the global regulation of miRNA targets of different subcellular fractions such as miRNA-seq. One example, proximity-specific ribosome profiling uses a biotin ligase targeted to the ER to biotinylate ribosomes tagged with a biotin acceptor peptide (AviTag), and subjects this to ribosome profiling to map translational activity at the ER [513]. It can achieve high spatial specificity but is unable to detect non-translated mRNAs or non-coding RNAs. If applied to our investigation, this method could be used to look at ER-associated translation to compare miRNA-mediated translational repression at the ER versus cytoplasm.

Similar proximity-specific tagging utilises APEX, an engineered ascorbate peroxidase that catalyses the *in situ* biotinylation of proteins which can be targeted by genetic fusion to various subcellular sites of interest. Of most relevance are APEX-NES which contains a cytoplasmic nuclear export signal and ERM-APEX2 which targets the APEX to the ER cytosolic surface. The Ting lab found that APEX is also capable of directly biotinylating cellular RNAs *in situ*, and when combined with RNA-Seq can provide in depth analysis of the spatial transcriptome [514]. They showed that APEX-Seq is able to detect 90% of known membrane/secretory mRNAs including those lower in RNA abundance. Unlike ribosome profiling, which is only capable of capturing actively transcribing mRNAs, APEX-Seq is able to capture lncRNAs and antisense RNAs. Although it is beneficial that APEX-Seq is performed in living cells, it does mean that it is limited to RNAs accessible in the native environment, and as such cannot target RNAs trapped in macromolecular complexes. As such, APEX-Seq could be applied to Huh7 cells to look more precisely at the transcriptome at the ER and determine whether there is change following miR-122 inhibition and/or overexpression.

6.6 Investigation into the Role of Chromatin-Localised miR-122

In this thesis, to examine the localisation of miR-122 within the nucleus a fractionation method was used to separate chromatin and nucleoplasmic fractions from the cytoplasm which demonstrated the presence of miR-122 specifically in the chromatin fractions. Whilst the presence of miR-122 at the chromatin was ascertained through qRT-PCR and chromatin-associated fractionation, the possible role of miR-122 in these fractions remains unanswered. As the splicing and processing of pri-miRNAs occurs co-transcriptionally in the nucleus, miR-122 may have the ability to bind nascent transcripts and regulate transcription/co-transcriptional processing.

There have been a few studies that provide examples of nuclear miRNA regulating the biogenesis of other microRNA. For example, in mice, nuclear miR-709 has been shown to directly bind the pri-miR-15a/16-1 transcript to prevent its processing and maturation as a means to regulate cell apoptosis [376]. Most relevant is work done by Wang *et al* [379] published during the course of this study, where the use of luciferase reporters showed that miR-122 is capable of regulating the biogenesis of miR-21 by directly binding to the pri-miR-21 transcript thus blocking processing by the microprocessor complex. Their identification of miR-122 in nuclear fractions corroborates the presence seen in our chromatin fractions and provides a potential role of chromatin-localised miR-122.

Furthermore, there has been one example of autoregulation by a microRNA; in *C.elegans*, the microRNA Let-7 has been shown to bind to a complementary site in its own primary transcript to promote the downstream processing into mature Let-7 [378]. Initial analysis of the pri-miR-122 gene identified a potential seed match on the pri-miR-122 transcript, downstream of the pre-miR-122 hairpin, that could lead to autoregulation by miR-122 in a mechanism similar to Let-7 in *C.elegans*. As a result, a genome editing strategy using CRISPR/Cas9n was used to investigate this potential seed match with the aim of investigating whether

chromatin-localised miR-122 possesses the ability to autoregulate its biogenesis. Unfortunately screening of the genomic DNA showed wild-type sequences for the region of the pri-miR-122 gene spanning the miR-122 seed match with no incorporation of InDels. It would be possible to apply a different genome editing strategy to disrupt the potential seed match, which would not pose the same issues as the Nickase approach (as discussed in detail in Section 5.5), however assessment of the sequence downstream of the pre-miR-122 coding sequence by Dr Vladimir Ovchinnikov suggested this site was less conserved than previously thought.

Therefore, alternative methods could be used in future to investigate the role of miR-122 in the chromatin. For example, the Jopling lab possess a Huh7 cell line with the pre-miR-122 hairpin structure deleted (Huh7 Δ miR-122) that results in a knockout of mature miR-122. A comparison of the mature microRNA content of wildtype and Δ miR-122 Huh7 cells with miRNA-seq could identify miRNA that are potentially regulated by miR-122. Chromatin-fractionation of Huh7 and Huh7 Δ miR-122 cells could then be used to determine whether the processing of pri-microRNA transcripts is affected by miR-122. However, work performed during this thesis to investigate this was hindered by persistent problems with DNA contamination, despite comprehensive efforts to eliminate the source. This led to a high amount of background in the no template controls and false detection of the deleted hairpin structure in the Δ miR-122 cell line. In addition, immunoprecipitation of AGO(1-4) could also be used to identify which, if any, pri-microRNA transcripts are directly bound by miR-122 but again numerous attempts were made to optimise the AGO-IP method but were obstructed by DNA contamination.

6.7 Microscopy-based techniques for investigating the subcellular localisation of miRNA and RISC components

Whilst there is evidence for differences in the subcellular localisation of miRNA, there is not much known about what drives this. Some studies suggest that Exportin-1 may play a role in the export of nucleolar miRNAs to the cytoplasm as they found that treatment of cells with Leptomycin B (known to affect Exportin-1 dependent nuclear export) resulted in the abrogation of nucleolar miRNA transport to the cytoplasm [321, 392]. Furthermore, they found miR-29b contains a hexanucleotide sequence at its 3' end that directs its nuclear import and that when attached to other miRNAs or siRNAs also results in their nuclear localisation [321]. Preliminary data from the Jopling lab (data not shown) showed differential localisation of endogenous miR-122 and ectopically expressed miR-122 resulting from pre-miR-122 transfection, therefore it would be interesting to compare overexpression of miR-122 in miR-122 knockout cell lines and also different methods of expressing miR-122, i.e. plasmid, LNA mimic, or lentiviral constructs, to see if biogenesis of the mature miRNA has an effect on where it is localised.

In terms of ER-localised miRNA regulation, Stalder *et al*, showed that TRBP and Dicer localises to the ER membrane, and that active AGO2-loaded miRNAs associate with the ER membrane at the cytosolic membrane surface [351]. Furthermore, they showed that AGO2 associates with membranes through TBRP and that when the RISC loading complex (RLC) is bound to double stranded siRNA becomes enriched at the ER. Therefore, they propose RISC loading of AGO2 occurs at the ER, and that miRNA-mediated mRNA regulation occurs at the cytosolic side of the ER. Whilst the methods used in that paper and this thesis were indirect measures of miRNA presence in difference subcellular fractions, immunofluorescence *in situ* hybridisation (IF/ISH) could be used to detect and localise miRNA and RISC components in cells directly. This could be combined with Bodipy, an ER-tracking immunofluorescence dye, which would allow the comparison of miRNA levels and AGO association at the ER to be studied [515].

Previously in the Jopling lab, *in situ* hybridisation (ISH) was used to identify a subset of miR-122 in the nucleus of Huh7 cells but due to the small size of miRNA standard ISH would be unable to resolve the localisation of miR-122 at the chromatin level to corroborate the presence of miR-122 we detected using qRT-PCR. However, RNAscope® assay technology has been recently developed to detect individual microRNA at single-molecule sensitivity [516]. Unfortunately, this technology was not available during this PhD project but could be used to validate what was seen with the fractionation/qRT-PCR approach and for further investigation into miR-122 localisation in Huh7 cells.

6.8 Concluding Remarks

As there has been no direct comparison of miRNA function in different subcellular sites in human cells this project aimed to do so by applying subcellular fractionation methods and generating luciferase reporters to compare regulation between subcellular compartments.

In this thesis, a series of ER-translated luciferase reporters for miR-122 regulation at the 3' UTR and 5' UTR were successfully generated and their regulation by miR-122 was compared with that of equivalent reporters translated in the cytoplasm ultimately providing evidence for differential regulation via 3' UTR sites at the ER and cytoplasm in some, but not all, tested reporters. In addition, a membrane fractionation method was applied to compare the effects of miR-122 inhibition and overexpression on known miR-122 targets in ER-associated and cytoplasmic fractions, observing some differences in regulation between the ER and cytoplasm. Alternative methods for isolating the ER are available for future investigations, and a comparison of the global regulation of mRNA at the ER by miR-122 could be examined using techniques such as APEX-seq.

Furthermore, the presence of miR-122 in the chromatin was detected using a chromatin-associated fractionation method and CRISPR/Cas9n genome modification was used to investigate the possibility of miR-122 autoregulation at the chromatin. However, more work needs to be done to elucidate the role miR-122 may play in chromatin.

In future, microscopy-based techniques such as immunofluorescence *in situ* hybridisation (IF/ISH) or RNAscope® could be used to detect and localise miRNA and RISC components in cells directly allowing the comparison of miRNA levels and AGO association at the subcellular sites to be investigated.

Ultimately the investigations achieved in this thesis provide new understanding of the subcellular localisation of miR-122 in Huh7 cells, demonstrating differences in miR-122 regulation at the ER and cytoplasm, and generated tools for the further analysis of miR-122 activity at different subcellular sites.

References

1. Roeder, R.G., *The role of general initiation factors in transcription by RNA polymerase II*. Trends Biochem Sci, 1996. **21**(9): p. 327-35.
2. Grunberg, S. and S. Hahn, *Structural insights into transcription initiation by RNA polymerase II*. Trends Biochem Sci, 2013. **38**(12): p. 603-11.
3. Colgan, D.F. and J.L. Manley, *Mechanism and regulation of mRNA polyadenylation*. Genes Dev, 1997. **11**(21): p. 2755-66.
4. Proudfoot, N.J. and G.G. Brownlee, *3' non-coding region sequences in eukaryotic messenger RNA*. Nature, 1976. **263**(5574): p. 211-4.
5. Mandel, C.R., Y. Bai, and L. Tong, *Protein factors in pre-mRNA 3'-end processing*. Cell Mol Life Sci, 2008. **65**(7-8): p. 1099-122.
6. Brown, K.M. and G.M. Gilmartin, *A mechanism for the regulation of pre-mRNA 3' processing by human cleavage factor Im*. Mol Cell, 2003. **12**(6): p. 1467-76.
7. Terns, M.P. and S.T. Jacob, *Role of poly(A) polymerase in the cleavage and polyadenylation of mRNA precursor*. Mol Cell Biol, 1989. **9**(4): p. 1435-44.
8. Kuhn, U., et al., *Poly(A) tail length is controlled by the nuclear poly(A)-binding protein regulating the interaction between poly(A) polymerase and the cleavage and polyadenylation specificity factor*. J Biol Chem, 2009. **284**(34): p. 22803-14.
9. Filipowicz, W., et al., *A protein binding the methylated 5'-terminal sequence, m7GpppN, of eukaryotic messenger RNA*. Proc Natl Acad Sci U S A, 1976. **73**(5): p. 1559-63.
10. Shatkin, A.J., *Capping of eucaryotic mRNAs*. Cell, 1976. **9**(4 PT 2): p. 645-53.
11. Izaurrealde, E., et al., *A nuclear cap binding protein complex involved in pre-mRNA splicing*. Cell, 1994. **78**(4): p. 657-68.
12. Lewis, J.D., et al., *A nuclear cap-binding complex facilitates association of U1 snRNP with the cap-proximal 5' splice site*. Genes Dev, 1996. **10**(13): p. 1683-98.
13. Izaurrealde, E., et al., *A cap binding protein that may mediate nuclear export of RNA polymerase II-transcribed RNAs*. J Cell Biol, 1992. **118**(6): p. 1287-95.
14. Flaherty, S.M., et al., *Participation of the nuclear cap binding complex in pre-mRNA 3' processing*. Proc Natl Acad Sci U S A, 1997. **94**(22): p. 11893-8.
15. Herzelt, L., et al., *Splicing and transcription touch base: co-transcriptional spliceosome assembly and function*. Nat Rev Mol Cell Biol, 2017. **18**(10): p. 637-650.
16. Carmody, S.R. and S.R. Wenthe, *mRNA nuclear export at a glance*. J Cell Sci, 2009. **122**(Pt 12): p. 1933-7.
17. Grifo, J.A., et al., *New initiation factor required for globin mRNA translation*. The Journal of Biological Chemistry, 1982. **258**(9): p. 5804-5810.
18. Sonenberg, N., et al., *Eukaryotic mRNA cap binding protein: purification by affinity chromatography on sepharose-coupled m7GDP*. Proc. Natl. Acad. Sci. USA, 1979. **76**(9): p. 4345-4349.
19. Liu, W., et al., *Structural basis for nematode eIF4E binding an m(2,2,7)G-Cap and its implications for translation initiation*. Nucleic Acids Res, 2011. **39**(20): p. 8820-32.
20. Altmann, M., et al., *The Saccharomyces cerevisiae translation initiation factor Tif3 and its mammalian homologue, eIF-4B, have RNA annealing activity*. EMBO J., 1995. **14**(15): p. 3820-3827.

21. Korneeva, N.L., et al., *Interaction between the NH2-terminal domain of eIF4A and the central domain of eIF4G modulates RNA-stimulated ATPase activity.* J Biol Chem, 2005. **280**(3): p. 1872-81.
22. Hentze, M.W., *eIF4G: a multipurpose ribosome adaptor.* Science, 1997. **275**: p. 500-501.
23. Deo, R.C., et al., *Recognition of polyadenylate RNA by the poly(A)-binding protein.* Cell, 1999. **98**: p. 835-845.
24. Marcotrigiano, J., et al., *A conserved HEAT domain within eIF4G directs assembly of the translation initiation machinery.* Molecular Cell, 2001. **7**: p. 193-203.
25. Tarun, S.Z. and A.B. Sachs, *Association of the yeast poly(A) tail binding protein with translation initiation factor eIF-4G.* EMBO J., 1996. **15**(24): p. 7168-7177.
26. Wells, S.E., et al., *Circularization of mRNA by eukaryotic translation initiation factors.* Molecular Cell, 1998. **2**: p. 135-140.
27. Derry, M.C., et al., *Regulation of Poly(A)-binding protein through PABP-interacting proteins.* Cold Spring Harb Symp Quant Biol, 2006. **71**: p. 537-543.
28. Tarun, S.Z., et al., *Translation initiation factor eIF4G mediates in vitro poly(A) tail-dependent translation.* Proc Natl Acad Sci U S A, 1997. **94**: p. 9046-9051.
29. Merrick, W.C., *Cap-dependent and cap-independent translation in eukaryotic systems.* Gene, 2004. **332**: p. 1-11.
30. Shin, B., et al., *Uncoupling of initiation factor eIF5B/IF2 GTPase and translational activities by mutations that lower ribosome affinity.* Cell, 2002. **111**: p. 1015-1025.
31. Kozak, M., *Pushing the limits of the scanning mechanism for initiation of translation.* Gene, 2002. **299**: p. 1-34.
32. Marintchev, A., et al., *Topology and regulation of the human eIF4A/4G/4H helicase complex in translation initiation.* Cell, 2009. **136**(3): p. 447-60.
33. Siridechadilok, B., et al., *Structural roles for human translation factor eIF3 in initiation of protein synthesis.* Science, 2005. **310**(5753): p. 1513-5.
34. Huang, H., et al., *GTP hydrolysis controls stringent selection of the AUG start codon during translation initiation in Saccharomyces cerevisiae.* Genes Dev, 1997. **11**: p. 2396-2413.
35. Schmitt, E., B. S., and Y. Mechulam, *The large subunit of initiation factor aiF2 is a close structural homologue of elongation factors.* EMBO 2002. **21**(7): p. 1821-1832.
36. Zuk, D. and A. Jacobson, *A single amino acid substitution in yeast eIF-5A results in mRNA stabilization.* EMBO J., 1998. **17**(10): p. 2914-2925.
37. Schneider-Poetsch, T., et al., *Garbled messages and corrupted translations.* Nat Chem Biol, 2010. **6**(3): p. 189-198.
38. Dever, T.E. and R. Green, *The Elongation, Termination and Recycling Phases of Translation in Eukaryotes.* Cold Spring Harbor perspectives in biology, 2012. **4**(7): p. a013706-a013706.
39. Pelletier, J. and N. Sonenberg, *Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA.* Nature, 1988. **334**: p. 320-325.
40. Jang, S.K., et al., *A segment of the 5' nontranslated region of encephalomyocarditis virus RNA directs internal entry of ribosomes during In Vitro translation.* J Vir, 1988. **62**(8): p. 2636-2643.
41. Pelletier, J. and N. Sonenberg, *Internal binding of eukaryotic ribosomes on poliovirus RNA: Translation in HeLa cell extracts.* J Vir, 1989. **63**(1): p. 441-444.
42. Pisarev, A.V., N.E. Shirokikh, and C.U. Hellen, *Translation initiation by factor-independent binding of eukaryotic ribosomes to internal ribosomal entry sites.* C R Biol, 2005. **328**(7): p. 589-605.

43. Kolupaeva, V.G., et al., *Translation eukaryotic initiation factor 4G recognises a specific structural element within the internal ribosome entry site of encephalomyocarditis virus RNA*. J Biol Chem, 1998. **273**(29): p. 18599-18604.
44. Johannes, G., et al., *Identification of eukaryotic mRNAs that are translated at reduced cap binding complex eIF4F concentrations using a cDNA microarray*. PNAS, 1999. **96**(23): p. 13118-13123.
45. Hellen, C.U.T. and P. Sarnow, *Internal ribosome entry site in eukaryotic mRNA molecules*. Genes Dev, 2001. **15**: p. 1593-1612.
46. Komar, A.A. and M. Hatzoglou, *Cellular IRES-mediated translation: the war of ITAFs in pathophysiological states*. Cell Cycle, 2011. **10**(2): p. 229-40.
47. Godet, A.C., et al., *IRES Trans-Acting Factors, Key Actors of the Stress Response*. Int J Mol Sci, 2019. **20**(4).
48. Mailliot, J. and F. Martin, *Viral internal ribosomal entry sites: four classes for one goal*. Wiley Interdiscip Rev RNA, 2018. **9**(2).
49. Walter, P. and G. Blobel, *Translocation of proteins across the endoplasmic reticulum. II. Signal recognition protein (SRP) mediates the selective binding to microsomal membranes of in-vitro-assembled polysomes synthesizing secretory protein*. J Cell Biol, 1981. **91**(2 Pt 1): p. 551-6.
50. Williams, E.J., C. Pal, and L.D. Hurst, *The molecular evolution of signal peptides*. Gene, 2000. **253**(2): p. 313-22.
51. Meyer, D.I., E. Krause, and B. Dobberstein, *Secretory protein translocation across membranes-the role of the "docking protein"*. Nature, 1982. **297**(5868): p. 647-50.
52. Gorlich, D., et al., *A mammalian homolog of SEC61p and SECYp is associated with ribosomes and nascent polypeptides during translocation*. Cell, 1992. **71**(3): p. 489-503.
53. Osborne, A.R., T.A. Rapoport, and B. van den Berg, *Protein translocation by the Sec61/SecY channel*. Annu Rev Cell Dev Biol, 2005. **21**: p. 529-50.
54. Sanders, S.L., et al., *Sec61p and BiP directly facilitate polypeptide translocation into the ER*. Cell, 1992. **69**(2): p. 353-65.
55. Hessa, T., et al., *Recognition of transmembrane helices by the endoplasmic reticulum translocon*. Nature, 2005. **433**(7024): p. 377-81.
56. Hegde, R.S. and H.D. Bernstein, *The surprising complexity of signal sequences*. Trends Biochem Sci, 2006. **31**(10): p. 563-71.
57. Matsuoka, K., et al., *COPII-coated vesicle formation reconstituted with purified coat proteins and chemically defined liposomes*. Cell, 1998. **93**(2): p. 263-75.
58. Lee, M.C., et al., *Bi-directional protein transport between the ER and Golgi*. Annu Rev Cell Dev Biol, 2004. **20**: p. 87-123.
59. Barlowe, C., et al., *COPII: a membrane coat formed by Sec proteins that drive vesicle budding from the endoplasmic reticulum*. Cell, 1994. **77**(6): p. 895-907.
60. Jackson, C.L., *Mechanisms of transport through the Golgi complex*. J Cell Sci, 2009. **122**(Pt 4): p. 443-52.
61. Xu, D. and J.C. Hay, *Reconstitution of COPII vesicle fusion to generate a pre-Golgi intermediate compartment*. J Cell Biol, 2004. **167**(6): p. 997-1003.
62. Hinnebusch, A.G. and J.R. Lorsch, *The mechanism of eukaryotic translation initiation: new insights and challenges*. Cold Spring Harb Perspect Biol, 2012. **4**(10).
63. Pavitt, G.D. and D. Ron, *New insights into translational regulation in the endoplasmic reticulum unfolded protein response*. Cold Spring Harb Perspect Biol, 2012. **4**(6).
64. Roux, P.P. and I. Topisirovic, *Regulation of mRNA translation by signaling pathways*. Cold Spring Harb Perspect Biol, 2012. **4**(11).

65. Uttam, S., et al., *eIF4E-Dependent Translational Control: A Central Mechanism for Regulation of Pain Plasticity*. *Front Genet*, 2018. **9**: p. 470.
66. Saxton, R.A. and D.M. Sabatini, *mTOR Signaling in Growth, Metabolism, and Disease*. *Cell*, 2017. **168**(6): p. 960-976.
67. Kessler, S.H. and A.B. Sachs, *RNA recognition motif 2 of yeast Pab1p is required for its functional interaction with eukaryotic translation initiation factor 4G*. *Mol Cell Biol*, 1998. **18**(1): p. 51-7.
68. Imataka, H., A. Gradi, and N. Sonenberg, *A newly identified N-terminal amino acid sequence of human eIF4G binds poly(A)-binding protein and functions in poly(A)-dependent translation*. *EMBO J*, 1998. **17**(24): p. 7480-9.
69. Cakmakci, N.G., et al., *SLIP1, a factor required for activation of histone mRNA translation by the stem-loop binding protein*. *Mol Cell Biol*, 2008. **28**(3): p. 1182-94.
70. Sonenberg, N. and A.G. Hinnebusch, *Regulation of translation initiation in eukaryotes: mechanisms and biological targets*. *Cell*, 2009. **136**(4): p. 731-45.
71. Minshall, N., et al., *CPEB interacts with an ovary-specific eIF4E and 4E-T in early Xenopus oocytes*. *J Biol Chem*, 2007. **282**(52): p. 37389-401.
72. Villaescusa, J.C., et al., *Cytoplasmic Prep1 interacts with 4EHP inhibiting Hoxb4 translation*. *PLoS One*, 2009. **4**(4): p. e5213.
73. Price, N.T., et al., *Identification of the phosphorylation sites in elongation factor-2 from rabbit reticulocytes*. *FEBS Lett*, 1991. **282**(2): p. 253-8.
74. Ovchinnikov, L.P., et al., *Three phosphorylation sites in elongation factor 2*. *FEBS Lett*, 1990. **275**(1-2): p. 209-12.
75. Hizli, A.A., et al., *Phosphorylation of eukaryotic elongation factor 2 (eEF2) by cyclin A-cyclin-dependent kinase 2 regulates its inhibition by eEF2 kinase*. *Mol Cell Biol*, 2013. **33**(3): p. 596-604.
76. Jacks, T. and H.E. Varmus, *Expression of the Rous sarcoma virus pol gene by ribosomal frameshifting*. *Science*, 1985. **230**(4731): p. 1237-42.
77. Firth, A.E. and I. Brierley, *Non-canonical translation in RNA viruses*. *J Gen Virol*, 2012. **93**(Pt 7): p. 1385-1409.
78. Manktelow, E., K. Shigemoto, and I. Brierley, *Characterization of the frameshift signal of Edr, a mammalian example of programmed -1 ribosomal frameshifting*. *Nucleic Acids Res*, 2005. **33**(5): p. 1553-63.
79. Wills, N.M., et al., *A functional -1 ribosomal frameshift signal in the human paraneoplastic Ma3 gene*. *J Biol Chem*, 2006. **281**(11): p. 7082-8.
80. Benham, A.M., *Protein secretion and the endoplasmic reticulum*. *Cold Spring Harb Perspect Biol*, 2012. **4**(8): p. a012872.
81. Zhang, G., M. Hubalewska, and Z. Ignatova, *Transient ribosomal attenuation coordinates protein synthesis and co-translational folding*. *Nat Struct Mol Biol*, 2009. **16**(3): p. 274-80.
82. Cabrita, L.D., C.M. Dobson, and J. Christodoulou, *Protein folding on the ribosome*. *Curr Opin Struct Biol*, 2010. **20**(1): p. 33-45.
83. Dever, T.E., J.D. Dinman, and R. Green, *Translation Elongation and Recoding in Eukaryotes*. *Cold Spring Harb Perspect Biol*, 2018. **10**(8).
84. Braun, J.E., et al., *GW182 proteins directly recruit cytoplasmic deadenylase complexes to miRNA targets*. *Mol Cell*, 2011. **44**(1): p. 120-33.
85. Chen, C.Y., et al., *Ago-TNRC6 triggers microRNA-mediated decay by promoting two deadenylation steps*. *Nat Struct Mol Biol*, 2009. **16**(11): p. 1160-6.

86. Brown, C.E. and A.B. Sachs, *Poly(A) tail length control in Saccharomyces cerevisiae occurs by message-specific deadenylation*. Molecular and Cellular Biology, 1998. **18**(11): p. 6548-6559.
87. Yamashita, A., et al., *Concerted action of poly(A) nucleases and decapping enzyme in mammalian mRNA turnover*. Nat Struct Mol Biol, 2005. **12**(12): p. 1054-63.
88. Zheng, D., et al., *Deadenylation is prerequisite for P-body formation and mRNA decay in mammalian cells*. J Cell Biol, 2008. **182**(1): p. 89-101.
89. Webster, M.W., et al., *mRNA Deadenylation Is Coupled to Translation Rates by the Differential Activities of Ccr4-Not Nucleases*. Mol Cell, 2018. **70**(6): p. 1089-1100 e8.
90. Sun, M., et al., *Global analysis of eukaryotic mRNA degradation reveals Xrn1-dependent buffering of transcript levels*. Mol Cell, 2013. **52**(1): p. 52-62.
91. Pavanello, L., et al., *The central region of CNOT1 and CNOT9 stimulates deadenylation by the Ccr4-Not nuclease module*. Biochem J, 2018. **475**(21): p. 3437-3450.
92. Maryati, M., B. Airhihen, and G.S. Winkler, *The enzyme activities of Caf1 and Ccr4 are both required for deadenylation by the human Ccr4-Not nuclease module*. Biochem J, 2015. **469**(1): p. 169-76.
93. Steiger, M., et al., *Analysis of recombinant yeast decapping enzyme*. RNA, 2003. **9**(2): p. 231-8.
94. Piccirillo, C., R. Khanna, and M. Kiledjian, *Functional characterization of the mammalian mRNA decapping enzyme hDcp2*. RNA, 2003. **9**(9): p. 1138-47.
95. van Dijk, E., et al., *Human Dcp2: a catalytically active mRNA decapping enzyme located in specific cytoplasmic structures*. EMBO J, 2002. **21**(24): p. 6915-24.
96. Li, Y. and M. Kiledjian, *Regulation of mRNA decapping*. Wiley Interdiscip Rev RNA, 2010. **1**(2): p. 253-65.
97. Ozgur, S., M. Chekulaeva, and G. Stoecklin, *Human Pat1b connects deadenylation with mRNA decapping and controls the assembly of processing bodies*. Mol Cell Biol, 2010. **30**(17): p. 4308-23.
98. Pilkington, G.R. and R. Parker, *Pat1 contains distinct functional domains that promote P-body assembly and activation of decapping*. Mol Cell Biol, 2008. **28**(4): p. 1298-312.
99. Chang, J.H., S. Xiang, and L. Tong, *Structures of 5'-3' Exoribonucleases*. Enzymes, 2012. **31**: p. 115-29.
100. Collier, J. and R. Parker, *Eukaryotic mRNA decapping*. Annu Rev Biochem, 2004. **73**: p. 861-90.
101. Till, D.D., et al., *Identification and developmental expression of a 5'-3' exoribonuclease from Drosophila melanogaster*. Mech Dev, 1998. **79**(1-2): p. 51-5.
102. Sato, Y., et al., *Cloning and characterization of human Sep1 (hSEP1) gene and cytoplasmic localization of its product*. DNA Res, 1998. **5**(4): p. 241-6.
103. Bashkirov, V.I., et al., *A mouse cytoplasmic exoribonuclease (mXRN1p) with preference for G4 tetraplex substrates*. J Cell Biol, 1997. **136**(4): p. 761-73.
104. Shobuike, T., et al., *Characterization of cDNA encoding mouse homolog of fission yeast dhp1+ gene: structural and functional conservation*. Nucleic Acids Res, 1995. **23**(3): p. 357-61.
105. Nagarajan, V.K., et al., *XRN 5'-->3' exoribonucleases: structure, mechanisms and functions*. Biochim Biophys Acta, 2013. **1829**(6-7): p. 590-603.
106. Anderson, J.S. and R.P. Parker, *The 3' to 5' degradation of yeast mRNAs is a general mechanism for mRNA turnover that requires the SKI2 DEVH box protein and 3' to 5' exonucleases of the exosome complex*. EMBO J, 1998. **17**(5): p. 1497-506.

107. Wang, Z. and M. Kiledjian, *Functional link between the mammalian exosome and mRNA decapping*. Cell, 2001. **107**(6): p. 751-62.
108. Liu, Q., J.C. Greimann, and C.D. Lima, *Reconstitution, activities, and structure of the eukaryotic RNA exosome*. Cell, 2006. **127**(6): p. 1223-37.
109. Lebreton, A., et al., *Endonucleolytic RNA cleavage by a eukaryotic exosome*. Nature, 2008. **456**(7224): p. 993-6.
110. Schaeffer, D., et al., *The exosome contains domains with specific endoribonuclease, exoribonuclease and cytoplasmic mRNA decay activities*. Nat Struct Mol Biol, 2009. **16**(1): p. 56-62.
111. Schneider, C., et al., *The N-terminal PIN domain of the exosome subunit Rrp44 harbors endonuclease activity and tethers Rrp44 to the yeast core exosome*. Nucleic Acids Res, 2009. **37**(4): p. 1127-40.
112. Liu, H., et al., *The scavenger mRNA decapping enzyme DcpS is a member of the HIT family of pyrophosphatases*. EMBO J, 2002. **21**(17): p. 4699-708.
113. Houseley, J., J. LaCava, and D. Tollervey, *RNA-quality control by the exosome*. Nat Rev Mol Cell Biol, 2006. **7**(7): p. 529-39.
114. Schmid, M. and T.H. Jensen, *The exosome: a multipurpose RNA-decay machine*. Trends Biochem Sci, 2008. **33**(10): p. 501-10.
115. Lebreton, A. and B. Seraphin, *Exosome-mediated quality control: substrate recruitment and molecular activity*. Biochim Biophys Acta, 2008. **1779**(9): p. 558-65.
116. Yang, F. and D.R. Schoenberg, *Endonuclease-mediated mRNA decay involves the selective targeting of PMR1 to polyribosome-bound substrate mRNA*. Mol Cell, 2004. **14**(4): p. 435-45.
117. Chernokalskaya, E., et al., *A polysomal ribonuclease involved in the destabilization of albumin mRNA is a novel member of the peroxidase gene family*. RNA, 1998. **4**(12): p. 1537-48.
118. Hollien, J. and J.S. Weissman, *Decay of endoplasmic reticulum-localized mRNAs during the unfolded protein response*. Science, 2006. **313**(5783): p. 104-7.
119. Yoshida, H., et al., *XBP1 mRNA is induced by ATF6 and spliced by IRE1 in response to ER stress to produce a highly active transcription factor*. Cell, 2001. **107**(7): p. 881-91.
120. Conti, E. and E. Izaurralde, *Nonsense-mediated mRNA decay: molecular insights and mechanistic variations across species*. Curr Opin Cell Biol, 2005. **17**(3): p. 316-25.
121. Garcia-Maurino, S.M., et al., *RNA Binding Protein Regulation and Cross-Talk in the Control of AU-rich mRNA Fate*. Front Mol Biosci, 2017. **4**: p. 71.
122. Andrei, M.A., et al., *A role for eIF4E and eIF4E-transporter in targeting mRNPs to mammalian processing bodies*. RNA, 2005. **11**(5): p. 717-27.
123. Cougot, N., S. Babajko, and B. Seraphin, *Cytoplasmic foci are sites of mRNA decay in human cells*. J Cell Biol, 2004. **165**(1): p. 31-40.
124. Lykke-Andersen, J., *Identification of a human decapping complex associated with hUpf proteins in nonsense-mediated decay*. Mol Cell Biol, 2002. **22**(23): p. 8114-21.
125. Ingelfinger, D., et al., *The human LSm1-7 proteins colocalize with the mRNA-degrading enzymes Dcp1/2 and Xrn1 in distinct cytoplasmic foci*. RNA, 2002. **8**(12): p. 1489-501.
126. Meister, G., et al., *Identification of novel argonaute-associated proteins*. Curr Biol, 2005. **15**(23): p. 2149-55.
127. Eystathioy, T., et al., *The GW182 protein colocalizes with mRNA degradation associated proteins hDcp1 and hLSm4 in cytoplasmic GW bodies*. RNA, 2003. **9**(10): p. 1171-3.
128. Sen, G.L. and H.M. Blau, *Argonaute 2/RISC resides in sites of mammalian mRNA decay known as cytoplasmic bodies*. Nat Cell Biol, 2005. **7**(6): p. 633-6.

