

# Investigating microRNAs during cardiac development in the chick

Dominique McCormick

Doctor of Philosophy

University of East Anglia

School of Biological Sciences

June 2014

## Abstract

The heart, the first organ to function in developing vertebrate embryos<sup>1</sup>, is affected by congenital defects in roughly 1% of the worldwide human population. These abnormalities frequently require treatment for the entire lifespan of the affected individual<sup>2,3</sup>.

microRNAs are an exciting class of non-coding regulatory genes that are known to be essential for the heart development of many organisms<sup>4,5</sup>. A miRNA becomes incorporated into the miRNA-induced silencing complex, which binds to a target mRNA transcript and inhibits its translation. Dysregulation of miRNAs is associated with failing human hearts<sup>6</sup>. More than 2,500 mature human miRNAs are currently annotated, and with each miRNA predicted to potentially target several hundred transcripts, the further study of these molecules is vital for the complete understanding of cardiac biology.

A miRNA expression profile in the chick heart, a model of human heart development, over several stages of embryonic development was generated by next generation sequencing. Two novel chick miRNAs were discovered, and 31 known animal miRNAs not yet annotated in the chick genome were discerned. The predicted targets of three miRNAs thought to be essential for heart development, miR-130, miR-138, and miR-499, were examined *in vitro*. Twelve genes were confirmed as targets of at least one of these three miRNAs by luciferase reporter assay, including genes in the Wnt signalling pathway which is essential at specific stages of heart development<sup>7</sup>.

The roles of miR-1, known to be indispensable for heart formation and function, were also studied in the developing chick heart. Two targets of miR-1, FZD7 and RAR $\beta$ , had been identified *in vitro*<sup>8</sup>. This study showed that inhibition of miR-1 in the HH14 chick resulted increased FZD7 and RAR $\beta$  protein levels, as well as in delayed development, possibly due to a loss of inhibition of canonical Wnt signalling in the epicardium.

# Table of Contents

<b>1</b>	<b>INTRODUCTION</b>	<b>1</b>
<b>1.1</b>	<b>Chick heart development</b>	<b>1</b>
1.1.1	Up to HH4	2
1.1.2	HH5 to HH7	3
1.1.3	HH8 to HH10	3
1.1.4	HH11 to HH13	4
1.1.5	HH14 to HH16	5
1.1.6	HH17 to HH20	6
1.1.7	HH21 to HH26	7
<b>1.2</b>	<b>Congenital heart diseases</b>	<b>7</b>
<b>1.3</b>	<b>Wnt signalling is essential for heart formation and development</b>	<b>7</b>
1.3.1	Wnt / $\beta$ -catenin pathway	9
1.3.2	Non-canonical Wnt pathways	9
1.3.3	Wnt signalling is context-specific	10
1.3.4	Wnt signalling is essential for heart development	11
<b>1.4</b>	<b>Factors involved in chick heart development</b>	<b>12</b>
1.4.1	Specification	13
1.4.2	Determination and heart field formation	15
1.4.3	Patterning within the bilateral heart fields	15
1.4.4	Differentiation and heart tube formation	16
<b>1.5</b>	<b>MicroRNAs</b>	<b>17</b>
1.5.1	miRNA biogenesis	17
1.5.2	miRNA targeting principles	19
1.5.3	miRNA regulation of gene expression	19
<b>2</b>	<b>MATERIALS AND METHODS</b>	<b>21</b>
<b>2.1</b>	<b>Embryo development</b>	<b>21</b>
<b>2.2</b>	<b>Dissection of embryos</b>	<b>21</b>
2.2.1	For whole mount in situ hybridization (WISH)	21

2.2.2	For RNA purification	21
2.2.3	For protein purification	21
<b>2.3</b>	<b>Purification and quantification of DNA, RNA and protein</b>	<b>22</b>
2.3.1	Plasmid DNA purification	22
2.3.2	DNA quantification	22
2.3.3	Total RNA purification	22
2.3.4	RNA quantification	23
2.3.5	Protein purification	23
2.3.6	Protein quantification	23
<b>2.4</b>	<b>In situ hybridisation</b>	<b>23</b>
2.4.1	mRNA probe synthesis	23
2.4.1.1	Plasmid preparation	23
2.4.1.2	RNA synthesis	24
2.4.2	mRNA whole mount in situ hybridisation	24
2.4.2.1	Hybridization	24
2.4.2.2	Post-hybridization washes	25
2.4.2.3	Antibody	25
2.4.2.4	Post-antibody washes	25
2.4.2.5	NBT/BCIP detection	25
2.4.3	LNA whole mount in situ hybridisation	26
<b>2.5</b>	<b>Cryosectioning of WISH embryos</b>	<b>26</b>
<b>2.6</b>	<b>In ovo microinjection into HH14 chick heart</b>	<b>26</b>
2.6.1	AntagomiR and RCAS virus injections	26
<b>2.7</b>	<b>Library preparation for Next Generation Sequencing</b>	<b>27</b>
2.7.1	TruSeq library preparation	27
2.7.1.1	mirVana small RNA enrichment	27
2.7.1.2	Illumina TruSeq Small RNA library preparation	28
2.7.2	Illumina HiSeq Sequencing	28
<b>2.8</b>	<b>Bioinformatics analysis</b>	<b>28</b>
2.8.1	Raw sequence mapping and filtering	29
2.8.2	Profiling of mapped reads to miRBase	29
2.8.3	Fold change profiles	29

2.8.4	Novel miRNA prediction	30
2.8.5	Novel miRNA target prediction	30
2.8.6	FunDo Disease Ontology analysis	31
<b>2.9</b>	<b>miRNA target selection</b>	<b>32</b>
2.9.1	TargetScan target identification	32
2.9.2	GO terms analysis of miRNA targets identified by TargetScan	32
2.9.3	Design of primers for 3' UTR amplification of predicted targets	32
<b>2.10</b>	<b>Experimental evaluation of putative miRNA targets</b>	<b>34</b>
2.10.1	PCR of 3' UTRs	34
2.10.1.1	Reverse Transcription	34
2.10.1.2	PCR	35
2.10.1.3	Agarose gel electrophoresis and purification of PCR products	35
2.10.2	Generating miRNA-sensing luciferase constructs	36
2.10.2.1	Ligation into pGEM-T Easy	36
2.10.2.2	Transformation into DH5 $\alpha$ E.coli competent cells	36
2.10.2.3	Plasmid purification	37
2.10.2.4	Restriction of amplified 3' UTRs	37
2.10.2.5	Ligation into modified pGL3 vector	37
2.10.2.6	Colony PCR identification of constructs containing 3' UTR inserts	38
2.10.3	Generating pGL3 mutant constructs	38
2.10.3.1	Mutagenesis PCR of pGL3 constructs	39
2.10.3.2	Gibson assembly of PCR products	40
2.10.4	Luciferase assays of miRNA sensing constructs	41
2.10.4.1	Cell culture	41
2.10.4.2	Cell transfection	41
2.10.4.3	Luciferase assay	42
2.10.4.4	Normalisation of luciferase assay data	42
<b>2.11</b>	<b>Western blotting</b>	<b>43</b>
<b>2.12</b>	<b>Northern blotting</b>	<b>43</b>
2.12.1	Polyacrylamide gel electrophoresis	44
2.12.2	Semi-dry transfer and cross-linking	44
2.12.3	Probing	45
2.12.3.1	Pre-hybridisation	45
2.12.3.2	$\gamma$ <sup>32</sup> P-ATP end labelling of complementary oligonucleotides	45

<b>2.13</b>	<b>Image J quantification of Western and Northern blots</b>	<b>46</b>
<b>2.14</b>	<b>Fold</b>	<b>46</b>
<b>3</b>	<b>SMALL RNA SEQUENCING OF CHICK CARDIAC AND CARIOGENIC TISSUE</b>	<b>47</b>
<b>3.1</b>	<b>Introduction</b>	<b>47</b>
3.1.1	Dissection of tissues for heart tissue-enriched miRNA sequencing	47
3.1.2	Next Generation Sequencing technologies	49
3.1.3	Techniques for assessing microRNA expression	50
3.1.4	NGS Library preparation and sequencing	52
3.1.5	Previous sequencing experiments	55
<b>3.2</b>	<b>Results</b>	<b>57</b>
3.2.1	Illumina TruSeq libraries were enriched for 22-nucleotide RNAs	57
3.2.2	Many animal miRNAs are expressed during chick cardiogenesis	59
3.2.3	A large fraction of known chicken miRNAs are expressed from gastrulation	60
3.2.4	Samples cluster according to their known animal miRNA expression	62
3.2.5	miRNA expression is dominated by two miRNAs	63
3.2.6	Enrichment of predicted target genes of top expressed miRNAs	66
3.2.7	Unique predicted target genes of top expressed miRNAs	70
<b>3.3</b>	<b>Discussion</b>	<b>72</b>
3.3.1	Illumina TruSeq libraries were enriched for 22-nucleotide RNAs	72
3.3.2	Many animal miRNAs are expressed during chick cardiogenesis	72
3.3.3	A large fraction of known chicken miRNAs are expressed from gastrulation	73
3.3.4	Samples cluster according to their known animal miRNA expression	73
3.3.5	miRNA expression is dominated by two miRNAs	74
3.3.6	Enrichment of predicted target genes of top expressed miRNAs	76
3.3.7	Unique predicted target genes of top expressed miRNAs	78
<b>4</b>	<b>EVALUATION OF MIRNAS IMPORTANT DURING HEART DEVELOPMENT</b>	<b>80</b>
<b>4.1</b>	<b>Introduction</b>	<b>80</b>
4.1.1	miRNAs are essential for heart development and function	80
4.1.2	miRNAs influence proliferation and differentiation	81
4.1.3	MicroRNAs involved in vasculogenesis and angiogenesis	81
4.1.4	miRNAs involved in patterning the heart	82

4.1.5	Other miRNAs	82
<b>4.2</b>	<b>Results</b>	<b>85</b>
4.2.1	Few miRNAs have large increases in expression	85
4.2.2	Creating a profile of miRNAs in Heart by subtraction of NT sample values	88
4.2.3	miRNAs with a role in heart development and function were detected	91
4.2.4	11 Known heart miRNAs correlate well by fold change	96
4.2.5	PTEN is an enriched target of most of the miRNAs with functions in the heart	99
4.2.6	Selected correlated miRNA fold change patterns	101
4.2.7	Validating targets of miR-130, miR-138 and miR-499	105
<b>4.3</b>	<b>Discussion</b>	<b>115</b>
4.3.1	Few miRNAs have large increases in expression	115
4.3.2	Creating a profile of miRNAs in Heart by subtraction of NT sample values	117
4.3.3	miRNAs with a role in heart development and function were detected	117
4.3.4	11 Known heart miRNAs correlate well by fold change	118
4.3.5	PTEN is an enriched target of most of the miRNAs with functions in the heart	119
4.3.6	Selected correlated miRNA fold change patterns	121
4.3.7	Validating targets of miR-130, miR-138 and miR-499	123
<b>5</b>	<b>NOVEL MIRNAS</b>	<b>129</b>
5.1	Introduction	129
5.2	Results	131
5.3	Discussion	137
<b>6</b>	<b>THE ROLE OF MIR-1 DURING EARLY CHICK HEART DEVELOPMENT</b>	<b>139</b>
6.1	Introduction	139
6.1.1	miR-1	139
6.1.2	Frizzled 7	146
6.1.3	Retinoic acid receptor $\beta$	148
6.2	Results	152
6.2.1	FZD7 and RAR $\beta$ mRNA is expressed in the developing chick heart	152
6.2.2	miR-1 reduces FZD7 and RAR $\beta$ protein levels in the heart	154

6.2.3	Reduction of miR-1 impairs embryonic growth	156
6.2.4	Reduction of growth is not seen when miR-206 is inhibited	158
<b>6.3</b>	<b>Discussion</b>	<b>160</b>
<b>7</b>	<b>GENERAL DISCUSSION AND CONCLUDING REMARKS</b>	<b>164</b>
<b>7.1</b>	<b>Study aims</b>	<b>164</b>
<b>7.2</b>	<b>miRNA expression profile</b>	<b>165</b>
7.2.1	Known animal miRNAs not annotated in the chicken genome were identified by NGS	165
7.2.2	miR-130 is highly expressed in heart tissue samples	165
7.2.3	miR-26, miR-30 and miR-92 were highly expressed across all stages	166
<b>7.3</b>	<b>miRNAs important during heart development</b>	<b>166</b>
7.3.1	Finding highly expressed miRNAs by generation of a fold change profile	166
7.3.2	Establishment of a heart miRNA expression profile, generated by subtracting miRNA fold change values in matched NT samples from Heart samples	167
7.3.3	Validation of miRNA targets	167
<b>7.4</b>	<b>Novel miRNAs</b>	<b>168</b>
<b>7.5</b>	<b>miR-1 affects growth during early chick heart development</b>	<b>169</b>
<b>8</b>	<b>REFERENCES</b>	<b>171</b>
	<b>APPENDIX A</b>	<b>180</b>
	<b>APPENDIX B</b>	<b>192</b>
<b>B.1</b>	<b>Oligonucleotide sequences</b>	<b>192</b>
B.1.1	Illumina adapter sequences	192
<b>B.2</b>	<b>AntagomiR sequences</b>	<b>193</b>
<b>B.3</b>	<b>Primers for amplifying 3' UTR fragments of predicted target genes</b>	<b>193</b>
<b>B.4</b>	<b>Mutagenesis primers for generation of mutant reporter constructs with nucleotides in the predicted target site mutated</b>	<b>194</b>

<b>B.5</b>	<b>Northern blot oligos</b>	<b>196</b>
 <b>APPENDIX C</b>		 <b>199</b>
<b>C.1</b>	<b>Cloned 3' UTR sequences</b>	<b>199</b>
C.1.1	ATG7 3' Region with miR-499 (green) and miR-138 (pink) target site	199
C.1.2	ENAH A 3' Region with miR-499 target site	200
C.1.3	ENAH B 3' Region with two miR-499 target sites	200
C.1.4	ETS1 3' Region with miR-499 target site	201
C.1.5	FZD8 3' Region with miR-499 target site	202
C.1.6	LEF1 3' Region with miR-130 target site	202
C.1.7	NFIB A 3' Region with two miR-138 target sites	203
C.1.8	NFIB B 3' Region with miR-130 (blue) and miR-138 (pink) target sites	204
C.1.9	NR3C1 3' Region with miR-138 target site	205
C.1.10	PPP3CB 3' Region with miR-499 target site	205
C.1.11	PRKACB 3' Region with miR-130 target site	206
C.1.12	QKI 3' Region with miR-130 (blue) and miR-499 (green) target sites	207
C.1.13	ROCK2 3' Region with miR-130 (blue) and miR-138 (pink) target sites	208
C.1.14	TCF12 3' Region with miR-138 (pink) and miR-499 (green) target sites	208
C.1.15	VANGL2 3' Region with miR-138 (pink) and miR-499 (green) target sites	210
C.1.16	ZEB2 3' Region with miR-130 (blue) and miR-138 (pink) target sites	210
 <b>C.2</b>	 <b>Mutated target sites</b>	 <b>212</b>
C.2.1	ATG7 miR-138 mutant	212
C.2.2	ATG7 miR-499 mutant	212
C.2.3	ETS1 miR-499 mutant	212
C.2.4	FZD8 miR-499 mutant	212
C.2.5	LEF1 miR-130 mutant	212
C.2.6	NFIB B miR-130 mutant	212
C.2.7	NFIB B miR-138 mutant	213
C.2.8	NR3C1 miR-138 mutant	213
C.2.9	PPP3CB miR-499 mutant	213
C.2.10	PRKACB miR-130 mutant	213
C.2.11	ROCK2 miR-130 mutant	213
C.2.12	TCF12 miR-138 mutant	213
C.2.13	ZEB2 miR-130 mutant	213

<b>APPENDIX D</b>	<b>214</b>
D.1 Expression of FZD7 mRNA during early stages of chicken development	214
D.2 Exploration of miR-1 dysregulation during heart development	215
<b>APPENDIX E</b>	<b>221</b>

## Table of Figures

FIGURE 1.1: A MODEL OF A STAGE HH3 AND A HH5-7 CHICK EMBRYO	2
FIGURE 1.2: ADAPTED SCHEMATIC DRAWINGS OF HH14 TO HH17 CHICK EMBRYOS.	5
FIGURE 1.3: SCHEMATIC DRAWING OF A CROSS-SECTION THROUGH THE LOOPING HEART	6
FIGURE 1.4: A SUMMARY OF THE CURRENTLY KNOWN REQUIREMENTS FOR CANONICAL WNT AND NON-CANONICAL WNT SIGNALLING IN THE DEVELOPING EMBRYONIC HEART UP TO CHICK STAGE HH20	12
FIGURE 1.5: ORIGINS OF CARDIAC INDUCING AND INHIBITING FACTORS	14
FIGURE 1.6: THE BIOGENESIS OF MIRNAS.	18
FIGURE 2.1: PGL3 LUCIFERASE MIRNA-SENSING CONSTRUCT.	39
FIGURE 3.1: TISSUE DISSECTION PERFORMED PRIOR TO SMALL RNA ENRICHMENT.	49
FIGURE 3.2: A BRIEF OVERVIEW OF ILLUMINA TRUSEQ SMALL RNA LIBRARY PREPARATION AND PAIRED END SEQUENCING.	53
FIGURE 3.3: KNOWN MIRNA PROFILE IN ADULT MOUSE HEARTS SHOW A FEW MIRNAS DOMINATE SEQUENCING SPACE.	56
FIGURE 3.4: SIZE DISTRIBUTION OF SEQUENCED SMALL-RNA ENRICHED TOTAL RNA.	58
FIGURE 3.5: KNOWN MIRNAS ACCOUNT FOR MOST OF THE READS FROM THE MIRNA SIZED FRACTION OF ALMOST ALL THE SAMPLES, AND MANY KNOWN CHICKEN MIRNAS WERE DETECTED.	61
FIGURE 3.6: REPRESENTATION OF FIVE RANGES OF KNOWN ANIMAL MIRNA NORMALISED READS.	63
FIGURE 3.7: MIRNA EXPRESSION WAS DOMINATED BY TWO MIRNAS AND THREE MIRNAS WERE CONSISTENTLY EXPRESSED AT A VERY HIGH LEVEL IN ALL NINE SAMPLES.	65
FIGURE 3.8 THREE TOP EXPRESSED MIRNAS ARE COMMON TO ALL SAMPLES AND MIR-10 EXPRESSION INCREASES RAPIDLY DURING DEVELOPMENT.	68
FIGURE 3.9: ANALYSIS OF PREDICTED TARGET GENES OF THE THREE TOP EXPRESSED MIRNAS COMMON TO ALL NINE SAMPLES.	69
FIGURE 3.10: A LARGE FRACTION OF THE PREDICTED TARGETS OF MIR-26 AND MIR-92 THAT ARE EXPRESSED IN THE HUMAN HEART ARE ALSO PREDICTED TARGETS OF THE OTHER TWO HIGHLY DETECTED MIRNAS.	69
FIGURE 4.1: THE MATURE SEQUENCE AND GENOME LOCATION OF MIR-130 (A), MIR-138 (B), AND MIR-499 (C) IN THE CHICK.	82
FIGURE 4.2: OVERVIEW OF ANALYSES PERFORMED IN CHAPTER 4.	85
FIGURE 4.3: THE FOLD CHANGE OF NORMALISED READS PER MILLION FOR EACH KNOWN ANIMAL MIRNA DETECTED IN THE NINE SAMPLES COMPARED TO READS AT HH5-7.	86
FIGURE 4.4: THE FOLD CHANGE PATTERNS OF SIX OF THE 14 MIRNAS WITH A LARGER THAN 100 FOLD INCREASE OVER THEIR READS DETECTED AT HH5-7 IN AT LEAST ONE SAMPLE.	88

FIGURE 4.5 MIRNAS KNOWN TO BE INVOLVED IN HEART DEVELOPMENT AND FUNCTION FROM CURRENT LITERATURE CLUSTER WELL BASED ON THEIR FOLD CHANGE PATTERNS ACROSS ALL SAMPLES.	94
FIGURE 4.6: THE PERCENTAGE OF THE TOTAL NORMALISED READS DUE TO MIRNAS WITH KNOWN ROLES IN THE HEART.	95
FIGURE 4.7: AN ASSESSMENT OF THE EXTENT TO WHICH 11 MIRNAS WITH KNOWN ROLES IN HEART DEVELOPMENT AND FUNCTION CORRELATE BASED ON THEIR FOLD CHANGE PATTERNS ACROSS ALL SAMPLES.	98
FIGURE 4.8: PTEN IS AN ENRICHED TARGET OF MIRNAS KNOWN TO AFFECT HEART DEVELOPMENT AND FUNCTION, AS WELL AS OF THEIR CORRELATED KNOWN ANIMAL MIRNAS.	100
FIGURE 4.9: NORMALISED READS FOLD CHANGE OVER LEVELS RECORDED AT HH5-7 FOR MIRNAS THAT CORRELATE WELL WITH SELECTED MIRNAS WITH KNOWN ROLES IN HEART DEVELOPMENT AND FUNCTION.	102
FIGURE 4.10: MIR-100, LET-7, MIR-125 AND MIR-99 ARE CLUSTERED IN THE CHICKEN GENOME.	104
FIGURE 4.11: THE PREDICTED TARGET SITES IN THE 3' UTRS OF THREE GENES WERE SHOWN TO NOT BE TRUE TARGET SITES BY LUCIFERASE REPORTER ASSAYS.	107
FIGURE 4.12: THE EFFECTS OF TWO MIRNAS, MIR-138 AND MIR-499, ON LUCIFERASE REPORTER CONSTRUCT CONTAINING THE ATG7 3' UTR.	108
FIGURE 4.13: REPORTER ASSAY RESULTS OF ELEVEN GENES FOUND TO CONTAIN MIRNA TARGET SITES IN THEIR 3' UTRS.	111
FIGURE 4.14: THE 3' UTR OF ETS1 HAS A SECOND PREDICTED TARGET SITE FOR MIR-499.	111
FIGURE 4.15: NORTHERN BLOT ANALYSIS OF MIR-130 EXPRESSION IN THE DEVELOPING CHICK HEART.	112
FIGURE 4.16: MIR-499 IS IMPLICATED TO REGULATE MIR-126 BY TARGETING ETS1.	116
FIGURE 4.17: THE ATTENUATION OF PTEN, A NEGATIVE REGULATOR OF GROWTH AND ANGIOGENESIS, BY MIRNAS.	120
FIGURE 5.1: NOVEL CHICK MIRNAS IDENTIFIED FROM SEQUENCING DATA.	132
FIGURE 5.2: ASSESSMENT TO DETERMINE WHETHER NOVEL MIRNAS HAVE BEEN ANNOTATED IN MIRBASE.	132
FIGURE 5.3: GENOMIC LOCATION OF PUTATIVE 12.	133
FIGURE 5.4: GENOMIC LOCATION OF PUTATIVE 14.	133
FIGURE 5.5: GENOMIC LOCATION OF PUTATIVE 10.	134
FIGURE 5.6: EXPRESSION OF THREE PREDICTED NOVEL MIRNAS IN NINE CHICK EMBRYO NGS SAMPLES.	135
FIGURE 6.1: CHICKEN MUSCLE-SPECIFIC MICRORNAS.	141
FIGURE 6.2: THE FOLD CHANGE PATTERN OF MIR-1 OVER ITS LEVELS IN THE HH5-7 SAMPLE.	145
FIGURE 6.3: FZD7 EXPRESSION IN THE CHICK DURING LATER STAGES OF DEVELOPMENT.	153
FIGURE 6.4: RARB EXPRESSION IN THE CHICK DURING EARLY STAGES OF DEVELOPMENT.	154
FIGURE 6.5: MIR-1 TRANSLATIONALLY INHIBITS FZD7 AND RARB IN THE DEVELOPING CHICK HEART.	155

