Gene editing and specific MicroRNA inhibition

Lataniotis, Lazaros

Awarding institution:
King's College London

The copyright of this thesis rests with the author and no quotation from it or information derived from it may be published without proper acknowledgement.

END USER LICENCE AGREEMENT

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International licence. https://creativecommons.org/licenses/by-nc-nd/4.0/

You are free to:

- Share: to copy, distribute and transmit the work

Under the following conditions:

- Attribution: You must attribute the work in the manner specified by the author (but not in any way that suggests that they endorse you or your use of the work).
- Non Commercial: You may not use this work for commercial purposes.
- No Derivative Works - You may not alter, transform, or build upon this work.

Any of these conditions can be waived if you receive permission from the author. Your fair dealings and other rights are in no way affected by the above.

Take down policy

If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
Gene Editing and Specific MicroRNA Inhibition

Lazaros Lataniotis, MSc

School of Cardiovascular Medicine & Sciences
British Heart Foundation Centre of Research Excellence

August 19

Submitted for the degree of Doctor of Philosophy

Supervisors: Dr Anna Zampetaki
Prof Manuel Mayr

King’s College London, London, UK
Abstract

Background: MicroRNAs (miRNAs) are short, non-coding RNAs that are involved in the post-transcriptional regulation of gene expression. They are important regulators of diverse physiological and pathophysiological processes. MiRNA families and clusters are two key features in miRNA biology. Previous work on miRNAs in abdominal aortic aneurysm has focused on miR-195, a member of the miR-15 family. MiR-195 exhibits high homology with the other members of the family and therefore, its manipulation, overexpression or inhibition, inevitably affects the expression of other miRNAs of the family.

Objective: The current project explores the use of CRISPR/Cas9 as a gene editing tool to delineate the expression of miRNA families.

Methods and Results: We employed CRISPR/Cas9, an RNA-guided system able to provide highly specific alterations of targeted sequences in the genome of eukaryotic cells. We designed RNA guides and validated their efficiency in editing the mir-195 locus in various cell types while no off-target mutagenesis was observed. Quantification of the levels of miR-195 and other members of the miR-15 family, revealed a significant decrease of miR-195 and in the expression of the other miRNA in the cluster, miR-497. Three more miRNA clusters composed of miRNA members of the same or different family were analysed. Although no gene editing was detected in the miR-497 genomic locus, computational simulation revealed alteration in the three-dimensional structure of the pri-miR-497~195 that may affect its processing. RNA pull-down experiments, using wild type and an edited pri-miRNA of shorter length, were performed and the eluted proteins were analysed by Mass Spectrometry.

Conclusions: Our findings suggest that CRISPR/Cas9 is a powerful gene editing tool able to provide highly specific editing of the genomic locus of miRNAs, resulting in decrease of the levels of their mature transcripts. They also highlight different regulatory mechanisms in miRNA cluster regulation and function.
# Table of Contents

Abstract .......................................................................................................................... 2

Table of Contents .......................................................................................................... 3

Table of Figures ............................................................................................................. 7

Acknowledgments ......................................................................................................... 13

Declaration of independent work .................................................................................. 14

Abbreviations ................................................................................................................ 15

Chapter 1 INTRODUCTION .................................................................................... 19

1.1 MicroRNAs (miRNAs) ................................................................. 19

1.1.1 MiRNA biogenesis ............................................................................. 19

1.2 Abdominal Aortic Aneurysm (AAA) ................................................ 26

1.2.1 AAA pathogenesis ............................................................................. 27

1.2.2 SMCs in health and in AAA disease ................................................ 28

1.3 Gene editing ................................................................................................. 38

1.3.1 Clustered Regularly Interspaced Short Palindromic Repeat (CRISPR)/
CRISPR associated (Cas) ......................................................................................... 39

1.3.2 Limitations of CRISPR editing ............................................................ 44

1.3.3 CRISPR/Cas9 and miRNAs ................................................................. 47

Chapter 2 AIMS OF THE STUDY .............................................................................. 48

Chapter 3 MATERIALS AND METHODS ................................................................. 49

3.1 Cell Culture ............................................................................................... 49

3.2 Single Guide RNA Design and Cloning .................................................... 50

3.3 In Vitro Transcription (IVT) of sgRNA ....................................................... 53
<table>
<thead>
<tr>
<th>Chapter 3</th>
<th>Section</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.4</td>
<td>DNA Extraction</td>
<td>55</td>
</tr>
<tr>
<td>3.5</td>
<td>Genomic PCR</td>
<td>55</td>
</tr>
<tr>
<td>3.6</td>
<td>PCR Clean Up</td>
<td>56</td>
</tr>
<tr>
<td>3.7</td>
<td>T7 Endonuclease I Assay</td>
<td>56</td>
</tr>
<tr>
<td>3.8</td>
<td>Restriction Enzyme Digestions</td>
<td>58</td>
</tr>
<tr>
<td>3.9</td>
<td>Lentiviral particle transduction</td>
<td>59</td>
</tr>
<tr>
<td>3.10</td>
<td>Lentiviral Infection</td>
<td>60</td>
</tr>
<tr>
<td>3.11</td>
<td>TA Cloning</td>
<td>60</td>
</tr>
<tr>
<td>3.12</td>
<td>RNA Extraction</td>
<td>61</td>
</tr>
<tr>
<td>3.13</td>
<td>MiRNA quantification</td>
<td>62</td>
</tr>
<tr>
<td>3.14</td>
<td>MiRNA Overexpression</td>
<td>62</td>
</tr>
<tr>
<td>3.15</td>
<td>Gene expression</td>
<td>63</td>
</tr>
<tr>
<td>3.16</td>
<td>Computational analysis</td>
<td>63</td>
</tr>
<tr>
<td>3.17</td>
<td>pLKO.1 cloning</td>
<td>64</td>
</tr>
<tr>
<td>3.18</td>
<td>RNA Pull-down assay</td>
<td>66</td>
</tr>
<tr>
<td>3.19</td>
<td>In-gel digestion and HPLC-MS/MS analysis</td>
<td>69</td>
</tr>
<tr>
<td>3.20</td>
<td>In-solution digestion and HPLC-MS/MS analysis</td>
<td>70</td>
</tr>
<tr>
<td>3.21</td>
<td>Software analysis of proteomic data</td>
<td>71</td>
</tr>
<tr>
<td>3.22</td>
<td>Immunoblotting</td>
<td>72</td>
</tr>
<tr>
<td>3.23</td>
<td>Statistical analysis</td>
<td>73</td>
</tr>
</tbody>
</table>

Chapter 4 EVALUATION OF THE DELIVERY METHOD FOR CRISPR/Cas9 AND ITS EFFICIENCY IN EDITING THE TARGETED SITES ................................. 74

<table>
<thead>
<tr>
<th>Chapter 4</th>
<th>Section</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1</td>
<td>Introduction</td>
<td>74</td>
</tr>
</tbody>
</table>
4.2 Experimental Design

4.3 Results

4.3.1 Gene Editing in Human Cells

4.3.2 Gene Editing in Mouse Cells

4.4 Discussion

4.4.1 Gene Editing in Human Cells

4.4.2 Gene Editing in Mouse Cells

Chapter 5 INVESTIGATION OF THE REGULATION AND FUNCTION OF CLUSTERED miRNAs WITH CRISPR/Cas9 EDITING

5.1 Introduction

5.2 Experimental Design

5.3 Results

5.3.1 Regulation of mmu-miR-195~497 cluster in VSMCs

5.3.2 Gene editing in different miRNA clusters

5.3.3 Gene editing of mmu-miR-17~92 cluster

5.3.4 Gene editing of mmu-miR-106b~25 cluster

5.4 Discussion

5.4.1 Editing of mmu-miR-143~145 cluster

5.4.2 Editing of mmu-miR-17~92 cluster

5.4.3 Editing of mmu-miR-106~25

Chapter 6 ASSESSMENT OF THE LEVELS OF RBPS THAT BIND PRIMARY MMU-MIR-195A

6.1 Introduction
6.2 Experimental Design 161
6.3 Results 162
  6.3.1 Differences in expression of RBPs 162
  6.3.2 Proteomic Analysis of the RNA Pull-down proteins 166
6.4 Discussion 178
  6.4.1 Expression levels of RBPs 179
  6.4.2 Proteomic Analysis 179

Chapter 7 GENERAL DISCUSSION ......................................................... 189
  7.1 Conclusions 189
  7.2 Limitations 195
  7.3 Future Outlook 199

BIBLIOGRAPHY .................................................................................. 201

SUPPLEMENT .................................................................................... 218
# Table of Figures

| Figure 1.1 | MiRNA biogenesis. | 20 |
| Figure 1.2 | Primary miRNA transcript processing. | 21 |
| Figure 1.3 | Pre-miRNA processing by Dicer. | 22 |
| Figure 1.4 | MiR-15 family. | 23 |
| Figure 1.5 | Angiogram of AAA. | 26 |
| Figure 1.6 | Hetero-cluster miR-143~145 controls VSMC differentiation. | 32 |
| Figure 1.7 | Schematic representation of miR-17~92 genomic locus cluster along with the genomic loci of its paralogue clusters. | 34 |
| Figure 1.8 | Expression levels of 5 members of the miR-15 family upon inhibition mmu-miR-195a. | 37 |
| Figure 1.9 | Classification scheme for Cas proteins. Two classes, and five types (the best described ones) of Cas proteins are presented along with | 40 |
| Figure 1.10 | CRISPR genomic locus. | 41 |
| Figure 1.11 | Precursor crRNA maturation. | 41 |
| Figure 1.12 | CRISPR/Cas9 in nature. | 42 |
| Figure 1.13 | CRISPR/Cas9 with sgRNA. | 43 |
| Figure 1.14 | DSB repair in eukaryotic cells. | 43 |
| Figure 3.1 | Vectors pX330 and LentiCRISPRv2. | 51 |
| Figure 3.2 | PCR assembly of the gRNA DNA template using the Target F1 and Target R1 oligonucleotides. | 53 |
| Figure 3.3 | T7 Endonuclease I digestion. | 57 |
| Figure 3.4 | SspI, Mwol, Hinfl, Bsp1286I and MefI restriction sites. | 59 |
| Figure 3.5 | pGEM®-T vector. | 61 |
| Figure 3.6 | Vector pLKO.1-puro cloning vector. | 64 |
| Figure 4.1 | In silico design of the single-guide RNA for hsa-miR-195 locus. | 79 |
Figure 4.2 Schematic representation of the editing of the hsa-miR-195 genomic locus. ................................................................. 80

Figure 4.3 Gene editing of the miR-195 genomic locus using plasmid vectors. .................................................................................... 82

Figure 4.4 T7 Endonuclease I assay for HUVECs upon lentiviral infection. ...................................................................................... 84

Figure 4.5 Lentiviral delivery of Cas9 and sgRNA ............................................................................................................................... 85

Figure 4.6 Schematic representation of the digestion products upon editing with guides sg195g1 and sg195g2. .............................................. 86

Figure 4.7 T7 Endonuclease I assay for hSMCs upon lentiviral infection. ...................................................................................... 87

Figure 4.8 Assessment of Cas9/RNPs approach in HEK293T cells. ...................................................................................... 88

Figure 4.9 Assessment of possible off-target mutagenesis upon lentiviral infection of HUVECs ........................................................ 89

Figure 4.10 In silico design of the single-guide RNA for the mmu-miR-195a locus ................................................................. 91

Figure 4.11 Schematic representation of the editing of the mmu-miR-195 genomic locus. ................................................................. 92

Figure 4.12 Schematic representation of the secondary structure of the mmu-miR-195a stem loop with guides. ........................................ 93

Figure 4.13 T7 Endonuclease I assay for mouse SMCs upon lentiviral infection with guides Gm2 and Gm4. ................................................................. 94

Figure 4.14 Schematic representation of the sizes of T7EI digestion products upon editing with guide Gm4 and Gm2. ........................................ 95

Figure 4.15 T7 Endonuclease I assay for mouse VSMCs upon lentiviral infection with guides Gm2 and GmA. ................................................................. 96

Figure 4.16 Schematic representation of the T7EI digestion products upon editing with guide GmA and Gm2. ................................................................. 97

Figure 4.17 Cas9 expression in VSMCs infected with LentiCRISPRv2 lentiviral particles. ................................................................. 98
Figure 4.18 Assessment of gene editing in mouse VSMCs stably expressing Cas9 ................................................................. 99
Figure 4.19. T7EI assay for mouse VSMCs stably expressing Cas9 after transfection with IVT guides ........................................... 99
Figure 4.20 Restriction enzyme digestion with SsP1 and MwoI ............ 101
Figure 4.21 Assessment of possible off-target mutagenesis upon lentiviral infection of mouse SMCs ............................................. 102
Figure 4.22 Assessment of possible off-target mutagenesis upon editing with IVT Gm2 .............................................................. 103
Figure 4.23. A comprehensive summary of the experimental approaches and their effects ............................................................... 105
Figure 5.1 Gene editing of the miR-195a locus using IVT sgRNA195m2 .... 117
Figure 5.2 Investigation of editing in the miR-497a locus with sg195m2. .... 118
Figure 5.3 Gene editing of the mmu-miR-195a locus using sgRNA195m3 .... 120
Figure 5.4 Assessment of off-target editing with sg195m3. ...................... 121
Figure 5.5 Investigation of editing in the miR-497a locus with sg195m3. ..... 122
Figure 5.6. Overexpression of miR-195a ........................................ 123
Figure 5.7 Predicted secondary structure of the mmu-miR-195a stem loop upon editing ......................................................... 124
Figure 5.8 Effect of CRISPR/Cas9 editing in the tertiary structure of mmu-miR-497~195 cluster ............................................................... 126
Figure 5.9 In silico design of the sgRNAs for mmu-miR-145a locus ............ 128
Figure 5.10 Schematic representation of the secondary structure of the mmu-miR-145a stem loop with guides ........................................ 129
Figure 5.11 Gene editing of the mmu-miR-145a locus .......................... 130
Figure 5.12 Assessment of the levels of mmu-miR-143 and mmu-miR-145a upon editing .............................................................. 131
Figure 5.13 Off-target editing in miR-143 locus .................................... 132
Figure 5.14 Assessment of the levels of pri-miR-143 and pri-miR-145a. 

Figure 5.15 Predicted secondary structure of mmu-miR-145a stem loop.

Figure 5.16 Expression of long non-coding RNA Carmn.

Figure 5.17 In silico design of the single-guide RNAs for mmu-miR-18 locus.

Figure 5.18 Schematic representation of the secondary structure of the miR-18a stem loop with guides.

Figure 5.19 Gene editing of the miR-18a locus.

Figure 5.20 Off-target editing in mmu-miR-17~92 locus.

Figure 5.21 Predicted secondary structure of miR-18a stem loop.

Figure 5.22 In silico design of the single-guide RNAs for mmu-miR-25 locus.

Figure 5.23 Schematic representation of the secondary structure of the mmu-miR-25 stem loop with the designed guide.

Figure 5.24 Gene editing of the miR-25 locus.

Figure 5.25 Off-target editing in mmu-miR-106b~25 locus.

Figure 5.26 Predicted secondary structure of mmu-miR-25 stem loop.

Figure 6.1 Assessment of the expression levels of a panel of RBPs.

Figure 6.2. Western blot analysis for HuR.

Figure 6.3. Titration experiment for the RNA template of the Pull-down.

Figure 6.4. Titration experiment for the amount of labelled RNA to be used for the Pull-down.

Figure 6.5. Western blot analysis for Dicer.

Figure 6.6. Distribution of RNA pulled down proteins.

Figure 6.7. Chart for the localization of the RNA pulled down proteins.

Figure 6.8 Number of RBDs in the pulled down proteins.

Figure 6.9. Volcano plot analysis for the RNA pulled down proteins that were differentially bound on the two transcripts.
Figure 6.10. Clusters of proteins that bind exclusively on the mutant transcript. .......................................................................................................................... 170

Figure 6.11. Analysis of the predicted protein-protein interactions for the protein Eri1. ........................................................................................................................................... 171

Figure 6.12. Analysis of the predicted protein-protein interactions for the protein Luc7I3.............................................................................................................................................. 172

Figure 6.13. Analysis of the predicted protein-protein interactions for the protein Tra2b. ............................................................................................................................................... 173

Figure 6.14. Analysis of the predicted protein-protein interactions for the cluster of proteins that bind only the mutant RNA transcript. ........................................................................................................... 174

Figure 6.15. Analysis of the molecular action for the predicted protein-protein interactions for the cluster of proteins that bind only the mutant RNA transcript. ........................................................................................................... 175

Figure 6.16. Analysis of the confidence for the predicted protein-protein interactions for the cluster of proteins that bind only the mutant RNA transcript. ........................................................................................................... 176

Figure 6.17. Cluster analysis of the identified proteins. .......................................................................................................................... 177
Table of Tables

Table 1. List of Lentiviruses used. ................................................................. 60

Table 2. Primary antibodies used for immunoblotting. .............................. 73
Acknowledgments

First and foremost, I would like to thank my first supervisor, Dr Anna Zampetaki, for her help during my technical training and for explaining the protocols employed in this project. Also, I would like to thank her for her great support and guidance throughout the project. Her technical expertise, knowledge and continuous patience was of high importance for the completion of the project.

I would also like to thank my second supervisor Prof. Manuel Mayr for his support, technical advice and significant input in my project.

My thanks also to all the members of the Cardiovascular Proteomics Lab, past and present; Dr Ruifang Lu, Dr Ursula Mayr, Dr Xiaoke Yin, Dr Javier Barallobre-Barreiro, Dr Marika Fava, Dr Gonca Suna, Ms Ferheen Baig, Dr Temo Barwari, Mr Eloi Haudebourg, Mr Marc Lynch, Mr Joshua Goulding, Dr Adam Fellows, Dr Konstantinos Theofilatos, Dr Marieke Rienks, Mr Lukas Schmidt, Mrs Timnit Tekie, Mr Sean Burnap, Dr Abishek Joshi, Dr Christian Schulte, Dr Elisa Duregotti, Mr Christian Cassel, Mr James Campbell and Miss Ella Louise Reed for their help in my lab work and support throughout my PhD.

Finally, I would like to thank my family for their encouragement to pursue a PhD and their support to its completion.
Declaration of independent work

I, the author of this thesis, declare that the work presented here was conducted solely by me.

.................................................................

Lazaros Lataniotis
Abbreviations

2D   two-dimensional
3D   three-dimensional
AAA  Abdominal Aortic Aneurysm
AAV  Adenoassociated Virus
ACN  Acetonitrile
ADAR1 Adenosine deaminases acting on RNA 1
AGO2 Argonaute 2
Apex1 apurinic/apyrimidinic (AP) endodeoxyribonuclease 1
a-SMA alpha smooth muscle actin (Acta2)
ATP  Adenosine Triphosphate
BIM  Bcl-2-like protein 11
BMP  bone morphogenetic protein
BMPR2 Bone morphogenetic protein receptor type II
bp   base pair
BSA  Bovine Serum Albumin
CAD  Coronary Artery Disease
Carmn cardiac mesoderm enhancer-associated non-coding RNA
Cas  CRISPR Associated
Cas9 CRISPR Associated 9
CRISPR Clustered Regular Interspaced Short Palindromic Repeat
crRNA CRISPR RNA
dCas9 Nuclease dead Cas9
DGCR8 Di George syndrome Critical Region 8
DMEM Dulbecco's Modified Eagle Medium
DNA Deoxyribonucleic Acid
dNTP deoxynucleotide triphosphate
ds double stranded
DSB Double Stranded Break
DSRM dsRNA binding motif
DTT  Dithiothreitol
dTTP deoxyThymidine Triphosphate
E, Coli Escherichia Coli
E2F1 E2F Transcription Factor 1
E2F2 E2F Transcription Factor 2
E2F3 E2F Transcription Factor 3
ECM Extracellular Matrix
EDTA ethylenediaminetetraacetic acid
EFS EF-1 Alpha Short
ELISA enzyme-linked immunosorbent assay
ER  Endoplasmatic Reticulum
FASTA FAST-ALL
FBS Foetal Bovine Serum
FDR False Discovery Rate
GTP Guanosine Triphosphate
HCD Higher energy collision dissociation
PEG  Polyethylene Glycol  
PEI  Polyethyleneimine  
PH  Pulmonary Hypertension  
P13K  phosphoinositide 3-kinase  
PIWI  P-element Induced Wlmyp  
pre-crRNA  premature CRISPR RNA  
pre-miRNA  precursor microRNA  
pri-miRNA  primary microRNA  
PTEN  Phosphatase and tensin homolog  
qPCR  quantitative real-time polymerase chain reaction  
Ran  RAs-related Nuclear protein  
RB  Retinoblastoma  
RBM6  RNA Binding Motif Protein 6  
RBPs  RNA Binding Proteins  
RISC  RNA Induced Silencing Complex  
RnaseIII  Ribonuclease III  
RNP  Ribonucleoprotein  
ROS  Reactive Oxygen Species  
rpm  revolutions per minute  
RRM  RNA Recognition Motif  
R-SMAD  Receptor SMAD  
RT  Reverse Transcription  
S  Synthetic  
SD  Standard Deviation  
SDS  Sodium Dodecyl Sulphate  
sgRNA  single guide RNA  
shRNA  short hairpin RNA  
SM22  Smooth Muscle 22  
SMAD  Sma genes and the Drosophila Mad, Mothers against decapentaplegic  
SMCs  Smooth Muscle Cells  
SNP  Single Nucleotide Polymorphism  
SpCas9  Streptococcus pyogenes Cas9  
ss  single stranded  
STRING  Search Tool for the Retrieval of Interacting Genes/Proteins  
T7EI  T7 Endonuclease I  
TAL  transcription activator-like  
TALENs  Transcription Activator-Like Effector Nucleases  
TBE  Tris/Borate/EDTA  
TbRI  transforming growth factor beta receptor I  
TbRII  transforming growth factor beta receptor II  
TbRIII  transforming growth factor beta receptor III  
TBS  Tris-buffered saline  
TBS-T  Tris-buffered saline-Tween  
TDP43  TAR DNA-binding protein 43  
TEAB  triethylammonium bicarbonate  
TFA  Trifluoroacetic acid  
TGF  Transform Growth Factor  
TGF-β  Transform Growth Factor-β
<table>
<thead>
<tr>
<th>Term</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>TGF-β1</td>
<td>Transforming Growth Factor beta-1</td>
</tr>
<tr>
<td>TGFBRII</td>
<td>Transform Growth Factor-β</td>
</tr>
<tr>
<td>TIMP</td>
<td>Tissue Inhibitors of Metalloproteinases</td>
</tr>
<tr>
<td>TLR2</td>
<td>Toll-like receptor 2</td>
</tr>
<tr>
<td>Tra2a</td>
<td>Transformer 2 Alpha Homolog</td>
</tr>
<tr>
<td>tracrRNA</td>
<td>Trans-activating crRNA</td>
</tr>
<tr>
<td>TRBP</td>
<td>TAR RNA binding protein</td>
</tr>
<tr>
<td>UTR</td>
<td>Untranslated Region</td>
</tr>
<tr>
<td>UV</td>
<td>Ultra Violet</td>
</tr>
<tr>
<td>VSMCs</td>
<td>Vascular Smooth Muscle Cells</td>
</tr>
<tr>
<td>WGS</td>
<td>Whole Genome Sequencing</td>
</tr>
<tr>
<td>WT</td>
<td>Wild Type</td>
</tr>
<tr>
<td>Zc3h10</td>
<td>Zinc finger ccch-type containing 10</td>
</tr>
<tr>
<td>ZFN</td>
<td>Zinc Finger Nucleases</td>
</tr>
<tr>
<td>α</td>
<td>alpha</td>
</tr>
<tr>
<td>β</td>
<td>beta</td>
</tr>
<tr>
<td>Δ</td>
<td>difference</td>
</tr>
</tbody>
</table>
Chapter 1 INTRODUCTION

1.1 MicroRNAs (miRNAs)

1.1.1 MiRNA biogenesis

MicroRNAs (miRNAs) are short, 19-24 nucleotide long, non-coding RNA molecules that regulate gene expression post-transcriptionally by binding target transcripts in animals through partial complementarity\(^1\). They represent the major subgroup of small ncRNAs, with more than 2000 miRNAs identified in the human genome\(^2\). In addition, they are evolutionary conserved and display distinct tissue expression profile\(^3\).

1.1.1.1 MiRNA transcription and processing

MiRNA genes are transcribed by RNA Polymerase II as long precursor molecules, called primary miRNAs (pri-miRNAs), which can encode single or multiple miRNAs, forming miRNA clusters\(^4\), (Figure 1.1). The pri-miRNAs consist of one or more hairpin-like structures, ~100nt long, which consist of a long stem with bulges and a terminal loop, (Figure 1.2). In a first step, they are processed by RNase III Drosha and its co-factor, DGCR8, which form the microprocessor complex\(^5\), to release a ~65nt long hairpin-shaped precursor (pre-miRNA)\(^6\). It has been previously shown that a very specific and precise recognition of the primary transcript is needed in order for the Drosha-DGCR8 complex to recognise and cleave the pri-miRNAs. More specifically, basal elements are used as rulers by Drosha while DGCR8 ensures fidelity of the processing by interacting with apical elements\(^7, 8\). Thus it becomes apparent that sequence changes, like the abolition of specific sequence motifs, such as a UG motif at the base of the hairpin, a UGU/GUG motif in the apical loop and a CNNC motif downstream of the hairpin, as well as structural alterations in the secondary structure of the hairpin of the pri-miRNA, such as a shorter terminal loop, the size and the position of the bulge, the apical loop size, and especially the single stranded state of the basal segments, influence the Drosha processing efficiency\(^7, 11\).
The pre-miRNA is, then, exported to the cytoplasm with the help of the Exportin5 and is further processed by another RNase III protein, Dicer, (Figure 1.1). Drosha cleavage generates a two-nucleotide-long overhang on the 3′-end of the pre-miRNA that is bound by Dicer. More specifically, the two cleavage sites of Dicer bind to the 3′-end of dsRNAs at a fixed distance, of 21–25 nucleotides in length (3′-counting rule), that is dependable on the type of Dicer. In mammals, an additional mechanism is employed by Dicer to determine the cleavage site of pre-miRNA; it binds to the non-stable 5′-phosphorylated end of the pre-miRNA and cleaves it, at a distance of 22nt away from the 5′-end (5′-counting rule) while a strongly paired 5′-end, through G-C base pairing, impairs its binding and further processing. Human Dicer shows two basic pockets in its PAZ domain that bind to the 5′-end and 3′-end of the pre-miRNA and are spatially arranged to occupy simultaneously the 5′-end and 3′-end of the pre-miRNA only in the case that there is a two-nucleotide 3′-overhang, (Figure 1.3).

Figure 1.1 MiRNA biogenesis. MiRNAs mature through a multistep process. They are transcribed into pri-miRNAs and processed to pre-miRNAs by Drosha. The pre-miRNA is exported to the cytoplasm, where it is further cleaved to the miRNA:miRNA* duplex by Dicer. Both strands can be loaded to the RISC complex for silencing the mRNA. From Winter et al., 2009.
Figure 1.2. Primary miRNA transcript processing. Drosha and DGCR8, the microprocessor complex, recognizes the single-stranded parts of the transcript, the length of the stem (~35 bp) and the terminal loop of the primary microRNA. Microprocessor measures the distances from the basal junction and the apical loop, ~11 bp and ~22 bp respectively, and Drosha cleaves the primary transcript. From Ha M. et al., 2014.
Dicer’s processing results in a short (19-24 nucleotide) miRNA:miRNA* duplex. In most cases, the miRNA strand is selected as the mature miRNA, that then binds the Argonaute 2 (Ago2) protein of the RNA induced Silencing Complex (RISC), while the miRNA* strand is degraded. Loaded into the RISC, miRNAs function as sequence-specific guides that direct the complex to mRNA transcripts, targeting mostly the 3’ untranslated region (UTR) but, in some cases, also the 5’-UTR and the coding region. Although the mechanism of miRNA:mRNA interaction is not well understood, in mammalian cells this it occurs through incomplete complementarity.

Nucleotides 2-8 at the 5’-end of the miRNA constitute its ‘seed region’ that plays a critical role in the interaction of miRNA:mRNA. Binding of the miRNA to the miRNA Response Elements (MRE) on its targets negatively regulates their expression by preventing their translation and triggering transcript degradation mainly through de-adenylation. Consequently, miRNAs induce post transcriptional gene silencing by either the translational repression of the mRNA, its de-adenylation or its degradation, process that is influenced by the percentage of complementarity between the miRNA and its target (Figure 1.1).
Because of the short length of the miRNA seed sequence and the partial complementarity that is needed for a miRNA to recognise and bind its corresponding targets, single miRNAs can regulate more than one gene. This partial complementarity enables miRNAs to target multiple transcripts that, in many cases, are effectors of the same signalling pathway thus creating a complex regulatory network and eliciting system wide biological responses. Thus, miRNAs with similar seed regions may exert control on set of genes with similar, but not identical sequence to differing degrees.

Moreover, miRNAs with similar sequences form families that are comprised of multiple members that share a common seed region and often display high homology in the rest of their sequence (Figure 1.4). This means that a set of miRNAs can suppress the same gene or set of genes. Additionally, the fact that miRNAs are organised in families imply important roles for them in regulating the expression levels of their mRNA targets with their specific expression further supplementing a flexibility in regulating. Furthermore, control on the expression of genes is partial and a single mRNA can have multiple binding sites for miRNAs.

![Figure 1.4. MiR-15 family](image)

Overall, miRNAs seem to be an evolutionary tool developed to absorb perturbations, reinforce the robustness of biological systems and ensure that gene expression occurs both at desirable levels and with appropriate timing. MiRNAs are largely considered to be the fine tuners of gene expression that exert mild effects under basal conditions and have more pronounced responses after stress. Intriguingly, several
miRNAs may target the same transcript forming a co-targeting network and implying redundancy in miRNA function\textsuperscript{26}.

Through regulating the expression of various effectors that function at different steps in complex biological pathways, miRNAs can coordinate important cellular processes such as cell differentiation, metabolism, and apoptosis\textsuperscript{23}. Thus, miRNAs are able, by combining mild effects on multiple individual mRNA targets, to act synergistically within a common pathway in order to exert a more pronounced outcome\textsuperscript{26}. Because miRNAs contribute to the gene expression in various burdensome diseases like cancer and cardiovascular diseases\textsuperscript{1}, they have emerged as an important field of study. Nonetheless, taking into consideration the complexity they present, studying the functions or manipulating their effects presents great challenges\textsuperscript{27}.

1.1.1.2 RNA Binding Proteins (RBPs) for the processing of miRNAs

RBPs are proteins that are capable of binding double or single stranded RNA in cells and participate in forming ribonucleoprotein complexes that are involved in gene expression\textsuperscript{28}. They participate in the process of the post transcriptional gene regulation of coding and non-coding RNAs involving the transport, maturation and stability of the transcripts. RBPs bind RNAs using sequence and/or structural motifs in RNA molecules via using RNA-binding domains (RBDs) such as the RNA recognition motif (RRM)\textsuperscript{29}, K homology (KH) domain\textsuperscript{30}, DEAD box helicase domain\textsuperscript{31}, double-stranded RNA-binding motif (DSRM) or a zinc-finger domain\textsuperscript{32}. Different mRNA-binding domain classes of RBPs are found in high frequency either in multiple repeats or in combination with other RBDs\textsuperscript{33}. The most prominent example is that of ssRBD such as RPMs, KH domains and zinc-finger domains, that recognise motifs of 4-6 nucleotides and work in combinations\textsuperscript{34}. Nonetheless, recent findings in respect of the structure and the function of the RBPs through the study of RNP machines like the spliceosome\textsuperscript{35,36,37} and the ribosome\textsuperscript{38,39,40}, reveal the complexity of unconventional protein-RNA interactions that do not necessitate canonical RBDs.

RBPs exert effects on the miRNA biogenesis by interacting with Drosha, Dicer and specific miRNAs. RNase III proteins often interact with dsRBD proteins, such as Dcr-1 in \textit{Drosophila spp.} that binds one of the two isoforms of Loquacious (Loqs), Loqs-
PA and Loqs-PB, necessary for the production of most miRNAs, each of which contains three dsRBDs\(^{41,42}\). Furthermore, a set of proteins, such as p68 and p72 are essential for the processing of a subset of miRNAs by Drosha\(^{43}\) while proteins R-SMA\(\delta\), SMAD1–3, SMAD5 and p53 modulate the activity of the Microprocessor complex through interaction with p68\(^{44,45,46}\). TDP43 RBP was reported to increase Drosha’s stability\(^{47}\) and to promote Drosha and Dicer processing\(^{48}\). It has also been demonstrated that human Dicer interacts with TAR RNA-binding protein (TRBP), a homologue of Loqs-PB, in a similar way\(^{13,49}\) while cofactor PACT, that harbours a dsRBD, has also been shown to associate with mammalian Dicer\(^{50}\). Similarly, TRBP regulates the efficiency of processing of certain pre-miRNAs and regulates mature miRNA length\(^{51}\). Nonetheless, contrary to Loqs, mammalian TRBP or PACT are not essential for the Dicer-mediated processing of the pre-miRNA processing\(^{49}\).

Moreover, various proteins have been reported to selectively bind the terminal loop of pri-miRNA. RBPs heterogeneous nuclear ribonucleoprotein A1 (HNRNPA1) and KH-type splicing regulatory protein (KSRP) facilitate the processing by Drosha via binding to the terminal loop of pri-miR-18a and pri-let-7\(^{52,53}\). Moreover, pri-let-7 is also regulated by LIN28A and its paralogue LIN28B, that bind its terminal loop and arrests both Drosha and Dicer-mediated processing.

Additionally, negative feedback loops between miRNAs and RBPs that regulate their expression levels have been reported. Human DICER1 mRNA is targeted by let-7 miRNA\(^{54}\). RBPs partake in the regulation of the processing of pre-miRNA. KSRP, interacts with the terminal loop of various pre-miRNAs and modulates their maturation process through affecting the Dicer-mediated processing\(^{55}\) while LIN28 proteins, through binding to the terminal loop of pre-let-7, regulates the processing by Dicer by inducing oligouridylation of the pre-miRNA\(^{56}\). Additionally, modifications in the sequence of the stem loop of miRNA, through RNA editing with ADAR1, have demonstrated diminished maturation either at the processing by Drosha, as in the example of pri-miR-142\(^{57}\), or by Dicer, as in the example of pre-miR-151\(^{58}\).
1.2 Abdominal Aortic Aneurysm (AAA)

AAA is defined as a focal dilation of the wall of the abdominal aorta due to the loss of its elasticity, exceeding 50% of its normal size\textsuperscript{58}. It corresponds to a region of the abdominal aorta that has weakened wall and is dilated usually affecting the infra-renal segment\textsuperscript{60}, (Figure 1.5). Aneurysmal disease is more common in the elderly population (>65 years old) with its prevalence in western populations being 0.4–0.67%\textsuperscript{61}. Although mortality rate of AAA is improving because of the better intervention techniques\textsuperscript{62}, it still is a significant cause of death\textsuperscript{63,64}.

![Image of AAA angiogram](image)

**Figure 1.5. Angiogram of AAA.** 3D MR angiogram with horizontal line demonstrating the location of the infrarenal area that present AAA. The distance from the left renal artery is showcased with a double-headed arrow. From Kramer et al., 2007.

Risk factors for development of AAA include advanced age (>65 years old), gender, with males presenting higher prevalence, genetic predilection via family history, obesity and smoking\textsuperscript{66,67}.

AAA is an asymptomatic disease\textsuperscript{68}, a characteristic that renders it more dangerous since patients are unaware of their condition. A major implication of the disease is the
acute rupture of the aneurysm, which leads to heavy bleeding into the abdomen that can be fatal and is associated with up to 80% mortality\textsuperscript{61}. Given the burden of AAA and the lack of symptoms, that makes the diagnosis more difficult, there is a high interest for the discovery of new biomarkers that could indicate AAA development and rupture.

1.2.1 AAA pathogenesis

Although aneurysm formation is a complex process multiple histological features, important for the development of the disease, have been described such as vascular inflammatory processes, extracellular matrix (ECM) degradation and vascular smooth muscle cells (VSMCs) apoptosis\textsuperscript{69}. In more detail, chronic infiltration of inflammatory cells in the adventitial and medial layers of the aorta, elastin degeneration and medial attenuation have been reported to participate in the formation of the disease\textsuperscript{69}. These inflammatory stimuli together with mechanical injury that lead to phenotypic changes affect the proliferation, migration and apoptosis of the cells\textsuperscript{70}. The infiltration of macrophages, leukocytes and lymphocytes\textsuperscript{71}, that lead to a higher production of cytokines under the effects of which the inflammatory response progresses further, results in neovascularisation, loss of SMCs and matrix degradation\textsuperscript{72,73,74}. Further to this, ECM degradation seems to advance due to locally increased levels of matrix metalloproteinases (MMPs), cysteine and serine proteases which result in different degradation among collagen isoforms (type I and III) and the degradation of elastin which affects the elasticity of the aorta\textsuperscript{75}.

Moreover, both T-cells and B lymphocytes have been shown to be abundant in aneurysmal tissues and therefore they have been implicated in the progression of AAA process while depletion of B-cell has protective effects against the disease\textsuperscript{71}. Additionally, regulatory T-cells have a protective effect against the formation of AAA with the secretion of IL-10 and TGF-β being implicated by stabilising the aneurysm\textsuperscript{73} while CD4+ T-cells have been linked to the pathogenesis of AAA\textsuperscript{71}.
1.2.2 SMCs in health and in AAA disease

1.2.2.1 SMCs function and biology

VSMCs have been shown to play a major role in the pathogenesis of many disorders such as the aortic aneurysm formation, as they are key components of blood vessels\(^7^6\). They regulate the diameter of small arteries, controlling blood pressure and they are the main cells that constitute the aortic wall able to synthesise ECM proteins, participate in the regulation of proteases and protease inhibitors and recruit inflammatory cells upon cell injury. VSMCs do not proliferate at high rate, under physiological conditions, unless they undergo a phenotypic switch, from contractile to proliferating phenotype, as a response to stimuli like vascular injury and hypertension\(^7^7\). Additionally, the ability of VSMCs to contract and proliferate is affected by mechanical forces, ROS, ECM components and the TGF-β and SMAD3 signalling among others, changes that occur during aneurysmal formation\(^7^8,7^9,8^0\). In light of this, the pathophysiology of aortic aneurysms is associated with VSMC apoptosis and ECM proteins and elastin degradation, features that distinguish it from atherosclerosis albeit the common risk factors\(^8^1\). However, it is not clear how VSMCs participate in the formation of AAA. They can contribute to the development of the disease through inflammation and production of MMPs\(^8^2\), while a third pathway of differential production of ECM has been suggested. It has been reported in mouse models that aneurysm formation can be reversed through prevention of VSMC apoptosis with caspase inhibitor treatment\(^8^3\), data that support the important role of VSMCs in aneurysm pathogenesis. It, therefore, becomes clear that VSMCs can exert important control on the homeostasis of the aortic ECM.

Although VSMCs partake in the onset and development of aortic aneurysms through loss of synthetic capability, due to cell apoptosis or other mechanisms\(^8^4\), there is growing evidence that aortic VSMCs are also capable of participating in the degenerative process\(^8^5\). Degradation of the ECM of the aortic wall by elastases has been considered to be predominantly due to the infiltration of inflammatory cells\(^8^6\), namely macrophages\(^8^4\) as well as neutrophils\(^8^7\) and mast cells\(^8^8\), as it has been shown by studies in murine models\(^8^9\). However, it has also been reported that apoptosis of aortic VSMCs may contribute to the pathogenesis of AAA via reducing the repair rate of the damaged aortic wall\(^9^0\) while AAA-derived cells demonstrated increased levels of proteases and cysteine classes that participate in the elastolysis\(^8^2\).
More specifically, a lack of upregulation of Tissue Inhibitor of Metalloproteinase (TIMP) -1 and -2 was reported while there was considerable upregulation of TIMP-3 whilst there is an increased MMP activity data that support an imbalanced proteolytic state. These findings suggest that VSMCs can also participate in the formation of AAA through degradation of ECM via elastolytic activity. In addition, evaluation of interactions of VSMC with macrophages demonstrated a significant increase in the elastolytic activity of the AAA-derived cells. Further studies showed that in the case of AAA, the protein levels of MMP-9 and MMP-2 protein levels are substantially increased compared to non-AAA cells. Elevated levels of MMP-2 and MMP-9 activity have been reported in the aneurysmal part of the mouse aorta after AngII infusion. MMP-9 is the most abundant gelatinolytic MMP and is secreted in high levels in AAA tissue with MMP-9-deficient mice demonstrating normal levels of the elastic lamellae in the aortic wall and, interestingly, resistance to the development of aneurysm. These data suggest a posttranscriptional regulation in normal cells that is abrogated in the context of AAA. A meta-analysis on serum and plasma levels has implicated higher levels of circulating MMP-9 to AAA while the macrophage origin of the MMP-9 was shown when mice were infused with non-MMP9-deficient macrophages, reversing the protection effects on AAA. Taken together, these data demonstrate that VSMCs, in the case of AAA, present a pro-elastolytic phenotype that is further pronounced in the presence of activated macrophages. This appears to be the effect of differentiated post-transcriptional control of MMP-9 synthesis resulting in activation of elastolytic MMPs.

1.2.2.2 MiRNAs and AAA

The maintenance of the physiological function of the cardiovascular system is of high importance for mammals while miRNA networks orchestrate cellular homeostasis, vascular inflammation, angiogenesis and metabolism and have been implicated in the control of important biological processes in the cardiovascular system and in many aspects of the development and pathophysiology of cardiovascular diseases. A plethora of miRNAs, from the same or different families, have been implicated in various cardiovascular diseases, a feature that demonstrates the significant role the miRNA networks play in the physiology of the heart and blood vessels as well as the sensitivity they both demonstrate to disease. Studies have reported
perturbations in the levels of a vast number of miRNAs in both humans and mice models for many cardiovascular diseases such as heart failure, myocardial infarction, atherosclerosis and AAA\textsuperscript{26, 98}.

MiRNAs have also been strongly implicated in the formation and development of the AAA. Gain- and loss-of-function assays have demonstrated a direct link of specific miRNAs with AAA and highlighted the potential they have as novel targets for therapeutic applications\textsuperscript{99, 100, 101, 102}. Accumulating data, demonstrate that they are important regulators of the function of vascular cell homeostasis in atherosclerotic and inflammatory processes. As an example of the role of miRNAs in the function of SMCs, postnatal Dicer deletion led to loss of contractile differentiation for VSMCs and to lower blood pressure\textsuperscript{103}. A variety of miRNAs have been connected with remodelling of vascular cells including miR-126, miR-10a, miR-143~145 and the miR-15/16 families\textsuperscript{26, 104}.

In more detail, in the vasculature, inhibition of miR-29 has been correlated with the reduction of the aortic aneurysm formation\textsuperscript{105}. Additionally, members of the miR-15 family have been implicated in collagen remodelling and the postnatal silencing of elastin in aortas\textsuperscript{101, 106}. Intriguingly, miR-195, a member of the miR-15 family, is known to share many of the same targets as miR-29b\textsuperscript{101} with the seed sequence of miR-15 miRNAs (5’-AGCAGC-3’) differing only by one nucleotide from that of the miR-29 family miRNAs (5’-AGCAC-3’). Proteomic analysis of the secretome of murine aortic SMCs revealed that miR-195 targets a group of ECM proteins, including elastin, collagens and proteoglycans. In mice treated with antagomiR-195, higher aortic elastin expression was associated with an increase of MMP-2 and MMP-9 while in human plasma the presence of AAA and aortic diameter was inversely correlated to the levels of miR-195\textsuperscript{98}.

1.2.2.3 MiR-195 and AAA

MiR-195 is a member of the miR-15 family and is located on chromosome 17 in humans and chromosome 11 in mice (Supplementary Table 1). It is transcribed together with miR-497 in one primary transcript, forming a homo-cluster since both miRNAs are members of the same miRNA family and they share the same seed sequence and high homology throughout their sequence (Figure 1.4). Although miR-
15 family and specifically miR-195 is not studied as intensely as other miRNAs, it is known to participate in various disorders like cardiovascular diseases and tumorigenesis\textsuperscript{98, 107}.

The ECM is an important component of the structure of aortic wall that determines the formation of aneurysm by the dilatation of the aortic wall\textsuperscript{108}. AAAs are characterised by decreased levels of elastin and high turnover of collagen. Various miRNAs have been implicated in collagen turnover and as a result also in the development of the disease, with the members of the miR-15 family participating in postnatal silencing of elastin\textsuperscript{101}. The murine miR-15 family consists of 6 highly conserved microRNAs (miR-15a, miR-15b, miR-16-1, miR-16-2, miR-195, and miR-497), which are clustered in 3 separate chromosomes. They have a common seed region (AGCAGCA) and varying degrees of sequence homology in the 3’ region of the mature microRNA\textsuperscript{109,98}. Expression of miR-195 and miR-497, miRNAs that are transcribed in the same primary transcript, is altered in human aortic specimens\textsuperscript{110} while various ECM transcripts have been shown to be directly targeted by miR-195 in H4 cancer\textsuperscript{111} results that were confirmed \textit{in vivo} in murine models where antagomiRs to miR-195 upregulated elastin levels\textsuperscript{98}.

Although miR-195 is not extensively studied in the context of AAA\textsuperscript{98}, data from preliminary studies suggest that downregulation of the miRNA lead to de-repression of MMP-2 while it is implicated in the downregulation of Smad3 levels by directly targeting it, thus leading to decreased proliferation of VSMCs, and increased expression of Collagen III\textsuperscript{112}. From the above, it becomes clear that miR-195 is strongly correlated with the onset and the development of the disease although further studies are needed to conclusively demonstrate the connection.

\subsection*{1.2.2.4 MiR-143~145 in AAA}

MiR-143 and miR-145 are two highly conserved miRNAs that form a hetero-cluster with their genes located on chromosome 5 in humans and chromosome 18 in mice (Supplementary Table 1). They are transcribed as one primary transcript although the two miRNAs do not share any sequence homology and belong to different miRNA families.
MiR-145 is highly expressed in VSMCs of healthy rat carotid arteries\textsuperscript{113} while expression of both miRNAs of the cluster are observed during the differentiation of stem cells to VSMCs\textsuperscript{114}. The miR-143~145 cluster has been implicated both in the differentiation and proliferation of the VSMCs since it has been reported that downregulation of their expression leads to higher proliferation of the cells with upregulation of their levels having the opposite effect\textsuperscript{114}. ELK1 and KLF4 are among the validated targets of miR-143 and miR-145, respectively, and both factors control the differentiation of the VSMCs. In more detail, ELK1 promotes the proliferation of the cells while KLF4 regulates the contractile phenotype of the cells. Targeting by the miR-143~145 cluster modulates the phenotypic switch\textsuperscript{114,115}.

![Diagram](image)

**Figure 1.6. Hetero-cluster miR-143~145 controls VSMC differentiation.** VSMC differentiation is regulated by Myocardin and myocardin-related transcription factors which also regulate the expression of miR-143 and miR-145 through a feedback loop. Inhibition of KLF4 miR-145 leads to SMC differentiation while targeting of ELK1 by miR-143 inhibits proliferation of SMCs. From Welten et al., 2016.

Studies in murine models where an IRES-lacZ promoter was fused to the pre-miR-143 sequence demonstrated that the expression of mmu-miR-143 and mmu-mR-145 in heterozygous mice was active in the heart in early developmental stages but at later stages was exclusively expressed in SMCs in the cardiovascular system, bladder and lungs, among other organs, while their expression levels were stable across adulthood\textsuperscript{78}. These data unveiled miR-143~145 as a cell-type specific cluster for the regulation of the SMCs.
Although data from AAA tissues showed that there was no difference in the levels of neither miR-143 nor miR-145 when compared to healthy tissue, studies on intracranial aneurysm and thoracic aorta aneurysm demonstrated a decrease in the levels of both miRNAs\textsuperscript{116,117,118}. Additionally, data from double knock-out mice revealed differences both in the structure and the phenotype of the VSMCs of the aorta as well as an increase in the migration and proliferation rate of the cells\textsuperscript{78,119}. More specifically, a significant decrease in the numbers of contractile VSMCs were observed by electron microscopy and a phenotypic switch to a synthetic phenotype was demonstrated together with the decrease in the size of the cells, alterations that contributed to a thinner size of the aorta. These data, together with the fact that neither the number nor the size of VSMCs were differentiated in other arteries of adult knock-out mice, suggest that the differences observed in the aorta were due to the phenotypic switch of the cells and to their lack of ability to adjust to the environment\textsuperscript{78}.

In fact, accumulation of synthetic VSMCs in knock-out vessels has been implicated in the loss of the vessels' ability to develop the same contractile forces. Synthetic VSMCs, contrary to the proliferative ones, demonstrate contractility defects that influence the response of the mutant vessels while various mechanisms have been implicated in the development of this phenotype. These suggest that miR-143 and miR-145 exert control the phenotypic switch to synthetic cells of smaller size that display reduced response to receptor-mediated signals\textsuperscript{78}.

1.2.2.5 MiR-17~92 cluster in cardiovascular diseases

MiR-17~92 is a well-studied cluster that has been implicated in many diseases including tumorigenesis, aging, neurodegenerative as well as cardiovascular diseases\textsuperscript{120,121}. It is located on chromosome 13 in humans and chromosome 14 in mice (Supplementary Table 1) and is transcribed in one primary transcript, pri-miR-17~92, harbouring 6 miRNA hairpins, miR-17, miR-18a, miR-19a, miR-19b-1, miR-20a and miR-92a-1 that comprise four different miRNA families\textsuperscript{122}.

Two paralogues of the main cluster have also been described (Figure 1.7), the miR-106b~25, located on chromosome 7 in humans, (7q22.1), in the 13th intron of the MCM7 gene, and chromosome 5 in mice, and the miR-106a/363 cluster, located on chromosome X in both organisms (Supplementary Table 1). In more detail, miR-106b~25 consists of six miRNAs miR-106a, miR-18b, miR-20b, miR-19b-2, miR-92a-
2 and miR-363 (Figure 1.7, middle panel) with miR-106b~25 cluster comprising three miRNAs, miR-106b, miR-93 and miR-25 (Figure 1.7, lower panel). Both miR-17~92 and miR-106b~25 are expressed abundantly in multiple tissues although miR-106a~363 is expressed at lower levels\textsuperscript{120, 123}. MiR-17~92 cluster, together with the paralogue ones, presents a combined total of 15 miRNAs that form four miRNA families grouped according to their seed sequences, the miR-17 family, the miR-18 family, the miR-19 family and the miR-92 family.

Upregulation of miR-19a, miR-19b and miR-92a is implicated in the development of AAA, implying a differential expression for these miRNAs due to post-transcriptional regulation\textsuperscript{124}. Additionally, different miRNAs of the cluster show implication in the control of different pathways of the pathophysiology of AAA and the development of the disease\textsuperscript{125-127}. More specifically, phosphatase and tensin homolog (PTEN) is one of the validated targets of the cluster as shown in \textit{in vitro} experiments\textsuperscript{128}. MREs for miR-19a and miR-19b-1 have been reported in the 3’-UTR of PTEN while introduction of the two miRNAs in miR-17~92-deficient cells repressed PTEN expression levels\textsuperscript{120}.