129. Liu, J., et al., *MicroRNA-dependent localization of targeted mRNAs to mammalian P-bodies*. Nat Cell Biol, 2005. **7**(7): p. 719-23.
130. Decker, C.J. and R. Parker, *P-bodies and stress granules: possible roles in the control of translation and mRNA degradation*. Cold Spring Harb Perspect Biol, 2012. **4**(9): p. a012286.
131. Bhattacharyya, S.N., et al., *Relief of microRNA-mediated translational repression in human cells subjected to stress*. Cell, 2006. **125**(6): p. 1111-24.
132. Brengues, M., D. Teixeira, and R. Parker, *Movement of eukaryotic mRNAs between polysomes and cytoplasmic processing bodies*. Science, 2005. **310**(5747): p. 486-9.
133. Malone, C.D. and G.J. Hannon, *Molecular evolution of piRNA and transposon control pathways in Drosophila*. Cold Spring Harb Symp Quant Biol, 2009. **74**: p. 225-34.
134. Lee, R.C., R.L. Feinbaum, and V. Ambros, *The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14*. Cell, 1993. **75**: p. 843-854.
135. Reinhart, B.J., et al., *The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans*. Nature, 2000. **403**: p. 901-906.
136. Ambros, V., *microRNAs: Tiny regulators with great potential*. Cell, 2001. **107**: p. 823-826.
137. Lau, N.C., et al., *An abundant class of tiny RNAs with probable regulatory roles in Caenorhabditis elegans*. Science, 2001. **294**(5543): p. 858-862.
138. Lee, R.C. and V. Ambros, *An extensive class of small RNAs in Caenorhabditis elegans*. Science, 2001. **294**(5543): p. 862-864.
139. Hammond, S.M., *An overview of microRNAs*. Adv Drug Deliv Rev, 2015. **87**: p. 3-14.
140. Kozomara, A. and S. Griffiths-Jones, *miRBase: annotating high confidence microRNAs using deep sequencing data*. Nucleic Acids Res, 2014. **42**(Database issue): p. D68-73.
141. Selbach, M., et al., *Widespread changes in protein synthesis induced by microRNAs*. Nature, 2008. **455**(7209): p. 58-63.
142. Lim, L.P., et al., *Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs*. Nature, 2005. **433**: p. 769-733.
143. Wu, L. and J.G. Belasco, *Micro-RNA Regulation of the Mammalian lin-28 Gene during Neuronal Differentiation of Embryonal Carcinoma Cells*. Molecular and Cellular Biology, 2005. **25**(21): p. 9198-9208.
144. Meister, G. and T. Tuschl, *Mechanisms of gene silencing by double-stranded RNA*. Nature, 2004. **431**(7006): p. 343-9.
145. Hafner, M., et al., *Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP*. Cell, 2010. **141**(1): p. 129-41.
146. Lytle, J.R., T.A. Yario, and J.A. Steitz, *Target mRNAs are repressed as efficiently by microRNA-binding sites in the 5' UTR as in the 3' UTR*. Proc Natl Acad Sci U S A, 2007. **104**(23): p. 9667-72.
147. Xu, W., et al., *Identifying microRNA targets in different gene regions*. BMC Bioinformatics, 2014. **15 Suppl 7**: p. S4.
148. Doench, J.G. and P.A. Sharp, *Specificity of microRNA target selection in translational repression*. Genes Dev, 2004. **18**(5): p. 504-11.
149. Mello, C.C. and D. Conte, Jr., *Revealing the world of RNA interference*. Nature, 2004. **431**(7006): p. 338-42.
150. Lippman, Z. and R. Martienssen, *The role of RNA interference in heterochromatic silencing*. Nature, 2004. **431**(7006): p. 364-70.
151. Golden, D.E., V.R. Gerbasi, and E.J. Sontheimer, *An inside job for siRNAs*. Mol Cell, 2008. **31**(3): p. 309-12.

152. Vazquez, F., et al., *Endogenous trans-acting siRNAs regulate the accumulation of Arabidopsis mRNAs*. Mol Cell, 2004. **16**(1): p. 69-79.
153. Ryther, R.C., et al., *siRNA therapeutics: big potential from small RNAs*. Gene Ther, 2005. **12**(1): p. 5-11.
154. Chang, J., et al., *miR-122, a mammalian liver-specific microRNA, is processed from hcr mRNA and may downregulate the high affinity cationic amino acid transporter CAT-1*. RNA Biol, 2004. **1**(2): p. 106-13.
155. Roush, S. and F.J. Slack, *The let-7 family of microRNAs*. Trends Cell Biol, 2008. **18**(10): p. 505-16.
156. Bentwich, I., et al., *Identification of hundreds of conserved and nonconserved human microRNAs*. Nat Genet, 2005. **37**(7): p. 766-70.
157. Sempere, L.F., et al., *The phylogenetic distribution of metazoan microRNAs: insights into evolutionary complexity and constraint*. J Exp Zool B Mol Dev Evol, 2006. **306**(6): p. 575-88.
158. Massirer, K.B. and A.E. Pasquinelli, *The evolving role of microRNAs in animal gene expression*. Bioessays, 2006. **28**(5): p. 449-52.
159. Stark, A., et al., *Animal MicroRNAs confer robustness to gene expression and have a significant impact on 3'UTR evolution*. Cell, 2005. **123**(6): p. 1133-46.
160. Baek, D., et al., *The impact of microRNAs on protein output*. Nature, 2008. **455**(7209): p. 64-71.
161. Ebert, M.S. and P.A. Sharp, *Roles for microRNAs in conferring robustness to biological processes*. Cell, 2012. **149**(3): p. 515-24.
162. Plasterk, R.H., *Micro RNAs in animal development*. Cell, 2006. **124**(5): p. 877-81.
163. Yoo, A.S. and I. Greenwald, *LIN-12/Notch activation leads to microRNA-mediated down-regulation of Vav in C. elegans*. Science, 2005. **310**(5752): p. 1330-3.
164. Hornstein, E., et al., *The microRNA miR-196 acts upstream of Hoxb8 and Shh in limb development*. Nature, 2005. **438**(7068): p. 671-4.
165. Niwa, R. and F.J. Slack, *The evolution of animal microRNA function*. Curr Opin Genet Dev, 2007. **17**(2): p. 145-50.
166. Nguyen, H.T. and M. Frasch, *MicroRNAs in muscle differentiation: lessons from Drosophila and beyond*. Curr Opin Genet Dev, 2006. **16**(5): p. 533-9.
167. Ason, B., et al., *Differences in vertebrate microRNA expression*. Proc Natl Acad Sci U S A, 2006. **103**(39): p. 14385-9.
168. Mann, M., et al., *An NF-kappaB-microRNA regulatory network tunes macrophage inflammatory responses*. Nat Commun, 2017. **8**(1): p. 851.
169. Taganov, K.D., et al., *NF-kappaB-dependent induction of microRNA miR-146, an inhibitor targeted to signaling proteins of innate immune responses*. Proc Natl Acad Sci U S A, 2006. **103**(33): p. 12481-6.
170. He, Y., et al., *MiR-146a regulates IL-6 production in lipopolysaccharide-induced RAW264.7 macrophage cells by inhibiting Notch1*. Inflammation, 2014. **37**(1): p. 71-82.
171. Lu, J., et al., *MicroRNA expression profiles classify human cancers*. Nature, 2005. **435**(7043): p. 834-8.
172. Calin, G.A., et al., *Frequent deletions and down-regulation of micro- RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia*. Proc Natl Acad Sci U S A, 2002. **99**(24): p. 15524-9.
173. Cimmino, A., et al., *miR-15 and miR-16 induce apoptosis by targeting BCL2*. Proc Natl Acad Sci U S A, 2005. **102**(39): p. 13944-9.

174. Calin, G.A., et al., *MiR-15a and miR-16-1 cluster functions in human leukemia*. Proc Natl Acad Sci U S A, 2008. **105**(13): p. 5166-71.
175. Feng, Z., et al., *Tumor suppressor p53 meets microRNAs*. J Mol Cell Biol, 2011. **3**(1): p. 44-50.
176. Chang, T.C., et al., *Widespread microRNA repression by Myc contributes to tumorigenesis*. Nat Genet, 2008. **40**(1): p. 43-50.
177. Fazi, F., et al., *Epigenetic silencing of the myelopoiesis regulator microRNA-223 by the AML1/ETO oncoprotein*. Cancer Cell, 2007. **12**(5): p. 457-66.
178. Walz, A.L., et al., *Recurrent DGCR8, DROSHA, and SIX homeodomain mutations in favorable histology Wilms tumors*. Cancer Cell, 2015. **27**(2): p. 286-97.
179. Melo, S.A., et al., *A genetic defect in exportin-5 traps precursor microRNAs in the nucleus of cancer cells*. Cancer Cell, 2010. **18**(4): p. 303-15.
180. Pampalakis, G., et al., *Down-regulation of dicer expression in ovarian cancer tissues*. Clin Biochem, 2010. **43**(3): p. 324-7.
181. Zhu, D.X., et al., *Downregulated Dicer expression predicts poor prognosis in chronic lymphocytic leukemia*. Cancer Sci, 2012. **103**(5): p. 875-81.
182. Karube, Y., et al., *Reduced expression of Dicer associated with poor prognosis in lung cancer patients*. Cancer Sci, 2005. **96**(2): p. 111-5.
183. Volinia, S., et al., *Reprogramming of miRNA networks in cancer and leukemia*. Genome Res, 2010. **20**(5): p. 589-99.
184. Pan, X., Z.X. Wang, and R. Wang, *MicroRNA-21: a novel therapeutic target in human cancer*. Cancer Biol Ther, 2010. **10**(12): p. 1224-32.
185. O'Donnell, K.A., et al., *c-Myc-regulated microRNAs modulate E2F1 expression*. Nature, 2005. **435**(7043): p. 839-43.
186. Dews, M., et al., *Augmentation of tumor angiogenesis by a Myc-activated microRNA cluster*. Nat Genet, 2006. **38**(9): p. 1060-5.
187. Olive, V., et al., *miR-19 is a key oncogenic component of mir-17-92*. Genes Dev, 2009. **23**(24): p. 2839-49.
188. Hong, L., et al., *The miR-17-92 cluster of microRNAs confers tumorigenicity by inhibiting oncogene-induced senescence*. Cancer Res, 2010. **70**(21): p. 8547-57.
189. Cai, X., C.H. Hagedorn, and B.R. Cullen, *Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs*. RNA, 2004. **10**: p. 1957-1966.
190. Lee, Y., et al., *MicroRNA genes are transcribed by RNA polymerase II*. EMBO, 2004. **23**: p. 4051-4060.
191. Bortolin-Cavaille, M.L., et al., *C19MC microRNAs are processed from introns of large Pol-II, non-protein-coding transcripts*. Nucleic Acids Res, 2009. **37**(10): p. 3464-73.
192. Poliseno, L., et al., *Identification of the miR-106b~25 microRNA cluster as a proto-oncogenic PTEN-targeting intron that cooperates with its host gene MCM7 in transformation*. Sci Signal, 2010. **3**(117): p. ra29.
193. Ventura, A., et al., *Targeted deletion reveals essential and overlapping functions of the miR-17 through 92 family of miRNA clusters*. Cell, 2008. **132**(5): p. 875-86.
194. Lee, Y., et al., *MicroRNA maturation: stepwise processing and subcellular localisation*. EMBO, 2002. **21**(17): p. 4663-4670.
195. Morlando, M., et al., *Primary microRNA transcripts are processed co-transcriptionally*. Nature Structural & Molecular Biology, 2008. **15**(9): p. 902-909.
196. Gromak, N., et al., *Drosha regulates gene expression independently of RNA cleavage function*. Cell Rep, 2013. **5**(6): p. 1499-510.

197. Lee, Y., et al., *The nuclear RNase III Drosha initiates microRNA processing*. Nature, 2003. **425**: p. 415-419.
198. Han, J., et al., *The Drosha-DGCR8 complex in primary microRNA processing*. Genes Dev, 2004. **18**(24): p. 3016-27.
199. Han, J., et al., *Molecular basis for the recognition of primary microRNAs by the Drosha-DGCR8 complex*. Cell, 2006. **125**(5): p. 887-901.
200. Yi, R., et al., *Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs*. Genes Dev, 2003. **17**(24): p. 3011-6.
201. Bohnsack, M.T., K. Czaplinski, and D. Gorlich, *Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs*. RNA, 2004. **10**(2): p. 185-91.
202. Lund, E., et al., *Nuclear export of microRNA precursors*. Science, 2004. **303**(5654): p. 95-98.
203. Bernstein, E., et al., *Role for a didentate ribonuclease in the initiation step of RNA interference*. Nature, 2001. **409**: p. 363-366.
204. Hutvagner, G., et al., *A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA*. Science, 2001. **293**(5531): p. 834-838.
205. Ketting, R.F., et al., *Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in C. elegans*. Genes Dev, 2001. **15**(20): p. 2654-2659.
206. Khvorova, A., A. Reynolds, and S.D. Jayasena, *Functional siRNAs and miRNAs exhibit strand bias*. Cell, 2003. **115**: p. 209-216.
207. Gregory, R.I., et al., *Human RISC couples microRNA biogenesis and posttranscriptional gene silencing*. Cell, 2005. **123**(4): p. 631-40.
208. Martinez, J., et al., *Single-stranded antisense siRNAs guide target RNA cleavage in RNAi*. Cell, 2002. **110**: p. 563-574.
209. Ruby, J.G., C.H. Jan, and D.P. Bartel, *Intronic microRNA precursors that bypass Drosha processing*. Nature, 2007. **448**(7149): p. 83-86.
210. Xie, M., et al., *Mammalian 5'-capped microRNA precursors that generate a single microRNA*. Cell, 2013. **155**(7): p. 1568-80.
211. Cheloufi, S., et al., *A dicer-independent miRNA biogenesis pathway that requires Ago catalysis*. Nature, 2010. **465**(7298): p. 584-9.
212. Xie, Z., et al., *Expression of Arabidopsis MIRNA genes*. Plant Physiol, 2005. **138**(4): p. 2145-54.
213. Corcoran, D.L., et al., *Features of mammalian microRNA promoters emerge from polymerase II chromatin immunoprecipitation data*. PLoS One, 2009. **4**(4): p. e5279.
214. Megraw, M., et al., *MicroRNA promoter element discovery in Arabidopsis*. RNA, 2006. **12**(9): p. 1612-9.
215. Ma, L., et al., *miR-9, a MYC/MYCN-activated microRNA, regulates E-cadherin and cancer metastasis*. Nat Cell Biol, 2010. **12**(3): p. 247-56.
216. Ben-Ami, O., et al., *A regulatory interplay between miR-27a and Runx1 during megakaryopoiesis*. Proc Natl Acad Sci U S A, 2009. **106**(1): p. 238-43.
217. Gruber, J.J., et al., *Ars2 links the nuclear cap-binding complex to RNA interference and cell proliferation*. Cell, 2009. **138**(2): p. 328-39.
218. Morlando, M., et al., *FUS stimulates microRNA biogenesis by facilitating co-transcriptional Drosha recruitment*. EMBO J, 2012. **31**(24): p. 4502-10.
219. Tang, X., et al., *Phosphorylation of the RNase III enzyme Drosha at Serine300 or Serine302 is required for its nuclear localization*. Nucleic Acids Res, 2010. **38**(19): p. 6610-9.

220. Tang, X., et al., *Glycogen synthase kinase 3 beta (GSK3beta) phosphorylates the RNAase III enzyme Drosha at S300 and S302*. PLoS One, 2011. **6**(6): p. e20391.
221. Bail, S., et al., *Differential regulation of microRNA stability*. RNA, 2010. **16**(5): p. 1032-9.
222. Ding, X.C. and H. Grosshans, *Repression of C. elegans microRNA targets at the initiation level of translation requires GW182 proteins*. EMBO J, 2009. **28**(3): p. 213-22.
223. Kai, Z.S. and A.E. Pasquinelli, *MicroRNA assassins: factors that regulate the disappearance of miRNAs*. Nat Struct Mol Biol, 2010. **17**(1): p. 5-10.
224. Hata, A. and B.N. Davis, *Control of microRNA biogenesis by TGFbeta signaling pathway-A novel role of Smads in the nucleus*. Cytokine Growth Factor Rev, 2009. **20**(5-6): p. 517-21.
225. Lee, Y., et al., *MicroRNA maturation: stepwise processing and subcellular localization*. EMBO J, 2002. **21**(17): p. 4663-70.
226. Newman, M.A. and S.M. Hammond, *Emerging paradigms of regulated microRNA processing*. Genes Dev, 2010. **24**(11): p. 1086-92.
227. Lee, Y., et al., *The role of PACT in the RNA silencing pathway*. EMBO J, 2006. **25**(3): p. 522-32.
228. Slezak-Prochazka, I., et al., *MicroRNAs, macrocontrol: regulation of miRNA processing*. RNA, 2010. **16**(6): p. 1087-95.
229. Paroo, Z., et al., *Phosphorylation of the human microRNA-generating complex mediates MAPK/Erk signaling*. Cell, 2009. **139**(1): p. 112-22.
230. Chendrimada, T.P., et al., *TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing*. Nature, 2005. **436**(7051): p. 740-4.
231. Meister, G., et al., *Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs*. Mol Cell, 2004. **15**(2): p. 185-97.
232. Okamura, K., et al., *Distinct roles for Argonaute proteins in small RNA-directed RNA cleavage pathways*. Genes Dev, 2004. **18**(14): p. 1655-66.
233. Hu, Q., et al., *Dicer and Ago3-dependent generation of retinoic acid-induced DR2 Alu RNAs (riRNAs) regulates human stel cell proliferation*. Nat Struct Mol Biol, 2012. **19**(11): p. 1168-1175.
234. Winter, J. and S. Diederichs, *Argonaute-3 activates the let-7a passenger strand microRNA*. RNA Biol, 2013. **10**(10): p. 1631-43.
235. Modzelewski, A.J., et al., *AGO4 regulates entry into meiosis and influences silencing of sex chromosomes in the male mouse germline*. Dev Cell, 2012. **23**(2): p. 251-64.
236. Bartel, D.P., *MicroRNAs: target recognition and regulatory functions*. Cell, 2009. **136**(2): p. 215-33.
237. Czech, B. and G.J. Hannon, *Small RNA sorting: matchmaking for Argonautes*. Nat Rev Genet, 2011. **12**(1): p. 19-31.
238. Hock, J. and G. Meister, *The Argonaute protein family*. Genome Biol, 2008. **9**(2): p. 210.
239. Seok, H., et al., *MicroRNA Target Recognition: Insights from Transcriptome-Wide Non-Canonical Interactions*. Mol Cells, 2016. **39**(5): p. 375-81.
240. Wang, H.W., et al., *Structural insights into RNA processing by the human RISC-loading complex*. Nat Struct Mol Biol, 2009. **16**(11): p. 1148-53.
241. MacRae, I.J., et al., *In vitro reconstitution of the human RISC-loading complex*. Proc Natl Acad Sci U S A, 2008. **105**(2): p. 512-7.
242. Lee, H.Y. and J.A. Doudna, *TRBP alters human precursor microRNA processing in vitro*. RNA, 2012. **18**(11): p. 2012-9.
243. Chakravarthy, S., et al., *Substrate-Specific Kinetics of Dicer-Catalyzed RNA Processing*. Journal of Molecular Biology, 2010. **404**(3): p. 392-402.

244. Behm-Ansmant, I., *mRNA degradation by miRNAs and GW182 requires both CCR4:NOT deadenylase and DCP1:DCP2 decapping complexes*. *Genes & Development*, 2006. **20**(14): p. 1885-1898.
245. Liu, J., et al., *A role for the P-body component GW182 in microRNA function*. *Nat Cell Biol*, 2005. **7**(12): p. 1261-6.
246. Eulalio, A., E. Huntzinger, and E. Izaurralde, *GW182 interaction with Argonaute is essential for miRNA-mediated translational repression and mRNA decay*. *Nat Struct Mol Biol*, 2008. **15**(4): p. 346-53.
247. Jakymiw, A., et al., *Disruption of GW bodies impairs mammalian RNA interference*. *Nat Cell Biol*, 2005. **7**(12): p. 1267-74.
248. Schirle, N.T. and I.J. MacRae, *The crystal structure of human argonaute 2*. *Science*, 2012. **336**(6084): p. 1037-1040.
249. Pfaff, J., et al., *Structural features on argonaute-GW182 protein interactions*. *PNAS*, 2013. **110**(40): p. 3370-3779.
250. Hock, J., et al., *Proteomic and functional analysis of Argonaute-containing mRNA-protein complexes in human cells*. *EMBO Rep*, 2007. **8**(11): p. 1052-60.
251. Landthaler, M., et al., *Molecular characterization of human Argonaute-containing ribonucleoprotein complexes and their bound target mRNAs*. *RNA*, 2008. **14**(12): p. 2580-96.
252. Miles, W.O., et al., *Pumilio facilitates miRNA regulation of the E2F3 oncogene*. *Genes Dev*, 2012. **26**(4): p. 356-68.
253. James, V., et al., *LIM-domain proteins, LIMD1, Ajuba, and WTIP are required for microRNA-mediated gene silencing*. *Proc Natl Acad Sci U S A*, 2010. **107**(28): p. 12499-504.
254. Jannot, G., et al., *The ribosomal protein RACK1 is required for microRNA function in both C. elegans and humans*. *EMBO Rep*, 2011. **12**(6): p. 581-6.
255. Iwasaki, S., et al., *Hsc70/Hsp90 chaperone machinery mediates ATP-dependent RISC loading of small RNA duplexes*. *Mol Cell*, 2010. **39**(2): p. 292-9.
256. Johnston, M., et al., *HSP90 protein stabilizes unloaded argonaute complexes and microscopic P-bodies in human cells*. *Mol Biol Cell*, 2010. **21**(9): p. 1462-9.
257. La Rocca, G., et al., *In vivo, Argonaute-bound microRNAs exist predominantly in a reservoir of low molecular weight complexes not associated with mRNA*. *Proc Natl Acad Sci U S A*, 2015. **112**(3): p. 767-72.
258. Olejniczak, S.H., et al., *Long-lived microRNA-Argonaute complexes in quiescent cells can be activated to regulate mitogenic responses*. *Proc Natl Acad Sci U S A*, 2013. **110**(1): p. 157-62.
259. Christie, M., et al., *Structure of the PAN3 pseudokinase reveals the basis for interactions with the PAN2 deadenylase and the GW182 proteins*. *Mol Cell*, 2013. **51**(3): p. 360-73.
260. Ding, L., et al., *The developmental timing regulator AIN-1 interacts with miRISCs and may target the argonaute protein ALG-1 to cytoplasmic P bodies in C. elegans*. *Mol Cell*, 2005. **19**(4): p. 437-47.
261. Rehwinkel, J., et al., *A crucial role for GW182 and the DCP1:DCP2 decapping complex in miRNA-mediated gene silencing*. *RNA*, 2005. **11**(11): p. 1640-7.
262. Jonas, S. and E. Izaurralde, *Towards a molecular understanding of microRNA-mediated gene silencing*. *Nat Rev Genet*, 2015. **16**(7): p. 421-33.
263. Braun, J.E., E. Huntzinger, and E. Izaurralde, *A molecular link between miRISCs and deadenylases provides new insight into the mechanism of gene silencing by microRNAs*. *Cold Spring Harb Perspect Biol*, 2012. **4**(12).

264. Fabian, M.R., et al., *Mammalian miRNA RISC recruits CAF1 and PABP to affect PABP-dependent deadenylation*. Mol Cell, 2009. **35**(6): p. 868-80.
265. Chekulaeva, M., et al., *miRNA repression involves GW182-mediated recruitment of CCR4-NOT through conserved W-containing motifs*. Nat Struct Mol Biol, 2011. **18**(11): p. 1218-26.
266. Nishihara, T., et al., *miRISC recruits decapping factors to miRNA targets to enhance their degradation*. Nucleic Acids Res, 2013. **41**(18): p. 8692-705.
267. Barisic-Jager, E., et al., *HPat a decapping activator interacting with the miRNA effector complex*. PLoS One, 2013. **8**(8): p. e71860.
268. Wakiyama, M., et al., *Let-7 microRNA-mediated mRNA deadenylation and translational repression in a mammalian cell-free system*. Genes Dev, 2007. **21**(15): p. 1857-62.
269. Fukaya, T., H.O. Iwakawa, and Y. Tomari, *MicroRNAs block assembly of eIF4F translation initiation complex in Drosophila*. Mol Cell, 2014. **56**(1): p. 67-78.
270. Mathonnet, G., et al., *MicroRNA inhibition of translation initiation in vitro by targeting the cap-binding complex eIF4F*. Science, 2007. **317**: p. 1764-1767.
271. Meijer, H.A., et al., *Translational repression and eIF4A2 activity are critical for microRNA-mediated gene regulation*. Science, 2013. **340**(6128): p. 82-85.
272. Waghray, S., et al., *Xenopus CAF1 requires NOT1-mediated interaction with 4E-T to repress translation in vivo*. RNA, 2015. **21**(7): p. 1335-45.
273. Fukao, A., et al., *MicroRNAs trigger dissociation of eIF4A1 and eIF4AII from target mRNAs in humans*. Mol Cell, 2014. **56**(1): p. 79-89.
274. Schweingruber, C., et al., *Identification of Interactions in the NMD Complex Using Proximity-Dependent Biotinylation (BioID)*. PLoS One, 2016. **11**(3): p. e0150239.
275. Meijer, H.A., et al., *Translational repression and eIF4A2 activity are critical for microRNA-mediated gene regulation*. Science, 2013. **340**(6128): p. 82-5.
276. Meijer, H.A., et al., *DEAD-box helicase eIF4A2 inhibits CNOT7 deadenylation activity*. Nucleic Acids Res, 2019. **47**(15): p. 8224-8238.
277. Wilczynska, A., et al., *eIF4A2 drives repression of translation at initiation by Ccr4-Not through purine-rich motifs in the 5'UTR*. Genome Biol, 2019. **20**(1): p. 262.
278. Mathys, H., et al., *Structural and biochemical insights to the role of the CCR4-NOT complex and DDX6 ATPase in microRNA repression*. Mol Cell, 2014. **54**(5): p. 751-65.
279. Chen, Y., et al., *A DDX6-CNOT1 complex and W-binding pockets in CNOT9 reveal direct links between miRNA target recognition and silencing*. Mol Cell, 2014. **54**(5): p. 737-50.
280. Kuzuoglu-Ozturk, D., et al., *miRISC and the CCR4-NOT complex silence mRNA targets independently of 43S ribosomal scanning*. EMBO J, 2016. **35**(11): p. 1186-203.
281. Chekulaeva, M., et al., *miRNA repression involves GW182-mediated recruitment of CCR4-NOT through conserved W-containing motifs*. Nat Struct Mol Biol, 2008. **18**(11): p. 1218-1226.
282. Cooke, A., A. Prigge, and M. Wickens, *Translational repression by deadenylases*. J Biol Chem, 2010. **285**(37): p. 28506-13.
283. Bawankar, P., et al., *NOT10 and C2orf29/NOT11 form a conserved module of the CCR4-NOT complex that docks onto the NOT1 N-terminal domain*. RNA Biol, 2013. **10**(2): p. 228-44.
284. Chu, C.Y. and T.M. Rana, *Translation repression in human cells by microRNA-induced gene silencing requires RCK/p54*. PLoS Biol, 2006. **4**(7): p. e210.
285. Nishimura, T., et al., *The eIF4E-Binding Protein 4E-T Is a Component of the mRNA Decay Machinery that Bridges the 5' and 3' Termini of Target mRNAs*. Cell Rep, 2015. **11**(9): p. 1425-36.

286. Kamenska, A., et al., *Human 4E-T represses translation of bound mRNAs and enhances microRNA-mediated silencing*. Nucleic Acids Res, 2014. **42**(5): p. 3298-313.
287. Ozgur, S., et al., *Structure of a Human 4E-T/DDX6/CNOT1 Complex Reveals the Different Interplay of DDX6-Binding Proteins with the CCR4-NOT Complex*. Cell Rep, 2015. **13**(4): p. 703-711.
288. Kamenska, A., et al., *The DDX6-4E-T interaction mediates translational repression and P-body assembly*. Nucleic Acids Res, 2016. **44**(13): p. 6318-34.
289. Jinek, M., et al., *Structural insights into the human GW182-PABC interaction in microRNA-mediated deadenylation*. Nat Struct Mol Biol, 2010. **17**(2): p. 238-40.
290. Huntzinger, E., et al., *Two PABPC1-binding sites in GW182 proteins promote miRNA-mediated gene silencing*. EMBO J, 2010. **29**(24): p. 4146-60.
291. Kozlov, G., et al., *Structural basis of binding of P-body-associated proteins GW182 and ataxin-2 by the Mlle domain of poly(A)-binding protein*. J Biol Chem, 2010. **285**(18): p. 13599-606.
292. Huntzinger, E., et al., *The interactions of GW182 proteins with PABP and deadenylases are required for both translational repression and degradation of miRNA targets*. Nucleic Acids Res, 2013. **41**(2): p. 978-94.
293. Zekri, L., et al., *The silencing domain of GW182 interacts with PABPC1 to promote translational repression and degradation of microRNA targets and is required for target release*. Mol Cell Biol, 2009. **29**(23): p. 6220-31.
294. Zekri, L., D. Kuzuoglu-Ozturk, and E. Izaurralde, *GW182 proteins cause PABP dissociation from silenced miRNA targets in the absence of deadenylation*. EMBO J, 2013. **32**(7): p. 1052-65.
295. Hendrickson, D.G., et al., *Concordant regulation of translation and mRNA abundance for hundreds of targets of a human microRNA*. PLoS Biol, 2009. **7**(11): p. e1000238.
296. Guo, H., et al., *Mammalian microRNAs predominantly act to decrease target mRNA levels*. Nature, 2010. **466**(7308): p. 835-40.
297. Eichhorn, S.W., et al., *mRNA destabilization is the dominant effect of mammalian microRNAs by the time substantial repression ensues*. Mol Cell, 2014. **56**(1): p. 104-15.
298. Ding, X.C. and H. Großhans, *Repression of C. elegans microRNA targets at the initiation level of translation requires GW182 proteins*. EMBO, 2009. **28**: p. 213-222.
299. Djuranovic, S., A. Nahvi, and R. Green, *miRNA-mediated gene silencing by translational repression followed by mRNA deadenylation and decay*. Science, 2012. **336**(6078): p. 237-40.
300. Bethune, J., C.G. Artus-Revel, and W. Filipowicz, *Kinetic analysis reveals successive steps leading to miRNA-mediated silencing in mammalian cells*. EMBO Rep, 2012. **13**(8): p. 716-23.
301. Bazzini, A.A., M.T. Lee, and A.J. Giraldez, *Ribosome profiling shows that miR-430 reduces translation before causing mRNA decay in zebrafish*. Science, 2012. **336**(6078): p. 233-7.
302. Iwasaki, S., T. Kawamata, and Y. Tomari, *Drosophila argonaute1 and argonaute2 employ distinct mechanisms for translational repression*. Mol Cell, 2009. **34**(1): p. 58-67.
303. Eulalio, A., et al., *Deadenylation is a widespread effect of miRNA regulation*. RNA, 2009. **15**(1): p. 21-32.
304. Piao, X., et al., *CCR4-NOT deadenylates mRNA associated with RNA-induced silencing complexes in human cells*. Mol Cell Biol, 2010. **30**(6): p. 1486-94.
305. Kundu, P., et al., *HuR protein attenuates miRNA-mediated repression by promoting miRISC dissociation from the target RNA*. Nucleic Acids Res, 2012. **40**(11): p. 5088-100.

306. Banerjee, S., P. Neveu, and K.S. Kosik, *A coordinated local translational control point at the synapse involving relief from silencing and MOV10 degradation*. *Neuron*, 2009. **64**(6): p. 871-84.
307. Vasudevan, S., Y. Tong, and J.A. Steitz, *Switching from repression to activation: MicroRNAs can up-regulate translation*. *Science*, 2007. **318**(5858): p. 1931-1934.
308. Orom, U.A., F.C. Nielsen, and A.H. Lund, *MicroRNA-10a binds the 5'UTR of ribosomal protein mRNAs and enhances their translation*. *Mol Cell*, 2008. **30**(4): p. 460-71.
309. Jangra, R.K., M. Yi, and S.M. Lemon, *Regulation of hepatitis C virus translation and infectious virus production by the microRNA miR-122*. *J Virol*, 2010. **84**(13): p. 6615-25.
310. Jopling, C.L., S. Schutz, and P. Sarnow, *Position-dependent function for a tandem microRNA miR-122-binding site located in the hepatitis C virus RNA genome*. *Cell Host Microbe*, 2008. **4**(1): p. 77-85.
311. Lewis, A., *Investigations of miR-122 function, stability and localisation*, in *School of Pharmacy*. 2014, University of Nottingham.
312. Li, Z.Y., et al., *Positive regulation of hepatic miR-122 expression by HNF4alpha*. *J Hepatol*, 2011. **55**(3): p. 602-611.
313. Mortimer, S.A. and J.A. Doudna, *Unconventional miR-122 binding stabilizes the HCV genome by forming a trimolecular RNA structure*. *Nucleic Acids Res*, 2013. **41**(7): p. 4230-40.
314. Roberts, A.P., A.P. Lewis, and C.L. Jopling, *miR-122 activates hepatitis C virus translation by a specialized mechanism requiring particular RNA components*. *Nucleic Acids Res*, 2011. **39**(17): p. 7716-29.
315. Sedano, C.D. and P. Sarnow, *Hepatitis C virus subverts liver-specific miR-122 to protect the viral genome from exoribonuclease Xrn2*. *Cell Host Microbe*, 2014. **16**(2): p. 257-64.
316. Randall, G., et al., *Cellular cofactors affecting hepatitis C virus infection and replication*. *Proc Natl Acad Sci U S A*, 2007. **104**(31): p. 12884-9.
317. Ahmed, C.S., et al., *Eukaryotic translation initiation factor 4AII contributes to microRNA-122 regulation of hepatitis C virus replication*. *Nucleic Acids Res*, 2018.
318. Libri, V., et al., *Regulation of microRNA biogenesis and turnover by animals and their viruses*. *Cell Mol Life Sci*, 2013. **70**(19): p. 3525-44.
319. Gatfield, D., et al., *Integration of microRNA miR-122 in hepatic circadian gene expression*. *Genes Dev*, 2009. **23**(11): p. 1313-26.
320. van Rooij, E., et al., *Control of stress-dependent cardiac growth and gene expression by a microRNA*. *Science*, 2007. **316**(5824): p. 575-9.
321. Hwang, H.W., E.A. Wentzel, and J.T. Mendell, *A hexanucleotide element directs microRNA nuclear import*. *Science*, 2007. **315**(5808): p. 97-100.
322. Avraham, R., et al., *EGF decreases the abundance of microRNAs that restrain oncogenic transcription factors*. *Sci Signal*, 2010. **3**(124): p. ra43.
323. Ruegger, S. and H. Grosshans, *MicroRNA turnover: when, how, and why*. *Trends Biochem Sci*, 2012. **37**(10): p. 436-46.
324. Katoh, T., et al., *Selective stabilization of mammalian microRNAs by 3' adenylation mediated by the cytoplasmic poly(A) polymerase GLD-2*. *Genes Dev*, 2009. **23**(4): p. 433-8.
325. Andrade, J.M. and C.M. Arraiano, *PNPase is a key player in the regulation of small RNAs that control the expression of outer membrane proteins*. *RNA*, 2008. **14**(3): p. 543-51.
326. Das, S.K., et al., *Human polynucleotide phosphorylase selectively and preferentially degrades microRNA-221 in human melanoma cells*. *Proc Natl Acad Sci U S A*, 2010. **107**(26): p. 11948-53.