FIGURE 6.6 INHIBITION OF MIR-1 BY AM1 AT HH14 RESULTED IN SIGNIFICANT DEVELOPMENTAL DELAY NOT SEEN UPON INJECTION WITH A CONTROL ANTAGOMIR (SCR206) NOR FOLLOWING INHIBITION OF MIR-206 BY AM206.	157
FIGURE APX A.1: CORRELATION OF SAMPLES BY KNOWN ANIMAL MIRNA NORMALISED READS.	181
FIGURE APX A.2: CORRELATION OF SAMPLES BY KNOWN ANIMAL MIRNA FOLD CHANGE OVER HH5-7.	185
FIGURE APX D.1: FZD7 EXPRESSION IN THE CHICK DURING EARLY STAGES OF DEVELOPMENT.	214
FIGURE APX D.2: AM1 AND SCR206 INJECTIONS INTO THE HH11 AND HH20 HEART <i>IN OVO</i> DID NOT RESULT IN A DEVELOPMENTAL DELAY AS OBSERVED WHEN MIR-1 WAS INHIBITED AT HH14.	215
FIGURE APX D.3: AM1 CARDIAC MICROINJECTIONS AT HH20 DID NOT RESULT IN DEVELOPMENTAL DELAY.	216
FIGURE APX D.4: SCR206 MICROINJECTIONS AT HH20.	216
FIGURE APX D.5: ANTAGOMIR-1 DISTRIBUTION <i>IN OVO</i> .	217
FIGURE APX D.6: DEVELOPMENTAL DELAY WAS OBSERVED FOLLOWING INHIBITION OF MIR-1 AT HH14.	218
FIGURE APX D.7: WIDE DISTRIBUTION OF THE CONTROL ANTAGOMIR, SCR206, WAS SEEN AFTER <i>IN OVO</i> MICROINJECTION INTO HH14 CHICK HEARTS.	218
FIGURE APX D.8: SCR206 INJECTED EMBRYOS DID NOT DISPLAY DEVELOPMENTAL DELAY.	219
FIGURE APX D.9: MICROINJECTION WITH FITC LABELLED ANTAGOMIR-206 (AM206) INTO HH14 HEARTS <i>IN OVO</i> RESULTED IN A DISTRIBUTION OF THE ANTAGOMIR THROUGHOUT THE EMBRYOS.	219
FIGURE APX D.10: INHIBITION OF MIR-206 AT HH14 DID NOT RESULT IN DEVELOPMENTAL DELAY.	220

## Table of Tables

TABLE 2.1: A SUMMARY OF THE NATURE OF THE 3' UTR SEQUENCES CLONED INTO MIRNA SENSING LUCIFERASE CONSTRUCTS.	34
TABLE 3.1: A SUMMARY OF THE EVENTS TAKING PLACE DURING HEART DEVELOPMENT BETWEEN STAGE 5 (HH5) AND STAGE 23 (HH23) IN THE CHICKEN EMBRYO.	48
TABLE 3.2: SUMMARY OF THE CAPACITIES OF THE MOST WIDELY USED SEQUENCING PLATFORMS.	50
TABLE 3.3: BREAKDOWN OF THE TOTAL NUMBER OF DISTINCT MIRNAS MATCHED TO ANIMAL MIRNAS IN MIRBASE RELEASE 20 BY SAMPLE.	59
TABLE 3.4: COMPARISON OF PREDICTED TARGET GENES OF TOP EXPRESSED MIRNAS FOR SAMPLES HH5-7, HH8-10 AND HH11-13.	66
TABLE 3.5: COMPARISON OF PREDICTED TARGET GENES OF TOP EXPRESSED MIRNAS FOR SAMPLES HH14-16 HEART, HH17-20 HEART AND HH21-23 HEART.	67
TABLE 3.6: COMPARISON OF PREDICTED TARGET GENES OF TOP EXPRESSED MIRNAS FOR SAMPLES HH14-16 NT, HH17-20 NT AND HH21-23 NT.	67
TABLE 3.7: SIGNIFICANT GO TERMS UNIQUE TO THREE GROUPS OF SAMPLES	70
TABLE 3.8: OVERLAPPING PREDICTED TARGET GENES, KNOWN TO BE EXPRESSED IN THE HEART, OF AT LEAST TWO OF MIR-26, MIR-30, AND MIR-92 ARE SIGNIFICANTLY ASSOCIATED WITH DISEASE OF THE HEART, AMONGST OTHER DISEASES.	77
TABLE 4.1: A FEW MIRNAS HAVE LEVELS MORE THAN 100 TIMES INCREASED OVER THEIR LEVELS AT HH5-7	87
TABLE 4.2: BREAKDOWN OF THE PROFILE GENERATED WHEN SUBTRACTING MIRNA FOLD CHANGE VALUES IN NT FROM THOSE OF THE CORRESPONDING HEART SAMPLES.	89
TABLE 4.3: BREAKDOWN OF NUMBERS OF MIRNAS THAT CORRELATE TO 26 KNOWN HEART MIRNAS BASED ON A H – NT FOLD CHANGE PROFILE ESTABLISHED FOR THE HH14-16, HH17-20 AND HH21-23 SAMPLES.	90
TABLE 4.4: 26 MIRNAS IDENTIFIED FROM A LITERATURE SEARCH HAVE ROLES IN HEART DEVELOPMENT AND FUNCTION.	92
TABLE 4.5: THIRTEEN MIRNAS WERE FOUND TO HAVE SLIGHTLY INCREASED READS IN THE HH14-16 NT SAMPLE OVER THAT DETECTED IN THE HH14-16 H SAMPLE.	95
TABLE 4.6: ELEVEN MIRNAS WERE SELECTED FOR FURTHER INVESTIGATION BASED ON A LIST OF 26 MIRNAS KNOWN TO HAVE ROLES IN HEART DEVELOPMENT AND FUNCTION.	96
TABLE 4.7: SIGNIFICANTLY ENRICHED COMMON GO TERMS OF THE PREDICTED AND CONFIRMED TARGET GENES OF MIR-101, MIR-190 AND MIR-19.	103
TABLE 4.8: FEW OF THE PREDICTED OR CONFIRMED TARGET GENES OF MIR-130 ARE ALSO PREDICTED OR CONFIRMED TARGETS OF MIR-10.	105

TABLE 4.9: MANY MIRNAS WITH KNOWN ROLES IN HEART DEVELOPMENT AND FUNCTION HAVE CONSERVED PREDICTED TARGET SITES IN THE 3' UTR OF <i>QKI</i> .	127
TABLE 5.1: THE GENOMIC LOCATIONS AND MATURE SEQUENCES OF THE FOURTEEN POTENTIAL MIRNAS PREDICTED BY SIMON MOXON USING MIRCAT <sup>42</sup> .	131
TABLE 5.2: THE SEED 6-MER SEQUENCES OF PUTATIVE MIRNAS USED TO CHECK FOR TARGETS IN ALL CHICKEN 3' UTRS.	134
TABLE 5.3: MANY POSSIBLE TARGET SITES IN A LARGE NUMBER OF CHICKEN GENES WITH HUMAN ORTHOLOGUES WERE IDENTIFIED FOR THE THREE PREDICTED MIRNAS CONFIRMED IN CHICK.	135
TABLE 5.4: SIGNIFICANT GO TERMS ASSOCIATED WITH THE PREDICTED TARGET GENES OF THREE PUTATIVE MIRNAS CONFIRMED IN THE DEVELOPING CHICK.	136
TABLE APX A.1: ANIMAL MIRNAS SEQUENCED FROM CHICKEN TISSUE BUT NOT YET FORMALLY ANNOTATED AS CHICKEN MIRNAS IN MIRBASE.	180
TABLE APX A.2: MIRNAS DETECTED IN A SINGLE SAMPLE ONLY.	180
TABLE APX A.3: OVERLAPPING PREDICTED TARGETS OF MIR-26, MIR-30, MIR-92 AND MIR-10 THAT ARE KNOWN TO BE EXPRESSED IN THE HUMAN HEART.	182
TABLE APX A.4: OVERLAPPING PREDICTED TARGETS OF MIR-26, MIR-30, MIR-92 AND MIR-10 THAT ARE KNOWN TO BE EXPRESSED IN THE HUMAN HEART.	183
TABLE APX A.5: DISEASES SIGNIFICANTLY ASSOCIATED WITH PREDICTED OVERLAPPING TARGET GENES OF MIR-26, MIR-30, AND MIR-92 THAT ARE ALSO KNOWN TO BE EXPRESSED IN THE HEART.	184
TABLE APX A.6: MIRNAS FOUND TO HAVE FOLD CHANGE LEVELS GREATER THAN 10 FOLD HIGHER AFTER SUBTRACTING MIRNA READS DETECTED IN NT SAMPLES FROM THAT DETECTED IN HEART SAMPLES.	186
TABLE APX A.7: MIRNAS FOUND TO HAVE SIGNIFICANTLY SIMILAR PATTERNS OF EXPRESSION BY FOLD CHANGE ACROSS THE HH14-16, HH17-20 AND HH21-23 TIME POINTS, AFTER SUBTRACTING MIRNA READS DETECTED IN NT SAMPLES FROM THAT DETECTED IN H SAMPLES.	187
TABLE APX A.8: MIRNAS FOUND TO HAVE SIGNIFICANTLY SIMILAR PATTERNS OF EXPRESSION BY FOLD CHANGE ACROSS THE HH14-16, HH17-20 AND HH21-23 TIME POINTS, AFTER SUBTRACTING MIRNA READS DETECTED IN NT SAMPLES FROM THAT DETECTED IN H SAMPLES.	188
TABLE APX A.9: MIRNAS FOUND TO HAVE SIGNIFICANTLY SIMILAR PATTERNS OF EXPRESSION BY FOLD CHANGE ACROSS THE HH14-16, HH17-20 AND HH21-23 TIME POINTS, AFTER SUBTRACTING MIRNA READS DETECTED IN NT SAMPLES FROM THAT DETECTED IN H SAMPLES.	189
TABLE APX A.10: CONSERVED MIRNAS FOUND TO CORRELATE SIGNIFICANTLY TO 11 KNOWN HEART MIRNAS.	190
TABLE APX A.11: CONSERVED MIRNAS FOUND TO CORRELATE SIGNIFICANTLY TO 11 KNOWN HEART MIRNAS.	191



## Acknowledgements

**Prof. Andrea Münsterberg**, for her constant guidance and support, as well as her exemplary leadership in all aspects of scientific research.

**Prof. Tamas Dalmay** for sharing his expertise in miRNA biology and sequencing methodologies.

**Dr Kasia Goljanek-Whysall**, who was always willing to listen and help plan future work and whose assistance and support in the laboratory was invaluable.

**Dr Simon Moxon**, who performed some of the bioinformatic analyses contained in this thesis, for always being approachable and willing and able to give advice.

**Dr Tim Grocott**, whose love of research makes him an excellent source of inspiration in the laboratory.

**Prof. Ian Clark** for guidance through the transfer process.

**Dr Natalie Crowe, Linh Le, and Dr Tracey Swingler** for lots of help with R and with luciferase assays.

**Dr Alistair McCormick**, for many scientific discussions and fantastic ideas as well as general patience with an often grumpy wife.

**Dr Grant Wheeler and Dr Mohammad Hajhosseini** for many helpful suggestions during lab presentations.

**Dr Guy Wheeler and Dr Ping Xu** for lots of help with northern blot assays.

**Past and current members of the Münsterberg and Wheeler laboratories:**

Dr Niki Kennerley and Kim Hanson for chats over many cups of tea.

Dr Vicky Hatch and Dr Adam Hendry, who never objected to incessant conversations on interesting science fiction movies and books.

Andy Loveday, for always knowing where everything in BIO and the BMRC is.

Amy Lyall, for not batting an eye at weekly demands for competent cells.

Camille Viaut, Dr James McColl, Dr Geoff Mok, Abdul Alrefaei, Dr Esther Camp-Dotlic, Dr Amanda Barber, Marlene Jahnke, Ayisha Ahmed, Chris Ford and Dr Connie Ochmann for being fantastic team mates during these three years.

**Dr Deborah Gill and Dr Stephanie Sumner-Jones**, without whom I would never have considered studying towards a PhD.

My dad, whose interest in and enthusiasm for all things science sparked my own.

## Abbreviations

$\alpha$ MHC	$\alpha$ Myosin heavy chain
AM1	AntagomiR-1
AM206	AntagomiR-206
AP1	Activating Protein 1
APS	Ammonium Persulphate
ATG7	Autophagy related 7
ATP	Adenosine Triphosphate
$\beta$ MHC	$\beta$ Myosin heavy chain
BBR	Boehringer Blocking Reagent
BCIP	5-Bromo-4-Chloro-3'-Indolyphosphate p-Toluidine Salt
bHLH	Beta helix loop helix
BMP	Bone Morphogenic Factor
bp	base pair
BSA	Bovine Serum Albumin
cDNA	Complementary DNA
<i>C. elegans</i>	<i>Caenorhabditis elegans</i>
CAM-1	Cell Adhesion Molecule 1
CHAPS	3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate
CPCs	Cardiac progenitor cells
Cre	Causes recombination
CNC	Cardiac Neural Crest
Cx43	Connexin 43
CYP26A1	Cytochrome P450 26 A1
<i>D. melanogaster</i>	<i>Drosophila melanogaster</i>
DCM	Dilated cardiomyopathy
DEPC	Diethyl Bicarbonate
DIG	Digoxigenin
DKK1	Dickkopf-related protein 1
DMEM	Dulbecco's Modified Eagle Medium

DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide
DTT	Dithiothreitol
Dvl1	Disheveled 1
E	Embryonic day
<i>E.coli</i>	<i>Escherichia coli</i>
EDC	1-ethyl-3-(3dimethyl laminoopropyl) carbodiimide
EDTA	Trypsin-ethylenediaminetetraacetic acid
Egfl7	Epidermal Growth Factor like domain 7
emPCR	Emulsion polymerase chain reaction
EMT	Epithelial to mesenchymal transition
ENAH	Enabled homolog (Drosophila)
Epo	Erythropoietin
ERG	Ether-a-go-go-related gene
ESC	Embryonic stem cell
EtOH	Ethanol
ETS1	v-ets avian erythroblastosis virus E26 oncogene homolog 1
EVH1	Enabled/vasodilator-stimulated phosphoprotein homolog 1
FBS	Foetal Bovine Serum
FGF	Fibroblast Growth Factor
FITC	Fluorescein isothiocyanate
FZD7	Frizzled 7
FZD8	Frizzled 8
<i>G. gallus</i>	<i>Gallus gallus</i>
Gga	Gallus gallus
GATA2	GATA-binding protein 2
Gb	Gigabases
GFP	Green Fluorescent Protein
H <sub>2</sub> O	Water
H <sub>2</sub> O <sub>2</sub>	Hydrogen peroxide
HAND2	Heart- and neural crest derivatives-expressed protein 2
HDAC4	Histone deacetylase 4

HH	Hamilton Hamburger
HRP	Horseradish peroxidase
Igf2	Insulin-like growth factor 2
IPTG	Isopropyl $\beta$ -D-1-thiogalactopyranoside
Irx5	Iroquois-class homeodomain protein
ISH	<i>In situ</i> hybridization
Isl1	Insulin gene enhancer protein
Jnk	c-Jun NH2-terminal kinase
kb	Kilo base pair
Kcnd2	Potassium voltage-gated channel subfamily D member 2
KCNQ1	Potassium voltage-gated channel, KQT-like subfamily, member 1
Kir2.1	Inward-rectifier potassium ion channel
LEF1	Lymphoid enhancer binding factor 1
let-7	Lethal-7
LNA	Locked Nucleic Acid
LVAD	Left ventricular assist device
MAPK	Mitogen activated protein kinase
MCS	Multiple Cloning Site
Mef2	Myocyte enhancing factor 2
MeOH	Methanol
MI	Myocardial infarction
miR	microRNA
miRNA	microRNA
miRNP	miRNA-containing Ribonucleoprotein complex
MRF	Myogenic Regulatory Factor
mRNA	Messenger RNA
Myh6	Myosin heavy chain 6
Myh7	Myosin heavy chain 7
Myh15	Myosin heavy chain
NaOAc	Sodium acetate
NBT	Nitro-Blue Tetrazolium Chloride

NEB	New England Biolabs
NFAT	Nuclear factor of activated T-cells
NFIB	Nuclear factor I B
NR3C1	Nuclear receptor subfamily 3 group C member 1
NGS	Next Generation Sequencing
nPTB	Polypyrimidine tract-binding protein
nt	Nucleotide
°C	Degrees Celsius
PAGE	Polyacrylamide Gel Electrophoresis
PBS	Phosphate Buffered Saline
PBST	Phosphate Buffered Saline Tween-20
PCR	Polymerase Chain Reaction
PDGFRA	Platelet-derived growth factor receptor $\alpha$
PFA	Paraformaldehyde
PI3K	Phosphatidylinositide 3-kinase
PIK3R2	PI3-kinase regulatory subunit- $\beta$
piRNA	Piwi-interacting RNA
PIWI	P-element induced Wimpy testis
PNK	Polynucleotide Kinase
PPP3CB	Calcineurin A, $\beta$ isozyme
pre-miRNA	Precursor miRNA
pri-miRNA	Primary miRNA
PRKACB	cAMP-dependent protein kinase catalytic subunit beta of PKA
PTEN	Phosphatase and tensin homolog
QKI	QKI KH domain containing RNA binding protein
qPCR	Quantitative polymerase chain reaction
RA	Retinoic acid
Raldh2	Retinaldehyde dehydrogenase 2
Rarg	Retinoic acid receptor gamma
RIPA	Radio Immuno Precipitation Assay
RISC	RNA-induced silencing complex

RNA	Ribonucleic acid
ROCK2	Rho-associated coiled-coil containing protein kinase 2
RT	Room Temperature
RT-PCR	Reverse Transcription PCR
RXR	Retinoid X receptor
Scr206	Scrambled 206
SDS	Sodium Dodecyl Sulphate
sFRPs	Secreted Frizzled-related proteins
SOX9	Sex determining region Y box 9
SPRED1	Sprouty-related protein EVH1 domain containing 1
SRF	Serum Response Factor
SSC	Sodium Chloride/Sodium Citrate
SV40	simian virus 40
TBE	Tris/Borate/EDTA
TBS	Tris Buffered Saline
TBST	Tris Buffered Saline Tween-20
TCF12	Transcription factor 12
TEMED	N, N, N, N'- Tetra-methyl-ethylenediamine
TGF $\beta$	Transforming Growth Factor $\beta$
T <sub>M</sub>	Melting temperature
UTR	Untranslated region
VANGL2	VANGL planar cell polarity protein 2
VEGF	Vascular Endothelial Growth Factor
VSD	Ventricular septal defects
VSMCs	Vascular Smooth Muscle Cells
WISH	Wholemound <i>in situ</i> hybridisation
Wnt	Wingless-type MMTV integration site
WT1	Wilms tumor 1 transcription factor
X-gal	5-bromo-4-chloro-3-indolyl-b-D-galactopyranoside
<i>X.laevis</i>	<i>Xenopus laevis</i>
ZEB2	Zinc Finger E-box binding homeobox 2

## **Statement**

This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with the author and that use of any information derived there from must be in accordance with current UK Copyright Law. In addition, any quotation or extract must include full attribution.

# 1 Introduction

## 1.1 Chick heart development

During vertebrate development, the heart is the first organ to form and function<sup>1</sup>. An estimated 1% of the worldwide population is born with congenital defects to the heart, many of whom will require treatment during their entire lifespan<sup>2,3</sup>. Cardiovascular disease is a very frequent cause of adult morbidity and mortality in humans, affecting even those patients not born with known heart abnormalities<sup>2</sup>. The study of heart development is essential to further our understanding of, and leading to prevention of, the factors influencing diseases of the heart. The very early heart is a simple contractile tube composed of striated muscle, called myocardium, which pumps oxygen and nutrients through the developing embryo. Mammalian and chicken embryos share many of the basic mechanisms of early heart development<sup>9</sup> and, like mammals, the fully developed chicken heart has four chambers<sup>1</sup>. Along with their relatively short developmental period, the ease with which fertilized eggs can be obtained and the many molecular techniques that can be used to study heart development *in ovo*, this similarity makes the use of *G. gallus* as a model organism for heart development very attractive. A summary of incubation times that correspond to chick developmental stages, the Hamburger-Hamilton stages of development<sup>10</sup>, can be found in **Table Apx E.1**. This section aims to give a brief overview of some of the major events during early chick heart development that were examined for miRNA expression by Next Generation Sequencing.

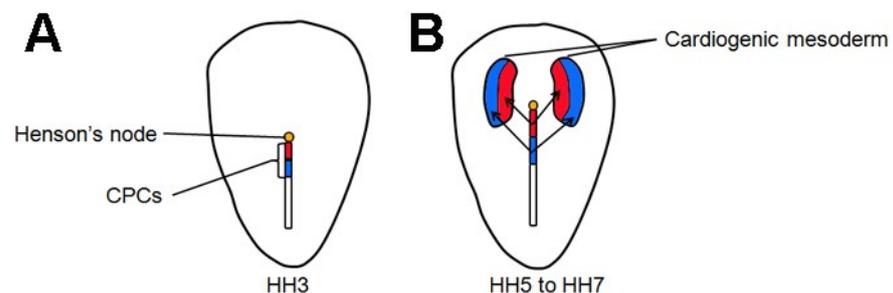
The first microRNA genes identified, *lin-4* and *let-7*, were discovered during studies into the effect of loss-of-function mutations which caused notable defects in developmental timing in *Caenorhabditis elegans*<sup>11-13</sup>. The depletion of the miRNA-processing enzyme Dicer in mouse embryos resulted in the arrest of development during gastrulation, demonstrating that microRNA function is vital for successful early vertebrate development<sup>14,15</sup>. The importance of microRNAs during development has further been indicated by studies showing that these molecules have essential roles during *C. elegans*

and *D. melanogaster* development, both as on-off switches and as fine-tuners of gene expression<sup>16</sup>. A study performed in chick embryos by examining the expression patterns of microRNAs by locked nucleic acid (LNA) in situ hybridization (ISH) found 75 microRNAs to be expressed consistently in specific regions and/or cell layers, confirming the extent of their role during chick development<sup>17</sup>.

Given the wide-ranging effects of miRNAs during development, this study aimed to discover miRNAs that are involved in the early stages of heart development by determining a profile of miRNA expression patterns during these early stages in the heart and surrounding tissues, and testing the effects of a selected few miRNAs.

### 1.1.1 Up to HH4

The epiblast and the hypoblast are the two cell layers that chick embryos consist of before the start of gastrulation. Cardiac progenitors have been identified in the epiblast prior to gastrulation, and the first embryonic cells to gastrulate include these progenitors<sup>18</sup>. After undergoing an epithelial-to-mesenchymal transformation, cells in the epiblast enter the primitive streak and form the mesodermal layer<sup>19</sup>. By HH3 cardiac progenitor cells (CPCs) can be found in the rostral half of the primitive streak, from an area caudal to Hensen's node, with at least two discrete subpopulations present that will contribute to either the cardiomyocyte or the endocardial cell lineage (**Figure 1.1A**)<sup>18</sup>.



**Figure 1.1:** A model of a stage HH3 and a HH5-7 chick embryo, showing the position of the cardiac precursor cells (CPCs) in the primitive streak (**A**) and indicating the path they follow when migrating out of the streak (**B**). Hensen's node is shown in yellow. The relative positions of cell populations that will give rise to the atria, atrioventricular canal and left ventricle (blue) and those that will form the right ventricle, conus and truncus (red) are indicated<sup>20</sup>.

Cells that will eventually form the anterior structures of the heart are located more rostrally (**Figure 1.1**, red) to those that will form the posterior structures (**Figure 1.1**, blue)<sup>21</sup>. CPCs have left the primitive streak by HH4, migrating to form the cardiogenic regions indicated in **Figure 1.1B**.

### 1.1.2 HH5 to HH7

The CPCs have migrated out of the primitive streak into the anterior mesoderm by HH4, and the bilateral primary heart fields are established in the splanchnic mesoderm layer located directly above the endoderm by HH5 – HH7<sup>18,19</sup>. The primary heart field stretches from slightly above Hensen's node to about a quarter of the length of the primitive streak and laterally almost to the edge of the area pellucida at HH4<sup>18,19</sup>. Although specified, CPCs are not yet completely committed to a cardiac fate; it has been shown that cells still present in the caudal primitive streak at HH5 can form heart tissue if left in contact with the anterior endoderm, indicating that cardiac inducing factors arise from the anterior endoderm at this stage<sup>18</sup>. Once the CPCs have migrated out of the primitive streak and formed the primary heart fields at HH5 (**Figure 1.1B**), the transcription factor *Nkx2.5* is expressed in cells that have a ventricular fate<sup>21</sup>.

Between HH6 and HH7, concurrent with neurulation and the formation of the first three somites, CPCs become increasingly committed to their fate and by HH7 terminal myocardial differentiation markers can be detected in the primary heart fields<sup>18</sup>. Cardiac tissue development is strongly inhibited by anterior neural tissue and more weakly inhibited by posterior neural tissue. Between HH5 and HH8 small groups of cells delaminate from the splanchnic mesoderm and become situated between the developing myocardium and the endoderm; these presumptive endocardial cells then migrate to form the endocardial heart tubes<sup>18,19</sup>.