However, studies on the vascular expression of miRNAs miR-19a and miR-19b-1 are limited and provide controversial results. Although studies in murine endothelial differentiation of stem cells have reported increased miR-19 levels, antagonmir administration did not affect endothelial markers expression\textsuperscript{129}. Additionally, miR-92a
is highly expressed in endothelial cells with this overexpression inhibiting angiogenesis in ischemic conditions\textsuperscript{130}. Studies in endothelial cells isolated from patients with coronary artery disease (CAD) demonstrated elevated levels of both miR-17 and miR-92a\textsuperscript{131} while studies in patients with acute coronary syndrome demonstrated increased levels of miR-19a contrary to the results from patients with CAD where a decrease in miR-19a levels was observed compared to healthy controls\textsuperscript{132}.

TGF-β has been shown by numerous studies to be implicated in every step of the development of AAA and TGF-β signaling has been shown to be down-regulated in human AAA\textsuperscript{133} while in murine models it inhibits AAA\textsuperscript{134}. TGF-β has been implicated in the inhibition of infiltration of inflammatory cells\textsuperscript{135} while it promotes elastin and collagen formation\textsuperscript{136}. By doing so, TGF-β hinders the degradation of the ECM of the aortic wall by inhibiting the TNF-α-induced MMP-9 expression\textsuperscript{137}.

Interestingly, the miRNAs of the cluster show antagonistic effects with regards to the regulation of various targets and pathways. MiR-18a, miR-19a, miR-17 and miR-20 have been reported to target the TGF-β tumour suppressor pathway, thus facilitating tumour angiogenesis, while miR-17 and miR-20 target the E2F transcription factor family\textsuperscript{138-142}. In contrast to the effects mentioned above, miR-92a, which resides in the same cluster and is transcribed in the same primary transcript, has been shown to have antiangiogenic activity\textsuperscript{130}, features that highlight the regulation that the miRNA clusters exert in the same pathways and how they can control the equilibrium in these pathways by targeting different proteins that have contrasting effects.

With regards to SMCs, studies in mouse models of Pulmonary Hypertension (PH) showed that knock out of miR-17\textemdash 92 cluster in SMCs attenuated hypoxia-induced PH whilst reconstitution of the expression of the cluster restored the condition. Additionally, it was demonstrated that miR-17\textemdash 92 regulates SMC phenotype via the PDLIM5/transforming growth factor TGF-β 3/Smad3 pathway\textsuperscript{143}. Furthermore, miR-17\textemdash 92 is able to act on TGF-β and multiple TGF-β receptors, by inducing TGF-β 3 signaling to restore the differentiated SMC phenotype.
Discrepancies in the proliferation of VSMCs result in vascular remodeling and serve a key role in several vascular disorders, including atherosclerosis\textsuperscript{144}, pulmonary artery hypertension\textsuperscript{145} and AAA\textsuperscript{146}. However, the molecular mechanism underlying VSMC proliferation remains unclear. Studies have indicated that miRNA may also have roles in AAA through controlling the physiology and proliferation of VSMCs\textsuperscript{121}. Studies have revealed that the miR-17~92 cluster promotes VSMC proliferation in a murine model\textsuperscript{147} and mediates inhibition of VSMC proliferation via bone morphogenetic protein receptor type II (BMPR2)\textsuperscript{148} while overexpression of miR-17 with mimics demonstrated that miR-17 stimulates VSMC proliferation. MiR-17 also targets the 3'-UTR of retinoblastoma (RB) protein, which regulates the proliferation rate of the VSMCs\textsuperscript{149}. These results suggested that overexpression of miR-17 stimulated VSMC proliferation and enhanced cell cycle progression through promoting G1/S transition by increasing the levels of PCNA and E2F1.

Excess proliferation of VSMCs and inflammation have a key role in several vascular disorders while proliferation of VSMCs under inflammation is regulated by NF-\kappa B p65/microRNA-17/RB pathway activation\textsuperscript{150}. It has been previously shown that NF-\kappa B regulates miR-143 expression, enhancing the metastasis of hepatocellular carcinoma\textsuperscript{151}. The transcription factor NF-\kappa B is activated in and is a master regulator of the vascular inflammatory response in many vascular diseases, including AAA formation\textsuperscript{69}. NF-\kappa B p65 signalling directly regulates the miR-17 promoter activity, through potential p65 binding sites in the upstream regulatory region of the miR-17 gene, while suppression of RB levels, by activation of p65, was de-repressed when miR-17 was downregulated with inhibitors\textsuperscript{150}. NF-\kappa B p65 regulates the expression of miR-17, while miR-17 directly targets RB which modulates the proliferation of VSMCs through interacting with E2Fs. Taken together, these data support that VSMC proliferation is regulated by the NF-\kappa B p65/miR-17/RB pathway\textsuperscript{150}.

1.2.2.6 Limitations in miRNA research

Although, modulating miRNAs expression enables us to elucidate their role in disease, the specific inhibition of certain miRNAs can be cumbersome. So far, there are two common approaches employed in the research of miRNA inhibition: the antimiRs and the antagomiRs.
AntimiRs are chemically modified, single stranded DNA oligonucleotides that are used to inhibit miRNAs. They can be partially or fully complementary to their corresponding miRNA and they work by hybridising to their target miRNAs, thus inhibiting their function\textsuperscript{152}. Nonetheless, due to their small size, the employment of antimiRs as a tool in miRNA studies carries the risk of affecting RNA molecules of even partial complementarity and miRNAs other than the targeted one, thus inducing unwanted inhibition of miRNAs of the same family (Figure 1.8).

Moreover, antagomiRs are synthetic, single-stranded RNA analogues that are used in animal models which are fully complementary to miRNAs. They are 3’ cholesterol-conjugated, to facilitate the uptake by the cells, and they can inhibit miRNAs by hybridisation\textsuperscript{153}. Systemic administration of the antagomiRs can effectively reduce the levels of the targeted miRNAs for a sufficient amount of time\textsuperscript{152}. However, given that miRNAs of the same family share the same seed sequence, it becomes apparent that antagomiRs may have adverse effects. They are unable of providing specific targeting, and thus inhibition, of only one miRNA but instead they result in inhibition of multiple miRNAs of one family.

![Figure 1.8. Expression levels of 5 members of the miR-15 family upon inhibition mmu-miR-195a.](image)

The levels of 5 members of the miR-15 family were quantified with qPCR following transfection of mouse VSMCs with LNA-195a. Except for mmu-miR-15a, the other miRNAs show a downregulation with mmu-miR-195a, mmu-miR-16 and mmu-miR-497a showing statistical significance. Lataniotis et al., 2017.

Taking into consideration the short length of the mature miRNAs (19-24 nucleotides) and that miRNAs of the same family share the highly homologous 7 nucleotide seed
sequence, it becomes evident that in order to sufficiently inhibit specific miRNAs, without affecting miRNAs of the same family, a different approach is needed. Given the higher diversity of the genomic loci among different miRNAs, targeting the genomic locus of a miRNA gene rather than the mature product seems to be a promising alternative. Editing the miRNA genes will subsequently affect the maturation process and will result in sufficient inhibition of specific miRNAs. Various gene editing platforms have been developed over the past years. Zinc-Finger Nucleases (ZFN), Transcription Activator-like Effector Nucleases (TALENs) and Clustered Regular Interspaced Short Palindromic Repeats (CRISPR) as well as studies in vitro and in multiple organism models have employed the above mentioned techniques to specifically inhibit miRNAs\textsuperscript{154, 155, 156, 157, 158}.

1.3 Gene editing

Gene editing is defined as DNA engineering that results in the deletion, insertion or insertion of specific sequences. Many platforms have been developed and used over the last few decades.

ZFNs was historically the first platform of gene editing of eukaryotic genome to be developed. It contains zinc-binding sequences in their DNA-binding domains and can provide specific recognition of sequences\textsuperscript{159, 160}. Zinc domains can be modified to recognise any triplet of DNA thus being able to target any DNA sequence in eukaryotic cells\textsuperscript{161}. Fused to bacterial endonuclease FokI and used in pairs, ZFNs can be employed to induce double stranded breaks (DSBs) effectively in any DNA sequence.

Alongside, TALENs employ TAL effector DNA-binding domains that is able to specifically recognise single nucleotides thus rendering their employment easier and faster than ZFNs leading to DSBs in living cells\textsuperscript{162}. In more detail, TALENs are able of being engineered in order to target any DNA sequence in the genome. For a DSB to be induced a pair of TALENs is engineered to recognise a 20nt DNA sequence. Each TALEN of the pair, fused to FokI, recognises and binds the sequence flanking the targeted position and induces a DSB\textsuperscript{163, 164}.
In addition to these techniques, a new and highly versatile RNA guided gene-editing technique, CRISPR/Cas, which is able to target DNA or RNA sequences with high specificity, has emerged as an attractive platform for specific gene editing. CRISPR was first described over 3 decades ago as a locus harbouring high numbers of short direct DNA repeats interspaced by short sequences. A series of publications demonstrated that CRISPR is an adaptive immune system that, in nature, is used by bacteria as a defence mechanism against phages, upon re-infection.

In the current study, the CRISPR/Cas9 platform was assessed as an alternative to the traditional miRNA inhibition techniques in order to inhibit specifically miRNAs by targeting the miRNA genomic locus rather than the mature RNA sequence.

1.3.1 Clustered Regularly Interspaced Short Palindromic Repeat (CRISPR)/ CRISPR associated (Cas)

CRISPR/Cas is an adaptive immune system, naturally found in ~50% of bacteria and 90% of archaea and was firstly described in 1987. The CRISPR locus consists of DNA sequences that encode for proteins and for a series of DNA repeats that are interspaced by non-repetitive DNA sequences originating from phages. According to the organisation of their locus and the Cas proteins that are used, CRISPR/Cas systems are classified in six types (I-VI), (Figure 1.9).
Figure 1.9 Classification scheme for Cas proteins. Two classes, and five types (the best described ones) of Cas proteins are presented along with a schematic representation of the operon and its organisation together with the domain architecture. For each type of Cas system the name of the gene is indicated for each subunit. Genes that are homologous are colour-coded and identified by a family name. Genes and gene regions encoding components of the interference module (crRNA or Cas9 proteins) are highlighted with a light brown background. The adaptation nucleases (cas1 and cas2) and cas6 are dispensable in subtypes III-A and III-B. Also the regions of cas9 that correspond to the RuvC-like nuclease (green), HNH nuclease (yellow), recognition lobe (purple) and PAM recognizing domain (pink) are depicted. The regions of cpf1 (also known as Cas12a) aside from the RuvC-like domain are shown in grey. Makarova et al., 2015.

Although there is a plethora of Cas proteins that are used in CRISPR systems in nature, the type II endonuclease Cas9 form Streptococcus pyogenes (SpCas9) is of particular interest since it is the Cas protein that is employed in the vast majority of studies. The reason for this is that although its relative big size (1,368 amino acids, 4.10 kbp)\(^{175}\) in type II CRISPR systems only one Cas protein is used, contrary to
the other types that utilise multiple proteins to induce the DSB\textsuperscript{176}, (Figure 1.10). Regarding the organisation of the locus, viral sequences, that derive from viral genome after unsuccessful infection, are incorporated in the CRISPR repeat-spacer array with the help of the adaptation proteins of the cas operon (Figure 1.10).

![Figure 1.10. CRISPR genomic locus.](image)

The type II CRISPR locus from \textit{Streptococcus pyogenes} is shown as an example. The CRISPR genomic locus is comprised of the Cas genes, the tracrRNA and of regularly interspaced repetitive sequences (spacers) that are processed to crRNA. The spacer sequences are complementary to viral DNA and are used for targeting viruses upon re-infection. From Doudna and Charpentier, 2014.

These sequences are transcribed into the precursor CRISPR RNA (pre-crRNA) that shares partially complementarity with the trans-activating crRNA (tracrRNA). This complementarity is used for the tracrRNA to bind the pre-crRNA in a duplex (pre-crRNA:tracrRNA) that is then recognised by RNase III and cleaves the pre-crRNA to mature crRNA\textsuperscript{176} (Figure 1.11). The tracrRNA is then bound on the Cas9 protein and is used as a tether for the crRNA which guides the endonuclease to target specific DNA sequences according to the its sequence\textsuperscript{179}.

![Figure 1.11. Precursor crRNA maturation.](image)

As immune system, CRISPR/Cas9 in \textit{Streptococcus pyogenes} functions by recognising and binding a tri-nucleotide sequence of the form 5’-NGG-3’, called
Protospacer Adjacent Motif (PAM). DNA of the invading virus. Following the PAM binding, Cas9, using the crRNA as a guide, interrogates the adjacent sequence, through Watson and Crick base pairing, and if there is enough complementarity, it induces a blunt-end double stranded break (DSB), using its activated endonuclease domains, 3-4 nucleotides upstream of the PAM. In order for the DSB to happen, the first 17-20 nucleotides of the 5′-end of the crRNA need to be complementary to the targeted sequence. Two separate endonuclease domains, HNH and RuvC, induces the DSB at 3 nucleotide upstream of the PAM (Figure 1.12).

**Figure 1.12. CRISPR/Cas9 in nature.** The crRNA:tracrRNA duplex is of paramount importance for the function of the CRISPR/Cas9 system. By binding to the crRNA through its complementarity, the tracrRNA keeps it bound to the Cas9 protein which uses it as a guide to recognise the targeted sequence. This targeted sequence needs to be followed by a 5′-NGG-3′sequence for the SpCas9 to induce the DSB 3-4nt upstream of the PAM site with the RuvC and HNH domains. Adapted from Doudna and Charpentier, 2014.

Furthermore, it has been demonstrated that humanised SpCas9 is still functional in eukaryotic cells when the crRNA:tracrRNA is substituted by a chimeric single guide RNA (sgRNA) produced by fusing the 3′-end of the crRNA with the 5′-end of the tracrRNA (Figure 1.13), paving the way for fast, easy and cheap gene editing. SgRNA is an easy to design and fast to produce 20nt long RNA molecule that can be designed to target any sequence and is able of guiding SpCas9 to target and cleave any DNA sequence located upstream of the PAM sequence.
A chimeric sgRNA can be used in the place of crRNA:tracrRNA to guide the Cas9 protein to the targeted sequence. Upon recognition, Cas9 unravels the dsDNA using the PAM sequence, the first 20nt of the 5’-end of the sgRNA are bound to the targeted sequence and the Cas9 cleaves both strands of the DNA with two distinct domains with endonuclease activity, RuvC and HNH. From Doudna and Charpentier, 2014.

After the DSB, two different pathways can be employed to repair it: (A) the error prone Non Homologous End Joining (NHEJ), which alters the DNA sequence and can lead to premature stop codons, and (B) the Homology Directed Repair (HDR), which requires a highly homologous DNA template and can be employed for the insertion of specific sequences in the edited site (Figure 1.14).

Upon cleavage from the CRISPR/Cas9, the cell is trying to repair the DSB by following two distinct pathways. (A) The first, NHEJ, is an error prone repair which can potentially lead to a premature stop codon thus prematurely terminating the mRNA transcription, subsequently resulting to a non-functioning protein. (B) The second, less frequent pathway, HDR, can be employed for the introduction of specific sequences, such as genes or stop codons. Adapted from Doudna and Charpentier, 2014.
CRISPR/Cas9, due to its versatility, has emerged as an alternative to Zinc Finger Nucleases (ZFNs) and the Transcription Activator-Like Effector Nucleases (TALENs) gene editing techniques\textsuperscript{181}. As an RNA-guided system, CRISPR/Cas9 is easier to use, faster to design and cheaper to produce compared to ZFNs and TALENs which rely on protein-DNA interaction. Moreover, since the 5'-NGG-3' motif is found in every 8 nucleotides, on average, in the human genome\textsuperscript{182}, CRISPR/Cas9, employing SpCas9, can be used to edit almost every genomic site. Adding to this, the discovery of new Cas proteins with different PAM requirements and different ways of inducing the DSB, expand the gene editing toolbox with new CRISPR/Cas systems that can be used to edit different genetic loci according to individual needs\textsuperscript{183}.

CRISPR’s versatility and ease of use has sparked a series of publications that expand on the arsenal of gene editing providing myriad alternative options to SpCas9 depending on the chemical nature of the molecule (DNA or RNA), the sequence of the locus targeted, the nature of the intended induced DSB (staggered or blunt end) and the intended editing (gene disruption, DNA base editing, RNA base editing or RNA degradation)\textsuperscript{183, 184, 185}.

1.3.2 Limitations of CRISPR editing

CRISPR/Cas9 technology has revolutionised gene editing providing a highly advantageous platform for editing almost any sequence in the genome of multiple organisms, including humans and mice. Therefore, it is considered as being a highly promising approach to develop therapies against multiple diseases\textsuperscript{186, 187} as well as an alternative to gene therapy, in order to battle a number of genetic diseases\textsuperscript{188, 189, 190, 191, 192}. However, there are major innate limitations of the technique that hinder its universal application for therapeutic purposes. Namely, two of the most important limitations are the possibility of the off-target editing in unwanted loci and the delivery of the system, either in \textit{in vitro} or \textit{in vivo} models.

1.3.2.1 Off-target editing

Regardless of the advances CRISPR/Cas9 has introduced in the field of gene editing, important concerns remain about the possibility of inducing unwanted editing in non-
targeted sites. Although various, preliminary studies in murine models did not show any off-target editing\textsuperscript{193, 194} the concerns about the precise editing of only the targeted sites have not yet convincingly addressed with further studying needed to decipher the full impact of CRISPR/Cas9 editing.

As described above, CRISPR/Cas9 employs a 20nt long sgRNA to target and edit DNA sequences\textsuperscript{179}. The complementarity between the guide and the targeted sequence activates the endonuclease domains of Cas9. Regardless of the high specificity of the sgRNA to recognise and bind the targeted sequence, it can tolerate a small number (varying from 3 to 5) of mismatches in its sequence\textsuperscript{195}. More specifically, the greater the distance of the mismatches from the 3'-end of the sgRNA the more tolerable the binding is, with nucleotides 2-8 (from the 3'-end) being the least tolerable\textsuperscript{196}. Taken the short length of the sgRNA together with the relative tolerance to mismatches Cas9 can withstand into consideration, it becomes apparent that is likely that Cas9 is guided to target and induce DSBs in multiple unintended sites.

Software employed to design sgRNAs, addressing the problem of off-target mutagenesis, provide a list of putative off-target sites according to the tolerance of sgRNAs to mismatches. With regards to research and, more importantly, therapeutic and clinical applications of the system, off-target editing is a major concern. Although initial studies have provided data about the mechanism behind the recognition of off-target sites\textsuperscript{196}, data that are used for the prediction of sites of unintended editing, additional reports \textit{in vitro} using mouse stem cells and differentiated human stem cells support that editing with CRISPR/Cas9 can lead to large deletions and sequence rearrangements\textsuperscript{197}. These changes can potentially lead to editing oncogenes and/or protein coding genes resulting in severe implications.

Various strategies have been used to address the issue of unwanted editing. These include using shorter, 17nt long sgRNA, shortening the guide at its 5'-end which leads to a guide less tolerable to mismatches\textsuperscript{195}, the addition of two Guanine nucleotides at the 5'-end of the sgRNA\textsuperscript{198}, titrating the concentration of the Cas9 protein used\textsuperscript{199}, replacing the WT Cas9 with D10 nickases, where only one of the endonuclease domain is active, and using two sgRNAs flanking a specific point thus leading to immediate repair of any unintended edited sites\textsuperscript{200}, fusing inactive Cas9 (dCas9) with
FokI nuclease$^{201}$ and using phage-derived anti-Cas9 inhibitors that prevent Cas9 from editing the targeted DNA, thus needing higher complementarity to induce editing$^{202}$. However, regardless of the numerous techniques that have been developed, the issue of the off-target editing still remains to be resolved.

1.3.2.2 Delivery of the CRISPR/Cas9 system

A second major limitation for the CRISPR/Cas9 platform is the delivery of the system to cells or organisms either at the embryo or adult stage. This limitation becomes even more prominent in therapeutic applications depending on the gene editing approach that is used to change the sequence of the disease-causing gene. As explained above, two pathways can be employed to repair the DSB that is induced upon editing with Cas9, the NHEJ or the HDR. The former pathway, which leads to the disruption of the sequence of the targeted gene, is used when silencing of a mutant gene is needed while the latter one is used for correcting a mutant gene back to its wild-type sequence. HDR based therapies demonstrate added advantages since they can address the off-target problems with the use of nickases as well as provide highly specific editing. However, the need to deliver Cas9 proteins, sgRNAs and the donor DNA simultaneously and in high efficiency in order for the HDR to occur, together with the low editing efficiency the pathway exhibits$^{189}$, renders this approach challenging.

Multiple in vitro techniques have been used with varying efficiencies. These include the use of liposomes$^{196}$, nucleofection$^{203}$, electroporation$^{204}$, lentiviral based delivery$^{158}$ and microinjections in embryos$^{205}$. Additionally, ex vivo approaches have been studied for the delivery of the system to cells isolated from mice or patients under in vitro conditions and the re-introduction of the edited cells to the individuals$^{206}$. Nonetheless, the in vivo delivery still remains the most promising approach for therapeutic applications.

One of the first techniques used for gene therapy was the employment of adeno-associated viruses (AAVs) which is to this day one of the most advanced technologies for in vivo editing$^{207}$. However, this technique presents various limitations since AAV vectors are big in size, thus hindering the delivery efficiency, they can induce off-target editing$^{208}$ and can trigger immune response due to the bacterial origin of the
Cas proteins. A second approach is that of the direct delivery of Cas9/RNP, either in vivo or in vitro, with the use of lipoproteins\textsuperscript{209, 210}. The advantages this approach presents are the ease at producing high levels of recombinant Cas proteins and the track record in protein therapeutics. Moreover, other in vivo deliveries have been developed. Hydrodynamic delivery of plasmid DNA, that harbours the sequence for Cas9, sgRNAs and donor DNA\textsuperscript{211}, is an alternative to lipoprotein or polyethyleneimine (PEI) delivery of macromolecules, approaches that have not worked in the past. Lastly, two new, highly promising techniques have been described. The first one employs gold nanoparticles, conjugated with DNA, complexed with Cas9 RNP and donor DNA, entering the cells under in vivo conditions via endocytosis (gold nanoparticles) while the second one uses a sialoglycoprotein receptor ligands that facilitate the receptor-mediated endocytosis of the system\textsuperscript{114, 212}.

1.3.3 CRISPR/Cas9 and miRNAs

The use of gene editing for inhibition of miRNAs has already been demonstrated in various studies\textsuperscript{213-215}. More specifically, with regards to CRISPR/Cas9, it has been reported that it can target the expression of miRNAs, although with contradicting results about its efficiency\textsuperscript{202, 216}. Moreover, in vivo studies showed that the platform can be used for the excision of a 53-kb fragment leading to knock-out mice to study the effects of the deletion of the largest miRNA cluster\textsuperscript{217}, the targeting of entire families\textsuperscript{218} while double knock out mice have also been reported\textsuperscript{219}. Further to this, CRISPR/Cas9 has also been used to elucidate both the pathway of the miRNA induced silencing\textsuperscript{220} as well as the miRNA biogenesis\textsuperscript{221}. 47
Chapter 2 AIMS OF THE STUDY

MiRNAs are important and powerful regulators of many cellular processes by targeting and modulating the expression levels of multiple mRNA targets that participate in various pathways. The study of the miRNAs, though, is hindered by the fact that they form families that are comprised of members that share identical seed sequences and high homology throughout their sequence. A new approach for the study of miRNAs using gene editing has emerged as a promising alternative to overcome these limitations.

In the current study CRISPR/Cas9 is employed as a powerful platform to induce editing in the miRNA genes in order to elucidate its application in the miRNA study and the role of miRNAs. More specifically, the aims of the study are:

1. To study whether gene editing with CRISPR/Cas9 is capable of replacing the traditional methods of inhibiting miRNAs, overcoming the non-specific inhibition of the miRNAs of the same family. Following this approach, instead of targeting the mature miRNAs, editing of their genomic locus was thought to provide specific knock out only of the targeted miRNA.

2. To establish a reliable workflow for inhibiting miRNAs. CRISPR/Cas9 is a highly versatile and easy to use platform for editing almost any locus in the genome. Valuable conclusions reached, through studying its application in the inhibition of miRNAs, were postulated to simplify miRNA inhibition with CRISPR/Cas9 and potentially add another tool in the arsenal of miRNA study if not replace the existing methods.

3. To study the regulatory mechanisms and function of miRNAs and more specifically of miR-195a. Various miRNAs have been implicated in cardiovascular diseases with miR-195a being among them. However, it is not as well studied as other miRNAs e.g. from the miR-17~92 cluster. The current study aimed to provide a better understanding of the maturation process of the primary transcript and more specifically of the proteins that recognise, bind and participate in the maturation process.
Chapter 3 MATERIALS AND METHODS

3.1 Cell Culture

Human Embryonic Kidney 293T cells

Human Embryonic Kidney (HEK) 293T cells (HEK293T) were at fetal age, obtained from ATCC (CRL-3216) and cultured according to the recommended protocol. The cells were cultured in Dulbecco’s Modified Eagle’s Medium (DMEM, ThermoFisher Scientific #12491-015), supplemented with 10% heat-inactivated Foetal Bovine Serum (FBS) and 1% penicillin/streptomycin (100 U/ml penicillin and 100 µg/ml streptomycin) at 37°C in a humidified atmosphere of 95% air/5% CO₂. HEK293T cells were subcultured with a ratio of 1:3 or 1:4 every 2-3 days up to passage 20.

Human Umbilical Vein Endothelial Cells

Human Umbilical Vein Endothelial Cells (HUVECs) were obtained from Lonza and cultured as described previously. HUVECs were plated on gelatin coated flasks (Sigma-Aldrich #9000-70-8, 0.04% in DMEM) and cultured in Medium 199 (ThermoFisher Scientific #11150059) supplemented with 10% FBS, 1% penicillin/streptomycin (100 U/ml penicillin and 100 µg/ml streptomycin) and 3 ng/ml endothelial cell growth factor (Sigma, St Louis, Mo # E9640), 1µg/ml endothelial growth supplement from bovine neural tissue (Sigma # E2759-5X15MG), 10 U/ml heparin (Sigma # E4643), 1.25µg/ml thymidine at 37°C in a humidified atmosphere of 95% air / 5% CO₂. HUVECs were subcultured at a ratio of 1:3 every other day to passage 12.

Human Smooth Muscle Cells

Human Aortic SMCs were obtained from ATCC (PCS-100-012), were of male gender, adult age and plated in gelatin coated flasks (0.04% gelatin in DMEM) and cultured
in DMEM, supplemented with 15% FBS and 1% penicillin/streptomycin (100 U/ml penicillin and 100 µg/ml streptomycin) at 37°C in a humidified atmosphere of 95% air/5% CO₂. Cells were passaged at a ratio of 1:3 every three days to passage 10.

**Mouse Smooth Muscle Cells**

VSMCs were isolated by enzymatic digestion of mouse aortas derived from C57BL mice (8-10 week old) as described previously. They were of male gender and cultured in Dulbecco's modified Eagle's Medium (DMEM, ThermoFisher Scientific, Runcorn, UK, #11965-084) supplemented with 10% heat-inactivated Foetal Bovine Serum (FBS, # 16000044), 2mM L-glutamine and 1% penicillin/streptomycin (100U/ml penicillin and 100 µg/ml streptomycin, #15240062), at 37°C in a humidified atmosphere of 95% air/5% CO₂. VSMCs were plated on gelatin coated flasks (0.04% in DMEM, ThermoFisher Scientific, Runcorn, UK). Cells were passaged at a ratio of 1:3 every other day up to passage 30.

### 3.2 Single Guide RNA Design and Cloning

The CRISPR DESIGN tool (http://crispr.mit.edu, no longer available), provided by the Zhang Lab, MIT 2015, was used to design the single guide RNAs. Briefly, the DNA sequence of the locus of the miRNA was submitted in FASTA format. After analysing the sequence, the tool’s algorithm returned a list of possible guides, ranked according to a quality score of inverse likelihood of off-target sites. It also provided information about the specific position of the sequence they target in the miR-195 locus, putative off-target sites, with their corresponding score, and the number and positions of the mismatches between the guides and their off-target sites. Moreover, CRISPR Finder on-line tool, provided by Wellcome Trust Sanger Institute, was used to identify additional off-target sites for the selected guides in order to eliminate the possibility of not detection due to algorithm bias. The highest scoring guides, which targeted sequences either in or close to the miRNA stem-loop, were selected.

After choosing the guides to be used, they were synthesized with overhangs which would facilitate the cloning into the backbone pX330 plasmid vector (Addgene
#42230), (Figure 3.1A) or LentiCRISPRv2 lentiviral vector (Addgene #52961), (Figure 3.1B).

**Figure 3.1 Vectors pX330 and LentiCRISPRv2.** (A) Map of vector pX330 with its corresponding elements. (B) Map of vector LentiCRISPRv2 with its corresponding elements.
In detail, to clone the guide sequence into the vectors the two oligos were synthesized in the form:

**Oligo1:**
\[5'\text{-CACCGNNNNNNNNNNNNNNNNNNN}-3'\]

**Oligo2:**
\[3'\text{-CNNNNNNNNNNNNNNNNNNNCAA}-5'\]

For cloning the oligos into the backbone vector a single step digestion-ligation protocol was used.

Firstly, the oligos were phosphorylated and annealed. For the reaction, 1μl Oligo1 (100μM), 1μl Oligo2 (100μM), 1μl 10x T4 Ligation Buffer (NEB M0308S), 0.5μl T4 PNK (NEB) and 6.5μl nuclease free H\(_2\)O were added to a final volume of 10μl. The oligos were annealed in a thermocycler at 37°C for 30min, 95°C for 5min and then the temperature was ramped down to 25°C at 5°C/min. After the annealing, the oligos were diluted 250-fold (1:250) for the pX330 vector or 10-fold (1:10) for the LentiCRISPRv2 vector, respectively.

For the single digestion-ligation reaction, 1μl of vector (100ng), 2μl Phosphorylated and annealed diluted oligo duplex from the previous step, 2μl 10x Tango Buffer (NEB), 1μl FastDigest enzyme, 1μl DTT (10mM) (Sigma D9779), 1μl ATP (10mM), 0.5μl T7 DNA ligase (NEB M0318S) and 11.5μl nuclease free H\(_2\)O were added to a final volume of 20μl. The BbsI enzyme (ThermoFisher Scientific FD1014) and Esp31 (ThermoFisher Scientific ER0451) were used for the pX330 and LentiCRISPRv2 vectors, respectively. The digestion-ligation reaction was incubated in a thermocycler at 37°C for 5 min, 23°C for 5 min, repeated for 6 cycles and 15 cycles for the pX330 and LentiCRISPRv2 vectors, respectively.

The ligated product was treated with PlasmidSafe exonuclease to remove the unligated linear plasmids. For the PlasmidSafe reaction 11μl Ligation reaction (from the previous step), 1.5μl 10x PlasmidSafe Buffer, 1.5μl ATP (10mM) and 1μl PlasmidSafe
exonuclease (PlasmidSafe™ ATP-Dependent DNase kit #E3101K) were added to a final volume of 15μl. The reaction was incubated at 37°C for 30min.

Next, the ligated vectors were transformed into Stbl3 E. Coli competent cells (ThermoFisher Scientific #C7373-03). Briefly, 2μl of the PlasmidSafe treated product were added into a shot (20μl) of ice-cold chemically competent Stbl3 cells and incubated on ice for 30 min. They were then heat shocked at 42°C for 45sec and placed back on ice for 2 min. Next, 240μl of SOC medium was added to the transformed bacteria and incubated at 37°C for 40 min. Subsequently, 100μl of bacteria were plated on an LB agar petri dish containing ampicillin (50μg/ml) and incubated at 37°C overnight.

3.3 In Vitro Transcription (IVT) of sgRNA

For the IVT, the GeneArt™ Precision gRNA Synthesis Kit (Invitrogen #A29377) was used to generate single guide RNAs (sgRNAs), according to manufacturer’s recommendations. The DNA template of the sgRNA was PCR assembled as shown in Figure 3.2. This DNA template was then used to generate the sgRNAs by IVT. All IVT Target primer sets are provided in Supplementary Table 2.

![Figure 3.2. PCR assembly of the gRNA DNA template using the Target F1 and Target R1 oligonucleotides.](http://www.thermofischer.com)

The oligonucleotides were used both as primers and templates together with the T7 primer mix and the tracr fragment for the IVT of the sgRNA. Figure from [http://www.thermofischer.com](http://www.thermofischer.com). Date accessed 15/08/2019.
PCR Assembly

The following PCR was used for the assembly of the gRNA DNA template: 12.5µl of 2x Pfusion™ High-Fidelity PCR Master Mix, 1µl of Tracr Fragment + T7 Primer Mix, 1µl of Target F1/R1 oligonucleotide mix (0.3µM) and 10.5µl of nuclease free H₂O were added to a final volume of 25µl. The following PCR programme was run on a Veriti® 60-well Thermal Cycler: initial denaturation at 98°C for 10sec and then underwent 32 cycles of 98°C for 5sec, 55°C for 15sec and then a final extension step at 72°C for 1min.

IVT Reaction

After the DNA template assembly, IVT was used for the generation of the gRNA. The reaction mix for each sample contained 8µl of NTP mix (100mM), 6µl of gRNA DNA template from the previous step, 4µl of 5x TranscriptAid™ Reaction Buffer and 2µl of TranscriptAid™ Enzyme Mix to a final volume of 20µl. The samples were incubated on a Veriti® 60-well Thermal Cycler form Thermo Fisher Scientific at 37°C for 3h. As a final step, to remove the DNA template, 1µl of DNase I (1U/µl) was added to the reaction and was incubated at 37°C for a further 30min. After the DNA template digestion, the gRNA was purified using the gRNA Clean Up Kit following the manufacturer’s recommendations.

Transfection

For transfection experiments, 293T cells were trypsinised and re-suspended into complete medium with no antibiotics. Subsequently, cells were counted using a haemocytometer and plated in 24 well plates at a cell density of 240,000 cells/ml. A total of 500µl of cell suspension was plated per well.

On the following day, cells were transfected with plasmid vectors using ViaFect™ Transfection Reagent by Promega (Catalog No E4981). Two mastermixes (A and B) were prepared. For mastermix A, 500ng of the plasmid were mixed with 25µl of DMEM per 24 well and for mastermix B, 2µl of ViaFect™ were added to 25µl of DMEM
per transfection. After incubating at RT for 5 min the two mastermixes were combined and incubated for further 20 min. Afterwards the transfection mastermix was added to the corresponding well. All transfection reactions were performed in duplicates.

After 24 hours of the transfection, the conditioned medium was supplemented with 500μl of complete medium. Two days later, HEK293T cells were detached by pipetting and 500μl of the cell suspension was re-plated to a 6 well plate, supplemented with 1.5ml of complete media. Cells were returned to the incubator and maintained for an additional 3 days. Additional DNA samples were harvested on Day5.

3.4 DNA Extraction

For the DNA extraction cells were pelleted and lysed in 300μl of QuickExtract™ DNA Extraction Solution 1.0 (Epicentre #QE09050). The samples were heated at 65°C for 20 min and then at 96°C for 10 min. The DNA concentration was measured using the NanoDrop 1000 (ThermoFisher Scientific) device.

3.5 Genomic PCR

For the amplification of the genomic locus of miR-195, primers were designed using the online tool OligoPerfect™ Designer by Thermo Fisher Scientific and Kapa HiFi PCR Kit Hot Start from Kapa Biosystems (Kapa Biosystems #KK2502) was used. The possible editing of the off-target sites, as provided by http://crispr.mit.edu, was also tested. The sequences of the genomic locus of the off-target sites were retrieved from www.ensebml.org, using the coordinates of the possible off-target sites as provided by the CRISPR DESIGN online tool. Details for the primer sequences are provided in Supplementary Table 3.

The reaction mix for each sample contained 9μl of 5x Kapa GC Buffer, 1.35μl of Forward Primer (10μM), 1.35μl of Reverse Primer (10μM), 1.35μl of dNTPs (10mM), 6μl of DNA template (50ng/μl), 0.45μl of Kapa Polymerase and 25.5μl of nuclease free H₂O to a final volume of 45μl.
The following PCR programme was run on a Veriti® 60-well Thermal Cycler for the miR-195 genomic locus (both human and mouse): initial denaturation and enzyme activation at 95°C for 5 min and then 30 cycles of 98°C for 20sec, 62°C for 15sec, 72°C for 20sec, then were heated at 72°C for 1min. Amplicons were visualized on a 2% agarose gel in 1xTBE buffer supplemented with Safe-view (10μl/150ml) run at 200V for 30 min.

### 3.6 PCR Clean Up

The PCR product was purified using the Thermo Scientific GeneJET Gel Extraction and DNA Cleanup Micro Kit (Catalog Number #K0702), following the manufacturer’s recommendations. Briefly, the volume of the PCR product was adjusted to 200μl with nuclease free H₂O. Then, 100μl of binding buffer were added followed by 300μl of 100% ethanol. The solution was mixed thoroughly by pipetting, transferred to a DNA purification Micro Column and centrifuged at 14,000 × g for 1 min. The flow-through was discarded and 200μl of Pre-wash Buffer was added to the column before being centrifuged at 14,000 × g for 1 min. The flow-through was discarded and 700μl of Wash Buffer was added to the column before being centrifuged at 14,000 × g for 1 min, a step that was repeated for a second time. The column was then transferred to a clean 1.5ml micro-centrifuge tube and centrifuged at 14,000 × g for 1 min for the removal of any residual buffer. The column was then transferred to a clean 1.5ml micro-centrifuge tube and 10μl of nuclease free water were added to the column for three min incubation. For the elution of the purified DNA product, the column was centrifuged at 14,000 × g for 1 min. The DNA concentration of the purified PCR product was measured using a NanoDrop 1000.

### 3.7 T7 Endonuclease I Assay

To assess the gene editing efficiency, the T7 Endonuclease I assay (T7EI) was employed as described recently. Briefly, 90ng of the purified PCR product were diluted to a volume of 14μl, in nuclease free H₂O, and supplemented with 3.5μl 5x KAPA HiFi GC Buffer. PCR products were denatured by heating to 95°C for 10min
(Figure 3.3A) and then re-annealed by slowly ramping down the temperature to 25°C by a rate of 4.5% per minute (Figure 3.3B).

![Diagram of T7 Endonuclease I digestion](image)

**Figure 3.3. T7 Endonuclease I digestion.** T7EI recognises and cleaves mismatched DNA. (A) The PCR amplicons were denatured and (B) re-annealed randomly. (C) Upon T7 endonuclease I digestion, the mismatched strands were cleaved at the edited site (red squares). The fragments were then analysed to assess the genome editing efficiency.

After the re-annealing, 2μl of 10x NEB Buffer 2 (NEB #B7002S) and 0.5μl of the T7EI (NEB #M0302S) were added to each sample, to a final volume of 20μl for a further 1h incubation at 37°C. As a final step, 4μl of 6x Purple Loading Buffer (NEB #B7024S) were added and the samples were heated to 70°C for 10min (Figure 3.3C).

To visualize the T7EI digestion products, a 2% agarose gel in 1xTBE buffer was prepared. The samples were loaded and run at 170V for 50 min alongside a 1kb Plus DNA ladder (ThermoFisher Scientific #10787018). After the run, the gel was, rinsed twice with 1x TBE Buffer. The gel was subsequently stained with SYBR® Gold Nucleic Acid Gel Stain (ThermoFisher Scientific #S-11494) in 1xTBE buffer for 30min, protected from light. The gel was visualized using the Ettan DIGE Imager by GE Healthcare.
For the quantification of the T7EI results, the ImageJ software was used (freely available at http://imagej.net). The band intensity of the uncut DNA substrate, the cleaved products and background was measured. The values of the band intensities were subtracted against the value of the background and the editing efficiency was calculated using the following formula:

\[
\text{Indel (\%) = } (1 - \sqrt{1 - \frac{B+C}{A+B+C}}) \times 100,
\]

where, \(A\)=uncut DNA substrate, \(B, C\)= digestion products

### 3.8 Restriction Enzyme Digestions

The SspI and MwoI digestions were set up in order to detect mutations in the genomic loci of mouse miR-195. More specifically, the SspI restriction enzyme identifies the sequence 5’-AATATT-3’ and the MwoI identifies the sequences 5’-GCNNNNNNNGC-3’ (Figure 3.4). Successful gene editing will interrupt these two sites, thus preventing the cleavage by the restriction enzymes. For the digestion, 8µl of the PCR reaction was used as a template in a digestion that contained 2µl 10x CutSmart (NEB #B7204S), 0.5µl of restriction enzyme and 9.5µl of nuclease free H₂O to a final volume of 20µl. The reaction was incubated in a thermocycler at 37°C for the SspI and at 60°C for the MwoI for 1h.

Restriction enzymes (all from NEB) were also used to assess gene editing with their corresponding sites outlined in Supplementary Table 4. Digestion reactions for all other restriction enzymes were incubated at 37 °C for 1 h in CutSmart Buffer with the exception of SspI that was incubated in NEB Buffer 2.1. Digestion products were visualised in 2% agarose gel in 1xTBE buffer.
3.9 Lentiviral particle transduction

To confer Cas9 expression in VSMCs, we used the LentiCRISPRv2 vector (Addgene, #52961) that encodes a *Streptococcus pyogenes* Cas9 (SpCas9) under the control of an elongation 1a short promoter (EFS)\(^{224}\). Lentiviral particles were produced using the lentiviral vector LentiCRISPRv2 and packaging plasmids pMD2.G and psPAX2 (Addgene, #12259, #12260) as described previously\(^{225}\). HEK293T cells were seeded in T75 flasks. When they reached a confluency of 70-80%, they were transduced with lentiviral vectors for the production of the lentiviral particles. Briefly, two mastermixes were prepared. Mastermix A consisted of 590μl of Opti-MEM® I Reduced Serum Media (ThermoFisher Scientific #31985070) and 10μg of the lentiCRISPRv2 vector harbouring the guide sequence of interest, 5μg of the pMD2.G envelope plasmid (Addgene #12259) and 7.5μg of the psPAX2 plasmid (Addgene #12260) and Mastermix B consisted of 600μl of Opti-MEM® I Reduced Serum Media and 50μl of ViaFect™. The mastermixes were incubated for 5 min at RT and then combined and incubated for further 20 min at RT. In the meantime, the medium of the HEK293T cells was replaced with 6ml of Opti-MEM® I Reduced Serum Media per flask. The transfection complexes were subsequently added to the cells and incubated at 37°C. Following a 6h incubation, the medium was replaced with 8ml of DMEM with 10% FBS and antibiotics supplemented with 1% Bovine Serum Albumin (Sigma-Aldrich #9048-46-8). After 60h, the lentiviral supernatant was collected, centrifuged at 1000 rpm for 5min and filtered through 0.45μm VWR Syringe Filters #28145481. A p24
antigen ELISA (Cell Biolabs) was used to determine the viral titer. The resulting solution was either used immediately or stored at -80°C.

### 3.10 Lentiviral Infection

VSMCs were seeded at T25 flasks at a density of 200,000 cells per flask. The next day cells were lentiviral infected using $1.6-2.4 \times 10^7$ TU/ml (Table 1). The media were supplemented with 8µg/µl of Polybrene (Sigma-Aldrich #28728-55-4) and cells were incubated for 24h with the viral suspension. The next day, the lentiviral supernatant was removed, discarded and replaced with 5ml of complete medium.

Table 1. List of Lentiviruses used.

<table>
<thead>
<tr>
<th>Lentiviruses</th>
<th>Titer</th>
</tr>
</thead>
<tbody>
<tr>
<td>LentiCrisprv2 (Addgene #52961)</td>
<td>$1.7-2.3 \times 10^7$ TU/ml</td>
</tr>
<tr>
<td>pLKO.1 (Addgene #8453)</td>
<td>$1.6-2.4 \times 10^7$ TU/ml</td>
</tr>
</tbody>
</table>

### 3.11 TA Cloning

PCR amplicons from the edited samples was sequenced to determine mutation patterns. Purified, PCR amplified DNA from each sample was ligated to the pGEM®-T Vector (Promega #A3600). For the ligation reaction 5µl of 2x Rapid Ligation Reaction Buffer, 2µl of PCR purified product, 2µl of pGEM®-T vector (Figure 3.5) and 1µl of T4 DNA ligase (Promega, Catalog No M1801) were used. The reaction was incubated at RT for 1 hour.
Subsequently, 2.5μl of the ligation product were added to 50μl of JM109™ Chemically Competent E. Coli (Promega #L1001) and incubated on ice for 30 min. They were then heat shocked at 42°C for 45 sec and placed on ice for 2 min. A total of 240μl of SOC medium was added to the transformed bacteria that were incubated at 37°C for further 20 min. An ampicillin LB agar petri dish was precoated with 40μl of Xgal (50mg/ml) (Promega #V3941) and 40μl of IPTG (100mM) (Promega #V3955) and 70μl of bacteria culture were subsequently plated. The dish was incubated at 37°C, overnight.

Blue and white bacterial colonies were formed overnight, with blue colonies corresponding to the bacteria that were transformed with re-ligated, empty vectors. White bacterial colonies correspond to bacteria that were transformed with the vector conferring the PCR product. Individual bacterial colonies were picked and PCR screened as described above. Mutation patterns were identified by Sanger sequencing.

3.12 RNA Extraction

RNA was extracted using the miRNAeasy mini kit (Qiagen #74104). Briefly, cells were lysed using 700μl of QIAzol® Lysis Reagent. Following a brief incubation at ambient temperature, 140μl of chloroform was added to each sample and the solution was mixed vigorously for 15 sec. The samples were incubated for 5 min at RT and were
then centrifuged for 15 min at 12,000 x rpm at 4°C. Only 280µl of the upper aqueous phase were transferred to a new collection tube and 420µl of 100% ethanol were added and mixed thoroughly and transferred to an RNeasy® Mini column in a 2ml collection tube, washed according to the company’s recommendations. Total RNA was eluted using 25µl of nuclease free H2O and the RNA concentration was measured using a NanoDrop 1000.

3.13 MiRNA quantification

To assess miRNA expression RNA was reverse transcribed using the Megaplex Primer Pools from ThermoFisher (Cat. #4401009) as described previously.98 A total of 100ng of RNA were reverse transcribed using the Megaplex Rodent Primer Pools A. The reaction was performed according to the company’s recommendations (0.8µl of Pooled Primers were combined with 0.2µl of 100mM dNTPs with dTTP, 0.8µl of 10x Reverse-Transcription Buffer, 0.9µl of MgCl2 (25mM), 1.5µl of Multiscribe Reverse-Transcriptase and 0.1µl of RNAsin (20U/µl) to a final volume of 7.5µl). The RT-PCR reaction was set as follows: 16°C for 2min, 42°C for 1min and 50°C for 1 sec for 40 cycles and then incubation at 85°C for 5min using a Veriti® thermocycler (Life Technologies). The RT reaction products were diluted to 1ng/µl corresponding RNA and stored at -20°C. Taqman miRNA assays were used to assess the expression of individual miRNAs. For the quantitative Polymerase Chain Reaction (qPCR), 2.25ng of Megaplex reverse transcription product were combined with 0.25µl of 20x Taqman miRNA Assay (Life Technologies) and 2.5µl of the 2x Taqman Universal PCR Master Mix No Amp Erase UNG (ThermoFisher Scientific #4324018) to a final volume of 5µl. QPCR was performed on a Viia7 thermocycler at 95°C for 10 min, followed by 40 cycles of 95°C for 15sec and 60°C for 1min. U6 was used as a normalization control.

3.14 MiRNA Overexpression

Cells were plated at 60–70% confluency on the day before transfection. Mouse miRNA mimics and a non-targeting control were obtained by ThermoFisher Scientific
and transfected at a final concentration of 20nM using Lipofectamine RNAiMAX (ThermoFisher Scientific, Runcorn, UK) as described previously.

### 3.15 Gene expression

QPCR was used to assess the gene expression levels. In these studies, 1μg of RNA was reversed transcribed into cDNA using the High Capacity Reverse Transcriptase kit (ThermoFisher Scientific, Runcorn, UK). Prior to pri-miRNA assessment RNA samples were treated with DNase for 30min at 37°C, to remove any genomic DNA contamination in the preparations. For all genes, Taqman Assays were used with the exception of Carmn that was assessed using specific primers (Supplementary Table 3) and SyBr Select Mastermix (TermoFisher Scientific, Runcorn, UK). Beta Actin was used as a normalisation control.

### 3.16 Computational analysis

Secondary structures without pseudoknots were generated by RNAfold (Vienna RNA tools, http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi) using the Turner model (Standard settings), while secondary structures with pseudoknots were generated by vsfold5 (Chiba Institute of Technology, Japan, http://www.rna.it-chiba.ac.jp/~vsfold/vsfold5/) and Settings: 37°C temperature; Kuhn length 6; Jacobson-Stockmayer gamma=1.75 and contiguous stems=6. Visualisation of the secondary structures with pseudoknots was performed using the Pseudoviewer software (Inha University, S Korea http://wilab.inha.ac.kr/ pseudoviewer/). The RNAComposer (Poland, http://rnacomposer.cs.put.poznan.pl/) was used to generate the pbd-files through molecular RNA simulation from vsfold5 output. The RNAComposer output pbd-files were then input to pyMOL to generate high quality 3D images. MiR-497a stem loop is depicted in green and its terminal loop in magenta. MiR-195a stem loop is depicted in red and its terminal loop in yellow.
3.17 pLKO.1 cloning

The pLKO.1 vector was used for pri-miRNA overexpression experiments, since it is the recommended vector for expression of shRNAs under a T7 promoter (Figure 3.6).

For the cloning of the DNA sequence of mmu-miR-497~195, pGEM®-T vectors harbouring WT or mutant sequences were used as template. These sequences were resulted from editing of the genomic locus of the primary transcript. For cloning purposes, new guides harbouring sequences corresponding to the restriction sites of AgeI and EcoRI were designed (Supplementary Table 3). For the PCR amplification of the DNA sequence of the primary transcript, a specific programme was used on a Veriti® 60-well Thermal Cycler.

More specifically, two PCR reactions were prepared for each transcript (WT and mutants) to final volume of 45µl. The reaction mix for each sample contained 9µl of 5x Kapa GC Buffer, 1.35µl of Forward Primer (10µM), 1.35µl of Reverse Primer
(10μM), 1.35μl of dNTPs (10mM), 2μl of DNA template (50ng/μl), 0.45μl of Kapa Polymerase and 29.5μl of nuclease free H2O to a final volume of 45μl.

The following PCR programme was run on a Veriti® 60-well Thermal Cycler for the miR-195 genomic locus (both human and mouse): initial denaturation and enzyme activation at 95°C for 5 min and then 5 cycles of 98°C for 20sec, 57°C for 20sec, 72°C for 40sec, followed by 30 cycles of 98°C for 20sec, 62°C for 20sec, 72°C for 40sec then were heated at 72°C for 1min. Next, the two reactions were combined and cleaned-up, following the same protocol as explained above, with the final elution being in 15μl of H2O. Then, the eluted DNA was run on a 0.8% agarose gel in 1xTBE buffer supplemented with Safeview (10μl/150ml) at 170V for 50 min. Following this, gel purification of the amplicons was performed using the Thermo Scientific GeneJET Gel Extraction and DNA Cleanup Micro Kit (Catalog Number #K0702) following the company’s recommendations. In brief, the part of the gel that engulfed the DNA band of the correct size was excised using a scalpel and was transferred to a 1.5mL Eppendorf tube. Next, 200μl of Extraction Buffer were added and the gel mixture was incubated at 58°C for 10min. Then, 200μl of ethanol were added to the tube and mixed by pipetting. The mixture was afterwards transferred to a DNA purification Micro Column and was centrifuged for 60sec at 14,000 x g. Next, the column was washed with 200μl of Prewash Buffer and centrifuged for 60sec at 14,000 x g. Then, 700μl of Wash Buffer was added and centrifuged twice. The empty column was then centrifuged and then transfer to an empty microcentrifuge tube. In a final step, 15μl were added to the empty column to elute the DNA.