327. Bak, R.O. and J.G. Mikkelsen, *miRNA sponges: soaking up miRNAs for regulation of gene expression*. Wiley Interdiscip Rev RNA, 2014. **5**(3): p. 317-33.
328. Ebert, M.S., J.R. Neilson, and P.A. Sharp, *MicroRNA sponges: competitive inhibitors of small RNAs in mammalian cells*. Nat Methods, 2007. **4**(9): p. 721-6.
329. Ebert, M.S. and P.A. Sharp, *MicroRNA sponges: progress and possibilities*. RNA, 2010. **16**(11): p. 2043-50.
330. Reichel, M., J. Li, and A.A. Millar, *Silencing the silencer: strategies to inhibit microRNA activity*. Biotechnol Lett, 2011. **33**(7): p. 1285-92.
331. Alkan, A.H. and B. Akgul, *Endogenous miRNA Sponges*. Methods Mol Biol, 2022. **2257**: p. 91-104.
332. Rutnam, Z.J. and B.B. Yang, *The non-coding 3' UTR of CD44 induces metastasis by regulating extracellular matrix functions*. J Cell Sci, 2012. **125**(Pt 8): p. 2075-85.
333. Wang, J., et al., *CREB up-regulates long non-coding RNA, HULC expression through interaction with microRNA-372 in liver cancer*. Nucleic Acids Res, 2010. **38**(16): p. 5366-83.
334. Poliseno, L., et al., *A coding-independent function of gene and pseudogene mRNAs regulates tumour biology*. Nature, 2010. **465**(7301): p. 1033-8.
335. Kulcheski, F.R., A.P. Christoff, and R. Margis, *Circular RNAs are miRNA sponges and can be used as a new class of biomarker*. J Biotechnol, 2016. **238**: p. 42-51.
336. Han, C., et al., *Regulation of microRNAs function by circular RNAs in human cancer*. Oncotarget, 2017. **8**(38): p. 64622-64637.
337. Song, Y.Z. and J.F. Li, *Circular RNA hsa_circ_0001564 regulates osteosarcoma proliferation and apoptosis by acting miRNA sponge*. Biochem Biophys Res Commun, 2018. **495**(3): p. 2369-2375.
338. Chen, L.L. and L. Yang, *Regulation of circRNA biogenesis*. RNA Biol, 2015. **12**(4): p. 381-8.
339. Memczak, S., et al., *Circular RNAs are a large class of animal RNAs with regulatory potency*. Nature, 2013. **495**(7441): p. 333-8.
340. Leung, A.K. and P.A. Sharp, *Quantifying Argonaute proteins in and out of GW/P-bodies: implications in microRNA activities*. Adv Exp Med Biol, 2013. **768**: p. 165-82.
341. Liu, J., et al., *MicroRNA-dependent localization of targeted mRNAs to mammalian P-bodies*. Nat Cell Biol, 2005. **7**(7): p. 719-723.
342. Eulalio, A., et al., *P-body formation is a consequence, not the cause, of RNA-mediated gene silencing*. Mol Cell Biol, 2007. **27**(11): p. 3970-81.
343. Aizer, A., et al., *Quantifying mRNA targeting to P-bodies in living human cells reveals their dual role in mRNA decay and storage*. J Cell Sci, 2014. **127**(Pt 20): p. 4443-56.
344. Cikaluk, D.E., et al., *GERp95, a membrane-associated protein that belongs to a family of proteins involved in stem cell differentiation*. Mol Biol Cell, 1999. **10**: p. 3357-3372.
345. Barbato, C., et al., *Dicer expression and localization in post-mitotic neurons*. Brain Res, 2007. **1175**: p. 17-27.
346. Tahbaz, N., et al., *Characterization of the interactions between mammalian PAZ PIWI domain proteins and Dicer*. EMBO Rep, 2004. **5**(2): p. 189-94.
347. Suzawa, M., et al., *Comprehensive Identification of Nuclear and Cytoplasmic TNRC6A-Associating Proteins*. J Mol Biol, 2017. **429**(21): p. 3319-3333.
348. Li, S., et al., *MicroRNAs inhibit the translation of target mRNAs on the endoplasmic reticulum in Arabidopsis*. Cell, 2013. **153**(3): p. 562-74.
349. Wu, P.H., M. Isaji, and R.W. Carthew, *Functionally diverse microRNA effector complexes are regulated by extracellular signaling*. Mol Cell, 2013. **52**(1): p. 113-23.

350. Antoniou, A., et al., *The dynamic recruitment of TRBP to neuronal membranes mediates dendritogenesis during development*. EMBO Rep, 2018. **19**(3).
351. Stalder, L., et al., *The rough endoplasmatic reticulum is a central nucleation site of siRNA-mediated RNA silencing*. EMBO J, 2013. **32**(8): p. 1115-27.
352. Gibbings, D.J., et al., *Multivesicular bodies associate with components of miRNA effector complexes and modulate miRNA activity*. Nat Cell Biol, 2009. **11**(9): p. 1143-9.
353. Lee, Y.S., et al., *Silencing by small RNAs is linked to endosomal trafficking*. Nat Cell Biol, 2009. **11**(9): p. 1150-6.
354. Bose, M., et al., *Spatiotemporal Uncoupling of MicroRNA-Mediated Translational Repression and Target RNA Degradation Controls MicroRNP Recycling in Mammalian Cells*. Mol Cell Biol, 2017. **37**(4).
355. Cortez, M.A., et al., *MicroRNAs in body fluids--the mix of hormones and biomarkers*. Nat Rev Clin Oncol, 2011. **8**(8): p. 467-77.
356. Morel, L., et al., *Neuronal exosomal miRNA-dependent translational regulation of astroglial glutamate transporter GLT1*. J Biol Chem, 2013. **288**(10): p. 7105-16.
357. Squadrito, M.L., et al., *Endogenous RNAs modulate microRNA sorting to exosomes and transfer to acceptor cells*. Cell Rep, 2014. **8**(5): p. 1432-46.
358. Ghosh, S., et al., *Polysome arrest restricts miRNA turnover by preventing exosomal export of miRNA in growth-retarded mammalian cells*. Mol Biol Cell, 2015. **26**(6): p. 1072-83.
359. Morel, E., *Endoplasmic Reticulum Membrane and Contact Site Dynamics in Autophagy Regulation and Stress Response*. Front Cell Dev Biol, 2020. **8**: p. 343.
360. Wei, Y., et al., *Origin of the Autophagosome Membrane in Mammals*. Biomed Res Int, 2018. **2018**: p. 1012789.
361. Gibbings, D., et al., *Selective autophagy degrades DICER and AGO2 and regulates miRNA activity*. Nat Cell Biol, 2012. **14**(12): p. 1314-21.
362. Dignam, J.D., R.M. Lebovitz, and R.G. Roeder, *Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei*. Nucleic Acids Res, 1983. **11**(5): p. 1475-89.
363. Hwang, H., E.A. Wentzel, and J.T. Mendell, *A hexanucleotide element directs microRNA nuclear import*. Science, 2007. **315**(5808): p. 97-100.
364. Castanotto, D., et al., *A stress-induced response complex (SIRC) shuttles miRNAs, siRNAs, and oligonucleotides to the nucleus*. Proc Natl Acad Sci U S A, 2018. **115**(25): p. E5756-E5765.
365. Liao, J.Y., et al., *Deep sequencing of human nuclear and cytoplasmic small RNAs reveals an unexpectedly complex subcellular distribution of miRNAs and tRNA 3' trailers*. PLoS One, 2010. **5**(5): p. e10563.
366. Jeffries, C.D., H.M. Fried, and D.O. Perkins, *Nuclear and cytoplasmic localization of neural stem cell microRNAs*. RNA, 2011. **17**(4): p. 675-86.
367. Gagnon, K.T., et al., *RNAi factors are present and active in human cell nuclei*. Cell Rep, 2014. **6**(1): p. 211-21.
368. Robb, G.B., et al., *Specific and potent RNAi in the nucleus of human cells*. Nat Struct Mol Biol, 2005. **12**(2): p. 133-7.
369. Ohrt, T., et al., *Fluorescence correlation spectroscopy and fluorescence cross-correlation spectroscopy reveal the cytoplasmic origination of loaded nuclear RISC in vivo in human cells*. Nucleic Acids Res, 2008. **36**(20): p. 6439-49.
370. Weinmann, L., et al., *Importin 8 is a gene silencing factor that targets argonaute proteins to distinct mRNAs*. Cell, 2009. **136**(3): p. 496-507.

371. Sarshad, A.A., et al., *Argonaute-miRNA Complexes Silence Target mRNAs in the Nucleus of Mammalian Stem Cells*. Mol Cell, 2018. **71**(6): p. 1040-1050 e8.
372. Chi, S.W., et al., *Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps*. Nature, 2009. **460**(7254): p. 479-86.
373. Fasanaro, P., et al., *An integrated approach for experimental target identification of hypoxia-induced miR-210*. J Biol Chem, 2009. **284**(50): p. 35134-43.
374. Leucci, E., et al., *microRNA-9 targets the long non-coding RNA MALAT1 for degradation in the nucleus*. Sci Rep, 2013. **3**: p. 2535.
375. Hansen, T.B., et al., *miRNA-dependent gene silencing involving Ago2-mediated cleavage of a circular antisense RNA*. EMBO J, 2011. **30**(21): p. 4414-22.
376. Tang, R., et al., *Mouse miRNA-709 directly regulates miRNA-15a/16-1 biogenesis at the posttranscriptional level in the nucleus: evidence for a microRNA hierarchy system*. Cell Res, 2012. **22**(3): p. 504-15.
377. Norman, K.L., et al., *Precursor microRNA-122 inhibits synthesis of Insig1 isoform mRNA by modulating polyadenylation site usage*. RNA, 2017. **23**(12): p. 1886-1893.
378. Zisoulis, D.G., et al., *Autoregulation of microRNA biogenesis by let-7 and Argonaute*. Nature, 2012. **486**(7404): p. 541-4.
379. Wang, D., et al., *Nuclear miR-122 directly regulates the biogenesis of cell survival oncomiR miR-21 at the posttranscriptional level*. Nucleic Acids Res, 2018. **46**(4): p. 2012-2029.
380. Kim, D.H., et al., *MicroRNA-directed transcriptional gene silencing in mammalian cells*. Proc Natl Acad Sci U S A, 2008. **105**(42): p. 16230-5.
381. Younger, S.T., A. Pertsemlidis, and D.R. Corey, *Predicting potential miRNA target sites within gene promoters*. Bioorg Med Chem Lett, 2009. **19**(14): p. 3791-4.
382. Tan, Y., et al., *Transcriptional inhibition of Hoxd4 expression by miRNA-10a in human breast cancer cells*. BMC Mol Biol, 2009. **10**: p. 12.
383. Adilakshmi, T., I. Sudol, and N. Tapinos, *Combinatorial action of miRNAs regulates transcriptional and post-transcriptional gene silencing following in vivo PNS injury*. PLoS One, 2012. **7**(7): p. e39674.
384. Zardo, G., et al., *Polycombs and microRNA-223 regulate human granulopoiesis by transcriptional control of target gene expression*. Blood, 2012. **119**(17): p. 4034-46.
385. Benhamed, M., et al., *Senescence is an endogenous trigger for microRNA-directed transcriptional gene silencing in human cells*. Nat Cell Biol, 2012. **14**(3): p. 266-75.
386. Modarresi, F., et al., *Inhibition of natural antisense transcripts in vivo results in gene-specific transcriptional upregulation*. Nat Biotechnol, 2012. **30**(5): p. 453-9.
387. Morris, K.V., et al., *Bidirectional transcription directs both transcriptional gene activation and suppression in human cells*. PLoS Genet, 2008. **4**(11): p. e1000258.
388. Matsui, M., et al., *Promoter RNA links transcriptional regulation of inflammatory pathway genes*. Nucleic Acids Res, 2013. **41**(22): p. 10086-109.
389. Allo, M., et al., *Control of alternative splicing through siRNA-mediated transcriptional gene silencing*. Nat Struct Mol Biol, 2009. **16**(7): p. 717-24.
390. Ameyar-Zazoua, M., et al., *Argonaute proteins couple chromatin silencing to alternative splicing*. Nat Struct Mol Biol, 2012. **19**(10): p. 998-1004.
391. Bai, B., H. Liu, and M. Laiho, *Small RNA expression and deep sequencing analyses of the nucleolus reveal the presence of nucleolus-associated microRNAs*. FEBS Open Bio, 2014. **4**: p. 441-9.
392. Li, Z.F., et al., *Dynamic localisation of mature microRNAs in Human nucleoli is influenced by exogenous genetic materials*. PLoS One, 2013. **8**(8): p. e70869.

393. Politz, J.C.R., F. Zhang, and T. Pederson, *MicroRNA-206 colocalizes with ribosome-rich regions in both the nucleolus and cytoplasm of rat myogenic cells*. Proceedings of the National Academy of Sciences, 2007. **104**(2): p. 684-684.
394. Politz, J.C., E.M. Hogan, and T. Pederson, *MicroRNAs with a nucleolar location*. RNA, 2009. **15**(9): p. 1705-15.
395. Marcon, E., et al., *miRNA and piRNA localization in the male mammalian meiotic nucleus*. Chromosome Res, 2008. **16**(2): p. 243-60.
396. Castanotto, D., et al., *CRM1 mediates nuclear-cytoplasmic shuttling of mature microRNAs*. Proc Natl Acad Sci U S A, 2009. **106**(51): p. 21655-9.
397. Reyes-Gutierrez, P., J.C. Ritland Politz, and T. Pederson, *A mRNA and cognate microRNAs localize in the nucleolus*. Nucleus, 2014. **5**(6): p. 636-42.
398. Lung, B., et al., *Identification of small non-coding RNAs from mitochondria and chloroplasts*. Nucleic Acids Res, 2006. **34**(14): p. 3842-52.
399. Kren, B.T., et al., *MicroRNAs identified in highly purified liver-derived mitochondria may play a role in apoptosis*. RNA Biol, 2009. **6**(1): p. 65-72.
400. Bian, Z., et al., *Identification of mouse liver mitochondria-associated miRNAs and their potential biological functions*. Cell Res, 2010. **20**(9): p. 1076-8.
401. Barrey, E., et al., *Pre-microRNA and mature microRNA in human mitochondria*. PLoS One, 2011. **6**(5): p. e20220.
402. Sripada, L., et al., *Systematic analysis of small RNAs associated with human mitochondria by deep sequencing: detailed analysis of mitochondrial associated miRNA*. PLoS One, 2012. **7**(9): p. e44873.
403. Bandiera, S., et al., *Nuclear outsourcing of RNA interference components to human mitochondria*. PLoS One, 2011. **6**(6): p. e20746.
404. Zhang, X., et al., *MicroRNA directly enhances mitochondrial translation during muscle differentiation*. Cell, 2014. **158**(3): p. 607-19.
405. Das, S., et al., *Nuclear miRNA regulates the mitochondrial genome in the heart*. Circ Res, 2012. **110**(12): p. 1596-603.
406. Das, S., et al., *miR-181c regulates the mitochondrial genome, bioenergetics, and propensity for heart failure in vivo*. PLoS One, 2014. **9**(5): p. e96820.
407. Farina, K.L., et al., *Two ZBP1 KH domains facilitate beta-actin mRNA localization, granule formation, and cytoskeletal attachment*. J Cell Biol, 2003. **160**(1): p. 77-87.
408. Huttelmaier, S., et al., *Spatial regulation of beta-actin translation by Src-dependent phosphorylation of ZBP1*. Nature, 2005. **438**(7067): p. 512-5.
409. Ashraf, S.I., et al., *Synaptic protein synthesis associated with memory is regulated by the RISC pathway in Drosophila*. Cell, 2006. **124**(1): p. 191-205.
410. Watts, M., et al., *MicroRNA-210 Regulates Dendritic Morphology and Behavioural Flexibility in Mice*. Mol Neurobiol, 2021. **58**(4): p. 1330-1344.
411. Schrott, G.M., et al., *A brain-specific microRNA regulates dendritic spine development*. Nature, 2006. **439**(7074): p. 283-9.
412. Dhir, A., et al., *Microprocessor mediates transcriptional termination of long noncoding RNA transcripts hosting microRNAs*. Nat Struct Mol Biol, 2015. **22**(4): p. 319-27.
413. Lagos-Quintana, M., et al., *Identification of tissue-specific microRNAs from mouse*. Curr Biol, 2002. **12**(9): p. 735-9.
414. Esau, C., et al., *miR-122 regulation of lipid metabolism revealed by in vivo antisense targeting*. Cell Metab, 2006. **3**(2): p. 87-98.

415. Elmen, J., et al., *Antagonism of microRNA-122 in mice by systemically administered LNA-antimiR leads to up-regulation of a large set of predicted target mRNAs in the liver*. Nucleic Acids Res, 2008. **36**(4): p. 1153-62.
416. Krutzfeldt, J., et al., *Silencing of microRNAs in vivo with 'antagomirs'*. Nature, 2005. **438**(7068): p. 685-9.
417. Castoldi, M., et al., *The liver-specific microRNA miR-122 controls systemic iron homeostasis in mice*. J Clin Invest, 2011. **121**(4): p. 1386-96.
418. Burchard, J., et al., *microRNA-122 as a regulator of mitochondrial metabolic gene network in hepatocellular carcinoma*. Mol Syst Biol, 2010. **6**: p. 402.
419. Hsu, S.H., et al., *MicroRNA-122 regulates polyploidization in the murine liver*. Hepatology, 2016. **64**(2): p. 599-615.
420. Coulouarn, C., et al., *Loss of miR-122 expression in liver cancer correlates with suppression of the hepatic phenotype and gain of metastatic properties*. Oncogene, 2009. **28**(40): p. 3526-36.
421. Hsu, S.H., et al., *Essential metabolic, anti-inflammatory, and anti-tumorigenic functions of miR-122 in liver*. J Clin Invest, 2012. **122**(8): p. 2871-83.
422. Tsai, W.C., et al., *MicroRNA-122 plays a critical role in liver homeostasis and hepatocarcinogenesis*. J Clin Invest, 2012. **122**(8): p. 2884-97.
423. Hornshoj, H., et al., *Pan-cancer screen for mutations in non-coding elements with conservation and cancer specificity reveals correlations with expression and survival*. NPJ Genom Med, 2018. **3**: p. 1.
424. Wang, B., H. Wang, and Z. Yang, *MiR-122 inhibits cell proliferation and tumorigenesis of breast cancer by targeting IGF1R*. PLoS One, 2012. **7**(10): p. e47053.
425. Chen, Y., et al., *miR-122 targets NOD2 to decrease intestinal epithelial cell injury in Crohn's disease*. Biochem Biophys Res Commun, 2013. **438**(1): p. 133-9.
426. Bandiera, S., et al., *miR-122--a key factor and therapeutic target in liver disease*. J Hepatol, 2015. **62**(2): p. 448-57.
427. Barajas, J.M., et al., *The role of miR-122 in the dysregulation of glucose-6-phosphate dehydrogenase (G6PD) expression in hepatocellular cancer*. Sci Rep, 2018. **8**(1): p. 9105.
428. Du, W., et al., *TAp73 enhances the pentose phosphate pathway and supports cell proliferation*. Nat Cell Biol, 2013. **15**(8): p. 991-1000.
429. Martinez-Outschoorn, U.E., et al., *Cancer metabolism: a therapeutic perspective*. Nat Rev Clin Oncol, 2017. **14**(1): p. 11-31.
430. Jiang, P., et al., *p53 regulates biosynthesis through direct inactivation of glucose-6-phosphate dehydrogenase*. Nat Cell Biol, 2011. **13**(3): p. 310-6.
431. Sengupta, D., et al., *Regulation of hepatic glutamine metabolism by miR-122*. Mol Metab, 2020. **34**: p. 174-186.
432. Wei, S., et al., *HNF-4alpha regulated miR-122 contributes to development of gluconeogenesis and lipid metabolism disorders in Type 2 diabetic mice and in palmitate-treated HepG2 cells*. Eur J Pharmacol, 2016. **791**: p. 254-263.
433. Zeng, Y., et al., *G6PC3, ALDOA and CS induction accompanies mir-122 down-regulation in the mechanical asphyxia and can serve as hypoxia biomarkers*. Oncotarget, 2016. **7**(46): p. 74526-74536.
434. Dubuisson, J., F. Penin, and D. Moradpour, *Interaction of hepatitis C virus proteins with host cell membranes and lipids*. Trends Cell Biol, 2002. **12**(11): p. 517-23.
435. Suzuki, T., *Hepatitis C Virus Replication*. Adv Exp Med Biol, 2017. **997**: p. 199-209.
436. Jopling, C.L., et al., *Modulation of Hepatitis C virus RNA abundance by a liver-specific microRNA*. Science, 2005. **309**(5740): p. 1577-1581.

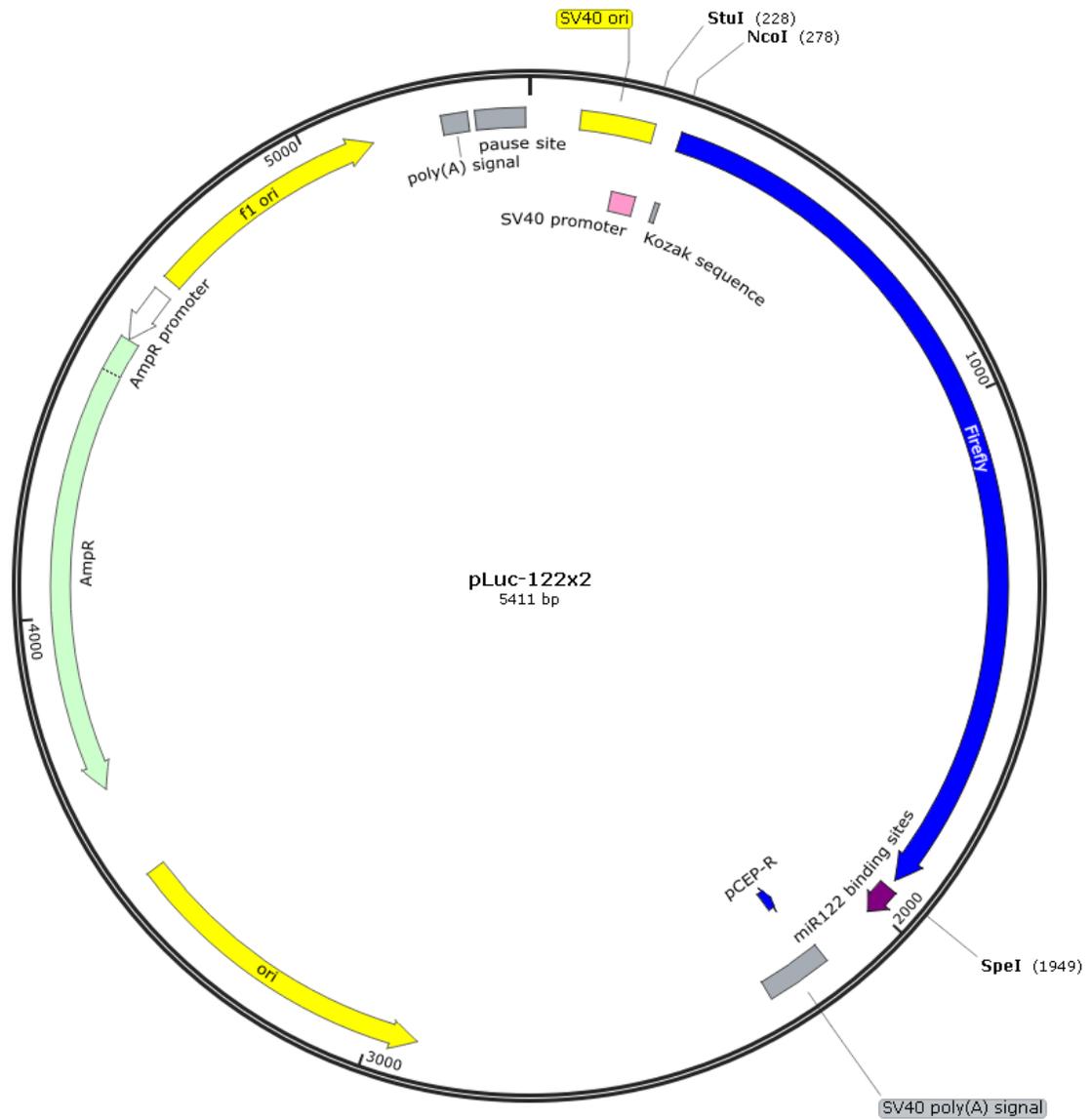
437. Henke, J.I., et al., *microRNA-122 stimulates translation of hepatitis C virus RNA*. EMBO J, 2008. **27**(24): p. 3300-10.
438. Bradrick, S.S., S. Nagyal, and H. Novatt, *A miRNA-responsive cell-free translation system facilitates isolation of hepatitis C virus miRNP complexes*. RNA, 2013. **19**(8): p. 1159-69.
439. Masaki, T., et al., *miR-122 stimulates hepatitis C virus RNA synthesis by altering the balance of viral RNAs engaged in replication versus translation*. Cell Host Microbe, 2015. **17**(2): p. 217-28.
440. Ahmed, C.S., et al., *Eukaryotic translation initiation factor 4AII contributes to microRNA-122 regulation of hepatitis C virus replication*. Nucleic Acids Res, 2018. **46**(12): p. 6330-6343.
441. Wilson, J.A., et al., *Human Ago2 is required for efficient microRNA 122 regulation of hepatitis C virus RNA accumulation and translation*. J Virol, 2011. **85**(5): p. 2342-50.
442. Machlin, E.S., P. Sarnow, and S.M. Sagan, *Masking the 5' terminal nucleotides of the hepatitis C virus genome by an unconventional microRNA-target RNA complex*. Proc Natl Acad Sci U S A, 2011. **108**(8): p. 3193-8.
443. Li, Y., et al., *Competing and noncompeting activities of miR-122 and the 5' exonuclease Xrn1 in regulation of hepatitis C virus replication*. Proc Natl Acad Sci U S A, 2013. **110**(5): p. 1881-6.
444. Pang, P.S., et al., *Structural map of a microRNA-122: hepatitis C virus complex*. J Virol, 2012. **86**(2): p. 1250-4.
445. Zhang, Y., Z. Wang, and R.A. Gemeinhart, *Progress in microRNA delivery*. J Control Release, 2013. **172**(3): p. 962-74.
446. Rupaimoole, R. and F.J. Slack, *MicroRNA therapeutics: towards a new era for the management of cancer and other diseases*. Nat Rev Drug Discov, 2017. **16**(3): p. 203-222.
447. Simonson, B. and S. Das, *MicroRNA Therapeutics: the Next Magic Bullet?* Mini Rev Med Chem, 2015. **15**(6): p. 467-74.
448. Ma, L., et al., *Therapeutic silencing of miR-10b inhibits metastasis in a mouse mammary tumor model*. Nat Biotechnol, 2010. **28**(4): p. 341-7.
449. Yoo, B., et al., *Combining miR-10b-Targeted Nanotherapy with Low-Dose Doxorubicin Elicits Durable Regressions of Metastatic Breast Cancer*. Cancer Res, 2015. **75**(20): p. 4407-15.
450. Janssen, H.L., et al., *Treatment of HCV infection by targeting microRNA*. N Engl J Med, 2013. **368**(18): p. 1685-94.
451. Jones, D., *Setbacks shadow microRNA therapies in the clinic*. Nat Biotechnol, 2018. **36**(10): p. 909-910.
452. Wang, W.X., et al., *The expression of microRNA miR-107 decreases early in Alzheimer's disease and may accelerate disease progression through regulation of beta-site amyloid precursor protein-cleaving enzyme 1*. J Neurosci, 2008. **28**(5): p. 1213-23.
453. Segura, M.F., et al., *Melanoma MicroRNA signature predicts post-recurrence survival*. Clin Cancer Res, 2010. **16**(5): p. 1577-86.
454. Yu, S.L., et al., *MicroRNA signature predicts survival and relapse in lung cancer*. Cancer Cell, 2008. **13**(1): p. 48-57.
455. Mishra, P.J., *Non-coding RNAs as clinical biomarkers for cancer diagnosis and prognosis*. Expert Rev Mol Diagn, 2014. **14**(8): p. 917-9.
456. Mishra, P.J. and G. Merlino, *MicroRNA reexpression as differentiation therapy in cancer*. J Clin Invest, 2009. **119**(8): p. 2119-23.

457. Lawrie, C.H., et al., *MicroRNA expression distinguishes between germinal center B cell-like and activated B cell-like subtypes of diffuse large B cell lymphoma*. *Int J Cancer*, 2007. **121**(5): p. 1156-61.
458. Wang, J., et al., *MicroRNAs in plasma of pancreatic ductal adenocarcinoma patients as novel blood-based biomarkers of disease*. *Cancer Prev Res (Phila)*, 2009. **2**(9): p. 807-13.
459. Huang, Z., et al., *Plasma microRNAs are promising novel biomarkers for early detection of colorectal cancer*. *Int J Cancer*, 2010. **127**(1): p. 118-26.
460. Xing, L., et al., *Early detection of squamous cell lung cancer in sputum by a panel of microRNA markers*. *Mod Pathol*, 2010. **23**(8): p. 1157-64.
461. Chen, X., et al., *Characterization of microRNAs in serum: a novel class of biomarkers for diagnosis of cancer and other diseases*. *Cell Res*, 2008. **18**(10): p. 997-1006.
462. Anadol, E., et al., *Circulating microRNAs as a marker for liver injury in human immunodeficiency virus patients*. *Hepatology*, 2015. **61**(1): p. 46-55.
463. Rabinowits, G., et al., *Exosomal microRNA: a diagnostic marker for lung cancer*. *Clin Lung Cancer*, 2009. **10**(1): p. 42-6.
464. Kuwabara, Y., et al., *Increased microRNA-1 and microRNA-133a levels in serum of patients with cardiovascular disease indicate myocardial damage*. *Circ Cardiovasc Genet*, 2011. **4**(4): p. 446-54.
465. Street, J.M., et al., *Urine Exosomes: An Emerging Trove of Biomarkers*. *Adv Clin Chem*, 2017. **78**: p. 103-122.
466. Salehi, M. and M. Sharifi, *Exosomal miRNAs as novel cancer biomarkers: Challenges and opportunities*. *J Cell Physiol*, 2018. **233**(9): p. 6370-6380.
467. Chen, J.J., et al., *Potential Roles of Exosomal MicroRNAs as Diagnostic Biomarkers and Therapeutic Application in Alzheimer's Disease*. *Neural Plast*, 2017. **2017**: p. 7027380.
468. Lindenbach, B.D. and C.M. Rice, *Unravelling hepatitis C virus replication from genome to function*. *Nature*, 2005. **436**(7053): p. 933-8.
469. Labun, K., et al., *CHOPCHOP v3: expanding the CRISPR web toolbox beyond genome editing*. *Nucleic Acids Res*, 2019. **47**(W1): p. W171-W174.
470. Dye, M.J., N. Gromak, and N.J. Proudfoot, *Exon tethering in transcription by RNA polymerase II*. *Mol Cell*, 2006. **21**(6): p. 849-59.
471. Jagannathan, S., C. Nwosu, and C.V. Nicchitta, *Analyzing mRNA localization to the endoplasmic reticulum via cell fractionation*. *Methods Mol Biol*, 2011. **714**: p. 301-21.
472. Reid, D.W. and C.V. Nicchitta, *Diversity and selectivity in mRNA translation on the endoplasmic reticulum*. *Nat Rev Mol Cell Biol*, 2015. **16**(4): p. 221-31.
473. Reid, D.W. and C.V. Nicchitta, *Primary role for endoplasmic reticulum-bound ribosomes in cellular translation identified by ribosome profiling*. *J Biol Chem*, 2012. **287**(8): p. 5518-27.
474. Jagannathan, S., et al., *De novo translation initiation on membrane-bound ribosomes as a mechanism for localization of cytosolic protein mRNAs to the endoplasmic reticulum*. *RNA*, 2014. **20**(10): p. 1489-98.
475. Kyono, K., M. Miyashiro, and I. Taguchi, *Human eukaryotic initiation factor 4AII associates with hepatitis C virus NS5B protein in vitro*. *Biochem Biophys Res Commun*, 2002. **292**(3): p. 659-66.
476. Hall, M.P., et al., *Engineered luciferase reporter from a deep sea shrimp utilizing a novel imidazopyrazinone substrate*. *ACS Chem Biol*, 2012. **7**(11): p. 1848-57.
477. Banerji, J., S. Rusconi, and W. Schaffner, *Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences*. *Cell*, 1981. **27**(2 Pt 1): p. 299-308.