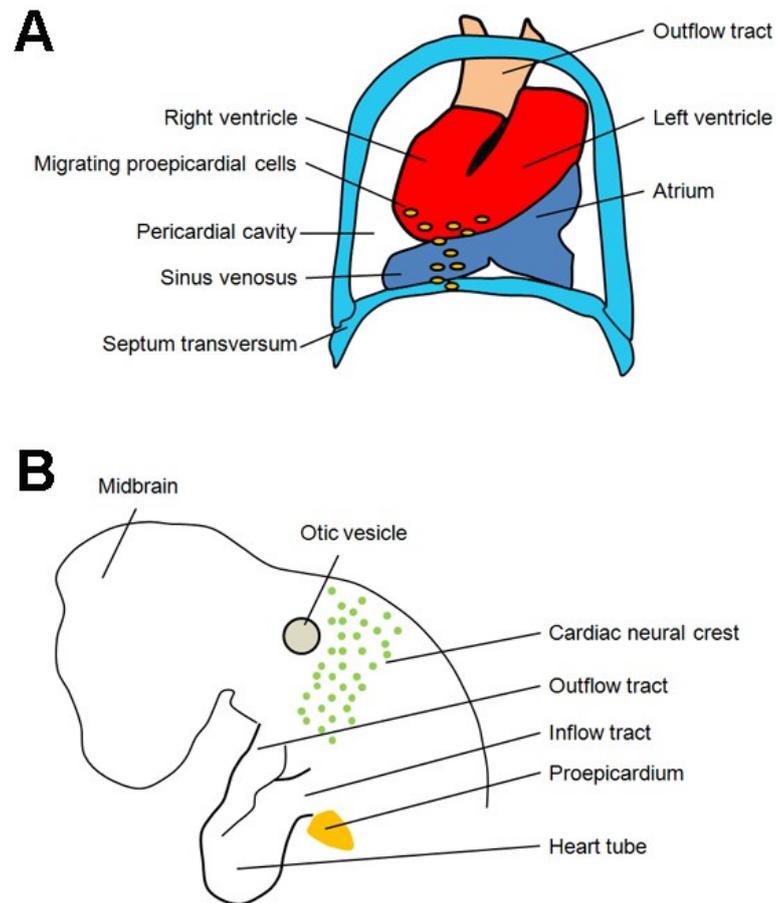
### 1.1.3 HH8 to HH10

Between HH7 and HH8 the primary heart field extends up to the level of the first somite and by HH8 it is level with the fourth somite<sup>18</sup>. At HH8, just before the bilateral heart

fields fuse, pre-migratory cardiac neural crest (CNC) cells are located in the most dorsal part of the embryo, between the mid-otic placode and level with the primary heart field at the caudal limit of somite 3<sup>18</sup>. The cardiogenic mesoderm, up to now bilateral, starts the process of fusing at the midline during HH8 and is fused at the midline by HH9, whilst at the same time cardiomyocytes start the process of differentiation<sup>18,19,22</sup>. CNC cells start to migrate from the neural crest towards the outflow tract aorticopulmonary (AP) septum and towards the pharyngeal arch arteries at HH9 and continue their migration during HH10<sup>18</sup>. Outflow tract septation depends on these migrating CNC cells. By HH10 both the endocardial tubes and the myocardium have fused, and the heart tube begins to beat despite a portion of the outflow tract close to the ventral pharynx still being open<sup>18</sup>.

#### **1.1.4 HH11 to HH13**

The heart tube continues to lengthen and starts to loop to the right until HH12 as a result of inward migration, via the inflow tract (**Figure 1.2A**, darker blue) and the outflow tract (**Figure 1.2A**, orange), of newly differentiated myocardium<sup>18,19,22</sup>. Once myocardial differentiation starts, progenitor cells stop proliferating for some time and therefore growth of the heart at this stage results mostly from this inward migration despite the fast cycling time of myocardial progenitor cells<sup>22</sup>. At HH12 the late migration of CNC cells into the pharyngeal arch arteries starts, eventually resulting in the remodelling of the pharyngeal arch artery<sup>18</sup>.



**Figure 1.2: Adapted schematic drawings of HH14 to HH17 chick embryos.**<sup>18</sup> **A:** Cardiac progenitors add to the looping heart by migrating in through the inflow (darker blue) and outflow (orange) tracts. The migrating proepicardial cells, shown in yellow, originate from the proepicardial organ in the splanchnic mesoderm and migrate over the myocardium and will eventually cover the heart. **B:** This lateral drawing of an HH17 chick embryo shows cardiac neural crest cells (green) migrating through the branchial arches 3, 4 and 6. The proepicardial organ is shown in yellow.

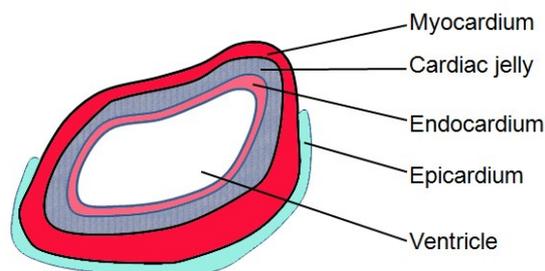
### 1.1.5 HH14 to HH16

Further cardiac looping, resulting in the c-shaped heart transforming into an s-shaped heart, starts around HH13 with the primitive atria being positioned closer to the outflow tract and the primitive ventricular bend moving to a more caudal position relative to the forming atria (**Figure 1.2A**)<sup>18</sup>. Between HH13 and HH15 the dorsal mesocardium, which had up to now attached the heart tube to the ventral pharynx, breaks and disappears making this movement of the heart tube possible<sup>18,22</sup>. By HH13 CNC no longer produces CNC cells; the late migrating CNC cells reach aortic arch 3<sup>18</sup>. At about HH14, once anterograde circulation has started, ventricular myocardial cells become proliferative again and as a

result the ventricles distend outward<sup>22</sup>. However, myocardial cells in the atrioventricular (AV) canal and in the outflow tract do not proliferate in a similar way<sup>22</sup>. Proepicardial cells migrate from the proepicardial organ, over the developing myocardium of the heart, to form the epicardium as well as smooth muscle cells (**Figure 1.2A, B**)<sup>18</sup>. The process of atrial septation starts at HH16<sup>18</sup>. Additional cardiac looping moves the sinus venosus from a position caudal to the atria, to one dorsal to the atria between HH14 – HH18<sup>18</sup>.

### 1.1.6 HH17 to HH20

By HH17 to HH18 the epicardium, myocardium, cardiac jelly and endocardium layers of the heart have formed (**Figure 1.3**)<sup>18</sup>. The myocardium consists of an outer layer, the highly mitotic compact myocardium, and an inner layer, the less mitotic trabeculated myocardium<sup>18</sup>. Research has shown that the compact myocardium is necessary for myocardial growth; this layer becomes slightly thicker between HH12 and HH21 but a second phase of proliferation between HH34 and HH40 increases the myocardial mass massively. The signalling molecules responsible for cardiomyocyte proliferation are thought to originate in the epicardium, but the patterning of the trabeculated myocardium is regulated by the endocardium<sup>18</sup>.



**Figure 1.3: Schematic drawing of a cross-section through the looping heart**, showing the four layers formed by HH17 – HH18: the myocardium (red), cardiac jelly (grey), endocardium (pink) and epicardium layers (blue)<sup>18</sup>.

The remodelling of the developing heart by apoptosis starts at HH19 and follows the path of epicardial coverage, in the tissue bridge between the myocardium and the proepicardial organ as well as in the outflow tract<sup>18</sup>.

### **1.1.7 HH21 to HH26**

By HH20 the migrating CNC cells reach the outflow tract, migration of myocardium into the outflow tract is finished and cardiac valve formation begins<sup>18</sup>. Cardiac septation and the late phase of cardiac looping begins at HH24, and apoptosis is still observed in the outflow tract, in the ventricles as well as in the atria<sup>18</sup>. Although a large number of changes still occur between this stage and adulthood, such as the completion of valve development, the formation of the coronary vasculature and the maturation of the cardiac conduction system, for the purposes of this study only these early stages of heart development will be considered in detail.

## **1.2 Congenital heart diseases**

Despite the frequency of congenital heart defects in the human population, affecting almost 1% of live births and resulting in the mortality or morbidity of many patients, the genetic cause of these defects remain largely unexplained. Current evidence suggests that these defects are caused by the improper formation of the heart and vessels as a result of a combination of factors<sup>23</sup>. A severe and common form of congenital heart disease arises due to a large ventricular septal defect (VSD) which is formed when the conal septum is misaligned during development<sup>24</sup>. This disease, Tetralogy of Fallot, occurs in between five to seven live human births per 10,000, and generally requires surgical intervention early on in life. miR-421 is upregulated in ventricular tissue from human infants with Tetralogy of Fallot, and negatively affects a key regulator of the Notch signalling pathway, SOX4. SOX4 is required for the formation of the outflow tract and therefore increased expression of miR-421 was implicated as one part of the combination of factors that lead to Tetralogy of Fallot.

## **1.3 Wnt signalling is essential for heart formation and development**

The highly conserved Wnt signalling pathway is central to wide-ranging processes in embryonic and adult stages of many organisms<sup>7,25,26</sup>. This signalling pathway is known to

function from the very start of cardiac development, and is essential for patterning the heart<sup>7,19,27,28</sup>. This project discovered miRNAs that regulate miRNA-sensing constructs containing the 3' UTRs of several components of the Wnt signalling pathway (**Chapter 4**), strongly suggesting that miRNAs have essential roles during early heart development by modulating such components.

The Wnt genes, more than 100 of which have been annotated thus far, regulate many developmental processes, determine cell fate, influence differentiation and cell migration, and establish polarity in tissues<sup>29,30</sup>. These genes are very widely conserved and have been identified in organisms ranging from *C. elegans* to *H. sapiens*. The proteins encoded by the Wnt genes all contain a signal sequence and a pattern of 23 cysteine residues that does not differ greatly between genes. Large families of genes encode the ligands and receptors that make up Wnt signalling, giving rise to a very large number of possible interactions. Current thinking suggests that all Wnt proteins are secreted and act by binding to receptors on the cell surface of the secreting cells or any cells with appropriate receptors nearby. In addition, associations of receptors and ligands can cause several different reactions and furthermore Wnt ligands can also act to induce other signalling pathways by activating intracellular messengers such as c-Jun NH2-terminal kinase (Jnk) and Src kinases<sup>30</sup>.

Four different responses to Wnt-receptor binding are described in literature, although not all have been completely characterised. Broadly, these four responses can be divided into three categories: binding of a Wnt ligand to a Frizzled (FZD) receptor, to the receptor tyrosine kinase (RTK) Ror or to a Ryk RTK. The RTK signalling pathways are incompletely understood. Canonical  $\beta$ -catenin / TCF signalling is inhibited when Wnt ligands bind Ror2, and Jnk can be activated as a result of this binding, although other interaction partners during this process have yet to be identified<sup>30</sup>. Wnt binding to Ryk RTK results in activation of Src proteins, and has been studied mostly during neuronal development<sup>30</sup>. The cellular response when Wnt binds to FZD can be subdivided into two categories, one in which an LRP co-receptor is present and another where possible co-receptors have not yet been identified. Wnt – FZD signalling is discussed in **1.3.1** and **1.3.2** below.

### 1.3.1 Wnt / $\beta$ -catenin pathway

Canonical Wnt signalling is defined as signalling that disrupts the establishment of the body axis in *Xenopus* embryos along with an increase in the levels of  $\beta$ -catenin and, as a result,  $\beta$ -catenin signalling. In this pathway Wnt protein signals on the extracellular surface are bound by the 7-transmembrane receptors in the Frizzled (FZD) family, along with their LRP co-receptors<sup>26</sup>. In the absence of Wnt a destruction complex associated with LRP, consisting of Axin and adenomatous polyposis coli (APC), GSK-3 and Casein kinase, binds  $\beta$ -catenin and phosphorylates it leading to the transfer of  $\beta$ -catenin to the proteasome and its destruction by the ubiquitin pathway. However, when Wnt binds to FZD the binding allows FZD to deactivate the signal transduction protein Disheveled (Dvl), thus preventing Dvl from inhibiting the action of glycogen synthase kinase-3 enzyme (GSK-3). The inactivation of GSK-3 prevents the degradation of  $\beta$ -catenin, allowing it to travel into the nucleus and activate transcription factors from the T-cell factor / lymphocyte enhancer factor (TCF / LEF) family, which leads to the activation of Wnt-responsive genes.

At the plasma membrane  $\beta$ -catenin,  $\alpha$ -catenin and cadherin proteins can form a complex that regulates cell adhesion by its effect on adherens junctions<sup>26</sup>. The canonical Wnt/ $\beta$ -catenin pathway described above is known to control cell proliferation by the activation of cell cycle genes such as c-myc and cyclinD1<sup>7</sup>.

### 1.3.2 Non-canonical Wnt pathways

The planar cell polarity (PCP) and Wnt /  $\text{Ca}^{2+}$  pathways are non-canonical signalling pathways and neither requires the involvement of  $\beta$ -catenin<sup>7</sup>. These pathways are known to inhibit the  $\beta$ -catenin dependent canonical Wnt signalling and in vertebrates, although not in flies, Wnt still binds FZD to initiate non-canonical signalling<sup>7</sup>. The binding of Wnt to a FZD receptor that is not associated with the co-receptor LRP results in Dvl being allowed to form a complex which activates the GTPases Rho and Rac, leading to the activation of Rho kinase and Jnk<sup>7</sup>. This in turn causes the activation of Rho-associated kinase (ROCK) and results in modifications to the cytoskeleton, or regulation of transcription by the activation of Activating Protein 1 (AP1) by Jnk. Furthermore, actin modifications can be

caused by Dvl and Rac1 association as a result of Wnt binding<sup>7</sup>. The PCP signalling pathway is known to control migration of neural crest cells<sup>31</sup>.

In some cases an increase in the intracellular concentration of Ca<sup>2+</sup> ions can result when the Wnt ligand binds its FZD receptor. In the Wnt/Ca<sup>2+</sup> pathway this release of Ca<sup>2+</sup> ions causes the activation of calcium-sensitive enzymes such as calcineurin (CaCN), which regulates Nuclear factor of activated T-cells (NFAT) and as a result, gene transcription.

### 1.3.3 Wnt signalling is context-specific

Attempts have been made in the past to classify Wnt ligands and FZD receptors as either canonical or non-canonical. For example, Wnt5A is known to be a non-canonical ligand as it did not induce signalling through  $\beta$ -catenin / TCF when exposed to different mammalian FZD receptor that were ectopically expressed in *D. melanogaster*<sup>30</sup>. However, when Wnt5A was recently added to 293 HEK cells that express both FZD4, its receptor, and LRP5, canonical signalling was induced. This suggests that the failure of Wnt5A to signal in a canonical manner in the early *D. melanogaster* experiments may have been a result of a failure of Wnt5A to recruit a LRP co-receptor. In addition, maternally contributed Wnt11 acts canonically to increase  $\beta$ -catenin levels, whereas Wnt11 is otherwise known to induce non-canonical signalling. Therefore it is possible that Wnt ligands are either canonical or non-canonical as a result of the signals and factors locally available.

Furthermore, the cellular responses to Wnt signalling are influenced by the nature of the receptor. The extent to which the receptor activates  $\beta$ -catenin signalling is known to depend on the C-terminus sequence of the FZD receptor. In *D. melanogaster* exchanging the C-terminal sequence of FZD and FZD2, that induce PCP and  $\beta$ -catenin / TCF signalling respectively, has led to a reversal of their signalling<sup>30</sup>.

Evidence has been found in *C. elegans* that Cell Adhesion Molecule 1 (CAM-1), a Ror2 homolog, interacts with several Wnt ligands and may act as a sponge to sequester Wnt molecules and therefore modulate the range and concentration of secreted Wnts, affecting

the balance of signals encountered by the cell and thus adding another point of control over the cellular response.

Finally, LRP and Wnt-FZD interactions are prevented by the combined association of the transmembrane receptor Kremen and the secreted protein Dkk<sup>30</sup>. When Dkk binds to LRP in the presence of Kremen, LRP is transferred into the cell. However, when Dkk is not present Kremen has the opposite effect, ensuring that LRP is located on the cell surface. The effect of these interactions is that Wnt /  $\beta$ -catenin signalling is promoted when Dkk is absent and Kremen is present, and inhibited in the presence of Dkk and Kremen.

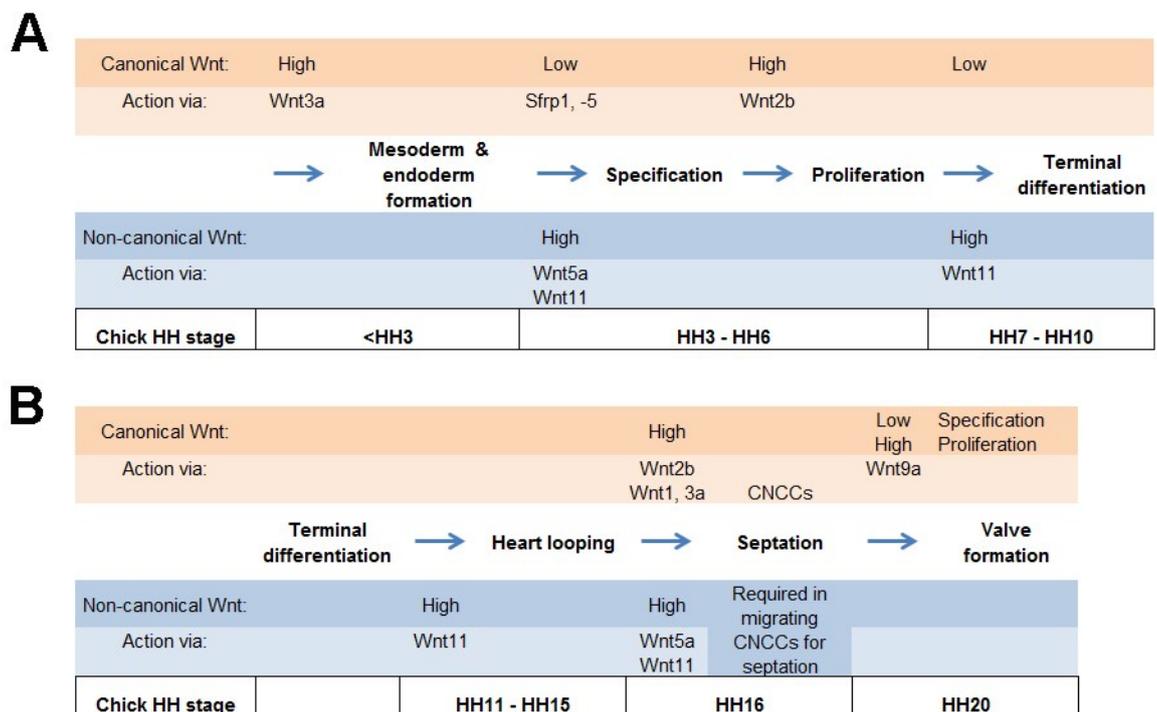
The factors influencing canonical and non-canonical Wnt signalling discussed above illustrate some of the complexity of this signalling pathway. Both canonical and non-canonical Wnt signalling influences heart development, and a description of signalling in the context of specific stages of heart development is discussed in **1.3.4**.

#### **1.3.4 Wnt signalling is essential for heart development**

Canonical and non-canonical Wnt signalling are both known to be essential at certain points during cardiogenesis<sup>7</sup>. The requirement for each type of Wnt signalling becomes more difficult to elucidate as heart development proceeds and the heart becomes more patterned. Broadly speaking, canonical Wnt signalling induces proliferation whereas non-canonical signalling controls differentiation in some tissues and EMT and migration in others (**Figure 1.4**). Canonical Wnt signalling via Wnt3a is high to induce mesoderm and endoderm formation prior to HH3 (**Figure 1.4A**). To induce specification of cardiac fate in cardiogenic mesoderm, non-canonical Wnt signalling is mediated through Wnt5a and Wnt11 from HH3 – HH6. Proliferation is induced by canonical Wnt2b once cardiogenic precursors have been specified, whereas non-canonical Wnt11 induces terminal differentiation from HH7 – HH10. Non-canonical signalling via Wnt11 is still required from HH11 to HH15 in the looping heart (**Figure 1.4B**). As the heart becomes more patterned, the roles of canonical and non-canonical Wnt signalling increasingly depend on the cell type investigated. Non-canonical signalling through Wnt5a and Wnt11 allows EMT and migration of cardiac neural crest cells that will partake in septation of the heart,

whereas canonical Wnt signalling induces cardiac neural crest cells to proliferate, through Wnt1 and Wnt3a. Canonical Wnt signalling (Wnt2b) is required for primary atrial septation, as the genetic knockout of Wnt2a in mice resulted in mutant hearts lacking atrial septation. Wnt9a allows canonical Wnt signalling to activate proliferation of endocardial cells.

The role of Wnt signalling in the development of the heart is discussed in more detail below (1.4).



**Figure 1.4: A summary of the currently known requirements for canonical Wnt and non-canonical Wnt signalling in the developing embryonic heart up to chick stage HH20<sup>7</sup>.** **A:** Canonical Wnt signalling is alternately high and low during mesoderm and endoderm formation, specification, proliferation and terminal differentiation. **B:** Non-canonical signalling leads to heart looping, whereas high canonical and high non-canonical Wnt signalling results in septation. Valve formation proceeds when canonical Wnt signalling is low.

## 1.4 Factors involved in chick heart development

The four major milestones of heart development are specification, determination, patterning and differentiation. The CPCs that migrate from the rostral half of the primitive

streak to the lateral mesoderm between HH3 and HH5 are specified to a cardiac fate and will continue on this path even if placed in a neutral environment<sup>18,19</sup>. A higher level of commitment to the cardiac lineage is achieved during determination, between HH5 and HH7, when the cardiogenic mesoderm cannot be induced to deviate from its fate even if placed in an antagonistic environment<sup>18,19,21</sup>. The bilateral heart fields resolve into regions with specific fates between HH7 and HH9 during the process of patterning, and the expression of molecular and genetic markers that characterise the specific cell type marks differentiation, which occurs between HH9 and HH11<sup>19</sup>.

#### 1.4.1 Specification

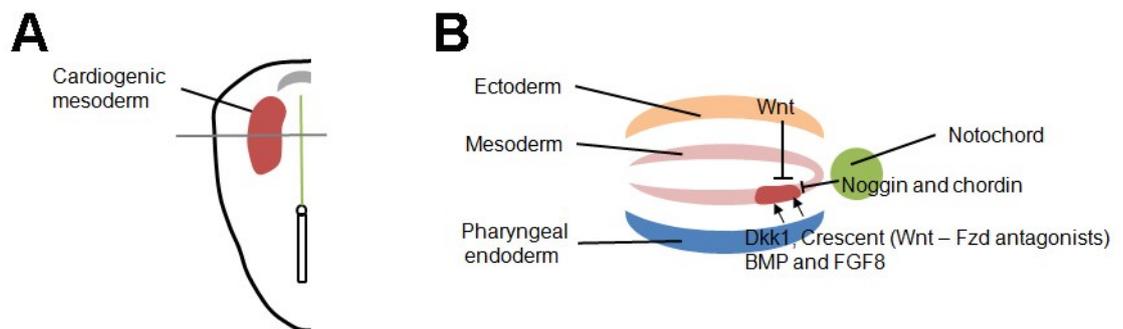
Cranio-lateral endoderm is the source of most of the factors that facilitate myocardial specification, including bone morphogenetic proteins (BMPs), fibroblast growth factors (FGFs) and transforming growth factor beta (TGF $\beta$ ) family members. Canonical Wnt signalling has been shown to prevent cardiac specification, whilst non-canonical Wnt and Sonic hedgehog (shh) signalling induces it<sup>19,27,28</sup>.

Components of the BMP signalling cascade are expressed together in the bilateral heart fields. In chick both the hypoblast and anterior endoderm have been found to express *Bmp2* and *Bmp4*, and the cardiac transcription factors *Nkx2.5* and *Gata4* are upregulated in response to *Bmp2*<sup>19</sup>. Furthermore, when BMP signalling is inhibited by the antagonist *noggin*, cardiogenic mesoderm is not formed. Together these observations strongly suggest that myocardial induction is dependent on BMP signalling.

In regions where BMP signalling is present, FGFs cause cardiac specification<sup>21</sup>. FGF8 is essential for right ventricle and outflow tract formation, and is expressed in the mesoderm during gastrulation, in the pharyngeal endoderm and cardiogenic mesoderm, and in the outflow myocardium following specification<sup>19,21</sup>. In the bilateral heart fields *Nkx2.5* and *myocyte enhancer factor 2C (Mef2C)* are expressed partly as a result of FGF8 expression in the endoderm.