Following the clean-up, two digestions were set up with Agel restriction enzyme. In more detail, a reaction with 5μl of Buffer O, 2μl of Agel, with 15μl of PCR product and 28μl of H2O was set-up for the gel purified PCR product. A second reaction with 5μl of Buffer O, 2μl of Agel, 9.09μl of pLKO.1 vector (a total amount of 5000ng) and 33.91μl of H2O was set-up and they were both incubated at 37°C for 2 hours.

Following these digestions, the products were cleaned-up as per standard protocol eluting in 15μl of H2O. A second step of digestions was set-up with 5μl of CutSmart, 2μl of EcoRI, 15μl of PCR clean-up product and 28μl of H2O and were incubated at 37°C for both the PCR product and the vector.
As a last step, the digested products were cleaned-up as per standard protocol and the final concentration was measured using the NanoDrop 1000.

For the ligation reaction, 1µl of 10x T4 ligase buffer, 240ng of pLKO.1, 420ng of PCR clean-up and 1µl of T4 ligase were added to a final volume of 10µl and they were incubated at RT overnight.

### 3.18 RNA Pull-down assay

#### Cell lysate preparation

RNA Pull-down experiments were conducted to assess how changes in the tertiary structure of the pri-miR-497~195, due to CRISPR/Cas9 editing, might affect its recognition and binding from RNA binding proteins. For the RNA Pull-down experiment, VSMC were cultured in T175 flasks to a confluency of 90% before they were harvested. In more detail, the culture medium was removed from the flasks and the flasks were washed off with 4ml of phosphate-buffered saline (PBS; Sigma, cat. no. D8537). Next, the cells were scraped, and the cell suspension was placed in a 50ml falcon tube before it was centrifuged at 1,000 rpm for 5min. After spinning, the supernatant was removed and 800µl of Pierce IP Lysis buffer (ThermoFisher #87787) supplemented with protease inhibitor (Roche Complete Mini, EDTA-free, cat. no. 11836170), phosphatase and RNAse inhibitors (Life Technologies) were added and the solution was incubated on ice for 30min while mixing occasionally. After the 30min had elapsed, the lysate was centrifuged at 13,000 rpm for 20min at 4°C, the supernatant was removed and stored at -20°C.

#### IVT of primary transcript

For the RNA Pull-down experiments, primary transcripts corresponding to the wild type or the mutated sequences conferring a 16-nucleotide deletion, were PCR amplified with the use of forward primers containing the T7 promoter. Afterwards, the
PCR products were run on a 0.8% agarose gel in 1xTBE buffer (ThermoFisher #15581-028), stained with Safe-view 10µl/150ml (NBS Biologicals, NBS-SV5), at 170V for 50min. The products were then gel purified using the GeneJET Genomic DNA Purification Kit (#K0721). The DNA concentration of the gel purified products was measured using the NanoDrop 1000 (ThermoFisher Scientific) device and 400ng of DNA were used for the IVT reaction to a final volume of 20µl as described above in Chapter 3.3. The samples were incubated on a Veriti® 60-well Thermal Cycler form ThermoFisher Scientific at 37°C for 2h. Finally, 1µl of DNase I (1U/µl) was added to the reaction and was incubated at 37°C for a further 30min to remove the DNA template. Afterwards, the RNA was purified using the RNA Clean Up Kit following the manufacturer’s recommendations (ThermoFisher Scientific). Briefly, after adjusting the volume to 200µl with nuclease-free water, 100µl of Binding Buffer were added and mixed thoroughly. Then 300µl of ethanol were added before the mixture was transferred to a GeneJet™ RNA Purification Micro Column and centrifuged at 14,000 x g for 60 seconds. After discarding the flow through, 700µl of Wash Buffer 1 was added and the column was centrifuged at 14,000 x g for 60 seconds. Next, 700µl of Wash Buffer 2 was added and the column was centrifuged again at 14,000 x g for 60 seconds. The empty purification column was centrifuged to remove any residual Buffer and 10µl of nuclease-free water was used to elute the RAN. After the clean-up, the concentration was measured with the NanoDrop 1000 (ThermoFisher Scientific) and 50pmol of cleaned-up RNA was used for labeling in a final volume of 5µl.

RNA labeling

For the RNA labeling the Pierce RNA 3' End Desthiobiotinylation Kit (#20163) was used. In more detail, the RNA was supplemented to a final volume of 6.25µl, 0.5 µl of DMSO were added and it was incubated on a Veriti® 60-well Thermal Cycler form Thermo Fisher Scientific at 85°C for 10min and then reannealed by ramping down the temperature to 10°C.

For the labelling reaction 1.75µl of nuclease-free water, 3µl of 10x RNA Ligase Reaction Buffer, 1µl of RNase inhibitor, 6.25µl of 50pmol of reannealed RNA, 1µl of Biotinylated Cytidine Bisphosphate, 2µl of T4 RNA ligase and 15µl of PEG 30% were mixed in the above order to a final volume of 30µl. The reaction was carefully mixed
and incubated on a Veriti® 60-well Thermal Cycler from ThermoFisher Scientific at 16°C overnight.

After the overnight incubation, 70µl of nuclease-free water were added to the ligation reaction. As a next step, 100µl of chloroform were added in order to extract the RNA. The mixture was vortexed briefly and then centrifuged for 5min at 12,000 rpm to separate the phases. The aqueous phase was carefully removed and transferred to a 1.5ml Eppendorf tube. Then, 10µl of 5M NaCl, 1µl of glycogen and 300µl of ice-cold 100% ethanol were added before RNA being precipitated at -20°C for 2h. In the meantime, the cell lysate was precleared. In detail, 200µl (4ng/µl) of cell lysate per sample was used and 20µl of Pierce™ Streptavidin Magnetic Beads (1:10 of the lysate volume). A magnetic rack was used to gather the magnetic beads against the wall of the tube and washed twice with Tris-buffered saline (TBS, Product No. 28379, containing 0.05% Tween-20 Detergent). The beads were then re-suspended in 20µl of Tris and applied to the lysate. They were put in the cold room (4°C) at rotation for 2h. After the 2h had elapsed for the RNA precipitation, it was centrifuged at 13,000 × g for 15 minutes at 4°C. The supernatant was carefully removed and the pellet was washed with 300µl of ice-cold 70% ethanol carefully not to detach the pellet. The ethanol was carefully removed and the tubes flipped upside down for 5min to air-dry the pellet. Afterwards, the pellet was re-suspended in 20µl of re-suspension buffer (490µl of 20mM Tris-buffered saline supplemented with 10µl of 5M NaCl). They were incubated on a Veriti® 60-well Thermal Cycler from Thermo Fisher Scientific at 65°C for 10min before the temperature was ramped down to 4°C by 1.5% rate.

In the meantime, the RNA binding mastermix was prepared. In more detail, 20µl of Protein-RNA Binding Buffer, 60µl of 50% Glycerol, 1µl of 5M NaCl, 100µl of pre-cleared lysate and 19µl of nuclease-free water were mixed per sample. In a 1.5ml Eppendorf tubes, 200µl of the mastermix and 10µl of the labelled RNA were added and incubated at 4°C with rotation for 2h. In the meantime, 20µl of beads per sample were pre-cleared. In detail, a magnetic rack was used to gather the beads against the wall of the tube and they were washed twice with Tris. They were then re-suspended in 20µl of Tris and added to the RNA-protein samples and incubated for 2h at 4°C with rotation. Afterwards, a magnetic rack was used to gather the beads against the wall of the tube, the depleted supernatant was removed and kept at -20°C and the beads were washed off with 250µl of wash buffer for three times. Then, 40µl of elution
buffer was added, mixed well by vortexing and incubated at 37°C with mild agitation. With the help of a magnetic rack the beads were gathered against the wall of the tube and the eluate was removed and stored at -20°C.

3.19 In-gel digestion and HPLC-MS/MS analysis

To analyse the RNA Pull-down eluate, 20µl of each sample were loaded on to pre-cast 10-well NuPAGE Novex 4-12% gradient Bis-Tris polyacrylamide gels (Life Technologies), with 7µl of Novex Sharp Pre-stained Protein Standard (Invitrogen, Life Technologies) loaded as marker. Separation was achieved by electrophoresis for approximately 90min at 180V. Protein profiles of the gels were visualised by silver staining. All steps were performed with gentle shaking. After the electrophoresis, the gel was rinsed with ddH₂O and was incubated in fixing solution of 5% acetic acid and 50% methanol solution for 30min, at room temperature, followed by three brief washes in ddH₂O, and rehydration in ultrapure H₂O overnight at 4°C. The next day, gels were incubated in sensitising solution (1.25 mM sodium thiosulphate 5-hydrate) for 1min, washed for 1min with H₂O for three times and stained in 6mM silver nitrate solution for 30min at 4°C. After three brief washes in H₂O, gels were incubated in developing solution (280mM sodium carbonate, 5% formaldehyde) for approximately 5min until the protein bands are clearly discernable. Gels were washed in 5% acetic acid for 10min to stop the developing, followed by three washes in H₂O and placed in clear plastic bags to be scanned using a calibrated scanner (GS-800, Bio-Rad). Each gel lane was divided in twelve sections of various sizes, according to the amount of protein as judged by Silver staining, using a scalpel, and transferred to a 96-well plate for automated de-staining and tryptic digestion, using an Investigator ProGest (Digilab) robotic digestion system according to the manufacturer’s protocol. The obtained peptide solution was frozen at -80°C for at least 2 h and lyophilised under vacuum at -55°C for approximately 5h in a Christ Alpha 1-2 LD Freeze Dryer. The lyophilised peptides were re-suspended in 40 µl of 2% acetonitrile, 0.05% trifluoroacetic acid (TFA) in HPLC-grade H₂O and frozen at -20°C until mass spectrometric analysis.
3.20 In-solution digestion and HPLC-MS/MS analysis

After the preliminary results of the in-gel digestion were assessed an in-solution digestion was used due to the low number of proteins that were pulled-down. The eluates were denatured by the addition of 9M urea, 2M thiourea in a 1:2 ratio (final concentration 6M urea, 2M thiourea) and were vortexed briefly. Samples were then reduced by adding 100mM dithiothreitol (DTT; final concentration 10mM) followed by incubation at 37°C for 1 hour shaking at 600rpm. The samples were then cooled down to room temperature before alkylation of proteins using 500mM iodoacetamide (IAA, final concentration 50mM) followed by incubation in the dark for 1h at room temperature. Proteins were then precipitated using pre-chilled (-20°C) acetone in a 6:1 volume ratio, overnight. The next day, samples were briefly vortexed and centrifuged at 0°C for 40min. the protein pellets were fully dried using a SpeedVac Concentrator (ThermoFisher Scientific, Savant SPD131DDA) for approximately 30min. For protein digestion, samples were re-suspended in 40µl trypsin solution, containing 0.01µg/µl trypsin in 0.1 M triethylammonium bicarbonate (TEAB), pH 8.5, assuming 20µl, before briefly vortexing and centrifuging them. Proteins were digested overnight at 37°C under agitation (240rpm). The digestion was stopped by acidification of the samples with 10% v/v trifluoroacetic acid (TFA; final concentration 1%). Peptide samples were purified using a 96-well C18 spin plate (MicroSpin, Harvard Apparatus, cat. no. 745617). The resin was activated using 200 µl methanol and centrifuged at 1000 x g for 1 minute. Wash steps included 200 µl of 80% acetonitrile (ACN), 0.1% TFA in H₂O, and three equilibration steps using 200µl of 1% ACN, 0.1% TFA in H₂O with centrifugation (1000 x g for 1 minute) after each step. Samples were loaded onto the resin and centrifuged at 2250 x g for 1 minute; the flow through was reloaded onto the resin a second time and centrifugation repeated. The resin was then washed three times with 200µl 1% acetonitrile (ACN), 0.1% TFA in H₂O (centrifugation at 2250 x g for 1 minute). Finally, the samples were eluted with 170 µl of 50% acetonitrile, 0.1% TFA in H₂O (centrifugation at 1000 x g for 1 minute); this step was repeated, combining the collected eluate. Samples were then dried down using the SpeedVac Concentrator for approximately 5 hours. Dried peptide samples were reconstituted with TFA (0.05%) in ACN (2%) (Thermo Scientific, 51101) and separated by a nanoflow LC system (Dionex UltiMate 3000 RSLC nano). Samples were injected onto a nano-trap column (EASY-Spray PepMap® RSLC C18, 2µm 100Å, 75µm x 50cm) (Thermo Scientific, ES803), at a flow rate of 25 µL/min for three minutes, using 2% ACN, 0.1% FA in H₂O.
The following nano-LC gradient was then used to separate the peptides at 0.25 µL/min: 0–5 min, 4–10% B; 5–75 min, 10–30% B; 78–80 min, 30–40% B; 85–89 min, 99% B; 90–120 min 4% B, where A=0.1% formic acid (FA) in H₂O, B=80% ACN, 0.1% FA in H₂O. The nano column (Acclaim® PepMap100 C18, 50 cm x 75 µm, 3µm, 100 Å) was set at 45°C and coupled to a nanospray source (Picoview, New Objective, US). The temperature of the analytical column was set at 45°C and connected to an EASY-Spray ion source (Thermo Scientific, ES081). Spectra were collected from an Orbitrap mass analyser (Q Exactive HF, Thermo Scientific) using full Mass Spectrometry (MS) mode (resolution of 60,000 at 200 m/z) over the range 350–1600 m/z. Data-dependent MS2 scan was performed using Quadrupole isolation in Top15 mode using higher energy collision dissociation (HCD) fragmentation in each full MS scan (resolution of 15,000 at 200 m/z) with dynamic exclusion enabled. Data-dependent MS2 scan was performed using the top fifteen ions in each full MS scan (resolution of 17,500 at 200 m/z) with dynamic exclusion enabled. Proteome Discoverer software (Thermo Scientific, version 2.2.0.388) was used to search the raw data files against the mouse database (mouseSP201701, 16,844 protein entries) using Mascot (version 2.6.0, Matrix Science).

The mass tolerance for precursor ions and fragment ions was set at 10 ppm and 20 mmu, respectively. Trypsin was used as the digestion enzyme with up to two missed cleavages allowed. Variable modifications included acetylation of N-terminus, oxidation of methionine residues and deamination of asparagine in the presence of O₁₈ water. MS/MS-based peptide and protein identifications were validated after filtering for a peptide probability of greater than 95%, a protein probability of greater than 99% and at least two proteins per sample group. Data was normalised to the total peptide amount to take into account inter-sample variations in abundances and scaled on all average. The normalised total precursor intensity was used for quantification.

### 3.21 Software analysis of proteomic data

For the analysis of the interactions of the proteins that were identified with Mass Spectrometry a set of the software were used.
STRING (Search Tool for the Retrieval of Interacting Genes/Proteins) software provides information on known and predicted protein-protein interactions based on experimental evidences and multiple other sources. An additional advantage is that the interactions are presented according their confidence following statistical advantage.

Cytoscape, an open source software, employs protein-protein interaction data from large databases and provides information on the functionality of the proteins, their interaction networks and link the networks to databases of functional annotations. Moreover, Cytoscape offers a comprehensive visualisation of the interactions.

Lastly, Reactome, an open-source and curated database of pathways, integrates user-supplied data as well as data from a set of databases to perform a pathway and reaction analysis combined with statistical analysis of the results.

3.22 Immunoblotting

Western Blot analysis was used for the validation of the Mass Spectrometry data. Eluate from RNA Pull-down was separated on a pre-cast 10-well NuPAGE Novex 4-12% gradient Bis-Tris polyacrylamide gels (Life Technologies) as described above in 3.18. For immunoblotting, proteins were transferred from the polyacrylamide gel to a nitrocellulose membrane in ice-cold transfer buffer (25mM Tris base and 200mM glycine dissolved in 20% methanol) for 2h at 350mA. Proteins were then stained with Ponceau S solution (SIGMA, cat. no. P7170) to confirm adequate transfer. After washing in TBS with 0.1% Tween-20 (TBS-T), membranes were blocked with 5% milk in TBS-T for 1h at room temperature. Membranes were washed briefly in TBS-T and incubated overnight in primary antibody solution (see list below, Table 2) at a 1:1000 dilution in 5% bovine serum albumin (BSA; Sigma, cat. no. A9418) in TBS-T at 4 °C while rotating. Membranes were then washed thrice for 10min in TBS-T, followed by incubation with a light chain specific horseradish peroxidase (HRP)-conjugated secondary antibody (Dako) to detect the primary antibody, in a 1:5000 dilution in 5% fat-free milk in TBS-T for 1h at room temperature. After three washes in TBS-T (10min each), membranes were developed using enhanced chemiluminescence (ECL, GE
Healthcare) on a Xograph processor with exposure time dependent on signal intensity.

Table 2. Primary antibodies used for immunoblotting.

<table>
<thead>
<tr>
<th>Primary antibody (species)</th>
<th>Company, catalogue no.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Larp1 (Rabbit)</td>
<td>Proteintech 13708-1-AP</td>
</tr>
<tr>
<td>IGF2BP2 (Rabbit)</td>
<td>Proteintech 11601-1-AP</td>
</tr>
<tr>
<td>hnRNPE1 (Rabbit)</td>
<td>Proteintech 145231-AP</td>
</tr>
<tr>
<td>DHX9 (Rabbit)</td>
<td>Proteintech 177211-AP</td>
</tr>
<tr>
<td>EIF2AK2 (Rabbit)</td>
<td>Proteintech 182441-AP</td>
</tr>
<tr>
<td>ILF3 (Rabbit)</td>
<td>Proteintech 198871-AP</td>
</tr>
<tr>
<td>Tra2b (Rabbit)</td>
<td>Abcam ab31353</td>
</tr>
</tbody>
</table>

3.23 Statistical analysis

Statistical analyses of the levels of the miRNAs of different clusters upon editing were performed with the ANOVA test with Dunnett post hoc test. Results are given as mean±SD. A value of $P<0.05$ was considered significant. Microsoft® Excel® (version 14.0.7208.5000) and GraphPad Prism (version 7.00) were used for statistical calculations. MS data were quantified using normalised total intensities and analysed as described previously. All other statistical analyses were performed with Microsoft® Excel® (version 14.0.7208.5000) and GraphPad Prism (version 7.00). FDR correction was used to correct for multiple testing for the statistical analysis of the Mass Spectrometry data. P-values<0.05 were considered significant with * signifying a p-value<0.05 and ** a p-value<0.01.
Chapter 4 EVALUATION OF THE DELIVERY METHOD FOR CRISPR/Cas9 AND ITS EFFICIENCY IN EDITING THE TARGETED SITES

4.1 Introduction

The extended miR-15 family comprises of 10 members (miR-107, miR-103, miR-15a, miR-15b, miR-16, miR-195, miR-497, miR-503, miR-424 and miR-646) of miRNAs which share an identical seed sequence (Figure 1.4).

The members of the family are clustered in three separate chromosomes, with miR-195, the miRNA of interest of the current study, being clustered together with miR-497. The inhibition of miR-195, with the use of antimiRs/antagomiRs, is not precise, due to the high homology of sequence that the two miRNAs of the same family share and their identical seed sequence, a feature that leads to the non-specific downregulation of multiple members of the same family.

In this project, CRISPR/Cas9 system was employed to edit the genomic locus of miR-195 to assess whether genome editing can result in specific inhibition of the miRNA without affecting the levels of the other members of the family. CRISPR/Cas9 has emerged as a fast, easy-to-use and highly specific method for editing DNA in eukaryotic cells although there are limitations, concerning the delivery of the system to cells, that need to be overcome. The objective of the study was to take advantage of the CRISPR/Cas9 system to target the miR-195 genomic locus in primary cells with high specificity. By using sgRNAs designed to target specific sites in the stem loop of miR-195, a part of the genomic locus of the miR-195 was edited, a change that was expected to impair Drosha’s ability to process the stem loop. In more detail, the aims were to:

- Establish a reliable workflow for miRNA inhibition using genome editing by:
  - Designing in silico and testing sgRNAs that target and edit miRNA genes,
  - Identifying the optimal delivery method for CRISPR/Cas9 in primary cells and
Determining the appropriate conditions for achieving high editing efficiency

- To genetically engineer, with the use of CRISPR/Cas9, stable cell lines of SMCs that harbour deletions of miR-195 and

- To assess putative off target mutagenesis of the guides tested.

The results presented in this chapter demonstrated that transfection of mouse VSMCs stably expressing Cas9 with IVT guides lead to efficient editing of the mmu-miR-195a genomic locus and downregulation of the levels of the miRNA.
4.2 Experimental Design

The present study aimed to explore the use of CRISPR/Cas9 in editing miRNA genes in primary cells in order to investigate their regulation and function. The focus of it was miR-195, a miRNA previously shown to participate in the onset and development of AAA. As an initial step, different strategies were tested to generate genetically engineered SMCs that harboured a deletion for miR-195. To achieve sufficient editing efficiency and take advantage of the latest technological developments in the field, a total of four experimental strategies were employed:

(A) Plasmid delivery of the CRISPR/Cas9 system,

(B) Lentiviral delivery of the CRISPR/Cas9 system,

(C) Cas9 ribonucleoproteins (Cas9/RNPs) and

(D) Transfection of VSMCs stably expressing Cas9, with IVT sgRNAs.

Firstly, after designing guides targeting the desired sites *in silico*, the plasmid delivery was assessed in human cells. By using plasmids, which harboured the corresponding guide sequences and the Cas9 gene, we tested their ability to edit the targeted site and their efficiency. The efficiency of editing was also assessed at different time points. By doing so, potential effects on the viability of the cells, as a result of the transfection or the editing, were assessed. Genomic DNA was also sequenced to identify the mutation patterns. However, the obtained results suggested a rather low efficiency and led to undertaking a different delivery technique.

As a second strategy, lentiviral particles were used to infect primary cells. Since the lentiviral infection results in the integration of the sequence of interest into the genome, we expected to have more efficient gene editing of the miR-195 locus. The second strategy was assessed in two different human cell types, HUVECs and human SMCs, with the results suggesting a substantial editing efficiency, and mouse VSMCs
with the efficiency reaching significant levels. However, the strategy was deemed as non-optimal because of the potential high off-target mutagenesis, due to the continuous presence of Cas9/RNP complexes in the cells, as well as due to the limitations of the lentiviral delivery.

To address these issues, a third strategy was employed. Transient transfection of cells with Cas9/RNPs was performed. Commercially available recombinant Cas9 proteins were used along with IVT sgRNA designed to target the human miR-195 and assessed in in vitro experiments. Additionally, HEK293T cells were transfected with Cas9/RNP complexes with the editing efficiency not reaching sufficient levels.

Furthermore, mouse VSMCs were infected with lentiviral particles that harboured only the Cas9 gene, thus generating stable cell lines that were stably expressing the Cas9 protein. The cells were then transfected with IVT guides. The approach was thought as the most suitable to overcome the limitations of transfecting cells, an approach that led to low efficiency of editing. After concluding to our approach of choice, the levels of the mature miRNA were assessed in mouse VSMCs to determine the impact of gene editing in the mmu-miR-195 inhibition. The effect on the expression levels of other members of the miR-15 family was also evaluated.

Finally, possible off-target mutagenesis was studied, with the experiments showing that there were no off-target sites for the guides used at the in silico predicted sites.
4.3 Results

4.3.1 Gene Editing in Human Cells

4.3.1.1 In Silico sgRNA Design for hsa-miR-195 and cloning

The CRISPR DESIGN Tool (http://crispr.mit.edu, Zhang Lab, MIT) was used to design the guides for the genomic locus of miR-195 (Figure 4.1). The DNA sequence corresponding to the locus of the miRNA gene was used as input sequence to the software. The following oligos, which targeted sites in the miR-195 stem loop (Figure 4.2), had the highest score:

sg195g1(G1): 5'-GGAAGCGAGTCTGCAATAT-3'

sg195g2(G2): 5'-TTGAGGCGAGTCTGCAATAT-3'
Figure 4.1 In silico design of the single-guide RNA for hsa-miR-195 locus. The guides used for the human miR-195 locus were designed with the help of the CRISPR DESIGN Tool (http://crispr.mit.edu, Zhang Lab, MIT). The guide sequence is depicted, with the corresponding score and the 5 most likely off target sites for guide (A) sg195G1 and (B) sg195G2.
Figure 4.2 Schematic representation of the editing of the hsa-miR-195 genomic locus. (A) The hsa-miR-195 genomic locus with the corresponding elements. The location of the 5p arm of the miRNA is highlighted in red and the location of the 3p arm of the miRNA is highlighted in blue. The SspI restriction enzyme sites are depicted as squares. (B) Schematic of the miR-195~497 cluster and the location of the DSBs. (C) The genomic DNA sequence of the miR-195 locus. The mature miR-195 is underlined and the seed sequence is highlighted with a red rectangular. The target sites of the two guides are shown with their corresponding PAM sequences highlighted in green. The positions, where the DSB occurred, are depicted with a purple triangle. G1: sg195G1 and G2: sg195G2.
4.3.1.2 Strategy A: Plasmid Delivery

To determine the efficiency of the selected guides, HEK293T cells were transfected with pX330 plasmids harbouring the sgRNAs (pX330G1 and pX330G2). PCR amplification of the genomic locus revealed two bands in the samples that were transfected with both plasmids (Figure 4.3A), a long one corresponding to the full size and a shorter one corresponding to a truncated product. The latter was a result of the simultaneous editing of the locus by both guides and the perfect re-ligation of the strands which resulted in the deletion of a 105bp long fragment as confirmed by sequencing. The gene editing efficiency of each treatment was assessed with the T7EI assay (Figure 4.3B). Each of the single guides demonstrated an efficiency which varied from 24-27% while the combination of both guides showed an efficiency of 34%. Editing efficiency did not change between Day2 and Day5 post transfection. Sanger sequencing was used to identify the mutation patterns. In total, a variety of patterns were observed: (i) DNA that was not edited, (ii) DNA that had a deletion of a 105bp long sequence upon editing from both guides and perfect re-ligation of the DNA strands, (iii) DNA with short indels at the site where the DSB occurred and (iv) DNA with a 160bp long deletion and an insertion starting at the site of G1 (Figure 4.3E).
Figure 4.3 Gene editing of the miR-195 genomic locus using plasmid vectors. (A) HEK293T cells were transfected with the pX330 plasmid harbouring Cas9 and sgRNA1 or sgRNA2 (pX330G1 and pX330G2, respectively). On Day 2 (D2) and Day 5 (D5) post transfection, DNA was harvested and the genomic locus of miR-195 was PCR amplified. Agarose gel electrophoresis revealed two amplicons of different size in the cells that were transfected with both guides simultaneously, one corresponding to the full-size product (577bp) and one to the truncated product (471bp). (B) T7EI assay was used for assessing the gene editing efficiency that ranged from 18-34%. (C) Blue and white colony screening was used to pick the colonies that were successfully transformed with the PCR amplicons. (D) PCR amplification of the isolated plasmids indicated the differences in the size of the cloned sequences. (E) Nine colonies were selected according to the size of the PCR amplicon and were Sanger sequenced. Dashes represent deletions and insertions are highlighted in blue. G1: sg195G1 and G2: sg195G2, n=1.
4.3.1.3 Strategy B: Lentiviral Delivery in Human Primary Cells

4.3.1.3.1 Gene Editing in HUVECs

To facilitate the delivery of the CRISPR/Cas9 system to primary cells lentiviral particles were used. HUVECs were infected simultaneously with lentiviral particles harbouring both sgRNAs (LEG1 and LEG2).

Titration experiments with PLUS™ Reagent, a reagent that enhances cationic lipid-mediated transfection of DNA into cultured eukaryotic cells, were conducted. Indeed, high editing efficiency was observed based on the intensity of the bands that corresponded to the truncated product (Figure 4.4, lanes denoted with asterisks). The T7EI confirmed the high editing efficiency. Moreover, the fact that, after the T7EI only 4 distinct bands were observed indicates perfect re-ligation of the DNA bands while, the marginally higher editing efficiency in the case PLUS™ Reagent was not used, demonstrated the efficient delivery of the system (Figure 4.4).

Next, in three independent experiments PCR amplification of the locus revealed two bands for the edited cells with the shorter one corresponding to the truncated product in accordance with the results for the plasmid delivery (Figure 4.5(A)). The T7EI assay revealed editing efficiency varying from 81-96% (Figure 4.5(B)). Sanger sequencing revealed 3 different mutation patterns: (i) DNA that had a deletion of a 105bp long sequence upon editing from both guides and perfect re-ligation of the two strands and (ii) DNA with short indels at only one or (iii) both sites of the DSB (Figure 4.5(C)).

The pattern of the DNA bands, after the T7EI digestion, differs among the samples infected with only one of the two guides and those infected with both guides. In the latter occasion, only four bands are observed: (i) the band corresponding to the full sequence, (ii) the band of the truncated sequence, (iii) the band corresponding to a 359bp long sequence and (iv) the band of the 112bp long sequence while the 465bp and 218bp long bands, which are noticed in the treatments with either LEG1 or LEG2, are not detected. These differences are due to the simultaneous cleavage of the sequence by the two guides which result in the generation of a diverse cleaving pattern (Figure 4.6).
Figure 4.4 T7 Endonuclease I assay for HUVECs upon lentiviral infection. HUVECs were infected with lentiviral particles harbouring Cas9 and sg195g1 or sg195g2 (LEG1 and LEG2, respectively). On Day 3 post infection, DNA was harvested and the genomic locus of miR-195 was PCR amplified. In the cells that were infected with both guides simultaneously PCR revealed two bands, one corresponding to the full-size product (577bp) and one to the truncated product (471bp), derived from perfect re-ligation of the genomic DNA following the two DSBs, n=2.
Figure 4.5 Lentiviral delivery of Cas9 and sgRNA. (A) HUVECs were infected with lentiviral particles harbouring Cas9 and sg195g1 or sg195g2 (LEG1 and LEG2, respectively). On Day 3 post infection, DNA was harvested and the genomic locus of miR-195 was PCR amplified. In the cells infected with both guides simultaneously PCR revealed two bands, one corresponding to the full-size product (577bp) and one to the truncated product (471bp). (B) The T7EI assay revealed an edited efficiency of 81-96%, in three independent experiments. (C) Mutation screening was assessed using Sanger sequencing. Dashes represent deletions, insertions are highlighted in blue. Exp=Experiment, (n=3).
Figure 4.6 Schematic representation of the digestion products upon editing with guides sg195g1 and sg195g2. Upon T7EI digestion of the miR-195 amplicon (A) gene editing using the LEG2 is predicted to generate two fragments of 112bp and 465bp, (B) gene editing using the LEG1 is predicted to generate two fragments of 218bp and 395bp and (C) gene editing using the LEG12 and LEG2 followed by perfect re-ligation will generate a truncated product consisting of the two fragments of 1112bp and 359bp.
4.3.1.3.2 Gene Editing in human SMCs

Human SMCs were infected with lentiviral particles harbouring the designed sgRNAs (LEG1 and LEG2). PCR amplification of the locus revealed two bands. Based on T7EI assay gene editing efficiency that varied from 70-80% was identified (Figure 4.7).

Figure 4.7 T7 Endonuclease I assay for hSMCs upon lentiviral infection. Human SMCs were infected with lentiviral particles harbouring Cas9 and sg195g1 or sg195g2 (LEG1 and LEG2, respectively). On Day 3 post infection, DNA was harvested and the genomic locus of miR-195 was PCR amplified. The cells infected with both guides simultaneously revealed, after PCR amplification, two bands, one corresponding to the full-size product (577bp) and one to the truncated product (471bp), (n=3).
4.3.1.4 Strategy C: Cas9/RNPs in HEK293T cells

As a third approach, Cas9/RNPs were used. *In vitro* transcribed guide G2 was used to form Cas9/RNP complexes. In order to test for the ability of the Cas9/RNP complex to cleave the targeted DNA sequence, an *in vitro* Cas9 digestion was performed which showed that Cas9 is able to cleave the PCR template when in complex with guide G2. The sizes of the two observed bands corresponded to the position of the cleavage by guide G2 and the digestion was complete (Figure 4.8A). Additionally, HEK293T cells were transfected with Cas9/RNPs to study the complexes’ ability to recognise and cleave the targeted DNA in cells. T7E1 showed an editing efficiency of 25% (Figure 4.8B).

![Figure 4.8](image)

**Figure 4.8 Assessment of Cas9/RNPs approach in HEK293T cells.** (A) Cas9/RNPs complexes were generated by incubating Cas9 protein (0.5μg) with *in vitro* transcribed guide sg195g2 (G2), (125ng). The Cas9/RNPs were tested *in vitro* for their ability to recognise and cut the targeted DNA. (B) HEK293T cells were transfected with Cas9/G2 RNP complexes and the editing efficiency was assessed using T7EI assay. N.D: Not Detected, (n=3). Representative results from only one experiment are shown.
4.3.1.5 Off-target Mutagenesis

Mutagenesis, in the most likely off-target sites (OFT), was assessed using the T7EI assay. DNA from lentiviral infected HUVECs that showed the highest editing efficiency was used. T7EI assay did not show any off-target editing. The non-specific bands, of shorter length, that were observed in some of the samples (Figure 4.9), did not result from gene editing, since they were observed in both the edited and unedited cells and their sizes did not correspond to the off-target cleavage of the LEGuides. OFT 2-5 for guide LEG1 and OFT 1, 2 and 4 for guide LEG2 were tested.

![Figure 4.9 Assessment of possible off-target mutagenesis upon lentiviral infection of HUVECs.](image)

The most likely off-target sites for both guides (OFT 1-5), as suggested by the CRISPR DESIGN Tool (http://crispr.mit.edu, Zhang Lab, MIT), were assessed by T7EI assay. No off-target mutagenesis was detected for neither of the LEG1 and LEG2. OFT: Off-target, N.D.: Not Detected, n=1.
4.3.2 Gene Editing in Mouse Cells

4.3.2.1 In Silico sgRNA Design for mmu-miR-195 and cloning

The CRISPR DESIGN Tool (http://crispr.mit.edu, Zhang Lab, MIT) was used to design the guides targeting the mmu-miR-195 genomic locus (Figure 4.10). As an input, the DNA sequence of the locus of the miRNA was used. Three oligos, which targeted sites within the miRNA sequence (Figure 4.11 and Figure 4.12), with the highest scores, were selected:

sg195m2 (Gm2): 5’-TGGAGCAGCAGGCAATAT-3’
sg195m4 (Gm4): 5’-CAGCACAGAAATATTGGCAT-3’
sg195mA (GmA): 5’-ATTTCTGTGCTGCTAGG-3’
Figure 4.10 In silico design of the single-guide RNA for the mmu-miR-195a locus. The guides used for editing the mouse miR-195 locus were designed with the help of the CRISPR DESIGN Tool (http://crispr.mit.edu, Zhang Lab, MIT) on-line tool. The guide sequence is depicted, with the corresponding score and the 5 most likely off target sites for guide (A) sg195m2, (B) sg195Gm4 and (C) sg195GmA.
Figure 4.11 Schematic representation of the editing of the mmu-miR-195 genomic locus. (A) The mmu-miR-195 genomic locus. The location of the 5p arm of the miRNA is highlighted in red and the location of the 3p arm of the miRNA is highlighted in blue. DSBs are depicted as purple triangles and SspI and MwoI restriction sites are depicted as squares. (B) Schematic of the mmu-miR-195-497 cluster. (C) Part of the DNA sequence of the miR-195 locus. The DNA sequence of the mature mmu-miR-195 is underlined. The target sites of the guides are shown with their corresponding PAM sequences highlighted in green. The positions where the DSB occurs are depicted as purple triangles. Gm2: sg195m2, Gm4: sg195m4 and GmA: sg195GmA.
Figure 4.12 Schematic representation of the secondary structure of the mmu-miR-195a stem loop with guides. Representation of the mmu-miR-195a stem loop with the detailed sequence, bulges, mismatches and apical loop along with the targeted DNA sequence and the PAM sequence for (A) sg195m2, (B) sg195m4 and (C) sg195mA. PAM sequence is highlighted in blue. The mature miRNA sequence is highlighted in purple and grey arrows indicate the corresponding position of the DSB on the stem loop which is the site of mutations induced by random indels following editing of the genomic locus in mutant cells.
4.3.2.2 Strategy B: Lentiviral delivery in mouse VSMCs

4.3.2.2.1 Lentiviral particles Gm2 and Gm4

Mouse VSMCs were infected with lentiviral particles harbouring the designed guides Gm2 and Gm4 (LEGm2 and LEGm4). On Day 3 after the infection, gene editing efficiency was assessed by T7EI assay and it varied from 11-44%, (Figure 4.13).

Figure 4.13 T7 Endonuclease I assay for mouse SMCs upon lentiviral infection with guides Gm2 and Gm4. Mouse SMCs were infected with lentiviral particles harbouring Cas9 and sg195m2 or sg195m4 (LEGm2 and LEGm4, respectively). On Day 3 post infection, DNA was harvested and the genomic locus of mmu-miR-195 was PCR amplified. T7EI assay was used to assess the efficiency of the editing in two independent experiments. Exp=Experiment, n=2.
Figure 4.14 Schematic representation of the sizes of T7EI digestion products upon editing with guide Gm4 and Gm2. T7EI digestion of the mmu-miR-195 amplicon upon editing with (A) guide LEGm4 is predicted to generate two fragments of 204bp and 326bp, (B) guide LEGm2 is predicted to generate two fragments of 228bp and 302bp and (C) both guides simultaneously, followed by perfect re-ligation, will generate a truncated product consisting of two fragments of 204bp and 302bp.
4.3.2.2.2 Lentiviral particles Gm2 and GmA

Mouse VSMCs were infected with lentiviral particles harbouring the designed sgRNAs (LEGm2 and LEGmA). PCR amplification of the DNA harvested from the cells revealed two bands for the infected mouse VSMCs while T7EI assay showed a gene editing efficiency of 32% and 26% for LEGm2 and LEGmA respectively and 48-51% for cells infected with both guides simultaneously (Figure 4.15). The pattern of the DNA bands, after the T7EI digestion, differs according to the treatment. When mouse SMCs were infected with both guides at the same time, four bands can be observed: (i) a band of the full sequence (530bp), (ii) a band of the truncated sequence (484bp), (iii) a band corresponding to a 302bp long sequence and (iv) a band of the 182bp long sequence. These differences in the pattern between the single use of the guides and the infection with both together are the result of the simultaneous cleavage of the genomic locus by the two guides simultaneously (Figure 4.16).

Figure 4.15 T7 Endonuclease I assay for mouse VSMCs upon lentiviral infection with guides Gm2 and GmA. Mouse VSMCs were infected with lentiviral particles harbouring Cas9 and guides sg195m2 or sg195mA expression (LEGm2 and LEGmA, respectively). On Day 3 post infection, DNA was harvested and the genomic locus of mmu-miR-195 was PCR amplified. The T7EI assay was used for assessing the efficiency of the editing, n=2.
Figure 4.16 Schematic representation of the T7EI digestion products upon editing with guide GmA and Gm2. Upon T7EI digestion of the mmu-miR-195 amplicon (A) gene editing using the LEGmA is predicted to generate two fragments of 182bp and 348bp, (B) gene editing using the LEGm2 is predicted to generate two fragments of 228bp and 302bp and (C) gene editing using the LEGmA and LEGm2 followed by perfect re-ligation will generate a truncated product consisting of two fragments of 182bp and 302bp.
4.3.2.3 Strategy D: Transfection of mouse VSMCs stably expressing Cas9 with sgRNAs

Mouse VSMCs stably expressing Cas9 were generated by infecting cells with lentiviral particles harbouring Cas9. Following infection, cells were selected with Puromycin treatment (2µg/ml) for 48h, (Figure 4.17).

![Lenti CRISPR v2 - + + + Cas9 B-actin](image)

*Figure 4.17 Cas9 expression in VSMCs infected with LentiCRISPRv2 lentiviral particles.* Lysates from mouse VSMCs infected with lentiviral particles harbouring Cas9 were harvested and blotted with antibody against the protein (ThermoFisher Scientific, #MA5-23519). Four different lines are presented. B-actin was used as a normalisation control. For comparison, a similar exposure time was applied.

Next, the ability of the Cas9/RNP complexes to form and cleave the targeted DNA sequence was tested by an *in vitro* Cas9 digestion. *In vitro* transcribed guides were used to form Cas9/RNP complexes with the digestion showing that Cas9 is able to cleave the PCR template when in complex with either Gm2 or GmA. The sizes of the observed bands corresponded to the position of the cleavage by the Gm2 or GmA and the digestion was nearly complete (Figure 4.18A). Additionally, mouse VSMCs infected with lentiviral particles harbouring Cas9, were transfected with IVT guides to assess the ability of the single guide RNAs to form complexes with the Cas9, recognise and cleave the targeted DNA in cells. Two conditions were tested, cells infected with lentiviral particles harbouring Cas9 but not treated with puromycin and selected cells. T7EI showed that only the selected cells had successful editing of the targeted locus with an efficiency of 22% (Figure 4.18B).
Figure 4.18. Assessment of gene editing in mouse VSMCs stably expressing Cas9. (A) Cas9/RNPs complexes were generated by incubating Cas9 protein (0.5μg) with in vitro transcribed guide (125ng). The Cas9/RNPs were tested in vitro for their ability to recognise and cut the targeted DNA. (B) Mouse VSMCs, stably expressing Cas9, were transfected with IVT guides GmA or Gm2. Two days post transfection, DNA was harvested and the genomic locus of mmu-miR-195 was PCR amplified. The T7EI assay was used for assessing the efficiency of the editing of 22%. The digested products corresponded to the expected sizes as described above. Gm2: sg195m2 and GmA: sg195mA, (n=3). Representative results from only one experiment are shown.

Selected mouse VSMCs were then transfected simultaneously with IVT guides GmA and Gm2. DNA was extracted on two different days, Day2 (D2) and Day7 (D7) and the editing efficiency showing levels that were stable across the two timepoints (Figure 4.19).

Figure 4.19. T7EI assay for mouse VSMCs stably expressing Cas9 after transfection with IVT guides. Mouse VSMCs that were infected with lentiviral particles harbouring Cas9 were selected with puromycin treatment and compared to non-selected cells. Gm2: sg195m2 and GmA: sg195mA. n=3, representative results from only one experiment are shown.

The mmu-miR-195 locus has two SspI restriction sites (5’-AATATT-3’) with single guide Gm2 disrupting one of them and 3 MwoI restriction sites (5’-GCNNNNNNNGC-3’) with single guide GmA disrupting one (Figure 4.11A). Mouse VSMCs stably
expressing Cas9 were transfected with IVT guides and SspI and MwoI restriction digestions were performed for rapid assessment of gene editing. After the simultaneous or single transfection of mouse VSMCs stably expressing Cas9 with single guides Gm2 and GmA, digestion with SspI or MwoI restriction enzymes were used to rapidly assess the success of the editing. According to the disruption of the digestion sites that each guide conferred, different patterns were observed. Transfection solely with single guide Gm2 and digestion with SspI resulted in three bands corresponding to 331bp and 199bp after the successful editing of the targeted site and the disruption of one of the two SspI sites (as depicted above in Figure 4.11), and bands of 303bp and 199bp after no editing of the two sites resulting in digestion of both of them. In the latter case, the digestion resulted in a third band of 28bp which diffused in the agarose gel and is non-discernible. Upon transfection with single guide GmA, no SspI site got disrupted. Simultaneous transfection with both guides and digestion with SspI showed the same pattern as previously described (Figure 4.20A). The 530bp corresponded to non-complete digestion of the full-length PCR product. Upon digestion with MwoI restriction enzyme, multiple bands for the non-edited PCR product were observed, corresponding to 184bp, 177bp, 102bp and a band of shorter length of 67bp. After, transfection with single guide GmA, only one MwoI got disrupted upon successful editing, resulting in the presence of a band 286bp long. Upon editing with both guides simultaneously, the pattern changed as a result of the perfect re-ligation of the DNA and the digestion of a shorter PCR product which resulted to a 244bp (Figure 4.20B). In both cases, editing with both guides but not digestion with either restriction enzyme, revealed a 484bp long band, which corresponded to the perfect re-ligation of the DNA after editing with the guides (right panel of the two figures A and B).
Figure 4.20 Restriction enzyme digestion with SspI and MwoI. Mouse VSMCs stably expressing Cas9 were transfected with IVT Gm2 or GmA. On Day3 post transfection, DNA was harvested and the genomic locus of miR-195 was PCR amplified. (A) Upon SspI digestion, only cells transfected with both guides (denoted with two asterisks) showed an undigested product of 484bp. Cells transfected only with Gm2 had a disruption of one of the SspI sites (one asterisk). (B) MwoI digestion showed a different pattern for transfection with Gm2 together with GmA and Gm2 only transfected cells (denoted with asterisks). Gm2: sg195m2 and GmA: sg195mA, n=2.
4.3.2.4 Off-target mutagenesis in mouse VSMCs

4.3.2.4.1 Off-target mutagenesis upon lentiviral infection

Mutagenesis, in the most likely off-target sites (OFT) as returned by the CRISPR online tool, was assessed using the T7EI assay. DNA from lentiviral infected mouse VSMCs did not show any off-target editing. The non-specific bands, of shorter length, that were observed in some of the samples (Figure 4.21), did not result from gene editing, since they were observed in both the edited and unedited cells and their sizes did not correspond to the predicted site of off-target cleavage of the LEGuides. The five most likely off-targts were tested for each guide (Figure 4.21).

![Figure 4.21 Assesment of possible off-target mutagenesis upon lentiviral infection of mouse SMCs. The five most likely off-target (OFT 1-5) for both guides (A) Gm2 and (B) GmA, as suggested by the CRISPR DESIGN Tool (http://crispr.mit.edu, Zhang Lab, MIT), were assessed by T7EI assay. No off-target mutagenesis was detected for neither of the LEGm2 and LEGmA. OFT: Off-target, ND: Non Detectable, NS: Non Specific, n=1.](image-url)
4.3.2.4.2 Off-target mutagenesis upon transfection of mouse VSMCs stably expressing Cas9 with IVT guides

For the assessment of putative off-target mutagenesis the 7 most likely off-target sites (OFT) as returned by the CRISPR DESIGN Tool (http://crispr.mit.edu, Zhang Lab, MIT) and CRISPR Finder (Wellcome Trust Sanger Institute), was assessed using the T7EI assay. DNA from mouse VSMCs stably expressing Cas9 were transfected with IVT single guide Gm2 and did not show any off-target editing. The digested observed bands, of shorter length (Figure 4.22), did not result from gene editing, since they were observed in both the edited and unedited cells and their sizes did not correspond to the predicted site of off-target cleavage of the guides.

![Figure 4.22 Assessment of possible off-target mutagenesis upon editing with IVT Gm2. T7EI assessment of gene editing in the loci of putative off-target sites for Gm2 as predicted by the CRISPR DESIGN Tool (http://crispr.mit.edu, Zhang Lab, MIT) and CRISPR Finder (Wellcome Trust Sanger Institute). OFT; Off-Target, ND: Non-Detectable, NS: Non Specific. Gm2: sg195m2, n=1.](image-url)
4.4 Discussion

CRISPR/Cas9 is a novel, versatile gene editing technique that appears to be able to overcome the problem of the unspecific miRNA inhibition, a limitation that hinders the study of miRNAs. In the present study, we investigated whether genome editing, with the use of CRISPR/Cas9, can be employed for specifically inhibiting miR-195 in primary cells by testing four different approaches. Although gene editing has been used in the past for the knock-out of miRNAs, to the best of my knowledge, this is the first time the CRISPR/Cas9 is used in primary cells to inhibit miR-195.

4.4.1 Gene Editing in Human Cells

For the current project, the first step undertaken was to design guides to edit the miR-195 genomic locus. The objective was to target two sites in the genomic locus of the hsa-miR-195 stem loop, one close to the seed sequence and one further upstream of the stem loop of the miRNA. It has already been shown that a combination of sgRNAs can be used for editing a locus more effectively. However, the short length of the DNA sequence corresponding to the stem loop together with the PAM requirements of the CRISPR/Cas9 was an important limiting factor since the chances of finding 5'-NGG-3' sequences in such a short sequence were poor. Moreover, although it has been shown that Cas9 is able to induce DSBs in different sites in short distances within the same genomic locus and lead to the deletion of the flanked sequence, the short distance between putative targeted sites can interfere with the ability of Cas9 protein to physically approach the sequences. Considering results reported in previous studies along with the CRISPR requirements, the guides were designed accordingly. A sequence of a total length of 250nt, flanking the hsa-miR-195 gene, was used as an input to the CRISPR Design Tool and because of the need to target specific sites in the miRNA stem loop, the guides were chosen primarily based on the position of the site they targeted and not necessarily on their ranking score (Figure 4.1, Figure 4.2). Four different delivery strategies with different advantages and limitations were tested (Figure 4.23).
4.4.1.1 Strategy A: Plasmid Delivery

Plasmid delivery to HEK293T cells was used to induce gene editing to the hsa-miR-195 locus. Plasmid delivery would have only transient expression of the Cas9 protein thus minimising the effects of the protein accumulation that could possibly lead to off-target effects\textsuperscript{235}. Nevertheless, according to our results, the efficiency of the gene editing, as the T7EI assay revealed, was rather modest. The highest percentage of indels was achieved after transfection with plasmids harbouring both guides, simultaneously, reaching a 34% on Day2. These results come in accordance with previous studies reporting higher efficiency upon editing with two guides\textsuperscript{233, 236}. Although T7EI assay is a quantitative technique used for assessing successful editing that cannot provide qualitative information and is thought to marginally underestimate its efficiency due to innate limitations of the assay, it is regarded as the gold standard assay. The sizes of the bands after the T7EI indicated the sites where editing was induced. In more detail, T7EI is an in vitro assay that is used to assess the editing by employing the T7 endonuclease which is able to recognise and cleave mismatches in double stranded DNA. The digestion products have sizes that correspond to the site of the edit (Figure 3.3). Formation of homoduplexes between edited strands or incomplete digestion by T7 endonuclease underestimate the editing efficiency and limit the assay\textsuperscript{208}.

Next, the results from T7EI were confirmed by DNA sequencing, which showed that the editing occurred specifically in the sites targeted by the designed guides. In line with published results\textsuperscript{237}, editing of the hsa-miR-195 genomic locus resulted in the formation mainly of deletions, rather than insertions, while, in the case of transfection with both guides simultaneously, either the formation of indels in both targeted sites
or the excision of the sequence flanked by the designed guides and perfect re-ligation of the DNA strands (Figure 4.3E). Further to this, it has already been reported that two guides can be used for the excision of DNA sequences as was observed in the current project\textsuperscript{200}. Of note is the fact that although the predicted length of the excised sequence was 106bp, Sanger sequencing revealed that in 50% of the blue and white colonies screened (data not shown) the length of the excised sequence was in fact 105bp. Although it could be the result of an insertion of an adenine following the editing, it seems more likely to be the result of Cas9 cutting 4nt upstream of the PAM sequence instead of 3nt, as was expected, in either of the two targeted sites. Although this discrepancy has been reported\textsuperscript{179}, due to sequencing limitations, there was not a way of deciphering at which site that discrepancy occurred. In more detail, both of the two trinucleotide sequences upstream of the PAM sequences of the two guides used are followed by an adenine. In both cases the expected outcome would have been the adenine to be excised although sequencing data did not confirm that. A single unexpected adenine was identified in all of the sequences where the 105nt excision occurred although, based on these results, we can conclude in which case DSB happened 4nt upstream of the PAM sequence. However, taking into consideration the high percentage of the transfection (data not shown), that was observed and the relatively modest percentage of the gene editing, the plasmid delivery did not appear as a favourable delivery system in primary cells. Nevertheless, plasmid delivery was confirmed as a suitable delivery approach to quickly test for the ability of the designed guides to edit the sequence of choice in human cells.