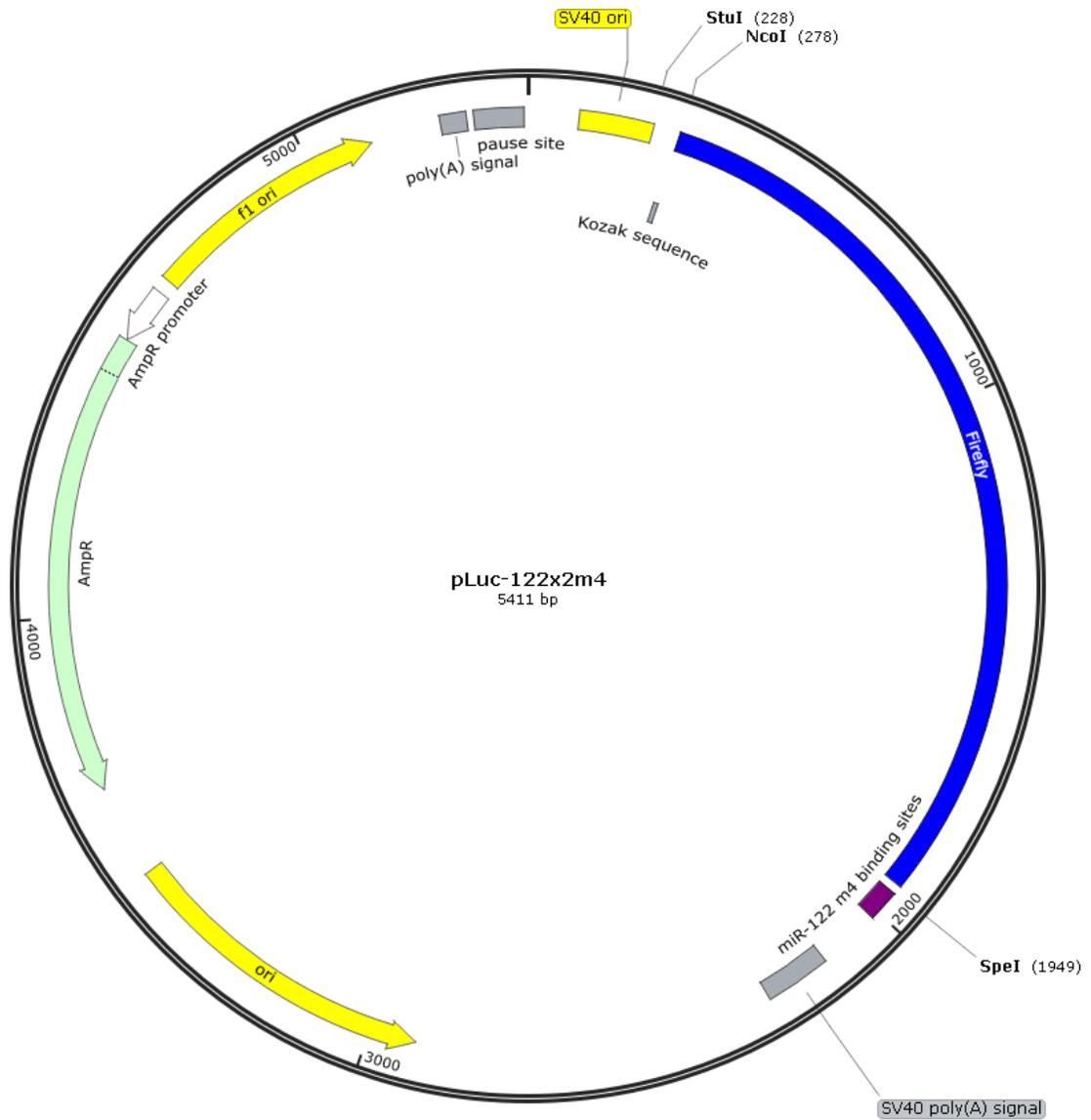
478. Wider, D. and D. Picard, *Secreted dual reporter assay with Gaussia luciferase and the red fluorescent protein mCherry*. PLoS One, 2017. **12**(12): p. e0189403.
479. Nagase, T., H. Koga, and O. Ohara, *Kazusa mammalian cDNA resources: towards functional characterization of KIAA gene products*. Brief Funct Genomic Proteomic, 2006. **5**(1): p. 4-7.
480. Ohara, O., et al., *Construction and characterization of human brain cDNA libraries suitable for analysis of cDNA clones encoding relatively large proteins*. DNA Res, 1997. **4**(1): p. 53-9.
481. Inouye, S. and T. Suzuki, *Protein expression of preferred human codon-optimized Gaussia luciferase genes with an artificial open-reading frame in mammalian and bacterial cells*. Protein Expr Purif, 2016. **128**: p. 93-100.
482. Chou, T.C. and R.L. Moyle, *Synthetic versions of firefly luciferase and Renilla luciferase reporter genes that resist transgene silencing in sugarcane*. BMC Plant Biol, 2014. **14**: p. 92.
483. Kudla, G., et al., *High guanine and cytosine content increases mRNA levels in mammalian cells*. PLoS Biol, 2006. **4**(6): p. e180.
484. Kusakabe, T., K. Motoki, and K. Hori, *Mode of interactions of human aldolase isozymes with cytoskeletons*. Arch Biochem Biophys, 1997. **344**(1): p. 184-93.
485. Li, X., et al., *Fructose-Bisphosphate Aldolase A Regulates Hypoxic Adaptation in Hepatocellular Carcinoma and Involved with Tumor Malignancy*. Dig Dis Sci, 2019. **64**(11): p. 3215-3227.
486. Semenza, G.L., et al., *Transcriptional regulation of genes encoding glycolytic enzymes by hypoxia-inducible factor 1*. J Biol Chem, 1994. **269**(38): p. 23757-63.
487. McDonald, K.K., et al., *A caveolar complex between the cationic amino acid transporter 1 and endothelial nitric-oxide synthase may explain the "arginine paradox"*. J Biol Chem, 1997. **272**(50): p. 31213-6.
488. Fernandez, J., et al., *Translation mediated by the internal ribosome entry site of the cat-1 mRNA is regulated by glucose availability in a PERK kinase-dependent manner*. J Biol Chem, 2002. **277**(14): p. 11780-7.
489. Huang, C.C., et al., *Temporal regulation of Cat-1 (cationic amino acid transporter-1) gene transcription during endoplasmic reticulum stress*. Biochem J, 2010. **429**(1): p. 215-24.
490. Banka, S. and W.G. Newman, *A clinical and molecular review of ubiquitous glucose-6-phosphatase deficiency caused by G6PC3 mutations*. Orphanet J Rare Dis, 2013. **8**: p. 84.
491. Pan, C.J., et al., *Transmembrane topology of glucose-6-phosphatase*. J Biol Chem, 1998. **273**(11): p. 6144-8.
492. Chen, Y. and X. Wang, *miRDB: an online database for prediction of functional microRNA targets*. Nucleic Acids Res, 2020. **48**(D1): p. D127-D131.
493. Liu, W. and X. Wang, *Prediction of functional microRNA targets by integrative modeling of microRNA binding and target expression data*. Genome Biol, 2019. **20**(1): p. 18.
494. Geladaki, A., et al., *Combining LOPIT with differential ultracentrifugation for high-resolution spatial proteomics*. Nat Commun, 2019. **10**(1): p. 331.
495. Queiroz, R.M.L., et al., *Comprehensive identification of RNA-protein interactions in any organism using orthogonal organic phase separation (OOPS)*. Nat Biotechnol, 2019. **37**(2): p. 169-178.
496. Huggett, J., et al., *Real-time RT-PCR normalisation; strategies and considerations*. Genes Immun, 2005. **6**(4): p. 279-84.

497. Wuarin, J. and U. Schibler, *Physical isolation of nascent RNA chains transcribed by RNA polymerase II: evidence for cotranscriptional splicing*. Mol Cell Biol, 1994. **14**(11): p. 7219-25.
498. Pandya-Jones, A. and D.L. Black, *Co-transcriptional splicing of constitutive and alternative exons*. RNA, 2009. **15**(10): p. 1896-908.
499. Ran, F.A., et al., *Genome engineering using the CRISPR-Cas9 system*. Nat Protoc, 2013. **8**(11): p. 2281-2308.
500. Chiang, T.W., et al., *CRISPR-Cas9(D10A) nickase-based genotypic and phenotypic screening to enhance genome editing*. Sci Rep, 2016. **6**: p. 24356.
501. Labun, K., et al., *CHOPCHOP v2: a web tool for the next generation of CRISPR genome engineering*. Nucleic Acids Res, 2016. **44**(W1): p. W272-6.
502. Gearing, M. *CRISPR 101: Cas9 Nickase Design and Homology Directed Repair*. 2018; Available from: <https://blog.addgene.org/crispr-101-cas9-nickase-design-and-homology-directed-repair>.
503. Hu, J.H., et al., *Evolved Cas9 variants with broad PAM compatibility and high DNA specificity*. Nature, 2018. **556**(7699): p. 57-63.
504. Shimatani, Z., et al., *Targeted base editing in rice and tomato using a CRISPR-Cas9 cytidine deaminase fusion*. Nat Biotechnol, 2017. **35**(5): p. 441-443.
505. Komor, A.C., et al., *Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage*. Nature, 2016. **533**(7603): p. 420-4.
506. Gaudelli, N.M., et al., *Programmable base editing of A*T to G*C in genomic DNA without DNA cleavage*. Nature, 2017. **551**(7681): p. 464-471.
507. Komor, A.C., A.H. Badran, and D.R. Liu, *Editing the Genome Without Double-Stranded DNA Breaks*. ACS Chem Biol, 2018. **13**(2): p. 383-388.
508. D'Souza, A.R. and M. Minczuk, *Mitochondrial transcription and translation: overview*. Essays Biochem, 2018. **62**(3): p. 309-320.
509. Voigt, F., et al., *Single-Molecule Quantification of Translation-Dependent Association of mRNAs with the Endoplasmic Reticulum*. Cell Rep, 2017. **21**(13): p. 3740-3753.
510. Maurel, M. and E. Chevet, *Endoplasmic reticulum stress signaling: the microRNA connection*. Am J Physiol Cell Physiol, 2013. **304**(12): p. C1117-26.
511. Malhi, H., *MICRORNAs IN ER STRESS: DIVERGENT ROLES IN CELL FATE DECISIONS*. Curr Pathobiol Rep, 2014. **2**(3): p. 117-122.
512. Leung, A.K., J.M. Calabrese, and P.A. Sharp, *Quantitative analysis of Argonaute protein reveals microRNA-dependent localization to stress granules*. Proc Natl Acad Sci U S A, 2006. **103**(48): p. 18125-30.
513. Jan, C.H., C.C. Williams, and J.S. Weissman, *Principles of ER cotranslational translocation revealed by proximity-specific ribosome profiling*. Science, 2014. **346**(6210): p. 1257521.
514. Fazal, F.M., et al., *Atlas of Subcellular RNA Localization Revealed by APEX-Seq*. Cell, 2019. **178**(2): p. 473-490 e26.
515. Bergstrom, F., et al., *Dimers of dipyrrometheneboron difluoride (BODIPY) with light spectroscopic applications in chemistry and biology*. J Am Chem Soc, 2002. **124**(2): p. 196-204.
516. Yin, V.P., *In Situ Detection of MicroRNA Expression with RNAscope Probes*. Methods Mol Biol, 2018. **1649**: p. 197-208.

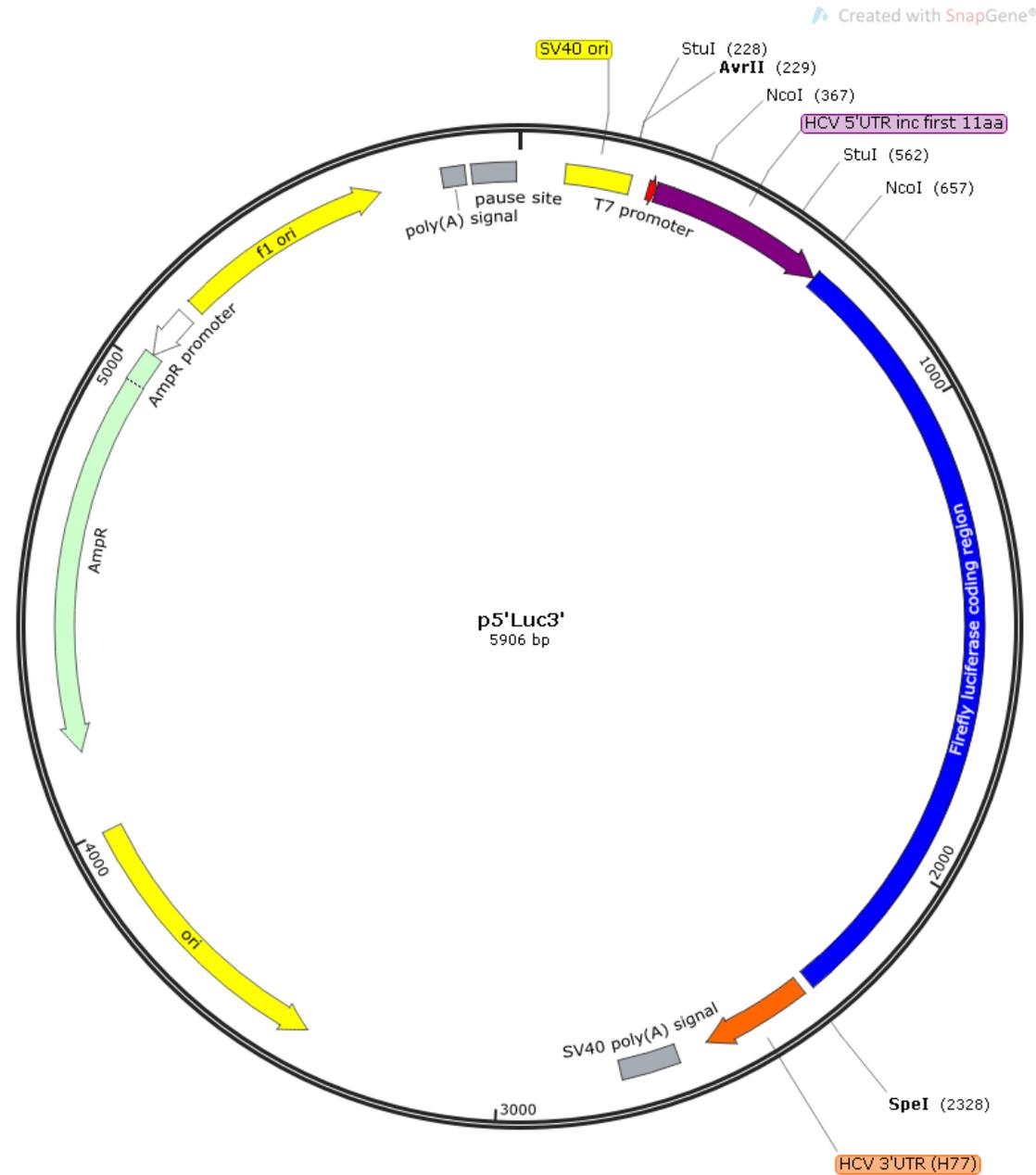
Supplementary Data



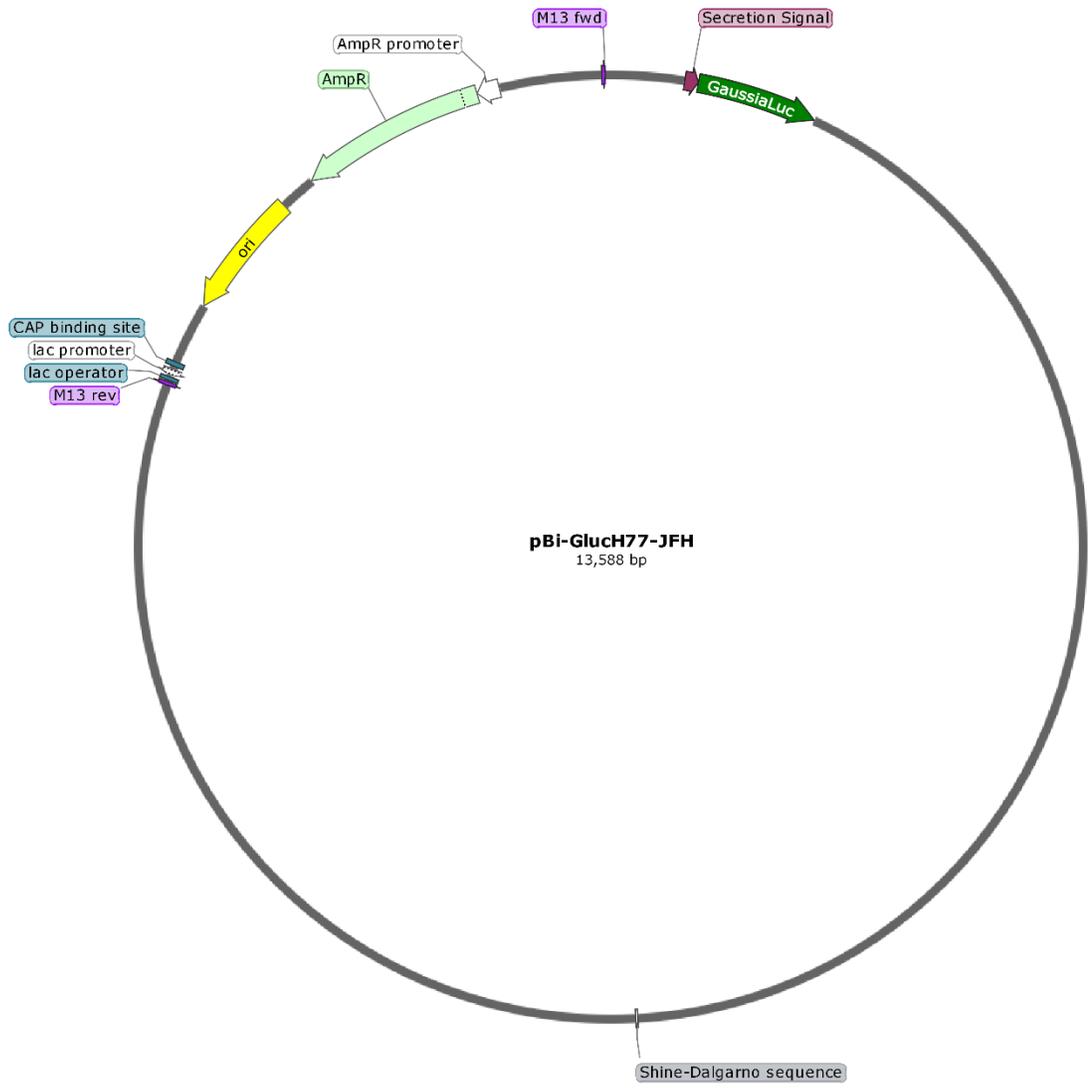
Supplementary Figure 1: Plasmid map of pLuc-122x2. Referred to in the main text as FlucWT. Shows the location of the Firefly luciferase coding region (Blue), miR-122 binding sites (Purple), SV40 promoter, and restriction sites for NcoI, StuI and SpeI.



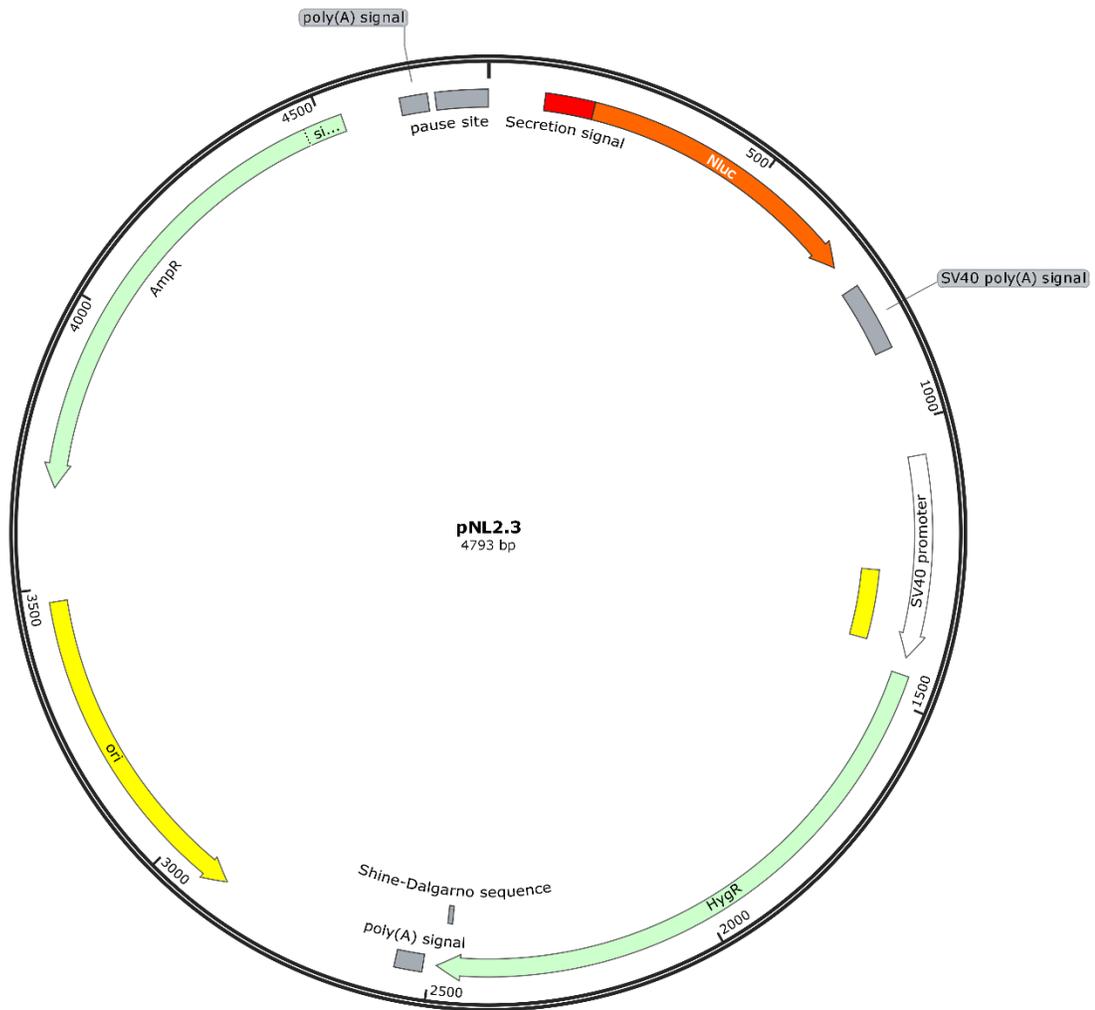
Supplementary Figure 2: Plasmid map of pLuc-122x2m4. Referred to in the main text as FlucM4. Shows the location of the Firefly luciferase coding region (Blue), mutant miR-122 binding sites (Purple), SV40 promoter, and restriction sites for NcoI, StuI and SpeI.



Supplementary Figure 3: Plasmid map of p5'Luc3'. Referred to in the main text as 5'Fluc3'. Shows the location of the Firefly luciferase coding region (Blue), 5'UTR from HCV (Purple), 3'UTR from HCV (Orange), SV40 promoter, and restriction sites for NcoI and SpeI.

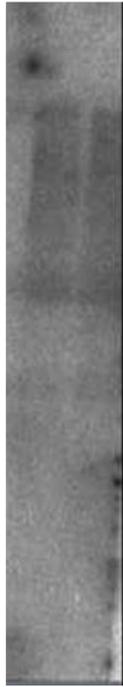


Supplementary Figure 4: Plasmid map of pBi-GlucH77-JFH. Shows the location of the Gaussia luciferase coding region (Green) and secretion signal (Purple).



Supplementary Figure 5: Plasmid map of pNL2.3. Shows the location of the NanoLuc luciferase coding region (Orange), secretion signal (Red) and qRT-PCR primers.

FlucM4



GlucM4



Supplementary Figure 6: Northern Blot analysis of Luciferase Reporters. Total RNA extracted from Huh7 cells 48 hours after transfection of luciferase reporters, probing for Firefly Luciferase (Fluc) and Gaussia Luciferase (Gluc).

Supplementary Table 1: NanoLuc Luciferase Codon Usage Score. Calculated by Christopher Roberts.

| Position | Input Codon | Input codon usage score | Optimal codon | Optimal codon usage score | Bias |
|----------|-------------|-------------------------|---------------|---------------------------|------|
| 1 | ATG | 1 | ATG | 1 | 0 |
| 2 | AAC | 0.53 | AAC | 0.53 | 0 |
| 3 | TCC | 0.22 | AGC | 0.24 | 0.02 |
| 4 | TTC | 0.54 | TTC | 0.54 | 0 |
| 5 | TCC | 0.22 | AGC | 0.24 | 0.02 |
| 6 | ACA | 0.28 | ACC | 0.36 | 0.08 |
| 7 | AGC | 0.24 | AGC | 0.24 | 0 |
| 8 | GCC | 0.4 | GCC | 0.4 | 0 |
| 9 | TTC | 0.54 | TTC | 0.54 | 0 |
| 10 | GGT | 0.16 | GGC | 0.34 | 0.18 |
| 11 | CCA | 0.28 | CCC | 0.32 | 0.04 |
| 12 | GTT | 0.18 | GTG | 0.46 | 0.28 |
| 13 | GCC | 0.4 | GCC | 0.4 | 0 |
| 14 | TTC | 0.54 | TTC | 0.54 | 0 |
| 15 | TCC | 0.22 | AGC | 0.24 | 0.02 |
| 16 | CTG | 0.4 | CTG | 0.4 | 0 |
| 17 | GGC | 0.34 | GGC | 0.34 | 0 |
| 18 | CTG | 0.4 | CTG | 0.4 | 0 |
| 19 | CTC | 0.2 | CTG | 0.4 | 0.2 |
| 20 | CTG | 0.4 | CTG | 0.4 | 0 |
| 21 | GTG | 0.46 | GTG | 0.46 | 0 |
| 22 | TTG | 0.13 | CTG | 0.4 | 0.27 |
| 23 | CCT | 0.29 | CCC | 0.32 | 0.03 |
| 24 | GCT | 0.27 | GCC | 0.4 | 0.13 |
| 25 | GCC | 0.4 | GCC | 0.4 | 0 |
| 26 | TTC | 0.54 | TTC | 0.54 | 0 |
| 27 | CCT | 0.29 | CCC | 0.32 | 0.03 |
| 28 | GCC | 0.4 | GCC | 0.4 | 0 |
| 29 | CCA | 0.28 | CCC | 0.32 | 0.04 |
| 30 | GTC | 0.24 | GTG | 0.46 | 0.22 |
| 31 | TTC | 0.54 | TTC | 0.54 | 0 |
| 32 | ACA | 0.28 | ACC | 0.36 | 0.08 |
| 33 | CTC | 0.2 | CTG | 0.4 | 0.2 |
| 34 | GAA | 0.42 | GAG | 0.58 | 0.16 |
| 35 | GAT | 0.46 | GAC | 0.54 | 0.08 |
| 36 | TTC | 0.54 | TTC | 0.54 | 0 |
| 37 | GTT | 0.18 | GTG | 0.46 | 0.28 |
| 38 | GGG | 0.25 | GGC | 0.34 | 0.09 |
| 39 | GAC | 0.54 | GAC | 0.54 | 0 |
| 40 | TGG | 1 | TGG | 1 | 0 |

| | | | | | |
|----|-----|------|-----|------|------|
| 41 | CGA | 0.11 | AGA | 0.21 | 0.1 |
| 42 | CAG | 0.73 | CAG | 0.73 | 0 |
| 43 | ACA | 0.28 | ACC | 0.36 | 0.08 |
| 44 | GCC | 0.4 | GCC | 0.4 | 0 |
| 45 | GGC | 0.34 | GGC | 0.34 | 0 |
| 46 | TAC | 0.56 | TAC | 0.56 | 0 |
| 47 | AAC | 0.53 | AAC | 0.53 | 0 |
| 48 | CTG | 0.4 | CTG | 0.4 | 0 |
| 49 | GAC | 0.54 | GAC | 0.54 | 0 |
| 50 | CAA | 0.27 | CAG | 0.73 | 0.46 |
| 51 | GTC | 0.24 | GTG | 0.46 | 0.22 |
| 52 | CTT | 0.13 | CTG | 0.4 | 0.27 |
| 53 | GAA | 0.42 | GAG | 0.58 | 0.16 |
| 54 | CAG | 0.73 | CAG | 0.73 | 0 |
| 55 | GGA | 0.25 | GGC | 0.34 | 0.09 |
| 56 | GGT | 0.16 | GGC | 0.34 | 0.18 |
| 57 | GTG | 0.46 | GTG | 0.46 | 0 |
| 58 | TCC | 0.22 | AGC | 0.24 | 0.02 |
| 59 | AGT | 0.15 | AGC | 0.24 | 0.09 |
| 60 | TTG | 0.13 | CTG | 0.4 | 0.27 |
| 61 | TTT | 0.46 | TTC | 0.54 | 0.08 |
| 62 | CAG | 0.73 | CAG | 0.73 | 0 |
| 63 | AAT | 0.47 | AAC | 0.53 | 0.06 |
| 64 | CTC | 0.2 | CTG | 0.4 | 0.2 |
| 65 | GGG | 0.25 | GGC | 0.34 | 0.09 |
| 66 | GTG | 0.46 | GTG | 0.46 | 0 |
| 67 | TCC | 0.22 | AGC | 0.24 | 0.02 |
| 68 | GTA | 0.12 | GTG | 0.46 | 0.34 |
| 69 | ACT | 0.25 | ACC | 0.36 | 0.11 |
| 70 | CCG | 0.11 | CCC | 0.32 | 0.21 |
| 71 | ATC | 0.47 | ATC | 0.47 | 0 |
| 72 | CAA | 0.27 | CAG | 0.73 | 0.46 |
| 73 | AGG | 0.21 | AGA | 0.21 | 0 |
| 74 | ATT | 0.36 | ATC | 0.47 | 0.11 |
| 75 | GTC | 0.24 | GTG | 0.46 | 0.22 |
| 76 | CTG | 0.4 | CTG | 0.4 | 0 |
| 77 | AGC | 0.24 | AGC | 0.24 | 0 |
| 78 | GGT | 0.16 | GGC | 0.34 | 0.18 |
| 79 | GAA | 0.42 | GAG | 0.58 | 0.16 |
| 80 | AAT | 0.47 | AAC | 0.53 | 0.06 |
| 81 | GGG | 0.25 | GGC | 0.34 | 0.09 |
| 82 | CTG | 0.4 | CTG | 0.4 | 0 |
| 83 | AAG | 0.57 | AAG | 0.57 | 0 |
| 84 | ATC | 0.47 | ATC | 0.47 | 0 |
| 85 | GAC | 0.54 | GAC | 0.54 | 0 |
| 86 | ATC | 0.47 | ATC | 0.47 | 0 |
| 87 | CAT | 0.42 | CAC | 0.58 | 0.16 |

| | | | | | |
|-----|-----|------|-----|------|------|
| 88 | GTC | 0.24 | GTG | 0.46 | 0.22 |
| 89 | ATC | 0.47 | ATC | 0.47 | 0 |
| 90 | ATC | 0.47 | ATC | 0.47 | 0 |
| 91 | CCG | 0.11 | CCC | 0.32 | 0.21 |
| 92 | TAT | 0.44 | TAC | 0.56 | 0.12 |
| 93 | GAA | 0.42 | GAG | 0.58 | 0.16 |
| 94 | GGT | 0.16 | GGC | 0.34 | 0.18 |
| 95 | CTG | 0.4 | CTG | 0.4 | 0 |
| 96 | AGC | 0.24 | AGC | 0.24 | 0 |
| 97 | GGC | 0.34 | GGC | 0.34 | 0 |
| 98 | GAC | 0.54 | GAC | 0.54 | 0 |
| 99 | CAA | 0.27 | CAG | 0.73 | 0.46 |
| 100 | ATG | 1 | ATG | 1 | 0 |
| 101 | GGC | 0.34 | GGC | 0.34 | 0 |
| 102 | CAG | 0.73 | CAG | 0.73 | 0 |
| 103 | ATC | 0.47 | ATC | 0.47 | 0 |
| 104 | GAA | 0.42 | GAG | 0.58 | 0.16 |
| 105 | AAA | 0.43 | AAG | 0.57 | 0.14 |
| 106 | ATT | 0.36 | ATC | 0.47 | 0.11 |
| 107 | TTT | 0.46 | TTC | 0.54 | 0.08 |
| 108 | AAG | 0.57 | AAG | 0.57 | 0 |
| 109 | GTG | 0.46 | GTG | 0.46 | 0 |
| 110 | GTG | 0.46 | GTG | 0.46 | 0 |
| 111 | TAC | 0.56 | TAC | 0.56 | 0 |
| 112 | CCT | 0.29 | CCC | 0.32 | 0.03 |
| 113 | GTG | 0.46 | GTG | 0.46 | 0 |
| 114 | GAT | 0.46 | GAC | 0.54 | 0.08 |
| 115 | GAT | 0.46 | GAC | 0.54 | 0.08 |
| 116 | CAT | 0.42 | CAC | 0.58 | 0.16 |
| 117 | CAC | 0.58 | CAC | 0.58 | 0 |
| 118 | TTT | 0.46 | TTC | 0.54 | 0.08 |
| 119 | AAG | 0.57 | AAG | 0.57 | 0 |
| 120 | GTG | 0.46 | GTG | 0.46 | 0 |
| 121 | ATC | 0.47 | ATC | 0.47 | 0 |
| 122 | CTG | 0.4 | CTG | 0.4 | 0 |
| 123 | CAC | 0.58 | CAC | 0.58 | 0 |
| 124 | TAT | 0.44 | TAC | 0.56 | 0.12 |
| 125 | GGC | 0.34 | GGC | 0.34 | 0 |
| 126 | ACA | 0.28 | ACC | 0.36 | 0.08 |
| 127 | CTG | 0.4 | CTG | 0.4 | 0 |
| 128 | GTA | 0.12 | GTG | 0.46 | 0.34 |
| 129 | ATC | 0.47 | ATC | 0.47 | 0 |
| 130 | GAC | 0.54 | GAC | 0.54 | 0 |
| 131 | GGG | 0.25 | GGC | 0.34 | 0.09 |
| 132 | GTT | 0.18 | GTG | 0.46 | 0.28 |
| 133 | ACG | 0.11 | ACC | 0.36 | 0.25 |
| 134 | CCG | 0.11 | CCC | 0.32 | 0.21 |

| | | | | | |
|-----|-----|------|-----|------|------|
| 135 | AAC | 0.53 | AAC | 0.53 | 0 |
| 136 | ATG | 1 | ATG | 1 | 0 |
| 137 | ATC | 0.47 | ATC | 0.47 | 0 |
| 138 | GAC | 0.54 | GAC | 0.54 | 0 |
| 139 | TAT | 0.44 | TAC | 0.56 | 0.12 |
| 140 | TTC | 0.54 | TTC | 0.54 | 0 |
| 141 | GGA | 0.25 | GGC | 0.34 | 0.09 |
| 142 | CGG | 0.2 | AGA | 0.21 | 0.01 |
| 143 | CCG | 0.11 | CCC | 0.32 | 0.21 |
| 144 | TAT | 0.44 | TAC | 0.56 | 0.12 |
| 145 | GAA | 0.42 | GAG | 0.58 | 0.16 |
| 146 | GGC | 0.34 | GGC | 0.34 | 0 |
| 147 | ATC | 0.47 | ATC | 0.47 | 0 |
| 148 | GCC | 0.4 | GCC | 0.4 | 0 |
| 149 | GTG | 0.46 | GTG | 0.46 | 0 |
| 150 | TTC | 0.54 | TTC | 0.54 | 0 |
| 151 | GAC | 0.54 | GAC | 0.54 | 0 |
| 152 | GGC | 0.34 | GGC | 0.34 | 0 |
| 153 | AAA | 0.43 | AAG | 0.57 | 0.14 |
| 154 | AAG | 0.57 | AAG | 0.57 | 0 |
| 155 | ATC | 0.47 | ATC | 0.47 | 0 |
| 156 | ACT | 0.25 | ACC | 0.36 | 0.11 |
| 157 | GTA | 0.12 | GTG | 0.46 | 0.34 |
| 158 | ACA | 0.28 | ACC | 0.36 | 0.08 |
| 159 | GGG | 0.25 | GGC | 0.34 | 0.09 |
| 160 | ACC | 0.36 | ACC | 0.36 | 0 |
| 161 | CTG | 0.4 | CTG | 0.4 | 0 |
| 162 | TGG | 1 | TGG | 1 | 0 |
| 163 | AAC | 0.53 | AAC | 0.53 | 0 |
| 164 | GGC | 0.34 | GGC | 0.34 | 0 |
| 165 | AAC | 0.53 | AAC | 0.53 | 0 |
| 166 | AAA | 0.43 | AAG | 0.57 | 0.14 |
| 167 | ATT | 0.36 | ATC | 0.47 | 0.11 |
| 168 | ATC | 0.47 | ATC | 0.47 | 0 |
| 169 | GAC | 0.54 | GAC | 0.54 | 0 |
| 170 | GAG | 0.58 | GAG | 0.58 | 0 |
| 171 | CGC | 0.18 | AGA | 0.21 | 0.03 |
| 172 | CTG | 0.4 | CTG | 0.4 | 0 |
| 173 | ATC | 0.47 | ATC | 0.47 | 0 |
| 174 | AAC | 0.53 | AAC | 0.53 | 0 |
| 175 | CCC | 0.32 | CCC | 0.32 | 0 |
| 176 | GAC | 0.54 | GAC | 0.54 | 0 |
| 177 | GGC | 0.34 | GGC | 0.34 | 0 |
| 178 | TCC | 0.22 | AGC | 0.24 | 0.02 |
| 179 | CTG | 0.4 | CTG | 0.4 | 0 |
| 180 | CTG | 0.4 | CTG | 0.4 | 0 |
| 181 | TTC | 0.54 | TTC | 0.54 | 0 |

| | | | | | |
|-----|-----|------|-----|--------------|------|
| 182 | CGA | 0.11 | AGA | 0.21 | 0.1 |
| 183 | GTA | 0.12 | GTG | 0.46 | 0.34 |
| 184 | ACC | 0.36 | ACC | 0.36 | 0 |
| 185 | ATC | 0.47 | ATC | 0.47 | 0 |
| 186 | AAC | 0.53 | AAC | 0.53 | 0 |
| 187 | GGA | 0.25 | GGC | 0.34 | 0.09 |
| 188 | GTG | 0.46 | GTG | 0.46 | 0 |
| 189 | ACC | 0.36 | ACC | 0.36 | 0 |
| 190 | GGC | 0.34 | GGC | 0.34 | 0 |
| 191 | TGG | 1 | TGG | 1 | 0 |
| 192 | CGG | 0.2 | AGA | 0.21 | 0.01 |
| 193 | CTG | 0.4 | CTG | 0.4 | 0 |
| 194 | TGC | 0.54 | TGC | 0.54 | 0 |
| 195 | GAA | 0.42 | GAG | 0.58 | 0.16 |
| 196 | CGC | 0.18 | AGA | 0.21 | 0.03 |
| 197 | ATT | 0.36 | ATC | 0.47 | 0.11 |
| 198 | CTG | 0.4 | CTG | 0.4 | 0 |
| 199 | GCG | 0.11 | GCC | 0.4 | 0.29 |
| 200 | TAA | 0.3 | TGA | 0.47 | |
| | | | | Score | 0.07 |