A non-canonical Wnt ligand expressed in the endoderm and mesoderm, Wnt11, inhibits canonical Wnt signalling and has been shown to induce myocardial differentiation<sup>19,21</sup>. When the Wnt molecules in the canonical signalling pathway that bind to Frizzled receptors are competitively bound by Wnt-antagonists Dickkopf 1 (Dkk1) and Crescent, the canonical signalling pathway is inhibited and myocardial specification is promoted<sup>19</sup>. In addition, canonical signalling via Wnt3A and Wnt8 can prevent cardiomyogenesis, whereas their inhibition promotes cardiac specification<sup>19,21</sup>. Anterior neural ectoderm inhibits cardiac specification more strongly than posterior neural ectoderm by the expression of canonical Wnts<sup>18,19</sup>. Along with its function as a repressor of cardiac lineages, canonical Wnt signalling also induces endoderm formation<sup>21</sup>.

Although specification of a cardiac fate depends on various factors, the bilateral heart fields are characterized by low canonical Wnt signalling and high levels of BMPs, as a result of signals originating in the endoderm and mesoderm (**Figure 1.5**)<sup>21</sup>.



**Figure 1.5: Origins of cardiac inducing and inhibiting factors**<sup>21,32</sup>. **A:** Schematic drawing of an early stage chick embryo with the cardiogenic mesoderm indicated in red and the notochord in green. **B:** Embryo in **A** viewed cut through the plane indicated by the grey line. The ectoderm, shown in orange, secretes canonical Wnt signals that inhibit cardiogenesis. BMP antagonists noggin and chordin are expressed in the notochord (green) and also inhibit cardiogenesis in the adjacent mesoderm (pink). The pharyngeal endoderm (blue) induces the nearby splanchnic mesoderm (red) to form cardiogenic mesoderm through BMP and FGF8 factors as well as the antagonists of Wnt signalling, that act by binding Frizzled receptors, Dkk1 and Crescent<sup>21</sup>. Somatic mesoderm and endoderm express non-canonical Wnt11, inducing a cardiac fate in the target cells<sup>19,21,32</sup>.

Mesp1 and Mesp2, basic helix-loop-helix (bHLH) transcription factors, are involved in the migration of the CPCs from the primitive streak into the anterolateral mesoderm<sup>21</sup>.

### 1.4.2 Determination and heart field formation

Several transcription factors, including *Nkx2.5* and members of the GATA and myocyte enhancer factor 2 (MEF2) families, become activated in the bilateral heart field cells which results in their determination to a cardiac fate<sup>19</sup>.

In both mice and chicks the NK homeobox gene *Nkx2.5* is expressed very early after CPCs have migrated to form the bilateral heart fields, and is still expressed in adult hearts<sup>19,21</sup>. In mice, *Nkx2.5* knockout causes foetal mortality due to outflow and inflow tract malformations although cardiac differentiation is not completely abrogated, as evidenced by the presence of functional cardiomyocytes in the heart tube<sup>19</sup>. *Nkx2.5* regulates the expression of a large number of cardiac-specific genes and therefore is essential for cardiogenesis, and is expressed in cells at HH5 that eventually form the ventricles of the chick heart<sup>21</sup>. A Smad transcription factor binding site, and hence BMP signalling, is essential for cardiac-specific expression of *Nkx2.5*<sup>19</sup>. The Smad transcription factor is thought to act together with the GATA4 factor to bring about expression of *Nkx2.5* in precardiac cells.

Three members of the GATA zinc-finger transcription factor family, GATA4-6, are expressed in the developing vertebrate heart. The deletion of *GATA4* resulted in a reduction in the number of cardiomyocytes as well as cardia bifida in mice, whereas its overexpression induced cardiac differentiation in ESCs<sup>21</sup>. Just as *GATA4* expression induces *Nkx2.5* expression, *Nkx2.5* expression also upregulates the expression of *GATA4*<sup>21,33</sup>.

### 1.4.3 Patterning within the bilateral heart fields

Fate mapping experiments have traced the origins of cardiac cells but the earliest molecular and genetic patterning markers of myocardial progenitors are first expressed some time after gastrulation<sup>19</sup>. The bilateral heart fields consist of the primary heart field, the anterior heart field, the first heart field and the second heart field. The primary heart field differentiates before the anterior heart field, is defined by *Fgf8* / *Fgf10* expression and

forms the inflow limbs of the heart, while the anterior heart field does not express *Fgf8* / *Fgf10* and forms the outflow limbs<sup>19</sup>. The first and second heart fields do not correspond exactly with the primary and anterior heart fields and are characterised by *Islet1* (*Isl1*) expression. The first heart field forms the outflow limb and a part of the atria whereas the second heart field gives rise to the right and left ventricles<sup>19</sup>. Mutation of *Isl1* leads to disrupted cardiac looping<sup>34</sup>. Additional smaller fields are created within the heart fields by the expression of genes such as *Tbx1*, *Tbx5*, *Tbx20* as well as *eHand* and *dHand*<sup>19</sup>.

#### 1.4.4 Differentiation and heart tube formation

Cardiomyocyte differentiation markers sarcomeric myosin (MF20) and cardiac troponin-I are first expressed at stage HH8 to HH9, in the centre of the heart field and then in the heart tube, marking the start of cardiomyocyte differentiation<sup>19</sup>. Cardiomyocytes will differentiate even if the heart fields do not join, as evidenced by the formation of two beating hearts in *cardia bifida*. The formation of the foregut, a pocket created when the anterior endoderm folds, and the heart tube has been shown to depend on extraembryonic endoderm<sup>19</sup>. Heart tube formation has been shown to rely on the expression of the small GTPase RhoA, which is highly expressed in HH6-HH9 chick anterior lateral plate mesoderm, heart primordia and anterior intestinal portal<sup>19</sup>.

Although the MADS-box protein serum response factor (SRF) is not muscle-specific, it associates with the transcription factor myocardin to activate cardiac muscle-specific genes. The expression of *SRF* is regulated by *Nkx2.5*<sup>21</sup>, which therefore manages growth and differentiation in cardiac cells directed by the action of *SRF* and myocardin. Cardiac differentiation gene expression is controlled at HH8 by the coexpression of *SRF*, *GATA4* and *Nkx2.5*<sup>21</sup>.

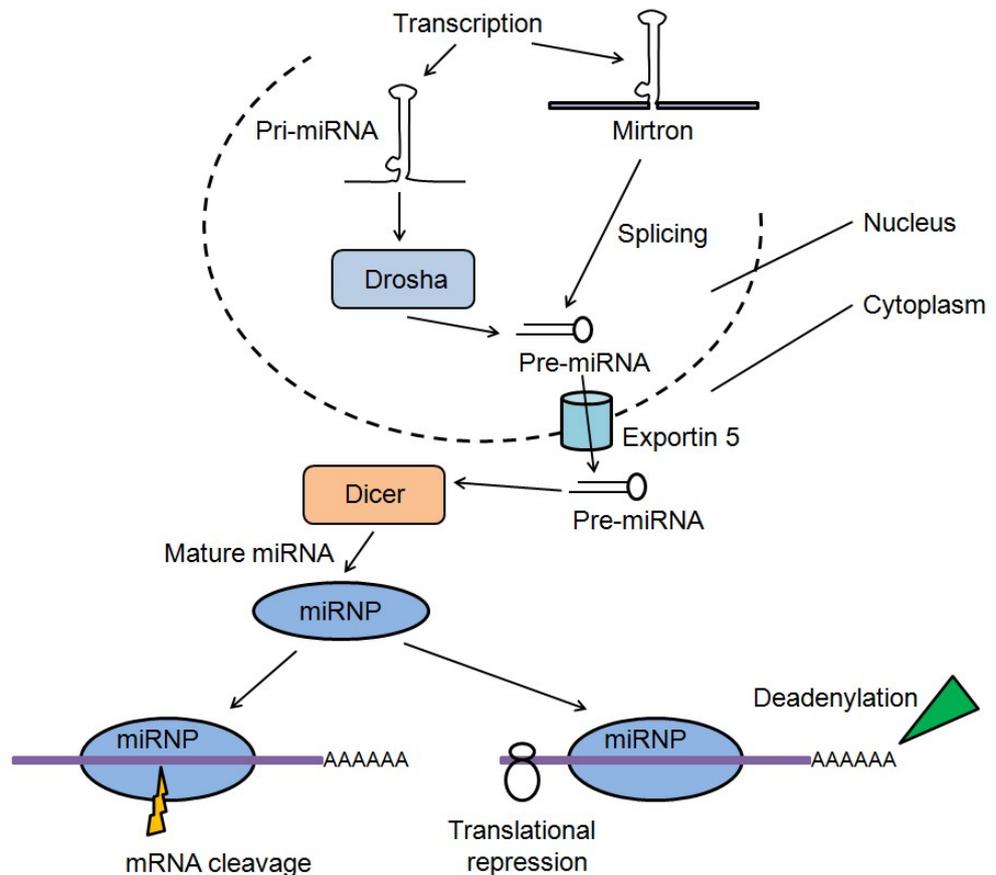
Given the pervasive expression of miRNAs during development, as well as their implication as essential regulators of many of the early processes of heart development, this work aimed to add to the current knowledge of miRNA expression during the early stages of heart development described above.

## 1.5 MicroRNAs

MicroRNAs (miRNAs) were first discovered during studies into the temporal regulation of development in *Caenorhabditis elegans* (*C. elegans*) and have since been identified to be highly expressed in many model organisms<sup>16</sup>. Although there are several types of small RNAs, including small interfering RNAs (siRNAs) and Piwi-interacting RNAs (piRNAs), only miRNAs are processed from hairpin structures that contain bulges where the strands are imperfectly paired (**Figure 1.6**)<sup>16</sup>. Despite not being translated into protein, mature ~20nt miRNAs control the expression of genes post-transcriptionally either by cleaving target mRNAs directly, by inducing degradation of the mRNA targets, or by slowing or stopping translation of mRNAs. A large number of all eukaryotic protein coding genes and most cellular processes are predicted to be controlled by miRNAs<sup>35</sup>. In mammals each conserved putative miRNA has an average of 300 conserved predicted targets and more than 50% of human 3' UTRs seem to have conserved miRNA binding sites<sup>16</sup>. With so many predicted targets, the elucidation of the biologically important functions of miRNAs is likely to prove essential for a more complete understanding of cellular function and development.

### 1.5.1 miRNA biogenesis

MicroRNA genes can be situated within introns of protein coding host genes or within intergenic regions. Once transcribed, primary microRNAs (pri-miRNAs) are processed inside the nucleus into ~70nt hairpin structures, called pre-miRNAs, by the RNase III endonuclease Drosha (**Figure 1.6**). Some intronic miRNAs do not get processed by Drosha but become pre-miRNAs during the splicing process undergone by the host protein coding gene<sup>36</sup>. Pre-miRNAs are transported into to the cytoplasm via the exportin5 channel. In the cytoplasm pre-miRNAs are cleaved into ~20bp miRNA/miRNA\* duplexes with a 2nt overhang on each end by Dicer, another RNase III enzyme. One of the mature miRNA strands, called the guide strand, is bound to a ribonucleoprotein (RNP) complex by becoming attached to the Argonaute protein, thus forming a miRNP complex, and the passenger strand is degraded<sup>16,36</sup>. It has been noted that the strand with the less stably bound 5' end is generally the guide strand<sup>36</sup>.



**Figure 1.6: The biogenesis of miRNAs.** Long pri-miRNA structures are cleaved in the nucleus by Droscha to shorter pre-miRNAs that are exported into the cytoplasm by Exportin 5. In the cytoplasm pre-miRNAs are cleaved ~20bp miRNA/miRNA\* duplexes with a 2bp overhang on either end. One strand is degraded and the other is incorporated into ribonucleoprotein complexes called micro-ribonucleoproteins (miRNPs). The miRNP complex can induce cleavage of its target mRNA if the miRNA is perfectly complementary to its target, as is generally the case in plants. When the miRNA is not perfectly complementary to its target mRNA, translation of the target mRNA is repressed.

If the miRNA is perfectly complementary to its target mRNA, the miRNP complex cleaves the target mRNA in the centre of the miRNA-mRNA duplex<sup>36</sup>. Most animal miRNAs are not perfectly complementary to their targets and their recruitment into miRNP complexes results in translational repression or deadenylation of the target mRNA. In the case of imperfectly complementary miRNAs, four mechanisms of translational repression are currently considered possible. When the miRNP complex binds to its target mRNA during protein synthesis, ribosomes are impeded or knocked off the mRNA strand leading to slowed translation<sup>37</sup>. Some researchers speculate that miRNP complexes engage as yet unknown protease enzymes during translation that then degrade the nascent peptides<sup>36</sup>,

although no evidence supporting this degradation has been presented to date. Finally, target mRNAs are sent to processing bodies (P-bodies) for storage or degradation as a result of the miRNP complex blocking the initiation of translation or inducing deadenylation<sup>36</sup>.

### 1.5.2 miRNA targeting principles

Even though imperfect base pairing of a miRNA to its target mRNA is generally seen in animals, nucleotides 2 to 8 at the 5' end of the miRNA, called the seed region, do have to bind in a perfect and contiguous manner for the miRNA to target a specific mRNA<sup>38</sup>. The miRNA-mRNA association is further enhanced by perfect pairing of the miRNA nt 12-17 to its target mRNA<sup>36,38</sup>. Based on modelling the stability of possible complexes Bartel *et al* propose that the miRNA guide strand is bound by Argonaute in a way that forces nucleotides 2 to 8 on the 5' end of the miRNA strand into a position favourable for pairing with the target mRNA<sup>16</sup>. Pairing at nucleotides 12-17 then further enhances binding, although it is not essential for target recognition.

Several studies have concluded that the most effective miRNA target sites are located in the 3' untranslated region (UTR) of mRNA, specifically about 15 nucleotides downstream from the stop codon<sup>36,38</sup>. Better targeting is seen when the miRNA binding site is present near either end of the 3' UTR, especially in longer 3' UTRs<sup>38</sup>. The expression of target mRNAs with two or more miRNA binding sites is repressed in a non-cooperative manner except when two target sites were between 8 and 40 nucleotides apart, suggesting that the miRNAs cooperate to increase the amount of repression<sup>38</sup>.

### 1.5.3 miRNA regulation of gene expression

Each microRNA expressed by a specific cell will have a large number of predicted target sites complementary to its seed region, in a large number of mRNAs. Whether any of these miRNA-mRNA interactions result in biologically relevant effects will depend on many different factors. Some microRNAs function as on-off switches of target function, as when *let-7* targets *lin-41* during *C. elegans* development<sup>12</sup>. MicroRNA expression can be upregulated in response to an environmental or developmental signal and repress a

target that had previously been expressed in the cell to a level more optimal in this new state. Switch interactions can also be the result of a pre-existing microRNA repressing a newly upregulated target, as is the case when *miR-9a* represses *senseless* in *Drosophila melanogaster*<sup>16</sup>. If a protein should be downregulated by other factors in the cell, such as the non-muscle gene *Tropomyosin* in developing muscle, the continued presence of a microRNA, *miR-1* in this case, will ‘mop up’ any deviant transcripts and ensure that functional levels of this protein are not achieved<sup>16</sup>. The levels of a target can also be fine-tuned, when miRNA expression reduces the protein expression to a more optimal but still functional level. Neutral miRNA – mRNA interactions do not have observable biological effects as the protein levels remain in the optimal range despite the miRNA-mediated repression. MicroRNAs are essential regulators of gene function, and their roles during development certainly deserve further elucidation.

This study aimed to expand the current knowledge of miRNA expression during chicken cardiogenesis, as well as to further elucidate the effects of a known miRNA, miR-1, on the developing heart. In **Chapter 3** an analysis of the complete sequencing results, including that of sequences not known to be miRNAs, is presented. This includes analysis of the quality of the NGS libraries sequenced, as well as an investigation of the top detected miRNAs. Following a literature search several miRNAs known to affect heart development and function were identified. The fold change patterns of these known miRNAs are examined in **Chapter 4**, to present networks of miRNAs that behave in ways significantly correlated to the changes in these known miRNAs. Furthermore, the biological effects of a subset of miRNAs on reporter constructs containing 3’ UTRs of several genes are summarised in **Chapter 4**. Novel miRNAs are validated (**Chapter 5**), and the effect of miR-1 dysregulation *in ovo* is presented (**Chapter 6**).

## **2 Materials and methods**

### **2.1 Embryo development**

After laying, eggs were stored at 16°C and then incubated at 37°C until the desired stage had been reached.

### **2.2 Dissection of embryos**

#### **2.2.1 For whole mount in situ hybridization (WISH)**

Embryos were dissected into cold DEPC-treated PBS. The extraembryonic regions were removed from embryos older than HH12. 4% PFA was added to flattened embryos and replaced with cold 4% PFA after 5 minutes, then left at 4°C overnight. Embryos were washed twice for 5 minutes in PBT, then twice for 5 minutes in MeOH. Embryos were stored at -20°C in MeOH for no longer than one month.

#### **2.2.2 For RNA purification**

Depending on the stage, cardiogenic mesoderm or cardiac tissue was dissected under cold DEPC-treated PBS. Dissected tissue was transferred into Trizol kept on ice during the dissection process; tubes were left at -80°C after dissection.

#### **2.2.3 For protein purification**

Cardiogenic mesoderm or cardiac tissue was dissected under cold DEPC-treated PBS and then transferred to tubes containing DEPC-treated PBS and kept on ice. The tubes were spun in a microcentrifuge at 2,000 rpm at 4°C for 5 minutes. The supernatant was removed and cold RIPA buffer added so that x mg tissue equalled x µl RIPA buffer. The tubes were vortexed thoroughly, then left on ice for 30 minutes with regular vortexing. Tubes were

spun at maximum rpm for 10 minutes and the supernatant was transferred to a new tube. Isolated protein was stored at -20°C for up to one month.

## **2.3 Purification and quantification of DNA, RNA and protein**

### **2.3.1 Plasmid DNA purification**

Plasmids were purified by QIAprep Spin Miniprep Kit (Qiagen) according to manufacturer's instructions.

### **2.3.2 DNA quantification**

DNA was quantified using a standard NanoDrop spectrophotometer;  $^{260}/_{280}$  ratios of 1.8 – 2.0 indicated good quality DNA.

### **2.3.3 Total RNA purification**

Chloroform was added to dissected tissue in trizol to 20% of the trizol volume. The tube was gently shaken for 15 seconds and incubated at room temperature for 4 minutes before being spun at maximum rpm for 20 minutes at 4°C. The clear aqueous phase was transferred to a new tube. Isopropanol was added to the aqueous phase to a volume of 50% of the original trizol volume and 2µl 15mg/ml Glycoblue was added. The solution was incubated at room temperature for 10 minutes and spun at maximum rpm for 15 minutes at 4°C. The supernatant was removed from the pellet and discarded. Freshly prepared 70% EtOH was used to wash the pellet by vortexing before spinning at maximum rpm for 5 minutes at 4°C. The supernatant was removed and the pellet dried at 50°C. The pellet was dissolved in a small amount of nuclease-free water and the quality of the RNA was checked by running a small aliquot on a 1% agarose gel. Samples were stored at -80°C.

### **2.3.4 RNA quantification**

RNA was quantified using a standard NanoDrop spectrophotometer;  $^{260/280}$  ratios of 1.8 – 2.0 indicated good quality RNA.

### **2.3.5 Protein purification**

Protein was purified from tissue as detailed in 2.3. above.

### **2.3.6 Protein quantification**

Protein was quantified by Bradford assay. Fresh dilutions of BSA in deionised water ranging in concentration from 2 ng/ $\mu$ l to 10 ng/ $\mu$ l were prepared and used as standards. 800 $\mu$ l sample was added to 200 $\mu$ l 5x Dye Reagent concentrate and left at room temperature in the dark for 5 minutes. Absorbance was measured at 595nm and the concentration of protein in each sample was determined.

## **2.4 In situ hybridisation**

### **2.4.1 mRNA probe synthesis**

#### **2.4.1.1 Plasmid preparation**

A small amount of the required DNA glycerol stock was streaked onto a LB agar plate containing an appropriate antibiotic and incubated upside down at 37°C overnight. A single isolated colony was picked and added to 10ml LB medium containing the relevant antibiotic, and grown at 37°C and 250 rpm overnight. Plasmid was isolated by QiaPrep Miniprep kit, according to manufacturer's instructions. 1 $\mu$ g of purified plasmid was linearized using the appropriate enzyme and buffer at 37°C for at least 2 hours. A small amount of digestion mix was run on a 1% agarose gel next to a small amount of uncut DNA to confirm successful digestion. DNA was precipitated from this solution by adding 0.3M NaOAc and 3 volumes of EtOH, leaving at -80°C for 1 hour and spinning at

maximum rpm for 30 minutes at 4°C. The supernatant was removed and the pellet washed with freshly prepared 70% EtOH by vortexing vigorously. The tube was spun at maximum rpm for 30 minutes at 4°C, the supernatant was removed and the pellet was dried at 50°C. Once dry, the pellet was resuspended in 8µl water.

#### **2.4.1.2 RNA synthesis**

1µg linearized DNA, 1x Transcription buffer, 0.5x RNasin, 10µM DTT, 1x NTP-DIG or 1x NTP-FITC, 0.5x Transcription enzyme (T3, T7 or SP6) and 2µl nuclease-free water were mixed and incubated at 37°C for two hours. The mixture was purified using the Illustra MicroSpin G-25 column purification kit, according to kit instructions. A small aliquot of eluate was checked on a 1% agarose gel to confirm successful synthesis; if both the RNA and DNA bands were not observed the probe synthesis was considered to have failed and the process was repeated. The eluate was left at 80°C for 5 minutes to denature the RNA, immediately placed into an equal amount of formamide and then added to 10ml Hybridisation buffer. Probes were stored at -20°C.

#### **2.4.2 mRNA whole mount in situ hybridisation**

##### **2.4.2.1 Hybridization**

Embryos were washed in 75% MeOH/PBT for 5 minutes, in 50% MeOH/PBT for 5 minutes, in 25% MeOH/PBT for 5 minutes and then twice for 5 minutes in PBT. Embryos older than HH18 were bleached in 6% H<sub>2</sub>O<sub>2</sub>/PBT for one hour. Embryos were washed in PBT 3 times for 10 minutes each time and left in 4% PFA for 30 minutes, then rinsed twice with PBT. Pre-warmed Prehyb buffer at 65°C was added to the embryos and replaced with more warm Prehyb buffer after 10 minutes. The embryos were left to rock in Prehyb buffer at 65°C for 2 hours. RNA probe, in Prehyb buffer at a concentration between 100ng/ml and 1µg/ml, was preheated to 65°C and used to replace the Prehyb buffer covering the embryos after two hours. The embryos were left to rock in the probe / Prehyb solution at 65°C for at least 12 hours.

#### **2.4.2.2 Post-hybridization washes**

2x SSC / 0.1% CHAPS and 0.2x SSC / 0.1% CHAPS were pre-heated to 65°C. Embryos were washed three times in 2 x SSC / 0.1% CHAPS for 20 minutes each time, and then the same was repeated with 0.2% SSC / 0.1% CHAPS. Room temperature KTBT was used twice to wash the embryos, 10 minutes each time.

#### **2.4.2.3 Antibody**

Embryos were washed for 3 hours at room temperature in KTBT with 20% goat serum to block non-specific interactions. Antibodies were diluted in KTBT with 20% goat serum: 1/2000 dilution of  $\alpha$ -DIG or 1/4000 dilution of  $\alpha$ -FITC. Diluted antibody was added to the embryos and left to rock at 4°C overnight.

#### **2.4.2.4 Post-antibody washes**

Embryos were washed for 1 hour in KTBT at room temperature, five times. Embryos were left to rock in MABT at 4°C overnight.

#### **2.4.2.5 NBT/BCIP detection**

Freshly prepared NTMT buffer was used to wash the embryos, twice for 15 minutes at room temperature. The colouring solution was made up by combining 9 $\mu$ l NBT (75mg/ml in 70% DMF) and 7 $\mu$ l BCIP (50mg/ml in 100% DMF) per ml NTMT buffer. Embryos were washed in this colouring solution, protected from light, until the signal was optimal. The colouring reaction was temporarily stopped by rinsing twice with KTBT at room temperature and then washing in KTBT until the background colour had been removed. Pictures of embryos were taken in KTBT.