4.4.1.2 Strategy B: Lentiviral Delivery

In order to address the low efficiency of the plasmid delivery, lentiviral delivery of the CRISPR/Cas9 was chosen as an alternative approach and it was tested in primary human cells. Previous studies on HUVECs, using lentiviral particles, demonstrated that it is a powerful method\textsuperscript{238} and taking advantage of the high proliferation rate of the cells, preliminary experiments were conducted. It has been reported that CRISPR/Cas9 leads to higher editing efficiency in dividing cells\textsuperscript{203}. Upon division and more specifically during the mitotic (M) phase of the cell cycle, the nuclear envelope breaks down due to the karyokinesis. During this, the CRISPR/Cas9 system can more easily access the DNA and edit it. Since initial experiments were conducted in human cells, there was no need to design new guides. Firstly, optimisation experiments were
used to identify the optimal conditions of infection (Figure 4.4) and results from three independent experiments suggested that infection with lentiviral particles resulted in almost complete gene editing as shown in Figure 4.5. PCR amplification showed two bands of different sizes with the intensity of the shorter one, corresponding to the size of the truncated DNA upon successful editing, indicating high efficiency (Figure 4.5A). Indeed, T7EI confirmed that lentiviral infection had high levels of editing with the efficiency reaching a 98% occurrence (Figure 4.5B). These results can be attributed to both the high proliferation rate of the cells as well as the efficacy of the delivery. DNA sequencing established that editing occurred in the targeted sites while the cells showed the same editing patterns as with plasmid delivery (Figure 4.5C). According to our results, we observed a high percentage of targeted genomic deletions, while in the instances when the flanked sequence was not excised, indels were observed in the cleaved sites, in agreement with previous findings\textsuperscript{239}. Interesting are the changes of the pattern of the digested bands following T7EI digestion with the sizes of the bands showing as a mix following the editing with the two guides alone (Figure 4.5B). This can be explained by the fact that when editing with both guides occurred, a DNA sequence is excised and the DNA strands are perfectly re-ligated, not leading to any cleavage by T7EI, while editing at the corresponding sites with the two guides at different timepoints induce the formation of indels that upon cleavage resulted in fragments of different lengths, unique to this treatment (Figure 4.6).

Following the preliminary results on HUVECs and given that the study aimed at using human SMCs, since the apoptosis of SMCs is a key component in triggering the AAA, the system’s efficiency was investigated in human SMCs. Although lentiviral infection confirmed the high editing efficiency, with a percentage of 80% (Figure 4.7), miRNA quantification (data not shown) showed that the levels of the miR-195 in prolonged cell culture conditions were extremely low, thus leading to the conclusion that the human SMCs would not be an appropriate model for the study. Therefore, mouse VSMCs were used.
4.4.1.3 Strategy C: Cas9/RNPs

In a third approach, the use of Cas9/RNPs appeared as a favourable alternative to the previous strategies deployed because of its transient effects and to its quick turnover\textsuperscript{240}. Taking advantage of the latest developments in the CRISPR/Cas research and the reagents that became widely commercially available, such as recombinant Cas9 protein and IVT and synthetic sgRNAs, a new set of experiments was conducted. In order to test the ability of the Cas9 recombinant protein to bind the IVT guides and form RNP complexes and to recognise the targeted DNA, \textit{in vitro} experiments were conducted. Guide G2 was used because of the high editing efficiency it showed in preliminary experiments (Figure 4.3). Following IVT and purification of the guide, it was added to recombinant Cas9 \textit{in vitro} and was incubated with the corresponding DNA template leading to its cleavage at the targeted site, as the length of the digested bands indicated (Figure 4.8A). Next, HEK293T cells were transfected with Cas9/RNPs complexes with IVT guide G2, showed an efficiency of 25\% (Figure 4.8B). The fact that the editing efficiency was similar to the one observed after plasmid transfection with the same guide, suggested that for the hsa-miR-195 locus and the specific guide it was not possible to obtain a higher efficiency for Cas9/RNPs compared to plasmid delivery as described in the literature\textsuperscript{208}.

4.4.1.4 Off-target mutagenesis

One of CRISPR’s major limitation is the putative off-target mutagenesis as a result of the tolerance the guide shows in mismatches in their complementarity with the targeted DNA\textsuperscript{235}. In more detail, the sgRNA can recognise and bind DNA sequences that show up to 3nt differences in their complementarity, leading Cas9 to edit the off-target sites. This activity occurs more easily in the 5’-end of the guide, which is its distal end from the endonuclease domains of Cas9. This off-target activity can result in unwanted editing of the DNA sequence, although with lower efficiency\textsuperscript{203}. Given the importance of the gene editing to be specific together with the possibility of the off-target sites to be in onco-suppressor genes, the need to test for possible off-target mutagenesis becomes vital. An increasing number of studies suggest that CRISPR/Cas9 is likely to induce high off-target mutagenesis in multiple sites although with no conclusive data\textsuperscript{241}. To address this issue, the CRISPR design software provide a list of possible off-target sites as predicted by an algorithm that penalises
the suggested guides according to the number of mismatches between the guide and the putatively recognised DNA sequence, ranking them in an inverse order of off-target likelihood (Figure 4.1). To assess the off-target mutagenesis, primers flanking the 7 most likely off-target sites for guides G1 and G2, as predicted by CRISPR Online tool, were designed (Supplementary Table 5), with T7EI showing no editing at said sites (Figure 4.9). Taking into consideration the high efficiency of the guides used together with the fact that no off-target editing was detected, we can securely conclude in that guides G1 and G2 did not induce any off-target editing in these sites, although the possibility that the guides can have other sites cannot be excluded. Whole genome sequencing is the technique of choice for assessing possible off-target mutagenesis by scanning the whole genome for discrepancies in the sequence\textsuperscript{242} although, even in this case, no definitive results can be obtained due to the presence of SNPs\textsuperscript{243}.

4.4.2 Gene Editing in Mouse Cells

Mouse VSMCs appeared as a favourable alternative to human SMCs. As in human cells, the first step undertaken was the design of guides. Although human miR-195 genomic locus shares high sequence similarity with the corresponding mouse one, the prerequisite for the guides to be fully complementary to the targeted sequence meant that new guides needed to be designed. For the mouse VSMCs three guides were designed since PAM sequences were more common in the mouse genomic locus of miR-195a compared to the human one. However, due to the need of the guides to target sites in the stem loop of miR-195a, guides that did not have the highest score were selected (Figure 4.10, Figure 4.11 and Figure 4.12).

4.4.2.1 Strategy B: Lentiviral delivery

Lentiviral delivery of the system was tested and editing efficiency in two independent experiments with guides Gm2 and Gm4 reached 31% when both guides were used simultaneously (Figure 4.13). Of interest is the fact that the pattern of the edited bands changed according to the guides used (Figure 4.14). A second set of experiments with guides Gm2 and GmA showed that they were more efficient when used together (Figure 4.15). In more detail, although guide GmA showed lower efficiency compared
to Gm4 (26% and 48%, respectively), infection of mouse VSMCs together with lentiviral particles conferring Gm2 resulted in editing of 51% efficiency, higher than the one observed when Gm2 and Gm4 guides were used simultaneously. The difference between the efficiency when targeting with both guides and with each guide alone can possibly be explained by the fact that guides Gm2 and Gm4 target sites that are located closer together (24nt apart), thus hampering the accessibility of the sequence by the Cas9/RNP complexes. The observed differences in the editing efficiency between the two guides for mouse VSMCs and between the two species (human and mouse) SMCs can be explained by the fact that different guides were employed targeting the same or different genomic loci, respectively. Discrepancies among the efficiency of different guides, targeting the same locus, have been reported previously for experiments in the same cell type with some guides failing to induce any editing while it has also been illustrated that the gene editing efficiency can vary greatly among different loci. These observations can be possibly explained by the different organisation and accessibility of the various loci or by possible instability or secondary structures of the sgRNAs. Since lentiviral delivery was deemed to be the most effective, it was concluded that no significantly higher editing efficiency could be induced in mouse VSMCs.

4.4.2.2 Strategy D: Transfection of mouse VSMCs stably expressing Cas9 with IVT guides

Although lentiviral delivery of the system to mouse VSMCs was shown to be highly effective, the constant expression of Cas9 together with the guides was deemed to be a non-favourable feature. Moreover, lentiviral infection also presents substantial limitations and risks. Considering these, in order to circumvent the difficulties raised, a different approach was followed. Cells stably expressing Cas9 were generated by infecting mouse VSMCs with lentiviral particles harbouring only the gene for Cas9. Next, IVT sgRNAs GmA and Gm2 was used to test their ability to form Cas9/RNP complexes to form and cleave the targeted sequence in vitro, showing a successful and almost complete digestion with the sizes of the resulting bands corresponding to the sites of cleavage (Figure 4.18A). Following these results, optimisation of the infection showed that selection of the infected cells with the use of Puromycin resulted in successful editing compared to no selection (Figure 4.18B). Following these results, newly engineered and selected cells were transfected with IVT guides Gm2
and GmA, simultaneously, with T7EI showing a similar efficiency with that achieved with lentiviral infection (54% on Day 2, Figure 4.19), results indicative of the efficacy of this approach. In addition, the editing efficiency on two different timepoints (Day2 and Day7) indicated the viability of the edited cells.

In addition, the guides selected for editing mmu-miR-195a interrupted different restriction sites upon successful editing (Figure 4.11). That provided the opportunity to test, in a fast and easy way, with the use of restriction enzymes, the editing efficiency. In detail, since editing interrupted the restriction sites restriction enzymes recognised, formation of any indel would result in the corresponding site to stay intact upon digestion (Figure 4.20). Furthermore, the T7EI assay was employed to evaluate the level of efficiency.

All in all, by using this approach no off-target editing was expected to occur since the accumulation of the Cas9 was not expected to result in any off-target editing. The presence of sgRNAs, that could potentially form a complex with the stably expressed Cas9 and guide it to edit off-target sites, was transient. Not only Cas9 is incapable to target any sequence without the help of the sgRNA but, further to this, it is the interrogation of the targeted sequence and its binding by the sgRNA, that activates the endonuclease domains of the protein. In more detail, Cas9 endonuclease scans the genome for PAM sequences which it binds and releases rapidly if no complementarity is recognised between the sgRNAs and the adjacent to the PAM genomic DNA. Upon recognition of a highly complementary sequence, the conformation of the Cas9 changes, because of the DNA binding by the sgRNA, an alteration that brings the two nuclease domains of Cas9 (HNH and RuvC) closer together and activates them. Therefore, the transient presence of IVT guides in the cells would mean that the accumulated Cas9 would stay inactive. Thus, with this approach, gene editing occurs only upon transfection with the guides while their presence in the cells is transient. Taken together the above, transfecting VSMCs stably expressing Cas9 with IVT guides was concluded to be the most favourable delivery approach for the aim of the current study.

4.4.2.3 Off-target mutagenesis

To address the limitation of the unintended mutagenesis, the same approach, as previously described, was followed. Moreover, since a new delivery system was
tested, that of the transfection of mouse VSMCs stably expressing Cas9 with transiently present IVT sgRNA, the need to test for off-target editing was high. Firstly, the 5 most likely off-target sites for either guides, as returned by the CRISPR Design Tool (Figure 4.10), were tested in cells infected with lentiviral particles with no mutagenesis being detected (Figure 4.21). Furthermore, new guides were identified and tested for guide Gm2. Various software, freely available online, can be employed for the design of guides with different software using different algorithms to predict possible off-target sites. These algorithms penalise mismatches between the DNA sequence scanned and the suggested guides in a different way resulting in a list of possible off-target sites that vary among the different software. To address this, a set of 7 possible off-target sites was tested consisting of sites that were returned both by CRISPR Online Tool and CRISPR Finder, provided by Sanger Institute (Figure 4.22). By testing sites that appeared in both software the chances of false positive sites were eliminated. T7EI results in Cas9 expressing mouse VSMCs demonstrated that guide Gm2 did not induce any off-target mutagenesis.
Chapter 5 INVESTIGATION OF THE REGULATION AND FUNCTION OF CLUSTERED miRNAS WITH CRISPR/Cas9 EDITING

5.1 Introduction

MiRNAs provide rapid responses in physiological and non-physiological conditions to counteract rapid and variable fluctuations and safeguard the robustness of biological systems. MiRNA biogenesis is a multistep process that starts with the miRNAs being transcribed by RNA polymerase II as pri-miRNAs which encode for either single or multiple miRNAs. A large portion of miRNAs organise in miRNA clusters while computational analysis suggests that the existence of miRNA clusters is much higher than expected by chance. MiRNAs being transcribed in the same primary transcript are organised in homo-clusters, that consist of members of the same family, and hetero-clusters, that consist of miRNA from different families. Pri-miRNAs present a very specific secondary structure by folding into hairpins, that contain imperfectly base-paired stems. They are then processed via a two-step, sequential maturation process during which they get cleaved by two RNA III nucleases. During the canonical pathway, they are cleaved, in a first step in the nucleus by Drosha and DGCR8, resulting in the release of a premature miRNA and its export to the cytoplasm. In a second step the pre-miRNA is cleaved by Dicer, leading to a miRNA duplex and subsequently to a mature miRNA. A precise recognition of the pri-miRNA by the microprocessor have emerged as important for the fidelity of processing. Several determinants, such as the secondary structure of the stem loop and sequence motifs, need to interact in a co-ordinated response for pri-miRNA recognition and cleavage. These findings imply that CRISPR/Cas9 gene editing of miRNA genes can affect processing of the hairpin in a dual manner either through the alteration of its sequence or of its secondary structure.

In this chapter, the use of gene editing as a tool to understand the clustered miRNA regulation was examined. Previous experiments, outlined in Chapter 4, demonstrated that CRISPR/Cas9 system is a very precise tool, inducing indel formation only in the
vicinity of DSB. Intriguingly, it was shown that in mmu-miR-497~195 cluster mutations in the mmu-miR-195a stem loop can affect the expression of mmu-miR-497a that resides in the other hairpin of the cluster. It was hypothesised that the editing of the mmu-miR-195a exerts an effect on the maturation of mmu-miR-497a by affecting its processing by Drosha.

Following the literature and the results outlined in Chapter 4, the study focused on four clusters, the homo-cluster mmu-miR-497~195, the hetero-clusters mmu-miR-143~145 and mmu-miR-17~92 and cluster mmu-miR-106b~25, a paralogue of cluster miR-17~92. The objective was to determine whether editing the sequence of one miRNA of the cluster is able to affect the expression levels of the other clustered miRNAs by influencing their maturation.

The results presented in this Chapter demonstrated diverse regulatory mechanisms of expression for clustered miRNAs. Firstly, they confirmed that mutations in the genomic locus of mmu-miR-497~195, induced by CRISPR/Cas9 editing affect the expression levels of the mature mmu-miR-497a that resides in the other hairpin of the same cluster. Secondly, experiments in different clusters showed that although editing of clustered miRNAs is capable of affecting the expression levels of the other miRNAs of the cluster, these results cannot be extrapolated to all of the clusters. In some cases, editing of a clustered miRNA did not affect the expression of the rest of the miRNAs.
5.2 Experimental Design

The aim of the experiments presented in the current chapter was to elaborate on the findings outlined in Chapter 4, on editing of mmu-miR-195a, and to elucidate whether editing of clustered miRNAs can affect the expression levels of the other miRNAs of the cluster. Four different clusters were chosen to be studied.

Firstly, cluster mmu-miR-497~195 was further studied in more detail and a new guide was designed and used to confirm that the observed results were not guide specific. Next, the effects of editing were evaluated in the hetero-cluster mmu-miR-143~145\(^{114}\) which consists of two miRNAs, mmu-miR-143 and mmu-miR-145a, that are not members of the same miRNA family and do not share seed sequence similarities. The objective was to overcome the limitations in studying possible autoregulatory effects of miRNAs of homo-clusters and to assess whether editing of hetero-clusters shows different effects. Guides editing different sites in or close to the stem loop of mmu-miR-145a were designed while changes in its secondary structure upon editing was also evaluated. In addition, clusters mmu-miR-17~92 and mmu-miR-106b-25 were studied. Guides targeting mmu-miR-18 were designed and used and it was hypothesised that the abolition of its hairpin would affect the maturation only of the flanking miRNA hairpins as was the expected outcome of editing mmu-miR-25. However, results showed that the effect of genomic editing in miRNA clusters can differ.
5.3 Results

5.3.1 Regulation of mmu-miR-195~497 cluster in VSMCs

5.3.1.1 Editing of mmu-miR-195 locus with sg195m2

Mmu-miR-195a forms a homo-cluster together with mmu-miR-497 as they are member of the same miR-15 family and get transcribed in one primary transcript with the stem loops located within 500bp. For the editing of mmu-miR-195a locus, mouse VSMCs stably expressing Cas9 were transfected with guide sg195m2 targeting a site in the stem loop of the miRNA disrupting one of the two SspI sites located in the sequence of mmu-miR-195a (Figure 5.1A). Digestion with SspI restriction enzyme showed successful editing (Figure 5.1C) while T7EI confirmed the editing with the efficiency varying from 28-30% (Figure 5.1D). Sanger sequencing demonstrated the formation of a panel of indels, mainly deletions flanking the targeted area (Figure 5.1E). QPCR quantification showed a reduction in the expression levels of mmu-miR-195a by 55% and of mmu-miR-497a by 62%, with the levels of mmu-miR-15a, mmu-miR-15b and mmu-miR-16 not being affected (Figure 5.1F) confirming results shown above.
Figure 5.1 Gene editing of the miR-195a locus using IVT sgRNA195m2. (A) Schematic representation of the genomic locus of the mmu-miR-497~195 cluster with the two stem loops of mmu-miR-497a and mmu-miR-195a being depicted as empty rectangles. The corresponding site of editing is depicted as a grey triangle and the position of the guide upon editing as a grey line. (B) The T-Coffee multiple sequence alignment programme was used to compare the mouse miR-15 family members. Bottom Panel: mmu-miR-497a displays high homology with mmu-miR-497b. Mmu-miR-497b is encoded by the antisense strand of the mmu-miR-497a~195 cluster and does not harbour the same seed sequence. (C) SspI digestion results upon editing of mmu-miR-195a with sg195m2 (D) T7EI assay for the mmu-miR-195a locus following editing with sg195m2. ND: Non-detectable. P1 and P2: Genomic PCR primers. (E) Sanger sequencing showing the editing of the mmu-miR-195 locus. The presence of random indels, mainly deletions was detected and represented as dashes. The PAM sequence is highlighted in grey. Insertions are shown in bold small case fonts. N/A: Not Applicable. (F) Gene editing of the mmu-miR-195a locus resulted in the down-regulation of both mmu-miR-195a and mmu-miR-497a without affecting the expression of other members of the miR-15 family as assessed by qPCR. U6 was used as a normalization control. n=4, * p<0.05. Assessment of the levels of miR-497b showed that it is not expressed in mouse VSMCs.
5.3.1.2 Possible gene editing of miR-497a locus with sg195m2

Following the assessment of the levels of miR-497a miRNAs upon editing of the miR-195 with the sg195m2, T7EI and Sanger sequencing was used to investigate putative editing and formation of indels in the genomic locus of mmu-miR-497a with neither of the assays showing any editing (Figure 5.2).

![T7EI Assay Image](image)

**Figure 5.2 Investigation of editing in the miR-497a locus with sg195m2.** (A) T7EI assay showed no editing at the locus of mmu-miR-497a upon editing with sg195m2. (B) Clustal alignment of Sanger sequencing data of the genomic locus of the mmu-miR-497a stem loop following sg195m2 editing in VSMCs. Wt: unedited cells, 1-10 randomly picked individual colonies harbouring genomic PCR amplicons from sg195m2 edited VSMCs. P3, P4: Genomic PCR primers as depicted in Figure 5.1, (n=3).
5.3.1.3 Editing of mmu-miR-195a locus with sg195m3

Editing of mmu-miR-195a locus was induced with transfection with guide sg195m3 of mouse VSMCs stably expressing Cas9. The newly designed guide targeted a site in the stem loop of the miRNA disrupting an SspI site (Figure 5.3A and B). T7EI showed efficient editing with the efficiency varying from 24-33% (Figure 5.3C). Sanger sequencing demonstrated the formation of indels, mainly deletions (Figure 5.3D), while qPCR quantification showed a reduction in the expression levels of both mmu-miR-195a and of mmu-miR-497a, without affecting the levels of mmu-miR-15a, mmu-miR-15b and mmu-miR-16 (Figure 5.3E), as previously shown (Figure 5.1). Off-target mutagenesis assessment did not show any editing in the six most likely off-target sites (Figure 5.4) with Sanger sequencing data of the mmu-miR-497a locus not revealing any editing (Figure 5.5).
Figure 5.3 Gene editing of the mmu-miR-195a locus using sgRNA195m3. (A) Schematic representation of the genomic locus of the mmu-miR-497~195 cluster with the two stem loops of mmu-miR-497 and mmu-miR-195 being depicted as an empty rectangular. The corresponding site of editing with sg195m3 is depicted as a grey triangle. (B) Representation of the mmu-miR-195a stem loop with the detailed sequence along with the targeted DNA sequence and the PAM sequence for sg195m3. (C) T7EI assay for the miR-195a locus following editing with sg195m3. ND: Non-detectable. P1 and P2: Genomic PCR primers. (D) Sanger sequencing showing the editing of the mmu-miR-195a locus. The presence of random indels, mainly deletions was detected and represented as dashes. The PAM sequence is highlighted in grey. Insertions are shown in bold small case fonts. N/A: Not Applicable. (E) Gene editing of the mmu-miR-195a locus resulted in the down-regulation of both mmu-miR-195a and mmu-miR-497a without affecting the expression of other members of the miR-15 family as assessed by qPCR. U6 was used as a normalization control. n=4, * p<0.05.
Figure 5.4 Assessment of off-target editing with sg195m3. T7EI did not show any gene editing in the loci of the 6 most likely putative off-target sites for sg195m3 as predicted by the CRISPR DESIGN Tool (http://cripsr.mit.edu, Zhang Lab, MIT) and the CRISPR Finder (Wellcome Trust Sanger Institute). OFT: Off-Target, ND: Non-detectable, n=1.
Figure 5.5 Investigation of editing in the miR-497a locus with sg195m3. Clustal alignment of Sanger sequencing data of the miR-497a stem loop locus following sg195m3 editing in VSMCs. Wt: unedited cells, 1-10 randomly picked individual colonies harbouring genomic PCR amplicons from sg195m3 edited VSMCs. P3, P4: Genomic PCR primers as depicted in Figure 5.1.
5.3.1.4 Regulation of the mmu-miR-497a expression

In order to determine whether the downregulation of mmu-miR-497a occurred as a secondary effect due to the downregulation of mmu-miR-195a, gain of function experiments were conducted. Overexpression of mmu-miR-195a in mouse VSMCs with the use of miRNA mimics demonstrated no increase in the levels of mmu-miR-497a (Figure 5.6).

![Figure 5.6. Overexpression of miR-195a. QPCR quantification of the levels of mature mmu-miR-195a and mmu-miR-497a, following overexpression of miR-195a using miRNA mimics. U6 was used as a normalisation. n=3, *p-value<0.05, (ANOVA with Dunnett post hoc test).]
5.3.1.5 Secondary structure of mmu-miR-195a stem loop upon editing

Computational analysis was performed in collaboration with Dr Kathleen Steinhofel and Prof Andreas Albrecht. Analysis with RNAfold prediction programme was used to determine the secondary structure of the stem loop upon editing with guide sg195m2 (Figure 5.7A) and sg195m3 (Figure 5.7B). The structure of the wild-type mmu-miR-195a stem loop presented the expected bulges and apical loop while one nucleotide deletion with either guide resulted in a predicted formation with a more pronounced bulge. A 18nt deletion with guide sg195m2 (Figure 5.7A, Left Panel) resulted in a shorter stem and a bigger apical loop while a 16nt deletion with guide sg195m3 (Figure 5.7B, Right Panel) resulted in a predicted secondary structure with no evident stem and apical loop.

![Figure 5.7 Predicted secondary structure of the mmu-miR-195a stem loop upon editing](image)

Minimum free energy structure of the mmu-miR-195a stem loop following editing with (A) sg195m2 and (B) sg195m3, as assessed by the structure prediction software RNAfold. The base pairing probability, as indication of structural remodelling or dynamics, was used for colour-coding.

5.3.1.6 Tertiary structure of mmu-miR-497~195 primary transcript upon editing

For the investigation of the effects of editing with guide sg195m2 on the tertiary structure of mmu-miR-497-195 primary transcript, the RNAComposer fully automated
RNA structure modelling server was used. The secondary structure without pseudoknots, as returned by vsfold5 was used as an input for the unedited transcript (wt, Figure 5.8A), after a 1nt mutation (mut1, Figure 5.8B) and an 18nt deletion (mut5, Figure 5.8C). The three 3D simulations appear different with the wild type showing a prominent mmu-miR-195a stem loop and a compressed but still accessible miR-497a stem loop (Figure 5.8A) while the conformation of the 1nt deletion (mut1) had no effect on mmu-miR-195a stem loop but led to a mmu-miR-497a stem loop strongly attached to the main core of the primary transcript (Figure 5.8B). In the case of the 18nt deletion (mut5), profound differences were observed as no typical hairpin for pri-miR-195a could be detected and the whole structure was more strongly entangled with the main core. The miR-497a stem loop displayed a compact shape in closer proximity with the miR-195a stem loop and with no clear accessibility to the hairpin.
Figure 5.8 Effect of CRISPR/Cas9 editing in the tertiary structure of mmu-miR-497~195 cluster.

Computational simulation of the 3D structure (generated by vsfold5 - with pseudoknots – as input to RNAComposer, pdb-file visualised by PyMOL) of the pri-miR-497~195 in (A) unedited WT, (B) mutant harbouring a deletion of 1nt (mut1) and (C) a mutant harbouring a deletion of 18nt (mut5). Mmu-miR-497a stem loop is depicted in green and its terminal loop in magenta. Mmu-miR-195a stem loop is depicted in red and its terminal loop in yellow. A full movie of the 3D structure of the different transcripts can be accessed on https://nms.kcl.ac.uk/kks/projects/SREP-17-12031.php (Username: cluster Password: 195497).
5.3.2 Gene editing in different miRNA clusters

5.3.2.1 Gene editing in mmu-miR-143~145 cluster

5.3.2.1.1 In silico design of guides

To design the guides targeting the genomic locus of mmu-miR-145, the CRISPR DESIGN Tool (http://crispr.mit.edu, Zhang Lab, MIT) was employed (Figure 5.9). The DNA sequence, corresponding to the transcript annotated as miRNA stem loop in miRBasev21, was uploaded to the software. The two following oligos, which targeted sites upstream of the stem loop and in the mmu-miR-145 stem loop (Figure 5.10), were deemed as the most suitable ones:

sg145m1: 5’-AGGGCGTGGACGTGCTGA-3’

sg145m2: 5’-GATGCTAAGATGGGATTCC-3’
Figure 5.9 In silico design of the sgRNAs for mmu-miR-145a locus. The CRISPR DESIGN Tool (http://crispr.mit.edu, Zhang Lab, MIT) was used to design the guides for editing the mmu-miR-145 locus. In the above figures, the sequences of the two guides, their corresponding score and the 5 most likely off-target sites are depicted for (A) guide sg145m1 (B) and sg145m2.
Figure 5.10 Schematic representation of the secondary structure of the mmu-miR-145a stem loop with guides. Representation of the mmu-miR-145a stem loop with the detailed sequence for (A) sg145m1 and (B) sg195m2. PAM sequence is highlighted in blue. The mature miRNA sequence is highlighted in purple and grey arrows indicate the corresponding position of the DSB on the stem loop which is the site of mutations induced by random indels following editing of the genomic locus in mutant cells.

5.3.2.1.2 Editing of the mmu-miR-145a locus

Mmu-miR-145a is transcribed in one primary transcript together with mmu-miR-143 forming a hetero-cluster, since the two miRNAs are not members of the same miRNA family. For the editing of mmu-miR-145a locus, two guides were designed, sg145m1 and sg145m2 with the former one targeting a site 30bp upstream of the mmu-miR-145a stem loop and the latter one disrupting a HinfI site in the mmu-miR-145a stem loop (Figure 5.11A). The two miRNAs do not share the same seed sequence neither any sequence homology (Figure 5.11B). T7EI results indicate successful editing with guide sg145m1 demonstrating editing efficiency varying from 47-57% while guide sg145m2 showed an efficiency of 24-29% (Figure 5.11C). HinfI restriction digestion enzyme confirmed the editing with guide sg145m2 (Figure 5.11D) and Sanger sequencing demonstrated the formation of small indels, mainly deletions with either guide (Figure 5.11E).
Figure 5.11 Gene editing of the mmu-miR-145a locus. (A) The genomic locus of the miR-143~145 cluster with the corresponding sites of the DSBs induced by sg145m1 and sg145m2. (B) MiRNAs miR-145a and miR-143 do not share a common seed region or any sequence homology, as the T-Coffee multiple sequence alignment programme indicates. (C) T7EI assay in three independent experiments for the miR-145a locus after editing with either guide. ND: non-detectable, (n=3). (D) The sg145m2 targets the miR-145 genomic locus and disrupts a HinfI restriction site. Asterisks (*) indicate differential digestion products. (E) Sanger sequencing data for the edited mmu-miR-145a genomic locus using guides sg145m1 and sg145m2, respectively. The PAM sequence is highlighted in grey. Deletions are represented as dashes, insertions are shown in bold small case fonts. N/A: Not Applicable. P7 and P8: Genomic PCR primers.
The assessment of the effect of editing on the levels of the miRNAs with qPCR revealed a downregulation of both mmu-miR143 and mmu-miR-145a upon editing with either guide although guide sg145m2 showed a sharper effect (Figure 5.12). In more detail, guide sg145m1 resulted in a 28% downregulation of miR-143 and 33% downregulation of miR-145 while guide sg145m2 showed a downregulation of 50% and 89% for miR-143 and miR-145a, respectively.

![Figure 5.12 Assessment of the levels of mmu-miR-143 and mmu-miR-145a upon editing. QPCR quantification of the primiR-143~145 expression in sg145m2 edited cells. Pri-miRNA Taqman assays that are specifically designed to amplify flanking sequences within 500 base pairs on either side of the stem loop were used. Beta actin was used as a normalization control. n=3, *p<0.05 (ANOVA with Dunnett post hoc test).](image-url)
T7EI did not reveal any off-target effect in the mmu-miR-143 locus with either guide (Figure 5.13A), results that were confirmed by Sanger sequencing (Figure 5.13B).

**Figure 5.13 Off-target editing in miR-143 locus.** (A) T7EI assay showed no editing at the locus of mmu-miR-143 upon editing with sg145m1 or sg145m2. The asterisk (*) depicts a non-specific band as a result of the T7EI digestion that is not due to editing, (n=3). (B) Clustal alignment of Sanger sequencing data for the mmu-miR-143 stem loop locus following sg145m1 and sg145m2 editing, respectively. Wt: unedited cells, 1-10 randomly picked individual colonies harbouring genomic PCR amplicons from edited VSMCs. P5, P6: Genomic PCR primers as depicted in Figure 5.11.

The expression levels of pri-miR-143 and pri-miR-145a were quantified with qPCR in sg145m2 edited cells. Targeting of mmu-miR-145a led to a robust decline of the levels
of both pri-miR-143 and pri-miR-145a (Figure 5.14A) while transfection of mouse VSMCs with miR-145a mimics resulted in a significant increase in the expression of mmu-miR-143 (250%) accompanied by an increase in pri-miR-143 and pri-miR-145a of 162% and 165%, respectively. The levels of KLF4, as a validated target of miR-145a, were subsequently downregulated by 40% (Figure 5.14B).

RNAfold prediction programme demonstrated changes in the secondary structure of the stem loop upon editing with guide sg145m2 (Figure 5.15). The structure of the WT mmu-miR-145a stem loop presented the expected bulges and apical loop while 1nt deletion did not lead to any significant changes. On the other hand, a 12nt deletion with guide sg145m2 resulted in a stem with no bulges and the loss of the apical loop.

**Figure 5.14 Assessment of the levels of pri-miR-143 and pri-miR-145a.** (A) QPCR quantification of the pri-miR-143~145 expression in sg145m2 edited cells. Pri-miRNA Taqman assays that are specifically designed to amplify flanking sequences within 500 base pairs on either side of the stem loop were used. Beta actin was used as a normalisation control. n=3, *p < 0.05 (ANOVA with Dunnett post hoc test). (B) The expression of the mature mmu-miR-143 and pri-miR-143~145, following mmu-miR-145a overexpression with mmu-miR-145a mimics, was quantified by qPCR. Kruppel Like Factor 4 (KLF4), a validated target of miR-145a was also assessed. Beta actin was used as a normalization control. n=3, *p<0.05 (ANOVA with Dunnett post hoc test).
Figure 5.15 Predicted secondary structure of mmu-miR-145a stem loop. Minimum free energy structure of the mmu-miR145a stem loop as assessed by the structure prediction software RNAfold after editing with sg145m2. The base-pairing probability, as indication of structural remodelling or dynamics, was used for colour-coding.

The expression of the long non-coding RNA Carmn, that overlaps the miR-143~145a cluster and constitutes an independent transcription unit (Figure 5.16B), did not differ after editing with either of the guides (Figure 5.16A).
Figure 5.16 Expression of long non-coding RNA Carmn. (A) QPCR quantification of Carmn in sg145m2 edited cells. Beta actin was used as a normalisation control, (n=3), (ANOVA with Dunnett post hoc test). (B) Snapshot obtained from NCBI database (https://www.ncbi.nlm.nih.gov/gene/328968) depicting the genomic locus of Carmn (cardiac mesoderm enhancer-associated non-coding RNA) and mmu-miR-143~145 cluster.
5.3.3 Gene editing of mmu-miR-17~92 cluster

5.3.3.1 In silico design of the guide

CRISPR DESIGN Tool (http://crispr.mit.edu, Zhang Lab, MIT) was used to design the guides for the genomic locus of miR-18a, the (Figure 5.17). The DNA sequence corresponding to the transcript annotated in miRBasev21 was used as input sequence to the software. The following oligo, targeting a site in the stem loop of miR-18a (Figure 5.18), was concluded to be the most suitable one:

Sg18: 5’-TTATGCCAGAAGGAGCAGTT-3’

![In silico design of the single-guide RNAs for mmu-miR-18 locus.](image)

Figure 5.17 In silico design of the single-guide RNAs for mmu-miR-18 locus. The CRISPR DESIGN Tool (http://crispr.mit.edu, Zhang Lab, MIT) was used to design the guide for editing the mmu-miR-18 locus. In the figure, the sequence of the guide, the corresponding score and the 5 most likely off-target sites are depicted.

![Schematic representation of the secondary structure of the miR-18a stem loop with guides.](image)

Figure 5.18 Schematic representation of the secondary structure of the miR-18a stem loop with guides. Representation of the miR-18a stem loop with the detailed sequence and the corresponding site of the DSB with sg18. PAM sequence is highlighted in blue, the mature miRNA sequence is highlighted in purple and grey arrow indicates the corresponding position of the DSB on the stem loop which is the site of mutations induced by random indels following editing of the genomic locus in mutant cells.
5.3.3.2 Editing of miR-18a

Cluster mmu-miR-17~92 consists of 6 miRNAs (Figure 5.19A) that do not share the same seed sequence and any sequence homology (Figure 5.19B). Gene editing of miR-18a with guide sg18 resulted in an efficiency varying from 32-37%, as shown by T7EI (Figure 5.19C), while digestion with restriction enzyme Bsp1286I confirmed the editing (Figure 5.19D). Sanger sequencing demonstrated the formation of small indels (Figure 5.19E) and qPCR revealed the robust downregulation only of miR-18a levels by 88% (Figure 5.19F).

![Figure 5.19 Gene editing of the miR-18a locus.](image)

(A) The genomic locus of the miR-17~92 cluster with the corresponding sites of the DSB induced by sg18. (B) T-Coffee multiple sequence alignment programme indicates that the miRNAs of the cluster do not share a common seed region or any sequence homology. (C) T7EI assay in three independent experiments for the miR-18a locus after editing. ND: non-detectable. (D) The sg18 disrupts a Bsp1286I restriction site. Asterisks (*) indicate differential digestion products. (E) Sanger sequencing of the edited miR-18a genomic locus. The PAM sequence is highlighted in grey, deletions are represented as dashes, insertions are shown in bold small case fonts. N/A: Not Applicable. P9 and P10: Genomic PCR primers. (F) Gene editing of the miR-18a locus resulted in the down-regulation only of mmu-miR-18a without affecting the expression of other members of the cluster as assessed by qPCR. U6 was used as a normalization control. n=4, * p<0.05.
Sanger sequencing did not reveal any off-target formation of indels in the loci of either mmu-miR-17, miR-19a or miR-20a (Figure 5.20).

Figure 5.20 Off-target editing in mmu-miR-17~92 locus. Clustal alignment of Sanger sequencing data for the mmu-miR-17~92 cluster following editing. Top Panel: miR-17 locus, Middle Panel: miR-19a locus and Bottom Panel: miR-20a locus. Wt: unedited cells, randomly picked individual colonies harbouring genomic PCR amplicons from edited VSMCs. P9 and P10: Genomic PCR primers as depicted in Figure 5.19.
Computational analysis with RNAfold prediction programme was used for assessing the secondary structure of the stem loop upon editing with guide sg18 (Figure 5.21). The structure of the wild-type miR-18a stem loop presented the expected bulges and apical loop while neither a 2nt nor a 14nt deletion led to significant changes with the latter resulting in an apical loop of smaller size.

![Predicted secondary structure of miR-18a stem loop](image)

**Figure 5.21 Predicted secondary structure of miR-18a stem loop.** Minimum free energy structure of the miR-18a stem loop as assessed by the structure prediction software RNAfold upon editing. The base-pairing probability, as indication of structural remodelling or dynamics, was used for colour-coding.
5.3.4 Gene editing of mmu-miR-106b-25 cluster

5.3.4.1 In silico design of the guide

The guide targeting the genomic locus of mmu-miR-25 was designed in CRISPR DESIGN Tool (http://crispr.mit.edu, Zhang Lab, MIT), (Figure 5.22) by using the DNA sequence corresponding to the transcript annotated in miRBasev21 as the miRNA stem loop. The highest scoring oligo, targeting the stem of the mmu-miR-25 hairpin (Figure 5.23), was selected as the most suitable one:

![Figure 5.22 In silico design of the single-guide RNAs for mmu-miR-25 locus.](image)

The guide for editing the mmu-miR-25 locus was designed with the help of the CRISPR DESIGN Tool (http://crispr.mit.edu, Zhang Lab, MIT). In the figure, the sequence of the guide, the corresponding score and the 5 most likely off-target sites are depicted.
Figure 5.23 Schematic representation of the secondary structure of the mmu-miR-25 stem loop with the designed guide. Representation of the mmu-miR-25 stem loop with the detailed sequence and the corresponding site of the DSB with sg25. PAM sequence is highlighted in blue, the mature miRNA sequence is highlighted in purple and grey arrow indicate the corresponding position of the DSB on the stem loop which is the site of the formation of random indels following editing of the genomic locus.

MiR-25 is part of the miR-106b~25 cluster and is transcribed in one primary transcript together with mmu-miR-106b and mmu-miR-93. For the editing of mmu-miR-25, guide sg25 was designed, targeting a site in the mmu-miR-25 stem loop disrupting an MfeI site (Figure 5.24A). The two miRNAs do not share the same seed sequence neither any sequence homology (Figure 5.24B). T7EI results indicate successful editing with an indel formation efficiency varying from 20-27% (Figure 5.24C). MfeI restriction digestion enzyme confirmed the editing (Figure 5.24D) and Sanger sequencing demonstrated the formation of small indels, mainly deletions (Figure 5.24E). Assessment of the levels of expression of miR-25 showed a significant downregulation only of miR-25 levels by 49% (Figure 5.24F).
Figure 5.24 Gene editing of the miR-25 locus. (A) The genomic locus of the miR-106b–25 cluster with the corresponding sites of the DSB induced by sg25. (B) T-Coffee multiple sequence alignment program indicates that miRNAs miR-106b, miR-25 and miR-93 that form the cluster do not share a common seed region or any sequence homology. (C) T7EI assay in three independent experiments for the miR-25 locus after editing. ND: non-detectable. (D) The sg25 disrupts an MfeI restriction site. Asterisks (*) indicate differential digestion products. (E) Sanger sequencing of the edited miR-25 genomic locus. The PAM sequence is highlighted in grey, deletions are represented as dashes, and insertions are shown in bold small case fonts. N/A: Not Applicable. P11 and P12: Genomic PCR primers. (F) Gene editing of the miR-25 locus resulted in the down-regulation only of miR-25 without affecting the expression of other members of the cluster as assessed by qPCR. U6 was used as a normalization control. n=4, * p<0.05. Assessment of the levels of miR-497b showed that it is not expressed in mouse VSMCs.

Sanger sequencing did not reveal any off-target formation of indels in the loci of either miR-93 or miR-106b (Figure 5.25).
Figure 5.25 Off-target editing in mmu-miR-106b-25 locus. Clustal alignment of Sanger sequencing of the mmu-miR-106b-25 cluster following editing. Top Panel: miR-93 locus and Bottom Panel: miR-106b locus Wt: unedited cells, randomly picked individual colonies harbouring genomic PCR amplicons from edited VSMCs. P9 and P10: Genomic PCR primers as depicted above in Figure 5.24.

RNAfold prediction programme revealed a secondary structure of the stem loop upon editing with guide sg25 with the 1nt deletion not conferring any important changes in contrast to the 6nt deletion that resulted to a short stem loop and a significant change in the structure of the apical loop (Figure 5.26).
Figure 5.26 Predicted secondary structure of mmu-miR-25 stem loop. Minimum free energy structure of the miR-25 stem loop as assessed by the structure prediction software RNAfold upon editing. The base-pairing probability, as an indication of structural remodelling or dynamics, was used for colour-coding.
5.4 Discussion

MiRNA maturation is a complex process which asserts the fidelity of the biogenesis of miRNAs. In an initial step the miRNA genes are transcribed to a pri-miRNA before getting cleaved by Drosha. The fidelity of the processing is ensured by a very specific and precise recognition of the pri-miRNA by the microprocessor. During the microprocessor recognition of the primary transcript a trimeric complex is formed including one Drosha, that recognises basal elements of the transcript, thus serving as a ruler, and two DGCR8 adaptors that interact with the apical elements. This modular model predicts varying contributions of several determinants that need to interact in a coordinated response for pri-miRNA recognition. Elegant experimental approaches have identified additional recognition and processing features. Secondary structures such as stem length, hairpin pairing, bulge size and position and apical loop size has been shown to contribute to effective miRNA biogenesis. Moreover, sequence motifs, such as a UG motif at the base of the hairpin, a UGU/GUG motif in the apical loop and a CNNC motif downstream of the hairpin, can enhance processing. Importantly, these primary sequence motifs exert their effects in some pri-miRNAs but not in others and are thought to have an additive effect in pri-miRNA processing. An intricate set of rules and modifications that are preferentially utilised in miRNAs over non-miRNA hairpins were also uncovered. These findings led to the prediction and experimental validation for a significant role of single-nucleotide polymorphisms (SNPs) in altering pri-miRNA processing and miRNA biogenesis in several occasions. Intriguingly they also imply that CRISPR/Cas9 gene editing of miRNA genes can affect processing of the hairpin in a dual manner either by altering the sequences or the tertiary structure.

Adding to the complexity of miRNA biogenesis, RBP can also modify various steps of the processing post-transcriptionally thus affecting the miRNA maturation. Additionally, an optimal length for the stem of the hairpin and two bulge-depleted regions in the stem, regions that may function as protein-interacting surfaces, were proposed as necessary for the RBP:miRNA interactions.

MiRNA clusters, based on their genomic locus organisation and regulation, can be divided into homo-clusters, that are composed of members of the same miRNA family and get transcribed together as one primary transcript, and hetero-clusters that are
comprised of miRNAs that belong to different families. In regarding how they exert their regulation, homo-clusters are capable of controlling their targets in a single step thus resulting in a rapid regulation, contrary to hetero-clusters that typically control their targets in multiple steps or targets and in a more delayed response\textsuperscript{248}. Thus, miRNAs are able to control the expression of multiple proteins, by binding their corresponding mRNAs, that function at different steps in various complex biological pathways. Through this, miRNAs can exert powerful effects on important processes of cell growth, differentiation, and apoptosis\textsuperscript{7, 23, 250}. Therefore, albeit the minor effects that an individual miRNA may have on a specific target, the combined effects of clustered miRNAs on multiple mRNA targets, of a common pathway, can be more pronounced and efficient.
Downregulation of the levels of mature mmu-miR-497a

In the current chapter, gene editing with CRISPR/Cas9 was assessed as a tool to understand the regulation of clustered miRNAs. Preliminary results demonstrated the effectiveness of CRISPR/Cas9 for the precise editing of the DNA sequence of miRNAs inducing indel formation in the vicinity of the DSB. As discussed previously, editing of mmu-miR-195a led to the downregulation of the expression levels of the edited miRNA as well as those of mmu-miR-497a, the miRNA that resides on the other hairpin of the same cluster.

After establishing the most effective delivery approach to edit the genomic loci of miRNAs, additional experiments were conducted in mouse VSMCs to edit mmu-miR-195a to decipher whether the downregulation was either the direct result of the editing or of the impaired processing of the pri-miRNA. The experiments focused on the genomic locus of the pri-miR-497~195 (Figure 5.1A). Guide sg195m2 was selected for the new set of experiments since it had previously shown the highest efficiency of the set of guides tested and Sanger sequencing had shown that the use of one guide resulted in the formation of various indels. Furthermore, compared to other guides tested, sg195m2 targeted a site in the stem loop of the miRNA but did not change the sequence of the mature miRNA and offered a better insight in the processing of the primary transcript of mmu-miR-195a. By not disrupting the sequence of the mature miRNA but, in contrast, by inducing indels in the stem loop, editing with sg195m2 can lead to an altered primary transcript with an intact mature sequence. Thus, the editing would lead to differential processing of the primary transcript and downregulation of the levels of mature mmu-miR-195a while in the case the maturation was not hindered the processing would result in a mature mmu-miR-195a with the WT sequence able to function physiologically.

Newly edited mouse VSMCs with guide sg195m2 were generated and restriction digestion showed successful editing (Figure 5.1C, D). Sanger sequencing confirmed the formation of indels (Figure 5.1D) and qPCR quantification, with more replicates than previously shown (n=4,) demonstrated the statistically significant downregulation of the levels of both the mature mmu-miR-195a and mmu-miR-497a (Figure 5.1E.), without affecting the levels of the other members of the miR-15 family, results that came in accordance with the results outlined in 4.4.2.2 in Chapter 4.
Nonetheless, the above results are not indicative of the mechanism that led to the downregulation of the levels of the mature mmu-miR-497a. Despite computational analysis not showing any complementarity between the DNA sequence of the genomic locus of mmu-miR-497a and sg195m2, the possibility of off-target mutagenesis could not be excluded, especially in light of the published studies reporting editing in off-target sites that were not predicted by online software. Moreover, the off-target editing of mmu-miR-497a due to the proximal editing by Cas9 of the targeted sites in mmu-miR-195a, could not be excluded. Although sequencing data showed precise editing in the vicinity of the DSBs induced by the designed guides due to innate PCR limitations and its incapability of amplifying long stretches of DNA with high fidelity, the genomic locus of the pri-miR-497~195 was tested in two independent sequencing experiments. Following the sequencing of the genomic locus of mmu-miR-195a, -PCR primers designed to amplify the genomic locus of mmu-miR-497a (Figure 5.1A) were used and T7EI showed no specific editing (Figure 5.2A) with Sanger sequencing confirming its absence (Figure 5.2B).

Further to this, sg195m3, as a newly designed guide with no sequence homology with sg195m2, induced gene editing with an overall efficiency that approximated that of sg195m2 (Figure 5.3C) resulting in statistically significant downregulation only of the levels of both the mature mmu-miR-195a and mmu-miR-497a (Figure 5.3D) with Sanger sequencing revealing the formation of indels in the vicinity of the induced DSB in the genomic locus of mmu-miR-195a. Noteworthy, no editing in a panel of off-target sites, predicted by both the CRISPR Online Tool and the CRISPR Finder was detected (Figure 5.4). Following the same approach, as outlined above, T7EI was used to assess unexpected off-target mutagenesis in the genomic locus of the mmu-miR-497a, with the results showing no editing (Figure 5.5).

Mmu-miR-195a and mmu-miR-497a form a cluster on chromosome 11 (Figure 5.1A and Figure 5.3A) and are transcribed as one primary transcript, with both stem loops located within 500bp. While no editing was detected in the genomic locus of mmu-miR-497a, the downregulation of the miRNAs needed to be further investigated. Gain and loss of function experiments could be employed in order to determine whether the mature downregulation of mmu-miR-497a may occur as a secondary effect due to reduced levels of mature mmu-miR-195a. In the case of the mmu-miR-497~195
cluster however, the sequence similarities between the two miRNAs indicate that the inhibitor can effectively silence both miRNAs and thus no conclusive data could be obtained. Overexpression of a miRNA can be achieved using short synthetic oligonucleotides that function as mature miRNA without requiring any processing by the Drosha or Dicer complexes. Following mmu-miR-195a overexpression using miRNA mimics, no differences on the expression of the mature miR-497a were observed (Figure 5.6), suggesting that the mmu-miR-497a expression is not controlled by miR-195a.

The tertiary structure of the primary miRNA transcripts can affect the accessibility of the stem loops by RNase III RBPs Drosha and Dicer and thus miRNA biogenesis may depend on this structure. Therefore, a computational analysis was conducted to assess whether structural constraints could hinder the processing of the mutant pri-miRNAs. The analysis for mmu-miR-195a focused on mutant 1 that harboured a deletion of 1 nucleotide (mut1), mutant 5 that harboured a deletion of 18 nucleotides (mut5) and the unedited transcript (WT). Three software prediction programmes (RNAfold, Sfold, Co-fold) were used and all of them returned identical results (Figure 5.7). Despite 1nt deletion not demonstrating any prominent changes in the secondary structure of the mmu-miR-195a hairpin, the 18nt deletion resulted in a wider apical loop and the abolition of a bulge in the stem of the hairpin (Figure 5.7A). Similarly, in the case of editing with sg195m3, 1nt deletion resulted in subtle changes whereas a 16nt deletion in the DNA sequence of the mature mmu-miR-195a led to substantial changes in the structure of the hairpin, with the elimination of the apical loop and the severe shortening of the stem (Figure 5.7B).