Supplementary Table 2: Firefly Luciferase Codon Usage Score. Calculated by Christopher Roberts.

| Position | Input Codon | Input codon usage score | Optimal codon | Optimal codon usage score | Bias |
|----------|-------------|-------------------------|---------------|---------------------------|------|
| 1 | ATG | 1 | ATG | 1 | 0 |
| 2 | GAA | 0.42 | GAG | 0.58 | 0.16 |
| 3 | GAC | 0.54 | GAC | 0.54 | 0 |
| 4 | GCC | 0.4 | GCC | 0.4 | 0 |
| 5 | AAA | 0.43 | AAG | 0.57 | 0.14 |
| 6 | AAC | 0.53 | AAC | 0.53 | 0 |
| 7 | ATA | 0.17 | ATC | 0.47 | 0.3 |
| 8 | AAG | 0.57 | AAG | 0.57 | 0 |
| 9 | AAA | 0.43 | AAG | 0.57 | 0.14 |
| 10 | GGC | 0.34 | GGC | 0.34 | 0 |
| 11 | CCG | 0.11 | CCC | 0.32 | 0.21 |
| 12 | GCG | 0.11 | GCC | 0.4 | 0.29 |
| 13 | CCA | 0.28 | CCC | 0.32 | 0.04 |
| 14 | TTC | 0.54 | TTC | 0.54 | 0 |
| 15 | TAT | 0.44 | TAC | 0.56 | 0.12 |
| 16 | CCG | 0.11 | CCC | 0.32 | 0.21 |
| 17 | CTG | 0.4 | CTG | 0.4 | 0 |
| 18 | GAA | 0.42 | GAG | 0.58 | 0.16 |
| 19 | GAT | 0.46 | GAC | 0.54 | 0.08 |
| 20 | GGA | 0.25 | GGC | 0.34 | 0.09 |
| 21 | ACC | 0.36 | ACC | 0.36 | 0 |
| 22 | GCT | 0.27 | GCC | 0.4 | 0.13 |
| 23 | GGA | 0.25 | GGC | 0.34 | 0.09 |
| 24 | GAG | 0.58 | GAG | 0.58 | 0 |
| 25 | CAA | 0.27 | CAG | 0.73 | 0.46 |
| 26 | CTG | 0.4 | CTG | 0.4 | 0 |
| 27 | CAT | 0.42 | CAC | 0.58 | 0.16 |
| 28 | AAG | 0.57 | AAG | 0.57 | 0 |
| 29 | GCT | 0.27 | GCC | 0.4 | 0.13 |
| 30 | ATG | 1 | ATG | 1 | 0 |
| 31 | AAG | 0.57 | AAG | 0.57 | 0 |
| 32 | AGA | 0.21 | AGA | 0.21 | 0 |
| 33 | TAC | 0.56 | TAC | 0.56 | 0 |
| 34 | GCC | 0.4 | GCC | 0.4 | 0 |
| 35 | CTG | 0.4 | CTG | 0.4 | 0 |
| 36 | GTT | 0.18 | GTG | 0.46 | 0.28 |
| 37 | CCT | 0.29 | CCC | 0.32 | 0.03 |
| 38 | GGA | 0.25 | GGC | 0.34 | 0.09 |
| 39 | ACA | 0.28 | ACC | 0.36 | 0.08 |
| 40 | ATT | 0.36 | ATC | 0.47 | 0.11 |
| 41 | GCT | 0.27 | GCC | 0.4 | 0.13 |
| 42 | TTT | 0.46 | TTC | 0.54 | 0.08 |

| | | | | | |
|----|-----|------|-----|------|------|
| 43 | ACA | 0.28 | ACC | 0.36 | 0.08 |
| 44 | GAT | 0.46 | GAC | 0.54 | 0.08 |
| 45 | GCA | 0.23 | GCC | 0.4 | 0.17 |
| 46 | CAT | 0.42 | CAC | 0.58 | 0.16 |
| 47 | ATC | 0.47 | ATC | 0.47 | 0 |
| 48 | GAG | 0.58 | GAG | 0.58 | 0 |
| 49 | GTG | 0.46 | GTG | 0.46 | 0 |
| 50 | GAC | 0.54 | GAC | 0.54 | 0 |
| 51 | ATC | 0.47 | ATC | 0.47 | 0 |
| 52 | ACT | 0.25 | ACC | 0.36 | 0.11 |
| 53 | TAC | 0.56 | TAC | 0.56 | 0 |
| 54 | GCT | 0.27 | GCC | 0.4 | 0.13 |
| 55 | GAG | 0.58 | GAG | 0.58 | 0 |
| 56 | TAC | 0.56 | TAC | 0.56 | 0 |
| 57 | TTC | 0.54 | TTC | 0.54 | 0 |
| 58 | GAA | 0.42 | GAG | 0.58 | 0.16 |
| 59 | ATG | 1 | ATG | 1 | 0 |
| 60 | TCC | 0.22 | AGC | 0.24 | 0.02 |
| 61 | GTT | 0.18 | GTG | 0.46 | 0.28 |
| 62 | CGG | 0.2 | AGA | 0.21 | 0.01 |
| 63 | TTG | 0.13 | CTG | 0.4 | 0.27 |
| 64 | GCA | 0.23 | GCC | 0.4 | 0.17 |
| 65 | GAA | 0.42 | GAG | 0.58 | 0.16 |
| 66 | GCT | 0.27 | GCC | 0.4 | 0.13 |
| 67 | ATG | 1 | ATG | 1 | 0 |
| 68 | AAA | 0.43 | AAG | 0.57 | 0.14 |
| 69 | CGA | 0.11 | AGA | 0.21 | 0.1 |
| 70 | TAT | 0.44 | TAC | 0.56 | 0.12 |
| 71 | GGG | 0.25 | GGC | 0.34 | 0.09 |
| 72 | CTG | 0.4 | CTG | 0.4 | 0 |
| 73 | AAT | 0.47 | AAC | 0.53 | 0.06 |
| 74 | ACA | 0.28 | ACC | 0.36 | 0.08 |
| 75 | AAT | 0.47 | AAC | 0.53 | 0.06 |
| 76 | CAC | 0.58 | CAC | 0.58 | 0 |
| 77 | AGA | 0.21 | AGA | 0.21 | 0 |
| 78 | ATC | 0.47 | ATC | 0.47 | 0 |
| 79 | GTC | 0.24 | GTG | 0.46 | 0.22 |
| 80 | GTA | 0.12 | GTG | 0.46 | 0.34 |
| 81 | TGC | 0.54 | TGC | 0.54 | 0 |
| 82 | AGT | 0.15 | AGC | 0.24 | 0.09 |
| 83 | GAA | 0.42 | GAG | 0.58 | 0.16 |
| 84 | AAC | 0.53 | AAC | 0.53 | 0 |
| 85 | TCT | 0.19 | AGC | 0.24 | 0.05 |
| 86 | CTT | 0.13 | CTG | 0.4 | 0.27 |
| 87 | CAA | 0.27 | CAG | 0.73 | 0.46 |
| 88 | TTC | 0.54 | TTC | 0.54 | 0 |
| 89 | TTT | 0.46 | TTC | 0.54 | 0.08 |

| | | | | | |
|-----|-----|------|-----|------|------|
| 90 | ATG | 1 | ATG | 1 | 0 |
| 91 | CCG | 0.11 | CCC | 0.32 | 0.21 |
| 92 | GTG | 0.46 | GTG | 0.46 | 0 |
| 93 | TTG | 0.13 | CTG | 0.4 | 0.27 |
| 94 | GGC | 0.34 | GGC | 0.34 | 0 |
| 95 | GCG | 0.11 | GCC | 0.4 | 0.29 |
| 96 | TTA | 0.08 | CTG | 0.4 | 0.32 |
| 97 | TTT | 0.46 | TTC | 0.54 | 0.08 |
| 98 | ATC | 0.47 | ATC | 0.47 | 0 |
| 99 | GGA | 0.25 | GGC | 0.34 | 0.09 |
| 100 | GTT | 0.18 | GTG | 0.46 | 0.28 |
| 101 | GCA | 0.23 | GCC | 0.4 | 0.17 |
| 102 | GTT | 0.18 | GTG | 0.46 | 0.28 |
| 103 | GCG | 0.11 | GCC | 0.4 | 0.29 |
| 104 | CCC | 0.32 | CCC | 0.32 | 0 |
| 105 | GCG | 0.11 | GCC | 0.4 | 0.29 |
| 106 | AAC | 0.53 | AAC | 0.53 | 0 |
| 107 | GAC | 0.54 | GAC | 0.54 | 0 |
| 108 | ATT | 0.36 | ATC | 0.47 | 0.11 |
| 109 | TAT | 0.44 | TAC | 0.56 | 0.12 |
| 110 | AAT | 0.47 | AAC | 0.53 | 0.06 |
| 111 | GAA | 0.42 | GAG | 0.58 | 0.16 |
| 112 | CGT | 0.08 | AGA | 0.21 | 0.13 |
| 113 | GAA | 0.42 | GAG | 0.58 | 0.16 |
| 114 | TTG | 0.13 | CTG | 0.4 | 0.27 |
| 115 | CTC | 0.2 | CTG | 0.4 | 0.2 |
| 116 | AAC | 0.53 | AAC | 0.53 | 0 |
| 117 | AGT | 0.15 | AGC | 0.24 | 0.09 |
| 118 | ATG | 1 | ATG | 1 | 0 |
| 119 | GGC | 0.34 | GGC | 0.34 | 0 |
| 120 | ATT | 0.36 | ATC | 0.47 | 0.11 |
| 121 | TCG | 0.05 | AGC | 0.24 | 0.19 |
| 122 | CAG | 0.73 | CAG | 0.73 | 0 |
| 123 | CCT | 0.29 | CCC | 0.32 | 0.03 |
| 124 | ACC | 0.36 | ACC | 0.36 | 0 |
| 125 | GTG | 0.46 | GTG | 0.46 | 0 |
| 126 | GTG | 0.46 | GTG | 0.46 | 0 |
| 127 | TTC | 0.54 | TTC | 0.54 | 0 |
| 128 | GTT | 0.18 | GTG | 0.46 | 0.28 |
| 129 | TCC | 0.22 | AGC | 0.24 | 0.02 |
| 130 | AAA | 0.43 | AAG | 0.57 | 0.14 |
| 131 | AAG | 0.57 | AAG | 0.57 | 0 |
| 132 | GGG | 0.25 | GGC | 0.34 | 0.09 |
| 133 | TTG | 0.13 | CTG | 0.4 | 0.27 |
| 134 | CAA | 0.27 | CAG | 0.73 | 0.46 |
| 135 | AAA | 0.43 | AAG | 0.57 | 0.14 |
| 136 | ATT | 0.36 | ATC | 0.47 | 0.11 |

| | | | | | |
|-----|-----|------|-----|------|------|
| 137 | TTG | 0.13 | CTG | 0.4 | 0.27 |
| 138 | AAC | 0.53 | AAC | 0.53 | 0 |
| 139 | GTG | 0.46 | GTG | 0.46 | 0 |
| 140 | CAA | 0.27 | CAG | 0.73 | 0.46 |
| 141 | AAA | 0.43 | AAG | 0.57 | 0.14 |
| 142 | AAG | 0.57 | AAG | 0.57 | 0 |
| 143 | CTC | 0.2 | CTG | 0.4 | 0.2 |
| 144 | CCA | 0.28 | CCC | 0.32 | 0.04 |
| 145 | ATC | 0.47 | ATC | 0.47 | 0 |
| 146 | ATC | 0.47 | ATC | 0.47 | 0 |
| 147 | CAA | 0.27 | CAG | 0.73 | 0.46 |
| 148 | AAA | 0.43 | AAG | 0.57 | 0.14 |
| 149 | ATT | 0.36 | ATC | 0.47 | 0.11 |
| 150 | ATT | 0.36 | ATC | 0.47 | 0.11 |
| 151 | ATC | 0.47 | ATC | 0.47 | 0 |
| 152 | ATG | 1 | ATG | 1 | 0 |
| 153 | GAT | 0.46 | GAC | 0.54 | 0.08 |
| 154 | TCT | 0.19 | AGC | 0.24 | 0.05 |
| 155 | AAA | 0.43 | AAG | 0.57 | 0.14 |
| 156 | ACG | 0.11 | ACC | 0.36 | 0.25 |
| 157 | GAT | 0.46 | GAC | 0.54 | 0.08 |
| 158 | TAC | 0.56 | TAC | 0.56 | 0 |
| 159 | CAG | 0.73 | CAG | 0.73 | 0 |
| 160 | GGA | 0.25 | GGC | 0.34 | 0.09 |
| 161 | TTT | 0.46 | TTC | 0.54 | 0.08 |
| 162 | CAG | 0.73 | CAG | 0.73 | 0 |
| 163 | TCG | 0.05 | AGC | 0.24 | 0.19 |
| 164 | ATG | 1 | ATG | 1 | 0 |
| 165 | TAC | 0.56 | TAC | 0.56 | 0 |
| 166 | ACG | 0.11 | ACC | 0.36 | 0.25 |
| 167 | TTC | 0.54 | TTC | 0.54 | 0 |
| 168 | GTC | 0.24 | GTG | 0.46 | 0.22 |
| 169 | ACA | 0.28 | ACC | 0.36 | 0.08 |
| 170 | TCT | 0.19 | AGC | 0.24 | 0.05 |
| 171 | CAT | 0.42 | CAC | 0.58 | 0.16 |
| 172 | CTA | 0.07 | CTG | 0.4 | 0.33 |
| 173 | CCT | 0.29 | CCC | 0.32 | 0.03 |
| 174 | CCC | 0.32 | CCC | 0.32 | 0 |
| 175 | GGT | 0.16 | GGC | 0.34 | 0.18 |
| 176 | TTT | 0.46 | TTC | 0.54 | 0.08 |
| 177 | AAT | 0.47 | AAC | 0.53 | 0.06 |
| 178 | GAA | 0.42 | GAG | 0.58 | 0.16 |
| 179 | TAC | 0.56 | TAC | 0.56 | 0 |
| 180 | GAT | 0.46 | GAC | 0.54 | 0.08 |
| 181 | TTT | 0.46 | TTC | 0.54 | 0.08 |
| 182 | GTG | 0.46 | GTG | 0.46 | 0 |
| 183 | CCA | 0.28 | CCC | 0.32 | 0.04 |

| | | | | | |
|-----|-----|------|-----|------|------|
| 184 | GAG | 0.58 | GAG | 0.58 | 0 |
| 185 | TCC | 0.22 | AGC | 0.24 | 0.02 |
| 186 | TTC | 0.54 | TTC | 0.54 | 0 |
| 187 | GAT | 0.46 | GAC | 0.54 | 0.08 |
| 188 | AGG | 0.21 | AGA | 0.21 | 0 |
| 189 | GAC | 0.54 | GAC | 0.54 | 0 |
| 190 | AAG | 0.57 | AAG | 0.57 | 0 |
| 191 | ACA | 0.28 | ACC | 0.36 | 0.08 |
| 192 | ATT | 0.36 | ATC | 0.47 | 0.11 |
| 193 | GCA | 0.23 | GCC | 0.4 | 0.17 |
| 194 | CTG | 0.4 | CTG | 0.4 | 0 |
| 195 | ATC | 0.47 | ATC | 0.47 | 0 |
| 196 | ATG | 1 | ATG | 1 | 0 |
| 197 | AAC | 0.53 | AAC | 0.53 | 0 |
| 198 | TCC | 0.22 | AGC | 0.24 | 0.02 |
| 199 | TCT | 0.19 | AGC | 0.24 | 0.05 |
| 200 | GGA | 0.25 | GGC | 0.34 | 0.09 |
| 201 | TCT | 0.19 | AGC | 0.24 | 0.05 |
| 202 | ACT | 0.25 | ACC | 0.36 | 0.11 |
| 203 | GGT | 0.16 | GGC | 0.34 | 0.18 |
| 204 | CTG | 0.4 | CTG | 0.4 | 0 |
| 205 | CCT | 0.29 | CCC | 0.32 | 0.03 |
| 206 | AAA | 0.43 | AAG | 0.57 | 0.14 |
| 207 | GGT | 0.16 | GGC | 0.34 | 0.18 |
| 208 | GTC | 0.24 | GTG | 0.46 | 0.22 |
| 209 | GCT | 0.27 | GCC | 0.4 | 0.13 |
| 210 | CTG | 0.4 | CTG | 0.4 | 0 |
| 211 | CCT | 0.29 | CCC | 0.32 | 0.03 |
| 212 | CAT | 0.42 | CAC | 0.58 | 0.16 |
| 213 | AGA | 0.21 | AGA | 0.21 | 0 |
| 214 | ACT | 0.25 | ACC | 0.36 | 0.11 |
| 215 | GCC | 0.4 | GCC | 0.4 | 0 |
| 216 | TGC | 0.54 | TGC | 0.54 | 0 |
| 217 | GTG | 0.46 | GTG | 0.46 | 0 |
| 218 | AGA | 0.21 | AGA | 0.21 | 0 |
| 219 | TTC | 0.54 | TTC | 0.54 | 0 |
| 220 | TCG | 0.05 | AGC | 0.24 | 0.19 |
| 221 | CAT | 0.42 | CAC | 0.58 | 0.16 |
| 222 | GCC | 0.4 | GCC | 0.4 | 0 |
| 223 | AGA | 0.21 | AGA | 0.21 | 0 |
| 224 | GAT | 0.46 | GAC | 0.54 | 0.08 |
| 225 | CCT | 0.29 | CCC | 0.32 | 0.03 |
| 226 | ATT | 0.36 | ATC | 0.47 | 0.11 |
| 227 | TTT | 0.46 | TTC | 0.54 | 0.08 |
| 228 | GGC | 0.34 | GGC | 0.34 | 0 |
| 229 | AAT | 0.47 | AAC | 0.53 | 0.06 |
| 230 | CAA | 0.27 | CAG | 0.73 | 0.46 |

| | | | | | |
|-----|-----|------|-----|------|------|
| 231 | ATC | 0.47 | ATC | 0.47 | 0 |
| 232 | ATT | 0.36 | ATC | 0.47 | 0.11 |
| 233 | CCG | 0.11 | CCC | 0.32 | 0.21 |
| 234 | GAT | 0.46 | GAC | 0.54 | 0.08 |
| 235 | ACT | 0.25 | ACC | 0.36 | 0.11 |
| 236 | GCG | 0.11 | GCC | 0.4 | 0.29 |
| 237 | ATT | 0.36 | ATC | 0.47 | 0.11 |
| 238 | TTA | 0.08 | CTG | 0.4 | 0.32 |
| 239 | AGT | 0.15 | AGC | 0.24 | 0.09 |
| 240 | GTT | 0.18 | GTG | 0.46 | 0.28 |
| 241 | GTT | 0.18 | GTG | 0.46 | 0.28 |
| 242 | CCA | 0.28 | CCC | 0.32 | 0.04 |
| 243 | TTC | 0.54 | TTC | 0.54 | 0 |
| 244 | CAT | 0.42 | CAC | 0.58 | 0.16 |
| 245 | CAC | 0.58 | CAC | 0.58 | 0 |
| 246 | GGT | 0.16 | GGC | 0.34 | 0.18 |
| 247 | TTT | 0.46 | TTC | 0.54 | 0.08 |
| 248 | GGA | 0.25 | GGC | 0.34 | 0.09 |
| 249 | ATG | 1 | ATG | 1 | 0 |
| 250 | TTT | 0.46 | TTC | 0.54 | 0.08 |
| 251 | ACT | 0.25 | ACC | 0.36 | 0.11 |
| 252 | ACA | 0.28 | ACC | 0.36 | 0.08 |
| 253 | CTC | 0.2 | CTG | 0.4 | 0.2 |
| 254 | GGA | 0.25 | GGC | 0.34 | 0.09 |
| 255 | TAT | 0.44 | TAC | 0.56 | 0.12 |
| 256 | TTG | 0.13 | CTG | 0.4 | 0.27 |
| 257 | ATA | 0.17 | ATC | 0.47 | 0.3 |
| 258 | TGT | 0.46 | TGC | 0.54 | 0.08 |
| 259 | GGA | 0.25 | GGC | 0.34 | 0.09 |
| 260 | TTT | 0.46 | TTC | 0.54 | 0.08 |
| 261 | CGA | 0.11 | AGA | 0.21 | 0.1 |
| 262 | GTC | 0.24 | GTG | 0.46 | 0.22 |
| 263 | GTC | 0.24 | GTG | 0.46 | 0.22 |
| 264 | TTA | 0.08 | CTG | 0.4 | 0.32 |
| 265 | ATG | 1 | ATG | 1 | 0 |
| 266 | TAT | 0.44 | TAC | 0.56 | 0.12 |
| 267 | AGA | 0.21 | AGA | 0.21 | 0 |
| 268 | TTT | 0.46 | TTC | 0.54 | 0.08 |
| 269 | GAA | 0.42 | GAG | 0.58 | 0.16 |
| 270 | GAA | 0.42 | GAG | 0.58 | 0.16 |
| 271 | GAG | 0.58 | GAG | 0.58 | 0 |
| 272 | CTG | 0.4 | CTG | 0.4 | 0 |
| 273 | TTT | 0.46 | TTC | 0.54 | 0.08 |
| 274 | CTG | 0.4 | CTG | 0.4 | 0 |
| 275 | AGG | 0.21 | AGA | 0.21 | 0 |
| 276 | AGC | 0.24 | AGC | 0.24 | 0 |
| 277 | CTT | 0.13 | CTG | 0.4 | 0.27 |

| | | | | | |
|-----|-----|------|-----|------|------|
| 278 | CAG | 0.73 | CAG | 0.73 | 0 |
| 279 | GAT | 0.46 | GAC | 0.54 | 0.08 |
| 280 | TAC | 0.56 | TAC | 0.56 | 0 |
| 281 | AAG | 0.57 | AAG | 0.57 | 0 |
| 282 | ATT | 0.36 | ATC | 0.47 | 0.11 |
| 283 | CAA | 0.27 | CAG | 0.73 | 0.46 |
| 284 | AGT | 0.15 | AGC | 0.24 | 0.09 |
| 285 | GCG | 0.11 | GCC | 0.4 | 0.29 |
| 286 | CTG | 0.4 | CTG | 0.4 | 0 |
| 287 | CTG | 0.4 | CTG | 0.4 | 0 |
| 288 | GTG | 0.46 | GTG | 0.46 | 0 |
| 289 | CCA | 0.28 | CCC | 0.32 | 0.04 |
| 290 | ACC | 0.36 | ACC | 0.36 | 0 |
| 291 | CTA | 0.07 | CTG | 0.4 | 0.33 |
| 292 | TTC | 0.54 | TTC | 0.54 | 0 |
| 293 | TCC | 0.22 | AGC | 0.24 | 0.02 |
| 294 | TTC | 0.54 | TTC | 0.54 | 0 |
| 295 | TTC | 0.54 | TTC | 0.54 | 0 |
| 296 | GCC | 0.4 | GCC | 0.4 | 0 |
| 297 | AAA | 0.43 | AAG | 0.57 | 0.14 |
| 298 | AGC | 0.24 | AGC | 0.24 | 0 |
| 299 | ACT | 0.25 | ACC | 0.36 | 0.11 |
| 300 | CTG | 0.4 | CTG | 0.4 | 0 |
| 301 | ATT | 0.36 | ATC | 0.47 | 0.11 |
| 302 | GAC | 0.54 | GAC | 0.54 | 0 |
| 303 | AAA | 0.43 | AAG | 0.57 | 0.14 |
| 304 | TAC | 0.56 | TAC | 0.56 | 0 |
| 305 | GAT | 0.46 | GAC | 0.54 | 0.08 |
| 306 | TTA | 0.08 | CTG | 0.4 | 0.32 |
| 307 | TCT | 0.19 | AGC | 0.24 | 0.05 |
| 308 | AAT | 0.47 | AAC | 0.53 | 0.06 |
| 309 | TTA | 0.08 | CTG | 0.4 | 0.32 |
| 310 | CAC | 0.58 | CAC | 0.58 | 0 |
| 311 | GAA | 0.42 | GAG | 0.58 | 0.16 |
| 312 | ATT | 0.36 | ATC | 0.47 | 0.11 |
| 313 | GCT | 0.27 | GCC | 0.4 | 0.13 |
| 314 | TCT | 0.19 | AGC | 0.24 | 0.05 |
| 315 | GGT | 0.16 | GGC | 0.34 | 0.18 |
| 316 | GGC | 0.34 | GGC | 0.34 | 0 |
| 317 | GCT | 0.27 | GCC | 0.4 | 0.13 |
| 318 | CCC | 0.32 | CCC | 0.32 | 0 |
| 319 | CTC | 0.2 | CTG | 0.4 | 0.2 |
| 320 | TCT | 0.19 | AGC | 0.24 | 0.05 |
| 321 | AAG | 0.57 | AAG | 0.57 | 0 |
| 322 | GAA | 0.42 | GAG | 0.58 | 0.16 |
| 323 | GTC | 0.24 | GTG | 0.46 | 0.22 |
| 324 | GGG | 0.25 | GGC | 0.34 | 0.09 |

| | | | | | |
|-----|-----|------|-----|------|------|
| 325 | GAA | 0.42 | GAG | 0.58 | 0.16 |
| 326 | GCG | 0.11 | GCC | 0.4 | 0.29 |
| 327 | GTT | 0.18 | GTG | 0.46 | 0.28 |
| 328 | GCC | 0.4 | GCC | 0.4 | 0 |
| 329 | AAG | 0.57 | AAG | 0.57 | 0 |
| 330 | AGG | 0.21 | AGA | 0.21 | 0 |
| 331 | TTC | 0.54 | TTC | 0.54 | 0 |
| 332 | CAT | 0.42 | CAC | 0.58 | 0.16 |
| 333 | CTG | 0.4 | CTG | 0.4 | 0 |
| 334 | CCA | 0.28 | CCC | 0.32 | 0.04 |
| 335 | GGT | 0.16 | GGC | 0.34 | 0.18 |
| 336 | ATC | 0.47 | ATC | 0.47 | 0 |
| 337 | AGG | 0.21 | AGA | 0.21 | 0 |
| 338 | CAA | 0.27 | CAG | 0.73 | 0.46 |
| 339 | GGA | 0.25 | GGC | 0.34 | 0.09 |
| 340 | TAT | 0.44 | TAC | 0.56 | 0.12 |
| 341 | GGG | 0.25 | GGC | 0.34 | 0.09 |
| 342 | CTC | 0.2 | CTG | 0.4 | 0.2 |
| 343 | ACT | 0.25 | ACC | 0.36 | 0.11 |
| 344 | GAG | 0.58 | GAG | 0.58 | 0 |
| 345 | ACT | 0.25 | ACC | 0.36 | 0.11 |
| 346 | ACA | 0.28 | ACC | 0.36 | 0.08 |
| 347 | TCA | 0.15 | AGC | 0.24 | 0.09 |
| 348 | GCT | 0.27 | GCC | 0.4 | 0.13 |
| 349 | ATT | 0.36 | ATC | 0.47 | 0.11 |
| 350 | CTG | 0.4 | CTG | 0.4 | 0 |
| 351 | ATT | 0.36 | ATC | 0.47 | 0.11 |
| 352 | ACA | 0.28 | ACC | 0.36 | 0.08 |
| 353 | CCC | 0.32 | CCC | 0.32 | 0 |
| 354 | GAG | 0.58 | GAG | 0.58 | 0 |
| 355 | GGG | 0.25 | GGC | 0.34 | 0.09 |
| 356 | GAT | 0.46 | GAC | 0.54 | 0.08 |
| 357 | GAT | 0.46 | GAC | 0.54 | 0.08 |
| 358 | AAA | 0.43 | AAG | 0.57 | 0.14 |
| 359 | CCG | 0.11 | CCC | 0.32 | 0.21 |
| 360 | GGC | 0.34 | GGC | 0.34 | 0 |
| 361 | GCG | 0.11 | GCC | 0.4 | 0.29 |
| 362 | GTC | 0.24 | GTG | 0.46 | 0.22 |
| 363 | GGT | 0.16 | GGC | 0.34 | 0.18 |
| 364 | AAA | 0.43 | AAG | 0.57 | 0.14 |
| 365 | GTT | 0.18 | GTG | 0.46 | 0.28 |
| 366 | GTT | 0.18 | GTG | 0.46 | 0.28 |
| 367 | CCA | 0.28 | CCC | 0.32 | 0.04 |
| 368 | TTT | 0.46 | TTC | 0.54 | 0.08 |
| 369 | TTT | 0.46 | TTC | 0.54 | 0.08 |
| 370 | GAA | 0.42 | GAG | 0.58 | 0.16 |
| 371 | GCG | 0.11 | GCC | 0.4 | 0.29 |

| | | | | | |
|-----|-----|------|-----|------|------|
| 372 | AAG | 0.57 | AAG | 0.57 | 0 |
| 373 | GTT | 0.18 | GTG | 0.46 | 0.28 |
| 374 | GTG | 0.46 | GTG | 0.46 | 0 |
| 375 | GAT | 0.46 | GAC | 0.54 | 0.08 |
| 376 | CTG | 0.4 | CTG | 0.4 | 0 |
| 377 | GAT | 0.46 | GAC | 0.54 | 0.08 |
| 378 | ACC | 0.36 | ACC | 0.36 | 0 |
| 379 | GGG | 0.25 | GGC | 0.34 | 0.09 |
| 380 | AAA | 0.43 | AAG | 0.57 | 0.14 |
| 381 | ACG | 0.11 | ACC | 0.36 | 0.25 |
| 382 | CTG | 0.4 | CTG | 0.4 | 0 |
| 383 | GGC | 0.34 | GGC | 0.34 | 0 |
| 384 | GTT | 0.18 | GTG | 0.46 | 0.28 |
| 385 | AAT | 0.47 | AAC | 0.53 | 0.06 |
| 386 | CAA | 0.27 | CAG | 0.73 | 0.46 |
| 387 | AGA | 0.21 | AGA | 0.21 | 0 |
| 388 | GGC | 0.34 | GGC | 0.34 | 0 |
| 389 | GAA | 0.42 | GAG | 0.58 | 0.16 |
| 390 | CTG | 0.4 | CTG | 0.4 | 0 |
| 391 | TGT | 0.46 | TGC | 0.54 | 0.08 |
| 392 | GTG | 0.46 | GTG | 0.46 | 0 |
| 393 | AGA | 0.21 | AGA | 0.21 | 0 |
| 394 | GGT | 0.16 | GGC | 0.34 | 0.18 |
| 395 | CCT | 0.29 | CCC | 0.32 | 0.03 |
| 396 | ATG | 1 | ATG | 1 | 0 |
| 397 | ATT | 0.36 | ATC | 0.47 | 0.11 |
| 398 | ATG | 1 | ATG | 1 | 0 |
| 399 | TCC | 0.22 | AGC | 0.24 | 0.02 |
| 400 | GGT | 0.16 | GGC | 0.34 | 0.18 |
| 401 | TAT | 0.44 | TAC | 0.56 | 0.12 |
| 402 | GTA | 0.12 | GTG | 0.46 | 0.34 |
| 403 | AAC | 0.53 | AAC | 0.53 | 0 |
| 404 | AAT | 0.47 | AAC | 0.53 | 0.06 |
| 405 | CCG | 0.11 | CCC | 0.32 | 0.21 |
| 406 | GAA | 0.42 | GAG | 0.58 | 0.16 |
| 407 | GCG | 0.11 | GCC | 0.4 | 0.29 |
| 408 | ACC | 0.36 | ACC | 0.36 | 0 |
| 409 | AAC | 0.53 | AAC | 0.53 | 0 |
| 410 | GCC | 0.4 | GCC | 0.4 | 0 |
| 411 | TTG | 0.13 | CTG | 0.4 | 0.27 |
| 412 | ATT | 0.36 | ATC | 0.47 | 0.11 |
| 413 | GAC | 0.54 | GAC | 0.54 | 0 |
| 414 | AAG | 0.57 | AAG | 0.57 | 0 |
| 415 | GAT | 0.46 | GAC | 0.54 | 0.08 |
| 416 | GGA | 0.25 | GGC | 0.34 | 0.09 |
| 417 | TGG | 1 | TGG | 1 | 0 |
| 418 | CTA | 0.07 | CTG | 0.4 | 0.33 |