### **2.4.3 LNA whole mount in situ hybridisation**

As for mRNA probes (2.4.1), but using a hybridisation temperature set to the LNA probe RNA  $T_m - 30^\circ\text{C}$ . LNA probes were pre-absorbed three times in day 10 embryos as described in 2.4.2.1

## **2.5 Cryosectioning of WISH embryos**

After WISH, embryos were washed twice in PBST for 5 minutes each time before being fixed in 4% PFA for 2 hours at room temperature. Embryos were then washed twice for 5 minutes each time in PBST and left to rock in a 30% sucrose solution at  $4^\circ\text{C}$  overnight. Embryos were placed in sectioning tubes and covered in O.C.T. embedding medium (Miles Inc.) for at least 2 hours. After repositioning embryos using needles and forceps, the O.C.T. tubes were rapidly frozen in a EtOH dry ice bath and left at  $-20^\circ\text{C}$  overnight.  $40\mu\text{m}$  sections were cut at  $-20^\circ\text{C}$ , sections were transferred to Superfrost Plus slides (Fischer Scientific) and dried at room temperature. The slides were then washed in  $\text{dH}_2\text{O}$  and mounted in Hydromount (National Diagnostics). The slides were dried overnight and an upright microscope (Zeiss) was used for microscopic analysis. Images were captured using AxioVision software.

## **2.6 In ovo microinjection into HH14 chick heart**

### **2.6.1 AntagomiR and RCAS virus injections**

Glass capillaries (1.0mm O.D. x 0.78 mm I.D.; Harvard apparatus, UK) were used to prepare microinjection needles on a horizontal Micropipette Puller (P-30, Sutter Instrument Co., CA). The closed tip of each needle was removed with fine forceps and  $8\mu\text{l}$  injectant was placed in each needle. Custom cholesterol-conjugated and FITC labelled antagomiRs supplied by Dharmacon. AntagomiRs were diluted to 0.1mM with RNase-free water before injection, whereas resuspended RCAS virus was used without further dilution. After removing eggshell to uncover the embryo, the shell membrane was removed and the heart

visualised. The needle was positioned using a micromanipulator and the heart tube filled with injectant by exertion of pressure. A small amount of PBS containing 0.1% Penicillin Streptomycin was pipetted onto the embryo and the hole closed with clear tape. Embryos were left at 37°C for 24 hours and harvested as described in 2.2.

## **2.7 Library preparation for Next Generation Sequencing**

### **2.7.1 TruSeq library preparation**

#### **2.7.1.1 mirVana small RNA enrichment**

Total RNA was enriched for small RNAs using the mirVana mRNA Isolation kit as per manufacturer's instructions (LifeTechnologies, AM1560). 50 – 100 µg of total RNA was mixed with 5 volumes of lysis buffer and 1/10 volume of miRNA homogenate additive added. The tube was vortexed thoroughly and left on ice for 10 minutes. 1/3 volume of EtOH was added to the mixture and the solution was vortexed. The mixture was pipetted onto a Filter Cartridge and centrifuged at 5,000 x g for 1 minute. The filtrate was collected and 2/3 volume room temperature EtOH was added; the mixture was vortexed thoroughly. The mixture was then applied to a new Filter Cartridge and centrifuged at 5,000 x g for 1 minute. The filtrate was discarded and the Filter Cartridge washed by applying miRNA Wash Solution 1 to the Filter Cartridge and spinning it at 5,000 x g for 1 minute. The filtrate was discarded, Wash Solution 2/3 was added and the tube was again spun at 5,000 x g for 1 minute. This step was repeated for a second time. The tube was spun for 1 minute at 10,000 x g before the Filter Cartridge was transferred to a new tube. Elution solution, heated to 95°C, was added to the Filter Cartridge and left to incubate at room temperature for 2 minutes. The tube was spun at 10,000 x g and the eluted RNA recovered. The enriched samples were quantified with a Qubit RNA High Sensitivity Assay kit (Life Technologies) and the quality of the small RNA-enriched samples was checked with a BioAnalyzer PicoRNA assay (Agilent).

### **2.7.1.2 Illumina TruSeq Small RNA library preparation**

Sequencing libraries were generated using mirVana small RNA-enriched samples, following the Illumina TruSeq Small RNA library preparation protocol. All incubations were done on a thermocycler to ensure precise control of the temperature. The 3' adapter was ligated to small RNA-enriched samples by truncated T4 RNA Ligase 2 at 28°C by incubating for one hour with ligation buffer and RNase inhibitor. The 5' adapter was ligated immediately by incubating the 3' ligated mixture, T4 RNA Ligase and ATP at 28°C for one hour. The 3' and 5' ligated mixture was reverse transcribed at 50°C for one hour by adding 1x First strand buffer, 1.25mM dNTPs, 20mM DTT, RNase Inhibitor and 5x diluted SuperScript II Reverse Transcriptase. PCR amplification of the cDNA was performed using the supplied PCR mix, an RNA PCR primer and an RNA PCR Primer Index, specific to each biological sample. 11 cycles of 10 seconds at 98°C, 30 seconds at 60°C and 15 seconds at 72°C were performed. Each library was purified on its own 1.0mm 6% Novex TBE PAGE gel; the section between 145 bp and 160 bp was excised and broken up into very small pieces by spinning through a 1.5ml Eppendorf tube with three holes made in the bottom into a 2.0ml Eppendorf tube at maximum RPM on a benchtop centrifuge. 200µl nuclease-free water was added to the gel pieces of each library and the slurry was left to rock at room temperature for 24 hours. The gel slurry was transferred onto a 5µm filter and spun for 10 seconds at 600 x g. The eluate was left at -80°C.

### **2.7.2 Illumina HiSeq Sequencing**

Small RNA libraries were sent to the Oxford Genomics Centre (Wellcome Trust Centre for Human Genetics) for 50bp paired end sequencing on the HiSeq 2000 platform (Illumina). Each library had been barcoded allowing all samples to be sequenced in a single lane.

## **2.8 Bioinformatics analysis**

Some bioinformatic analyses were performed by Dr. Simon Moxon (at the time working in Dr Grant Wheeler's lab). Unless otherwise indicated, all bioinformatic analyses were performed by the author.

### 2.8.1 Raw sequence mapping and filtering

Simon Moxon converted raw sequencing FASTQ files to FASTA format and removed adapter sequences with custom Perl scripts. Only reads that mapped perfectly by PatMaN<sup>39</sup> to the chicken genome (Galgal4<sup>40</sup>) were retained.

### 2.8.2 Profiling of mapped reads to miRBase

Initially Simon Moxon profiled mapped reads to miRBase<sup>41</sup> release 19 with the UEA Small RNA Workbench tool, miRProf<sup>42</sup>. However, the author subsequently profiled the mapped reads to miRBase release 20 using miRProf. All known miRNA analyses done in this work were a result of this profiling against miRBase release 20. All normalised reads larger than 0 reads per million were retained and used in subsequent analyses. miRProf settings were adjusted to return all matching animal miRNAs. Unless otherwise indicated, all data handling steps were performed using Microsoft Excel 2007.

### 2.8.3 Fold change profiles

Fold change profiles were generated by dividing the normalised reads per million values for each miRNA at each stage by the value recorded at HH5-7. It was considered that the HH5-7 sample represented a good baseline of miRNA expression against which to measure the levels detected at the remaining stages. Where the level of miRNA at HH5-7 was 0, the normalised read of the miRNA at each stage was returned without further division.

Correlations of fold change patterns were performed with GraphPad Prism version 5.01. A correlation matrix was generated and r values calculated for every pair of data sets. P values were calculated by Pearson's chi-squared test, assuming a two-tailed distribution at a 95% confidence interval. A hierarchical cluster of known heart important miRNAs (**Figure 4.5**) was prepared using the statistical package R by the default Complete Linkage method, with the script shown below.

```
### Read in data frame:
```

```

A <- read.csv("Known heart miRs fold change.csv")
### Indicate rownames and make new data frame with the assigned row names
rnamesA <- A[,1]          # assign labels in column 1 to "rnames"
A_data <- data.frame(A[,2:ncol(A)]) # transform columns
rownames(A_data) <- rnamesA      # assign row names
### Prepare hierarchical cluster
hc = hclust(dist(A_data))
### Load package ape
library(ape)
### Set colours for leaves:
edgecol <- rep('black', 60)
edgecol[c(1:3)] <- "darkorange2"
edgecol[c(5)] <- "forestgreen"
edgecol[c(7)] <- "lightseagreen"
edgecol[c(10,41)] <- "navy"
edgecol[c(11:19)] <- "deepskyblue"
edgecol[c(20:40)] <- "deepskyblue3"
edgecol[c(42:43)] <- "dodgerblue3"
edgecol[c(44:50)] <- "darkorchid"
### Plot basic tree
plot(as.phylo(hc), edge.width=2,edge.color=edgecol,cex = 0.9, label.offset = 2,
      tip.color="black",font=1,no.margin=TRUE)

```

#### 2.8.4 Novel miRNA prediction

Simon Moxon predicted novel miRNAs based on their sequence, their genomic location and the predicted folded structure of surrounding genomic sequence using the UEA Small RNA workbench tool miRCat<sup>42</sup>.

#### 2.8.5 Novel miRNA target prediction

Predicted target sites of novel miRNAs discovered in this screen were identified by searching for matching 6nt sequences in all the annotated 3' UTRs of the chicken genome (Galgal4), using custom R scripts written by the author. Chromosomes were searched in turn with a representative R script shown below.

```

utr3 <- read.DNAStringSet ("Chromosome 1 mart_export.txt")
P16mer <- vcountPattern ("CGCTCG", utr3)
P26mer <- vcountPattern ("CCTCTA", utr3)
P36mer <- vcountPattern ("CATTCA", utr3)
P46mer <- vcountPattern ("GGCACA", utr3)
P56mer <- vcountPattern ("ACAGAT", utr3)
P66mer <- vcountPattern ("CCCTGA", utr3)
P76mer <- vcountPattern ("GCGCCA", utr3)
P86mer <- vcountPattern ("AGCCAA", utr3)
P96mer <- vcountPattern ("CTGATG", utr3)
P106mer <- vcountPattern ("CCAGAA", utr3)
P116mer <- vcountPattern ("CATCCC", utr3)
P126mer <- vcountPattern ("CAGCGC", utr3)
P136mer <- vcountPattern ("TCTGCA", utr3)
P146mer <- vcountPattern ("AGCCCC", utr3)
### Split results by names
k <- strsplit (names(utr3), "\\ ")
### Link targets with sequences
m=c()
for (i in k) {
  m=rbind(m,i[1])}
### Combine search results with columns
#Targets <- cbind(PutativeOne6mer,m)
### Combine search results with columns for multiple searches
Targets <-
cbind(P16mer,P26mer,P36mer,P46mer,P56mer,P66mer,P76mer,P86mer,P96mer,P106m
er,P116mer,P126mer,P136mer,P146mer,m)
### Save
write.csv(Targets, file="Chromosome 1 All putative miRNAs.csv")

```

### 2.8.6 FunDo Disease Ontology analysis

Genes were analysed by FunDo using the web version<sup>43</sup> and genes found to be associated with diseases were returned. The Disease Ontology and peer-reviewed evidence from GeneRIF<sup>44</sup> was used for the statistical analysis<sup>45</sup>. As for Gene Ontology analysis, Fisher's exact test was used to evaluate the significance of each disease association<sup>46</sup>.

## **2.9 miRNA target selection**

### **2.9.1 TargetScan target identification**

TargetScan Release 6.2 (June 2012) was searched for conserved and non-conserved targets of selected chicken miRNAs. Three lists of targets of the selected miRNAs were generated: those that were known to be involved in Wnt signalling, those with known roles in heart development following a literature search, and those that were more speculative and were targeted by more than one of the selected miRNAs. The final selection of 14 targets were made based on the ease with which their 3' UTRs could be amplified from cDNA. This was done as many of the 3' UTRs of the chicken genome are not well annotated and therefore a resistance to PCR amplification could indicate that the 3' UTR is not correctly annotated, belongs to a splice variant expressed at very low levels, or may be too challenging to amplify due to high AT-richness.

### **2.9.2 GO terms analysis of miRNA targets identified by TargetScan**

gProfiler, an online toolset allowing profiling of gene lists based on Gene Ontology, was used to cluster predicted target genes of each miRNA (January 2014). Once targets of interest had been selected, the chicken orthologues of these genes were identified using the g:Orth function of gProfiler. Any statistical significance assigned to associated GO terms is a result of gProfiler analysis. It should be noted that although GO terms analysis is useful for trying to predict certain trends, the gene lists are often devoid of context and should be accepted with caution.

### **2.9.3 Design of primers for 3' UTR amplification of predicted targets**

Where available, the 3' UTR sequence of predicted target genes was downloaded from ENSEMBL 71 (April 2013; Galgal4) and the predicted target site identified by EMBOSS Needle, an online tool that allows sequence alignment of two selected sequences at a time. For target genes where a 3' UTR was not known, or where the predicted target site was

located downstream of the end of the annotated 3' UTR, genomic DNA sequence downstream of the last known intron or exon was used to design primers.

**Table 2.1: A summary of the nature of the 3' UTR sequences cloned into miRNA sensing luciferase constructs.**

<b>Gene</b>	<b>Known UTR?</b>
FZD8	No; used sequence downstream of coding region.
ATG7	No; used sequence downstream of coding region.
VANGL2	No; used sequence downstream of coding region.
QKI	No; used sequence downstream of coding region.
ZEB2	No; used sequence downstream of coding region.
ENAH	Yes
TCF12	Yes
PPP3CB	Yes
LEF1	Yes
ETS1	Yes
PRKACB	Yes
NFIB	Only partly annotated; used sequence downstream of end of annotated region.
NR3C1	Only partly annotated; used sequence downstream of end of annotated region.
ROCK2	Only partly annotated; used sequence downstream of end of annotated region.

Primers were designed using Primer3Plus to include the whole known 3' UTR, or at least 1,000bp of sequence if not. The restriction site sequence of BglII was added to the 5' end of each forward primer and of NheI to the 5' end of each reverse primer, to allow subcloning of the fragment into a modified pGL3 vector<sup>47</sup>.

## **2.10 Experimental evaluation of putative miRNA targets**

### **2.10.1 PCR of 3' UTRs**

#### **2.10.1.1 Reverse Transcription**

Total RNA was purified from whole HH18 and HH20 chicken embryos by phenol chloroform isoamyl alcohol as described in **2.3.3**. cDNA was synthesised from this RNA using SuperScript II Reverse Transcriptase (Invitrogen).

20µl Reverse transcription reaction mix:

200ng Random primers  
2µg RNA  
1µl dNTPs (10mM each)  
4µl RNase-free H<sub>2</sub>O  
4µl 5x First Strand Buffer  
2µl DTT (100mM)  
1µl RNasin (40U/µl)  
1µl SuperScript II Reverse Transcriptase (200U/µl)

The reaction mixture was incubated on a thermocycler at 42°C for 50 minutes, then 70°C for 15 minutes before freezing at -20°C. Two tubes were prepared for each sample; one with Reverse Transcriptase added (RT+) and one without Reverse Transcriptase (RT-) to check whether the RNA was contaminated with DNA.

### **2.10.1.2 PCR**

Aliquots from both RT+ and RT- tubes were amplified for each sample. cDNA was amplified with the Expand High Fidelity PCR System (Roche).

50µl PCR amplification mix:

1µl dNTPs (10mM each)  
37.75µl dH<sub>2</sub>O  
1.5µl Forward and Reverse primer mix (10µM each)  
2µl RT+ or RT- mix  
5µl 10x Expand High Fidelity Buffer with 15mM MgCl<sub>2</sub>  
2.6U Expand Enzyme

The mixtures were amplified as shown below:

94°C 2 minutes  
25 cycles of:  
94°C 15 seconds  
57°C 30 seconds  
72°C 2 minutes  
Then,  
72°C 7 minutes

### **2.10.1.3 Agarose gel electrophoresis and purification of PCR products**

The PCR products were purified by 0.7% agarose (Melford laboratories). The size of each product was evaluated by molecular size marker (Roche). Products were cut from the gel

and purified by GeneJET Gel Extraction Kit (Thermo Scientific) according to manufacturer's instructions. None of the RT- control reactions showed any amplification.

## **2.10.2 Generating miRNA-sensing luciferase constructs**

### **2.10.2.1 Ligation into pGEM-T Easy**

Purified PCR fragments were ligated into pGEM-T Easy using the pGEM-T Easy Vector System I (Promega).

10µl Ligation mixture:

5µl	2x Rapid Ligation Buffer
37.5ng	pGEM-T Easy Vector
1µl	T4 DNA Ligase
3.25µl	Purified PCR product

Mixtures were left at room temperature for 2 hours.

### **2.10.2.2 Transformation into DH5α E.coli competent cells**

75µl DH5α *E.coli* chemically competent cells were thawed on ice for 30 minutes. pGEM-T Easy ligation mixtures were simultaneously cooled on ice for 30 minutes. The entire ligation mixture was added to the competent cells and mixed by gentle flicking. Tubes were placed in a heat block at 37°C for 30 seconds and then left on ice for 2 minutes. 800µl LB broth was added and tubes were left shaking at 37°C and 200RPM for 60 minutes. After incubation tubes were gently spun down at 3,000 RPM for 3 minutes and the pellet resuspended in 200µl broth. 40µl 20mg/ml X-gal (Thermo Scientific) and 4µl 1M IPTG (Thermo Scientific) was added to the cell suspension and the entire volume was plated onto a LB-Agar plate containing carbenicillin and incubated at 37°C overnight. A single white colony was picked from each plate and cultured overnight in 10ml LB broth, at 37°C and 200RPM.

### **2.10.2.3 Plasmid purification**

pGEM-T Easy constructs were purified by QIAprep Spin Miniprep Kit (Qiagen) according to manufacturer's instructions. DNA was quantified as described in **2.3.2**.

### **2.10.2.4 Restriction of amplified 3' UTRs**

pGEM-T Easy constructs were digested with BglII and NheI restriction enzymes (Promega) to release the amplified 3' UTRs.

Digestion mix:

2µg pDNA  
3µl Buffer B  
xµl dH<sub>2</sub>O to make the reaction up to 30µl  
1µl NheI Enzyme  
1µl BglII Enzyme

Reaction mixtures were left in a heat block set to 37°C for 2 hours. The restricted 3' UTR fragments were isolated by agarose gel electrophoresis as described in **2.10.1.3** and quantified as described in **2.3.2**.

### **2.10.2.5 Ligation into modified pGL3 vector**

A modified pGL3 vector, created by deleting the MCS upstream of the SV40 promoter in the pGL3 Control cloning vector (Promega) and inserting a MCS downstream of the luciferase stop codon, was cultured and isolated as described in **2.10.2.3**. The empty vector was digested with BglII and NheI restriction enzymes and purified as described in **2.10.2.4** and quantified as described in **2.3.2**. Purified 3' UTR fragments (**2.10.2.4**) were ligated into the modified pGL3 vector.

Ligation mix:

5µl 2 x Rapid Ligation Buffer  
66ng Linear empty pGL3 vector  
xµl Purified 3' UTR fragment  
1µl T4 DNA Ligase  
yµl dH<sub>2</sub>O to make the reaction up to 10µl

A ratio of 6 moles of insert to 1 mole of empty pGL3 vector was used. The ligation mix was left at room temperature overnight. The entire ligation mix was transformed into DH5 $\alpha$  *E.coli* cells and cultured as described in **2.10.2.2**, without the use of X-Gal.

#### **2.10.2.6 Colony PCR identification of constructs containing 3' UTR inserts**

Seven colonies from each plate were checked for the presence of the specific insert by PCR, with the original pGEM-T Easy construct serving as a PCR control.

20 $\mu$ l PCR mix:

10 $\mu$ l 2x BioMix Red (Bioline)  
1 $\mu$ l Forward and reverse primer mix used to amplify 3' UTR (10 $\mu$ M each)  
9 $\mu$ l dH<sub>2</sub>O

Cycling conditions:

94°C 4 minutes  
30 cycles of:  
94°C 30 seconds  
55°C 30 seconds  
72°C 1 minute  
Then,  
72°C 7 minutes

Each colony was numbered and transferred to a new LB Agar carbenicillin plate and incubated at 37°C overnight. PCR products were assessed on a 0.7% agarose gel. One colony from each plate, known to contain the insert after PCR, was cultured as described in **9.2.2** and purified as described in **2.10.2.3**. The sequence of the insert was checked by Sanger sequencing using pGL3 sequencing primers:

pGL3 Sequencing primer 1:

5'-CCTCATAAAGGCCAAGAA-3'

pGL3 Sequencing primer 2:

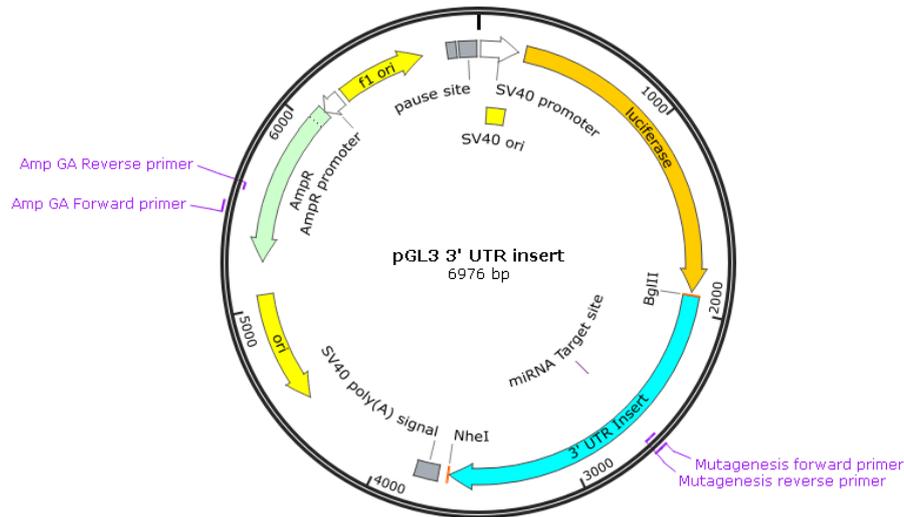
5'-CTCATCAATGTATCTTATCATGTC-3'

#### **2.10.3 Generating pGL3 mutant constructs**

pGL3 mutant constructs were generated using Gibson Assembly which generates a construct from PCR products with overlapping regions.

### 2.10.3.1 Mutagenesis PCR of pGL3 constructs

A pair of primers placed in the ampicillin resistance gene were used as universal primers in all mutagenesis cloning. Primers were designed over the miRNA target site which contained mismatched nucleotides chosen to create an enzyme restriction site out of the miRNA target site (**Figure 2.1**).



**Figure 2.1: pGL3 Luciferase miRNA-sensing construct.** An example of a modified pGL3 vector containing an amplified 3' UTR insert, showing the position of mutagenesis primers.

The two halves of each pGL3 construct were amplified using Phusion High-Fidelity DNA Polymerase (Finnzymes, NEB).

PCR mix and method:

2.5 $\mu$ l Forward primer (10 $\mu$ M), 2.5 $\mu$ l Reverse primer (10 $\mu$ M) and 20ng pDNA were combined in a PCR tube.

Mastermix:

31.5 $\mu$ l dH<sub>2</sub>O  
10 $\mu$ l 5 x Phusion HF buffer  
1 $\mu$ l dNTPs (10mM each)  
1U Phusion polymerase

43 $\mu$ l mastermix was added to the primers and pDNA mix by gentle pipetting.

Cycling conditions:

98°C 30 seconds  
20 cycles of:  
98°C 10 seconds  
59°C 30 seconds

72°C X minutes  
Then,  
72°C 10 minutes  
4°C 2 minutes

An extension time of approximately 1 minute per 1kb of PCR product was used. 5U DpnI restriction enzyme (Promega) was added to the PCR mixture immediately after the final PCR step. The mixture was incubated on a thermocycler at 37°C for 2 hours to digest any remaining pDNA. PCR products were purified on 1% agarose gels as described in **2.10.1.3** and quantified as described in **2.3.2**.

### **2.10.3.2 Gibson assembly of PCR products**

The two complementary halves of each plasmid were assembled using Gibson Assembly mastermix (NEB). PCR products were combined in a 1:1 molar ratio, with 100ng of the larger product being used.

Gibson Assembly reaction mixture:

A $\mu$ l PCR product A  
B $\mu$ l PCR product B  
10 $\mu$ l 2x Gibson Assembly mastermix  
x $\mu$ l dH<sub>2</sub>O to make reaction volume up to 20 $\mu$ l

The Gibson Assembly mixture was incubated on a thermocycler at 50°C for 60 minutes. 2 $\mu$ l Gibson Assembly mixture was then transformed into 150 $\mu$ l DH5 $\alpha$  *E.coli* cells as described in **2.10.2.2** and purified as described in **2.10.2.3**. A diagnostic digest of the purified mutant pDNA was done to confirm successful mutation.