Furthermore, the 3D simulation highlighted even more pronounced differences. In the wt, the mmu-miR-195a stem loop was prominent while the mmu-miR-497a stem loop more compressed but clearly accessible. In the mut1, the conformation of mmu-miR-195a stem loop was similar although the single nucleotide mutation had an effect leading to a mmu-miR-497a stem loop strongly attached to the main core and not as accessible as in WT. Nevertheless, profound differences were observed in the conformation of mut5. No typical hairpin for pri-miR-195a could be detected and the structure was more strongly entangled with the main core. The mmu-miR-497a stem loop displayed a compact shape in close proximity with the mmu-miR-195a stem loop but relatively distal from the main core and with no clear accessibility to the hairpin.
(Figure 5.8). Overall, the 3D remodelling indicated a clear difference between the entire structure of wt and mut5 and to a lesser extent also between wt and mut1 with accessibility of the mmu-miR-497a stem loop being affected in both mutants. These results suggested that extensive deletions (in this case 18nt) in the genomics locus of mmu-miR-195 stem loop can alter the tertiary structure of the entire mmu-miR-497~195 transcript.

Although these results cannot safely lead to definitive conclusions about the mechanism through which the editing of the mmu-miR-195a gene can affect the levels of the mature mmu-miR-497a, the possibility of impairing the maturation in a dual manner directly through sequence and structural alterations and disruptions cannot be excluded. Computational analysis showed that disrupting the mmu-miR-195a stem loop can impose tertiary constraints that influence the processing of the entire pri-miR-497~195 resulting in downregulation of the expression levels of both miRNAs.

In the case of the mmu-miR-195a stem loop, the UGU/GUG motif in the apical loop is not present implying that the sequence is not important for the maturation of the hairpin. Nonetheless, editing with CRISPR/Cas9 led to the direct deletion of the terminal loop of the hairpin affecting its secondary and tertiary structure (Figure 5.7). These findings are in line with the literature that reports that mutations in the apical loop of miRNA hairpins lead to less effective processing and an autoregulatory role of the tertiary structure252,253.

### 5.4.1 Editing of mmu-miR-143~145 cluster

Next, to determine the effect of gene editing in a hetero-cluster the study focused on the mmu-miR-143~145 cluster114. This locus spans approximately 1400bp in mice and harbours mmu-miR-143 and mmu-miR-145a, two miRNAs that do not share sequence similarities (Figure 5.11A and B) but regulate essential functions for VSMCs growth, differentiation, and contractility78,254.

Following the same rationale presented previously, two sgRNAs targeting the mmu-miR-145a locus were designed using the CRISPR Online Tool (Figure 5.9). The
guides used were selected according to the site they targeted with guide sg145m1 generating a DSB 30bp upstream of the stem loop of mmu-miR-145a and guide sg145m2 targeting directly the stem loop, disrupting a HinfI restriction site, (Figure 5.10 and Figure 5.11A). T7EI results indicated that sgRNA145m1 was more efficient in inducing gene editing (Figure 5.11C) while digestion with HinfI confirmed the editing (Figure 5.11D). Sanger sequencing demonstrated small indel formation (1–40bp) with either sgRNA (Figure 5.11E), in accordance with previous results from mmu-miR-497~195 cluster.

QPCR quantification showed that the gene editing had a robust effect on miRNA expression. Editing with guide sg145m1 led to largely similar and significant downregulation of the levels of both the mmu-miR-143 (28%) and mmu-miR-145a (33%). A strong inhibition of mature mmu-miR-145a and mmu-miR-143 (89% and 50%, respectively) expression was observed following sg145m2 application (Figure 5.12). Interestingly, despite the high levels of gene editing induced by the sgRNAm1 guide the downregulation in mature mmu-miR-145a and mmu-miR-143 expression was significant but lower than when guide sgRNA145m2 was used, suggesting that targeting directly the stem loop is a more efficient strategy for inhibiting miRNA expression (Figure 5.12). Further support to this notion was provided by Sanger sequencing. The sg145m1 guide led to indels that did not disrupt the mmu-miR-145a stem loop sequence, while sg145m2 guide resulted in mutations within the hairpin. These results showed that the effects of targeting directly the stem loop of the miRNA is a more effective strategy of downregulating its expression since a DSB induced within the miRNA sequence can result in the formation of indels and the direct disruption of its secondary structure contrary to the effects of sg145m1 which induced a DSB in a substantial distance from the stem loop.

Despite the significant decrease in mmu-miR-143 expression, no editing of the mmu-miR-143 genomic locus was observed with either guide (Figure 5.13A), as was expected. Mutation screening revealed no indels in the mmu-miR-143 locus in edited cells with either guide (Figure 5.13B).

Two different approaches, previously reported, have shown that mutations in the pri-miRNA sequence of mmu-miR-143~145 can be tolerated provided that they do not disrupt critical structural elements. In more detail, introducing loxP sites for Cre-
mediated recombination in the genomic loci of mmu-miR-143 or mmu-miR-145a resulted in mutant mice that do not exhibit any effect on the expression of any of the two miRNAs in either case\textsuperscript{254}. This approach deleted the stem–loop sequence of the pre-miRNA for each of the miRNAs, or the genomic sequence for both miRNAs, and replaced it an exogenous DNA. Interestingly though, replacing the genomic region of mmu-miR-143 with the exogenous sequence of the LacZ reporter not only disrupted the expression of mmu-miR-143 but also led to effective silencing of the mmu-miR-145a expression\textsuperscript{118}. This construct permitted the cells that produced miR-143 and miR-145 to be visualised through staining with β-galactosidase. From these experiments, it becomes apparent that while both approaches disrupt the miRNA gene, the introduction of a reporter in the DNA sequence of the primary transcript renders it unsuitable for downstream processing and hinders the maturation of mmu-miR145a despite the presence of an intact mmu-miR-145a stem loop. Although not conclusive, these reports imply a critical role of the tertiary structure as a critical determinant in the maturation of clustered miRNAs. Furthermore, they indicate that the introduction of mutations in the pri-miR-143~145 sequence per se does not hamper miRNA maturation\textsuperscript{254}. This is in line with findings presented in the current study that sg145m1, which induces sequence changes upstream of the mmu-miR-145a stem loop, had not but a minimal effect on the expression of the clustered miRNAs inducing effective editing of the locus, a conclusion that becomes even more evident when the indel efficiency of the two guides are compared to their final effects on the expression of the miRNAs. 

To elaborate on the regulatory mechanisms involved, the expression of pri-miR-143 and pri-miR-145a in sg145m2 edited cells was quantified. To this end, commercially available assays, designed to target the flanking sequences on either side of the miRNA stem loop, were used. Effectively, the assays employed to quantify the pri-miR-143 and pri-miR-145a targeted the hairpin of mmu-miR-143 and mmu-miR-145a, respectively. A sharp decrease was observed indicating that targeting mmu-miR-145a affects the expression of the entire pri-miR-143~145 (Figure 5.14A), in line with the results for homo-cluster mmu-miR-497~195.

In order to determine whether mature miR-145a expression exerts transcriptional control on the miR-143~145 cluster, mmu-miR-145a was overexpressed in VSMCs using miRNA mimics. These are short synthetic oligonucleotides that function as mature miRNA without requiring any processing by the Drosha or Dicer complexes. High levels of mmu-miR-145a led to increased miR-143 expression (250%) that was
accompanied by a coordinated increase in pri-miR-143 and miR-145a expression (162% and 165% respectively), indicating transcriptional regulation. As expected the levels of KLF4, a validated target of miR-145a\textsuperscript{14} were significantly downregulated (40%) following miR-145a overexpression (Figure 5.14B).

These findings provide a helpful insight in the co-ordination in the maturation of clustered miRNAs. In the case of the mmu-miR-143~145 cluster an interdependency in the mature miRNA expression was demonstrated. The results outlined above were indicative of a transcriptional regulation of mmu-miR-145a on mmu-miR-143 and a feed-forward loop that reinforces the expression of the entire cluster. Gene editing of the mmu-miR-145a stem loop led to decreased expression of both the miRNAs of the cluster while mmu-miR-145a mimics led to increased levels of pri-miR-143 and mature mmu-miR-143. Thus, a direct evidence for the role of mature mmu-miR-145a in triggering the expression of the cluster mmu-miR-143~145 is provided. Although it is difficult to dissect whether the differences in mmu-miR-143 expression are due to structural changes or due to transcriptional regulation triggered by the reduced mmu-miR-145a levels, the two possibilities are not mutually exclusive. It is important to point out that mmu-miR-143 and mmu-miR-145a show no sequence homology suggesting that they do not have common targets\textsuperscript{14}. Nonetheless, the two miRNAs participate in a regulatory network that controls cytoskeletal remodelling and phenotypic switching of VSMCs under pathological conditions\textsuperscript{254}. This underlines the added value of miRNA clusters that provide an effective mechanism of cellular response and may explain the evolutionary pressure for their sequence conservation.

Next, RNAfold revealed that induction of 1nt deletion in the hairpin of mmu-miR-145a with sg145m2 did not change the secondary structure while a 12nt deletion changed the structure significantly (Figure 5.15). Intriguingly, the expression of Carmn (cardiac mesoderm enhancer-associated non-coding RNA), a long non-coding RNA overlapping the mmu-miR-143~145 cluster that constitutes an independent transcription unit\textsuperscript{226} did not differ in mmu-miR-145a edited cells (Figure 5.16), confirming that only the primary transcript (pri-miR-143~145) and not the wider locus is affected.

With regards to the tertiary structure of the cluster miR-143~145, as its length is about 1400bp, it cannot be analysed using the algorithms currently available for the
computational simulation. The development of more elegant tools to predict the tertiary structure of longer sequences will provide further insights.

Although not adequate results that support the notion were obtained, it is tempting to speculate that the downregulation of the clustered miRNA is due to both a disruption of a feed-forward loop as well as of the hindering of the processing of the primary transcript due to alterations in its tertiary structure.
5.4.2 Editing of mmu-miR-17~92 cluster

Next, the study focused on a highly studied miRNA clusters in the mouse genome, the miR-17~92, spanning 800 bp. The miR-17~92 cluster plays a pivotal role in the cardiovascular system and in cancer and mediates processes such as angiogenesis\textsuperscript{120,122}. This cluster consists of six miRNAs that can be grouped in four families based on their seed sequence (Figure 5.19A and B). Intriguingly, two paralogs, the miR-106a~363 and miR-106b~25 clusters are believed to be derived from a series of duplication and deletion events during vertebrate evolution have been identified\textsuperscript{255}.

Gene editing using sg18 that targets the miR-18a stem loop (Figure 5.19A) was efficient (32–37\%) as assessed by T7EI assay (Figure 5.19B) and disrupted the Bsp1286I restriction site (Figure 5.19D). Sanger sequencing revealed the presence of small deletions in the miR-18a stem loop (Figure 5.19E) that were predicted to induce a differential secondary conformation (Figure 5.21). As previously, these deletions were very specific and occurred only in the vicinity of the DSB. No mutations were identified in a 1 kb region that encodes the cluster (Figure 5.20). Interestingly, in contrast to the results presented for clusters mmu-miR-497~195 and mmu-miR-143~145, in the case of mmu-miR-17~92 cluster, although the expression of miR-18a diminished significantly (88\%) there was no impact on miR-17, miR-19a and miR-20a levels, miRNAs that are all located on the same cluster (Figure 5.19F). Additionally, discrepancies between the editing efficiency as evaluated by T7EI and the levels of downregulation of mmu-miR-18 showcased the limitations of the assay for assessing the impact of editing. Although underestimation of the indel efficiency is expected these results demonstrate that T7EI is not as sensitive as a technique. The expression of miR-19b-1 and miR-92a-1 could not be assessed as they are also encoded by a second miRNA cluster on chromosome X (Supplementary Table 1).

Intriguingly, analysing Sanger sequencing data with RNAfold did not show any significant secondary structure alterations despite the deletion of a long sequence, although these results are not conclusive. Thus, in the case of mmu-miR-18 it is tempting to speculate that is neither the structural changes of the targeted hairpin nor the deletion of sequence motifs but, possibly, the deletion of miRNA specific sequences that led to the downregulation of the levels of the mature miRNA.
Nonetheless, there is also the possibility that the 6nt deletion that was chosen to be assessed with RNAfold had no effect neither on the structure nor on the expression levels of the miRNA.

In agreement with the results shown above, it has been reported that pri-miR-17~92 shows a highly specific globular, tertiary structure that affects the processing of the transcript by Drosha\textsuperscript{252, 253}. More specifically, Chaulk et al have demonstrated that pri-miR-17~92 takes a folded structure of globular shape that results in the internalisation of the 3’-end of the transcript while the 5’-end of the cluster folds on top of it, leading to a less efficient processing of the internalised miRNAs\textsuperscript{253, 256}. Moreover, a non-miRNA containing stem loop in the core of the primary transcript, showing a high sequence conservation, has been reported\textsuperscript{256}. This stem loop interacts with other hairpins in the transcript, namely miR-19b, and affects the processing of the whole structure. Interestingly, physiological interactions of the two hairpins lead to low levels of miR-92a while disruption of the primary transcript’s structure, due to mutations, leads to enhanced expression\textsuperscript{252, 256}. In addition, this structure affects the maturation of other miRNAs of the 3’ core suppressing the expression of miR-19b and miR-18a\textsuperscript{256} while removal of miR-19b resulted in the upregulation of miR-92a and, to a lesser extent, of miR-18a. These data also indicate that miR-18a hairpin, although not part of the 3’ core, is internalised upon folding of the primary transcript.

These data further support the notion that miRNA processing is complex and highly sensitive in tertiary structure changes and showcase how sequence alterations, with the subsequent disruption of the structure, can affect miRNA maturation. Taken together, with the data outlined above, they show how CRISPR/Cas9 editing can be employed for the downregulation of different miRNAs of the same cluster with the use of specific guides depending on the intended result. Nonetheless, editing of the cluster’s DNA sequence although successful, was shown not to exert any effects on the maturation of the clustered miRNAs in contrast to previous results, highlighting that the processing differs among miRNA clusters.

5.4.3 Editing of mmu-miR-106~25
Similar results, as discussed in 5.4.2, were obtained when the stem loop of miR-25 was targeted in the miR-106b~25 cluster (Figure 5.22 and Figure 5.23). Effective editing was observed with sg25, as determined by both the T7EI assay (20–27%, Figure 5.24C) and the disruption of the MfeI restriction digestion site (Figure 5.24D). Sanger sequencing demonstrated a mutation pattern of small indels (Figure 5.24E), while qPCR quantification showed a significant downregulation of miR-25 expression (49%, Figure 5.24F) with the levels of miR-93 and miR-106b not being affected. Off-target mutagenesis assessment showed no editing for the miR-93 and miR-106b loci (Figure 5.25). The DNA sequence alterations led to a secondary structure of the miR-25 stem loop with a disrupted terminal loop (Figure 5.26).

Interestingly, interdependency in expression does not seem to be a common feature in all clustered miRNAs. Gene editing in mmu-miR-497~195 homo-cluster and mmu-miR-143~145 hetero-cluster revealed that targeting of one of the hairpins of the cluster led to downregulation of both of the clustered miRNAs. Although overexpression of mmu-miR-195a did not result in enhanced levels of mature mmu-miR-497a, suggesting no interdependency in the mmu-miR-497~195 cluster, in the case of mmu-miR-143~145 cluster the levels of mature mmu-miR-143 were affected by those of mmu-miR-145.

However, targeting of miR-18a and mmu-miR-25 revealed that the expression of other miRNAs in the corresponding clusters are not affected by changes in the structure of the targeted hairpins. No compensatory increase of these miRNAs was observed in edited cells either. Thus, it seems that while coordinated expression of miRNAs in clusters is a shared feature, feed-forward and feedback loops are in place only when a synergistic effect and combined regulation of multiple pathways is required.

Overall, it has been shown that despite the sequence diversity of pri-miRNAs, mutations can be well tolerated, provided that they do not disrupt critical elements such as stem length, bulge position and terminal loops. Furthermore, it was demonstrated that although miRNAs in a cluster form a co-transcriptional unit their expression is not always interdependent.
In conclusion, these data confirmed that the CRISPR/Cas9 editing technology is extremely precise, since indels occurred only in the vicinity of the DSB while no mutations were identified in the regions encoding the cluster. All in all, CRISPR/Cas9 emerged as a robust system for miRNA inhibition that is able of revealing novel regulatory mechanisms for clustered miRNA expression.
Chapter 6 ASSESSMENT OF THE LEVELS OF RBPS THAT BIND PRIMARY MMU-MIR-195A

6.1 Introduction

RBPs are important regulators of various steps in the expression of genes either co- or post-transcriptionally and they participate in multiple steps such as RNA splicing, capping, polyadenylation, export and translation\textsuperscript{7,257}. Following the same recognition patterns, RBPs participate in the maturation process of the miRNAs by controlling their post-transcriptional processing, tightly and in various steps such as the primary transcript cleavage, the export, the pre-miRNA cleavage, the stability control and the loading to RISC complex.

The regulation of the maturation of miRNAs is a multistep process that is controlled both by the rate and the efficiency of the recognition and the binding of the RNase III nucleases Drosha and Dicer. This efficiency is influenced by both structural and sequence characteristics of the miRNA precursors\textsuperscript{7}. Further to this, RBPs are able of recognising specific sequence motifs in precursor miRNAs and regulate the maturation according to the cell type or exogenous signals\textsuperscript{258-260}. In addition, RBPs have been shown to be capable of binding short single-stranded or double-stranded RNA\textsuperscript{1,260-262}.

Only a small portion of the putative RBPs have been studied in eukaryotic cells with regards to their function with the reported results not being conclusive due to the diverse experimental approaches followed and the cell line specific RNA interactions that the RBPs showed\textsuperscript{221,263}. Some of the discrepancies among the data reported cannot be conclusively explained but could be due to discrepancies between \textit{in vitro} and \textit{in vivo} conditions, different binding conditions, and/or the proteins \textit{per se} analysed\textsuperscript{221}.

The cell type specific interactions between RBPs and miRNAs are an aspect of the proteins that restricts their study and the universal application of the conclusions drawn from the studies while some RBPs recognise multiple hairpin-forming RNAs,
indicating a wider range of interacting repertoire. Nonetheless, in the latter case, the recognition was specific only to hairpin structures but not generic dsRNAs, implying that other co-factors, and not the secondary structure alone, are necessary for the recognition of the miRNAs. This duality that the RBPs present in their function suggests that they might have specific context or cofactor-dependent effects on their targets.

Multiple RBPs have been implicated in the biogenesis of the miRNAs such as KSRP and TRIM25 in the regulation of let-7, MSI1/2 in miR-7 processing. However, an exhaustive approach to identify and describe RBPs that bind and regulate miRNAs during all the stages of their biogenesis and turnover has yet to be conducted. This would provide the needed information for a comprehensive picture of the dynamic interactome of pri-/pre-miRNAs to be drawn. Taken together, these data highlight the importance of the in vitro studies for the identification of the interactions between RBPs and miRNAs, although the limitations of the approach need to be also taken into consideration.

Although studies have reported data regarding recognition of specific RNA sequence motifs and secondary structures for various miRNAs during their maturation, namely miRNAs that participate in tumorigenesis, miR-195 is one of the least studied ones. In light of this, in the current chapter the project focused in establishing a workflow to identify the RBPs that bind the pri-miR-195a and participate in its maturation. Additionally, a comprehensive panel of RBPs that bind the mutant pri-miRNA was assessed in order to elucidate the downregulation of the miRNA levels upon editing with CRISPR/Cas9.
6.2 Experimental Design

The scope of the current chapter was to establish the RBPs that recognise and bind either the WT or the mutant pri-miR-195a, and to compare the panels of the proteins that bound differentially on the two primary transcripts. The objective of this approach was to provide in vitro data for proteins that can participate in the maturation process of the primary transcript of the miRNA. Moreover, by using a mutant primary transcript, the goal of the study was to decipher the alterations in the processing of the transcript that can lead to differential expression. In order to do so, RNA molecules were in vitro transcribed.

In more detail, DNA sequences of pri-miR-195a that corresponded to either the WT or a mutant sequence, the latter one harbouring a 16nt deletion upon editing with sg195m3, were in vitro transcribed as described in Chapter 3.18 and were labelled with biotin. Next, they were bound to streptavidin magnetic beads and they were used as baits to bind and pull-down proteins from mouse VSMC lysates, proteins that are capable of recognising and binding the two distinct RNA molecules. The eluted proteins were then analysed with Mass Spectrometry (MS) with the aim to be validated by an independent approach, like Western Blotting. These experiments were performed in collaboration with Prof Manuel Mayr and the hypothesis to be tested was that, since no editing of the genomic locus of the mmu-miR-497 was identified, the downregulation observed was the result of the differential processing of the whole primary transcript due to a new panel of RBPs that recognise and bind the mutant pri-miR-497~195 that either restrained its processing or led to its cleavage.

It is the first time, to the best of my knowledge, that a pull-down experiment was conducted aiming to identify the network of RBPs that recognise the pri-miR-195a. According to the data received and analysed, a higher number of RBPs bound the mutant transcript, despite the shorter sequence and the decreased number of the putative sequence motifs that can potentially be recognised and bound. Lastly, analysis of the networks of the identified proteins, with the use of online tools, demonstrated that these proteins participate in the differential splicing of the mRNAs and the clearance of mRNAs.
6.3 Results

6.3.1 Differences in expression of RBPs

The levels of a panel of 14 RBPs that have been previously identified to participate in the maturation process of miRNAs\textsuperscript{221}, and share RNA binding motifs with the primary transcript of mmu-mir-195a, were tested. In more detail, three different edited sequences were expressed, with the use of pLKO.1 vector, that harboured 1nt, 4nt and 126nt deletions. VSMCs were infected with lentiviral particles with either an empty pLKO.1 vector, a vector with the unedited sequence of the primary transcript or with vectors harbouring each of the edited sequence resulting in their overexpression. Assessment of the expression levels of the 14 RBPs without showing any statistically significant difference in their expression. The proteins with their corresponding RNA binding motifs are depicted in Supplementary Table 6.

![Graph showing expression levels of RBPs](image)

Figure 6.1 Assessment of the expression levels of a panel of RBPs. The level of 14 RBPs were quantified in VSMCs after transfection with pLKO.1 expressing mutants of mmu-miR-195a harbouring different stretches of deletion.
6.3.1.1 Validation of the Mass Spectrometry data using Western Blot

6.3.1.1.1 Titration of the Pull-down conditions

Titration experiments were conducted to identify the optimal conditions for the RNA Pull-down assays. IVT RNA transcripts corresponding to the WT or the mutant pri-miR-497~195 were tested in two individual Pull-down experiments. Blotting against human antigen R (HuR) showed that the two transcripts were able to bind to the protein and pull it down in the conditions tested.

![Western blot analysis for HuR](image)

**Figure 6.2. Western blot analysis for HuR.** Two independent experiments were used to assess the ability of the IVT transcripts to recognise and pull-down HuR protein (lanes 1-2 and lanes 3-4, respectively). The negative and positive RNA controls that were provided with the kit were used as controls for the Pull-down. As negative control a poly(A) RNA and as positive control the 3’ untranslated-region of androgen receptor (UC-rich region) was used. WT: wild type transcript, mt: mutant transcript.

In addition to the first set of titration experiments, further optimisation revealed that regardless of the amount of DNA used in the IVT, the efficiency of the Pull-down was not affected. This indicated that the IVT reaction was not affected by the amount of DNA. Furthermore, the streptavidin beads bound to the positive RNA control, following elution under the recommended conditions, were boiled with sample buffer containing SDS, and assessment of the eluted proteins, demonstrated that the initial elution, although efficient, is not complete (Figure 6.3).
Figure 6.3. Titration experiment for the RNA template of the Pull-down. Different amount of DNA was used for the IVT reaction, preceding the labelling and the Pull-down. An IVT RNA that corresponded to GFP was used as a negative control alongside the negative and positive RNA controls that were provided with the kit. Input was used as a control for the antibody together with the eluate from the streptavidin beads bound to the labelled positive or negative RNA controls.

Further optimisation of the Pull-down assay reached the conclusion that there are no detectable differences in the efficiency of the Pull-down when either 25 or 50pmol were used while the absence of 30% PEG during the labelling reaction was detrimental for the assay (Figure 6.4). In addition, blotting with Dicer antibody showed the same efficiency pattern although the differences were less pronounced (Figure 6.5).

Figure 6.4. Titration experiment for the amount of labelled RNA to be used for the Pull-down. Different amounts of labelled RNA were tested to assess the optimal conditions for the assay.
Figure 6.5. Western blot analysis for Dicer. Different amounts of labelled RNA were tested to assess the optimal conditions for the assay. They were then tested for their ability to identify and pull-down Dicer.
6.3.2 Proteomic Analysis of the RNA Pull-down proteins

6.3.2.1 Mass Spectrometry analysis of the RNA Pull-down proteins

The proteins that were pulled down using the IVT transcripts corresponding to the WT and mutant pri-miR-497~195 were analysed using MS. The eluted proteins were then injected onto the nano-LC-MS/MS system for data dependent analysis on a Q Exactive™ Plus Quadrupole-Orbitrap MS. Overall, 766 proteins were detected.

Protein abundance was measured using normalised TIC of MS/MS spectra in Scaffold (version 4.8.7). In the untargeted proteomics analysis, 139 proteins were considered quantifiable with the following criteria: the protein was present in a minimum of 2 samples in at least 1 group. Statistical analysis of the MS results revealed a number of proteins of which about 45% were RNA binding and 55% non-RNA binding as annotated in Uniprot (Figure 6.6). Moreover, their localization was reviewed, with the majority of them being found in the nucleus (42% (Figure 6.7) as well as the number of their RNA Binding Domains (RBDs).

![Distribution of RNA pulled down proteins](image)

**Figure 6.6. Distribution of RNA pulled down proteins.** The proteins that were pulled down with either of the two transcripts (WT or mutant) were grouped with regards to their ability to bind RNA or not.
Figure 6.7. Chart for the localization of the RNA pulled down proteins. UNIPROT annotation was used to investigate the localization of the pulled down proteins.
Figure 6.8 Number of RBDs in the pulled down proteins. Of the proteins that were pulled down, the RBDs were grouped according to the number of the RBDs that they harbour as per UNIPROT annotation.

Statistical analysis of the MS results revealed 5 proteins that are bound to the mutant transcript in higher abundancy than to the WT one.
Figure 6.9. Volcano plot analysis for the RNA pulled down proteins that were differentially bound on the two transcripts. Statistical analysis of the Mass Spectrometry data revealed 5 proteins that bound the mutant transcript in significantly higher abundance. Red dots signify a p-value <0.05.
In addition, the proteins that bound exclusively to the mutant transcript were analysed and 12 proteins were identified to bind to RNA, 4 of which participate in the formation of complexes that are involved in the pre-mRNA splicing (Figure 6.10).
6.3.2.2 In silico analysis of the interactions of the pulled down RBPs

*In silico* analysis of the proteins that bound solely to the mutant IVT transcript of the pri-miR-497~195, with the use of STRING software, showed experimentally confirmed interactions between the identified proteins (Figure 6.11, Figure 6.12 and Figure 6.13).

![Graph showing protein-protein interactions](image)

**Figure 6.11. Analysis of the predicted protein-protein interactions for the protein Eri1.** The protein-protein interactions as returned by the STRING software are depicted. Protein Eri1 and Dicer1 are highlighted as two of the proteins that were identified by Mass Spectrometry analysis of the pulled down eluate. Magenta lines indicate experimentally confirmed interactions, dark lines co-expression and light green text-mining association.
Figure 6.12. Analysis of the predicted protein-protein interactions for the protein Luc7l3. The protein-protein interactions, as returned by the STRING software, show that proteins the identified Luc7l3 and Prpf19 proteins (highlighted in red ovals) are experimentally determined to interact. Magenta lines indicate experimentally confirmed interactions, dark lines co-expression, light green text-mining association and light blue protein homology.
Figure 6.13. Analysis of the predicted protein-protein interactions for the protein Tra2b. The protein-protein interactions, as returned by the STRING software, show the proteins Tra2b and Prpf19 as identified by MS (highlighted in red ovals) are experimentally determined to interact. Magenta lines indicate experimentally confirmed interactions, dark lines co-expression, light green text-mining association, light blue protein homology and blue gene co-occurrence.
**In silico** analysis of the identified proteins in cluster revealed a less complex but more specific network for the proteins with most of them showing experimentally confirmed interactions. Of note, all of the proteins that participate in the processing of the pre-mRNA (Figure 6.14, red oval) showed interactions with each other and with a part of the generically RNA binding proteins, such as Larp1 and Ddx21 (Figure 6.14).

**Figure 6.14. Analysis of the predicted protein-protein interactions for the cluster of proteins that bind only the mutant RNA transcript.** Interactions of the cluster of proteins that bind only the mutant RNA transcript as returned by STRING. Highlighted in red oval shapes are the proteins that participate in the pre-mRNA splicing and in blue rectangular the ones that are RNA binding, as annotated by UNIPROT.
Further *in silico* analysis with STRING showed that the interactions among the identified proteins are mainly due to binding, with the new analysis revealing a subset of proteins that bind together and create a network of pre-mRNA splicing proteins (Figure 6.15).

**Figure 6.15.** Analysis of the molecular action for the predicted protein-protein interactions for the cluster of proteins that bind only the mutant RNA transcript. Assessment of the molecular action of the proteins that were pulled down. Blue lines indicate binding, black reaction, purple catalysis and grey unspecified interaction. No line indicates no action. Highlighted in red oval shapes are the proteins that participate in the pre-mRNA splicing and in blue rectangular the ones that are RNA binding, as annotated by UNIPROT.
The confidence for the predicted protein-protein interactions, as returned by the STRING, showed binding of high confidence for the cluster of the pre-mRNA binding and processing RBPs (Figure 6.16).

Figure 6.16. Analysis of the confidence for the predicted protein-protein interactions for the cluster of proteins that bind only the mutant RNA transcript. Assessment of the confidence for the functional association of the proteins. The weight of the grey lines is indicative of the confidence level varying from high (0.9) to low (0.15), indicating the estimated likelihood that the depicted interaction is specific and meaningful. No grey line indicates no association. Highlighted in red oval shapes are the proteins that participate in the pre-mRNA splicing and in blue rectangular the ones that are RNA binding, as annotated by UNIPROT.
To confirm the protein interactions indicated by STRING and to study the possible pathways that these proteins participate in, Reactome software was used. Except for Eri1 and Dicer1, that interact with each other but do not show to participate in any pathway that is shared by the other proteins, the rest of them display an interaction network. Tra2b, Prpf19 and Luc7l3 interact with each other and the first two with Ddx21 and Prpf31, respectively, proteins that form a different cluster with Larp1 and Sart3. Proteins Tra2b, Prpf19 and Prpf31, interact with each other and partake in an mRNA splicing pathway with a statistical significance of $p<0.01$.

**Figure 6.17. Cluster analysis of the identified proteins.** Reactome was used to identify interactions between the pulled down proteins and the pathway they participate in. The different proteins, depicted in circles, are coloured according to the interacting groups they participate in. The black lines indicate interaction between proteins while the underlined proteins, according to the Reactome analysis, partake in mRNA splicing pathway with statistical significance of a $p$-value $<0.01$. 

**mRNA Splicing – Major Pathway**

$p$-value $<0.01$
6.4 Discussion

It is well established that the biogenesis of miRNAs is a multistep process that presents high complexity and is regulated by a number of RBPs\textsuperscript{21}. They participate in every step of the maturation of the miRNAs, from recognition and cleavage of the primary transcript, to the export of the pre-miRNA hairpins to the cytoplasm, the further processing to a miRNA duplex and the loading to RISC\textsuperscript{263, 266-268}. After these initial steps, the RISC complex binds the corresponding mRNA targets, via imperfect Watson and Crick complementarity between the mature miRNA and the targeted mRNAs and controls their expression\textsuperscript{1, 40, 221}.

Although they play a central role in the maturation of the miRNAs, only a small fraction (<5%) of these are characterised, contrary to RBPs that participate in the general processing of the miRNAs, such as Drosha and Dicer, that are very well described\textsuperscript{263, 269, 270}. Taking into consideration the plethora of the miRNAs, the complicated clusters they form, the interdependent regulation of their transcription and the complexity of the miRNA networks it becomes apparent that more RBPs than just Drosha and Dicer participate in a targeted way in the biogenesis of the miRNAs. Furthermore, studies report specific subsets of RBPs to interact with distinct miRNAs in specific cell types\textsuperscript{40, 221}. As an example of the specific role that RBPs play in miRNA biogenesis, TRIM71/LIN41 has been shown to regulate the expression of miR-29a alone, in two different cell lines while the downregulation of the protein affects the levels of the miR-29a targets in a functional approach that provides data regarding how RBPs can both affect directly the maturation of miRNAs and indirectly the silencing of the targeted mRNAs\textsuperscript{221}.

To gain deeper knowledge in regard to the processing of the miRNAs and how it gets affected by both sequence and tertiary changes in the pri-miRNAs, further identification of the RBPs that regulate the miRNA biogenesis is needed. To this end, in this chapter, the study focused on characterising the RBPs that interact with the WT pri-miR-195a and to elucidate the mechanism that result in the downregulation of both miRNAs of the cluster, mmu-miR-497a and mmu-miR-195a, upon editing of the latter.
6.4.1 Expression levels of RBPs

As a first approach, a panel of RBPs that have been previously described\(^{221}\) to participate in the maturation process of miRNAs and share RNA binding domains that recognise sequence motifs located in the primary transcript of mmu-miR-195a was assessed upon miRNA overexpression with pLKO.1 (Figure 6.1). After editing the genomic locus of mmu-miR-195a, DNA sequences harbouring deletions of various lengths (1bp, 4bp and 126bp) were PCR amplified and cloned into pLKO.1 vectors. QPCR quantification did not show any significant changes in the expression levels of the RBPs when compared to the levels of the cells at baseline, infected with lentiviral particles with an empty vector. These results indicated that the differences of the levels of mature mmu-miR-497a and mmu-miR-195a upon editing were not due to the downregulation of the expression of certain RBPs but a result of the editing per se and the differential processing of the primary transcript.

6.4.2 Proteomic Analysis

MS is an unbiased screening technique for single protein analysis that offers a platform for the thorough examination of multiple proteins from tissue or cell lysates. An additional advantage that it presents is the unbiased detection of proteins since there is no need for antibodies designed to recognise specific epitopes to be employed.

In the present study, MS was employed to identify differential binding of RBPs to the mutant primary transcript of mmu-miR-195a~497 cluster that result in the downregulation of the mmu-miR-497a expression levels, two miRNAs that reside on the same cluster and are transcribed as one primary transcript. Results already presented above in Chapter 5, have shown that no editing occurs within the genomic locus of mmu-miR-497a.

To decipher the changes in the maturation process that led to the differential expression of the miRNAs, RNA Pull-down assays were conducted and the eluted proteins were analysed. As an initial step, the Western Blotting was tested using the conditions following the company’s recommendations. In more detail, two RNA
molecules are provided with the kit, a positive control, that corresponds to the 3'-untranslated region androgen receptor RNA which is recognised and bound by HuR, and a negative control of a poly(A) sequence consisting of 25nt that are not recognised by the protein. In the initial experiment, the RNA transcripts biotin labelled together with the positive and negative RNA control and, following the pull-down, a HuR mouse monoclonal antibody was used to detect the protein (Figure 6.2). Results showed that, under the tested conditions, HuR is able to bind all the transcripts, even including the negative control although with significantly lower efficiency. Noteworthy, the mutant transcript in both of the two independent experiments presented, although bound by HuR, was bound with an efficiency almost as low as that of the negative control and notably lower than that of the WT. Interestingly, the 16nt deletion does not contain any AU element, which means that the difference in the binding efficiency that was observed was not due to the excision of any AU sequences but possibly due to a different tertiary structure that does not let the transcript to get accessed by the RBPs. These results contradict the previously presented MS results\textsuperscript{221}, that showed that Tial1, that binds AU-rich elements, is enriched in the eluate of proteins pulled down with IVT mutant RNA, indicating that the recognition of the AU sequence motifs is not enough for the interaction between the transcripts and the RBPs. In more detail, the differential binding of the WT and mutant transcript by HuR cannot be attributed to sequence differences since the number of AU motifs in the sequence of the two transcripts are the same. Therefore, the differential binding can only be explained either by differences in the tertiary structure of the mutant transcript or the inability of other necessary co-factors to bind the transcript.

Further titration experiments, concluded in that the amount of DNA, corresponding either to the WT or the mutant sequence, does not affect the efficiency of the in vitro transcription. As shown in Figure 6.3, regardless of the amount of DNA that was in vitro transcribed, blotting with HuR antibody did not demonstrate any differences in the amount of the RBP that was pulled down. Negative RNA control did not show any binding of the protein. Moreover, streptavidin magnetic beads that bound labelled positive RNA controls were incubated for a second time with sample buffer containing SDS, at 95°C, elution conditions that were more extreme than the recommended ones. Western blotting revealed a distinct, although diffuse band of HuR protein, results that showed that with elution under the recommended conditions, not the whole amount of proteins were eluted. The intensity of the band of the positive control, compared to the intensity of the band of the input, demonstrated that the pull-down,
although highly specific, did not enrich for the pulled down proteins. On top of that, the fact that we were able to detect pulling down of HuR with an artificial positive control transcript indicates that binding of RBPs to the bio-transcript does not mean RBP depletion of the sample. Further blotting with beta-actin revealed a distinct band only for the input, showing the specificity of the pull-down. Next, following labelling of the IVT RNA transcripts, different amounts of labelled transcripts were used for the pull-down with the results demonstrating that half of the amount of the recommended labelled RNA to be used had the same results while the use of smaller amount of labelled RNA improved the results. Absence of PEG in the labelling experiment showed detrimental results for the pull-down. PEG provides a hydrophobic environment to the RNA molecules improving the precipitation efficiency which in its absence decreases. Lastly, blotting with antibody against Dicer, showed that regardless of the amount of labelled RNA used, Dicer was pulled down only for the WT transcript (Figure 6.5). Although Dicer, according to the canonical pathway of the maturation of miRNAs binds to the pre-miRNA in the cytoplasm, the lack of cell compartmentalisation exposed the primary transcript to Dicer which resulted in its binding, contrary to the MS results that showed binding only of the mutant transcript by Dicer. Nonetheless, although these results revealed the binding of Dicer, they cannot provide any insight of the nature of interactions. More specifically, the binding could have been detected due to the physiological recognition of the pre-miRNA stem loop, as part of the primary transcript, by Dicer or due to the participation of the protein in a complex that bind the transcript.

The proteomics data outlined in the current chapter, offer a unique description of the RBPs that interact with both the WT and the mutant transcript. In total, 766 proteins were identified of which only 141 were quantifiable according to quality control of the data (Supplementary Table 7). Of them, about 45% are annotated in UNIPROT database as RNA binding proteins (Figure 6.6). The relatively small percentage of the identified RBPs, compared to the total number of proteins that were pulled down, can be indicative either of innate limitations of the pull-down assay or of specific protein-protein interactions between RBPs and other co-factors that led to non RBPs to be pulled down in complexes. More specifically, this essay was performed in vitro, and high activity of background binding was expected. Furthermore, it should be considered that the differences between conditions of the presented assay and the physiological conditions of the processing of the miRNAs are notably different. Firstly, during the maturation of the miRNAs, cell compartmentalisation plays a pivotal role.
in the process, with primary transcript being exported from the nucleus to the cytoplasm after the initial cleavage by Drosha. In the case of the RNA Pull-down, the lack of compartmentalisation could have potentially led to different, non-physiological and not related interactions between RBPs and co-factors, thus hindering the ability of specific RBPs to freely recognise and bind the pri-miR-497~195, while resulting in more proteins being pulled down. Secondly, the concentration of the primary transcript, which is used for the pull-down, is significantly higher than under physiological condition. The transcription of the pri-miR-497~195 is tightly regulated, by various regulatory factors that expose the primary transcript to the maturation machinery in waves, thus its high, non-physiological concentration could affect its binding by RBPs and subsequently the proteins that are identified by MS, as it was previously described\textsuperscript{221,260}.

Following these results, the identified RBPs were further analysed regarding their localisation as annotated by UNIPROT. Of the proteins identified, 50% are localised either in the nucleus or the cytosol while the vast majority of them (40%) are nuclear proteins, as expected (Figure 6.7). Under the physiological pathway of the maturation of miRNAs, the primary transcript is, in a first step, bound by the RBPs localised in the nucleus. With this in mind, it was expected that the primary transcript would be recognised by nuclear proteins even in the lack of cell compartmentalisation.

Further statistical analysis of the data showed that despite the various changes that were noted between the proteins pulled down with the WT and the mutant primary transcript, only 5 of them were statistically significant (Figure 6.9), with the rest of the proteins either showing small differences between the two groups or no statistical significance, results that lead to the conclusion that these proteins bound the two transcripts at the same level. Noteworthy, all of the statistically significant proteins were bound with higher affinity to the mutant transcript, contrary to the expected results. It was thought that the 16nt deletion in the sequence of the mutant transcript would, on the one hand, result in the abolition of sequence motifs able to be recognised by the RBPs and, on the other hand, a non-physiological tertiary structure that would restrain the RBPs from approaching and binding the transcript. Furthermore, although not statistical significance was identified, a trend was observed for the majority of the pull-down proteins to bind the primary transcript with higher affinity when compared to the WT transcript (Figure 6.9, lower right panel). Of the 5
statistically significant proteins identified, 3 showed a smaller than a 2-fold increase. Tial1 is an RBP that has not been previously described in detail, with an AU-rich binding element, Ran is a GTP-ase that is involved in the nucleo-cytoplasmic transportation of proteins and RNA molecules and Apex1 is an endoribonuclease that recognises ssRNA molecules and exerts control on their metabolism. Moreover, proteins Sec63 and Copb2 showed a 2.5 and 4-fold increase, respectively, in their affinity to the mutant transcript.

In fact, it has been proposed that Tial1 binds both to mRNA and ssDNA via its RRM domains (RRM1 and RRM2) located in its C terminus. It participates in the splicing regulation and translational repression of the mRNA, via the alteration of their half-life, and transcription regulation of the DNA\textsuperscript{271-274}. Nevertheless, the precise mechanisms that Tial1 employs to bind to and thereby repress the translation of specific target mRNAs has not been specified. Tial-1 has been shown to assist the aggregation of complexes, such as eIF2 and eIF5, in discrete cytoplasmic foci called stress granules that inhibit the translation of the bound mRNAs\textsuperscript{275}. It is thought that via this interaction both the mRNA stability and the translation is regulated during stress the composition of mRNA RNPs and their subsequent binding and processing with the translation or degradation machineries\textsuperscript{276, 277}. In our experiments, Tial1 was detected to bind a non-coding transcript therefore it is thought to partake in the degradation of the transcript.

Ran is a predominantly nuclear GTPase protein, a member of the Ras superfamily\textsuperscript{278}. It is associated with different proteins and depending on the binding of a Guanine (RanGTP or RanGDP) it is located either in the cytoplasm or the nucleus\textsuperscript{279}. It regulates the nuclear trafficking and the translocation of mRNA\textsuperscript{280}.

Apex1 is a multifunctional protein that has been implicated in binding of ssDNA, DNA/RNA hybrids and ssRNA molecules and participates in the response to oxidative stress with its two functions being the DNA repair and regulation of transcriptional factors. Functions as an endodeoxyribonuclease in the DNA repair during oxidative stress. Apex1 is involved in the regulation of transcription by repressing it together with HNRNPL and partakes in the metabolism of ssRNA. It is associated with proteins non-sense mediated decay-1 (NMD1) and Y-box protein 1 (YBX1) that regulate the alternative splicing of pre-mRNA\textsuperscript{281, 282}. 

183
Furthermore, Sec63 has been implicated in the transport of proteins through the endoplasmatic reticulum (ER) while Copb2 participates in the transportation of proteins from the ER to the Golgi network in association with GTP-binding proteins\textsuperscript{283}.

Taken together these results indicate that the mutant transcript follows a different processing pathway. Although further \textit{in silico} analysis with STRING online software did not reveal any interaction between the identified proteins (data not shown), according to their annotation they can be organised in two groups. To be specific, Tial1 and Apex1 are RBPs that recognise specific motifs (AU-rich and UACA respectively) in ssRNA molecules and partake in their processing. Their function is indicative, although not conclusive, of how the editing of mmu-miR-497~195a locus, through a 16nt deletion in the mmu-miR-195a hairpin, can affect its processing. These proteins, being identified in higher abundance bound to the mutant transcript, could potentially be due to differential processing that leads to the clearance of the non-physiological transcript. Furthermore, a second cluster, comprised of Ran, Sec63 and Copb2, regulates the identification and transportation of proteins to Golgi network. Although, these results seem not to be relevant to the processing of RNAs and miRNAs, they should be analysed under the light of both the presence of the non-physiological pri-miRNA and the lack of compartmentalisation. The RBPs that bind the mutant transcript could form a complex that potentially attract a group of proteins that do not naturally bind the RBPs organised around the WT pri-miR-497~195 transcript. Thus, these results could be suggestive of a pathway that does not reflect any canonical processing and is not followed under physiological conditions in cells.

Interestingly, a set of proteins that bound only the mutant transcript was identified (Figure 6.10, Supplementary Table 7). Nonetheless, the possibility that the primary transcript forms a new structure, because of the different sequence, cannot be excluded. Of them, a small subset of 12 proteins were annotated to bind RNA molecules. More specifically, two subsets were identified according to UNIPROT annotation. A subset of 8 proteins that are generally RBPs, such as Dicer1 which binds primary transcripts and leads to specific cleavage, was identified and a subset of 4 proteins that is involved in the pre-mRNA splicing. The representative spectra of the identified proteins are shown in Supplementary Figure 1 and Supplementary Figure 2. Although these interactions could be due to new sequences being created,
analysis of the sequence of the mutant primary transcript did not show any known RNA binding motifs to be created.

Further analysis of the associations of these proteins with online software, revealed functional interactions between the proteins and between the two clusters. STRING software, a free online database of known and predicted functional and physical protein-protein interactions\textsuperscript{284}, was used for a preliminary analysis to highlight the functional relationship of the proteins. STRING uses experimental evidence, database information, text-mining from PubMed abstracts\textsuperscript{285}, co-expression data\textsuperscript{284}, neighbourhood, ortholog analysis and phylogenetic contribution genome-based prediction data\textsuperscript{286} to reveal interactions indicative of biological pathways. When proteins were analysed individually, interactions showed a group of proteins that have been experimentally shown to interact. More specifically, in the case of Eri1 (Figure 6.11), the proteins it interacts with include Dicer1, a type III ribonuclease that plays a pivotal role in the maturation of miRNAs. Eri1 is an RNA exonuclease that has been shown to participate in the degradation of histone mRNA along with the 3’ overhangs of siRNAs. Additionally, Dicer1’s function of cleaving miRNAs, as double stranded RNAs that form short hairpin structures, is well described. Taken together, these data suggest that either the two proteins work together conferring similar functions or that Dicer’s activity supplements Eri1 with cleaving the RNA molecules that the latter recognises. Furthermore, Luc7-like 3 protein, that is involved in the RNA splicing has been experimentally shown to interact with Prpf19 (Figure 6.12), a core component of complexes participating in pre-mRNA splicing, while Tra2b, a sequence specific RBP that participate in pre-mRNA splicing, interacts with Prpf19 according to experimentally proved data (Figure 6.13). These results are suggestive of a possible interaction between two groups of proteins (Eri1-Dicer1 and Luc7l3-Prpf19-Tra2b) that recognise the mutant transcript and cleave it.

Intriguingly, when the 12 RBPs that bind solely the mutant transcript, were input in the software together, a network of experimental validated protein-protein interactions was highlighted (Figure 6.14). The cluster analysis revealed that 7 of the input proteins form a network of interacting RBPs while the interaction between Eri1 and Dicer1 was confirmed. Additionally, a second clustering became apparent when they were grouped on the basis of their function. Proteins that participate in the pre-mRNA splicing (Figure 6.14 red oval) form a cluster of closer interacting members that show
an association through proteins Tra2b and Prpf31 with the second cluster of Larp1, Ddx21 and Prpf31 (Figure 6.14 red rectangular). Further analysis, with STRING software, demonstrated co-operative binding among the pre-mRNA splicing proteins indicating either an interdependent binding of the RBPs of the clusters to the mutant transcript or increased binding due to these interactions (Figure 6.15), with the confidence for these interactions being high for the core of the cluster among Tra2b, Prpf19 and Prpf31 (Figure 6.16).

Further analysis was performed with Cytoscape, an open-source software platform for the analysis and visualisation of interactions of proteins and protein networks. Cytoscape analysis of the 12 proteins that bound solely to the mutant transcript returned the same results as STRING software with the formation of 2 distinct clusters that interact with each other through binding between Tra2b-Ddx21 and Prpf19-Prpf31 (Figure 6.17). Lastly, pathway analysis with Reactome, an online software for the detection of biological pathways, indicated that these 3 proteins participated in a pathway of mRNA splicing with statistical significance (Figure 6.17, Supplementary Table 8).

In light of these results, it becomes apparent that this set of proteins participate in a different than the canonical pathway of the processing of miRNAs. The newly identified interactions cannot be explained exclusively through the formation of new sequence motifs due to the 16nt deletion. Moreover, although according to UNIPROT annotation these pulled down proteins are RNA binding, the possibility of interactions with the mutant transcript due to association with other proteins that bind the mutant transcript because of a newly created tertiary structure, cannot be excluded. Furthermore, these possible binding interactions would explain the non-RBPs that were pulled down solely for the mutant transcript (Figure 6.10). In addition, results from two independent software showed that 3 proteins (Tra2b, Prpf19 and Prpf31) form a core of interactions with the transcript and pathway analysis implicated them with pre-mRNA splicing.

In more detail, Tra2b has been implicated in mammals with the activation of nonsense-mediated decay splicing upon recognition of a GAA repeats although not specific mechanism has been identified. Furthermore, Prpf19 protein has been shown to be a core component of PRP19C/Prp19 assembly, as part of the U4/U5/U6 tri-snRNP spliceosomal complex, and partakes both in its assembly and activity of
Interestingly, complex PRP19C does not contain any known RNA binding proteins. This feature suggests that the complex interacts with the pre-mRNA through other RNA binding factors, a characteristic that comes in accordance with our results. Additionally, Prpf31 is another component of the spliceosome.

Moreover, it seems that the mutant primary transcript gets processed through a different pathway. Indicative results from the in silico analysis of the interactions of the MS analysed proteins suggest that the mutant primary transcript takes part in a pathway that leads to its degradation and clearance from the cells. These results also explain the decrease in the levels of the mature miRNAs, mmu-miR-497a and mmu-miR-195a, since clearance of the primary transcript would lead to a decrease of their levels. Nevertheless, although indicative, these results cannot be conclusive. The possibility of these interactions not reflecting the conditions in cells upon editing of the targeted locus cannot be ruled out. In more detail, as explained above, the conditions of the pull-down experiments do not represent physiological conditions; thus, although in silico and statistical analysis identified a set of proteins that potentially participate in a different processing pathway the possibility that these interactions came up as significant due to innate limitations of the assay cannot be excluded. In more detail, as already explained, firstly the presence of a mutant transcript, harbouring a 16nt deletion, does not occur in nature and secondly the conditions of the assay do not represent cellular conditions. This could mean that the showed interactions do not represent any specific pathway that is followed by the cells but artefact interactions that is the result of non-physiological binding of the identified proteins to the mutant transcript.