| | | | | | |
|-----|-----|------|-----|------|------|
| 419 | CAT | 0.42 | CAC | 0.58 | 0.16 |
| 420 | TCT | 0.19 | AGC | 0.24 | 0.05 |
| 421 | GGA | 0.25 | GGC | 0.34 | 0.09 |
| 422 | GAC | 0.54 | GAC | 0.54 | 0 |
| 423 | ATA | 0.17 | ATC | 0.47 | 0.3 |
| 424 | GCT | 0.27 | GCC | 0.4 | 0.13 |
| 425 | TAC | 0.56 | TAC | 0.56 | 0 |
| 426 | TGG | 1 | TGG | 1 | 0 |
| 427 | GAC | 0.54 | GAC | 0.54 | 0 |
| 428 | GAA | 0.42 | GAG | 0.58 | 0.16 |
| 429 | GAC | 0.54 | GAC | 0.54 | 0 |
| 430 | GAA | 0.42 | GAG | 0.58 | 0.16 |
| 431 | CAC | 0.58 | CAC | 0.58 | 0 |
| 432 | TTC | 0.54 | TTC | 0.54 | 0 |
| 433 | TTC | 0.54 | TTC | 0.54 | 0 |
| 434 | ATC | 0.47 | ATC | 0.47 | 0 |
| 435 | GTT | 0.18 | GTG | 0.46 | 0.28 |
| 436 | GAC | 0.54 | GAC | 0.54 | 0 |
| 437 | CGC | 0.18 | AGA | 0.21 | 0.03 |
| 438 | CTG | 0.4 | CTG | 0.4 | 0 |
| 439 | AAG | 0.57 | AAG | 0.57 | 0 |
| 440 | TCT | 0.19 | AGC | 0.24 | 0.05 |
| 441 | CTG | 0.4 | CTG | 0.4 | 0 |
| 442 | ATT | 0.36 | ATC | 0.47 | 0.11 |
| 443 | AAG | 0.57 | AAG | 0.57 | 0 |
| 444 | TAC | 0.56 | TAC | 0.56 | 0 |
| 445 | AAA | 0.43 | AAG | 0.57 | 0.14 |
| 446 | GGC | 0.34 | GGC | 0.34 | 0 |
| 447 | TAT | 0.44 | TAC | 0.56 | 0.12 |
| 448 | CAG | 0.73 | CAG | 0.73 | 0 |
| 449 | GTG | 0.46 | GTG | 0.46 | 0 |
| 450 | GCT | 0.27 | GCC | 0.4 | 0.13 |
| 451 | CCC | 0.32 | CCC | 0.32 | 0 |
| 452 | GCT | 0.27 | GCC | 0.4 | 0.13 |
| 453 | GAA | 0.42 | GAG | 0.58 | 0.16 |
| 454 | TTG | 0.13 | CTG | 0.4 | 0.27 |
| 455 | GAA | 0.42 | GAG | 0.58 | 0.16 |
| 456 | TCC | 0.22 | AGC | 0.24 | 0.02 |
| 457 | ATC | 0.47 | ATC | 0.47 | 0 |
| 458 | TTG | 0.13 | CTG | 0.4 | 0.27 |
| 459 | CTC | 0.2 | CTG | 0.4 | 0.2 |
| 460 | CAA | 0.27 | CAG | 0.73 | 0.46 |
| 461 | CAC | 0.58 | CAC | 0.58 | 0 |
| 462 | CCC | 0.32 | CCC | 0.32 | 0 |
| 463 | AAC | 0.53 | AAC | 0.53 | 0 |
| 464 | ATC | 0.47 | ATC | 0.47 | 0 |
| 465 | TTC | 0.54 | TTC | 0.54 | 0 |

| | | | | | |
|-----|-----|------|-----|------|------|
| 466 | GAC | 0.54 | GAC | 0.54 | 0 |
| 467 | GCA | 0.23 | GCC | 0.4 | 0.17 |
| 468 | GGT | 0.16 | GGC | 0.34 | 0.18 |
| 469 | GTC | 0.24 | GTG | 0.46 | 0.22 |
| 470 | GCA | 0.23 | GCC | 0.4 | 0.17 |
| 471 | GGT | 0.16 | GGC | 0.34 | 0.18 |
| 472 | CTT | 0.13 | CTG | 0.4 | 0.27 |
| 473 | CCC | 0.32 | CCC | 0.32 | 0 |
| 474 | GAC | 0.54 | GAC | 0.54 | 0 |
| 475 | GAT | 0.46 | GAC | 0.54 | 0.08 |
| 476 | GAC | 0.54 | GAC | 0.54 | 0 |
| 477 | GCC | 0.4 | GCC | 0.4 | 0 |
| 478 | GGT | 0.16 | GGC | 0.34 | 0.18 |
| 479 | GAA | 0.42 | GAG | 0.58 | 0.16 |
| 480 | CTT | 0.13 | CTG | 0.4 | 0.27 |
| 481 | CCC | 0.32 | CCC | 0.32 | 0 |
| 482 | GCC | 0.4 | GCC | 0.4 | 0 |
| 483 | GCC | 0.4 | GCC | 0.4 | 0 |
| 484 | GTT | 0.18 | GTG | 0.46 | 0.28 |
| 485 | GTT | 0.18 | GTG | 0.46 | 0.28 |
| 486 | GTT | 0.18 | GTG | 0.46 | 0.28 |
| 487 | TTG | 0.13 | CTG | 0.4 | 0.27 |
| 488 | GAG | 0.58 | GAG | 0.58 | 0 |
| 489 | CAC | 0.58 | CAC | 0.58 | 0 |
| 490 | GGA | 0.25 | GGC | 0.34 | 0.09 |
| 491 | AAG | 0.57 | AAG | 0.57 | 0 |
| 492 | ACG | 0.11 | ACC | 0.36 | 0.25 |
| 493 | ATG | 1 | ATG | 1 | 0 |
| 494 | ACG | 0.11 | ACC | 0.36 | 0.25 |
| 495 | GAA | 0.42 | GAG | 0.58 | 0.16 |
| 496 | AAA | 0.43 | AAG | 0.57 | 0.14 |
| 497 | GAG | 0.58 | GAG | 0.58 | 0 |
| 498 | ATC | 0.47 | ATC | 0.47 | 0 |
| 499 | GTG | 0.46 | GTG | 0.46 | 0 |
| 500 | GAT | 0.46 | GAC | 0.54 | 0.08 |
| 501 | TAC | 0.56 | TAC | 0.56 | 0 |
| 502 | GTC | 0.24 | GTG | 0.46 | 0.22 |
| 503 | GCC | 0.4 | GCC | 0.4 | 0 |
| 504 | AGT | 0.15 | AGC | 0.24 | 0.09 |
| 505 | CAA | 0.27 | CAG | 0.73 | 0.46 |
| 506 | GTA | 0.12 | GTG | 0.46 | 0.34 |
| 507 | ACA | 0.28 | ACC | 0.36 | 0.08 |
| 508 | ACC | 0.36 | ACC | 0.36 | 0 |
| 509 | GCG | 0.11 | GCC | 0.4 | 0.29 |
| 510 | AAA | 0.43 | AAG | 0.57 | 0.14 |
| 511 | AAG | 0.57 | AAG | 0.57 | 0 |
| 512 | TTG | 0.13 | CTG | 0.4 | 0.27 |

| | | | | | |
|-----|-----|------|-----|------|------|
| 513 | CGC | 0.18 | AGA | 0.21 | 0.03 |
| 514 | GGA | 0.25 | GGC | 0.34 | 0.09 |
| 515 | GGA | 0.25 | GGC | 0.34 | 0.09 |
| 516 | GTT | 0.18 | GTG | 0.46 | 0.28 |
| 517 | GTG | 0.46 | GTG | 0.46 | 0 |
| 518 | TTT | 0.46 | TTC | 0.54 | 0.08 |
| 519 | GTG | 0.46 | GTG | 0.46 | 0 |
| 520 | GAC | 0.54 | GAC | 0.54 | 0 |
| 521 | GAA | 0.42 | GAG | 0.58 | 0.16 |
| 522 | GTA | 0.12 | GTG | 0.46 | 0.34 |
| 523 | CCG | 0.11 | CCC | 0.32 | 0.21 |
| 524 | AAA | 0.43 | AAG | 0.57 | 0.14 |
| 525 | GGT | 0.16 | GGC | 0.34 | 0.18 |
| 526 | CTT | 0.13 | CTG | 0.4 | 0.27 |
| 527 | ACC | 0.36 | ACC | 0.36 | 0 |
| 528 | GGA | 0.25 | GGC | 0.34 | 0.09 |
| 529 | AAA | 0.43 | AAG | 0.57 | 0.14 |
| 530 | CTC | 0.2 | CTG | 0.4 | 0.2 |
| 531 | GAC | 0.54 | GAC | 0.54 | 0 |
| 532 | GCA | 0.23 | GCC | 0.4 | 0.17 |
| 533 | AGA | 0.21 | AGA | 0.21 | 0 |
| 534 | AAA | 0.43 | AAG | 0.57 | 0.14 |
| 535 | ATC | 0.47 | ATC | 0.47 | 0 |
| 536 | AGA | 0.21 | AGA | 0.21 | 0 |
| 537 | GAG | 0.58 | GAG | 0.58 | 0 |
| 538 | ATC | 0.47 | ATC | 0.47 | 0 |
| 539 | CTC | 0.2 | CTG | 0.4 | 0.2 |
| 540 | ATA | 0.17 | ATC | 0.47 | 0.3 |
| 541 | AAG | 0.57 | AAG | 0.57 | 0 |
| 542 | GCC | 0.4 | GCC | 0.4 | 0 |
| 543 | AAG | 0.57 | AAG | 0.57 | 0 |
| 544 | AAG | 0.57 | AAG | 0.57 | 0 |
| 545 | GGC | 0.34 | GGC | 0.34 | 0 |
| 546 | GGA | 0.25 | GGC | 0.34 | 0.09 |
| 547 | AAG | 0.57 | AAG | 0.57 | 0 |
| 548 | ATC | 0.47 | ATC | 0.47 | 0 |
| 549 | GCC | 0.4 | GCC | 0.4 | 0 |
| 550 | GTG | 0.46 | GTG | 0.46 | 0 |
| 551 | TAA | | | | |

Score 0.10

Supplementary Table 3: Gaussia Luciferase Codon Usage Score. Calculated by Christopher Roberts.

| Position | Input Codon | Input codon usage score | Optimal codon | Optimal codon usage score | Bias |
|----------|-------------|-------------------------|---------------|---------------------------|------|
| 1 | ATG | 1 | ATG | 1 | 0 |
| 2 | GGA | 0.25 | GGC | 0.34 | 0.09 |
| 3 | GTC | 0.24 | GTG | 0.46 | 0.22 |
| 4 | AAA | 0.43 | AAG | 0.57 | 0.14 |
| 5 | GTT | 0.18 | GTG | 0.46 | 0.28 |
| 6 | CTG | 0.4 | CTG | 0.4 | 0 |
| 7 | TTT | 0.46 | TTC | 0.54 | 0.08 |
| 8 | GCC | 0.4 | GCC | 0.4 | 0 |
| 9 | CTG | 0.4 | CTG | 0.4 | 0 |
| 10 | ATC | 0.47 | ATC | 0.47 | 0 |
| 11 | TGC | 0.54 | TGC | 0.54 | 0 |
| 12 | ATC | 0.47 | ATC | 0.47 | 0 |
| 13 | GCT | 0.27 | GCC | 0.4 | 0.13 |
| 14 | GTG | 0.46 | GTG | 0.46 | 0 |
| 15 | GCC | 0.4 | GCC | 0.4 | 0 |
| 16 | GAG | 0.58 | GAG | 0.58 | 0 |
| 17 | GCC | 0.4 | GCC | 0.4 | 0 |
| 18 | AAG | 0.57 | AAG | 0.57 | 0 |
| 19 | CCC | 0.32 | CCC | 0.32 | 0 |
| 20 | ACC | 0.36 | ACC | 0.36 | 0 |
| 21 | GAG | 0.58 | GAG | 0.58 | 0 |
| 22 | AAC | 0.53 | AAC | 0.53 | 0 |
| 23 | AAC | 0.53 | AAC | 0.53 | 0 |
| 24 | GAA | 0.42 | GAG | 0.58 | 0.16 |
| 25 | GAC | 0.54 | GAC | 0.54 | 0 |
| 26 | TTC | 0.54 | TTC | 0.54 | 0 |
| 27 | AAC | 0.53 | AAC | 0.53 | 0 |
| 28 | ATC | 0.47 | ATC | 0.47 | 0 |
| 29 | GTG | 0.46 | GTG | 0.46 | 0 |
| 30 | GCC | 0.4 | GCC | 0.4 | 0 |
| 31 | GTG | 0.46 | GTG | 0.46 | 0 |
| 32 | GCC | 0.4 | GCC | 0.4 | 0 |
| 33 | AGC | 0.24 | AGC | 0.24 | 0 |
| 34 | AAC | 0.53 | AAC | 0.53 | 0 |
| 35 | TTC | 0.54 | TTC | 0.54 | 0 |
| 36 | GCG | 0.11 | GCC | 0.4 | 0.29 |
| 37 | ACC | 0.36 | ACC | 0.36 | 0 |
| 38 | ACG | 0.11 | ACC | 0.36 | 0.25 |
| 39 | GAT | 0.46 | GAC | 0.54 | 0.08 |
| 40 | CTC | 0.2 | CTG | 0.4 | 0.2 |

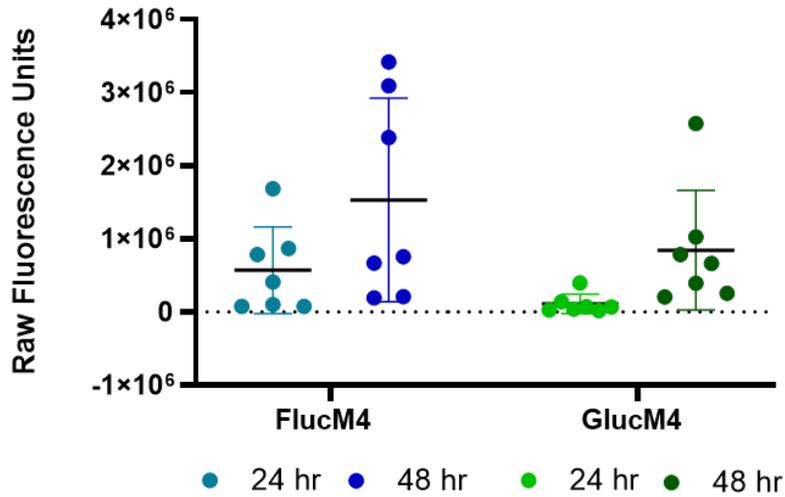
| | | | | | |
|----|-----|------|-----|------|------|
| 41 | GAT | 0.46 | GAC | 0.54 | 0.08 |
| 42 | GCT | 0.27 | GCC | 0.4 | 0.13 |
| 43 | GAC | 0.54 | GAC | 0.54 | 0 |
| 44 | CGC | 0.18 | AGA | 0.21 | 0.03 |
| 45 | GGG | 0.25 | GGC | 0.34 | 0.09 |
| 46 | AAG | 0.57 | AAG | 0.57 | 0 |
| 47 | TTG | 0.13 | CTG | 0.4 | 0.27 |
| 48 | CCC | 0.32 | CCC | 0.32 | 0 |
| 49 | GGC | 0.34 | GGC | 0.34 | 0 |
| 50 | AAG | 0.57 | AAG | 0.57 | 0 |
| 51 | AAG | 0.57 | AAG | 0.57 | 0 |
| 52 | CTG | 0.4 | CTG | 0.4 | 0 |
| 53 | CCG | 0.11 | CCC | 0.32 | 0.21 |
| 54 | CTG | 0.4 | CTG | 0.4 | 0 |
| 55 | GAG | 0.58 | GAG | 0.58 | 0 |
| 56 | GTG | 0.46 | GTG | 0.46 | 0 |
| 57 | CTC | 0.2 | CTG | 0.4 | 0.2 |
| 58 | AAA | 0.43 | AAG | 0.57 | 0.14 |
| 59 | GAG | 0.58 | GAG | 0.58 | 0 |
| 60 | ATG | 1 | ATG | 1 | 0 |
| 61 | GAA | 0.42 | GAG | 0.58 | 0.16 |
| 62 | GCC | 0.4 | GCC | 0.4 | 0 |
| 63 | AAT | 0.47 | AAC | 0.53 | 0.06 |
| 64 | GCC | 0.4 | GCC | 0.4 | 0 |
| 65 | CGG | 0.2 | AGA | 0.21 | 0.01 |
| 66 | AAA | 0.43 | AAG | 0.57 | 0.14 |
| 67 | GCT | 0.27 | GCC | 0.4 | 0.13 |
| 68 | GGC | 0.34 | GGC | 0.34 | 0 |
| 69 | TGC | 0.54 | TGC | 0.54 | 0 |
| 70 | ACC | 0.36 | ACC | 0.36 | 0 |
| 71 | AGG | 0.21 | AGA | 0.21 | 0 |
| 72 | GGC | 0.34 | GGC | 0.34 | 0 |
| 73 | TGT | 0.46 | TGC | 0.54 | 0.08 |
| 74 | CTG | 0.4 | CTG | 0.4 | 0 |
| 75 | ATC | 0.47 | ATC | 0.47 | 0 |
| 76 | TGC | 0.54 | TGC | 0.54 | 0 |
| 77 | CTG | 0.4 | CTG | 0.4 | 0 |
| 78 | TCC | 0.22 | AGC | 0.24 | 0.02 |
| 79 | CAC | 0.58 | CAC | 0.58 | 0 |
| 80 | ATC | 0.47 | ATC | 0.47 | 0 |
| 81 | AAG | 0.57 | AAG | 0.57 | 0 |
| 82 | TGC | 0.54 | TGC | 0.54 | 0 |
| 83 | ACG | 0.11 | ACC | 0.36 | 0.25 |
| 84 | CCC | 0.32 | CCC | 0.32 | 0 |
| 85 | AAG | 0.57 | AAG | 0.57 | 0 |
| 86 | ATG | 1 | ATG | 1 | 0 |
| 87 | AAG | 0.57 | AAG | 0.57 | 0 |

| | | | | | |
|-----|-----|------|-----|------|------|
| 88 | AAG | 0.57 | AAG | 0.57 | 0 |
| 89 | TTC | 0.54 | TTC | 0.54 | 0 |
| 90 | ATC | 0.47 | ATC | 0.47 | 0 |
| 91 | CCA | 0.28 | CCC | 0.32 | 0.04 |
| 92 | GGA | 0.25 | GGC | 0.34 | 0.09 |
| 93 | CGC | 0.18 | AGA | 0.21 | 0.03 |
| 94 | TGC | 0.54 | TGC | 0.54 | 0 |
| 95 | CAC | 0.58 | CAC | 0.58 | 0 |
| 96 | ACC | 0.36 | ACC | 0.36 | 0 |
| 97 | TAC | 0.56 | TAC | 0.56 | 0 |
| 98 | GAA | 0.42 | GAG | 0.58 | 0.16 |
| 99 | GGC | 0.34 | GGC | 0.34 | 0 |
| 100 | GAC | 0.54 | GAC | 0.54 | 0 |
| 101 | AAA | 0.43 | AAG | 0.57 | 0.14 |
| 102 | GAG | 0.58 | GAG | 0.58 | 0 |
| 103 | TCC | 0.22 | AGC | 0.24 | 0.02 |
| 104 | GCA | 0.23 | GCC | 0.4 | 0.17 |
| 105 | CAG | 0.73 | CAG | 0.73 | 0 |
| 106 | GGC | 0.34 | GGC | 0.34 | 0 |
| 107 | GGC | 0.34 | GGC | 0.34 | 0 |
| 108 | ATA | 0.17 | ATC | 0.47 | 0.3 |
| 109 | GGC | 0.34 | GGC | 0.34 | 0 |
| 110 | GAG | 0.58 | GAG | 0.58 | 0 |
| 111 | GCG | 0.11 | GCC | 0.4 | 0.29 |
| 112 | ATC | 0.47 | ATC | 0.47 | 0 |
| 113 | GTC | 0.24 | GTG | 0.46 | 0.22 |
| 114 | GAC | 0.54 | GAC | 0.54 | 0 |
| 115 | ATT | 0.36 | ATC | 0.47 | 0.11 |
| 116 | CCT | 0.29 | CCC | 0.32 | 0.03 |
| 117 | GAG | 0.58 | GAG | 0.58 | 0 |
| 118 | ATT | 0.36 | ATC | 0.47 | 0.11 |
| 119 | CCT | 0.29 | CCC | 0.32 | 0.03 |
| 120 | GGG | 0.25 | GGC | 0.34 | 0.09 |
| 121 | TTC | 0.54 | TTC | 0.54 | 0 |
| 122 | AAG | 0.57 | AAG | 0.57 | 0 |
| 123 | GAC | 0.54 | GAC | 0.54 | 0 |
| 124 | TTG | 0.13 | CTG | 0.4 | 0.27 |
| 125 | GAG | 0.58 | GAG | 0.58 | 0 |
| 126 | CCC | 0.32 | CCC | 0.32 | 0 |
| 127 | ATG | 1 | ATG | 1 | 0 |
| 128 | GAG | 0.58 | GAG | 0.58 | 0 |
| 129 | CAG | 0.73 | CAG | 0.73 | 0 |
| 130 | TTC | 0.54 | TTC | 0.54 | 0 |
| 131 | ATC | 0.47 | ATC | 0.47 | 0 |
| 132 | GCA | 0.23 | GCC | 0.4 | 0.17 |
| 133 | CAG | 0.73 | CAG | 0.73 | 0 |
| 134 | GTC | 0.24 | GTG | 0.46 | 0.22 |

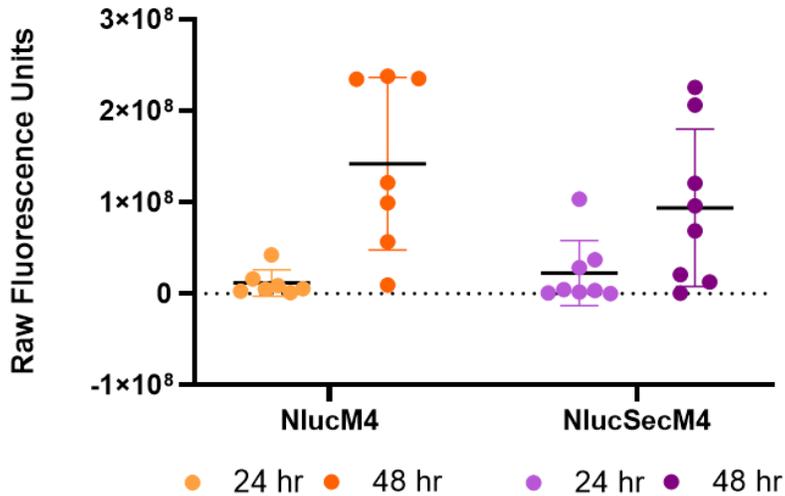
| | | | | | |
|-----|-----|------|-----|------|------|
| 135 | GAT | 0.46 | GAC | 0.54 | 0.08 |
| 136 | CTG | 0.4 | CTG | 0.4 | 0 |
| 137 | TGT | 0.46 | TGC | 0.54 | 0.08 |
| 138 | GTG | 0.46 | GTG | 0.46 | 0 |
| 139 | GAC | 0.54 | GAC | 0.54 | 0 |
| 140 | TGC | 0.54 | TGC | 0.54 | 0 |
| 141 | ACA | 0.28 | ACC | 0.36 | 0.08 |
| 142 | ACT | 0.25 | ACC | 0.36 | 0.11 |
| 143 | GGC | 0.34 | GGC | 0.34 | 0 |
| 144 | TGC | 0.54 | TGC | 0.54 | 0 |
| 145 | CTC | 0.2 | CTG | 0.4 | 0.2 |
| 146 | AAA | 0.43 | AAG | 0.57 | 0.14 |
| 147 | GGG | 0.25 | GGC | 0.34 | 0.09 |
| 148 | CTT | 0.13 | CTG | 0.4 | 0.27 |
| 149 | GCC | 0.4 | GCC | 0.4 | 0 |
| 150 | AAC | 0.53 | AAC | 0.53 | 0 |
| 151 | GTG | 0.46 | GTG | 0.46 | 0 |
| 152 | CAG | 0.73 | CAG | 0.73 | 0 |
| 153 | TGT | 0.46 | TGC | 0.54 | 0.08 |
| 154 | TCT | 0.19 | AGC | 0.24 | 0.05 |
| 155 | GAC | 0.54 | GAC | 0.54 | 0 |
| 156 | CTG | 0.4 | CTG | 0.4 | 0 |
| 157 | CTC | 0.2 | CTG | 0.4 | 0.2 |
| 158 | AAG | 0.57 | AAG | 0.57 | 0 |
| 159 | AAG | 0.57 | AAG | 0.57 | 0 |
| 160 | TGG | 1 | TGG | 1 | 0 |
| 161 | CTG | 0.4 | CTG | 0.4 | 0 |
| 162 | CCG | 0.11 | CCC | 0.32 | 0.21 |
| 163 | CAA | 0.27 | CAG | 0.73 | 0.46 |
| 164 | CGC | 0.18 | AGA | 0.21 | 0.03 |
| 165 | TGT | 0.46 | TGC | 0.54 | 0.08 |
| 166 | GCG | 0.11 | GCC | 0.4 | 0.29 |
| 167 | ACC | 0.36 | ACC | 0.36 | 0 |
| 168 | TTT | 0.46 | TTC | 0.54 | 0.08 |
| 169 | GCC | 0.4 | GCC | 0.4 | 0 |
| 170 | AGC | 0.24 | AGC | 0.24 | 0 |
| 171 | AAG | 0.57 | AAG | 0.57 | 0 |
| 172 | ATC | 0.47 | ATC | 0.47 | 0 |
| 173 | CAG | 0.73 | CAG | 0.73 | 0 |
| 174 | GGC | 0.34 | GGC | 0.34 | 0 |
| 175 | CAG | 0.73 | CAG | 0.73 | 0 |
| 176 | GTG | 0.46 | GTG | 0.46 | 0 |
| 177 | GAC | 0.54 | GAC | 0.54 | 0 |
| 178 | AAG | 0.57 | AAG | 0.57 | 0 |
| 179 | ATC | 0.47 | ATC | 0.47 | 0 |
| 180 | AAG | 0.57 | AAG | 0.57 | 0 |
| 181 | GGG | 0.25 | GGC | 0.34 | 0.09 |

| | | | | | |
|-----|-----|------|-----|--------------|------|
| 182 | GCC | 0.4 | GCC | 0.4 | 0 |
| 183 | GGT | 0.16 | GGC | 0.34 | 0.18 |
| 184 | GGT | 0.16 | GGC | 0.34 | 0.18 |
| 185 | GAC | 0.54 | GAC | 0.54 | 0 |
| 186 | TAA | 0.3 | TGA | 0.47 | |
| | | | | Score | 0.05 |

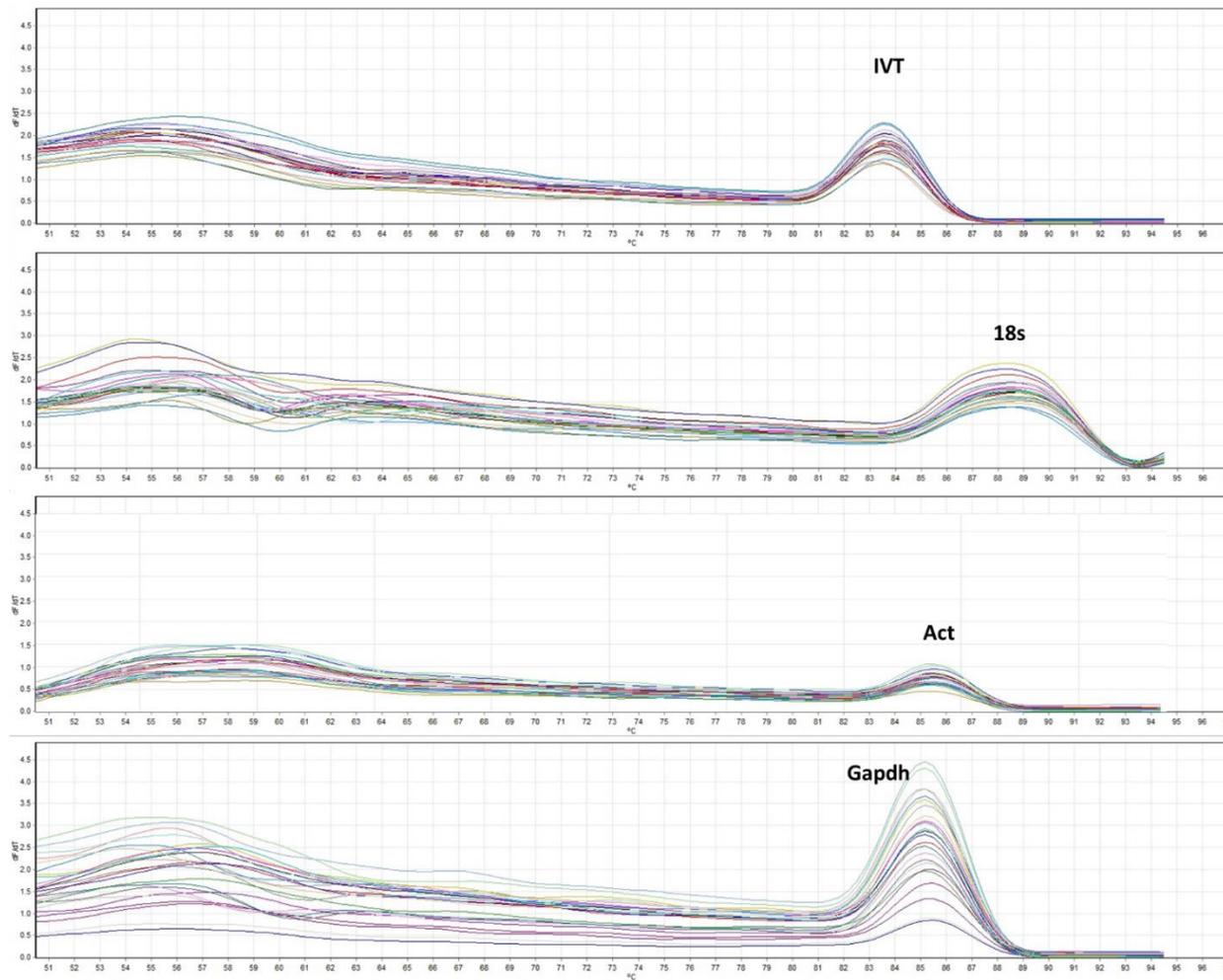
A.) Raw Fluorescence for FlucM4 and GlucM4 for pre-miR-122wt transfection



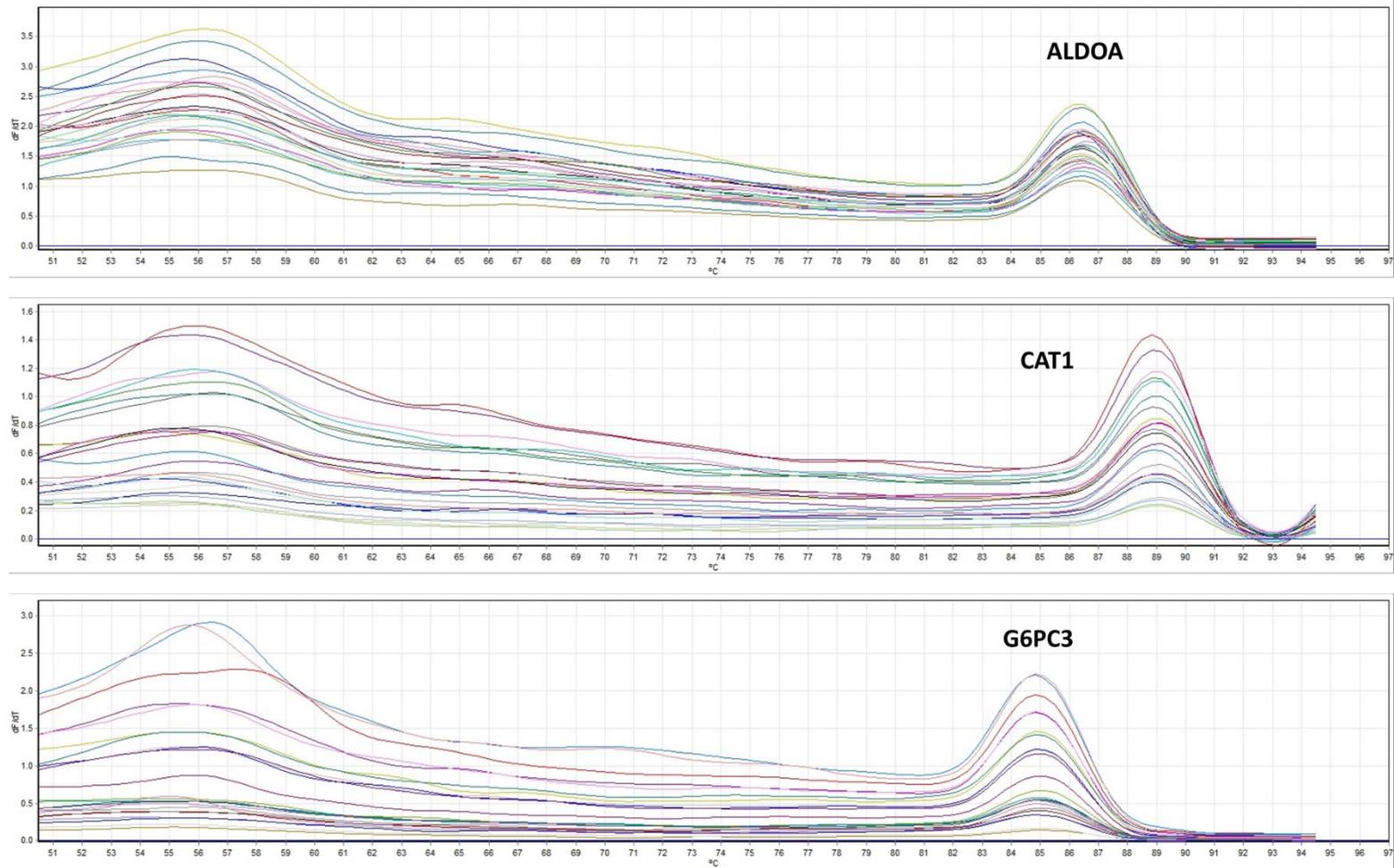
B.) Raw Fluorescence for NlucM4 and NlucSecM4 for pre-miR-122wt transfection