Digestion mix:

1 $\mu$ g pDNA  
2 $\mu$ l 10x Restriction buffer  
1 $\mu$ l Restriction enzyme  
x $\mu$ l dH<sub>2</sub>O to make reaction volume up to 20 $\mu$ l

Digestion mixtures were incubated at 37°C for 2 hours, and digestion checked on a 1% agarose gel.

## **2.10.4 Luciferase assays of miRNA sensing constructs**

### **2.10.4.1 Cell culture**

Chicken DF1 fibroblast cells were cultured in Dulbecco's modified Eagle medium (DMEM) containing GutaMAX, 1g/L D-glucose and Pyruvate (Life Technologies), with 10% fetal bovine serum (FBS) (Gibco, Life Technologies) and 1% penicillin-streptomycin (Gibco, Life Technologies) added on first use. Cells were passaged every other day by treatment with 1ml 0.25% Trypsin-EDTA (Gibco, Life Technologies) for 30 seconds at room temperature followed by removal of trypsin and further incubation at 37°C for approximately 2 minutes. Cells were then transferred to a new flask containing media at a 1:5 dilution. Cells were maintained in a humidified cell culture incubator at 37°C with 5% CO<sub>2</sub>.

### **2.10.4.2 Cell transfection**

Cells were detached from the flask surface as described in 2.10.4.1, and counted using a haemocytometer. Approximately 6,000 cells were added to each well of a 96-well plate and media added to a volume of 100µl. Cells were incubated for at least 12 hours to allow re-attachment to the plate surface. The media was removed and replaced with 50µl serum-free media. A transfection mix was prepared as described below.

Transfection mix 1:

100ng pGL3 pDNA  
50nM miRNA mimic  
25ng Renilla pDNA  
xµl Serum-free media to a volume of 25µl

Transfection mix 2:

25µl Serum-free media  
0.2µl Lipofectamine 2000 Transfection reagent (Life Technologies)

Tubes were left at room temperature for 5 minutes. Transfection mixes 1 and 2 were combined and left at room temperature for 20 minutes. 50µl of the combined transfection mix was added to a well of cells and cells were left to incubate at 37°C. After 6 hours the media was removed and replaced with media described in 2.10.4.1. but without 1%

penicillin-streptomycin. Cells left in humidified incubator for 24 hours. Either miScript miRNA mimics (Qiagen) or custom siRNA oligos (Sigma) were used as miRNA mimics. Custom siRNA oligos were designed to represent the endogenous miRNA/miRNA\* duplexes, complete with mismatches. Each construct was evaluated in the presence of the putative target miRNA mimic, a control miRNA mimic (miR-140) and without a miRNA mimic. Some 3' UTRs had more than one putative miRNA target; these constructs were also evaluated with a mixture of the two putative miRNA mimics.

#### **2.10.4.3 Luciferase assay**

A Dual-Luciferase Reporter Assay System kit was used to assess luciferase activity (Promega). This kit allows *Renilla* Luciferase to serve as an internal control of transfection efficiency.

Cells were washed twice with cold PBS and care was taken to remove all PBS. 60µl 1x Passive Lysis buffer was added to each well and the plate was rocked gently for at least 15 minutes. 10µl lysis solution and 50µl Luciferase Assay Reagent II were combined and photo emission was measured at 562 nm using a multi-label counter (Perkin Elmer). 50µl Stop & Glo reagent was added and the *Renilla* Luciferase photo emission measured at 562 nm.

#### **2.10.4.4 Normalisation of luciferase assay data**

Four biological replicates, each with three technical replicates, were performed. Each firefly luciferase assay reading was normalised to its *renilla* luciferase reading. The average activity of the plasmid in the absence of miRNA mimics was set to 100%. The activity of the constructs transfected with putative target miRNA mimic(s) and control miRNA mimics was normalised to the average activity of the plasmid in the absence of miRNA mimics. A student's t-test using standard deviation values was performed using GraphPad Prism version 5.01 for Windows, GraphPad Software, San Diego California USA, [www.graphpad.com](http://www.graphpad.com).

## 2.11 Western blotting

Loading buffer and 1 $\mu$ l 100mM DTT was added to 40 $\mu$ g protein from tissue lysate prepared as described in 2.3.5 and quantified as in 2.3.6, and denatured at 70°C for 10 minutes. Samples were loaded onto an 8% resolving polyacrylamide gel with a 5% stacking polyacrylamide layer and run at 100V for up to 2 hours.

Samples were then transferred onto a nitrocellulose membrane using a semidry blotter (BioRad) at 100V for 90 minutes. Following transfer the membrane was washed in TBST for 5 minutes and then blocked in freshly prepared 5% milk in TBST at room temperature for 2 hours. The membrane was washed in TBST for 15 minutes and then rocked in a 1:5,000 dilution of primary antibody in 5% milk in TBST overnight at 4°C. After washing the membrane four times for 10 minutes each time in TBST, a secondary antibody coupled to HRP (Jackson Laboratories) was applied at a 1:1,500 dilution for 1 hour at room temperature. Residual secondary antibody was removed by washing with TBST. Protein was detected by chemiluminescence subsequent to treatment of the membrane with 1.25mM luminol (Sigma) and 0.2mM  $\beta$ -coumaric acid (Sigma) in 100mM Tris-HCl pH 8.5, followed by imaging on a Fujifilm LAS 3000 imager (Fuji Photo Film Co., Ltd).

Membranes were stripped by washing in mild stripping solution.

Mild stripping solution:

15g Glycine  
1g SDS  
10ml Tween20  
xml Sterile H<sub>2</sub>O to 1L after adjusting the pH to 2.2.

Membranes were washed twice for 10 minutes in mild stripping solution, then twice in PBS for 10 minutes and finally twice in TBST for 5 minutes before being put back into 5% milk in TBST.

## 2.12 Northern blotting

Total RNA was purified and quantified from whole HH16 and HH20 embryos as described in 2.2.2, 2.3.3 and 2.3.4 and probed for microRNAs.

### 2.12.1 Polyacrylamide gel electrophoresis

A 15% denaturing polyacrylamide gel was prepared by dissolving urea in dH<sub>2</sub>O and combining with other reagents as described below:

2.1g	Urea
1.25ml	dH <sub>2</sub> O
1.85ml	40% Acrylamide solution
500μl	5x TBE
50μl	10% APS
2.5μl	TEMED

2x loading dye was added to 50μg total RNA and denatured at 90°C for 30 seconds before placing on ice.

2x loading dye mix:

50μl	0.5M EDTA
0.2g	2% Bromophenol Blue
0.01g	Xylene cyanol
9.5ml	Formamide

Samples were run at 100V for approximately 1 hour in 0.5x TBE running buffer.

### 2.12.2 Semi-dry transfer and cross-linking

RNA was transferred onto a membrane (Hybond NX, Amersham Biosciences) by semidry blotter at 300mA for 1 hour, and then crosslinked.

Crosslinking mix 1:

10ml	dH <sub>2</sub> O
122.5μl	12.5M 1-methylimidazole
10μl	HCl <sub>conc</sub>

Crosslinking mix 2:

10ml	Crosslinking mix 1
0.373g	EDC
2ml	dH <sub>2</sub> O

The membrane was placed on top of Whatman filter paper that had been soaked in crosslinking mix 2, sealed with clingfilm and left rocking at 60°C for 2 hours. The membrane was then washed in dH<sub>2</sub>O twice for 5 minutes at room temperature.

### 2.12.3 Probing

#### 2.12.3.1 Pre-hybridisation

ULTRAhyb-Oligo Hybridization Buffer (Life Technologies) was warmed to 37°C before being added to hybridisation tubes containing RNA membranes. Membranes were washed in hybridisation buffer in a HB-1000 hybridisation oven (UVP) at 37°C for 2 hours.

#### 2.12.3.2 $\gamma^{32}\text{P}$ -ATP end labelling of complementary oligonucleotides

$\gamma^{32}\text{P}$ -ATP labelled probes were generated by a T4 Polynucleotide Kinase (NEB) reaction.

Reaction mix:

2 $\mu\text{l}$	10 $\mu\text{M}$ Complementary oligo
2 $\mu\text{l}$	10x Reaction Buffer A
12 $\mu\text{l}$	Nuclease-free H <sub>2</sub> O
1 $\mu\text{l}$	T4 Polynucleotide Kinase
3 $\mu\text{l}$	$\gamma^{32}\text{P}$ -ATP

The mixture was incubated in a heat block at 37°C for an hour and unincorporated nucleotides were removed by centrifugation (MicroSpin G25 columns, Illustra). Sterile H<sub>2</sub>O was added to the probe to a final volume of 50 $\mu\text{l}$ , added to the hybridisation tubes containing the RNA membranes and left rotating overnight at 37°C. The membranes were washed with washing solution at 37°C.

Washing solution:

500 $\mu\text{l}$	20x SSC pH 5
500 $\mu\text{l}$	10% SDS
49ml	dH <sub>2</sub> O

The membranes were vigorously rinsed in washing solution, then washed twice for 20 minutes at 37°C and finally rinsed again with washing solution before being wrapped in clingfilm and exposed on a phosphorimaging plate (Fujifilm BAS cassette 2025). Plates were developed using a phosphorimager (Molecular imager FX pro plus, BIORAD). Membranes were stripped by washing twice in boiling 0.1% SDS for 20 minutes.

## **2.13 Image J quantification of Western and Northern blots**

Relative intensities of Western and Northern blot bands were calculated using ImageJ software<sup>48,49</sup>. The Rectangular Selections pane was used to draw a rectangle around the normalising band (Actin for Western blots and U6 for Northern blots). Once the selection was confirmed a second, identically sized rectangle was used to surround the target band (FZD7 or RAR $\beta$  for Western blots, miR-130 for Northern blots). The software automatically calculated the intensity of the coloured pixels in the identically sized areas and allowed the normalisation of target to loading control. This value was then expressed as a percentage and a comparison made between adjacent lanes. A student's t-test was performed on the Western blot results using GraphPad Prism version 5.01 for Windows, GraphPad Software, San Diego California USA, [www.graphpad.com](http://www.graphpad.com) and P values reported.

## **2.14 Fold**

## **3 Small RNA sequencing of chick cardiac and cardiogenic tissue**

### **3.1 Introduction**

A summary of the sequencing strategy employed in this project, along with bioinformatic analysis of the known animal miRNAs identified in the samples is presented in this chapter. Apart from removal of adapter sequences and removal of reads that did not map to the chicken genome, all work was performed by the author. Simon Moxon performed the adapter removal and the mapping as described in **2.8.1**.

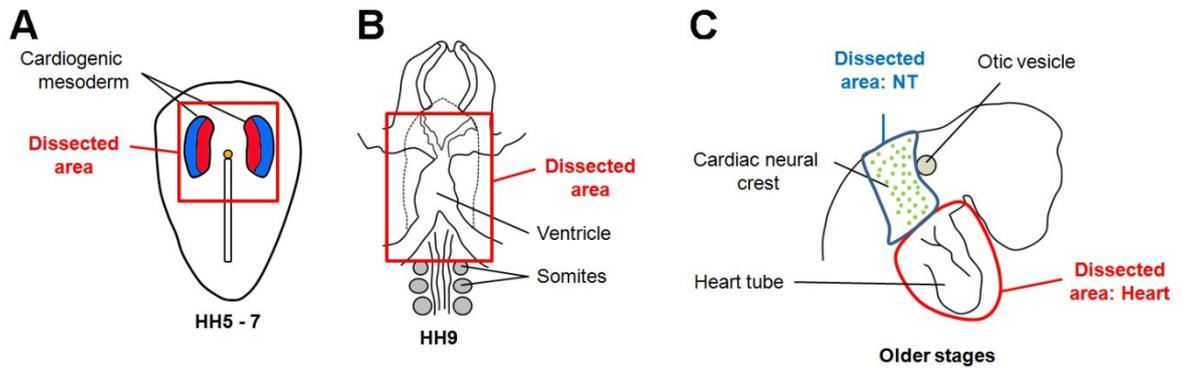
#### **3.1.1 Dissection of tissues for heart tissue-enriched miRNA sequencing**

In order to investigate the miRNA expression profile of the chick embryo during cardiogenesis, cardiac and cardiogenic tissue was dissected from embryos over a range of stages that represent many of the major early events in heart development (**Table 3.1**).

**Table 3.1: A summary of the events taking place during heart development between stage 5 (HH5) and stage 23 (HH23) in the chicken embryo.**

Period	Major heart development events
HH5-7	Cardiac progenitor cells migrate from the primitive streak, forming the bilateral heart fields. Terminal differentiation markers can be observed by HH7.
HH8-10	The heart fields fuse. Cardiomyocytes start to differentiate. Cardiac neural crest cells start to migrate from the neural crest to the outflow tract.
HH11-13	The heart tube loops from a C- to an S-shape. Growth during this time is mainly from migration of cells. The cardiac neural crest cells migrate into the pharyngeal arch arteries.
HH14-16	Myocardial cells proliferate in the ventricles, causing them to balloon out. The epicardium and smooth muscle cells start to form. Atrial septation starts.
HH17-20	The four layers of the heart (myocardium, cardiac jelly, endocardium and epicardium) have formed. The cardiac neural crest cells reach the outflow tract and valve formation begins.
HH21-23	Valve formation continues. Ventricular septation begins.

During the early stages of development (HH5 to HH7), cardiogenic mesoderm migrates from the primitive streak to form the bilateral heart fields (**Figure 3.1A**). Between stage HH8 and stage HH10 the bilateral heart fields fuse and the heart tube is formed (**Figure 3.1B**). During these early stages it is difficult to separate cardiogenic and cardiac tissue from other surrounding tissues and therefore the dissected area was chosen to include all forming heart regions but also contained other developing tissues. From stage HH14 onwards the heart projects away from the rest of the embryo, and it becomes more feasible to dissect only heart tissues. During these later stages the cardiac neural crest cells migrate towards the forming heart (**Figure 3.1C**). To capture these migrating cells, tissue indicated in the blue region of **Figure 3.1C** was dissected and labelled “NT”; these samples contain many different tissue types, including neural tube and head mesenchyme, in addition to cardiac neural crest cells.



**Figure 3.1: Tissue dissection performed prior to small RNA enrichment. A:** A schematic of a typical chick embryo between stages HH5 and HH7, with the migrating cardiogenic mesoderm indicated in blue and red. The dissected area is shown in the red box. **B:** A ventral view of a HH9 chick embryo. By HH9 the heart fields have fused and a heart tube has formed. **C:** A lateral view of an embryo past HH14. Dissected heart tissue is shown outlined in red, and a second sample containing migrating neural crest and other tissues, designated “NT”, is outlined in blue.

In total, nine samples were enriched for small RNA and used to prepare Illumina TruSeq indexed libraries: HH5-7, HH8-10, HH11-13, HH14-16 Heart, HH14-16 NT, HH17-20 Heart, HH17-20 NT, HH21-23 Heart, and HH21-23 NT. These libraries were sequenced by 50bp paired-end Illumina HiSeq 2000 sequencing.

### 3.1.2 Next Generation Sequencing technologies

The Human Genome Project was initiated in 1990 and aimed to sequence the entire human haploid genome by automated Sanger sequencing<sup>50</sup>. Despite the increase in sequencing throughput that resulted from improvements in automation, the Sanger sequencing method still relies on polymer separation of fluorescently labelled DNA fragments, limiting the number of samples assayed in parallel and thus limiting throughput<sup>51</sup>. The need for improvement drove the development of several Next Generation Sequencing (NGS) platforms, able to produce large amounts of sequencing data by a variety of different strategies<sup>50,51</sup>.

**Table 3.2: Summary of the capacities of the most widely used sequencing platforms.** Comparison of Illumina GAIIX<sup>52</sup>, Illumina HiSeq 2000<sup>52</sup>, Life Technologies 5500W<sup>53</sup> and Roche 454 FLX+ platforms<sup>54</sup>. The approximate total number of GAIIX, HiSeq 2000, Roche 454 and ABI SOLiD machines in the world<sup>55</sup> is indicated to give perspective on the popularity of the different NGS platforms.

Platform	Illumina GAIIX	Illumina HiSeq 2000	LifeTechnologies SOLiD 5500W	Roche 454 FLX+
<b>Template preparation</b>	Solid-phase PCR	Solid-phase PCR	Solid-phase PCR	Emulsion PCR
<b>NGS chemistry</b>	Reversible terminators	Reversible terminators	Ligation	Pyrosequencing
<b>Read length</b>	2 x 150bp	2 x 150bp	75bp + 25bp	700bp
<b>Run time</b>	10 days	11 days	7 days	1 day
<b>Gb per run</b>	30	600	160	0.7
<b>Machines worldwide</b>	500	826	~324	364

Bp: base pairs; Gb: gigabases.

Although not without their limitations<sup>51</sup>, these new technologies are rapidly being improved to increase throughput, reduce errors and decrease cost. Several platforms are currently in development<sup>56</sup>, and many benchtop machines are now available for smaller-scale use<sup>57,58</sup>. Three platforms are most commonly used for larger sequencing projects: the 454 sequencer by Roche, Life Technologies' SOLiD system as well as the Illumina GAIIX and HiSeq systems (**Table 3.2**). In addition, improvements in reagents chemistries and machine optics have allowed for steady gains in sequencing output and flexibility. For example, the Illumina HiSeq 2500 can be used for either rapid runs (600x10<sup>6</sup> reads per run, 2 x 150bp reads), or for high output runs (4x10<sup>9</sup> reads per run, 2 x 125bp reads)<sup>59</sup>.

### 3.1.3 Techniques for assessing microRNA expression

Three methods are currently used to assess expression of miRNAs: microarrays, quantitative PCR (qPCR) and NGS. Microarrays involve the hybridisation of labelled small RNAs to oligonucleotide probes on a solid surface<sup>60</sup>. The design of microRNA probes is often constrained by the short length of miRNAs, which can lead to inclusion of the entire mature miRNA sequence in the probe and therefore a large variation (>20°C) in the melting temperatures ( $T_M$ ) of the different probes<sup>61</sup>. Such a large variation in  $T_M$  then

results in cross-hybridisation and loss of specificity. This problem, along with difficulties in distinguishing between different members of a miRNA family, which can vary by as little as a single nucleotide, have been overcome by allowing the increase in hybridisation temperature as a result of the inclusion of Locked Nucleic Acid (LNA) nucleotides in the probe. Data analysis of microarray data is simple and well established, and for profiling experiments microarrays have much to offer. However, novel miRNAs cannot be readily detected by microarray.

qPCR for miRNAs can be done by preparing cDNA with a stem-loop primer, followed by either TaqMan or SYBR Green assays with a universal primer complementary to a portion of the stem loop sequence and a miRNA-specific forward primer<sup>62</sup>. The  $T_M$  of the forward primer can be adjusted by adding a tail, the length of which will depend on the  $T_M$  of the mature miRNA<sup>62</sup>, or by adding LNA nucleotides<sup>63</sup>. The design of miRNA-specific qPCR primers can be costly and it is often necessary to assay several genes in order to identify good normalisers for each experiment<sup>61</sup>. As with microarrays, qPCR cannot be used to discover novel miRNAs.

Sequencing of miRNAs, in contrast, does not require knowledge of the miRNA and therefore can be used to detect novel miRNAs. However, the preparation of samples for sequencing involves ligation and PCR amplification steps and at the moment is still plagued by biases introduced by these steps<sup>61,64</sup>. Finally, extensive computational knowledge is required to perform bioinformatic analyses on NGS data, although increasing numbers of user-friendly tools have recently become available<sup>42</sup>.

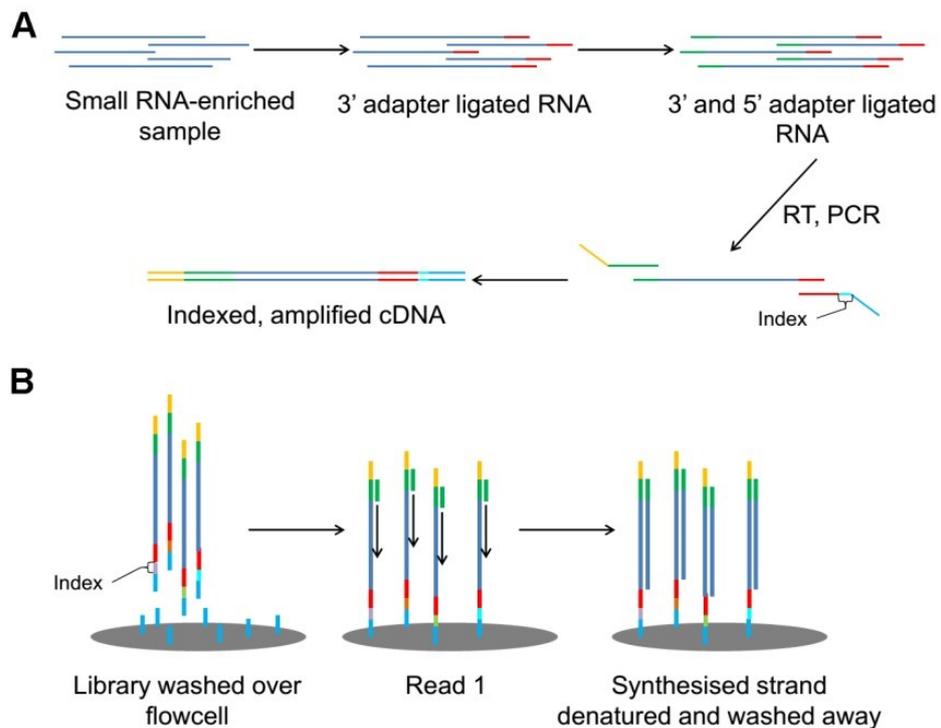
NGS was selected for this project as it represented the technology that offered the ability to both profile the expression of known miRNAs and to detect previously undiscovered miRNAs. The Illumina HiSeq 2000 platform was chosen for its accuracy, its sequencing yield and its flexibility which allowed all nine samples to be sequenced in a single lane. The High-Throughput Genomics Group at the Wellcome Trust Centre for Human Genetics in Oxford performed the Illumina HiSeq sequencing.

### 3.1.4 NGS Library preparation and sequencing

The general workflow from biological sample to sequencing result is similar for all three major NGS systems (**Table 3.2**): nucleic acid isolation, library preparation, template preparation, sequencing, and bioinformatic analysis. It is essential that pure DNA or RNA of good quality is used as an input into the library preparation process and to this end sample quality must be assessed. Library preparation involves the generation of a pool of DNA, the length of which will vary depending on the platform and the application, with sequencing primers ligated to each end of every strand. A small amount of this library is then immobilized either on beads, in the case of SOLiD and 454, or on a sequencing slide, in the case of both Illumina platforms<sup>50</sup>. Once immobilized, the library is amplified and then sequenced by extending primers by synthesis<sup>51</sup>.

Small RNA library preparation for the HiSeq platform comprises the ligation of a sequencing adapter to the 3' end of a small RNA-enriched sample followed by ligation of a different sequencing adapter to the 5' end of the sample (**Figure 3.2A**). cDNA synthesis and PCR amplification, with primers complementary to the ligated adapters, follows. The amplified products are then separated on a polyacrylamide gel allowing the isolation of the library, double stranded DNA copies of adapter ligated small RNA.

Once the library has been validated to contain a pool of the correct size DNA fragments by checking the size distribution of the library by fluorescent capillary gel electrophoresis on a BioAnalyzer machine, template preparation can begin. The double stranded DNA library is first denatured and then flowed over a solid surface, the flow cell, covered with primers covalently attached to the flow cell surface and complementary to the adapters ligated onto the library (**Figure 3.2B**).



**Figure 3.2: A brief overview of Illumina TruSeq small RNA library preparation and paired end sequencing.** **A:** Proprietary adapters are ligated onto the 3' ends and then the 5' ends of small RNA-enriched samples. After ligation a specific RT primer is used to make cDNA, which is then amplified with primers complementary to the ligated adapters. For barcoded libraries, indexed primers are used during the PCR step that add a six nucleotide sequence into the adapter sequence. Each sample is amplified with a different indexed primer, thus barcoding each biological sample. **B:** The indexed small RNA library is washed over the flow cell, which is covered with primers covalently attached to the flow cell surface, allowing library strands to attach to the flow cell. Sequencing is initiated by the binding of a sequencing primer complementary to the adapter sequence, and sequencing by synthesis follows. Once the sequencing reaction is complete, the synthesised strand is melted off of the template strand and washed away. At this point a primer complementary to the universal portion of the index is added, allowing the index to be sequenced. The synthesised strand is then again melted off. The template is flipped over and becomes attached to the flow cell surface by its formerly free end, allowing a sequencing primer complementary to the second adapter to bind and initiate another sequencing by synthesis reaction.