Although MS is a sensitive technique that is employed for both the identification of new proteins and of a set of known proteins it still presents limitations. It is a semi-quantitative technique that is capable of showing only the presence of proteins and not possible interactions between them. In addition, the need for the validation of the results with an independent technique led to the employment of Western Blotting to test for the proteins, as demonstrated by the analysis of MS data. To elucidate the pathway that lead to the downregulation of the mmu-miR-497a upon editing, the study focused, on the proteins that bind exclusively to the mutant transcript. The hypothesis to be tested was that the deletion changed the tertiary structure of the primary transcript in a way that it could not be processed by the maturation machinery either
due to its degradation, through a different pathway, or due to the binding of RBPs that hinder its recognition.

The experiments that were conducted to identify the optimal conditions of blotting did not lead to successful results while more antibodies were tested against proteins Larp1, IGF2BP2, hnRNPE1, DHX9, EIF2AK2, ILF3 Tra2b (Table 2) but did not allow quantification of the respective protein due to multiple nonspecific bands of multiple sizes being detected.

In conclusion, MS analysis of RNA pull-down with WT and mutant transcripts of pri-miR-497~195 revealed, for the first time, a set of proteins that bind the pri-miR-497~195 as well as RBPs that bind only the mutant transcript. Although the results did not lead to any firm conclusions with regards to the interaction between the proteins and the transcript they imply that alternative pathways may lead to the downregulation of the levels of mature mmu-miR-497a and mmu-miR-195a. Furthermore, as analysis of the 3D structure of the edited transcript revealed, the RNA Pull-down results could also indicate ineffective processing of the transcript. As the different elements of the primary transcript are not easily accessible upon editing, due to its non-canonical conformation, the binding of the identified pulled-down proteins could, potentially, signify differential processing that can explain the lower levels of mmu-miR-497. However, the lack of validation of the MS identified proteins with Western Blot indicate that further analysis and validation is needed in order to conclusively determine the RBPs that participate in the alternative processing of the mutant transcript.
Chapter 7 GENERAL DISCUSSION

7.1 Conclusions

The objective of the study was to specifically inhibit hsa-miR-195 in order to study its implication in aneurysm formation. To this end target the genomic locus of hsa-miR-195 stem loop was edited with the precision offered by CRISPR/Cas9. When designing the guides the short length of the gene was taken into consideration and the guides were chosen based on the position of the site they targeted. Three different strategies of delivery were studied.

Plasmid delivery to HEK293T cells, was tested as an initial method because of the transient expression of Cas9 minimising the possibility of off-target editing. DNA sequencing of the edited cells showed that the editing occurred specifically in the targeted sites, as expected, resulting mainly in deletions. However, taking into consideration the high percentage of the transfection that was observed and the relatively modest percentage of the gene editing, the plasmid delivery did not appear as favourable for primary cells, cells that are difficult to transfect.

In order to address the low efficiency observed with plasmid delivery, lentiviral delivery of the CRISPR/Cas9 was chosen and tested in primary human cells (HUVECs). Results from three independent experiments suggested that infection with lentiviral particles resulted in almost complete editing of the gene that can be attributed to both the high proliferation rate of the cells as well as the efficacy of the delivery.

In a third approach, the use of Cas9/RNPs appeared as an advantageous alternative to the previous strategies deployed because of its transient effects. The fact that, the editing efficiency was similar to the one observed after plasmid transfection with the same guide, lead to the conclusion that it was not possible to obtain a higher efficiency for Cas9/RNPs compared to plasmid delivery.
Next, after assessing the levels of miR-195 in human SMCs in prolonged cell culture conditions and because of their low levels we changed to mouse VSMCs. For the mouse VSMCs three guides were designed due to sequence differences. Lentiviral delivery of the system was tested based on the editing efficiency of the guides. Of interest is the fact that the pattern of the edited bands changed according to the guides used. The difference between the efficiency when targeting with different combination of guides and with each guide alone can possibly be explained by the relative position of the targeted sites influencing their accessibility by the Cas9/RNP complexes. Differences in the editing efficiency among the guides for mouse VSMCs and between the two species (human and mouse) SMCs can be explained by the fact that different guides were used.

Although lentiviral delivery was highly effective, the continuous expression of Cas9 together with the guides was deemed to be a non-favourable feature. Stably expressing Cas9 cells were generated by infecting mouse VSMCs with lentiviral particles harbouring only the gene for Cas9. Next, IVT sgRNAs were used to test their ability to form Cas9/RNP complexes showing a successful and almost complete digestion. By using this approach no off-target editing was expected to occur since the accumulation of the Cas9 was not expected to result in any off-target editing.

Next, to address the limitation of the off-target mutagenesis, 7 possible off-target sites as returned by both CRISPR Online Tool and CRISPR Finder were tested. By using the potential off-target sites that were returned by both software we eliminated the chances of false positive sites being false. T7EI assay demonstrated that none of the guides induced any off-target mutagenesis.

Next, to determine the effect of gene editing in a hetero-cluster the study focused on the mmu-miR-143~145 cluster, consisting of two miRNAs that do not share sequence similarities. Two sgRNAs targeting the mmu-miR-145a locus were designed with qPCR quantification showing that the gene editing had a robust effect on miRNA expression. Interestingly, despite the high levels of gene editing induced by the sgRNA1m1 guide, the downregulation in mature mmu-miR-145a and mmu-miR-143 expression was significant but lower than when guide sgRNA145m2 was used. These results indicate that targeting the stem loop is a more efficient strategy for inhibiting miRNA expression.
In order to determine whether mature miR-145a expression exerts transcriptional control on the miR-143~145 cluster, mmu-miR-145a was overexpressed in VSMCs using miRNA mimics. High levels of mmu-miR-145a led to increased miR-143 expression as well as pri-miR-143 and miR-145a expression. Although not conclusive, these results are indicative of a disruption of a feed-forward loop due to miRNA editing as well as of the hindering of the processing of the primary transcript.

Next, the study focused on a highly studied miRNA clusters in the mouse genome, the miR-17~92. Gene editing using sg18 was efficient and although the expression of miR-18a diminished significantly there was no impact on miR-17, miR-19a and miR-20a levels, miRNAs of the same cluster. Intriguingly, RNAfold did not show any significant secondary structure alterations upon editing.

Next, the stem loop of miR-25 was targeted in the miR-106b~25 cluster with qPCR quantification showing a significant downregulation of miR-25 expression with the levels of miR-93 and miR-106b not being affected.

These results show how editing with CRISPR/Cas9 can be employed for the downregulation of different miRNAs of the same cluster. Nonetheless, editing of the DNA sequence of clustered miRNAs showed that the maturation process differs among miRNA clusters with varying effect on the levels of the other miRNAs of the cluster. Additionally, these results showed that expression interdependency is not common in all clustered miRNAs. Gene editing of mmu-miR-497~195 and mmu-miR-143~145 led to downregulation of both of the clustered miRNAs. However, targeting of miR-18a and mmu-miR-25 revealed that the expression of other clustered miRNAs are not affected by editing of the cluster sequence.

In conclusion, these data confirmed that the CRISPR/Cas9 editing technology is extremely precise, since indels occurred only in the vicinity of the DSB while no mutations were identified in the regions encoding the cluster. All in all, CRISPR/Cas9 emerged as a robust system for miRNA inhibition that is able of revealing novel regulatory mechanisms for clustered miRNA expression.
Next, CRISPR/Cas9 was assessed as a tool to understand the maturation of clustered miRNAs. This is the first time, to the best of my knowledge, that the maturation process of the mmu-miR-195a is investigated. Following the preliminary results from editing the DNA sequence of the genomic locus of mmu-miR-195a and the unexpected downregulation of mmu-miR-497a, the study sought to elucidate the mechanism of this downregulation. In order to decipher whether the downregulation was either the direct result of the editing or of the impaired processing of the pri-miRNA, guide sg195m2 was used in the new set of experiments. While no editing was detected in the genomic locus of mmu-miR-497a, overexpression of mmu-miR-195a lead to no differences on the expression of the mature miR-497a suggesting that the mmu-miR-497a expression is not controlled by miR-195a.

The tertiary structure of the primary miRNA transcripts can affect the accessibility of the stem loops by Drosha and Dicer affecting miRNA biogenesis. Therefore, computational analysis assessed if the structure of pri-miRNAs could influence the processing of the mutant pri-miRNAs. Software prediction programmes were used and all of them returned identical results with a 18nt deletion resulting in a wider apical loop and the abolition of a bulge in the stem of the hairpin. Although these results cannot safely lead to definitive conclusions about the mechanism, the possibility of impairing the maturation in a dual manner directly through sequence and structural alterations and disruptions cannot be excluded.

Building on these results together with the alterations in the predicted tertiary structure of the edited primary transcript we postulated that these changes, as a result of the editing of the genomic locus with CRISPR/Cas9, led to differential processing. The biogenesis of miRNAs is a multistep and complex process regulated by a number of RBPs that participate in every step of the maturation. Nonetheless, only a small fraction are characterised. Given the plethora of miRNAs and the complicated clusters they form it becomes apparent that various RBPs partake in the miRNA biogenesis.

It has been previously shown that certain RBPs interact with specific miRNAs to regulate their maturation, with the two most prominent examples being Drosha and Dicer. Although previously published studies had reported a panel of RBPs that participate in the maturation of the miRNAs, pri-miR-497~195 had not been studied
in detail. To elucidate these differences, RNA pull-down experiments were conducted comparing the RBPs that recognise and bind the WT and a mutant primary transcript harbouring a 16nt deletion. The latter transcript, produced with CRISPR/Cas9 editing, was missing both the sequence of the 5p arm of the mmu-miR-195a as well as that of a bulge upstream of the DSB. Secondary structure prediction confirmed the effects of the editing on the structure of the miRNA while the tertiary structure was highly different compared to the WT one. These changes were deemed detrimental for the maturation of the primary transcript by affecting the approach by RBPs.

To gain deeper knowledge in the regulation of RBPs on the miRNA maturation this study focused on characterising the RBPs that interact with the WT pri-miR-195a. To this end, MS was employed to identify differential binding of RBPs to the mutant primary transcript of mmu-miR-195a~497.

The MS data presented, offer a unique description of the RBPs that interact with both the WT and the mutant transcript. In total, 45% of the identified proteins are annotated in UNIPROT database as RNA binding proteins with the small number of the RBPs being suggestive either of innate limitations of the assay or of specific protein-protein interactions between RBPs and other co-factors. Furthermore, 50% of the proteins identified are localised either in the nucleus or the cytosol with the vast majority of them being nuclear proteins, as expected. Further analysis showed that only 5 of the identified proteins were statistically significant.

Taken together these results indicate that the mutant transcript follows a different processing pathway. Although further in silico analysis did not show any interaction between the identified proteins, they can be organised in two groups. Tial1 and Apex1 are RBPs that recognise specific motifs in ssRNA molecules and help in their processing while Ran, Sec63 and Copb2, regulate the identification and transportation of proteins to Golgi network.

In addition, 12 RBPs that bind solely the mutant transcript were identified and when analysed with STRING, a network of experimental validated protein-protein interactions was highlighted. The analysis revealed that 7 of the input proteins form a network of interacting RBPs while a second clustering became apparent when they
were grouped on the basis of their function. Proteins that partake in pre-mRNA splicing form a group of closer interacting members results that were confirmed when the identified proteins were analysed with Cytoscape.

In light of these results, it becomes apparent that said proteins participate in a different than the canonical pathway of the processing of miRNAs. Moreover, results from the in silico analysis of the interactions of the MS analysed proteins suggest that the mutant primary transcript takes part in a pathway that leads to its degradation and clearance from the cells results that explain the decrease in the levels of both mmu-miR-497a and mmu-miR-195a.

In conclusion, MS analysis of RNA pull-down with WT and mutant transcripts of pri-miR-497~195 revealed a set of proteins that bind the pri-miR-497~195 together with RBPs that bind only the mutant transcript. Although the results are only indicative they suggest that alternative processing pathways may lead to the downregulation of the levels of mature mmu-miR-497a and mmu-miR-195a.

These results provide a helpful insight in the maturation process of the miRNAs in general and in that of pri-miR-497~195 specifically. They demonstrate that sequence changes lead to differential processing because the tertiary changes that they impose. This conclusion was drawn by the fact that firstly the RBPs that recognise the transcript seem to recognise specific structures and not sequence motifs and secondly no known motifs were created from the deletion. Moreover, pathway analysis of the proteins demonstrated that these proteins partake in the splicing step of miRNAs. This could mean that the structure of the mutant transcript hinders the recognition of the transcript from the miRNA processing machinery, resulting in the transcript not being recognising as a pre-miRNA leading to a new pathway getting triggered through which the mutant transcript gets degraded. However, these interactions that were described can be due to the non-physiological conditions of the assay, such as the lack of cell compartmentalisation, high concentration of primary transcript and the in vitro conditions of the interactions that do not correspond to the physiological ones.
7.2 Limitations

The current study investigates the use of CRISPR/Cas9 as a gene editing technique for the specific inhibition of miRNAs. It is the first time it has been shown that CRISPR/Cas9 can be used in primary mammalian cells for the inhibition of miRNAs. Although a reliable and robust workflow is provided, with regards to inducing gene editing in mouse VSMCs to dissect the regulation and function of clustered miRNAs, there are limitations that have to be taken into consideration.

An important limitation is the usage of mouse VSMCs and human SMCs cell line instead of human primary SMCs. The study focused on editing miRNA gene in human SMCs and, to that end, human SMCs cell lines were used for the initial experiments, since they were a limitless source of cells. Nonetheless, and despite the fact that the scope of the study was to examine the effects of editing miR-195 with CRISPR/Cas9 in human primary VSMCs, the low expression of the miRNAs of the miR-15 family, after prolonged culture, conditions that were needed in our study, was an important issue. Prolonged culture of VSMCs was needed in order for the cells to reach senescence and a phenotype as close to AAA. Due to this, the focus of the study shifted from human to mouse VSMCs that expressed the studied miRNAs at satisfactory levels. Additionally, due to the differences of the miRNA gene sequences between the two organisms new guides had to be designed and tested, guides that, due to sequence limitations, cannot be useful for human studies.

Moreover, the delivery of Cas9 to mouse VSMCs was an important step for the study, since they express all of the miRNAs that were evaluated. Although the approach that was finally undertaken resulted in engineering the genome of the cells and in the accumulation of Cas9 protein, that did not affect the phenotype or the proliferation rate (data not shown). Moreover, Cas9 was not expected to lead to any editing unless in complex with a sgRNA. In more detail, the binding between the endonuclease and the guide RNA is dependent on an interaction between the tracrRNA and the protein through a specific sequence. During the maturation process of the crRNA:tracrRNA, Cas9 recognises a sequence in the tracrRNA which binds and then uses the latter’s complementarity to crRNA to cleave it. The crRNA:tracrRNA duplex stays tethered to Cas9 with the help of tracrRNA$^{292}$. Since Cas9 is an exogenous protein of bacterial
origin it was expected that it would not get activated unless bound to a sgRNA introduced to the cells after transfection. Furthermore, the transient presence of the IVT sgRNA in the cells was thought to protect against potential off-target effects due to the high concentration of the Cas9. Indeed, no off-target effects were detected in any of the possible off-target sites, that were in silico predicted, for any of the guides tested.

Delivery of the system is one of the major limitations, especially when it comes to in vivo models. In order to address these difficulties various groups have developed new, more elegant techniques over the last years that help the uptake of the RNPs by the cells. Regardless, the technique employed in the present study was deemed suitable for the purpose it was used as sufficient editing occurred with no off-target effects in the 7 most likely targets, according to in silico analysis, and no other complications. Although the computationally predicted off-target sites that were tested, were confirmed by two software, the chances that the guides used could lead to non-identified editing in other sites still remains. Additionally, the list of the tested potential off-targets are by no means exhaustive. Hence, whole genome sequencing of edited compared to non-edited cells would possibly provide more convincing results.

Nevertheless, following the generation of mouse VSMCs stably expressing Cas9, the effects of editing the genomic loci of 4 different clusters were assessed. The clusters were chosen based on their implication in cardiovascular disease and their organisation in homo-clusters or hetero-clusters.

Editing of one of the miRNAs of the cluster led to unexpected downregulation of the second clustered miRNA in the case of the mmu-miR-497~195 and mmu-miR-143~145 clusters, while editing of one miRNA stem loop did not show any differences in the regulation of clustered miRNAs in the case of mmu-miR-17~92 and mmu-miR-106b~25. No off-target editing was detected in either case. Additionally, computational modelling suggests important changes in the structure of the primary transcript that was hypothesised to hinder its processing by the maturation machinery. The study concludes that the expression of the clustered miRNAs is not necessarily co-dependent as well as that clusters show strict tertiary structural requirements for their expression.
Although the study underscores the power of CRISPR/Cas9 in the knocking down of miRNAs over conventional methods that can target multiple related paralogs, the approach remains rather imprecise both in the investigation of the editing and its outcome. In regards to the former, the T7EI assay was used for the calculation of the editing efficiency as a technique has innate limitations. Firstly, it can underestimate the final efficiency due to fact that it cannot recognise homodimers of edited strands resulting in those strands not being cut and quantified. Secondly, the technique is based on gel separation of the bands and their intensity after the electrophoresis depends on both their size and the electrophoresis settings. In detail, the smaller sized bands during the electrophoresis get diffused easier and furthermore appear less intense due to less staining dye being bound by smaller DNA strands while high voltage of electrophoresis leads to the gel developing high temperatures and higher diffuse rate. Although the way that the technique was designed accounted for these limitations i.e. the PCR primers were designed so that the predicted cut site was upstream of the middle of the amplicon but as close as possible to it, in order for the two resulting bands to differ as little as possible from each other in size, and the used calculating formula adjusted for these discrepancies, the limitations still remain and need to be taken into consideration. However, at the time of the assessment of the editing efficiency of these experiments, the T7EI assay was the gold standard assay. In order to overcome the aforementioned limitations new techniques have been developed e.g. TIDE/ICE analysis or analysis of Amplicon sequencing data with bioinformatic tools. By employing the NHEJ pathway the formation of mutations, that result in the disruption of the miRNA sequence, is random, not only regarding the type (insertions or deletions) but also regarding the length and the exact position. Although an adequate number of colonies were picked and screened to reveal sequence changes upon induction of editing, these were but a small portion of the occurred alterations. A wider range of mutations was expected to have been induced. In the same line, the structure prediction models presented above are only indicative of a very small number of changes. However, these conclusions were based on a single 3D computational prediction, that of pri-miR-497~195, due to limitations of the software. The length of the cluster miR-143~145 is about 1400bp therefore it cannot be analysed using the currently available algorithms. In addition, changes in the secondary structure due to sequence motifs alterations could affect the expression of the clustered miRNAs.
To address these limitations the study focused on mmu-miR-195a, the miRNA of interest, and sought to decipher how the editing affects the processing of pri-miR-497~195. Proteins that bound to the WT and the mutant transcript were evaluated and analysed with MS. Analysis of the results showed a panel of proteins that bind preferentially to the mutant transcript. Further analysis revealed that the identified proteins bind RNA molecules based on their tertiary structure since, according to published data on their function, they do not recognise specific secondary structures or sequence motifs that were altered upon editing. These results indicate that the differential processing of the mutant transcript was due to changes in the tertiary and not the secondary structure or the abolition of any sequence motif.

Nonetheless, it is likely that the proteins identified do not represent interactions that occur in the edited cells. This could be for a number of reasons. Firstly, it is an in vitro assay which means that its conditions under which the proteins, approach, interrogate and bind the primary transcript are different. Additionally, the transcripts that were used in the assay were in vitro transcribed which means that the transcripts were isolated from possible post-transcriptional processing that could affect its binding by the RBPs and, subsequently, its maturation\textsuperscript{221}. In more detail, it has been shown that RNA-folding, especially for larger miRNA clusters, affects the accessibility of the processing machinery is important\textsuperscript{293, 294} therefore, isolating the primary transcripts from the cellular environment can result in differential folding and processing. The concentration of the transcripts was a lot higher than the one that occurs in vivo and the protein population of the cells get exposed to the transcript simultaneously which, together with the lack of cell compartmentalisation, means that the interactions between the proteins and the RNA transcripts are not necessarily canonical.

Moreover, although MS is a highly sensitive technique that can identify a vast number of proteins, even in low concentration, it is semi-quantitative. It can only provide data with regards to the detection of proteins and their relative, but not absolute, abundance. Given the above, WB was employed to confirm the MS results and further elucidate the abundance of proteins. Nonetheless, the validation of the MS data failed therefore additional validation is required to experimentally test the concept that emerged from the in silico analysis.
All in all, besides the limitations this study presents, it provides strong evidence of the suitability of the CRISPR/Cas9 editing to inhibit miRNAs, via permanent alterations of their genes, adds new concepts to the miRNA study and offers well-founded results regarding the processing of primary miRNA transcripts in primary mammalian cells.

### 7.3 Future Outlook

The current study offers an insightful groundwork on how editing with CRISPR/Cas9 can be employed for the inhibition of miRNAs. The problem of delivering the system in high efficiency was resolved and no more titration is needed in that front. Additionally, off-target editing seemed not to be an issue in the 7 most likely off-target sites. However, in light of new studies that have shown that CRISPR/Cas9 can result in non-predicted, unintended editing in various, random sites\textsuperscript{193,194}, further scrutiny is needed. In more detail, whole genome sequencing has emerged as the method of choice for investigating possible off-target editing. Further investigation using whole genome sequencing would clarify the unintended editing and resolve the issue.

The main aim of the study was to overcome the limitations of the traditional methods for inhibiting specific miRNAs. The results outlined above demonstrated that, although editing with CRISPR/Cas9 can lead to inhibiting miRNAs, the effects of editing differ according to the cluster that is targeted with certain clusters showing abolition of the expression of all the miRNAs that are co-transcribed. Clustered miRNAs have attracted intense interest in the biomedical research and the findings presented here on clustered miRNA regulation offer valuable insights. Secondary and tertiary structure prediction models have shown that the deletion of sequences of the hairpin of miRNAs result in different structures of the stem loop. However, given the small number of the edited sequences that were studied, the presented results can only be suggestive of the effects of the editing in the maturation of the miRNAs. Therefore, additional sequences harbouring deletions need to be analysed \textit{in silico} with regards to how sequence alterations change the secondary and tertiary structures of the hairpin and the primary transcript.

Proteins interact with RNA transcripts through recognition of specific binding motifs with these interactions being significantly influenced by the tertiary structure on the
RNA molecules. RNA pull-down assays are useful in studying these interactions by providing the possibility of extracting RBP–RNA complex for MS or western blotting by exploiting high affinity tags like biotin. Moreover, validation of the MS results with Western Blotting, as an independent method, was not possible since the antibodies used did not produce any conclusive results. Therefore, more titration experiments will resolve the issues raised and will provide a better understanding of the interactions of the identified proteins. In addition, given that this the first time that the maturation of the pri-miR-497~195 is being investigated, these results will provide a list of validated proteins that are important for the process. Quantification of the miRNA levels and MS analysis in experiments where these proteins will be knocked-down, will offer a valuable insight in their effects on the process.

Additionally, RNA pull-down assays and MS analysis of the eluted proteins demonstrated differential binding of certain RBPs. However, only one mutant sequence was analysed. New RNA pull-down assays with more mutant transcripts, harbouring deletions of various lengths and positions, excising various structural features of the hairpins need to be conducted to elucidate the contribution of each structural element in the recognition by RBPs. Thus, supplemental experiments are needed to decipher the mechanisms that affect the expression of the miRNAs especially since studies have shown that the secondary structure of the pri-miR-17~92 influences the processing of the transcript by following a highly specific sequence of cleavage of the clustered stem loops.

In general, the experimental approach outlined in the study can be proved to be useful for the clarification of the maturation of mmu-miR-195a, as well as other miRNAs, and the identification of the involved RBPs.
BIBLIOGRAPHY

33. Forstemann K, Tomari Y, Du T, Vagin VV, Denli AM, Bratu DP, Klattenhoff C, Theurkauf WE and Zamore PD. Normal microRNA maturation and germ-line stem cell


246. Cai X, Hagedorn CH and Cullen BR. Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. *RNA (New York, NY)*. 2004;10:1957-66.


296. https://github.com/staciawyman/cortado
SUPPLEMENT
## Supplementary Table 1. Details of mouse miRNA clusters

Data on miRNA clusters were obtained from the MiRBase database.

<table>
<thead>
<tr>
<th>miRNA</th>
<th>miRNA Clusters</th>
<th>Coordinates</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>mir-15a</td>
<td></td>
<td>chr14: 61632027-61632110 [-]</td>
<td>UAGCAGCACUAUAAUGGUUUGUG</td>
</tr>
<tr>
<td>mir-15b</td>
<td></td>
<td>chr3: 69009772-69009835 [+]</td>
<td>UAGCAGCACUAUAAUGGUUUAC</td>
</tr>
<tr>
<td>mir-16-2</td>
<td></td>
<td>chr3: 69009902-69009996 [+]</td>
<td>UAGCAGCACGUAUAAUUGGCG</td>
</tr>
<tr>
<td>mir-16-1</td>
<td></td>
<td>chr14: 61631880-61631972 [-]</td>
<td>UAGCAGCACGUAUAAUUGGCC</td>
</tr>
<tr>
<td>mir-195</td>
<td></td>
<td>chr11: 70235042-70235135 [+]</td>
<td>UAGCAGCACAGAAAUUGGC</td>
</tr>
<tr>
<td>mir-497a</td>
<td>miR-15</td>
<td>chr11: 70234717-70234800 [+]</td>
<td>CAGCAGCACACUGGUGGUUGUA</td>
</tr>
<tr>
<td>mir-497b</td>
<td></td>
<td>chr11: 70234692-70234816 [-]</td>
<td>CACCACAGUGGUUGGUUGGAGG</td>
</tr>
<tr>
<td>mir-143</td>
<td>miR-143-145</td>
<td>chr18: 61649196-61649258 [-]</td>
<td>GGUCAGAGCUCAGACUCUAG</td>
</tr>
<tr>
<td>mir-145</td>
<td></td>
<td>chr18: 61647825-61647894 [-]</td>
<td>GUCAGAGUUCCCAGAAUUCCU</td>
</tr>
<tr>
<td>mir-17</td>
<td>miR-17–92</td>
<td>chr14: 115043671-115043754 [+]</td>
<td>CAAGAGAGCUUACAGUGCAGGUAG</td>
</tr>
<tr>
<td>mir-20a</td>
<td></td>
<td>chr14: 115044157-115044263 [+]</td>
<td>UAAAGAGCUUACAGUGCAGGUAG</td>
</tr>
<tr>
<td>mir-106b</td>
<td></td>
<td>chr5: 138165737-138165818 [-]</td>
<td>UAAAGAGCUGACAGUGCAGAU</td>
</tr>
<tr>
<td>mir-93</td>
<td></td>
<td>chr5: 138165523-138165610 [-]</td>
<td>CAAAGAGCGUGUCGUGCAGGUAG</td>
</tr>
<tr>
<td>mir-92a1</td>
<td></td>
<td>chr14: 115044427-115044506 [+]</td>
<td>AGGUUUGGAAUUGUGCGAUCAGAGU</td>
</tr>
<tr>
<td>mir-25</td>
<td></td>
<td>chr5: 138165321-138165404 [-]</td>
<td>AGGCAGGAGCAUGGGCAUUUC</td>
</tr>
<tr>
<td>mir-19a</td>
<td></td>
<td>chr14: 115044000-115044081 [+]</td>
<td>UAGUUUGCAUGUUGUGCAGAUC</td>
</tr>
<tr>
<td>mir-19b-1</td>
<td></td>
<td>chr14: 115044305-115044391 [+]</td>
<td>AGUUUUGCAGGUGGUUGCAUCCAGA</td>
</tr>
<tr>
<td>mir-18a</td>
<td></td>
<td>chr14: 115043851-115043946 [+]</td>
<td>UAAGGUGCAUCUAGUGCAGAUAG</td>
</tr>
</tbody>
</table>
Supplementary Table 2. IVT primers. Sequences of the primers used for the IVT of the guides along with the coordinates that they target in the mouse genome.

<table>
<thead>
<tr>
<th>Assay</th>
<th>Primer</th>
<th>Sequence (5’ to 3’)</th>
<th>Coordinates</th>
</tr>
</thead>
<tbody>
<tr>
<td>IVT</td>
<td>F 195G1</td>
<td>TAATACGACTCACTATAGTTGAGGCAGAATTACTCC</td>
<td></td>
</tr>
<tr>
<td>IVT</td>
<td>R 195G1</td>
<td>TTCTAGCTCTAAAAACGGGAGTAGTTCTGCTCAA</td>
<td></td>
</tr>
<tr>
<td>IVT</td>
<td>F 195G2</td>
<td>TAATACGACTCACTATAGGAAGCGAGTCTGCAATAT</td>
<td></td>
</tr>
<tr>
<td>IVT</td>
<td>R 195G2</td>
<td>TTCTAGCTCTAAAAACATATTGGCGAGACTGCTTTCC</td>
<td></td>
</tr>
<tr>
<td>IVT</td>
<td>F 195m2</td>
<td>TAATACGACTCACTATAGGGATCTCTAAAACATATTGGCGAGACTGCTTTCC</td>
<td></td>
</tr>
<tr>
<td>IVT</td>
<td>R 195m2</td>
<td>TTCTAGCTCTAAAAACATATTGGCGAGACTGCTTTCC</td>
<td></td>
</tr>
<tr>
<td>IVT</td>
<td>F 195mA</td>
<td>TAATACGACTCACTATAGGGATCTCTAAAACATATTGGCGAGACTGCTTTCC</td>
<td></td>
</tr>
<tr>
<td>IVT</td>
<td>R 195mA</td>
<td>TTCTAGCTCTAAAAACATATTGGCGAGACTGCTTTCC</td>
<td></td>
</tr>
<tr>
<td>IVT</td>
<td>F 195m3</td>
<td>TAATACGACTCACTATAGGGATCTCTAAAACATATTGGCGAGACTGCTTTCC</td>
<td></td>
</tr>
<tr>
<td>IVT</td>
<td>R 195m3</td>
<td>TTCTAGCTCTAAAAACATATTGGCGAGACTGCTTTCC</td>
<td></td>
</tr>
<tr>
<td>IVT</td>
<td>F 145m1</td>
<td>TAATACGACTCACTATAGGGATCTCTAAAACATATTGGCGAGACTGCTTTCC</td>
<td></td>
</tr>
<tr>
<td>IVT</td>
<td>R 145m1</td>
<td>TTCTAGCTCTAAAAACATATTGGCGAGACTGCTTTCC</td>
<td></td>
</tr>
<tr>
<td>IVT</td>
<td>F 145m2</td>
<td>TAATACGACTCACTATAGGGATCTCTAAAACATATTGGCGAGACTGCTTTCC</td>
<td></td>
</tr>
<tr>
<td>IVT</td>
<td>R 145m2</td>
<td>TTCTAGCTCTAAAAACATATTGGCGAGACTGCTTTCC</td>
<td></td>
</tr>
<tr>
<td>IVT</td>
<td>F 18a</td>
<td>TAATACGACTCACTATAGGGATCTCTAAAACATATTGGCGAGACTGCTTTCC</td>
<td></td>
</tr>
<tr>
<td>IVT</td>
<td>R 18a</td>
<td>TAATACGACTCACTATAGGGATCTCTAAAACATATTGGCGAGACTGCTTTCC</td>
<td></td>
</tr>
<tr>
<td>IVT</td>
<td>F 25</td>
<td>TAATACGACTCACTATAGGGATCTCTAAAACATATTGGCGAGACTGCTTTCC</td>
<td></td>
</tr>
<tr>
<td>IVT</td>
<td>R 25</td>
<td>TAATACGACTCACTATAGGGATCTCTAAAACATATTGGCGAGACTGCTTTCC</td>
<td></td>
</tr>
</tbody>
</table>
Supplementary Table 3. Table of primers. Primers that were used for genomic PCR, cloning in pLKO.1 vector and qPCR for gene expression with the corresponding coordinates of the amplified genomic loci.

<table>
<thead>
<tr>
<th>Assay</th>
<th>Primer</th>
<th>Sequence (5' to 3')</th>
<th>Coordinates</th>
</tr>
</thead>
<tbody>
<tr>
<td>GENOMIC PCR</td>
<td>hsa-mir-195 F</td>
<td>GGCGCTTGTGACAAACTCTTCT</td>
<td>chr17:7017431-7017863</td>
</tr>
<tr>
<td>GENOMIC PCR</td>
<td>hsa-mir-195 R</td>
<td>GCTATTCCCGCATAAGCATC</td>
<td></td>
</tr>
<tr>
<td>GENOMIC PCR</td>
<td>mmu-mir-195 F (P1)</td>
<td>CACACACACACCGCTCAGGG</td>
<td>chr11:70234880-70235407</td>
</tr>
<tr>
<td>GENOMIC PCR</td>
<td>mmu-mir-195 R (P2)</td>
<td>CTGAGCCTTCCACCTGAC</td>
<td></td>
</tr>
<tr>
<td>GENOMIC PCR</td>
<td>mmu-mir-497 F (P3)</td>
<td>CCTGTGTCTTCCAGCATTTTTC</td>
<td>chr11:70234360-70234977</td>
</tr>
<tr>
<td>GENOMIC PCR</td>
<td>mmu-mir-497 R (P4)</td>
<td>GTATCAGACACCTGGGGGT</td>
<td></td>
</tr>
<tr>
<td>GENOMIC PCR</td>
<td>mmu-mir-143 F (P5)</td>
<td>GTCGCTGCTGATAAGAGA</td>
<td>chr18:61649137-61649567</td>
</tr>
<tr>
<td>GENOMIC PCR</td>
<td>mmu-mir-143 R (P6)</td>
<td>GCTATCCATGCAAACACTT</td>
<td></td>
</tr>
<tr>
<td>GENOMIC PCR</td>
<td>mmu-mir-145 F (P7)</td>
<td>CTTTCCAAGCCACTCAAGC</td>
<td>chr18:61647566-61648054</td>
</tr>
<tr>
<td>GENOMIC PCR</td>
<td>mmu-mir-145 R (P8)</td>
<td>GGAGCCGTCCTGATGTCTGG</td>
<td></td>
</tr>
<tr>
<td>GENOMIC PCR</td>
<td>mmu-mir-18 F (P9)</td>
<td>CCTGCTCAATGTGAGGCTT</td>
<td>chr14:115043309-115044386</td>
</tr>
<tr>
<td>GENOMIC PCR</td>
<td>mmu-mir-18 R (P10)</td>
<td>CCACAGTCAGTTTTTGACTG</td>
<td></td>
</tr>
<tr>
<td>GENOMIC PCR</td>
<td>mmu-mir-25 F (P11)</td>
<td>TTCTCCGACTTCCACCTG</td>
<td>chr5:138164956-138165932</td>
</tr>
<tr>
<td>GENOMIC PCR</td>
<td>mmu-mir-25 R (P12)</td>
<td>GCCACAAACAGTAGACAGCAA</td>
<td></td>
</tr>
<tr>
<td>pLKO.1 cloning</td>
<td>miR-497~195 F</td>
<td>GACCGGTTGCTGATGTCTGTTG</td>
<td>chr11:70048081-70048911</td>
</tr>
<tr>
<td>pLKO.1 cloning</td>
<td>miR-195 R</td>
<td>GATGAATTCAAAAAACTGAGCCTTCCACCTGAC</td>
<td></td>
</tr>
<tr>
<td>Gene expression</td>
<td>Carmn F</td>
<td>AATAGACTGAGCGCTCAGC</td>
<td></td>
</tr>
<tr>
<td>Gene expression</td>
<td>Carmn R</td>
<td>GTTCTCCTGACGAGAGG</td>
<td></td>
</tr>
<tr>
<td>Gene expression</td>
<td>Beta-actin F</td>
<td>CACAACTGGGACAGCATGAGGG</td>
<td></td>
</tr>
<tr>
<td>Gene expression</td>
<td>Beta-actin R</td>
<td>TTCATGAGGTAGTCAGTCTG</td>
<td></td>
</tr>
</tbody>
</table>
**Supplementary Table 4. Table of restriction sites.** The restriction sites, along with the digestion enzymes that are recognised by, for each miRNA and targeting guide used.

<table>
<thead>
<tr>
<th>miRNA</th>
<th>sgRNA</th>
<th>sgRNA sequence (5'-3’)</th>
<th>Restriction site (5'-3’)</th>
<th>Restriction enzyme</th>
</tr>
</thead>
<tbody>
<tr>
<td>mmu-miR-195a</td>
<td>sg195m2</td>
<td>TGGAGCAGCACACAGCCAATAT</td>
<td>AATATT</td>
<td>SspI</td>
</tr>
<tr>
<td></td>
<td>sg195mA</td>
<td>ATTTCTGTGCTGCTAGAGCC</td>
<td>GCNNNNNNNGC</td>
<td>MwoI</td>
</tr>
<tr>
<td>mmu-miR-145</td>
<td>sg145m2</td>
<td>GATGCTAAGATGGGGATTCC</td>
<td>GANTC</td>
<td>HinfI</td>
</tr>
<tr>
<td>mmu-miR-18</td>
<td>sg18</td>
<td>TTATGCCAGAAGGAGCAGTT</td>
<td>GDGCHC</td>
<td>Bsp1268I</td>
</tr>
<tr>
<td>mmu-miR-25</td>
<td>sg25</td>
<td>GCGGAGACTTGCGCAATTGC</td>
<td>CAATTG</td>
<td>MfeI</td>
</tr>
</tbody>
</table>
**Supplementary Table 5. Table of possible off-target sites.** The genomic loci that were possibly targeted by the guides used along with the mismatches (depicted in red), the primers used to study them and their coordinates.

<table>
<thead>
<tr>
<th>Assay</th>
<th>Gene</th>
<th>OFF-TARGET Sequence</th>
<th>Mismatches</th>
<th>Primer Sequence (5' to 3')</th>
<th>Mismatches</th>
</tr>
</thead>
<tbody>
<tr>
<td>GENOMIC PCR</td>
<td>OFFTARGET1 195m2</td>
<td>AGGAGCAGTGAGGCACTAT</td>
<td>3MMs [1:9:10]</td>
<td>OFT1 195M2 F GTGCTGGGATTACAGGCATTT</td>
<td>chr17::3133199</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>OFT1 195M2 R TCCTTGGATGAGGGGCTA</td>
<td></td>
</tr>
<tr>
<td>GENOMIC PCR</td>
<td>OFFTARGET2 195m2</td>
<td>AGGAGCAACAGGCACTAT</td>
<td>3MMs [1:8:17]</td>
<td>OFT2 195M2 F GCCCCAGAGTGAGTGAAAGT</td>
<td>chr8::10246601</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>OFT2 195M2 R TCAAGCCTCCTGAATGCACAC</td>
<td></td>
</tr>
<tr>
<td>GENOMIC PCR</td>
<td>OFFTARGET3 195m2</td>
<td>ATGAGCAGCAGGCACTAT</td>
<td>3MMs [1:2:15]</td>
<td>OFT3 195M2 F TCCACATGGAGCAGTGTGTT</td>
<td>chr14::73777792</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>OFT3 195M2 R GCTCCTGTTGCTAGGTGTA</td>
<td></td>
</tr>
<tr>
<td>GENOMIC PCR</td>
<td>OFFTARGET4 195m2</td>
<td>TGGAGCAACAGCAGGCACTAT</td>
<td>3MMs [8:9:11]</td>
<td>OFT4 195M2 F ATTCCTGGGAAATTTCCATC</td>
<td>chrX::+72165663</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>OFT4 195M2 R TGCACTGCTGTCACACAAA</td>
<td></td>
</tr>
<tr>
<td>GENOMIC PCR</td>
<td>OFFTARGET5 195m2</td>
<td>TGGAGCATCACAGGCACTAT</td>
<td>2MMs [8:17]</td>
<td>OFT5 195M2 F TATTTCTGCTGCTTTGCAC</td>
<td>chr2::+24952183</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>OFT5 195M2 R GCAGAGTGCTGTTGAGA</td>
<td></td>
</tr>
<tr>
<td>GENOMIC PCR</td>
<td>OFFTARGET6 195m2</td>
<td>TGCAAGCAGCAGGCACTAT</td>
<td>3MMs [3:11:17]</td>
<td>OFT6 195M2 F CGACTTGCAAGACCATGGAG</td>
<td>chr13::+35216456</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>OFT6 195M2 R CACCCACTCGGGACCTACTA</td>
<td></td>
</tr>
<tr>
<td>GENOMIC PCR</td>
<td>OFFTARGET7 195m2</td>
<td>AGGATTTAGCTGAGGCACTAT</td>
<td>4 MMs [1:5:6:17]</td>
<td>OFT7 195M2 F AGTGCTGGGCACTAGGACAAC</td>
<td>chr2::+91834510</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>OFT7 195M2 R CTGCTCTGTCAGGTTTGGA</td>
<td></td>
</tr>
<tr>
<td>GENOMIC PCR</td>
<td>OFFTARGET1</td>
<td>CTG TAGCA AACACAGAAATAT</td>
<td>2 MMs [3:9]</td>
<td>OFT1 195M3 F</td>
<td>TGGGT TGAAAGGTCTCTGAAG</td>
</tr>
<tr>
<td>GENOMIC PCR</td>
<td>OFFTARGET1</td>
<td>CTCTAGCAG GACAGAAATAA</td>
<td>2 MMs [10:20]</td>
<td>OFT2 195M3 F</td>
<td>TCTGCCCTCA GCTTTCCAAG</td>
</tr>
<tr>
<td>GENOMIC PCR</td>
<td>OFFTARGET3</td>
<td>TTAAAGCAG CACAGAAATAT</td>
<td>3 MMs [1:3:4]</td>
<td>OFT3 195M3 F</td>
<td>TCCCTGGACTTTGACATTC</td>
</tr>
<tr>
<td>GENOMIC PCR</td>
<td>OFFTARGET4</td>
<td>ACCTAGCAG CTCAAAATAT</td>
<td>3 MMs [1:2:11]</td>
<td>OFT4 195M3 F</td>
<td>TGGAGAAGGAGGAGGTCTGA</td>
</tr>
<tr>
<td>GENOMIC PCR</td>
<td>OFFTARGET5</td>
<td>GTCTACCAGA AACAGAAATAT</td>
<td>3 MMs [1:6:10]</td>
<td>OFT5 195M3 F</td>
<td>CCATT TTGCATGGAAGGATT</td>
</tr>
<tr>
<td>GENOMIC PCR</td>
<td>OFFTARGET6</td>
<td>CTCTGTCTGACAGAAATAT</td>
<td>3 MMs [5:6:8]</td>
<td>OFT6 195M3 F</td>
<td>GGGAGACTATT GTGCATGATTTTATT</td>
</tr>
</tbody>
</table>
### Supplementary Table 6. RNA binding proteins with their corresponding RNA binding motifs.

The panel of the RNA binding proteins with RNA binding motifs that could potentially bind the primary mmu-miR-195a transcript.

<table>
<thead>
<tr>
<th>RBP</th>
<th>RNA Binding motifs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hnrnpl</td>
<td>CACAC</td>
</tr>
<tr>
<td>Hnrnph2</td>
<td>GGAGG</td>
</tr>
<tr>
<td>RBM6</td>
<td>AUCCA</td>
</tr>
<tr>
<td>Hnrnpa1</td>
<td>UAGGG AGGG UGGAU</td>
</tr>
<tr>
<td>Hnrnpa2b1</td>
<td>GGAGG UGGAU</td>
</tr>
<tr>
<td>G3bp2</td>
<td>UGGAU</td>
</tr>
<tr>
<td>Tra2a</td>
<td>ACCCA UGUA</td>
</tr>
<tr>
<td>Cpsf3</td>
<td>GCAUG</td>
</tr>
<tr>
<td>Nudt21</td>
<td>AUCUU</td>
</tr>
<tr>
<td>Rbfox2</td>
<td>AGCG</td>
</tr>
<tr>
<td>Matr3</td>
<td>UGUNNNNNNUGU</td>
</tr>
<tr>
<td>Zc3h10</td>
<td>AGCG</td>
</tr>
<tr>
<td>Celf1</td>
<td>UGUNNNNNNUGU</td>
</tr>
</tbody>
</table>
### Supplementary Table 7. Statistical analysis of the MS identified proteins after RNA pull-down.

The information from the statistical analysis of the MS data regarding the proteins that were identified. The proteins are presented in order, according to their p-values, following FDR correction, while the p-values for the statistically significant proteins are highlighted in bold. The RBPs that bound only the mutant transcript are highlighted in grey.