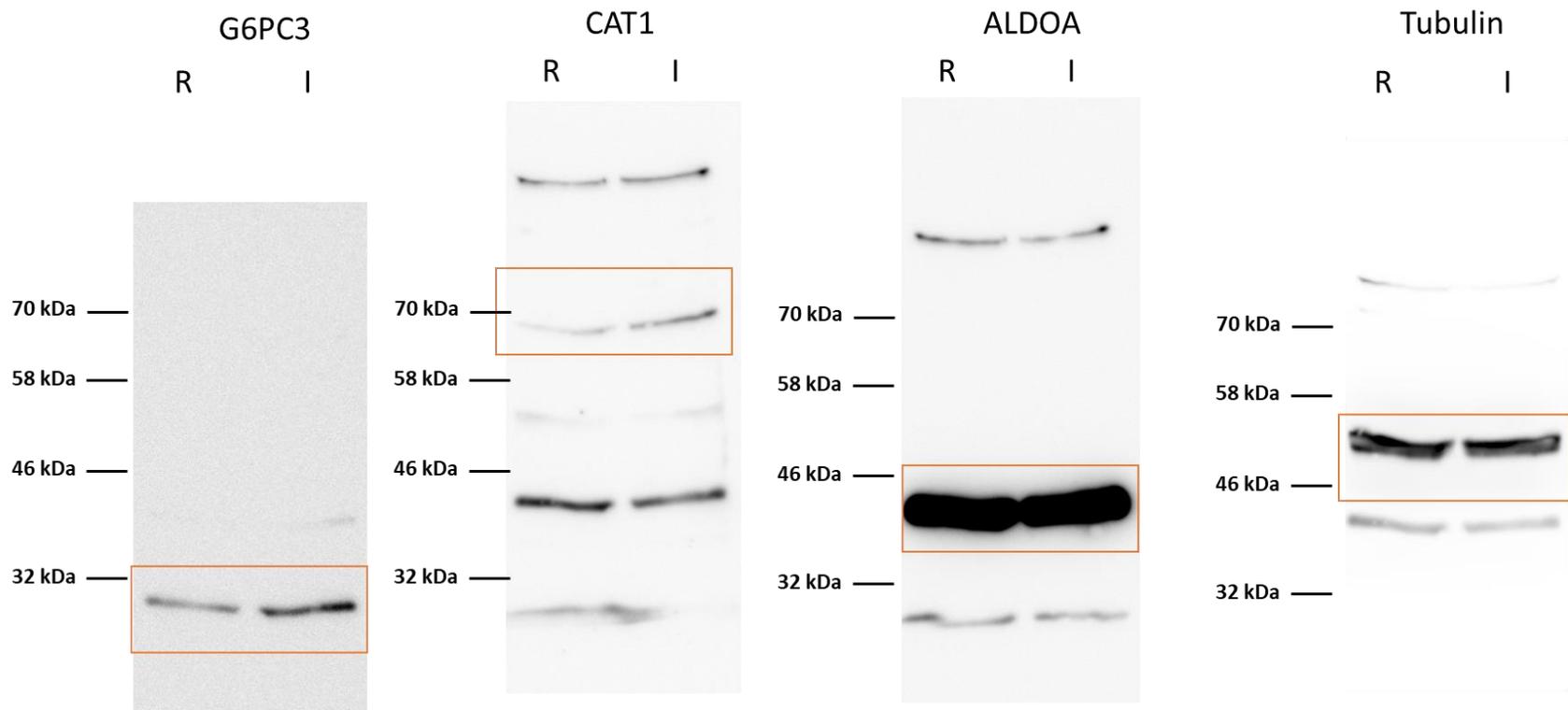
Supplementary Figure 7: Raw Fluorescence of Mutant Luciferase Reporters in Control Conditions. Reporter plasmids with mutant miR-122 binding sites were co-transfected with pre-miR-122wt, a synthetic pre-miR-122 wildtype oligonucleotide. The raw fluorescence units for **(A)** FlucM4 and GlucM4 and **(B)** NlucM4 and NlucSecM4 at 24 and 48 hours under these control conditions are shown. Data represents at least six independent experiments and error bars represent standard deviation.



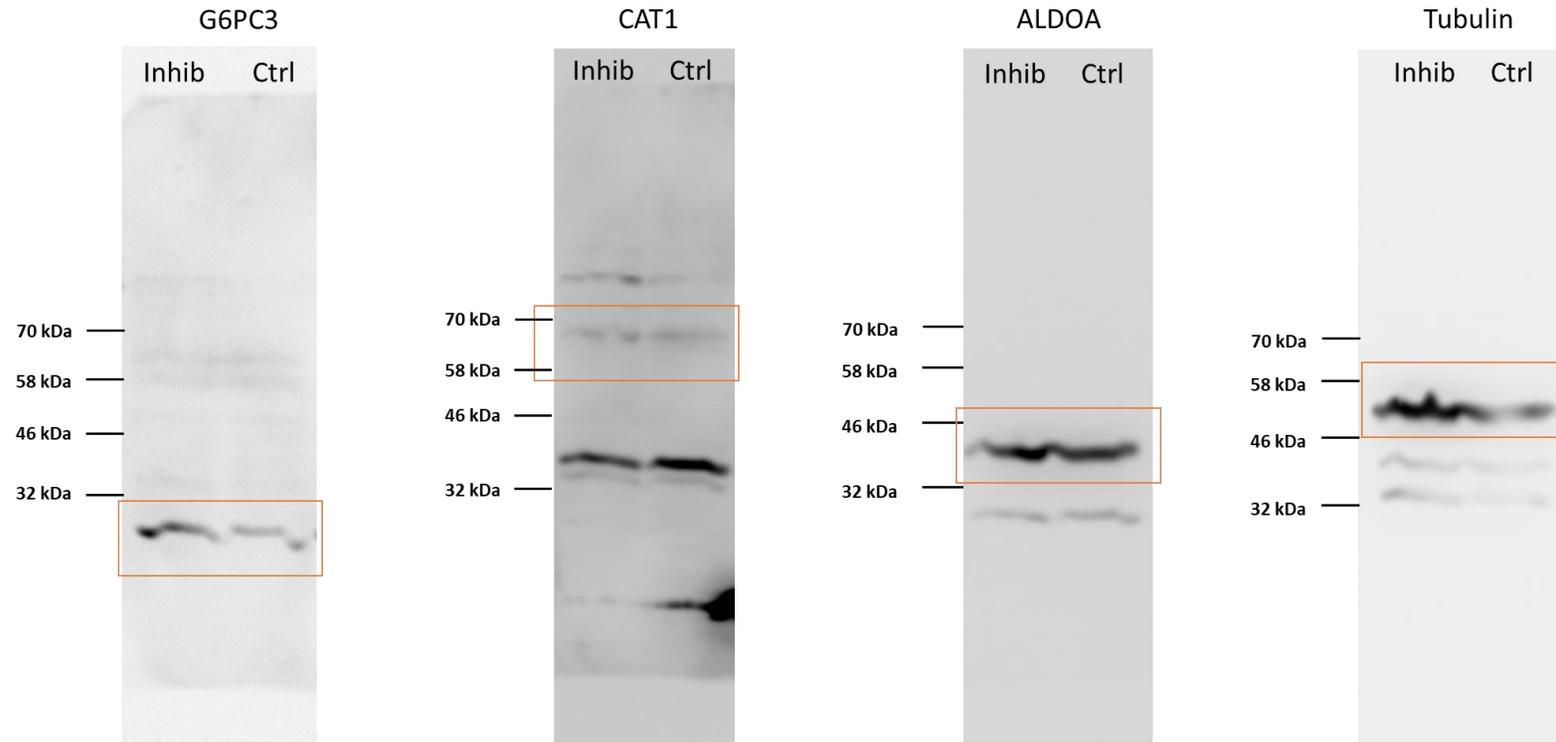
Supplementary Figure 8: Melt curves for qPCR primers targeting housekeeping genes 18s ribosomal rRNA (18s), Actin (Act), and glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and the spike-in control of *In Vitro* transcribed RNA (IVT).



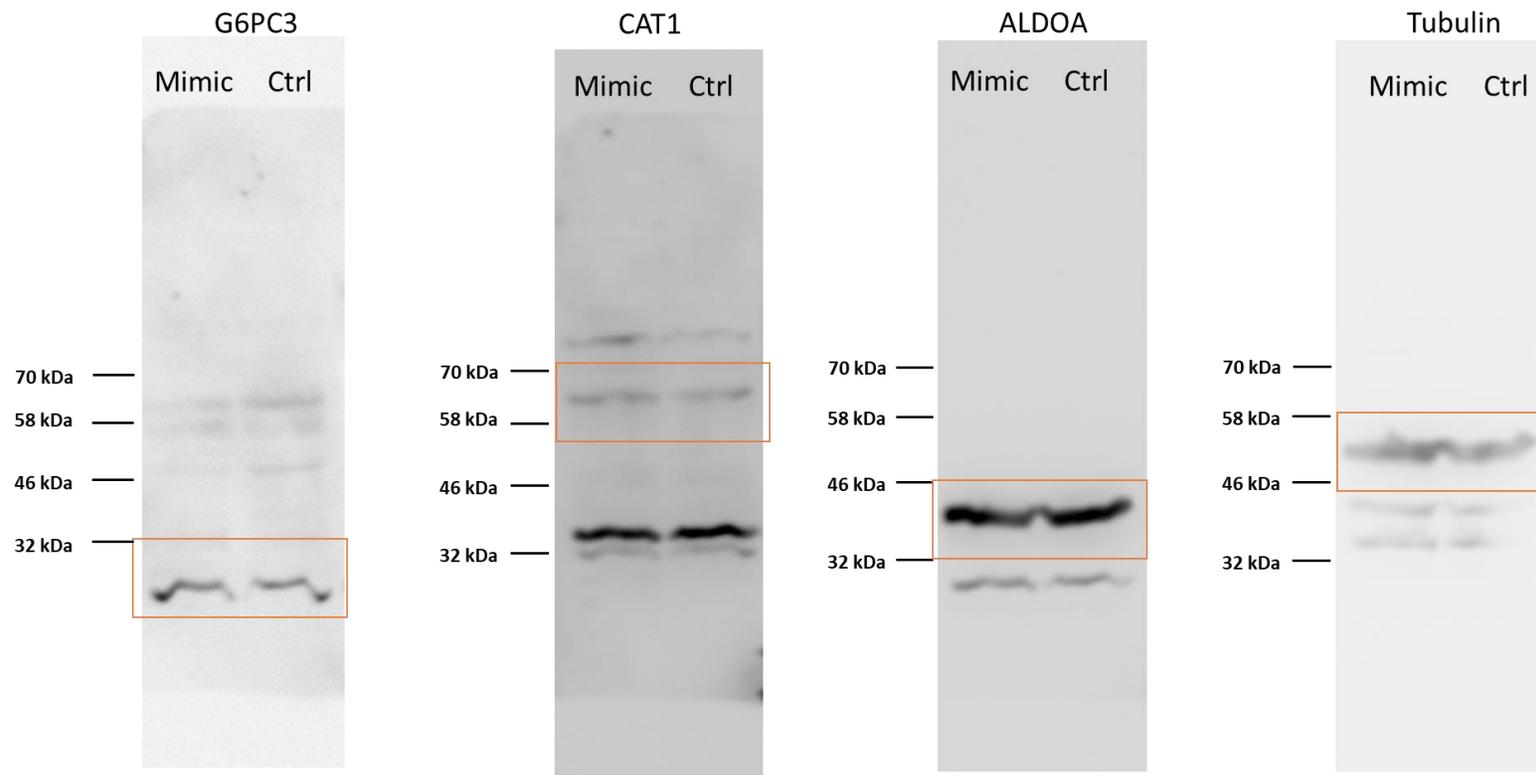
Supplementary Figure 9: Melt curves for qPCR primers targeting the mRNA known to be regulated by miR-122: Aldolase A (ALDOA), Cationic Amino Acid Transporter 1 (CAT1), and Glucose-6-Phosphatase Catalytic Subunit 3 (G6PC3).



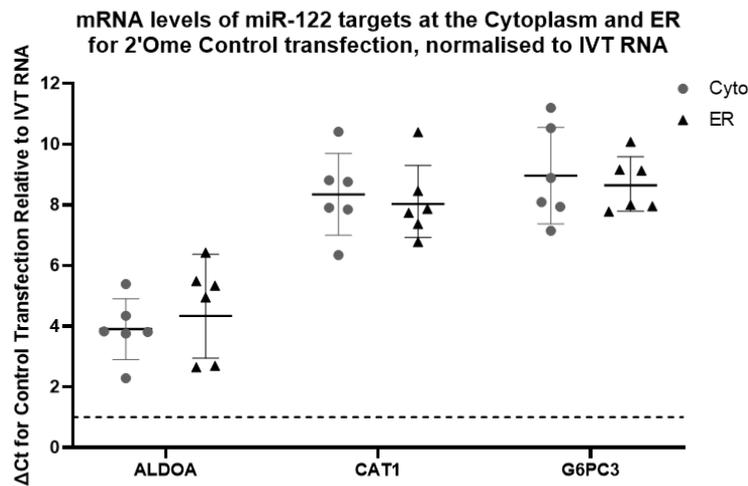
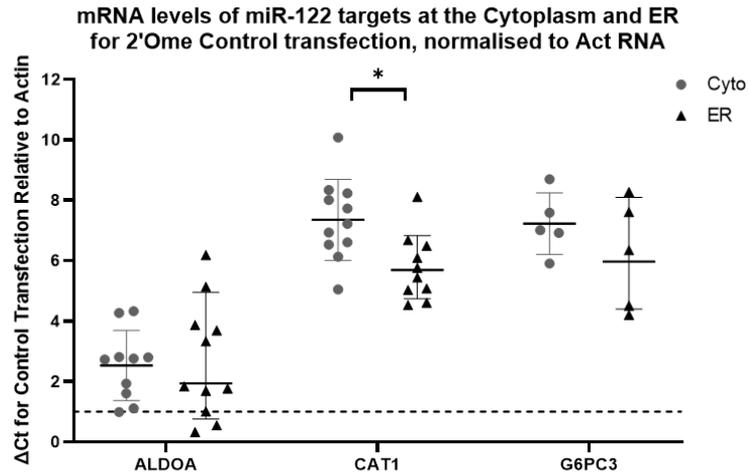
Supplementary Figure 10: Protein levels of miR-122 mRNA targets following inhibition of miR-122 with a 2'Ome Antisense Oligonucleotide. Total protein was extracted at 48 hours following transfection with a 2'Ome Antisense Oligonucleotide miR-122 inhibitor (I) or random control (R) and analysed by western blot with antibodies against ALDOA, CAT1 and G6PC3, with tubulin as a protein loading control with bands corresponding to the proteins of interest enclosed in an orange box.



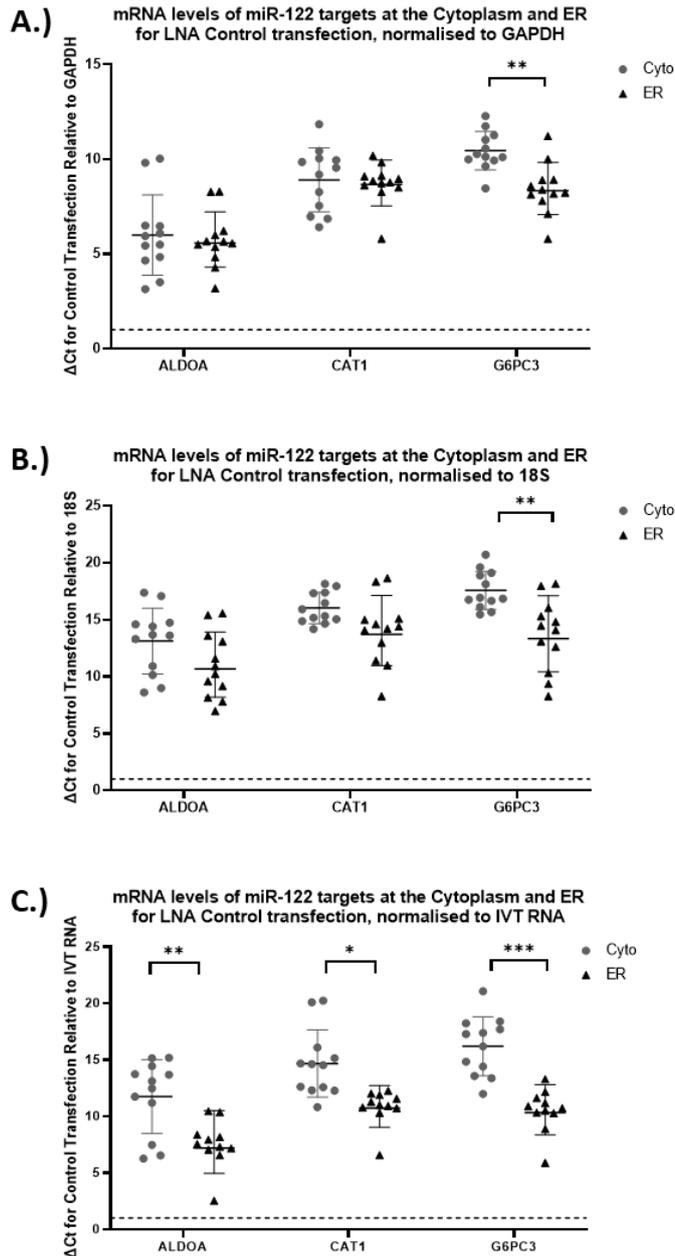
Supplementary Figure 11: Protein levels of miR-122 mRNA targets following inhibition of miR-122 with an LNA inhibitor. Total protein was extracted at 48 hours following transfection with an LNA miR-122 inhibitor (Inhib) or Negative control (Ctrl) and analysed by western blot with antibodies against ALDOA, CAT1 and G6PC3, with tubulin as a protein loading control with bands corresponding to the proteins of interest enclosed in an orange box.



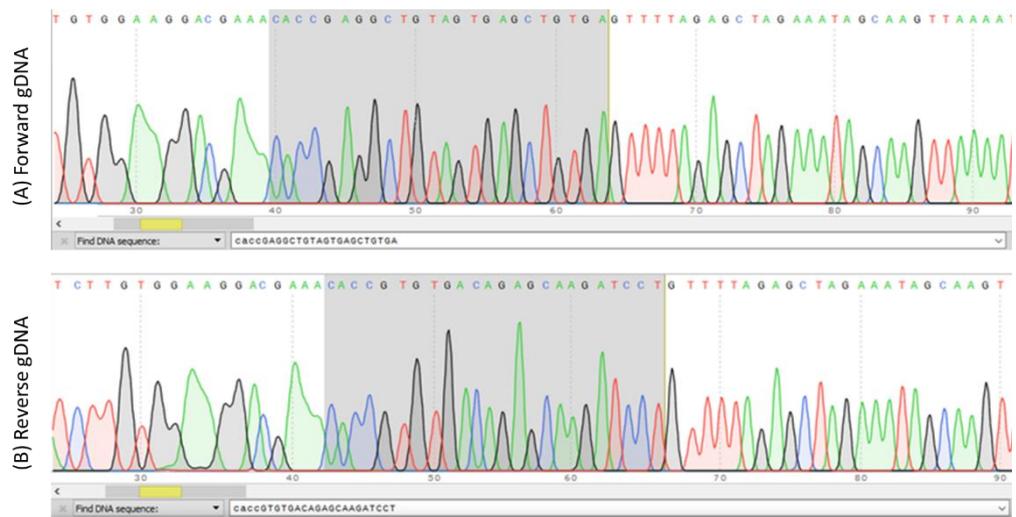
Supplementary Figure 12. Protein levels of miR-122 mRNA targets following overexpression of miR-122 with an LNA Mimic. Total protein was extracted at 48 hours following transfection with a miR-122 LNA mimic (Mimic) or negative control (Ctrl) and analysed by western blot with antibodies against ALDOA, CAT1 and G6PC3, with tubulin as a protein loading control with bands corresponding to the proteins of interest enclosed in an orange box.



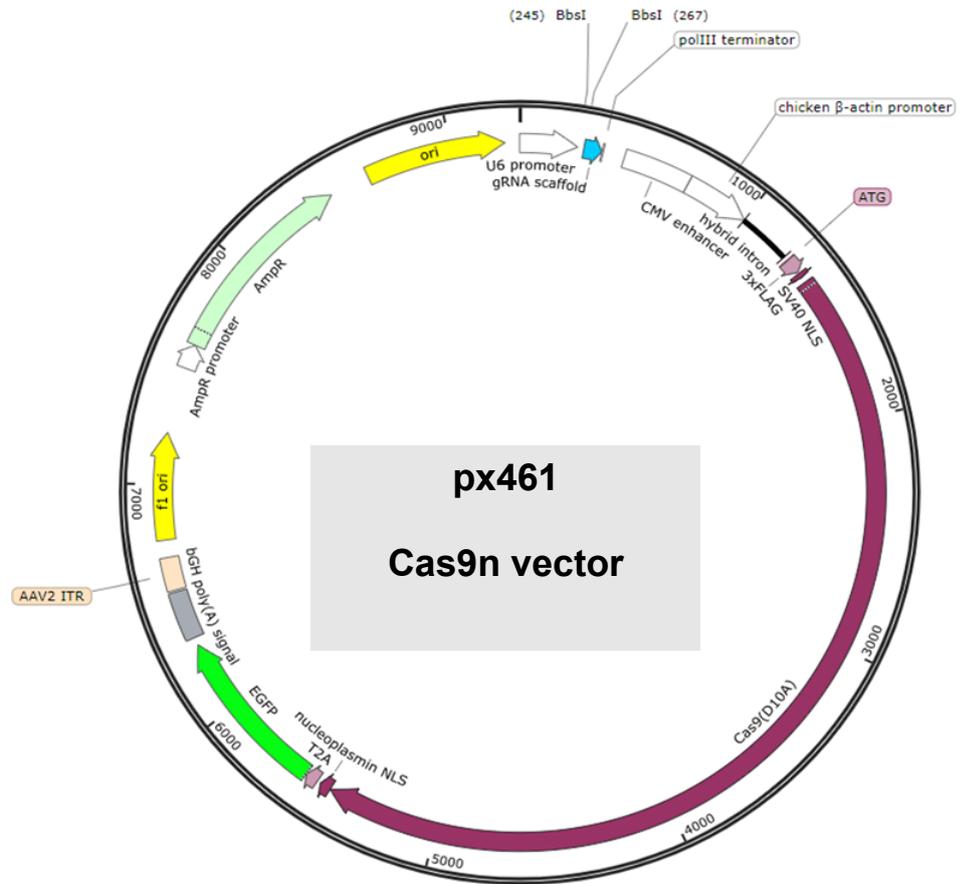
Supplementary Figure 13: Expression of miR-122 mRNA targets following transfection of a control 2'Ome oligonucleotide. Huh7 cells subjected to membrane fractionation to isolate cytoplasmic (Cyto, gray) and endoplasmic reticulum (ER, black)-rich fractions following transfection with a 2'Ome Oligonucleotide random control. qRT-PCR for these samples was used to detect the mRNA levels of miR-122 mRNA targets ALDOA, CAT1, and G6PC3, presented ΔCt (their raw Ct normalised to either **(A)** Actin, or **(B)** an IVT RNA spike-in). Student's T-test of mRNA levels normalised to Actin **(A)** determines a significant difference between cytoplasmic and ER rich fractions for G6PC3 (* $p < 0.05$). Graph represents data from at least five independent experiments plotted on a linear scale, with error bars representing the SD.



Supplementary Figure 14: Expression of miR-122 mRNA targets following transfection of a control LNA oligonucleotide. Huh7 cells subjected to membrane fractionation to isolate cytoplasmic (Cyto, gray) and endoplasmic reticulum (ER, black)-rich fractions following transfection with a LNA random control. qRT-PCR for these samples was used to detect the mRNA levels of miR-122 mRNA targets ALDOA, CAT1, and G6PC3, presented Δ Ct (their raw Ct normalised to either **(A)** GAPDH, **(B)** 18S rRNA, or **(C)** an IVT RNA spike-in). Students T-test of mRNA levels normalised to GAPDH **(A)** and 18S **(B)** determines a significant difference between cytoplasmic and ER rich fractions for G6PC3 (** $p < 0.01$). Students T-test of mRNA levels normalised to IVT RNA determines a significant difference between fractions for all three miR-122 targets (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.005$). Graph represents data from at least 11 independent experiments plotted on a linear scale, with error bars representing the SD.



Supplementary Figure 15: Sequence of both px461-sgDNA plasmids showing insertion of guide sequence. Sequencing using U6_F primer for (A) forward and (B) reverse guide with correct insert guide sequence highlighted in gray



Supplementary Figure 16: Map of the px461 Cas9n vector. Image taken from Snapgene Viewer [453]. The gRNA scaffolds (blue) are inserted between the Bbs1 double restriction sites upstream of the Cas9n-encoding gene (purple) and immediately following the U6 promoter (white). The fluorescent reporter eGFP (bright green) allows selection of successfully transfected cells via FACS.

Supplementary Figure 17: Conservation of potential miR-122 binding site across vertebrate species. Alignment of 800 nt sequences starting with the miR-122 pre-miRNA coding sequence from 18 eukaryotic genomes with the miR-122 seed sequence. Binding site indicated by *. Performed by Dr Vladimir Ovchinnikov.

```

Xtr      TGGAGTGTGACAATGGTGTGTTTGTGTC-AGAGCTATCAAACGCCATTATCACACTAATGAG
Dre      TGGAGTGTGACAATGGTGTGTTTGTATCATCTGTCGTCAAACGCCATTATCACACTAAATAG
GgaW     TGGAGTGTGACAATGGTGTGTTTGTGTC-CAATCTATCAAACGCCATTATCACACTAAATAG
GgaZ     TGGAGTGTGACAATGGTGTGTTTGTGTC-CAATCTATCAAACGCCATTATCACACTAAATAG
Aca      TGGAGTGTGACAATGGTGTGTTTGTATC-CAATCCGTCAAACGCCATTATCACACTAAATAG
Ami      TGGAGTGTGACAATGGTGTGTTTGTGTC-CAATCTATCAAACGCCATTATCACACTAAATAG
Cpi      TGGAGTGTGACAATGGTGTGTTTGTGTC-CAATCTATCAAACGCCATTATCACACTAAATAG
Mmu      TGGAGTGTGACAATGGTGTGTTTGTGTC-CAAACCATCAAACGCCATTATCACACTAAATAG
Mdo      TGGAGTGTGACAATGGTGTGTTTGTGTC-CAGTCTATCAAACGCCATTATCACACTAAATAG
Ete      TGGAGTGTGACAATGGTGTGTTTGTATT-CACAT-ATCAAACGCCATTATCACACTAAATAG
Cja      TGGAGTGTGACAATGGTGTGTTTGTGTC-TAAACTATCAAACGCCATTATCACACTAAATAG
Mml      TGGAGTGTGACAATGGTGTGTTTGTGTC-TAAACTATCAAACGCCATTATCACACTAAATAG
Hsa      TGGAGTGTGACAATGGTGTGTTTGTGTC-TAAACTATCAAACGCCATTATCACACTAAATAG
Ptr      TGGAGTGTGACAATGGTGTGTTTGTGTC-TAAATATCAAACGCCATTATCACACTAAATAG
Dno      TGGAGTGTGACAATGGTGTGTTTGTGTC-CACAATATCAAACGCCATTATCACACTAAATAG
Bta      TGGAGTGTGACAATGGTGTGTTTGTGTC-CAAACATCAAACGCCATTATCACACTAAATAG
Cfa      TGGAGTGTGACAATGGTGTGTTTGTGTC-CAAACATCAAACGCCATTATCACACTAAATAG
Pal      TGGAGTGTGACAATGGTGTGTTTGTGTC-CAAACATCAAACGCCATTATCACACTAAATAG
***** *                               ***** **

Xtr      CTACTGCAGGCACACTCACTCACTG-----GCTGTAGGCT-----CACTC
Dre      CCACGGT---GTGACCGCTCAACCT-----CCTCA
GgaW     CTACTGGTAGATGAGACATCAATTT-----GGAAAACATATGTGAATTGT
GgaZ     CTACTGGTAGATGAGACTTCAAGTCAGAGAAGCATGTAAGTTATA-GCGATATGTTTACC
Aca      CTACTGCTAGAGAAGCCATCGGTTC-----AGGTTTGGGGCAGGA
Ami      CTACTGTTAGATATGCCATCCATTTGCAGAAATAGGAGAGCTTTAAGCTATAAT-TTTTT
Cpi      CTACTGTTAGATAAGCCATCAGGTTCGAGAAGGAAGAGAGCTTTAAGCAATACATTTTCT
Mmu      CTACTGCTAGGCAATCCGTCCACTC-----CACGCGTGACTTGA--CGTCT
Mdo      CTACTGTTGGGCAATCCATCTGTTC----CAAGAAGAGCAGATTGAGCATTGGGCCATCT
Ete      CTACTGCTAGGCAACGCTCCCACCG-----GATACCCACCTTGG--CACCT
Cja      CTACTGCTAGGCAA---TCC-CTC-----GATAAATGTCTTGG--CATC-
Mml      CTACTACTAGGCAATCCTTCC-CTC-----GATAAATGTCTTGG--CATC-
Hsa      CTACTGCTAGGCAATCCTTCC-CTC-----GATAAATGTCTTGG--CATC-
Ptr      CTACTGCTAGGCAATCCTTCC-CTC-----AATAAATGTCTTGG--CATC-
Dno      CTACTGTTAGGCAATCCTTCCACTC-----AATAAATAACTTTG--CATCT

```

Bta C TACTGTTAGGCAATCCTTCCGCTC-----CCTGAAGGTGTTGA--CATCT
 Cfa C TACTGTTAGGCAATCCTTCTGCTC-----AATAAATGCCTTGG--CATCC
 Pal C TACTGTTAGGCAATCCTTCCGCTC-----GACAAATGCCTTGG--CATCT

* **

Xtr G G T C A C T G G C C G T A G G C T C A C T C G C T C ----- G C T G G C
 Dre G C C A A T C A G C G C C A - G A G A A C A G A T G ----- A C G T C G C A A C G G G C G G G A T T T T A C
 GgaW A G C T G A C G T G C G G A -----
 GgaZ C G C T G T A C T G A G G A - A T A A A T G A G C T T -- T T T T C C G T G T C A G G A T G A C T C T - T T C T T T T
 Aca G G A A G A A G A G A G A - A T A C A A C C G T T T ----- T C A A G A C A A C T C T T A C T T T T
 Ami T C C T G T A A T G A G G A - G T A T G A C A G T T T T G T T C A T G T A T G T C A T G A T G A C T T T G C T C T T T T
 Cpi G G C T G T G A T G A G G A - A T A A A G C A G T T T G T T C T T G T A T G T C A C G A T G ----- C T C T T T T
 Mmu G T T C T C C C G A G C A - A G A A --- G T T C ----- T G T C T T G T
 Mdo A C T G T G G A T G G A C A - A A A T G A A A G T T A ----- T G C T C A G
 Ete G T C T G G T G C A G G C A A G A G A G --- G T T C ----- A T C T G C T
 Cja A T T T G C T T T G A G C A - A G A A G --- G T T C ----- A T C T G A T
 Mml G T T T G C T T T G A G C A - A G A A G --- G T T C ----- A T C T A A T
 Hsa G T T T G C T T T G A G C A - A G A A G --- G T T C ----- A T C T G A T
 Ptr G T T T G C T T T G A G C A - A G A A G --- G T T C ----- A T C T G A T
 Dno G T T T G G T A T G G G C A C A G A G G --- G C T C ----- A T C T G A T
 Bta G C T C G G T A T G G G C A - G G A A A --- G T T C ----- A T C T T A C
 Cfa G C G T G G T A T G G G C A - G G G A G --- G T T C ----- C T C T T A C
 Pal G T T T G G T C T T G G C A - A G A A G --- G T T C ----- C T C T T A C

*

Xtr C G T A A G G C T C A C T C G C T C G C T G G ----- C C G T A A G G C T C A C T C G C T C G C T
 Dre G G C G G G A A T G T C A A A C G C G - C G T - G G A C T T C G T G A C T T T T T A A A -----
 GgaW A T C A A A C T C T A T A A C C A C A A A G T ----- A G T T A T A A A A T C A G ----- G T A T
 GgaZ G C A A G T G A A C C C A C A G T C A - T A C ----- G T G G G A A A A G C A C A C C C T G C C A T G
 Aca A T - G G A G T C A A A A G A C T A - T A C ----- A T T A A A T A A A ----- A A T A T G C T C T
 Ami A T A A G C A A T T C T A A G A T A A - T A T ----- G C A A T A A A T A A A ----- T A C T
 Cpi A T G A G C A A T G C T A A A A T A A - T G T ----- A C A A T A A A C A A A ----- C A C T
 Mmu G C T A G A G T T C T C C A T T G T G - T G T ----- A A T G T A T G A A T A A A ----- A T C T
 Mdo A T C A G T G C C T T C A C T C A T T - T G C A G G G G C C A A A T T A A T G G C C A A A - - A T G G T G G G C T G A T
 Ete C T C G G G G C T C T G G A C A T T - - C A C ----- A A T C A C T G G G T C A A ----- T G G C
 Cja A T C A G T G T T C C C A A T C - T G - T G T ----- A C T T A C A G A A T A A A T T A T G T G T T C T G T C T
 Mml - - - - - T G T T C C C A A T C T T G - T G T ----- A C T T A C T G A A T A A A ----- G T C T
 Hsa A T C A G T C T T C T C A A T C T T G - T G T ----- A C T T A C T G A A T A A A ----- G T C T
 Ptr A T C A G T G T T C C C A A T C T T G - T G T ----- A C T T A C T G A A T A A A ----- G T C T
 Dno A T C C G T G C T C A A A T C A T T - - T G T ----- A A T T C T T G A A T A A A ----- T C A C
 Bta A G C A G G G C T C C C A G T C C T G - T G T ----- C A C T G C T G A G G A A A ----- T T C T