The library will bind to these complementary primers and because the concentration of the library fragments added to the flow cell is closely controlled, several free primers will surround each bound library strand. The bound library strands will form a bridge by attaching to a neighbouring primer complementary to the second adapter on its free end. When a mix of unlabelled nucleotides and enzyme is flowed over the flow cell, library strands are copied. The DNA on the flow cell is then denatured and the original library strand is washed away, leaving only the copy of the library strand which is covalently attached to the flow cell surface. The flow cell is then treated to disrupt the bridge of single stranded DNA on one side only, and several cycles of bridge amplification PCR is

performed resulting in “monoclonal” clusters of DNA generated from a single parent library strand. The different clusters cover the flow cell, which will be used to sequence the template library.

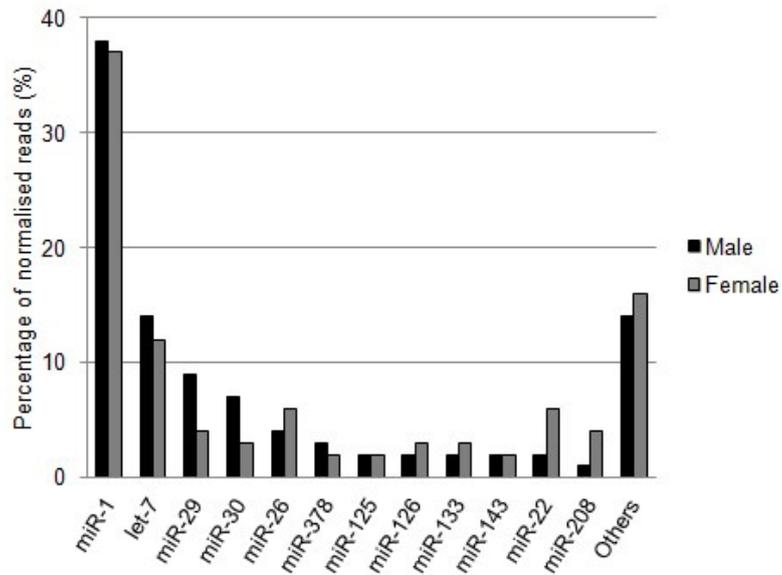
The template library on the HiSeq platform is sequenced by a method involving cyclic reversible terminators. A mixture of fluorescently labelled nucleotides, that have been chemically modified at their 3' end to prevent the addition of further nucleotides, is flowed over the flow cell<sup>50</sup>. This allows DNA polymerase to incorporate a nucleotide complementary to the template strand. The excess nucleotides are then washed away and the flow cell is imaged to detect which nucleotide was added. As the flow cell is covered in clusters from a single parent strand, the signal from a single cluster should be the same for each incorporation, therefore allowing the detection of the fluorescent signal. After the imaging step both the fluorescent dyes and terminating groups are cleaved off and washed away, and the cycle is repeated. The HiSeq platform is currently able to complete at least 150 cycles, although improvements are rapidly being made. Once 150 cycles of nucleotide incorporation has been completed, the DNA on the flow cell is denatured and each molecule is flipped over to become covalently attached to the flow cell by the adapter previously left free. Another 150 cycles of sequencing can then be done, giving results from both ends of the template strand. This is called paired end sequencing. Generally, 50 base pair (bp) paired end runs are more commonly used for small RNA sequencing.

Every flow cell contains eight separate lanes that allows the generation of up to eight separate template libraries. Furthermore, adapters can be made library-specific by the inclusion of a small identifying oligonucleotide that indexes each library<sup>65,66</sup>. These barcodes allow concurrent sequencing of templates generated from multiple libraries in a single flow cell lane, and can therefore greatly reduce the cost of a sequencing project. Once sequencing is complete, base calls are converted into sequence information corresponding to the library strand. The sequencing reads can then be aligned to a known reference genome, if one exists, and further analyses can be done to extract biologically relevant information.

### 3.1.5 Previous sequencing experiments

Chicken microRNA expression has been scrutinized by NGS in whole embryos at E11 (~HH37)<sup>67</sup>, in embryonic skeletal muscle at E10 (~HH36)<sup>68</sup>, in chicken liver at E15 (~HH41) and E20 (~HH45)<sup>69</sup>, in developing chicken adipocytes *in vitro*<sup>70</sup>, and in chicken embryonic fibroblasts *in vitro*<sup>71</sup>. Furthermore, NGS of somites from 3- to 5-day old chicken embryos affirmed the importance of tissue-specific sequencing in determining which microRNAs are highly expressed in such tissue, possibly indicating an important role in that tissue, and giving a starting point for the examination of the functions of the highly expressed microRNAs<sup>72</sup>. Although some highly expressed miRNAs such as miR-1 and miR-208 have been found to be essential for heart development or function, the NGS library preparation method is vulnerable to sequence-dependent ligation and PCR bias (discussed further in **3.3.5**). This work focussed on comparing expression profiles across different samples, as the ligation and PCR bias is most likely constant for each miRNA.

An Illumina GA NGS profiling experiment conducted in 2009 suggested that the 18 most highly expressed miRNAs in adult murine heart tissue account for more than 90% of cardiac miRNAs<sup>73</sup>. A summary of this profile, with the same miRNAs from different loci combined to simplify, is presented in **Figure 3.3**. In adult mouse hearts miR-1 was reported to be the most sequenced miRNA, taking up more than a third of all reads. Several of the miRNAs shown in **Figure 3.3** are known to be essential for heart development and function including miR-1<sup>4</sup>, miR-126<sup>74</sup>, miR-133<sup>4</sup>, miR-143<sup>75</sup>, and miR-208<sup>6</sup>.



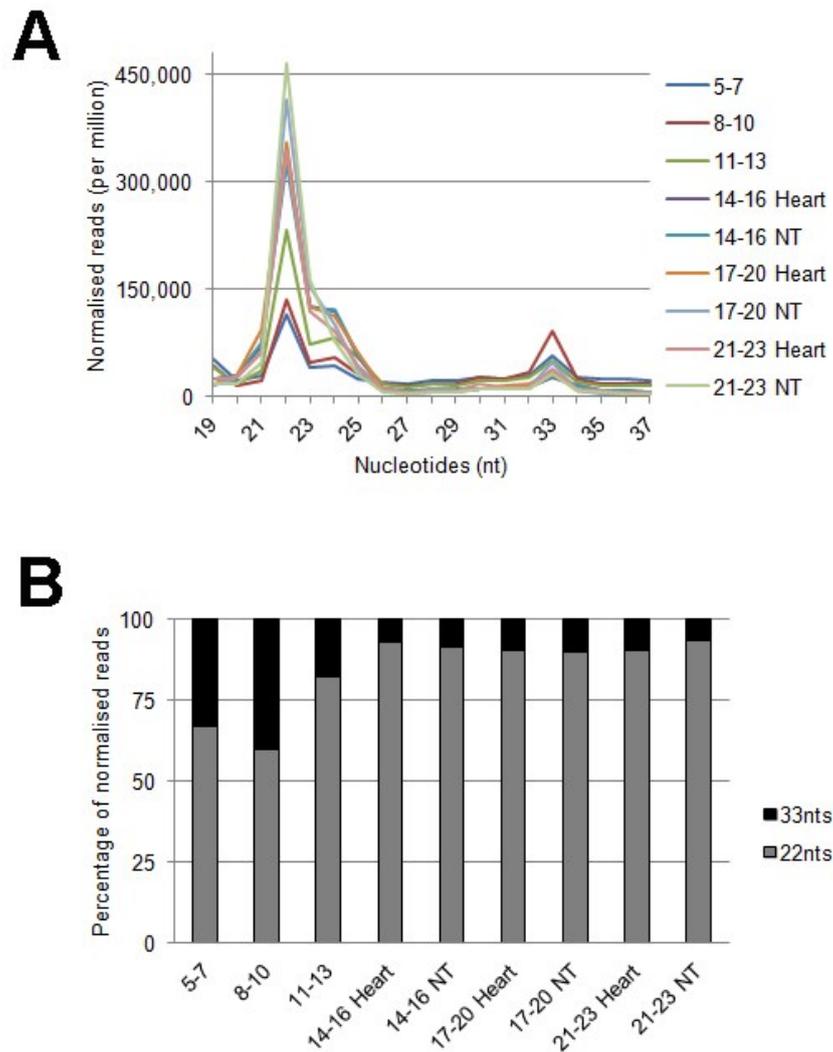
**Figure 3.3: Known miRNA profile in adult mouse hearts show a few miRNAs dominate sequencing space.** Summary of the percentage normalised reads taken up by the 12 most highly detected miRNAs in male and female adult mouse hearts, as determined by a sequencing study conducted by another group of researchers<sup>73</sup>.

However, very little is currently known about the specific miRNA expression profiles during the major events of early heart development, and elucidation of such a profile would assist in discovering miRNAs that may play an important role in forming the early heart.

## 3.2 Results

### 3.2.1 Illumina TruSeq libraries were enriched for 22-nucleotide RNAs

Simon Moxon converted raw sequencing FASTQ files to FASTA format and removed adapter sequences with custom Perl scripts. Only reads that mapped perfectly by PatMaN<sup>39</sup> to the chicken genome (Galgal4<sup>40</sup>) were retained. The vast majority of reads had a length of 22 nucleotides (**Figure 3.4**), indicating that the prepared libraries contained a fraction of RNA corresponding to miRNAs in size. A further peak, smaller but broadly around 33 nucleotides, most probably contains Piwi-interacting small RNAs (piRNAs)<sup>76</sup> that appear to be more highly expressed in the early stages of chick development (**Figure 3.4**).



**Figure 3.4: Size distribution of sequenced small-RNA enriched total RNA. A:** The peak at 22 nucleotides corresponds to mature miRNA length. A second, broader but much smaller peak at 33 nucleotides is most probably Piwi-interacting RNAs<sup>76</sup>. **B:** The percentage ratio of 22nt reads compared to 33nt reads for each sample shows the the relative peak size of the 33nt fraction is much larger in the HH5-7 and HH8-10 samples.

A comparison of the contribution of 22nt reads versus 33nt reads for each sample shows that 33nt sequences are more abundant in the early stages of chick development (**Figure 3.4B**). This profile is similar to that seen in the literature<sup>77</sup>, confirms that libraries containing mostly miRNA-sized sequences were successfully generated, and implicates Piwi-interacting RNAs in processes occurring between HH5 and HH8 in the chicken embryo (**3.3.1**).

### 3.2.2 Many animal miRNAs are expressed during chick cardiogenesis

Initially existing miRNAs from miRBase<sup>41</sup> release 19 were profiled across the developmental time series with the UEA Small RNA Workbench tool, miRProf<sup>42</sup>, by Simon Moxon. Subsequently, trimmed reads were profiled against existing miRNAs from miRBase release 20 by the author (2.8.2). miRProf searches for perfect matches to all animal miRNAs, allowing the identification of miRNAs that have yet to be annotated in some model organisms. The miRNA results from analysis of these NGS data with miRProf were compared to a list of known chicken miRNAs using a custom R script<sup>78</sup> written by the author, which resulted in the discovery of 31 known animal miRNAs not yet formally identified as chicken miRNAs in miRBase (Table Apx A.1). Three of these (miR-143, miR-182 and miR-363) were highly expressed in one or more samples (see 3.2.5 later). A list of the number of animal miRNAs that have not yet formally been added to miRBase as chicken miRNAs is given for each sample in the final column of Table 3.3.

**Table 3.3: Breakdown of the total number of distinct miRNAs matched to animal miRNAs in miRBase release 20 by sample.** The total number of known chicken miRNAs is shown in column three, with the percentage of the total number of known chicken miRNAs represented in each sample indicated in column four. miRBase release 20 contains a total of 636 known chicken miRNAs. Column five shows the number of miRNAs that match perfectly to miRNAs known in other animals that have not yet been added to the list of known chicken miRNAs.

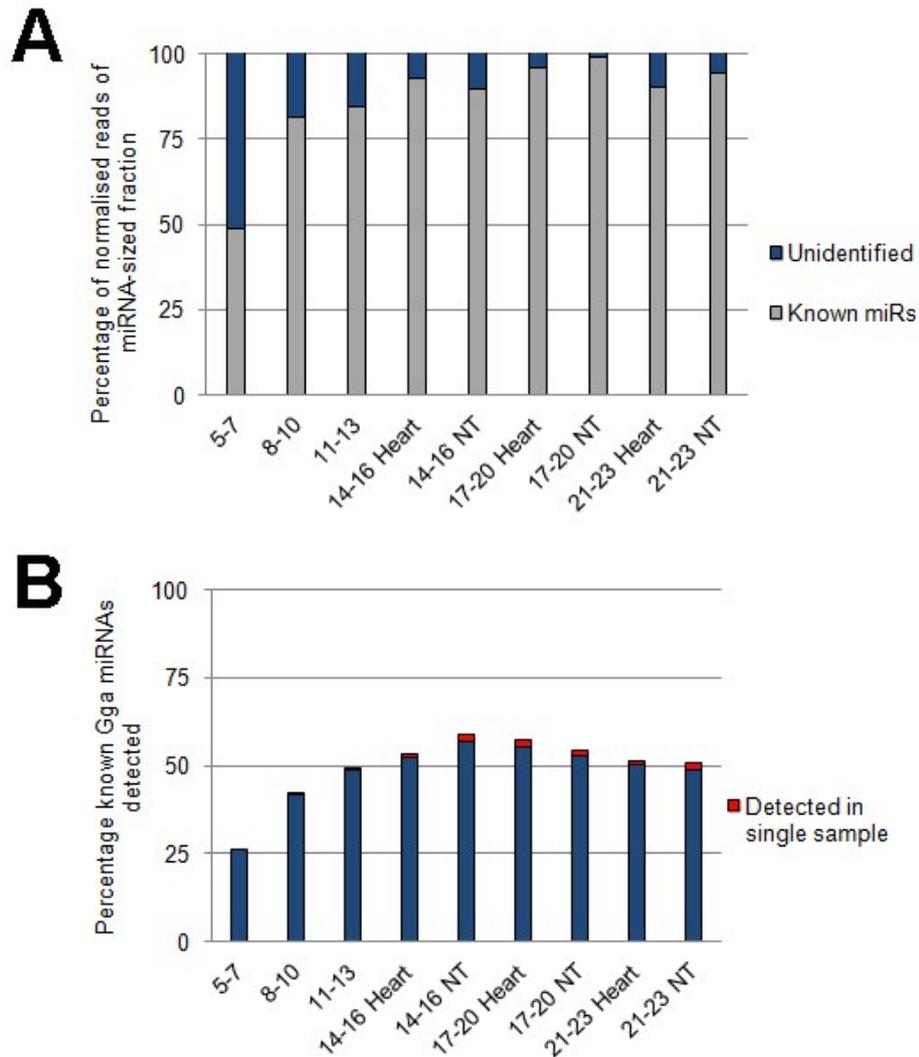
Sample	miRNAs by miRProf	Gga miRNAs	% of 636 known Gga miRNAs in sample	Non-Gga miRNAs in sample
5-7	183	165	26%	18
8-10	282	264	42%	18
11-13	327	309	49%	18
14-16 Heart	355	333	52%	22
14-16 NT	385	363	57%	22
17-20 Heart	375	351	55%	24
17-20 NT	357	336	53%	21
21-23 Heart	342	321	50%	21
21-23 NT	332	311	49%	21

The discovery of conserved animal miRNAs expressed in the chicken embryo not previously annotated adds several miRNAs that can be added to the miRBase chicken miRNA database. As discussed in 3.3.2, two of these miRNAs are known to be essential

for heart development (miR-143) and are upregulated in failing human hearts (miR-182), making their discovery in the chicken embryo very interesting.

### **3.2.3 A large fraction of known chicken miRNAs are expressed from gastrulation**

The total normalised reads in the miRNA-sized fraction of each sample was mostly assigned to known animal miRNAs (**Figure 3.5A**), apart from the 51% of reads not assigned in sample HH5-7. Only 26% of the 636 known chicken miRNAs were expressed during early development (HH5-7, **Table 3.3**). However, the percentage of total chicken miRNAs increased rapidly with development to reach a peak of 57% in the HH14-16 NT sample. This increase in the number of miRNAs detected as development proceeds has subsequently been found by other researchers<sup>77</sup>, and indicates that a large number of known chicken miRNAs are already expressed in HH5-7 embryos.



**Figure 3.5: Known miRNAs account for most of the reads from the miRNA sized fraction of almost all the samples, and many known chicken miRNAs were detected. A:** The percentage of normalised reads from the miRNA sized fraction that were found to be due to known animal miRNAs is shown in grey, with the remaining unlabelled reads shown in blue. **B:** The percentage of known chicken miRNAs detected in each sample. Early in development (HH5-7) only 26% of the 636 known chicken miRNAs were detected. However, the fraction of known chicken miRNAs detected soon rose to a maximum of 57% (HH14-16 NT). A small portion of miRNAs were detected in only one of the nine samples at low levels, as indicated by the red caps on the appropriate bars.

A small number of miRNAs were detected at only one time point, always at less than 1 read per million (**Figure 3.5B**). The average number of reads for miRNAs detected at more than one time point ranged from 1,418 (median 4.19) to 2,587 (median 1.95) reads per million. The percentage of miRNAs only present in one sample varied from 0% in HH5-7 to 2.4% in HH17-20 Heart. A complete list of miRNAs detected at single time points, along with their raw reads numbers, can be found in **Table Apx A.2**. In total 63 miRNAs were detected at single time points. Although the detection of these miRNAs at specific

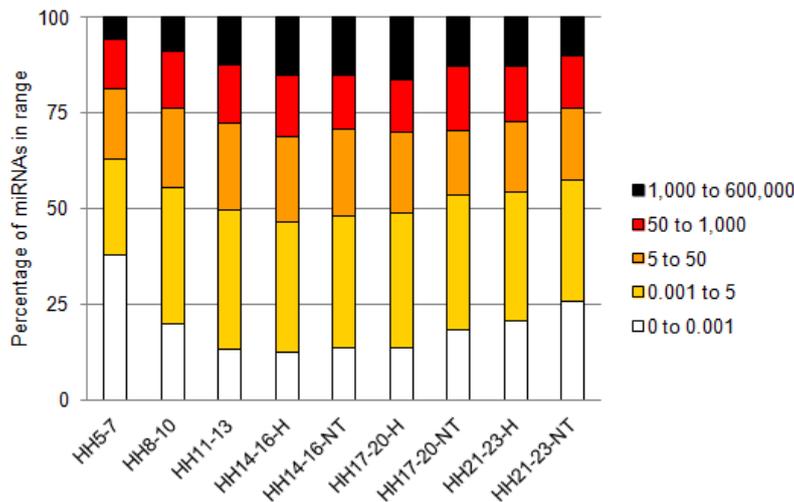
stages could indicate that they have functions particular to those stages, only five were found to be conserved in humans and therefore further investigations were focussed on miRNAs that could be profiled over several stages (3.3.3).

#### 3.2.4 Samples cluster according to their known animal miRNA expression

To determine whether miRNA expression profiles across all stages assayed showed gradual changes, the correlation of samples based on their known animal miRNA expression profiles, as matched by miRProf<sup>42</sup> to miRBase<sup>41</sup> release 20, was assessed using a custom R<sup>78</sup> script written by the author and is displayed as a heatmap in **Figure Apx A.1**.

Early samples (HH5-7, HH8-10 and HH11-13) clustered together, and were more closely related to other slightly older stages (HH14-16 Heart, HH14-16 NT and HH17-20 Heart) than to the oldest stages. The later heart sample, HH21-23 Heart, clustered together with the older “neural tube” samples (HH17-20 NT and HH21-23 NT). This indicates that most miRNA levels changed gradually as development proceeded, as was expected from the mixed temporal nature of the dissected tissues.

Sorting miRNAs by normalised read size gives a summary of the levels at which most miRNAs were detected (**Figure 3.6**). Many miRNAs were detected at quite low levels between 0 and 0.001 reads per million, particularly at the earliest time point (HH5-7, 38%). More than 25% of miRNAs were detected at between 0.001 and 5 reads per million, while between 6% (HH5-7) and 16% (HH17-20 Heart) of miRNAs had levels ranging from 1,000 and 600,000 reads per million.



**Figure 3.6: Representation of five ranges of known animal miRNA normalised reads.** The percentage of total miRNAs detected broken down by their normalised read numbers. Less than 17% of miRNAs detected had normalised read numbers in excess of 1,000 reads per million. The largest fraction (more than 25% in all samples) had normalised reads between 0.001 and 5 reads per million. White: 0 to 0.001 reads per million. Yellow: 0.001 to 5 reads per million. Orange: 5 to 50 reads per million. Red: 50 to 1,000 reads per million. Black: 1,000 to 600,000 reads per million.

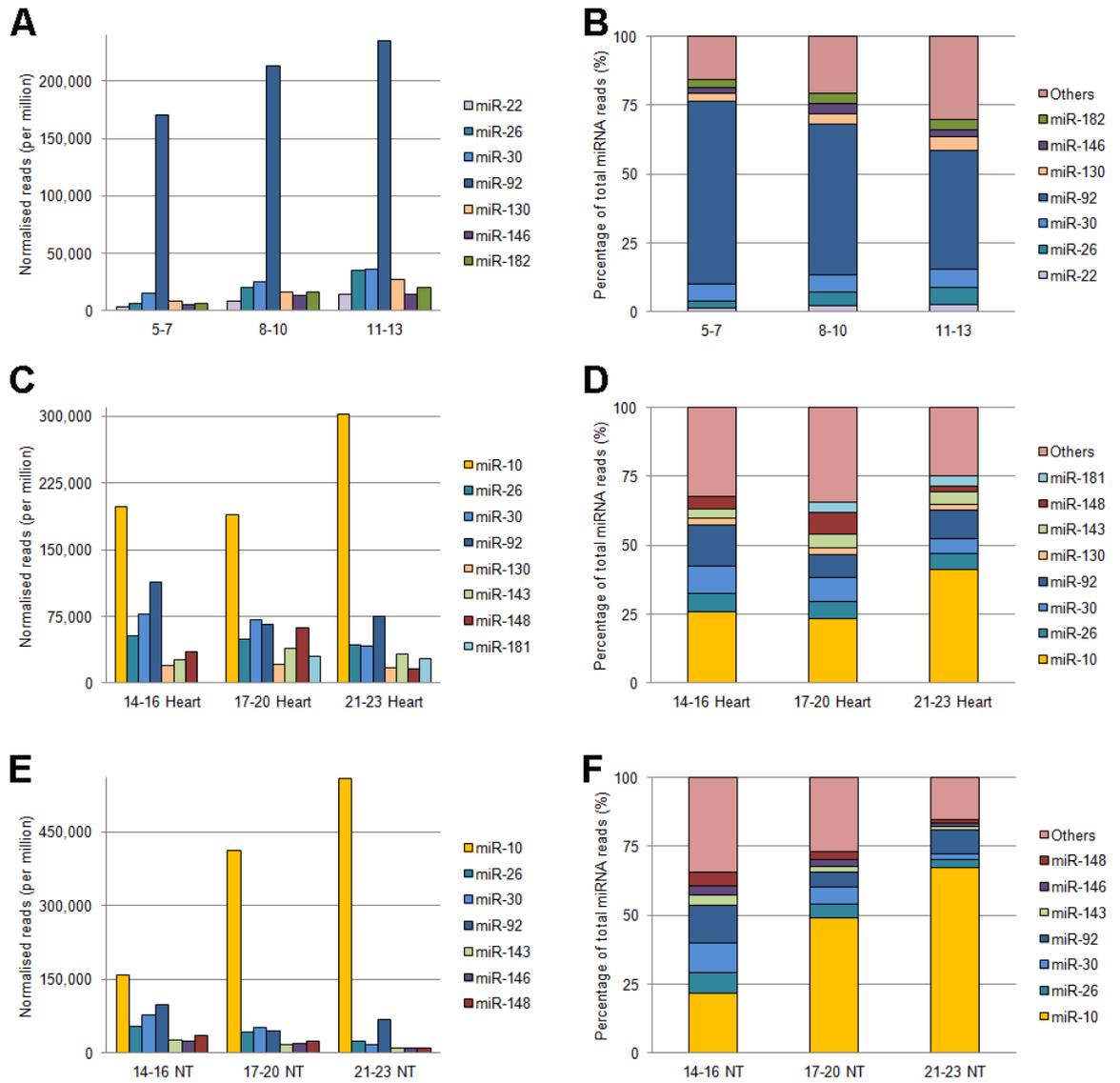
Previous sequencing experiments on adult mouse hearts found that 18 miRNAs represent 90% of the total sequencing reads. As discussed in 3.3.5 this abundance of very few miRNAs may be a result of ligation and PCR bias that greatly influenced earlier sequencing experiments, and has not been eliminated yet. Since the detected normalised read levels of a miRNA may be influenced by ligation or PCR bias, it is difficult to draw conclusions of function based on such reads. A better strategy comprises focussing on patterns of change over several time points, as the bias should remain constant for each miRNA. However, an analysis of the top expressed miRNAs (3.3.5 and 3.2.6) highlighted several predicted targets that are common to different miRNAs found to be highly expressed.