<table>
<thead>
<tr>
<th>PROTEIN NAME</th>
<th>ACCESSION NUMBER</th>
<th>UNIPROT ID</th>
<th>MW (kDa)</th>
<th>MAX sequence coverage</th>
<th>UNIQUE PEPTIDES</th>
<th>TOTAL PSMs</th>
<th>CONTROL (Mean±SD)</th>
<th>MUTANT (Mean±SD)</th>
<th>P VALUE</th>
<th>FOLD CHANGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coatomer subunit beta OS=Mus musculus GN=Copb2 PE=1 SV=2</td>
<td>Q9DC23</td>
<td>DNAJC10_MOUSE</td>
<td>68.5</td>
<td>19.4%</td>
<td>14</td>
<td>19</td>
<td>6109550±617951</td>
<td>10591300±6181539</td>
<td>0.117490648</td>
<td>2.85</td>
</tr>
<tr>
<td>DNAJC10_homolog subfamily C member 10 OS=Mus musculus GN=Dnaic10 PE=1 SV=2</td>
<td>Q8DFS9</td>
<td>EZR_MOUSE</td>
<td>69.4</td>
<td>16.6%</td>
<td>13</td>
<td>17</td>
<td>8582375±3232269</td>
<td>14253750±2956999</td>
<td>0.117490648</td>
<td>1.66</td>
</tr>
<tr>
<td>Replication factor C subunit 5 OS=Mus musculus GN=Rcs5 PE=1 SV=1</td>
<td>Q9JKB3</td>
<td>YB1_MOUSE</td>
<td>38.8</td>
<td>44.3%</td>
<td>10</td>
<td>14</td>
<td>10120000±1038752</td>
<td>14558000±836357</td>
<td>0.117490648</td>
<td>1.44</td>
</tr>
<tr>
<td>Radixin OS=Mus musculus GN=Rdx PE=1 SV=3</td>
<td>Q9H9A5</td>
<td>PABPN1_MOUSE</td>
<td>32.2</td>
<td>11.9%</td>
<td>4</td>
<td>4</td>
<td>7896850±564314</td>
<td>14105500±2056670</td>
<td>0.117490648</td>
<td>1.79</td>
</tr>
<tr>
<td>Ezrin OS=Mus musculus GN=Ezr7 PE=1 SV=3</td>
<td>Q9M8S5</td>
<td>XR1_MOUSE</td>
<td>104.2</td>
<td>9.19%</td>
<td>14</td>
<td>14</td>
<td>3337150±5604165</td>
<td>8627125±3290583</td>
<td>0.117490648</td>
<td>2.59</td>
</tr>
<tr>
<td>Y-box-binding protein 3 OS=Mus musculus GN=Ybx3 PE=1 SV=2</td>
<td>Q9Y1W8</td>
<td>FUBP1_MOUSE</td>
<td>68.5</td>
<td>24.3%</td>
<td>12</td>
<td>28</td>
<td>25273600±1039068</td>
<td>40113500±4730225</td>
<td>0.117490648</td>
<td>1.59</td>
</tr>
<tr>
<td>Cluster of Polyadenylate-binding protein 2 OS=Mus musculus GN=Pabpn1 PE=1 SV=3 (Polyadenylate-binding)</td>
<td>P97789</td>
<td>XR1_MOUSE</td>
<td>104.2</td>
<td>9.19%</td>
<td>14</td>
<td>14</td>
<td>3337150±5604165</td>
<td>8627125±3290583</td>
<td>0.117490648</td>
<td>2.59</td>
</tr>
<tr>
<td>S13' endonuclease 1 OS=Mus musculus GN=Snm1 PE=1 SV=1</td>
<td>Q9Y1M5</td>
<td>RBM11_MOUSE</td>
<td>42.1</td>
<td>9.28%</td>
<td>3</td>
<td>3</td>
<td>1089655±727860</td>
<td>2785285±891212</td>
<td>0.117490648</td>
<td>2.56</td>
</tr>
<tr>
<td>RNA binding motif protein, X-linked-like 1 OS=Mus musculus GN=Rbmxl PE=1 SV=1</td>
<td>Q9X1W8</td>
<td>FUBP1_MOUSE</td>
<td>68.5</td>
<td>24.3%</td>
<td>12</td>
<td>28</td>
<td>25273600±1039068</td>
<td>40113500±4730225</td>
<td>0.117490648</td>
<td>1.59</td>
</tr>
<tr>
<td>Proximal element-binding protein 1 OS=Mus musculus GN=Fup1 PE=1 SV=1</td>
<td>P67855</td>
<td>G3BP1_MOUSE</td>
<td>51.8</td>
<td>2.37%</td>
<td>1</td>
<td>2</td>
<td>7735275±3278167</td>
<td>16154600±2466521</td>
<td>0.117490648</td>
<td>2.09</td>
</tr>
<tr>
<td>Ras GTPase-activating protein-binding protein 1 OS=Mus musculus GN=Gbp1 PE=1 SV=1</td>
<td>Q62376</td>
<td>SNRP70_MOUSE</td>
<td>52</td>
<td>6.25%</td>
<td>2</td>
<td>2</td>
<td>1739742±1027569</td>
<td>504400±1015501</td>
<td>0.117490648</td>
<td>2.90</td>
</tr>
<tr>
<td>U1 small nuclear ribonucleoprotein 70 kDa OS=Mus musculus GN=Snmp70 PE=1 SV=2</td>
<td>Q8EP04</td>
<td>UPF1_MOUSE</td>
<td>123.9</td>
<td>1.96%</td>
<td>2</td>
<td>2</td>
<td>245387±591048</td>
<td>2093775±179931</td>
<td>0.117490648</td>
<td>8.53</td>
</tr>
<tr>
<td>PROTEIN NAME</td>
<td>ACCESSION NUMBER</td>
<td>UNIPROT ID</td>
<td>MW (kDa)</td>
<td>MAX-sequence coverage</td>
<td>UNIQUE PEPTIDES</td>
<td>TOTAL PSMs</td>
<td>CONTROL (Mean±SD)</td>
<td>MUTANT (Mean±SD)</td>
<td>P VALUE</td>
<td>FOLD CHANGE</td>
</tr>
<tr>
<td>--------------</td>
<td>------------------</td>
<td>------------</td>
<td>----------</td>
<td>-----------------------</td>
<td>----------------</td>
<td>------------</td>
<td>------------------</td>
<td>------------------</td>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>Heterogeneous nuclear ribonucleoprotein D0 OS=Mus musculus GN=Hnrnpd PE=1 SV=2</td>
<td>Q60668</td>
<td>HNRNPD_MOUSE</td>
<td>38.3</td>
<td>23.1%</td>
<td>7</td>
<td>15</td>
<td>324040000±11798612</td>
<td>409772505±3472172</td>
<td>0.11794068</td>
<td>1.26</td>
</tr>
<tr>
<td>Fibronectin OS=Mus musculus GN=Fn1 PE=1 SV=4</td>
<td>P11276</td>
<td>FN1_MOUSE</td>
<td>272.4</td>
<td>4.28%</td>
<td>8</td>
<td>8</td>
<td>14372155±1127426</td>
<td>2609565±538126</td>
<td>0.11794068</td>
<td>2.02</td>
</tr>
<tr>
<td>ATP-binding cassette sub-family E member 1 OS=Mus musculus GN=Abce1 PE=1 SV=1</td>
<td>P61222</td>
<td>ABCE1_MOUSE</td>
<td>67.3</td>
<td>3.51%</td>
<td>2</td>
<td>2</td>
<td>26085500±1716819</td>
<td>6178825±540335</td>
<td>0.11794068</td>
<td>2.37</td>
</tr>
<tr>
<td>Lypus La protein homolog OS=Mus musculus GN=Latp PE=1 SV=1</td>
<td>P23067</td>
<td>SS8_MOUSE</td>
<td>47.7</td>
<td>6.99%</td>
<td>3</td>
<td>3</td>
<td>14736750±2837240</td>
<td>2622500±4450531</td>
<td>0.11794068</td>
<td>1.78</td>
</tr>
<tr>
<td>Poly(ADP-ribose) polymerase 1 OS=Mus musculus GN=Parp1 PE=1 SV=3</td>
<td>P11103</td>
<td>PARP1_MOUSE</td>
<td>113</td>
<td>10.4%</td>
<td>8</td>
<td>8</td>
<td>795347.5±1217644</td>
<td>2808225±632890</td>
<td>0.11794068</td>
<td>3.51</td>
</tr>
<tr>
<td>Zinc finger protein 2 OS=Mus musculus GN=Znf2 PE=2 SV=1</td>
<td>Q8BIQ3</td>
<td>ZNF2_MOUSE</td>
<td>96.6</td>
<td>60%</td>
<td>3</td>
<td>4</td>
<td>3595650±107060</td>
<td>5628205±156055</td>
<td>0.11794068</td>
<td>1.57</td>
</tr>
<tr>
<td>Cytoplasmic dynein 1 light intermediate chain 2 OS=Mus musculus GN=Dync1l2 PE=1 SV=2</td>
<td>Q6PDL0</td>
<td>DYNCL2_MOUSE</td>
<td>54.2</td>
<td>3.46%</td>
<td>1</td>
<td>1</td>
<td>44328605±524229</td>
<td>11160275±3413432</td>
<td>0.11794068</td>
<td>2.52</td>
</tr>
<tr>
<td>Mossin OS=Mus musculus GN=Msn PE=1 SV=3</td>
<td>P26041</td>
<td>MSN_MOUSE</td>
<td>67.7</td>
<td>25.1%</td>
<td>15</td>
<td>27</td>
<td>71282752±2140250</td>
<td>13050725±1003194</td>
<td>0.11794068</td>
<td>1.83</td>
</tr>
<tr>
<td>60S ribosomal protein L15 OS=Mus musculus GN=Rpl15 PE=2 SV=4</td>
<td>Q8CZM2</td>
<td>RPL15_MOUSE</td>
<td>24.1</td>
<td>11.3%</td>
<td>2</td>
<td>3</td>
<td>14100000±26672696</td>
<td>22108750±15479489</td>
<td>0.1219124</td>
<td>1.57</td>
</tr>
<tr>
<td>DNA topoisomerase 1 OS=Mus musculus GN=Top1 PE=1 SV=2</td>
<td>Q04750</td>
<td>TOP1_MOUSE</td>
<td>90.8</td>
<td>10.6%</td>
<td>7</td>
<td>7</td>
<td>65480505±4261755</td>
<td>15203305±3234194</td>
<td>0.1251774</td>
<td>2.32</td>
</tr>
<tr>
<td>Poly(U)-binding-splicing factor Puf60 OS=Mus musculus GN=Puf60 PE=1 SV=2</td>
<td>Q3UEB3</td>
<td>PUF60_MOUSE</td>
<td>60.2</td>
<td>1.42%</td>
<td>1</td>
<td>1</td>
<td>39079000±4136619</td>
<td>7048650±3588780</td>
<td>0.13714767</td>
<td>1.80</td>
</tr>
<tr>
<td>Small nuclear ribonucleoprotein Sm D2 OS=Mus musculus GN=Snrp2 PE=1 SV=1</td>
<td>P62317</td>
<td>SNRPD2_MOUSE</td>
<td>13.5</td>
<td>8.47%</td>
<td>1</td>
<td>1</td>
<td>16837755±4309041</td>
<td>2776300±394620</td>
<td>0.13714767</td>
<td>1.68</td>
</tr>
<tr>
<td>Pre-mRNA-splicing factor ATP-dependent RNA helicase DHX15 OS=Mus musculus GN=Dhx15 PE=1 SV=2</td>
<td>Q35286</td>
<td>DHX15_MOUSE</td>
<td>90.9</td>
<td>9.81%</td>
<td>7</td>
<td>7</td>
<td>76808550±3219170</td>
<td>14339750±2326582</td>
<td>0.14414671</td>
<td>1.87</td>
</tr>
<tr>
<td>Extended synaptoplasm-1 OS=Mus musculus GN=Eys1f1 PE=1 SV=2</td>
<td>Q3UTR1</td>
<td>ESY1_MOUSE</td>
<td>121.5</td>
<td>21%</td>
<td>21</td>
<td>88</td>
<td>36279505±5829331</td>
<td>10550250±3228843±2917741</td>
<td>0.14414671</td>
<td>2.91</td>
</tr>
<tr>
<td>High mobility group protein B1 OS=Mus musculus GN=Hmgb1 PE=1 SV=2</td>
<td>P63158</td>
<td>HMGBI_MOUSE</td>
<td>24.9</td>
<td>36.3%</td>
<td>9</td>
<td>16</td>
<td>3350420±2848045</td>
<td>5613600±3564825</td>
<td>0.14414671</td>
<td>1.68</td>
</tr>
<tr>
<td>Selencoycte-specific elongation factor OS=Mus musculus GN=Eefsec PE=1 SV=2</td>
<td>Q8JHN4</td>
<td>EIFSEC_MOUSE</td>
<td>63.5</td>
<td>13.9%</td>
<td>8</td>
<td>8</td>
<td>1089645±717847</td>
<td>2676850±765831</td>
<td>0.14414671</td>
<td>2.46</td>
</tr>
<tr>
<td>Cell growth-regulating nuclear protein OS=Mus musculus GN=Lyar PE=1 SV=2</td>
<td>Q6B288</td>
<td>LYAR_MOUSE</td>
<td>43.7</td>
<td>11.3%</td>
<td>3</td>
<td>3</td>
<td>3061125±1490667</td>
<td>5972950±1137195</td>
<td>0.14414671</td>
<td>1.94</td>
</tr>
<tr>
<td>iRNA (cytosine(3A)-O5)-methyltransferase OS=Mus musculus GN=Mnu2 PE=1 SV=2</td>
<td>Q1HF20</td>
<td>NSUN2_MOUSE</td>
<td>85.4</td>
<td>16.8%</td>
<td>11</td>
<td>13</td>
<td>1381100±2515198</td>
<td>5454075±1261973</td>
<td>0.14414671</td>
<td>3.95</td>
</tr>
<tr>
<td>Cluster of Transcription elongation factor B poly peptide 3 OS=Mus musculus GN=Toalb3 PE=1 SV=3 (Transcription)</td>
<td>Q8CBE7</td>
<td>TCEBI_MOUSE</td>
<td>87.1</td>
<td>4%</td>
<td>3</td>
<td>7</td>
<td>18005550±3764890</td>
<td>6123150±1488485</td>
<td>0.15079522</td>
<td>3.49</td>
</tr>
<tr>
<td>Sarcoplasmin/elodaplastic reticulum calcium ATPase 2 OS=Mus musculus GN=Atp2a2 PE=1 SV=2</td>
<td>OS5143</td>
<td>ATP2A2_MOUSE</td>
<td>114.8</td>
<td>4.41%</td>
<td>4</td>
<td>4</td>
<td>10073755±420427</td>
<td>24422255±618034</td>
<td>0.15911865</td>
<td>2.42</td>
</tr>
<tr>
<td>Pep3-polyl-cis-trans isomerase FB35 OS=Mus musculus GN=Rplp3 PE=1 SV=2</td>
<td>Q62446</td>
<td>FKBP3_MOUSE</td>
<td>25.1</td>
<td>9.1%</td>
<td>1</td>
<td>1</td>
<td>12152500±2586753</td>
<td>14972750±2745745</td>
<td>0.15653494</td>
<td>1.23</td>
</tr>
<tr>
<td>X-ray repair cross-complementing protein 6 OS=Mus musculus GN=Xrc6 PE=1 SV=5</td>
<td>P23475</td>
<td>XRC6_MOUSE</td>
<td>69.4</td>
<td>14.8%</td>
<td>7</td>
<td>8</td>
<td>24038052±832444</td>
<td>4157005±1448647</td>
<td>0.15653494</td>
<td>1.71</td>
</tr>
<tr>
<td>Ribosomal L1 domain-containing protein 1 OS=Mus musculus GN=Ripl1d1 PE=1 SV=1</td>
<td>Q8BY70</td>
<td>RSL1D1_MOUSE</td>
<td>50.4</td>
<td>6.42%</td>
<td>2</td>
<td>3</td>
<td>77536.5±118701</td>
<td>1492500±587038</td>
<td>0.15653494</td>
<td>1.92</td>
</tr>
<tr>
<td>Eukaryotic translation initiation factor 2A OS=Mus musculus GN=Eif2a PE=1 SV=2</td>
<td>Q88BJ6</td>
<td>EIF2A_MOUSE</td>
<td>64.4</td>
<td>11.7%</td>
<td>5</td>
<td>6</td>
<td>24183000±8104034</td>
<td>38507005±7719760</td>
<td>0.15653494</td>
<td>1.59</td>
</tr>
<tr>
<td>Nuclear RNA export factor 1 OS=Mus musculus GN=Neft1 PE=1 SV=3</td>
<td>Q9JX7</td>
<td>NXF1_MOUSE</td>
<td>70.3</td>
<td>4.53%</td>
<td>2</td>
<td>2</td>
<td>42631250±923809</td>
<td>7240125±1322174</td>
<td>0.15653494</td>
<td>1.70</td>
</tr>
<tr>
<td>PROTEIN NAME</td>
<td>ACCESSION NUMBER</td>
<td>UNIPROT ID</td>
<td>MW (kDa)</td>
<td>MAX sequence coverage</td>
<td>UNIQUE PEPTIDES</td>
<td>TOTAL PSMs</td>
<td>CONTROL (Mean±SD)</td>
<td>MUTANT (Mean±SD)</td>
<td>P VALUE</td>
<td>FOLD CHANGE</td>
</tr>
<tr>
<td>--------------</td>
<td>------------------</td>
<td>------------</td>
<td>----------</td>
<td>----------------------</td>
<td>----------------</td>
<td>------------</td>
<td>------------------</td>
<td>------------------</td>
<td>---------</td>
<td>------------</td>
</tr>
<tr>
<td>Glycine/peptide N-tetradecanoyltransferase 1 OS=Mus musculus GN=Nmt1 PE=1 SV=1</td>
<td>O70310</td>
<td>NMT1_MOUSE</td>
<td>56.9</td>
<td>9.48%</td>
<td>5</td>
<td>5</td>
<td>1120882x593716</td>
<td>1984257.5x550122</td>
<td>0.159636494</td>
<td>1.77</td>
</tr>
<tr>
<td>CDK12A-interacting protein OS=Mus musculus GN=Cdk12aip PE=1 SV=1</td>
<td>Q6B7I2</td>
<td>CDK12AIP_MOUSE</td>
<td>59.7</td>
<td>7.64%</td>
<td>3</td>
<td>4</td>
<td>484522.5x645656</td>
<td>1166620x325712</td>
<td>0.159636494</td>
<td>2.41</td>
</tr>
<tr>
<td>iRNA-silencing endonuclease subunit 34 OS=Mus musculus GN=Tsna34 PE=2 SV=2</td>
<td>Q6BMZ5</td>
<td>TSEN34_MOUSE</td>
<td>34.2</td>
<td>12.3%</td>
<td>3</td>
<td>3</td>
<td>7459100x2698764</td>
<td>1059770x3094819</td>
<td>0.16635515</td>
<td>1.42</td>
</tr>
<tr>
<td>Polyribonucleotide 5'-hydroxylkinase Clp1 OS=Mus musculus GN=Clp1 PE=1 SV=1</td>
<td>Q9L9B</td>
<td>CLP1_MOUSE</td>
<td>47.7</td>
<td>6.35%</td>
<td>3</td>
<td>3</td>
<td>1192075x1234803</td>
<td>3176175x505991</td>
<td>0.16694312</td>
<td>2.66</td>
</tr>
<tr>
<td>Glutamate dehydrogenase 1, mitochondrial OS=Mus musculus GN=Gdh1 PE=1 SV=1</td>
<td>P26443</td>
<td>GLUD1_MOUSE</td>
<td>61.3</td>
<td>14.5%</td>
<td>6</td>
<td>8</td>
<td>3927550x1206429</td>
<td>6165450x1870370</td>
<td>0.17066624</td>
<td>1.57</td>
</tr>
<tr>
<td>Plasminogen activator inhibitor 1 RNA-binding protein OS=Mus musculus GN=Serpt1 PE=1 SV=2</td>
<td>Q6C9Y8</td>
<td>SERPT1_MOUSE</td>
<td>44.7</td>
<td>1.97%</td>
<td>1</td>
<td>1</td>
<td>1161127.5x1337669</td>
<td>2331306x649525</td>
<td>0.17066624</td>
<td>2.01</td>
</tr>
<tr>
<td>Exosome complex exonuclease RRP44 OS=Mus musculus GN=Exo1 PE=1 SV=4</td>
<td>Q9CS93</td>
<td>DIS3_MOUSE</td>
<td>108.8</td>
<td>3.44%</td>
<td>3</td>
<td>3</td>
<td>2704239x1595188</td>
<td>4132125x2150342</td>
<td>0.174433825</td>
<td>1.53</td>
</tr>
<tr>
<td>U2 small nuclear ribonucleoprotein A' OS=Mus musculus GN=Snrega1 PE=1 SV=2</td>
<td>P57784</td>
<td>SNERPA1_MOUSE</td>
<td>28.3</td>
<td>5.49%</td>
<td>1</td>
<td>1</td>
<td>420462.5x639106</td>
<td>1284472.5x361454</td>
<td>0.177750263</td>
<td>3.05</td>
</tr>
<tr>
<td>Nuclease-sensitive element-binding protein 1 OS=Mus musculus GN=Exo1 PE=1 SV=3</td>
<td>P62960</td>
<td>YBX1_MOUSE</td>
<td>35.7</td>
<td>59%</td>
<td>12</td>
<td>21</td>
<td>24221500x17312887</td>
<td>4523470x4829264</td>
<td>0.177750263</td>
<td>1.87</td>
</tr>
<tr>
<td>CAD protein OS=Mus musculus GN=Cad PE=1 SV=1</td>
<td>B2RQ6C</td>
<td>CAD_MOUSE</td>
<td>243.1</td>
<td>2.74%</td>
<td>5</td>
<td>5</td>
<td>204882.5x3095665</td>
<td>4526770x2670035</td>
<td>0.177750263</td>
<td>2.21</td>
</tr>
<tr>
<td>Cold shock domain-containing protein E1 OS=Mus musculus GN=Csf10 PE=1 SV=1</td>
<td>O91W50</td>
<td>CSF10_MOUSE</td>
<td>88.7</td>
<td>11.4%</td>
<td>8</td>
<td>8</td>
<td>2004912.5x1622819</td>
<td>4226925x394156</td>
<td>0.177750263</td>
<td>2.11</td>
</tr>
<tr>
<td>Gem-associated protein 5 OS=Mus musculus GN=Gemin5 PE=1 SV=2</td>
<td>Q8B17X</td>
<td>GEM5_MOUSE</td>
<td>166.5</td>
<td>1.26%</td>
<td>2</td>
<td>2</td>
<td>744720x1253023</td>
<td>1988962.5x525790</td>
<td>0.18091264</td>
<td>2.67</td>
</tr>
<tr>
<td>Citrate synthase, mitochondrial OS=Mus musculus GN=Cit PE=1 SV=1</td>
<td>Q6C2J6</td>
<td>CS_MOUSE</td>
<td>51.7</td>
<td>8.19%</td>
<td>3</td>
<td>3</td>
<td>3162525x1078201</td>
<td>3744250x697751</td>
<td>0.18091264</td>
<td>1.19</td>
</tr>
<tr>
<td>Serine-threonine kinase receptor-associated protein OS=Mus musculus GN=Strap PE=1 SV=2</td>
<td>Q8Z122</td>
<td>STRAP_MOUSE</td>
<td>38.4</td>
<td>14%</td>
<td>4</td>
<td>4</td>
<td>2273546x704776</td>
<td>938170.5x153694</td>
<td>0.18091264</td>
<td>4.12</td>
</tr>
<tr>
<td>Uncharacterized protein C6orf52 homolog OS=Mus musculus PE=1 SV=2</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>702542.5x261262</td>
<td>1444850x61948</td>
<td>0.1810901</td>
<td>1.82</td>
<td></td>
</tr>
<tr>
<td>Synaptic functional regulator FM1 OS=Mus musculus GN=Fmr1 PE=1 SV=1</td>
<td>P38922</td>
<td>FM1_MOUSE</td>
<td>68.9</td>
<td>7.98%</td>
<td>4</td>
<td>4</td>
<td>561637.5x1508914</td>
<td>2765697.5x579041</td>
<td>0.195375294</td>
<td>4.92</td>
</tr>
<tr>
<td>Constitutive coactivator of PPAR-gamma-like protein OS=Mus musculus GN=Fam120a PE=1 SV=2</td>
<td>Q6A0A9</td>
<td>FAM120A_MOUSE</td>
<td>121.6</td>
<td>0.809%</td>
<td>1</td>
<td>1</td>
<td>2043875x1749812</td>
<td>4329300x787791</td>
<td>0.208285124</td>
<td>2.12</td>
</tr>
<tr>
<td>Asparagine--RNA ligase, cytoplasmic OS=Mus musculus GN=Nars PE=1 SV=2</td>
<td>Q8BP47</td>
<td>NARS_MOUSE</td>
<td>64.2</td>
<td>10.9%</td>
<td>5</td>
<td>5</td>
<td>1350312.5x1294392</td>
<td>2670755x1729295</td>
<td>0.208285124</td>
<td>1.98</td>
</tr>
<tr>
<td>Double-stranded RNA-binding protein Staufen homolog 1 OS=Mus musculus GN=Stau1 PE=1 SV=1</td>
<td>Q5Z108</td>
<td>STAU1_MOUSE</td>
<td>53.9</td>
<td>36%</td>
<td>16</td>
<td>72</td>
<td>9758700x3222632</td>
<td>1326400x4468473</td>
<td>0.208285124</td>
<td>2.12</td>
</tr>
<tr>
<td>Bifunctional glutamate/proline--RNA ligase OS=Mus musculus GN=Eprn PE=1 SV=4</td>
<td>Q8C0C7</td>
<td>EPRN_MOUSE</td>
<td>170</td>
<td>11.1%</td>
<td>12</td>
<td>12</td>
<td>6020950x3032996</td>
<td>8958675x1339098</td>
<td>0.208285124</td>
<td>1.49</td>
</tr>
<tr>
<td>T-complex protein 1 subunit gamma OS=Mus musculus GN=Ctx PE=1 SV=1</td>
<td>P80316</td>
<td>CCTX3_MOUSE</td>
<td>60.6</td>
<td>24%</td>
<td>14</td>
<td>15</td>
<td>1219300x11068865</td>
<td>18662150x1340493</td>
<td>0.213042051</td>
<td>1.53</td>
</tr>
<tr>
<td>Cluster of Myxins-9 OS=Mus musculus GN=Myh9 PE=1 SV=4 (MYH9_MOUSE)</td>
<td>Q6VD05</td>
<td>MYH9_MOUSE</td>
<td>226.3</td>
<td>2.5%</td>
<td>4</td>
<td>4</td>
<td>1923217.5x432517</td>
<td>3443250x1355486</td>
<td>0.213042051</td>
<td>1.79</td>
</tr>
<tr>
<td>Prohibitin OS=Mus musculus GN=Phb PE=1 SV=1</td>
<td>P67776</td>
<td>PHB_MOUSE</td>
<td>29.8</td>
<td>15.1%</td>
<td>4</td>
<td>4</td>
<td>1489715x1016543</td>
<td>2013652.5x789384</td>
<td>0.21496364</td>
<td>1.35</td>
</tr>
<tr>
<td>40S ribosomal protein S5 OS=Mus musculus GN=Rps5 PE=1 SV=3</td>
<td>P67461</td>
<td>RPS5_MOUSE</td>
<td>22.9</td>
<td>28.9%</td>
<td>7</td>
<td>17</td>
<td>128127500x14194326</td>
<td>151727500x16230964</td>
<td>0.21496364</td>
<td>1.20</td>
</tr>
<tr>
<td>E3 ubiquitin ligase TRIM25 OS=Mus musculus GN=Trim25 PE=1 SV=2</td>
<td>Q61510</td>
<td>TRIM25_MOUSE</td>
<td>71.7</td>
<td>1.74%</td>
<td>1</td>
<td>1</td>
<td>711992.5x1476393</td>
<td>2553432.5x929563</td>
<td>0.21496364</td>
<td>3.59</td>
</tr>
<tr>
<td>PROTEIN NAME</td>
<td>ACCESSION NUMBER</td>
<td>UNIPROT ID</td>
<td>MW (kDa)</td>
<td>MAX-SEQUENCE COVERAGE</td>
<td>UNIQUE PEPTIDES</td>
<td>TOTAL PDMS</td>
<td>CONTROL (Mean±SD)</td>
<td>MUTANT (Mean±SD)</td>
<td>P VALUE</td>
<td>FOLD CHANGE</td>
</tr>
<tr>
<td>-------------</td>
<td>------------------</td>
<td>------------------</td>
<td>---------</td>
<td>------------------------</td>
<td>----------------</td>
<td>------------</td>
<td>-------------------</td>
<td>-----------------</td>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>Splicing factor, proline- and glutamine-rich OSH3 mouse muscle GN=Sppl PE=1 SV=1</td>
<td>Q8VU6</td>
<td>SFPQ_MOUSE</td>
<td>74.4</td>
<td>14%</td>
<td>10</td>
<td>13</td>
<td>15079600±1166019</td>
<td>2150500±s299643</td>
<td>0.21849634</td>
<td>1.83</td>
</tr>
<tr>
<td>Phosphatidylinositol-4-phosphate 5-kinase type-1 alpha OSH3 mouse muscle GN=Pip5K1a PE=1 SV=2</td>
<td>P70182</td>
<td>PIPSK1A_MOUSE</td>
<td>60.4</td>
<td>5.31%</td>
<td>3</td>
<td>3</td>
<td>487387.5±1100100</td>
<td>1585140±s182319</td>
<td>0.21849634</td>
<td>3.25</td>
</tr>
<tr>
<td>Putative ATP-dependent RNA helicase P101 OSH3 mouse muscle GN=D1Pass PE=1 SV=1</td>
<td>NA</td>
<td>D1PAS1_MOUSE</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>9920460±s2671356</td>
<td>15889750±s5177824</td>
<td>0.21849634</td>
<td>1.60</td>
</tr>
<tr>
<td>SRA stem-loop-interacting RNA-binding protein, mitochondrial OSH3 mouse muscle GN=Srip PE=1 SV=2</td>
<td>Q8Q7T7</td>
<td>SLRP_MOUSE</td>
<td>12.6</td>
<td>52.7%</td>
<td>8</td>
<td>47</td>
<td>2135775±s111873</td>
<td>3447575±s239464</td>
<td>0.21849634</td>
<td>1.61</td>
</tr>
<tr>
<td>Cluster of ATP-dependent RNA helicase DDXY OSH3 mouse muscle GN=DDXY PE=1 SV=2 (ATP-dependent)</td>
<td>Q62095</td>
<td>DDXY_MOUSE</td>
<td>73.4</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>39693750±s9857831</td>
<td>49679750±s10833320</td>
<td>0.21849634</td>
<td>1.25</td>
</tr>
<tr>
<td>DAZ-associated protein 1 OSH3 mouse muscle GN=Dazap1 PE=1 SV=2</td>
<td>Q9J6I</td>
<td>DAZAP1_MOUSE</td>
<td>43.2</td>
<td>8.87%</td>
<td>3</td>
<td>5</td>
<td>7904060±s2368501</td>
<td>3463255±s91297</td>
<td>0.220382266</td>
<td>4.38</td>
</tr>
<tr>
<td>Microtubule-associated protein 4 OSH3 mouse muscle GN=Map4 PE=1 SV=3</td>
<td>P7546</td>
<td>MAP4_MOUSE</td>
<td>117.4</td>
<td>2.49%</td>
<td>2</td>
<td>2</td>
<td>1777225±s675728</td>
<td>2707050±s513133</td>
<td>0.225760101</td>
<td>1.52</td>
</tr>
<tr>
<td>Fibrin-C OSH3 mouse muscle GN=Fnc PE=1 SV=3</td>
<td>Q8VK6X</td>
<td>FLNC_MOUSE</td>
<td>290.9</td>
<td>1.65%</td>
<td>3</td>
<td>3</td>
<td>2592425±s1813665</td>
<td>4140725±s187383</td>
<td>0.225760101</td>
<td>1.60</td>
</tr>
<tr>
<td>Calumenin OSH3 mouse muscle GN=Calu PE=1 SV=1</td>
<td>O35887</td>
<td>CALU_MOUSE</td>
<td>37</td>
<td>4.76%</td>
<td>1</td>
<td>1</td>
<td>1960400±s870734</td>
<td>3058575±s2015758</td>
<td>0.225760101</td>
<td>1.56</td>
</tr>
<tr>
<td>Prolyl-4-hydroxylyase subunit alpha-1 OSH3 mouse muscle GN=P4ha1 PE=1 SV=2</td>
<td>Q60715</td>
<td>P4HA1_MOUSE</td>
<td>60.9</td>
<td>11.8%</td>
<td>5</td>
<td>5</td>
<td>2313500±s109568</td>
<td>3829400±s1087413</td>
<td>0.228450181</td>
<td>1.66</td>
</tr>
<tr>
<td>Cap-specific mRNA (nucleoside-2'-O')-methyltransferase OSH3 mouse muscle GN=Cre1 PE=1 SV=1</td>
<td>Q5BDC3</td>
<td>CMTR1_MOUSE</td>
<td>95.6</td>
<td>14.6%</td>
<td>12</td>
<td>15</td>
<td>11128225±s4946442</td>
<td>16583500±s2426542</td>
<td>0.228450181</td>
<td>1.49</td>
</tr>
<tr>
<td>Cluster of Sodium/potassium-transporting ATPase subunit alpha-2 OSH3 mouse muscle GN=Ap4a2 PE=1 SV=1 (Sodium/potassium-transporting)</td>
<td>Q6PEP5</td>
<td>ATP1A2_MOUSE</td>
<td>112.2</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>2324925±s871731</td>
<td>3300125±s1004148</td>
<td>0.24274262</td>
<td>1.42</td>
</tr>
<tr>
<td>Slaphilococcal nuclease domain-containing protein 1 OSH3 mouse muscle GN=Sn21 PE=1 SV=1</td>
<td>Q78Y7</td>
<td>SND1_MOUSE</td>
<td>102</td>
<td>31.5%</td>
<td>26</td>
<td>36</td>
<td>9929300±s2833624</td>
<td>12503875±s4476015</td>
<td>0.247309751</td>
<td>1.26</td>
</tr>
<tr>
<td>Protein-glutamine gamma-glutamyltransferase II OSH3 mouse muscle GN=Tgcm3 PE=1 SV=2</td>
<td>Q08189</td>
<td>TGM3_MOUSE</td>
<td>77.3</td>
<td>1.88%</td>
<td>1</td>
<td>1</td>
<td>8528658±s21545</td>
<td>1951900±s620675</td>
<td>0.25693254</td>
<td>2.29</td>
</tr>
<tr>
<td>Polyproline tract-binding protein 3 OSH3 mouse muscle GN=Pp3b1 PE=1 SV=1</td>
<td>Q8BB7</td>
<td>TPB3_MOUSE</td>
<td>56.7</td>
<td>1.53%</td>
<td>1</td>
<td>1</td>
<td>6286500±s96275009</td>
<td>13652700±s2204263</td>
<td>0.25693254</td>
<td>2.17</td>
</tr>
<tr>
<td>Drabin OSH3 mouse muscle GN=Drn1 PE=1 SV=1</td>
<td>Q9QX56</td>
<td>DBN1_MOUSE</td>
<td>77.2</td>
<td>15%</td>
<td>8</td>
<td>36</td>
<td>1036400±s179034</td>
<td>1655930±s1056643</td>
<td>0.26333671</td>
<td>1.60</td>
</tr>
<tr>
<td>Protein PRRC2C OSH3 mouse muscle GN=Pr4o2c PE=1 SV=3</td>
<td>Q3T1H4</td>
<td>PRRC2C_MOUSE</td>
<td>310.7</td>
<td>0.492%</td>
<td>1</td>
<td>1</td>
<td>3027600±s1737539</td>
<td>4180900±s15067524</td>
<td>0.27983228</td>
<td>1.38</td>
</tr>
<tr>
<td>U1 snRNA-associated Sm-like protein LSml8 OSH3 mouse muscle GN=Smml8 PE=1 SV=3</td>
<td>Q6QM9M</td>
<td>LSM8_MOUSE</td>
<td>10.4</td>
<td>10.4%</td>
<td>1</td>
<td>1</td>
<td>5368528±s369424</td>
<td>8027550±s17236785</td>
<td>0.28528735</td>
<td>1.50</td>
</tr>
<tr>
<td>Interferon-activatable protein 204 OSH3 mouse muscle GN=IF204 PE=1 SV=1</td>
<td>Q9V192</td>
<td>ACOY_MOUSE</td>
<td>119.7</td>
<td>9.44%</td>
<td>10</td>
<td>10</td>
<td>48812±s378904</td>
<td>91890±s381354</td>
<td>0.30437115</td>
<td>1.88</td>
</tr>
<tr>
<td>ATP-citrate synthase OSH3 mouse muscle GN=Acty PE=1 SV=1</td>
<td>Q90405</td>
<td>NCL_MOUSE</td>
<td>76.7</td>
<td>24.9%</td>
<td>20</td>
<td>47</td>
<td>25658750±s892516</td>
<td>2977250±s5878866</td>
<td>0.30437115</td>
<td>1.16</td>
</tr>
<tr>
<td>Nucleolin OSH3 mouse muscle GN=Nld PE=1 SV=2</td>
<td>Q01140</td>
<td>COL1A2_MOUSE</td>
<td>129.5</td>
<td>30.8%</td>
<td>24</td>
<td>33</td>
<td>36566250±s19835242</td>
<td>4551600±s861173</td>
<td>0.33951034</td>
<td>1.25</td>
</tr>
<tr>
<td>Collagen alpha-2(I) chain OSH3 mouse muscle GN=Cotla2 PE=1 SV=2</td>
<td>Q61753</td>
<td>PNDGH_MOUSE</td>
<td>56.5</td>
<td>9.01%</td>
<td>4</td>
<td>4</td>
<td>2377930±s2461929</td>
<td>3971475±s2319231</td>
<td>0.343710701</td>
<td>1.67</td>
</tr>
<tr>
<td>D-3-phosphoglycerate dehydrogenase OSH3 mouse muscle GN=Pgdh PE=1 SV=3</td>
<td>Q8Q310</td>
<td>MATR3_MOUSE</td>
<td>94.6</td>
<td>41%</td>
<td>32</td>
<td>175</td>
<td>9027775±s3324674</td>
<td>10786575±s147817</td>
<td>0.343710701</td>
<td>1.19</td>
</tr>
<tr>
<td>Importin subunit beta-1 OSH3 mouse muscle GN=Kpnb1 PE=1 SV=2</td>
<td>P70168</td>
<td>KPNB1_MOUSE</td>
<td>97.1</td>
<td>2.4%</td>
<td>2</td>
<td>3</td>
<td>1551983±s147404</td>
<td>2203850±s1378683</td>
<td>0.338417651</td>
<td>1.42</td>
</tr>
<tr>
<td>PROTEIN NAME</td>
<td>ACCESION NUMBER</td>
<td>UNIPROT ID</td>
<td>MW (kDa)</td>
<td>MAX- sequence coverage</td>
<td>UNIQUE PEPTIDES</td>
<td>TOTAL PSMs</td>
<td>CONTROL (Mean±SD)</td>
<td>MUTANT (Mean±SD)</td>
<td>P VALUE</td>
<td>FOLD CHANGE</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------------</td>
<td>------------</td>
<td>----------</td>
<td>------------------------</td>
<td>----------------</td>
<td>------------</td>
<td>-----------------</td>
<td>-----------------</td>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>Tropomyosin alpha-3 chain OS=Mus musculus GN=Tpm3 PE=1 SV=3</td>
<td>P21107</td>
<td>TPM3_MOUSE</td>
<td>32.9</td>
<td>3.51%</td>
<td>1</td>
<td>1</td>
<td>3975250±1370069</td>
<td>5148675±1381974</td>
<td>0.33847651</td>
<td>1.30</td>
</tr>
<tr>
<td>Golgi apparatus protein 1 OS=Mus musculus GN=Glg1 PE=1 SV=1</td>
<td>Q61543</td>
<td>GLG1_MOUSE</td>
<td>133.6</td>
<td>3.66%</td>
<td>4</td>
<td>4</td>
<td>1230442.5±1258376</td>
<td>2013830±320118</td>
<td>0.33847651</td>
<td>1.64</td>
</tr>
<tr>
<td>Signal recognition particle 54 kDa protein OS=Mus musculus GN=Spr54 PE=1 SV=2</td>
<td>P14576</td>
<td>SRP54_MOUSE</td>
<td>55.7</td>
<td>18.8%</td>
<td>8</td>
<td>8</td>
<td>4255675±1341165</td>
<td>5580800±1064537</td>
<td>0.33847651</td>
<td>1.31</td>
</tr>
<tr>
<td>Spermatid perinuclear RNA-binding protein OS=Mus musculus GN=Spnbp PE=1 SV=1</td>
<td>Q81W1M</td>
<td>STRBP_MOUSE</td>
<td>73.7</td>
<td>3.42%</td>
<td>3</td>
<td>3</td>
<td>9562125±3739427</td>
<td>12641275±2395981</td>
<td>0.33847651</td>
<td>1.32</td>
</tr>
<tr>
<td>Switch-associated protein 70 OS=Mus musculus GN=Swap70 PE=1 SV=1</td>
<td>Q8628Q</td>
<td>SWAP70_MOUSE</td>
<td>69</td>
<td>6.67%</td>
<td>4</td>
<td>4</td>
<td>1353410±46661</td>
<td>1936375±1121862</td>
<td>0.34199549</td>
<td>1.43</td>
</tr>
<tr>
<td>Probable ribonuclease ZC3H12C OS=Mus musculus GN=Zc3H12c PE=1 SV=2</td>
<td>Q5DTV4</td>
<td>ZC12C_MOUSE</td>
<td>90.4</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>3796825±744511</td>
<td>4529225±4965523</td>
<td>0.34199549</td>
<td>1.19</td>
</tr>
<tr>
<td>Lysine--RNA Igase OS=Mus musculus GN=Kars PE=1 SV=1</td>
<td>Q89MN1</td>
<td>KARS_MOUSE</td>
<td>67.8</td>
<td>6.05%</td>
<td>3</td>
<td>3</td>
<td>4315350±2490045</td>
<td>5553350±981796</td>
<td>0.34199549</td>
<td>1.29</td>
</tr>
<tr>
<td>Cleavage and polyadenylation specificity factor subunit 7 OS=Mus musculus GN=Cps7 PE=1 SV=2</td>
<td>Q88TV2</td>
<td>CPS7_MOUSE</td>
<td>52</td>
<td>2.76%</td>
<td>1</td>
<td>1</td>
<td>4500825±1866643</td>
<td>6017375±3033532</td>
<td>0.34383751</td>
<td>1.34</td>
</tr>
<tr>
<td>Transitional endoplasmic reticulum ATPase OS=Mus musculus GN=Vcp PE=1 SV=4</td>
<td>Q01853</td>
<td>VCP_MOUSE</td>
<td>89.3</td>
<td>30%</td>
<td>20</td>
<td>108</td>
<td>2684565±1629636</td>
<td>396410±1416269</td>
<td>0.37132485</td>
<td>1.48</td>
</tr>
<tr>
<td>Dolichyl-diphosphooligosaccharide-protein glycosyltransferase subunit 1 OS=Mus musculus GN=Rgp1 PE=1 SV=1</td>
<td>Q81YQ5</td>
<td>RPN1_MOUSE</td>
<td>68.5</td>
<td>5.76%</td>
<td>3</td>
<td>3</td>
<td>1563612.5±1671729</td>
<td>2536352±675215</td>
<td>0.37305617</td>
<td>1.62</td>
</tr>
<tr>
<td>Phosphorylated adapter RNA export protein OS=Mus musculus GN=Phax PE=1 SV=1</td>
<td>Q5JU7T</td>
<td>PHAX_MOUSE</td>
<td>43.2</td>
<td>21%</td>
<td>8</td>
<td>9</td>
<td>1633185±875603</td>
<td>2094875±1177060</td>
<td>0.37323368</td>
<td>1.28</td>
</tr>
<tr>
<td>Serpin H1 OS=Mus musculus GN=Serpinh1 PE=1 SV=3</td>
<td>P19324</td>
<td>SERPINH1_MOUSE</td>
<td>46.5</td>
<td>10.3%</td>
<td>4</td>
<td>7</td>
<td>9921575±176038</td>
<td>1388720±170504</td>
<td>0.37601918</td>
<td>1.40</td>
</tr>
<tr>
<td>Procollagen galactosyltransferase 1 OS=Mus musculus GN=Colg1 PE=1 SV=2</td>
<td>Q8CBQ7</td>
<td>COLGALT1_MOUSE</td>
<td>71</td>
<td>7.46%</td>
<td>5</td>
<td>5</td>
<td>670372.5±892838</td>
<td>1287580±472404</td>
<td>0.38197146</td>
<td>1.92</td>
</tr>
<tr>
<td>ATP synthase subunit alpha, mitochondrial OS=Mus musculus GN=Atpsa1 PE=1 SV=1</td>
<td>Q03265</td>
<td>ATP5A1_MOUSE</td>
<td>59.7</td>
<td>9.04%</td>
<td>4</td>
<td>4</td>
<td>942850±2881096</td>
<td>1203175±3208360</td>
<td>0.38197146</td>
<td>1.28</td>
</tr>
<tr>
<td>RNA-binding protein FUS OS=Mus musculus GN=Fus PE=1 SV=1</td>
<td>P6S699</td>
<td>FUS_MOUSE</td>
<td>52.6</td>
<td>12.9%</td>
<td>5</td>
<td>9</td>
<td>2507457±3827819</td>
<td>3049725±6536074</td>
<td>0.38628683</td>
<td>1.22</td>
</tr>
<tr>
<td>Filamin-A OS=Mus musculus GN=Fmh PE=1 SV=5</td>
<td>Q88TM8</td>
<td>FLM_MOUSE</td>
<td>281</td>
<td>8.65%</td>
<td>16</td>
<td>17</td>
<td>2007000±4023688</td>
<td>2257250±487434</td>
<td>0.38628683</td>
<td>1.12</td>
</tr>
<tr>
<td>Cluster of Keratin, type II cytoskeletal 5 OS=Mus musculus GN=Krt5 PE=1 SV=1</td>
<td>Q822L2</td>
<td>KRT5_MOUSE</td>
<td>61.7</td>
<td>NA</td>
<td>43</td>
<td>22</td>
<td>95901000±369106601</td>
<td>33616500±251650289</td>
<td>0.38628683</td>
<td>0.35</td>
</tr>
<tr>
<td>14-3-3 protein batalpha OS=Mus musculus GN=Yaheb PE=1 SV=3</td>
<td>Q8COV9</td>
<td>YWHAB_MOUSE</td>
<td>28.1</td>
<td>NA</td>
<td>2</td>
<td>2</td>
<td>2008250±4286485</td>
<td>8912875±7497941</td>
<td>0.38628683</td>
<td>0.44</td>
</tr>
<tr>
<td>Junction plakoglobin OS=Mus musculus GN=Jup PE=1 SV=3</td>
<td>Q52537</td>
<td>JUP_MOUSE</td>
<td>84.6</td>
<td>6.44%</td>
<td>4</td>
<td>4</td>
<td>15238075±4266665</td>
<td>4652400±6482450</td>
<td>0.38628683</td>
<td>0.31</td>
</tr>
<tr>
<td>Intereukin enhancer-binding factor 3 OS=Mus musculus GN=If3 PE=1 SV=2</td>
<td>Q5Z1X4</td>
<td>IF3_MOUSE</td>
<td>96</td>
<td>15.8%</td>
<td>12</td>
<td>14</td>
<td>5859900±9145458</td>
<td>4796700±16924078</td>
<td>0.38628683</td>
<td>0.82</td>
</tr>
<tr>
<td>Desmoplakin OS=Mus musculus GN=Desp PE=1 SV=1</td>
<td>Q9Q5N7</td>
<td>DSP_MOUSE</td>
<td>332.7</td>
<td>1.25%</td>
<td>4</td>
<td>4</td>
<td>14629500±4696103</td>
<td>501695±6967817</td>
<td>0.38628683</td>
<td>0.82</td>
</tr>
<tr>
<td>AMP deaminase 2 OS=Mus musculus GN=Ampe2 PE=1 SV=1</td>
<td>Q9D8T5</td>
<td>AMPD2_MOUSE</td>
<td>92</td>
<td>NA</td>
<td>46</td>
<td>18</td>
<td>19893025±5571490</td>
<td>13332350±10528321</td>
<td>0.38628683</td>
<td>0.67</td>
</tr>
<tr>
<td>AHAP1 with GTPase, ANK repeat and PH domain-containing protein 3 OS=Mus musculus GN=Agap3 PE=1 SV=1</td>
<td>Q84YH5</td>
<td>AGAP3_MOUSE</td>
<td>97.9</td>
<td>3.85%</td>
<td>3</td>
<td>3</td>
<td>4134675±1183270</td>
<td>1588475±242522</td>
<td>0.38628683</td>
<td>0.38</td>
</tr>
<tr>
<td>Cluster of Ubiquitin-40S ribosomal protein S27a OS=Mus musculus GN=Rps27a PE=1 SV=2</td>
<td>P62983</td>
<td>RPS27A_MOUSE</td>
<td>17.9</td>
<td>30.1%</td>
<td>4</td>
<td>7</td>
<td>9635200±205778</td>
<td>5217800±3756753</td>
<td>0.38628683</td>
<td>0.54</td>
</tr>
<tr>
<td>Histone H4 OS=Mus musculus GN=H4t1H4a PE=1 SV=2</td>
<td>P62806</td>
<td>HIST4A_MOUSE</td>
<td>11.4</td>
<td>29.1%</td>
<td>3</td>
<td>3</td>
<td>39631250±1606502</td>
<td>2634200±1438573</td>
<td>0.38628683</td>
<td>0.66</td>
</tr>
<tr>
<td>Protein Name</td>
<td>Accession Number</td>
<td>Uniprot ID</td>
<td>MW (kDa)</td>
<td>Max-Sequence Coverage</td>
<td>Unique Peptides</td>
<td>Total PSMs</td>
<td>Control (Mean±SD)</td>
<td>Mutant (Mean±SD)</td>
<td>P Value</td>
<td>Fold Change</td>
</tr>
<tr>
<td>--------------------------------------------------</td>
<td>------------------</td>
<td>------------</td>
<td>----------</td>
<td>-----------------------</td>
<td>----------------</td>
<td>------------</td>
<td>------------------</td>
<td>------------------</td>
<td>----------------</td>
<td>-------------</td>
</tr>
<tr>
<td>Brain acid soluble protein 1 OS=Mus musculus GN=Basp1 PE=1 SV=3</td>
<td>Q81X3</td>
<td>BASP1_MOUSE</td>
<td>22.1</td>
<td>6.19%</td>
<td>1</td>
<td>1</td>
<td>1560100±693232</td>
<td>777797.5±1155073</td>
<td>0.3862883</td>
<td>0.50</td>
</tr>
<tr>
<td>Alpha-1-antitrypsin 1-1 OS=Mus musculus GN=Serpina1a PE=1 SV=4</td>
<td>P07758</td>
<td>SERPINA1A_MOUSE</td>
<td>46</td>
<td>NA</td>
<td>3</td>
<td>3</td>
<td>1419580±386392</td>
<td>472207.5±1023676</td>
<td>0.3862883</td>
<td>0.33</td>
</tr>
<tr>
<td>TAR DNA-binding protein 43 OS=Mus musculus GN=Tardbp PE=1 SV=1</td>
<td>Q621F2</td>
<td>TARDBP_MOUSE</td>
<td>44.5</td>
<td>17.1%</td>
<td>7</td>
<td>10</td>
<td>8313450±1086648</td>
<td>6663325±1945379</td>
<td>0.3862883</td>
<td>0.80</td>
</tr>
<tr>
<td>6-phosphogluconate dehydrogenase, decarboxylating OS=Mus musculus GN=Pgd PE=1 SV=3</td>
<td>Q8DCD0</td>
<td>PGD_MOUSE</td>
<td>53.2</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>1469482±4675104</td>
<td>876217.5±1293204</td>
<td>0.3862883</td>
<td>0.58</td>
</tr>
<tr>
<td>Fructose-biphosphate aldolase A OS=Mus musculus GN=Aldoa PE=1 SV=2</td>
<td>P05064</td>
<td>ALDOA_MOUSE</td>
<td>39.3</td>
<td>25%</td>
<td>8</td>
<td>10</td>
<td>8635075±5049243</td>
<td>7398400±5320273</td>
<td>0.3862883</td>
<td>0.86</td>
</tr>
<tr>
<td>Serinalbumin OS=Mus musculus GN=Alb PE=1 SV=3</td>
<td>P07724</td>
<td>ALB_MOUSE</td>
<td>68.6</td>
<td>15.8%</td>
<td>8</td>
<td>8</td>
<td>34353600±7076873</td>
<td>20988250±7118405</td>
<td>0.3862883</td>
<td>0.90</td>
</tr>
<tr>
<td>Ras-related protein Rab-8A OS=Mus musculus GN=Rab8a PE=1 SV=2</td>
<td>P63268</td>
<td>ACTG2_MOUSE</td>
<td>41.8</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>22829500±9672418</td>
<td>20988250±7118405</td>
<td>0.3862883</td>
<td>0.92</td>
</tr>
</tbody>
</table>

Cluster of Actin, gamma-enteric smooth muscle OS=Mus musculus GN=Actg2 PE=1 SV=1 (Actin,)