Cfa AACAGTGTCTCGGTCCTGTTGT-----TATTACTGAATAAA-----TCCA
 Pal AGCAGCGCTCCCAGTCATG-TGT-----AATTATGGAATAAA-----TTCT

*

Xtr GGCCGTAGGCTCACTCGCTCGC-----TGGCCGTAGGC-TCACTCGCT-----
 Dre -----CTTGAGTGTTTGA-----TGAATTGGA-----
 GgaW GTTT-----AATCAACGCTGCG-----CAGACGCCAGGTGCAAGGGGGA-----
 GgaZ CGCC-----TCCCCAGACTCCCAGTGCAGAGCACTCAAGAGTCACCTAAAGGAGTCACAG
 Aca GCTT-----TCCTAAAACACTACACACTGGGGTCTTTCTGTTTTTTCCAGA-----
 Ami GTGC-----TGTATGCATTGCTAGTTGTAAAAATGTCTTGT-TTAAGCACA-----
 Cpi CTGC-----TGTACGTACTTCCAGTTGTAAAACTCCTTTT-TAAAGTACA-----
 Mmu GGTC-----CTCTTGTGCTTAT-----AACTGTACAT-CTGACTGAC-----
 Mdo AATA-----TTTATGTGCTCTG-----AGAGCCCGAAT-TTTTCTCAA-----
 Ete AGTA-----TGTCTGTGCTGG-----AAACAACCTGT-CTATCTGGA-----
 Cja GGTT-----CTCTTGAGCTCAT-----AACCAATATGA-TTAGCTGAA-----
 Mml GGTT-----CTCTTGCACTCAT-----AACCAATATGA-TTAGCTGAA-----
 Hsa GGCT-----CTTTTGCACTCAT-----AACCAATATGG-TTAGCTGAA-----
 Ptr GGTT-----CTTTTGCACTCAT-----AACCAATATGA-TTAGCTGAA-----
 Dno AGTA-----CTTTTGTGCTCGT-----AAATAGTACAT-CTATTTGGA-----
 Bta GGTC-----CTTTCGTGCTCAT-----AACCAACATG--TCATCTGAA-----
 Cfa GGCC-----CGTTGGTGTGCTCGT-----AAGCAGCATG--TCATCCGAA-----
 Pal GGTC-----CTTTTGTGCTCAT-----AACCAATATG--TCATCTGAA-----

*

Xtr ---CGCTGGCCGTAGGCTCACTCGCTCGCTGGCCGTAGGCT-CACTCGCTCGCTGG---
 Dre ---GTCTGCTCAAATAACACAAGAAGTCCCTCGCTAGTGACCCAGTTG---CTAG---
 GgaW ---TAGCTC-----CTCCT-----AACTTGCACACTTTTCAGG
 GgaZ CACCAGTGTAGGAGTGCT---GATTTATAGGAGTCAGTGCAGCCG-----
 Aca -----GATATTTA---GGATTCG-----TATTTG---GCTGA---
 Ami ---CATTGTTATCAACTTA---ATGCTCCC-----TATTTG---CTCC---
 Cpi ---CATTGTAATGAACCTTATG---CTGTC-----CATTTG---CTCT---
 Mmu ---GCCTTC-----CCCT-----CGTGTG---CTAT---
 Mdo ---CATTG-----CTCT-----TAGCCA---TTGC---
 Ete ---TCCAGCTGT-----CGCCTG---CTGA---
 Cja ---CCCTGC-----CTCT-----CATGTG---CTAT---
 Mml ---CCCTGC-----CTCT-----CACATG---CTGT---
 Hsa ---CCCTGC-----CTCT-----CACATG---CTGT---
 Ptr ---CCCTGC-----CTCT-----CACATG---CTGT---
 Dno ---CTTGGC-----CTCT-----TACCCG---CAGT---
 Bta ---TGTGTC-----CTCTC-----CTCATG---CTGT---
 Cfa ---CTTTGC-----CTCT-----CGGATGTTCTCTGT---

Pal ---CTGTGC-----CTCT-----CACACG---CTAT---

 Xtr -----CCGTA-----GGCTCACTCGCTCGCTGGCCG-----
 Dre -----TATTA-----GAGTTGATAA-----
 GgaW AAAAAGCATGCTACAGATATAGGCTGAACTAATACATAATCAATAGTTGACAGGAATTCG
 GgaZ -----TGATA-----GCCTCTGCAAAACAAGGCTGGCC-----
 Aca -----GATTA-----TGCATCAGGTGGTGGGTCA-----
 Ami -----TACAA----AGAAGGGTAATCTTTCTTATATTCATGGGTTGTAA-----
 Cpi -----GACTG-----AGTCATGTTTATATTCATGGATTGTAA-----
 Mmu -----TGTGA-----AA-----
 Mdo -----CATCA-----ATAGATGGTTTCCCATAGGA-----
 Ete -----TGTA-----TGTCATACCCTAACAGATGG-----
 Cja -----TGTTA-----CATACCTTAGCTGGTGA-----
 Mml -----TGTTA-----CATACCTTAGCTGGTGA-----
 Hsa -----TGTTA-----CATACCTTAGCTGGTGA-----
 Ptr -----TGTTA-----CATACCTTAGCTGGTGA-----
 Dno -----TGTTA-----AATACTTAGCTGATGA-----
 Bta -----TGTTT-----TAAATAATGATTT-----
 Cfa -----TGTTA-----AACACGTTAGCTGGTGA-----
 Pal -----CGTTA-----AAGACTTAGCCGGTGA-----

 Xtr -----TAGGCTCAC-----TCGCTCGCTGGCCGT-----
 Dre -----TGAG-----TTAAACAATTTCTTT-----
 GgaW GATACATATTCATTACATTCCTGAGAACCATTAGCATTAGTCCCTCCCCCT-----
 GgaZ -----TGAGATTTTCT-----TTTTCT-----
 Aca -----TAAAGTTCTGTTTCAGGTCTCCTATCTCATCCCCTCTCCAT-----
 Ami -----TGATTTTATTTATATGTGTAAATGGTTTGTGCCTCTCCT-----
 Cpi -----TGTTTTTATTCAGCAATGCAAGTGCTGTGAATCTTCCC-----
 Mmu -----TGAA-----TAATGCT-----
 Mdo -----TGTAAGCTCCTCGAGGGCAGAGGCTTTG-----TTTTCTTTTCATT
 Ete -----TGAACCTTT-----TTGTCATTTTTTCCC-----
 Cja -----TGAATTTCC-----TTG-----TTTTCCC-----
 Mml -----TGAATTTTC-----TTG-----TTTTCCC-----
 Hsa -----TTAATTTTC-----TTG-----TTTTCCC-----
 Ptr -----TTAATTTTC-----TTG-----TTTTCCC-----
 Dno -----TGAATTTCT-----TTG-----CTTTTCT-----
 Bta -----TGAGTCTTT-----GTG-----TTTTCCC-----
 Cfa -----TGAATTTTG-----TTTTCT-----
 Pal -----TGAGTTTTTC-----TTG-----TTTTCT-----

*

Xtr -----AGGCTCAC-----
Dre -----GGATTGTAAGACATCAGAGAAG-----
GgaW -----CTCAT-----
GgaZ -----GTTGGA-----
Aca -----ACAACCCATAAACCTAACTT
Ami -----TTAGTTTTCTTCAGGGGAGACTTTTATTGAATTA-----
Cpi -----TTATTTAGATTTCTGT
Mmu -----ATAAAGGAAAAGACCAGTAAAATCCAACC
Mdo GTTGATCCCCAGCTTATCACAGTATCTTGTACAGAGTAGGGCTTAAT-AAATGCTTAT
Ete -----AAGCCAATAAA-----
Cja -----AAGTCAAT-AAA-----
Mml -----AAGTCAAT-AAC-----
Hsa -----AAGTCAAT-AAA-----
Ptr -----AAGTCAAT-AAA-----
Dno -----GAGTCAATAAA-----
Bta -----AAGTCAATAAA-----
Cfa -----GAGTCAATAAAA-----
Pal -----GAGTCAATAAA-----

Xtr TCGCTCGCTG-----GCCGTAGGCTCACTCACTCACTG-----GCA
Dre AATGACGCAG-----CTGTGTGAGGTAAACTACTTAATTCAAAA-----TGGAGGCAAGG
GgaW GCGTGTGTAG-----TGTGTCATGAGGGTCTCCTCCTGGTGGTGTGGGATGAAGG
GgaZ TTGGCCGTAG-----CTGTGGAGCACTGAAGCCAATCACCTGGTG-----GAGTCAAAG
Aca TTACATGCTG-----AGCCCTCTTCCCAACA-----TGACCATG
Ami CATCCTGCAG-----CATGTTTTGAAGCTGTTTTGTAAGTA-----
Cpi TAGGTTATAG-----CTGGCAACCTGCAGCCTCTATTTAAGCA-----GTGTG
Mmu TCTCCTATGG-----GACCCTCCCAACACTGACTACCTGACA-----
Mdo GAAATTGGGG-----TGTATTTCTGAGTCACGAAAAGGCACAC---ATGGCACAG
Ete ATGTCAGCAG-----GGACCCTCAGCTCAG-----GGGGAGG
Cja AAGATTGCAGACTGTGTATGGTGGCTGACGCCTGTAATCCCAGCACTTTGGGAGGCCAAA
Mml AAGATTGCAGACCAGGTGCGGTGGCTCAGCCTGTAATCCCAGCACTTTGGGAGGCCAAG
Hsa AAGATTGCAGACCAGGCACGGTGGCTCAGCCTGAAATCCCAGCACTTTGGGAGGCCAAG
Ptr AAGATTGCAGACCAGGCACGGTGGCTCATGCCTGAAATCCCAGCACTTTGGGAGGCCAAG
Dno AAGATTGCGA-----TGTCCCTCAAGGTACCTAGCCCAAGA-----CTGGG
Bta GCGATTACAC-----CGTGCCCCAGAGC----ACCCAAGCA-----CCAGGTTAGT
Cfa AATATGGCCG-----TGTCCCCAGAGC----ACCCAAGCA-----CAAGGCCAGG
Pal AATATCGCAG-----TGCCCCAGAGC----ATCCAGCA-----CAAGGCTAGG

Xtr TCAATTCAAGAGGCCCCAGGTGCAAACATTTTAGCCTATAGTATTCTCTGTTA-----
Dre T-----TTAAGCTTTTTGAGCATA-----TATAGTGGCATAAGATCGCTGATGTTGAG
GgaW CTCA--TCGTCTTCCT---CGCAAGGCTTGTAGTCTTCTCGCTGTAAAC-----
GgaZ CCTGGTTCTTATCTTGCACCAAAG----TAAAGCAGATATGTTTCAAACAA-----
Aca TACTGATTGTAGCAATGT--TCTACATGTGTTGTTTCAGTAGTACAACAAGCAA-----
Ami -----TTTAATGCCTGTAGTTAACAAGTTTTATTTTGAATATATTCTTCAATGAGACC
Cpi TCAGTATTTAATGCATTTAATTAACATGTTTTGCTCTGGAAGATATTCTTCAATTAT---
Mmu -----CCCAATTTAT-----TTCCAGCA-----
Mdo CATG-----CTATTAGACACAGGATGCCACAATAAGAATTTCTTGAAACAGTAGGAA
Ete CACG--TGGGACCTCTCACTCCAGGAATGGGGGGATGAAA-----CAAGCAGCCAGGAA
Cja GCAA--GTGGATTGCTTGAGCTCAGGAGTTCAAGACCAGTC-----TGGGCAG-----
Mml GCAG--GCGGATCACTTGAGCTCAGGAGTTCAAGTCCAGCC-----TGAGCAA-----
Hsa GCAG--GCGGATCACTTGAGCTCAGGAGTTCAAGTCCAGCC-----TGGGCAG-----
Ptr GCAG--GCGGATCACTTGAGCTCAGGAGTTCAAGTCCAGCC-----TGGGCAG-----
Dno CACG--CCATAATCATTGACTCAG-----
Bta CTCG--TGGTAATTGTTCAACCCACGCTTACTGAGCTGAAA-----TGGAGAG----GAA
Cfa CACG--TGGGAACTGTTCAACCCAGGCTTGTTAAGTTGAAA-----TGAATAGTTAGGGA
Pal CGCA--TGGAAACTGTTCAACCCATGCTTGTGAATTGAAA-----TGAATAGTAAAGAA

*

Xtr -----TTTAGTGTATTTTGGCCACATTAACAAAAACAACAT
Dre GATTTTCAGTCACAAAGCT-----TTTATGTAGTTTTTTT---TTTTTAAATATTTTTT
GgaW -----TTCTGAC---CCTGTGGGAGAGGCAC
GgaZ -----TAACTGCCTGAGATGATCGTGAT---TTT-CAAACAAAACAC
Aca ---CCTTGGACAGAAATCTTAAAGCCCATTTATATTATGCT-----AAATGC
Ami TATTGTTGTACTGTACAAAACAACTTATTGGATTTGTAAT---GGAACCAAATGAACAA
Cpi -----CTCATTGAAATGATCTC-----ACAA
Mmu -----CTAGGCCCACTCTCGAT---TTTGCTCTGAACGCTA
Mdo TGAAGCAGATACAGGAAAAGAAATC-----ATCTGAGA---AATGCAAATAGCTGGC
Ete AGCGCTAGGTAAAGAACA-----CTACAAGCAATCGAGGT---GATTCGAACTGCCAC
Cja -----CATCGTGAACCTGTGTC---TGTACAAAACACTACAA
Mml -----CATGGTGAACCGTGTC---TCTACAAAACACTACAA
Hsa -----CATGGTGAACCGTGTC---TCTACAAAACACTACAA
Ptr -----CATGGTGAACCGTGTC---TCTACAAAACACTACAA
Dno -----
Bta AATGGTAGATGCGGAAAACATAAGCTGGTTTTAAATCTAGGG---TTTGTAAACCACATGC
Cfa AGTGTTAGATCTGGAAAATAAAAAGCGGGTTGAACCTCAGT---TGTACAAACCATAGAC
Pal AGTGTTAGACACAGAAAACATAAACTAGGTTAAATCTAGGT---TATGCAAACCACAGGC

Xtr ATAATTT-----
Dre AAAATAGAAAG-----TGTGACTATTGCACCCATTTAT-----
GgaW CCCC-----CAAACCAGTTTCAGCTTGCCT-----
GgaZ CCTTCAG-----
Aca ATAGCATTTGGGCAT-----CAAAGCATTTATTTCTTCATT-----
Ami ATGATTTGCTGCCAACTCATCAAAGTGA-----
Cpi AGACCATTGTGGCAC-----TGAAATAAATTTGTTGGAGGT-----
Mmu G-----CGAGGCATTATTTCCATTCACTGTTGACACTGCCTCTGTG
Mdo TTATTATTTAGGTAGAAAATGAAGGTATGTGGCTTTTAT-CT-----TTTCATTATAA
Ete AGAGAGCTACCCAG-----GAAGGTGGGTGGGTCTCC-----
Cja AAATTAGCCAGGCAA-----GGTGGCATGCA---CCTGTAAT-----CC-----
Mml AAATTAGCCAGGCAT-----GGTGGAAATGTG---CCTGTAAT-----CC-----
Hsa AAATTAGCCAGGCAT-----GGTGGCATGTG---CCTGTAAT-----CT-----
Ptr AAATTAGCCAGGCAT-----GGTGGCATGTG---CCTGTAAT-----CT-----
Dno -----TAAGGTAGGTGGGTCTCCTGT-----TTTCATCA
Bta TTATAGGTAAAAAT-----GAAGGTATGTGGATCTTCTACT-----TGTCATCA
Cfa TTATAGGAACAAAAT-----TAAAGTATGCGGGCCTCACTT-----TCCACTGG
Pal TTATGGATGCGAAAT-----CGAGGTAACGGGGTCTCCATT-----CTCCATCA

Xtr -----CACTTTGATA-----TAATTCCTCT-----GCGGT
Dre -----GAACACATGAATA-----TTATAT-----ATATAAT
GgaW -----GTGGACATCTAATC-----ACCTCAT-----
GgaZ -----AATTCTGAT-----GAGAAGC
Aca -----AGCAAGTTAATATTTGCTGCTGTTTTAGTT-----TGGAAG-
Ami -----CAACAATATAAAA-----AAACTCAGGT-----GTCTGGC
Cpi -----CAGTTATACAA-----ATACAAC
Mmu GAGTTAGCCCTGAATCCAGCAGGTGGATG-----GCACAGTCC-----CTAAGAC
Mdo AATTCTTTTCAAAAGCACACTTGGTAATA-----AACTTATGG-----ATG--GC
Ete -----CTGCCCCACCAGA-----AACAAAGGCACTTCGGTGTGGC
Cja -----CAGCTACTTGGGA-----GGCTGAGGT-----GGGAAGA
Mml -----CAGCTACTTGAGA-----GGTTGAGGT-----GGGAGGA
Hsa -----CAGCTACTTGAGA-----GGCTGAGGT-----GGGAGGA
Ptr -----CAGCTACTTGAGA-----GGCTGAGGT-----GGGAGGA
Dno GAAAA-----AACTTACTAATA-----AACTTGCAC-----ATG--GC
Bta TGAAGA-----CGACTCTCCAATA-----AACTTGTAC-----ATG--GC
Cfa AAGAAAA-----CAACTGTCTGATA-----AACCTACAC-----CTG--GC
Pal TAAGAAAAACAACGCAAACTTCTCAGA-----GATTATAT-----ATG--GC

Xtr T-----CTGCTCCAGTACCGGTAGGGTACGAAATGTGTAAGGCGGCATTTTTAGG

Dre TATATATATAATTATCGAAATATGGAAGAATAACAGAATAAAATTTG--ATTGTTTTTTAC
GgaW -----TGCCAGACATCCTTGGGCTGTTCTCA-AACTCTTATCTTT
GgaZ T-----TCTACATCTAACAGGGAACGTGATTTATTCAGGTTTCAGCTCCGG
Aca -----AGAAGGGTTTTTGTGTTGTTGCTGTGA
Ami AGACATTAACAATGGGGCAACTGGGATCAGAGGAATTTATGCCCATCCCAACATC
Cpi G-----CACTGGAT---TCATGATTAAGAAAATGAATCAGGAGT-----A
Mmu C-----CAACT-----TCAGATGTCCATTGAATGTCCTAA--GTTGCTTTCCTT
Mdo T-----TAATTGGCCTTCTAGAAATGC-TAGCACACGTTG--ATTTCTATT--G
Ete T-----TCCCTGCCTTTTCAGAGAGTC-TAGCGTGTGTTGCGCACGCTACT--G
Cja T-----CAGTTAAGC--CCATGAAGTC-AAGGCTGCA-GTG--AGCTGTGAT--G
Mml T-----CAATTGAGC--CCATGAGGTT-GAGGCTGTA-GTG--AGCTGTGAT--G
Hsa T-----CAATTGAGC--CCACGAGGTT-GAGGCTGTA-GTG--AGCTGTGAT--G
Ptr T-----CAATTGAGC--CCATGAGGTT-GAGGCTGTA-GTG--AGCTGTGAT--G
Dno T-----TACCTGACATTTTCAGAGATTC-TAGAACATGTTG--GCTTCTATT--G
Bta C-----TAAATGAGCTTTCAGAAATTC-TAGAATGTGTTG--ATTTCTACT--G
Cfa T-----TAACTGACCTTTCAGAAATTC-TGGAACATGTTG--ATTTCTACT--G
Pal T-----TAAACGATCGTTCAGAAATTC-TAGAACATGTTG--ATTTCTACT--C

Xtr GGTTTGTATGCTTTTAATAAT---GATTTTTGAATAATGTGAG-----
Dre TTTTCTCTCTCTCTCTCTC-----TCT
GgaW ATGGCCTTCATCCTCCTCATTATCCCAGACCCAGTGTCTG-----CAATCC
GgaZ TCCCTTCTATTTTACAATGCTGTGATGATCTGAGTTCGGGGAGGT-----CACAA
Aca TTGACTTTTGATGAATAAGCT-----
Ami ATGTCCTCTGTGCCTTATACT-----
Cpi TTGCTGCCAAATCATCAAACCTGACGTGACGTTG-----
Mmu GTGTTTCTGACTCACCAGCT----ATAAACTGGGGTTCCACACCTGTGTTCTCAGGTC
Mdo ATTGTATTCAATTATTGTGCT---GATTTCTGTATT-TGTATGT-----GATATC
Ete ATCTTCTCGGCGACCAAGCT---GATTCCCGAATG-TGTGTGC-----TAGGTT
Cja GCACCGGGCACT-CCAGCCT---AAGTGACAAA-----GC-----AAGATC
Mml GCACCACTGTACT-CCAGCCT---GTGTGATAGA-----GC-----AAGATC
Hsa GTACCACTGCACT-CCAGCCT---GTGTGACAGA-----GC-----AAGATC
Ptr GTACCACTGCACT-CCAGCCT---GTGTGACAGA-----GC-----AAGATC
Dno ATTTTATTCAATTACCAACCT---GAATACAGAAGTATTAATGC-----TAGGTC
Bta ATTTTGATCAGTACCAGGCT---GACTGCTGAATT-TGGATGC-----TAGGTC
Cfa ATTTTAACCAATTACAGGGCT---GATCACTGAATT-TGTACGC-----CAAGCC
Pal ATTTTAATCCATTACCAGGCT---GATTACTGAATT-TGTATGC-----TAGGTT

;| ;:| | |

3p-guuugugguaacaguguga-ggu-5

Xtr -----TTTTGATCTTTTGGATTTTGTGTTTGGCTTG

Dre CTCTCTCTCTCTCTCTCTC-----TCTCT
GgaW CCATTCTTCTAACGACCTCTACATA-----AACTAT
GgaZ TTCCTCGTGAAATTCCTCTGAAAACCTTTGTGAAGAACAGGAGTCAGTG-----
Aca -----TCAGAA---TGGAGGGATAATTCCATTTTGG
Ami -----TCCCACAACAAG-----AGGAAA-----
Cpi -----TCAGAAAAGAAACCCGGGGCCCA
Mmu CAATTCTCT---TGCTTGAGTGGG---TCTAAAAATTCAGGAAAACGTG-----
Mdo CTCTTTCTTAGTGCCTCCTTAA---TTATC-----TTCTATGT-----
Ete CGGCTCCCTTC---TGTCTCCGTGAA---TCTCCA---TGAGAAACTCACACCTCAGCCTG
Cja C-----TGCTCAAAAA---TAAAA---TAAAAAGATTGCA---GAGTCC
Mml C-----TGCTCAAAAA---TGAAAA---TAAAAAGATTGCA---GAGTCC
Hsa C-----TGCTCAAAAA---TGAAAA---TAAAAAGATTGCA---GAGTTC
Ptr C-----TGCTCAAAAA---TGAAAA---TAAAAAGATTGCA---GAGTTC
Dno TGCTTCCTA---TGTCTCAATGAA---TTGTGA---TTAGAAAATTATATTTGCGACTGT
Bta C-----TGATGGG---TTTCCA---TCAGGAAATTACA--TGAACTAC
Cfa CTTTTCTTA---TGTTCATGAA---TTGCCA---TCAGGAAATTTTACATGAACTAC
Pal CTTCTTG----CCTCTCAATGAA---TTACCA---TCGGGAAACGATACGTGAGCTAT

Xtr CCC-----TTTAAATCTGGGGCAGTGTCCCATTCACCTTGG
Dre CTCTCTCTCTATATATATA----TATATATATATATATATATATATATATATA
GgaW CTCTAACTACCCCTATCCCAG----TTTCTGTG-----AGGCCCTCTCTC
GgaZ GTGGGAGCTCCA--AACAGG----TT-----ATGAAGGCCTGGCAA
Aca CTCTCAGCTCTGTACCACTCA-----GCTGTGTACATATTCATCATTAGCTAAGTGTG
Ami -----GAAACTTTGCCCA
Cpi GTCCTGCAACCCTTACGCAATAGTGTTTGGGCATGTGGTCTCATGGAAGCCAA---TG
Mmu -----TTGGCCAG----CTTATTCCAAAGGATAATTTAAAGGCAGAGTGGA
Mdo ---TGAGCAACTGAAAACAAC----TTTTGAAATT-----
Ete -----TCCCAGCAGTGAAGCTCACGCAGTTGCTG----
Cja CTCAGAGCACCCAA-CACAAG-----ACTAGGCACCAGGTGATCATTCAACCCATGCTTA
Mml CTCAGAGCACCCAA-CACAAG-----GCCAGGCACCAGGTAATCATTCAACCCATGCTTA
Hsa CTCAGAGCCCCAA-CACAAG-----GCCAGGCACCAGGTAATCATTCAACCCATGCTTA
Ptr CTCAGAGCCCCAA-CACAAG-----GCCAGGCACCAGGTAATCATTCAACCCATGCTTA
Dno GTGTGAACTGCCATAAGCAAA----TCTGCACATT-----CCAAAGCCA
Bta ATATGGGCTGTGAT-AGCGAA----TCTGAGCATT-----TTAAGCCAGTGTAG
Cfa GTGTGAGTTGTGAT-AGCAAA----TCTGAACATT-----TTAAACCTGAGCAAG
Pal GTGTGAGTTTCCAT-AGCAAA----TCTGACCACT-----TTAAACCCGAGC---

Xtr CTCATTTTGACTGGCCATTGAGGATCCTTCTTT----TGCTATAGTATAAAAC-----
Dre TATATATATGTTGATCAGT-ATTATCAGTATGGCATCATAATGAAATAAAAAAACTGCT

GgaW TCTTTT-----ACTGT-GGGACCCTTCTCT---CATCCTCATAGCAATACATTTTC

GgaZ CCACGCTGCACTGGAAAGC-----CTA---AATCCCACTGAATAATCCAAGACT

Aca AGTGAAATTACTGCTAACT-TGTCATACTCTGG---CAGGCAAACAACAAAGATCAATTT

Ami CGATTCCCAGATTAACAGT-GGGCCACTTCTTG---CAAACCTTGGTATTATCTTGGGCT

Cpi AAAGCTCGGGGTGTTCAAT-----ATCTTTA---AAAATGTAGGGTAAGATTATGAGC

Mmu GATATGTTCTCCGGGCACTGCTGCATATTTTTG---GCTACCTAGAAGCTCTACAAAGCT

Mdo -----GTT

Ete -----GGGCAGT-CGGATC--CCTTG---GGTGGTCTGGCTTGGTGCAAAGG-

Cja TAGAGTTGAAGTGGATAGT-TAGGAAAGTGTTA---GATGCAGAAGGCTCAGACTGGGTT

Mml TAGAGTTAAATGAATAGT-TAGGAAAGTGTTA---GATGCAGAAGGCTAAGACTGGGCT

Hsa TAGACTTAAATGAATAGT-TAGGAAAGTGTTA---GACGCAGAAGGCTAAGACTGGGCT

Ptr TAGACTTAAATGAATAGT-TAGGAAAGTGTTA---GACGCAGAAGGCTAAGACTGGGCT

Dno TACCGCTTCTGGGGATGGT-TGGACCA--CCTA---GGTGCTCTGGTTTGAAACACATT-

Bta CGCCTG-----GGACACT-TGGCCA---GCTG---GGTGGTCTGGTTTGAAACAC-GTT

Cfa CTTCCA-----GGACAAT-TGGACCAGC-TGG---GGTGTCTCAATTGAAATACTTTT

Pal -AGCTTCTG---GGACAGT-TGGACC--GGCTG---GGTGGTCTGGTTTGAAACAC-ACT

Xtr -----GTTCTGCCAGG--GAGGAAACAGCCATTTCTTTATTTCAGCTTTATTCTGTTATT

Dre CTAAGATTTATGTAAGAAAATATGACCAAATGTTTCATGT-----TT

GgaW CC---ACTGTACTGAGG-----AATAAATAAGCTTTTTT

GgaZ CA---GGCTCACAGAAT-----TA

Aca CC---ATCATTGCCATA-----TA

Ami CA---ACT-----TA

Cpi CCCTTACTCGTGTGAA-----TA

Mmu GT---CCTTCATAGATTTTATGGGTACCCCTTGGCATGGGTTTGATTTCAGTAGACAT-GA

Mdo AA---CTCTT-----TA

Ete -----AGCTCTGGGAGT-----GAGTGCTCAGAGCAGGGCCTGGGTTG

Cja AA---ATCTAGGTTATG----CAGACCGCAGGTTTATAGGTATAAATT-----TA

Mml AA---ATCTATGTTATG----CAGACCGCAAGCTTATCGGTACAAAAT-----TA

Hsa AA---ATCTATGTTATGCAGACAGACCGCAAGCTTATAGGTACAAAAT-----TA

Ptr AA---ATCTATGTTATG----CAGACCGCAAGCTTATAGGTACAAAAT-----TA

Dno -----CTCTCTCTCAGG-----TTTGTTTAAAGT-----TA

Bta CT---TTCTCTGAGAAG-----TTTAAAG-----TA

Cfa TT---TTCTCTGAGAG-----CATTTAAAG-----TG

Pal CT---TCCTCTGTGAGG-----TGTTTCAAG-----TA

Xtr GCATCTGTCCATCTCTTTCTACTCTGTATTTTCAGATAATAAAGGCAAATTC-----

Dre GATTGTGTTGTTTGCTGTTCTTT-----

GgaW ACATGTCAGGATGACTCCTTTCTTTTATAAGGAAACCCACAATAGTATGTGGG-----

GgaZ -----ACTGCTTTCTGATCATTGCAAGGG-----
Aca AATATTCCAAGGTGTGTATGTGCTTCATTAATAATCAATGGAGAGGATGTGA-----
Ami -----GTGGTTTTTTTTCCC-----
Cpi GTCCCTACAAATAGTTCTGTGAAGCT-----
Mmu AAATGTGCTGGCGACCCCAAGCCTCATGGACAGATAAGGAGGTGATGTACA-----
Mdo -----
Ete AAAGCCGGAGACTCTTGCCCTGGGGTCACCTGAAACCTCACCCACAGATAGGC-----
Cja AGGTGTGTGGGTCGTTTTCTTTTCATTGTTAAAAAAC-----
Mml AGGTGTATGGGTCCTTCTCTTTTCATCTTTAAAAATCCAACCTTCTAA-----
Hsa AGGTGTGTGGGTCCTTCTCTTTTCGTCGTTAAAAATCCA-----
Ptr AGGTGTGTGGGTCCTTCTCTTTTCGTCGTTAAAAATCCAACCT-----
Dno AGGAGTT---GTTTCTCTTTGTGCAGGGCTAAATGTCTGTCCCATGACTTGGGTAAATC
Bta AGAGGTA---GGTTCTGTCTACGCTGGGCTAAACTTCTCTCTGCGACTTGGCTAT---
Cfa AGGAGCA---GGTTCTGTTTGTCTGG-----
Pal AGAAGTA---GGTTCTGTTTGTGCAGAGC-----

Xtr -----
Dre -----
GgaW -----
GgaZ -----
Aca -----
Ami -----
Cpi -----
Mmu -----
Mdo -----
Ete -----
Cja -----
Mml -----
Hsa -----
Ptr -----
Dno AACTAGTATACGTATACTCAAGTTGACGGGATTGGATAGACCATCAGAAACCAGTTCATT
Bta -----
Cfa -----
Pal -----

Xtr -----
Dre -----
GgaW -----
GgaZ -----

| | |
|-----|---------------|
| Aca | ----- |
| Ami | ----- |
| Cpi | ----- |
| Mmu | ----- |
| Mdo | ----- |
| Ete | ----- |
| Cja | ----- |
| Mml | ----- |
| Hsa | ----- |
| Ptr | ----- |
| Dno | CAGTCATCAAAAA |
| Bta | ----- |
| Cfa | ----- |
| Pal | ----- |

PIPS Reflective Statement

Note to examiners:

The Professional Internship for PhD Students is a compulsory 3-month placement which must be undertaken by DTP students. It is usually centred on a specific project and must not be related to the PhD project. This reflective statement is designed to capture the skills development which has taken place during the student's placement and the impact on their career plans it has had.

PIPS Reflective Statement:

I undertook my three month placement within the Mechanistic Biology and Profiling (MBP) group, in Discovery Sciences at Astra Zeneca in Gothenburg, Sweden. During my internship, I performed biochemical and cellular screening assays to provide mechanistic and efficacy data for iMed projects.

My key aim when I started my placement mostly focused on wanting to understand the bigger picture within research, not just the academic side. During my time I was able to understand how the different departments (structural and biochemical assays, compound management, biophysics etc.) aid each other and make the research more streamlined and efficient and how a simple assay I performed as part of MBP fits into the bigger picture of iMed and drug discovery as a whole. I also gained an appreciation of the team dynamics for such a large research group and was able to compare this to the much smaller teams I was a part of during my undergraduate industrial placement in Belgium.

Even though I was performing new assays and using new equipment/techniques I had no prior experience with, my PhD has prepared me for such and I was able to get up to speed and be independent within such a short amount of time. As a result I have gained confidence and recognise I am more skilled than I would have otherwise realised. I was also able to improve my communication skills, for example presenting my results to chemists in a way suited to them rather than other biologists. In addition, I also had the challenge of conversing with non-native English speakers both at work and outside. It was also nice being able to compare my confidence with this placement to my industrial placement during my undergraduate degree.

Finally, I feel as though gaining experience of a wide range of biochemical/cellular techniques and of research within another large pharmaceutical company should put me in a good position for future job/postdoc applications as well as this placement giving me a chance to increase my professional network. Before I left, I ensured I had in-depth discussions with line managers and postdocs within MBP to critique my CV and discuss

future possible options, including the benefits of doing a postdoc even though I have a clear aim of entering industrial research rather than stay in academia. Overall, I was reminded of what I enjoy about research and found my motivation for my PhD.