<table>
<thead>
<tr>
<th>Protein Name</th>
<th>Accession Number</th>
<th>Uniprot ID</th>
<th>MW (kDa)</th>
<th>Max-Sequence Coverage</th>
<th>Unique Peptides</th>
<th>Total PSMs</th>
<th>Control (Mean±SD)</th>
<th>Mutant (Mean±SD)</th>
<th>P Value</th>
<th>Fold Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rho guanine nucleotide exchange factor 15 OS=Mus musculus GN=Arhgef15 PE=1 SV=1</td>
<td>ARHGEF15_MOUSE</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Actin-related protein 2/3 complex subunit 2 OS=Mus musculus GN=Arp2 PE=1 SV=3</td>
<td>ARPC2_MOUSE</td>
<td>34</td>
<td>7.67%</td>
<td>2</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Aspartyl/asparaginyl beta-hydroxylase OS=Mus musculus GN=Asph PE=1 SV=1</td>
<td>ASPH_MOUSE</td>
<td>83</td>
<td>18.2%</td>
<td>11</td>
<td>33</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Ataxin-2-like protein OS=Mus musculus GN=Atrx2i PE=1 SV=1</td>
<td>ATXN2L_MOUSE</td>
<td>111</td>
<td>21.9%</td>
<td>18</td>
<td>36</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Mitotic checkpoint serine/threonine-protein kinase BUB1 OS=Mus musculus GN=Bub1 PE=1 SV=1</td>
<td>BUB1_MOUSE</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Cullin-associated NEDD8-associated protein 1 OS=Mus musculus GN=Cand1 PE=1 SV=2</td>
<td>CAND1_MOUSE</td>
<td>136</td>
<td>3.33%</td>
<td>4</td>
<td>13</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Centrosomal protein of 170 kDa OS=Mus musculus GN=Cep170 PE=1 SV=2</td>
<td>CEP170_MOUSE</td>
<td>175</td>
<td>7.75%</td>
<td>9</td>
<td>12</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Calponin-3 OS=Mus musculus GN=Cnn3 PE=1 SV=1</td>
<td>CNNS3_MOUSE</td>
<td>36</td>
<td>20.6%</td>
<td>5</td>
<td>27</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Src substrate cortactin OS=Mus musculus GN=Ctn PE=1 SV=2</td>
<td>CTTN_MOUSE</td>
<td>61</td>
<td>16.8%</td>
<td>9</td>
<td>30</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Ehr domain-binding protein 1 OS=Mus musculus GN=Ehbp1 PE=1 SV=3</td>
<td>EHBP1_MOUSE</td>
<td>139</td>
<td>1.06%</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Elongator complex protein 3 OS=Mus musculus GN=Elt3 PE=1 SV=1</td>
<td>ELP3_MOUSE</td>
<td>62</td>
<td>8.04%</td>
<td>4</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Emerin OS=Mus musculus GN=Emd PE=1 SV=1</td>
<td>EMD_MOUSE</td>
<td>29</td>
<td>30.1%</td>
<td>1</td>
<td>18</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Enhancer of rudimentary homolog OS=Mus musculus GN=Erh PE=1 SV=1</td>
<td>ERH_MOUSE</td>
<td>12</td>
<td>42.3%</td>
<td>4</td>
<td>14</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Phenylalanine–RNA ligase beta subunit OS=Mus musculus GN=Farb PE=1 SV=2</td>
<td>FARSB_MOUSE</td>
<td>66</td>
<td>13.6%</td>
<td>8</td>
<td>14</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Gelsolin OS=Mus musculus GN=Gsn PE=1 SV=3</td>
<td>GSN_MOUSE</td>
<td>86</td>
<td>11.7%</td>
<td>6</td>
<td>18</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Host cell factor 1 OS=Mus musculus GN=Hcfc1 PE=1 SV=2</td>
<td>HCFC1_MOUSE</td>
<td>210</td>
<td>2.2%</td>
<td>4</td>
<td>15</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>PROTEIN NAME</td>
<td>ACCESSION NUMBER</td>
<td>MW (kDa)</td>
<td>MAX sequence coverage</td>
<td>UNIQUE PEPTIDES</td>
<td>TOTAL PSMs</td>
<td>CONTROL (Mean±SD)</td>
<td>MUTANT (Mean±SD)</td>
<td>P VALUE</td>
<td>FOLD CHANGE</td>
<td></td>
</tr>
<tr>
<td>-----------------------------------------------------------------------------</td>
<td>------------------</td>
<td>----------</td>
<td>-----------------------</td>
<td>----------------</td>
<td>------------</td>
<td>-------------------</td>
<td>------------------</td>
<td>----------</td>
<td>-------------</td>
<td></td>
</tr>
<tr>
<td>Heterochromatin protein 1-binding protein 3 OS=Mus musculus GN=Htbp3 PE=1 SV=1</td>
<td>HP1BP3_MOUSE</td>
<td>61</td>
<td>14.4%</td>
<td>7</td>
<td>12</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Peroxisomal multifunctional enzyme type 2 OS=Mus musculus GN=Hsd17bf PE=1 SV=3</td>
<td>HSD17B4_MOUSE</td>
<td>79</td>
<td>9.12%</td>
<td>6</td>
<td>15</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Lamin-B receptor OS=Mus musculus GN=Lmr PE=1 SV=2</td>
<td>LBR_MOUSE</td>
<td>71</td>
<td>10.5%</td>
<td>7</td>
<td>35</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Lipoma-preferred partner homolog OS=Mus musculus GN=Lpp PE=1 SV=1</td>
<td>LPP_MOUSE</td>
<td>66</td>
<td>23.2%</td>
<td>8</td>
<td>25</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Magnesium transporter protein 1 OS=Mus musculus GN=Magt1 PE=1 SV=1</td>
<td>MAGT1_MOUSE</td>
<td>38</td>
<td>11%</td>
<td>4</td>
<td>12</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Monofunctional C1-tetrahydrofolate synthase, mitochondrial OS=Mus musculus GN=Mthfld1 PE=1 SV=2</td>
<td>MTHFD1L_MOUSE</td>
<td>106</td>
<td>15.4%</td>
<td>12</td>
<td>25</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Phosphoenolpyruvate carboxykinase [GTP], mitochondrial OS=Mus musculus GN=Pck2 PE=1 SV=1</td>
<td>PCK2_MOUSE</td>
<td>71</td>
<td>16.7%</td>
<td>10</td>
<td>22</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Pleckstrin homology-like domain family B member 2 OS=Mus musculus GN=Rhdb2 PE=1 SV=2</td>
<td>PHLD2_MOUSE</td>
<td>141</td>
<td>19.6%</td>
<td>19</td>
<td>33</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Proline synthase co-transcribed bacterial homolog protein OS=Mus musculus GN=Prosc PE=1 SV=1</td>
<td>PROSC_MOUSE</td>
<td>30</td>
<td>23.4%</td>
<td>6</td>
<td>21</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Recombining binding protein suppressor of hairless OS=Mus musculus GN=Rbpi PE=1 SV=1</td>
<td>RBPI_MOUSE</td>
<td>59</td>
<td>16.3%</td>
<td>7</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>RuvB-like 1 OS=Mus musculus GN=Ruvb1 PE=1 SV=1</td>
<td>RUVBL1_MOUSE</td>
<td>50</td>
<td>14.5%</td>
<td>5</td>
<td>13</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Serina/threonine-protein kinase Sglk2 OS=Mus musculus GN=Sglk2 PE=1 SV=1</td>
<td>SGK2_MOUSE</td>
<td>NA</td>
<td>0%</td>
<td>1</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Na(+)/H(+) exchange regulatory cofactor NHE-RF1 OS=Mus musculus GN=Sla3r1 PE=1 SV=3</td>
<td>SLC9A3R1_MOUSE</td>
<td>39</td>
<td>21.7%</td>
<td>6</td>
<td>22</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Transmembrane emp24-domain-containing protein 2 OS=Mus musculus GN=Temed2 PE=1 SV=1</td>
<td>TMED2_MOUSE</td>
<td>23</td>
<td>4.48%</td>
<td>1</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>WD repeat and SOCS box-containing protein 2 OS=Mus musculus GN=Swsb2 PE=2 SV=2</td>
<td>WSB2_MOUSE</td>
<td>50</td>
<td>33.6%</td>
<td>10</td>
<td>33</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Lamin-associated polypeptide 2, isoforms beta/delta/epsilon/gamma OS=Mus musculus GN=Tmpo PE=1 SV=4</td>
<td>TMPO_MOUSE</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Endoribonuclease Dicer OS=Mus musculus GN=Dicer1 PE=1 SV=3</td>
<td>DICER1_MOUSE</td>
<td>217</td>
<td>6.21%</td>
<td>9</td>
<td>14</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>28 kDa heat- and acid-stable phosphoprotein OS=Mus musculus GN=Pdap1 PE=1 SV=1</td>
<td>PDAP1_MOUSE</td>
<td>21</td>
<td>31.5%</td>
<td>5</td>
<td>26</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>mRNA cap guanine-N7 methyltransferase OS=Mus musculus GN=Rnmt PE=1 SV=1</td>
<td>RNMT_MOUSE</td>
<td>53</td>
<td>9.25%</td>
<td>4</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>THUMP domain-containing protein 2 OS=Mus musculus GN=Thumpd2 PE=2 SV=3</td>
<td>THUMP2_MOUSE</td>
<td>58</td>
<td>15.7%</td>
<td>6</td>
<td>17</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Transform-2 protein homolog beta OS=Mus musculus GN=Trab2 PE=1 SV=1</td>
<td>TRAB2_MOUSE</td>
<td>34</td>
<td>8.68%</td>
<td>2</td>
<td>12</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>3'-5' exoribonuclease 1 OS=Mus musculus GN=Erf1 PE=1 SV=2</td>
<td>ER11_MOUSE</td>
<td>39</td>
<td>20.9%</td>
<td>5</td>
<td>17</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>La-related protein 1 OS=Mus musculus GN=Larp1 PE=1 SV=3</td>
<td>LARP1_MOUSE</td>
<td>121</td>
<td>36.5%</td>
<td>24</td>
<td>63</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Pre-mRNA-processing factor 19 OS=Mus musculus GN=Prpp19 PE=1 SV=1</td>
<td>PRPF19_MOUSE</td>
<td>55</td>
<td>8.53%</td>
<td>4</td>
<td>11</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>U4/U6 small nuclear ribonucleoprotein Pp31 OS=Mus musculus GN=Pp31 PE=1 SV=3</td>
<td>PRPF31_MOUSE</td>
<td>55</td>
<td>14.2%</td>
<td>5</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>PROTEIN NAME</td>
<td>ACCESSION NUMBER</td>
<td>UNIPROT ID</td>
<td>MW (kDa)</td>
<td>MAX-sequence coverage</td>
<td>UNIQUE PEPTIDES</td>
<td>TOTAL PSMs</td>
<td>CONTROL (Mean±SD)</td>
<td>MUTANT (Mean±SD)</td>
<td>P VALUE</td>
<td>FOLD CHANGE</td>
</tr>
<tr>
<td>-----------------------------------------------------------------------------</td>
<td>------------------</td>
<td>--------------</td>
<td>----------</td>
<td>------------------------</td>
<td>-----------------</td>
<td>------------</td>
<td>-------------------</td>
<td>-------------------</td>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>Squamous cell carcinoma antigen recognized by T-cells 3 OS=Mus musculus GN=Sant3 PE=1 SV=1</td>
<td></td>
<td>SART3_MOUSE</td>
<td>110</td>
<td>13.9%</td>
<td>11</td>
<td>41</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>GTP-AMP phosphotransferase AK3, mitochondrial OS=Mus musculus GN=AK3 PE=1 SV=3</td>
<td></td>
<td>AK3_MOUSE</td>
<td>25</td>
<td>35.7%</td>
<td>7</td>
<td>22</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Apoptosis inhibitor 5 OS=Mus musculus GN=Ap56 PE=1 SV=2</td>
<td></td>
<td>A5P_MOUSE</td>
<td>57</td>
<td>10.9%</td>
<td>5</td>
<td>20</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Caseinolytic peptidase B protein homolog OS=Mus musculus GN=Cpbd PE=1 SV=1</td>
<td></td>
<td>CLPB_MOUSE</td>
<td>76</td>
<td>11.5%</td>
<td>6</td>
<td>25</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Calponin-2 OS=Mus musculus GN=Cen2 PE=1 SV=1</td>
<td></td>
<td>CNN2_MOUSE</td>
<td>33</td>
<td>29.2%</td>
<td>8</td>
<td>36</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Cysteine and glycine-rich protein 1 OS=Mus musculus GN=Cspr1 PE=1 SV=3</td>
<td></td>
<td>CSRP1_MOUSE</td>
<td>21</td>
<td>28%</td>
<td>4</td>
<td>22</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Cluster of Cullin-4B OS=Mus musculus GN=Cull4b PE=1 SV=1 (CUL4B_MOUSE)</td>
<td></td>
<td>CUL4B_MOUSE</td>
<td>111</td>
<td>11.4%</td>
<td>12</td>
<td>39</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Endochilais differentiation-related factor 1 OS=Mus musculus GN=Edf1 PE=1 SV=1</td>
<td></td>
<td>EDF1_MOUSE</td>
<td>16</td>
<td>30.4%</td>
<td>4</td>
<td>21</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Band 4.1-like protein 2 OS=Mus musculus GN=Epbd412 PE=1 SV=2</td>
<td></td>
<td>EPB412_MOUSE</td>
<td>110</td>
<td>12.3%</td>
<td>10</td>
<td>25</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Leucine--RNA ligase, cytoplasmic OS=Mus musculus GN=Lars PE=1 SV=2</td>
<td></td>
<td>LARS_MOUSE</td>
<td>134</td>
<td>15.4%</td>
<td>15</td>
<td>36</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Methionine aminopeptidase 2 OS=Mus musculus GN=Metap2 PE=1 SV=1</td>
<td></td>
<td>METAP2_MOUSE</td>
<td>53</td>
<td>22%</td>
<td>8</td>
<td>29</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>(F-actin)-methionine sulfoxide oxidase MICAL3 OS=Mus musculus GN=Mical3 PE=1 SV=2</td>
<td></td>
<td>MICAL3_MOUSE</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Calcium uptake protein 1, mitochondrial OS=Mus musculus GN=Mcm1 PE=1 SV=1</td>
<td></td>
<td>MCM1_MOUSE</td>
<td>54</td>
<td>6.92%</td>
<td>3</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>1-phosphatidylinositol 4,5-bisphosphate phosphodiesterase beta-3 OS=Mus musculus GN=Pls3 PE=1 SV=2</td>
<td></td>
<td>PLCB3_MOUSE</td>
<td>139</td>
<td>6.89%</td>
<td>9</td>
<td>24</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>DNA-directed RNA polymerases I, II, and III subunit RPABC1 OS=Mus musculus GN=Pol2e PE=1 SV=1</td>
<td></td>
<td>POLR2E_MOUSE</td>
<td>25</td>
<td>19%</td>
<td>4</td>
<td>14</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Sphingosine-1-phosphate lyase 1 OS=Mus musculus GN=Spl1 PE=1 SV=1</td>
<td></td>
<td>SPL1_MOUSE</td>
<td>64</td>
<td>15.8%</td>
<td>8</td>
<td>28</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>UDP-glucuronosyltransferase 1-C OS=Mus musculus GN=Ugt1a7c PE=1 SV=1</td>
<td></td>
<td>UGT1A7C_MOUSE</td>
<td>60</td>
<td>14.3%</td>
<td>7</td>
<td>23</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Nuclear RNA helicase 2 OS=Mus musculus GN=Odx21 PE=1 SV=3</td>
<td></td>
<td>DDX21_MOUSE</td>
<td>94</td>
<td>35.1%</td>
<td>27</td>
<td>59</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Luc7-like protein 3 OS=Mus musculus GN=Luc73 PE=1 SV=1</td>
<td></td>
<td>LUC73_MOUSE</td>
<td>51</td>
<td>19.4%</td>
<td>4</td>
<td>22</td>
<td>0</td>
<td>0</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>
## Supplementary Table 8. Table of the pathways as returned by Reactome online software sorted according to the p-values of the identified interactions. The top six pathways, with a p-value<0.05 are highlighted in grey background while the ones with meaningful interactions are highlighted in bold.

<table>
<thead>
<tr>
<th>Pathway identifier</th>
<th>Pathway name</th>
<th>#Entities found</th>
<th>#Entities total</th>
<th>Interactors found</th>
<th>Interactors total</th>
<th>Entities in OR</th>
<th>Entities in ORF</th>
<th>#Reactions found</th>
<th>#Reactions total</th>
<th>Reactions ratio</th>
<th>Reactions ratio found</th>
<th>Submitted entities found</th>
<th>Mapped entities</th>
</tr>
</thead>
<tbody>
<tr>
<td>R-HSA-72162</td>
<td>mRNA Splicing – Major Pathway</td>
<td>3 168 6 633</td>
<td>0.0162498677</td>
<td>0.00004233</td>
<td>0.00002792</td>
<td>8 9</td>
<td>7.53E-04</td>
<td>Trx2;PyrF2;Pyr19</td>
<td>Q8IM5A</td>
<td>Q8IM17F2;P82065</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-72172</td>
<td>mRNA Splicing</td>
<td>3 168 6 636</td>
<td>0.016672164</td>
<td>0.00027696</td>
<td>0.00002792</td>
<td>8 9</td>
<td>8.09E-04</td>
<td>Trx2;PyrF2;Pyr19</td>
<td>Q8IM5A</td>
<td>Q8IM17F2;P82065</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-9017902</td>
<td>Noncanonical activation of NOTCH3</td>
<td>0 0 2 49</td>
<td>7.22E-04</td>
<td>0.00140675</td>
<td>0.00002792</td>
<td>1 2</td>
<td>1.76E-04</td>
<td></td>
<td>Q8IM5A</td>
<td>Q8IM17F2;P82065</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-150236</td>
<td>Transport of Mature mRNA derived from an Intron-Containing Transcript</td>
<td>0 76 3 110</td>
<td>0.00061043</td>
<td>0.00024916</td>
<td>0.00002792</td>
<td>1 4</td>
<td>3.52E-04</td>
<td></td>
<td>Q8IM5A</td>
<td>Q8IM17F2;P82065</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-73203</td>
<td>Processing of Capped Intron-Containing Pre-mRNA</td>
<td>3 343 6 725</td>
<td>0.01672208</td>
<td>0.0003362478</td>
<td>0.00002792</td>
<td>11 32</td>
<td>8.65E+03</td>
<td>Trx2;PyrF2;Pyr19</td>
<td>Q8IM5A</td>
<td>Q8IM17F2;P82065</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-45696</td>
<td>Small interfering RNA (siRNA) biogenesis</td>
<td>1 9 0 8</td>
<td>0.85E-04</td>
<td>0.00075035</td>
<td>0.00002792</td>
<td>3 5</td>
<td>4.48E-04</td>
<td>Diorc1</td>
<td>Q8IM5A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-77588</td>
<td>TSP54-Dependent Processing of Replication-Dependent Histone mRNA</td>
<td>0 11 1 2</td>
<td>9.94E-04</td>
<td>0.000189699</td>
<td>0.00002792</td>
<td>1 3</td>
<td>2.64E-04</td>
<td></td>
<td>Q8IM5A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-73202</td>
<td>Transport of Mature Transcript to Cytoplasm</td>
<td>0 85 3 228</td>
<td>0.06787538</td>
<td>0.01753478</td>
<td>0.00002792</td>
<td>1 13</td>
<td>0.00119527</td>
<td></td>
<td>Q8IM5A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-199101</td>
<td>Initiation of tRNA</td>
<td>0 2 1 22</td>
<td>1.81E-04</td>
<td>0.00140675</td>
<td>0.00002792</td>
<td>1 2</td>
<td>1.76E-04</td>
<td></td>
<td>Q8IM5A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-25027</td>
<td>MicroRNA (miRNA) biogenesis</td>
<td>1 24 0 1</td>
<td>0.03166552</td>
<td>0.00071252</td>
<td>0.00002792</td>
<td>3 8</td>
<td>7.05E-04</td>
<td>Dict1</td>
<td>Q8IM5A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-9013507</td>
<td>NOTCH3 Activation and Transmission of Signal to the Nucleus</td>
<td>0 25 2 120</td>
<td>0.02356929</td>
<td>0.001786257</td>
<td>0.00002792</td>
<td>1 15</td>
<td>0.001232327</td>
<td></td>
<td>Q8IM5A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-73080</td>
<td>RNA Polynuclease III Transcription Termination</td>
<td>0 23 1 5</td>
<td>0.03270743</td>
<td>0.00062975</td>
<td>0.00002792</td>
<td>1 2</td>
<td>1.76E-04</td>
<td></td>
<td>Q8IM5A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-73037</td>
<td>Processing of Capped Introns Pre-mRNA</td>
<td>0 26 1 6</td>
<td>0.0325778</td>
<td>0.01673507</td>
<td>0.00002792</td>
<td>1 10</td>
<td>8.91E-04</td>
<td></td>
<td>Q8IM5A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-454364</td>
<td>KSRF (KSRF) binds and destabilizes mRNA</td>
<td>0 17 2 170</td>
<td>0.01343172</td>
<td>0.03651658</td>
<td>0.00002792</td>
<td>1 5</td>
<td>4.46E-04</td>
<td></td>
<td>Q8IM5A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-167190</td>
<td>RNA Pol II CTD phosphorylation and interaction with CE during HIV infection</td>
<td>1 27 0 19</td>
<td>0.03437483</td>
<td>0.017743458</td>
<td>0.00002792</td>
<td>1 5</td>
<td>4.46E-04</td>
<td>Rienr</td>
<td>Q8IM5A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-77075</td>
<td>RNA Pol II CTD phosphorylation and interaction with CE</td>
<td>1 27 0 19</td>
<td>0.03437483</td>
<td>0.017743458</td>
<td>0.00002792</td>
<td>1 5</td>
<td>4.46E-04</td>
<td>Rienr</td>
<td>Q8IM5A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-749475</td>
<td>RNA Polynuclease III Absorbtion And Retraction Initiation</td>
<td>0 41 1 5</td>
<td>0.03271526</td>
<td>0.01643004</td>
<td>0.00002792</td>
<td>1 13</td>
<td>0.00114327</td>
<td></td>
<td>Q8IM5A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-457564</td>
<td>SUMOylation of RNA binding proteins</td>
<td>0 49 2 168</td>
<td>0.04842358</td>
<td>0.001673102</td>
<td>0.00002792</td>
<td>3 4</td>
<td>3.26E-04</td>
<td></td>
<td>Q8IM5A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-881604</td>
<td>MET activates RAS signaling</td>
<td>0 11 1 49</td>
<td>9.92E-04</td>
<td>0.017962005</td>
<td>0.00002792</td>
<td>1 10</td>
<td>8.81E-04</td>
<td></td>
<td>Q8IM5A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-9013507</td>
<td>Signaling by NOTCH3</td>
<td>0 49 2 193</td>
<td>0.04842358</td>
<td>0.017547568</td>
<td>0.00002792</td>
<td>1 37</td>
<td>0.00252958</td>
<td></td>
<td>Q8IM5A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-72186</td>
<td>mRNA Capping</td>
<td>0 26 0 34</td>
<td>0.03289327</td>
<td>0.016380112</td>
<td>0.00002792</td>
<td>7 11</td>
<td>9.66E-04</td>
<td>Rienr</td>
<td>O8IM5A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-830254</td>
<td>Metabolism of RNA</td>
<td>0 673 7 1453</td>
<td>0.0937050323</td>
<td>0.001643254</td>
<td>0.00002792</td>
<td>26 187</td>
<td>0.0714041417</td>
<td>Dict;Rienr;Trx2;PyrF2;Pyr19</td>
<td>Q8IM5A</td>
<td>Q8IM5A;Q8IM17F2;Q8IM54;Q8IM5A;Q8IM17F2;P82065</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-6762100</td>
<td>Gap-filling DNA repair synthesis and ligase in TC-NER</td>
<td>1 64 0 0</td>
<td>0.053777238</td>
<td>0.00420304</td>
<td>0.00002792</td>
<td>2 2</td>
<td>1.76E-04</td>
<td>Pyr19</td>
<td>Q8IM5A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-5323987</td>
<td>Defective Mitochordal Repair Associated With PKM2</td>
<td>0 2 1 67</td>
<td>1.81E-04</td>
<td>0.000189699</td>
<td>0.00002792</td>
<td>1 1</td>
<td>8.91E-04</td>
<td></td>
<td>Q8IM5A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

234
<table>
<thead>
<tr>
<th>Pathway Identifier</th>
<th>Pathway name</th>
<th># Entities Found</th>
<th># Entities total</th>
<th># Interactors Found</th>
<th># Interactors total</th>
<th>Entities/protein</th>
<th>Entities FOR</th>
<th># Reactions Found</th>
<th># Reactions total</th>
<th>Reactions ratio</th>
<th>Submitted Ent &amp; Reactions</th>
<th>Submitted entities found</th>
<th>Mapped entities</th>
</tr>
</thead>
<tbody>
<tr>
<td>R-HSA-474556</td>
<td>RNA Polymerase III Transcription</td>
<td>0</td>
<td>41</td>
<td>1</td>
<td>28</td>
<td>0.002731483</td>
<td>0.00273101</td>
<td>1</td>
<td>1</td>
<td>25</td>
<td>0.00233061</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-198275</td>
<td>NER Mediated Effects on Host Pathways</td>
<td>0</td>
<td>44</td>
<td>1</td>
<td>31</td>
<td>0.003872185</td>
<td>0.00380319</td>
<td>1</td>
<td>1</td>
<td>6</td>
<td>5.28E-04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-19826105</td>
<td>Dual incision in TC-NER</td>
<td>1</td>
<td>65</td>
<td>0</td>
<td>22</td>
<td>0.005895195</td>
<td>0.00445234</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>5.28E-04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-545998</td>
<td>Diseases of mismatch repair (MMR)</td>
<td>0</td>
<td>5</td>
<td>1</td>
<td>82</td>
<td>4.01E-04</td>
<td>0.34192324</td>
<td>1</td>
<td>1</td>
<td>6</td>
<td>5.28E-04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-629686</td>
<td>Interleukin-3, interleukin-5 and G-CSF signaling</td>
<td>0</td>
<td>48</td>
<td>2</td>
<td>257</td>
<td>0.09433353</td>
<td>0.24637849</td>
<td>1</td>
<td>1</td>
<td>38</td>
<td>0.00347133</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-5950548</td>
<td>Repairing of DNA Double Strand Breaks</td>
<td>0</td>
<td>6</td>
<td>1</td>
<td>84</td>
<td>5.42E-04</td>
<td>0.20579173</td>
<td>2</td>
<td>2</td>
<td>8</td>
<td>7.05E-04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-194461</td>
<td>Metabolism of non-coding RNA</td>
<td>0</td>
<td>55</td>
<td>1</td>
<td>45</td>
<td>0.04965243</td>
<td>0.25430646</td>
<td>1</td>
<td>1</td>
<td>6</td>
<td>5.28E-04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-191859</td>
<td>miRNP Assembly</td>
<td>0</td>
<td>55</td>
<td>1</td>
<td>45</td>
<td>0.04965243</td>
<td>0.25430646</td>
<td>1</td>
<td>1</td>
<td>6</td>
<td>5.28E-04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-490897</td>
<td>Asymmetric localization of FBP proteins</td>
<td>0</td>
<td>64</td>
<td>1</td>
<td>25</td>
<td>0.00277739</td>
<td>0.28313359</td>
<td>2</td>
<td>2</td>
<td>7</td>
<td>8.17E-04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-198253</td>
<td>Host interactions with influenza Factors</td>
<td>0</td>
<td>48</td>
<td>1</td>
<td>43</td>
<td>0.04933353</td>
<td>0.26153876</td>
<td>1</td>
<td>1</td>
<td>8</td>
<td>7.05E-04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-72149</td>
<td>mRNA Splicing - Minor Pathway</td>
<td>0</td>
<td>52</td>
<td>1</td>
<td>41</td>
<td>0.04964412</td>
<td>0.28423687</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td>4.49E-04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-5230555</td>
<td>Mismatch repair (MMR) directed by MSH2/MSH6 (MutSalpha)</td>
<td>0</td>
<td>14</td>
<td>1</td>
<td>82</td>
<td>0.01683088</td>
<td>0.26423687</td>
<td>1</td>
<td>1</td>
<td>10</td>
<td>8.15E-04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-5230566</td>
<td>Mismatch Repair</td>
<td>0</td>
<td>15</td>
<td>1</td>
<td>83</td>
<td>0.01951527</td>
<td>0.26423687</td>
<td>1</td>
<td>1</td>
<td>17</td>
<td>0.00447942</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-851987</td>
<td>RXR1 regulates estrogen receptor mediated transcription</td>
<td>0</td>
<td>6</td>
<td>1</td>
<td>81</td>
<td>5.42E-04</td>
<td>0.27416518</td>
<td>1</td>
<td>1</td>
<td>8</td>
<td>7.05E-04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-453141</td>
<td>Activation of the AP-1 family of transcription factors</td>
<td>0</td>
<td>10</td>
<td>1</td>
<td>103</td>
<td>9.00E-04</td>
<td>0.23012357</td>
<td>3</td>
<td>3</td>
<td>5</td>
<td>4.49E-04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-170866</td>
<td>ARMS-mediated activation</td>
<td>0</td>
<td>7</td>
<td>1</td>
<td>105</td>
<td>6.30E-04</td>
<td>0.21057384</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td>4.49E-04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-4894468</td>
<td>PTM Regulators Proteins Involved in RNA Processing</td>
<td>0</td>
<td>5</td>
<td>1</td>
<td>119</td>
<td>4.01E-04</td>
<td>0.25387841</td>
<td>2</td>
<td>2</td>
<td>7</td>
<td>6.17E-04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-676923</td>
<td>Formation of TC-NER Reunion Complex</td>
<td>1</td>
<td>53</td>
<td>0</td>
<td>75</td>
<td>0.04974649</td>
<td>0.26421302</td>
<td>2</td>
<td>2</td>
<td>7</td>
<td>6.17E-04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-211503</td>
<td>Gene Silencing by RNA</td>
<td>1</td>
<td>108</td>
<td>0</td>
<td>22</td>
<td>0.09749502</td>
<td>0.28465894</td>
<td>6</td>
<td>6</td>
<td>36</td>
<td>0.00317368</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-3009608</td>
<td>WHSC1-dependent internalization of Fzd4</td>
<td>0</td>
<td>15</td>
<td>1</td>
<td>116</td>
<td>0.01515417</td>
<td>0.28163666</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td>4.49E-04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-6292924</td>
<td>Intracellular transport</td>
<td>0</td>
<td>53</td>
<td>1</td>
<td>101</td>
<td>0.04974649</td>
<td>0.27057233</td>
<td>2</td>
<td>2</td>
<td>12</td>
<td>0.00353989</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-1617172</td>
<td>Transcription of the HIV genome</td>
<td>1</td>
<td>74</td>
<td>0</td>
<td>86</td>
<td>0.09880509</td>
<td>0.25317951</td>
<td>1</td>
<td>1</td>
<td>45</td>
<td>0.00365271</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-196981</td>
<td>Prolonged ERK activation events</td>
<td>0</td>
<td>12</td>
<td>1</td>
<td>121</td>
<td>0.01033236</td>
<td>0.27205435</td>
<td>1</td>
<td>1</td>
<td>11</td>
<td>9.68E-04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-4041258</td>
<td>Degradation of Dvl11</td>
<td>0</td>
<td>57</td>
<td>1</td>
<td>93</td>
<td>0.00547598</td>
<td>0.28579852</td>
<td>1</td>
<td>1</td>
<td>7</td>
<td>6.17E-04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-5764902</td>
<td>Deactivation of the beta-catenin/transforming complex</td>
<td>0</td>
<td>42</td>
<td>2</td>
<td>372</td>
<td>0.03370164</td>
<td>0.28364256</td>
<td>1</td>
<td>1</td>
<td>14</td>
<td>0.01231354</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-425913</td>
<td>Tissuepapin (TIP, 2F936) cists and destabilizes mRNA</td>
<td>0</td>
<td>17</td>
<td>1</td>
<td>134</td>
<td>0.01534712</td>
<td>0.29150571</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>3.52E-04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pathway identifier</td>
<td>Pathway name</td>
<td>#Entities found</td>
<td>#Entities total</td>
<td>#Interactors found</td>
<td>#Interactors total</td>
<td>#Entities ratio</td>
<td>#Interactors ratio</td>
<td>#Reactions</td>
<td>#Reactions ratio</td>
<td>Reactions ratio</td>
<td>Mapped entities</td>
<td>Submited entities</td>
<td>Mapped entities</td>
</tr>
<tr>
<td>--------------------</td>
<td>-------------------------------------------------------------------------------</td>
<td>-----------------</td>
<td>-----------------</td>
<td>-------------------</td>
<td>--------------------</td>
<td>-----------------</td>
<td>-------------------</td>
<td>--------------</td>
<td>----------------</td>
<td>-----------------</td>
<td>----------------</td>
<td>----------------</td>
<td>----------------</td>
</tr>
<tr>
<td>R-HSA-4765627</td>
<td>Transcription-Coupled Nucleoside Excision Repair (TC-NER)</td>
<td>1</td>
<td>78</td>
<td>0</td>
<td>80</td>
<td>0.007041019</td>
<td>0.007041019</td>
<td>11</td>
<td>0.00149319</td>
<td>0.00149319</td>
<td>Q8N3B4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-655524</td>
<td>Regulation of PTEN gene transcription</td>
<td>0</td>
<td>60</td>
<td>1</td>
<td>61</td>
<td>0.003466296</td>
<td>0.003466296</td>
<td>1</td>
<td>0.00125237</td>
<td>0.00125237</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-432832</td>
<td>MAPK targeted Nuclear events mediated by MAP kinase</td>
<td>0</td>
<td>31</td>
<td>1</td>
<td>32</td>
<td>0.027258852</td>
<td>0.027258852</td>
<td>3</td>
<td>0.00149319</td>
<td>0.00149319</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-147389</td>
<td>Signaling to ERRs</td>
<td>0</td>
<td>32</td>
<td>1</td>
<td>33</td>
<td>0.032888889</td>
<td>0.032888889</td>
<td>1</td>
<td>0.02411307</td>
<td>0.02411307</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-4766466</td>
<td>TPSD Regulates Transcription of DNA Repair Genes</td>
<td>0</td>
<td>65</td>
<td>1</td>
<td>104</td>
<td>0.003803145</td>
<td>0.003803145</td>
<td>1</td>
<td>0.00494732</td>
<td>0.00494732</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-433985</td>
<td>Bystander Response Factor 1 (BRF1) binds and deactivates mRNA</td>
<td>0</td>
<td>17</td>
<td>1</td>
<td>18</td>
<td>0.013541742</td>
<td>0.013541742</td>
<td>2</td>
<td>0.00282604</td>
<td>0.00282604</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-1912408</td>
<td>Pre-NOTCH Transcription and Translation</td>
<td>0</td>
<td>62</td>
<td>1</td>
<td>63</td>
<td>0.05387183</td>
<td>0.05387183</td>
<td>1</td>
<td>0.00278238</td>
<td>0.00278238</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-3028389</td>
<td>Signaling by Hip kinases</td>
<td>0</td>
<td>30</td>
<td>1</td>
<td>31</td>
<td>0.01930543</td>
<td>0.01930543</td>
<td>1</td>
<td>0.00624373</td>
<td>0.00624373</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-435521</td>
<td>Regulation of mRNAs stability by proteins that bind AU-rich elements</td>
<td>0</td>
<td>88</td>
<td>2</td>
<td>109</td>
<td>0.07845659</td>
<td>0.07845659</td>
<td>4</td>
<td>0.00289144</td>
<td>0.00289144</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-432994</td>
<td>MAP kinases activation</td>
<td>0</td>
<td>63</td>
<td>1</td>
<td>64</td>
<td>0.05874414</td>
<td>0.05874414</td>
<td>3</td>
<td>0.00291838</td>
<td>0.00291838</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-4767226</td>
<td>Major pathway of mRNA processing in the nucleus and cytoplasm</td>
<td>2</td>
<td>193</td>
<td>0</td>
<td>193</td>
<td>0.01920719</td>
<td>0.01920719</td>
<td>2</td>
<td>0.01576325</td>
<td>0.01576325</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-494546</td>
<td>Translucin-17 signaling</td>
<td>0</td>
<td>71</td>
<td>1</td>
<td>72</td>
<td>0.04056769</td>
<td>0.04056769</td>
<td>3</td>
<td>0.00330866</td>
<td>0.00330866</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-8886773</td>
<td>tRNA processing in the nucleus and cytoplasm</td>
<td>2</td>
<td>153</td>
<td>0</td>
<td>153</td>
<td>0.07124316</td>
<td>0.07124316</td>
<td>2</td>
<td>0.00129118</td>
<td>0.00129118</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-4613885</td>
<td>S100domain of DNA replication proteins</td>
<td>0</td>
<td>46</td>
<td>1</td>
<td>47</td>
<td>0.04333353</td>
<td>0.04333353</td>
<td>1</td>
<td>0.00680286</td>
<td>0.00680286</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-461665</td>
<td>Repression of VANT target genes</td>
<td>0</td>
<td>12</td>
<td>1</td>
<td>13</td>
<td>0.01633136</td>
<td>0.01633136</td>
<td>0</td>
<td>0.00282356</td>
<td>0.00282356</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-1912402</td>
<td>Pre-NOTCH Expression and Processing</td>
<td>0</td>
<td>79</td>
<td>1</td>
<td>80</td>
<td>0.03741109</td>
<td>0.03741109</td>
<td>1</td>
<td>0.00259365</td>
<td>0.00259365</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-5220224</td>
<td>B-Raf foci complex positively regulates mRNA expression</td>
<td>1</td>
<td>60</td>
<td>0</td>
<td>60</td>
<td>0.00166299</td>
<td>0.00166299</td>
<td>3</td>
<td>0.00284604</td>
<td>0.00284604</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-427273</td>
<td>Transcriptional regulation of pluripotent stem cells</td>
<td>0</td>
<td>36</td>
<td>1</td>
<td>37</td>
<td>0.03249717</td>
<td>0.03249717</td>
<td>1</td>
<td>0.00398586</td>
<td>0.00398586</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-973232</td>
<td>Ampullary fiber formation</td>
<td>0</td>
<td>71</td>
<td>1</td>
<td>72</td>
<td>0.04939678</td>
<td>0.04939678</td>
<td>1</td>
<td>0.00495128</td>
<td>0.00495128</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-72012</td>
<td>tRNA processing</td>
<td>2</td>
<td>203</td>
<td>0</td>
<td>203</td>
<td>0.01936328</td>
<td>0.01936328</td>
<td>2</td>
<td>0.01684031</td>
<td>0.01684031</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-4388492</td>
<td>FORCIE pathway</td>
<td>0</td>
<td>92</td>
<td>1</td>
<td>93</td>
<td>0.00635498</td>
<td>0.00635498</td>
<td>3</td>
<td>0.00213497</td>
<td>0.00213497</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-970445</td>
<td>Mixotic recombination</td>
<td>0</td>
<td>50</td>
<td>1</td>
<td>51</td>
<td>0.052552</td>
<td>0.052552</td>
<td>1</td>
<td>0.00568386</td>
<td>0.00568386</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-281774</td>
<td>Fcer1 mediated mRNAC activation</td>
<td>0</td>
<td>118</td>
<td>1</td>
<td>119</td>
<td>0.01821744</td>
<td>0.01821744</td>
<td>2</td>
<td>0.00461949</td>
<td>0.00461949</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-6037076</td>
<td>PTEN Regulation</td>
<td>0</td>
<td>139</td>
<td>2</td>
<td>141</td>
<td>0.01285424</td>
<td>0.01285424</td>
<td>3</td>
<td>0.00832017</td>
<td>0.00832017</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-5206338</td>
<td>Nucleoside Excision Repair</td>
<td>1</td>
<td>112</td>
<td>0</td>
<td>112</td>
<td>0.05863449</td>
<td>0.05863449</td>
<td>11</td>
<td>0.00317086</td>
<td>0.00317086</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-5220913</td>
<td>Positive regulatory role of mRNA expression</td>
<td>1</td>
<td>75</td>
<td>0</td>
<td>75</td>
<td>0.06077186</td>
<td>0.06077186</td>
<td>3</td>
<td>0.01670574</td>
<td>0.01670574</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Pathway identifier | Pathway name | #Entities found | #Entities total | #Reactions found | #Reactions total | Ranks ratio | Reactions found | Submitted entities found | Mapped entities
--- | --- | --- | --- | --- | --- | --- | --- | --- | ---
H-HSA-001559 | mTOR signaling | 0 | 40 | 1 | 284 | 0.003811085 | 0.00733895 | 1 | 28 | 0.003465038
H-HSA-988986 | Ovarian tumor domain-proteins | 0 | 36 | 1 | 274 | 0.00342032 | 0.004141401 | 1 | 14 | 0.00123154
H-HSA-255800 | Oxidative Stress Induced Senescence | 0 | 94 | 1 | 229 | 0.00480052 | 0.004141401 | 2 | 39 | 0.00342032
H-HSA-6384054 | Signaling by NET | 0 | 79 | 1 | 250 | 0.007131085 | 0.002341673 | 1 | 49 | 0.0041804
H-HSA-201491 | HIV Life Cycle | 0 | 202 | 3 | 937 | 0.01923584 | 0.0030006216 | 2 | 71 | 0.00825384
H-HSA-958494 | Beta-catenin independent WNT signaling | 0 | 143 | 1 | 210 | 0.013906323 | 0.005038769 | 3 | 51 | 0.004942035
H-HSA-416893 | Trafficking of GABA-containing AMPA receptors | 0 | 17 | 1 | 336 | 0.01134712 | 0.006602085 | 3 | 3 | 2.46E-04
H-HSA-73894 | DNA Repair | 1 | 206 | 2 | 880 | 0.007220377 | 0.007123589 | 14 | 308 | 0.007123589
H-HSA-373752 | Netrin-1 signaling | 0 | 50 | 1 | 322 | 0.004518389 | 0.005215945 | 1 | 27 | 0.0029565
H-HSA-162599 | Late Phase of HIV Life Cycle | 1 | 152 | 0 | 240 | 0.013641728 | 0.715168147 | 1 | 74 | 0.00518151
H-HSA-975338 | TSG101-mediated induction of NFAT and MAP kinase upon TLR9 or TLR4 activation | 0 | 91 | 1 | 312 | 0.00815221 | 0.716427453 | 3 | 47 | 0.00419675
H-HSA-407851 | MyoC2 cascade initiated on plasma membrane | 0 | 85 | 1 | 325 | 0.007370529 | 0.736617077 | 3 | 57 | 0.00523968
H-HSA-161940 | Toll-Like Receptor 10 (TLR10) Cascade | 0 | 85 | 1 | 335 | 0.00747552 | 0.736617077 | 3 | 58 | 0.00518782
H-HSA-161875 | Toll-Like Receptor 5 (TLR5) Cascade | 0 | 85 | 1 | 336 | 0.00747552 | 0.736617077 | 3 | 58 | 0.00518782
H-HSA-583728 | MAPK/AMPK signaling | 0 | 89 | 1 | 315 | 0.00834866 | 0.715045281 | 3 | 40 | 0.00523298
H-HSA-599386 | DNA Double Strand Break Response | 0 | 60 | 1 | 304 | 0.00516629 | 0.73504401 | 2 | 43 | 0.00279754
H-HSA-398719 | Trafficking of AMPA receptors | 0 | 31 | 1 | 362 | 0.01798592 | 0.73506233 | 3 | 4 | 7.92E-04
H-HSA-398621 | Glutamatergic binding, activation of AMPA receptor and synaptic plasticity | 0 | 31 | 1 | 362 | 0.01798592 | 0.73506233 | 3 | 9 | 7.92E-04
H-HSA-162024 | Inflammation | 0 | 168 | 1 | 245 | 0.01529636 | 0.7370743 | 1 | 59 | 0.00518684
H-HSA-407155 | MyoC2 dependent cascade initiated on endosome | 0 | 92 | 1 | 344 | 0.09305498 | 0.74283519 | 3 | 62 | 0.00541112
H-HSA-163191 | Toll-Like Receptor 7/8 (TLR7/8) Cascade | 0 | 92 | 1 | 344 | 0.09305498 | 0.74283519 | 3 | 63 | 0.00541112
H-HSA-162587 | HIV Life Cycle | 1 | 162 | 0 | 260 | 0.041715175 | 0.74283519 | 1 | 112 | 0.005186324
H-HSA-888010 | TGF2 (β2-family) regulates transcription of growth factors and their receptors | 0 | 13 | 1 | 367 | 0.01173463 | 0.74505772 | 2 | 18 | 0.001585484
H-HSA-161938 | Toll-Like Receptor 9 (TLR9) Cascade | 0 | 96 | 1 | 347 | 0.09866966 | 0.74947966 | 3 | 66 | 0.00831441
H-HSA-8882683 | Degradated CD40 ligands, multiple neurodegenerative pathways in Alzheimer’s disease models | 0 | 23 | 2 | 777 | 0.025760774 | 0.752749225 | 2 | 22 | 0.01937014
H-HSA-8883678 | Neurodegenerative Diseases | 0 | 23 | 2 | 777 | 0.025760774 | 0.752749225 | 2 | 22 | 0.01937014
<table>
<thead>
<tr>
<th>Pathway identifier</th>
<th>Pathway name</th>
<th>#Entities found</th>
<th>#Entities total</th>
<th>#Interactors found</th>
<th>#Interactors total</th>
<th>Entities prolific</th>
<th>Entities FOR</th>
<th>#Reactions found</th>
<th>Reactions total</th>
<th>#Reactions in total</th>
<th>pValue</th>
<th>#Submitted entities</th>
<th>#Mapped entities</th>
</tr>
</thead>
</table>
| R-HSA-8848751      | Regulation of PTEN stability and activity | 0 69 1 347 | 0.0982319454 0.751208945 0.751208945 2 13 | 0.011943707 
| R-HSA-187327       | Signaling by NTRK1 (TRKA) | 0 79 1 339 | 0.0713188844 0.752088906 0.752088906 1 85 | 0.004845034 
| R-HSA-186538       | Myo/BB Mat cascade initiated on plasma membrane | 0 95 1 360 | 0.06876329 0.753820418 0.753820418 3 63 | 0.005849149 
| R-HSA-187118       | Signaling by NOTCH | 0 203 2 634 | 0.010308161 0.708798249 0.708798249 2 151 | 0.012334999 
| R-HSA-186826       | DD3/BV8H1-mediated induction of interferon-alpha/beta | 0 77 1 357 | 0.096825145 0.758769322 0.758769322 1 46 | 0.006017902 
| R-HSA-186179       | Toll Like Receptor TLR1/TLR2 Cascade | 0 98 1 367 | 0.088647161 0.758503333 0.758503333 3 65 | 0.007235069 
| R-HSA-186188       | Toll Like Receptor TLR1/TLR2 Cascade | 0 95 1 367 | 0.088647161 0.758503333 0.758503333 3 65 | 0.007235069 
| R-HSA-181438       | Toll Like Receptor 3 (TLR3) Cascade | 0 98 1 367 | 0.088647161 0.758503333 0.758503333 3 67 | 0.003810249 
| R-HSA-186184       | Toll Like Receptor 3 (TLR3) Cascade | 0 97 1 362 | 0.0713188844 0.752088906 0.752088906 3 60 | 0.002846947 
| R-HSA-301619       | Estrogen-dependent gene expression | 0 129 2 742 | 0.019332355 0.775739242 0.775739242 2 64 | 0.008372767 
| R-HSA-303621       | ESRI-mediated signaling | 0 128 2 742 | 0.011276461 0.775739242 0.775739242 2 69 | 0.009276898 
| R-HSA-307561       | TRIF(TICAM)-mediated TLR4 signaling | 0 98 1 401 | 0.08487615 0.79660873 0.79660873 3 57 | 0.005283669 
| R-HSA-186185       | Myo/BB-independent TLR4 cascade | 0 98 1 401 | 0.08487615 0.79660873 0.79660873 3 59 | 0.005698564 
| R-HSA-186232       | Degradation of beta-catenin by the destruction complex | 0 82 1 388 | 0.07482735 0.788497566 0.788497566 1 22 | 0.001976814 
| R-HSA-245422       | Fc epsilon receptor (FCER1) signaling | 0 217 1 261 | 0.019801414 0.801407033 0.801407033 2 63 | 0.005491914 
| R-HSA-300801       | Signaling by Nuclear Receptor | 0 168 2 746 | 0.015269309 0.811799889 0.811799889 2 90 | 0.007274743 
| R-HSA-407605       | Neurotrophic degeneration | 1 400 0 0 | 0.04333632 0.809487968 0.809487968 1 10 | 8.81E-04 Faspl1 C10490 
| R-HSA-125704       | PPAR activator AKT signaling | 0 268 2 887 | 0.02684564 0.860804011 0.860804011 1 85 | 0.007487538 
| R-HSA-849654       | SUMOylation of intracellular receptors | 0 31 1 474 | 0.027985062 0.813862891 0.813862891 1 25 | 0.00203161 
| R-HSA-216365       | Epigenetic regulation of gene expression | 1 117 0 416 | 0.01026327 0.816401983 0.816401983 3 33 | 0.005863721 Ddx31 Q9NT20 
| R-HSA-130520       | Meiosis | 0 88 1 447 | 0.07495889 0.823132857 0.823132857 1 15 | 0.013231237 
| R-HSA-587703       | Ciliary Assembly | 0 201 1 383 | 0.01945707 0.782755625 0.782755625 2 50 | 0.004841322 
| R-HSA-8848755       | Intracellular signaling by second messengers | 0 306 2 704 | 0.027548716 0.848153574 0.848153574 3 104 | 0.009190574 
| R-HSA-187221       | Signaling by Wnt | 0 257 3 1123 | 0.026812314 0.845270443 0.845270443 6 158 | 0.012798681 
| R-HSA-14741865      | Reproduction | 0 114 1 452 | 0.012451585 0.841376749 0.841376749 1 24 | 0.002113579 
<p>| R-HSA-8848751       | Signaling by PTEN | 0 57 1 513 | 0.055147588 0.946807469 0.946807469 2 52 | 0.005832887 |</p>
<table>
<thead>
<tr>
<th>Pathway identifier</th>
<th>Pathway name</th>
<th>#Entities found</th>
<th>total</th>
<th>#Interactors found</th>
<th>total</th>
<th>#Interactors ratio</th>
<th>Entities ratio</th>
<th>Entities FDR</th>
<th>Entities FDR ratio</th>
<th>Entities FO</th>
<th>Entities FO ratio</th>
<th>#Interactors total</th>
<th>Interactors ratio</th>
<th>FDR ratio</th>
<th>#Submitted entities</th>
<th>Submitted entities ratio</th>
<th>#Map entities</th>
<th>Map entities ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pathway identifier</td>
<td>Pathway name</td>
<td>#Entities found</td>
<td>#Entities total</td>
<td>#Interactors found</td>
<td>#Interactors total</td>
<td>Entities pValue</td>
<td>Entities FDR</td>
<td>#Reactions found</td>
<td>#Reactions total</td>
<td>Reactions ratio</td>
<td>Submitted entities found</td>
<td>Mapped entities</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------</td>
<td>-------------------------------------------------</td>
<td>-----------------</td>
<td>-----------------</td>
<td>--------------------</td>
<td>--------------------</td>
<td>----------------</td>
<td>--------------</td>
<td>-----------------</td>
<td>-----------------</td>
<td>----------------</td>
<td>--------------------------</td>
<td>----------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-1643985</td>
<td>Disease</td>
<td>1</td>
<td>1148</td>
<td>4</td>
<td>2413</td>
<td>0.103830199</td>
<td>0.99927836</td>
<td>0.98927836</td>
<td>5</td>
<td>0.081919874</td>
<td>Rnmt</td>
<td>CH4HG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-183249</td>
<td>Innate Immune System</td>
<td>1</td>
<td>1180</td>
<td>2</td>
<td>1528</td>
<td>0.10027038</td>
<td>0.96343332</td>
<td>0.96343332</td>
<td>5</td>
<td>0.06837012</td>
<td>Plap1</td>
<td>O13HG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-1262215</td>
<td>Cytokine Signaling in immune system</td>
<td>0</td>
<td>771</td>
<td>3</td>
<td>2042</td>
<td>0.069616838</td>
<td>0.96416838</td>
<td>4</td>
<td>626</td>
<td>0.05664446</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-3703938</td>
<td>Transcriptional Regulation by TP53</td>
<td>0</td>
<td>367</td>
<td>1</td>
<td>1307</td>
<td>0.053131714</td>
<td>0.96925734</td>
<td>0.96925734</td>
<td>1</td>
<td>0.027215271</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-424275</td>
<td>Axon guidance</td>
<td>0</td>
<td>558</td>
<td>1</td>
<td>1279</td>
<td>0.050194085</td>
<td>0.97753932</td>
<td>0.97753932</td>
<td>2</td>
<td>0.020072404</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-1649710</td>
<td>Cell Cycle</td>
<td>0</td>
<td>621</td>
<td>1</td>
<td>1400</td>
<td>0.056621111</td>
<td>0.96821873</td>
<td>1</td>
<td>423</td>
<td>0.02725874</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-138582</td>
<td>Hemostasis</td>
<td>0</td>
<td>723</td>
<td>1</td>
<td>1294</td>
<td>0.065270138</td>
<td>0.96885925</td>
<td>1</td>
<td>327</td>
<td>0.02889396</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-74180</td>
<td>Gene expression (Transcription)</td>
<td>3</td>
<td>1416</td>
<td>3</td>
<td>3113</td>
<td>0.137832445</td>
<td>0.96913733</td>
<td>0.96913733</td>
<td>15</td>
<td>0.071982088</td>
<td>Rnmt;Scrat;Q125588</td>
<td>CH4HG;Q125588;Q125588</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-73687</td>
<td>RNA Polynase II Transcription</td>
<td>1</td>
<td>1214</td>
<td>3</td>
<td>2821</td>
<td>0.11013013</td>
<td>0.96924899</td>
<td>0.96924899</td>
<td>5</td>
<td>0.02627489</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-231366</td>
<td>General Transcription Pathway</td>
<td>0</td>
<td>1152</td>
<td>3</td>
<td>2913</td>
<td>0.13399978</td>
<td>0.96664263</td>
<td>0.96664263</td>
<td>4</td>
<td>0.00707274</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-1264758</td>
<td>Developmental Biology</td>
<td>0</td>
<td>1033</td>
<td>2</td>
<td>2268</td>
<td>0.06951016</td>
<td>0.98665210</td>
<td>0.98665210</td>
<td>4</td>
<td>0.045043521</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-1935256</td>
<td>Immune System</td>
<td>1</td>
<td>2265</td>
<td>4</td>
<td>3053</td>
<td>0.203569329</td>
<td>0.98964313</td>
<td>0.98964313</td>
<td>6</td>
<td>0.13868442</td>
<td>Plap1</td>
<td>O13HG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-597382</td>
<td>Post-translational protein modification</td>
<td>0</td>
<td>1415</td>
<td>2</td>
<td>2339</td>
<td>0.127742198</td>
<td>0.98662946</td>
<td>0.98662946</td>
<td>8</td>
<td>0.04833985</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-162592</td>
<td>Signal Transduction</td>
<td>0</td>
<td>2738</td>
<td>5</td>
<td>4507</td>
<td>0.247170639</td>
<td>0.96969829</td>
<td>0.96969829</td>
<td>20</td>
<td>0.18279081</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-HSA-329469</td>
<td>Metabolism of proteins</td>
<td>0</td>
<td>2111</td>
<td>2</td>
<td>3383</td>
<td>0.193275065</td>
<td>0.98669994</td>
<td>0.98669994</td>
<td>7</td>
<td>0.07027621</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Supplementary Figure 1. Representative MS/MS spectra for the proteins binding only the mutant transcript.
Supplementary Figure 2. Representative MS/MS spectra for the proteins binding only the mutant transcript.