### 3.2.5 miRNA expression is dominated by two miRNAs

Ranking known animal miRNAs by normalised read number revealed a single miRNA filled a large fraction of the sequencing space of each sample. A distinct change in miRNA expression pattern was observed between HH11-13 and HH14-16.

**Figure 3.7B** shows that during early stages in samples HH5-7, HH8-10 and HH11-13, miR-92 represented between 66% and 43% of the total known animal miRNA reads. Initially 84% of the total known miRNA reads were taken up by only seven miRNAs, but in HH11-13 this decreased to 70%. However, from HH14-16 miR-10 became the most ubiquitous miRNA in both heart and non-heart tissue samples (**Figure 3.7C** and **Figure 3.7D**), with miR-92 relegated to second or third place. Although miR-10 was the most sequenced miRNA in the HH14-16 H and HH14-16 NT samples, taking up about 25% of the total known animal miRNA reads, the increase in its expression was much larger in the non-heart tissue samples (**Figure 3.7F**). Investigations into known miR-10 functions revealed its intimate involvement in vasculogenesis and angiogenesis, processes that occur in both developing cardiac and neural tissue (**3.3.5**), explaining the large upregulation of this miRNA observed here.

miR-130 is dysregulated in diabetic mouse hearts<sup>79</sup> and consistently featured in the top 10 miRNAs by normalised expression in the early samples and in the later heart samples, although not in the later NT samples. Therefore this miRNA was selected for further biological investigation as described in **4.2.7**.



**Figure 3.7: miRNA expression was dominated by two miRNAs and three miRNAs were consistently expressed at a very high level in all nine samples. A:** The top miRNAs detected in samples HH5-7, HH8-10 and HH11-13 were dominated by miR-92. **B:** Although the total number of reads for miR-92 increases across these stages, the reads as a percentage of total miRNA reads decreased indicating that miRNA expression increased during these stages of development. **C,E:** miRNA-10 dwarfed the expression of other miRNAs in later heart and neural tube samples (HH14-16, HH17-20 and HH21-23 Heart and NT). **D,F:** miRNA-10 expression increased across these stages, especially in the NT samples where it eventually accounted for 68% of the total miRNA reads in the HH21-23 NT sample.

### 3.2.6 Enrichment of predicted target genes of top expressed miRNAs

Although the top ten miRNAs expressed in each sample stay mostly constant, an analysis was done to check whether any miRNAs not common to all the samples are still predicted to target the same genes. The overlap of predicted target genes of the top expressed miRNAs unique to three groups of samples, early development (HH5-7) and later development heart (HH14-23 H) and other tissues (HH14-23 NT), was assessed by comparison using custom R scripts written by the author.

The majority of top expressed miRNAs found in samples taken during the early stages of development (HH5-7, HH8-10, HH11-13) were common to all three samples.

**Table 3.4: Comparison of predicted target genes of top expressed miRNAs for samples HH5-7, HH8-10 and HH11-13.**

Samples compared	HH5-7	HH8-10	HH11-13
Number of predicted targets of top 10 expressed miRNAs	3,081	3,342	3,278
Number of top miRNAs unique to sample	1	0	2
Number of predicted targets of top miRNA(s) unique to sample	16		410
Genes targeted by unique miRNA(s) only	5		68
% Genes covered by other top expressed miRNAs	69%		83%

HH5-7 had one unique top expressed miRNA not found in the other two samples: miR-2954. miRBase release 20 contains only two entries for this miRNA, one for chicken and one for zebra finch. However (**Table 3.4**) shows that 69% of the predicted target genes of miR-2954 were also covered by the remaining top miRNAs, which were common to HH8-10 and HH11-13. Two top expressed miRNAs in sample HH11-13, miR-10 and miR-148, were not found in the top 10 miRNAs of either other sample. These miRNAs were predicted to target 410 genes, 83% of which were also targeted by the other common top miRNAs (**Table 3.4**).

During later stages of heart development one unique top expressed miRNA could be found in each of the three samples assayed (HH14-16 H, HH17-20 H and HH21-23 H).

**Table 3.5: Comparison of predicted target genes of top expressed miRNAs for samples HH14-16 Heart, HH17-20 Heart and HH21-23 Heart.**

Samples compared	HH14-16 H	HH17-20 H	HH21-23 H
Number of predicted targets of top 10 expressed miRNAs	3,480	3,046	3,379
Number of top miRNAs unique to sample	1	1	1
Number of predicted targets of top miRNA(s) unique to sample	112	6	28
Genes targeted by unique miRNA(s) only	37	1	11
% Genes covered by other top expressed miRNAs	67%	83%	61%

Even though miR-22, miR-126 and miR-100 were not common to all three samples, between 61% and 83% of their predicted target genes were also predicted to be targets of the mostly highly expressed miRNAs common to all three samples (**Table 3.5**).

Overall non-heart samples HH14-16 NT, HH17-20 NT and HH21-23 NT each showed high expression of a single unique miRNA, miR-22, miR182 and miR-100 respectively.

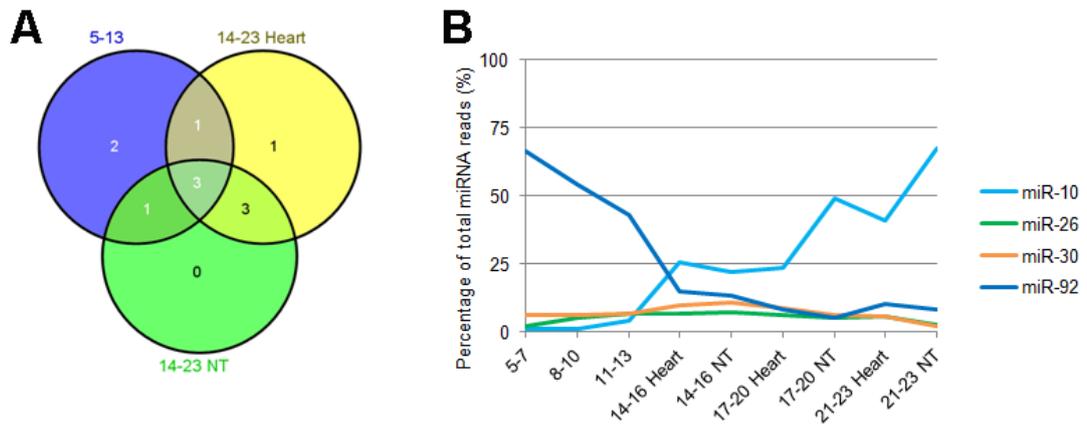
**Table 3.6: Comparison of predicted target genes of top expressed miRNAs for samples HH14-16 NT, HH17-20 NT and HH21-23 NT.**

Samples compared	HH14-16 NT	HH17-20 NT	HH21-23 NT
Number of predicted targets of top 10 expressed miRNAs	2,895	3,522	2,771
Number of top miRNAs unique to sample	1	1	1
Number of predicted targets of top miRNA(s) unique to sample	112	1121	28
Genes targeted by unique miRNA(s) only	36	567	11
% Genes covered by other top expressed miRNAs	68%	49%	61%

Although these miRNAs were unique when the top detected miRNAs of these three samples were compared, miR-22 is also a top expressed miRNA of sample HH14-16 H and miR-100 was also highly expressed in sample HH21-23 H. **Table 3.6** confirms that as in the matched heart samples (**Table 3.5**), more than 61% of the genes predicted to be targeted by miR-22 and miR-100 were also predicted targets of the miRNAs common to all three NT samples. Almost half of the predicted target genes of miR-182, a top expressed miRNA in sample HH21-23 NT only, were also predicted to be targeted by the remaining top expressed miRNAs found in all three samples. This suggests that the top expressed miRNAs may be functioning in concert to ensure the repression of a core set of targets.

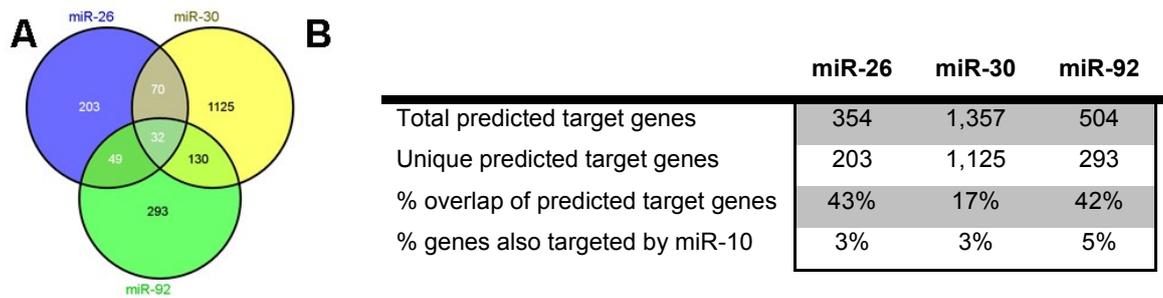
A comparison of the ten most highly detected miRNAs in all nine samples revealed that three miRNAs were common to all the samples (**Figure 3.8A**). Just over 40% of the predicted target genes of two of these common miRNAs, miR-26 and miR-92, were predicted target genes of one or both of the other two common miRNAs (**Figure 3.9**). However, almost four fifths of the predicted target genes of the third common top expressed miRNA, miR-30, were not common to miR-26 or miR-92. This analysis suggests that miR-26 and miR-92 may function redundantly to ensure the repression of their overlapping targets, whereas miR-30 does not (**3.3.6**).

As noted in **3.2.5**, miR-10 rapidly dominates the sequencing space from HH14-16 (**Figure 3.8B**). The overlap between predicted target genes of the three common highly expressed miRNAs and those of miR-10 is very low (**Figure 3.9**), ranging from 3% to 5%.



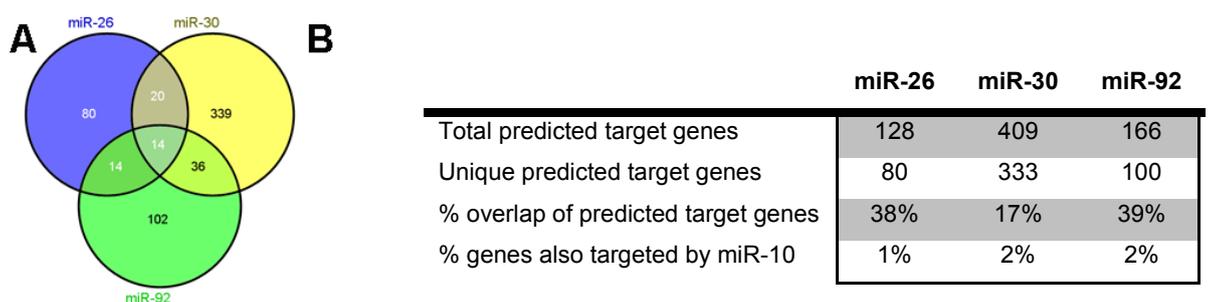
**Figure 3.8 Three top expressed miRNAs are common to all samples and miR-10 expression increases rapidly during development.** **A:** Venn diagram showing that three top miRNAs were common to all samples. **B:** Percentage of total reads by the three top miRNAs that were common to all samples: miR-26, miR-30 and miR-92. miR-92 initially dominated with 66% of total miRNAs in HH5-7, but quickly dropped to below 15% in older embryos. miR-26 and miR-30 both maintained relatively constant expression levels across all stages. miRNA-10 was rapidly upregulated from HH11-13.

As discussed in **3.3.5**, miR-10 is likely functioning to ensure normal vasculogenesis and angiogenesis in both cardiac and neural tissue and the rapid increase in its expression reiterates its importance in these tissues.



**Figure 3.9: Analysis of predicted target genes of the three top expressed miRNAs common to all nine samples.** A large portion of the predicted target genes of miR-26 are also predicted to be targeted by miR-92 and vice versa. However, only 17% of the predicted target genes of miR-30 are targeted by either miR-26 or miR-92.

The predicted target genes of miR-26, miR-30, miR-92 and miR-10 were filtered by genes known to be expressed in the human heart<sup>80</sup>. The Pattern Gene Database (PaGenBase) contains 7,187 genes known, collected from 26 array and NGS experiments, to be expressed in the human heart. Although the number of predicted target genes of miR-26, miR-30, miR-92 and miR-10 was more than halved, the extent to which targets of miR-26 and miR-92 were covered by one or more of the others remained very similar to that seen before filtering the predicted genes (**Figure 3.10B**, row 3). The percentage of shared targets between miR-10 and any of miR-26, miR-30 or miR-92 was very low (**Figure 3.10B**, row 4) after only considering predicted target genes known to be expressed in the heart.



**Figure 3.10: A large fraction of the predicted targets of miR-26 and miR-92 that are expressed in the human heart are also predicted targets of the other two highly detected miRNAs.** The overlap in predicted targets seen between miR-26, miR-30 and miR-92 remains of a similar size whether all predicted genes (**Figure 3.9**) or just those predicted genes known to be expressed in heart are considered (row three). A very small percentage of the predicted target genes of miR-26, miR-30 and miR-92 are also predicted target genes of miR-10 (row four).

A full list of the overlapping predicted targets that are also expressed in the human heart is shown in **Table Apx A.3** and **Table Apx A.4**. A core of 14 genes were found to be predicted targets of miR-26, miR-30 and miR-92. However, A GO terms analysis of these 14 genes did not return any significant associations confirming that the overlapping targeting of these genes is not connected to a single pathway or developmental process.

### 3.2.7 Unique predicted target genes of top expressed miRNAs

Many of the predicted target genes of the top expressed miRNAs miR-26, miR-30 and miR-92 were found to be targeted by at least two of these miRNAs. In order to assess whether these miRNAs target any unique genes or pathways, an analysis was performed on the predicted target genes unique to each of these three miRNAs.

To determine which target genes or pathways were targeted uniquely by top expressed miRNAs at each stage, the predicted target genes and their significant Gene Ontology (GO) terms<sup>81</sup> of the top ten sequenced miRNAs for each of the nine samples were identified. Custom R scripts written by the author were used to isolate those GO terms unique to three groups of samples: HH5-13 (HH5-7, HH8-10, HH11-13), HH14-23 H (HH14-16 H, HH17-20 H, HH21-23 H), and HH14-23 NT (HH14-16 NT, HH17-20 NT, HH21-23 NT).

**Table 3.7: Significant GO terms unique to three groups of samples:** early stages (HH5-7, HH8-10, HH11-13), later stages heart tissue (HH14-16 H, HH17-20 H, HH21-23 H) and later stages non-heart tissue (HH14-16 NT, HH17-20 NT, HH21-23 NT). The genes associated with the GO term are also shown.

Samples	Significant GO term	Genes
HH5-13	Signalling by TGFβ (REAC:170834)	SKI, SMAD2, SMAD6, SMAD7, SMURF1, TGFB2, ZFYVE9
HH14-23 Heart	Congenital malformation of the great arteries (HP:0011603)	ABCC9, ADAMTS18, BCOR, CD96, CHD7, COMT, DHCR24, DTNA, EP300, FANCF, FBN1, FBN2, FOXF1, GATA6, GJA1, GLI3, GPC3, KAT6B, MED13L, MYH11, NFIX, NPHP3, NSD1, PTPN11, RAB23, SALL1, SLX4, SMAD3, SOX2, TBX1, TP63, VANGL1, ZEB2, ZIC3, ZMPSTE24
HH14-23 NT	Positive regulation of cardiac muscle hypertrophy (GO:0010613)	EDN1, IGF1, AKAP6, MTPN, CAMK2D, PDE5A, ADRB1, IL6ST

Although there is large overlap between the targets of the top expressed miRNAs across all the sample groups, three unique and interesting significant groups of targets were identified from the short list of significant GO terms unique to each group (**Table 3.7**). The unique targets of the early stages group are involved in TGF $\beta$  signalling, whereas those of the later stages heart group have been associated with congenital malformations of the great arteries. Finally unique predicted target genes in the later stages non-heart tissue group have a significant association with positive regulation of cardiac muscle hypertrophy. Insulin-like growth factor 1 (IGF1) causes cardiac hypertrophy as a result of the activation of the phosphoinositide 3-kinase (PI3K) / AKT pathway, which is normally repressed by phosphatase and tensin homolog (PTEN)<sup>82</sup>. The effect of IGF1 is mediated through Parkinson protein 7 (PARK7) – induced inhibition of PTEN phosphatase in skeletal muscle<sup>83</sup>. Furthermore, Muscle A-kinase anchoring proteins (mAKAPs) are encoded by the *AKAP6* gene. Following their binding to calcineurin, NFATc3 is dephosphorylated and the resulting activation of NFAT transcriptional activity leads to cardiac myocyte hypertrophy<sup>84</sup>. Additionally, pathological hypertrophy has been linked to Ca<sup>2+</sup> / calmodulin-dependent protein kinase 2 delta (CAMK2D) and interleukin-6 signal transducer (IL6ST) function<sup>85,86</sup>, and endothelin 1 (EDN1) variations have been associated with an increased risk of developing left ventricular hypertrophy<sup>87</sup>.

This analysis of the most highly detected miRNAs during chick embryonic development has highlighted several essential roles played by these miRNAs, reinforcing their importance and generating a profile of their expression in the developing embryo not previously known (**3.3.6**). The most exciting finding relates to the large amount of overlap of predicted target genes of these highly expressed miRNAs, suggesting that these miRNAs may function to redundantly repress a core set of targets.

### 3.3 Discussion

#### 3.3.1 Illumina TruSeq libraries were enriched for 22-nucleotide RNAs

RNA sequencing libraries were successfully prepared from samples enriched for sequences 22nt long, containing mature miRNA transcripts (**Figure 3.4A**). A second broad peak around 33nts in size is consistent with sequencing results from other experiments<sup>77</sup> and is most likely due to piRNAs. Piwi-interacting RNAs are not processed by Dicer during their biogenesis and are largely known for their roles in epigenetic changes and maintenance of the gonads and germ cells of vertebrates<sup>77</sup>. The number of reads recorded at 33nts in the least developed samples, HH5-7 and HH8-10, are more than 33% of that recorded in the 22nt miRNA fraction, suggesting that the tissues captured at these stages express piRNAs at a high level. The focus of this work is the miRNA expression profile of the sequenced tissues, but this dataset could be probed for information on piRNA expression during the early stages of heart development in the future.

#### 3.3.2 Many animal miRNAs are expressed during chick cardiogenesis

The miRNA-sized fraction of almost all the samples contained mostly known animal miRNAs (**Figure 3.5A**), suggesting that most of the miRNAs expressed in the developing chick have been identified in at least one other animal. However, more than half of the reads in this miRNA-sized fraction in sample HH5-7 were not identified. At this stage it is not clear whether these unidentified sequences are novel miRNAs or fragments of degraded RNA. An assessment of the secondary structure of these sequences and their genomic surroundings may allow better classification of the unidentified sequences. However, the novel miRNA prediction performed by Simon Moxon (**Chapter 5**) did not suggest that this sample contain a large number of novel miRNAs and therefore it was considered unlikely that these sequences would be of interested in this study.

miRProf matching of the normalised reads from this dataset to all the animal miRNAs contained in miRBase release 20 revealed a total of 31 animal miRNAs not currently known in chick. Two of these, miR-143 and miR-182, are very highly expressed in the

developing chick embryo (**Figure 3.7**). miR-143 is known to play an important role during heart development in zebrafish<sup>75</sup> and miR-182 is upregulated in human patients with coronary artery disease<sup>88</sup>. Although not formally annotated yet, these 31 animal miRNAs match perfectly to those described and will be submitted for inclusion into the miRBase database.

### **3.3.3 A large fraction of known chicken miRNAs are expressed from gastrulation**

During the earliest stages profiled only 26% of the 636 known chick miRNAs available in miRBase release 20 were expressed, but this number rapidly increased to around 50% in the older samples (**Figure 3.5B**). These results are consistent with an increase in miRNA expression as development progresses which has been reported in the chick<sup>77</sup>, where small RNAs from repetitive elements were found to be more abundant than miRNAs before gastrulation. A small fraction (63 out of 636) of these miRNAs were detected at single time points, always at low levels, and only 5 are conserved in *Homo sapiens*. This implies that these miRNAs have limited biological roles, specifically in the context of discovery of potentially important human heart development regulators.

### **3.3.4 Samples cluster according to their known animal miRNA expression**

The normalised reads of the known animal miRNAs identified from miRBase release 20 were used to prepare a clustered heatmap in R (**Figure Apx A.1**). An analysis of the five ranges of normalised reads values represented as different colours on the heatmap showed that many miRNAs were detected at less than 0.001 reads per million, particularly at the earliest time point (HH5-7) where 38% of detected miRNAs fell into this lowest expression range (**Figure 3.6**). Few miRNAs (6% to 16%) had levels in the highest range of 1,000 to 600,000 reads per million. The largest fraction of miRNAs in each sample had levels between 0.001 and 5 reads per million, suggesting that the majority of miRNAs are expressed at low levels during the early stages of chick development. Based on normalised reads per million, the nine samples clustered into three groups: the early stages (HH5-7, HH8-10, HH11-13), and two later stages (HH14-16 H, HH14-16 NT, HH17-20 H and HH17-20 NT, HH21-23 H, HH21-23 NT). Interestingly the most developed heart tissue

sample was more similar to the older non-heart tissue samples than the younger heart tissue samples. This suggests a change in miRNA expression that might be very relevant to the developmental changes exclusive to this range of stages. For instance, during the earlier stages the cardiac progenitors are specified and start to differentiate whereas valve development and ventricular septation is known to take place in the heart from HH21 to HH23. The similarity in miRNA profiles of the HH14-16 and HH14-16 NT samples may be due to the contribution, as discussed in 4.2.3, of the cardiac neural crest which migrates through the tissue dissected in the HH14-16 NT sample (**Figure 1.2B**).

### 3.3.5 miRNA expression is dominated by two miRNAs

Two miRNAs were found to dominate over all other known animal miRNAs, with a distinct change in dominating miRNA observed between HH11-13 and HH14-16. In the early stages (HH5-7, HH8-10, HH11-13) miR-92 is the most abundant miRNA detected, taking up between 43% and 66% of the total known miRNA reads (**Figure 3.7A, B**). However, at HH14-16 miR-92 moves down to the second or third place, replaced by miR-10 as the most sequenced miRNA (**Figure 3.7C, D, E and F**). Although miR-10 is the most detected miRNA in all the older samples, it was found to increase more rapidly in non-heart NT samples than in heart samples. In the oldest samples, HH21-23 H and HH21-23 NT, miR-10 occupies 41% and 68% of the known miRNA reads respectively. miR-10 has known roles in the regulation of blood vessel formation, which takes place in the tissue dissected in Heart and NT samples.

Blood vessels are formed by either vasculogenesis, whereby the mesoderm gives rise to blood vessels following differentiation of cells into the endothelial lineage, or by angiogenesis, wherein new blood vessels form from existing vessels by proliferation<sup>89</sup>. VEGF signalling is essential for neurovascular development, during which blood vessels ingress into the neural tube<sup>90</sup>, as well as during vasculogenesis<sup>89</sup>. miR-10 is upregulated in hematopoietic stem cells<sup>91</sup> and has pro-angiogenic functions in zebrafish and human endothelial cells through its effects on Vascular Endothelial Growth Factor (VEGF) signalling<sup>92</sup>. Finally, the confirmed targets of miR-10 have four significant GO terms associated with vasculogenesis and angiogenesis: vasculature development (GO:0001944),

blood vessel development (GO:0001568), blood vessel morphogenesis (GO:0048514) and angiogenesis (GO:0001525). Given that vasculogenesis and angiogenesis are such essential parts of heart development and that miR-10 is intimately involved in regulating both processes, it is not surprising that miR-10 was found to be expressed at such a high level during these stages of embryonic development.

Interestingly a miRNA known to be dysregulated in diseased mouse hearts<sup>79</sup>, miR-130, was detected in the top 10 miRNAs by normalised expression in the early samples (HH5-7, HH8-10, HH11-13) and in the later heart samples (HH14-16 H, HH17-20 H, HH21-23 H). Although the effect of miR-130 on heart development is unstudied in current literature, this study suggested several tantalising targets of this miRNA when assayed by reporter assays (**Figure 4.13C, D, G, H, I, L**). The potential impact of miR-130 on these targets during heart development is discussed in **4.3.7**.

Similarly to a sequencing experiment conducted on adult mouse heart tissue previously<sup>73</sup>, a few miRNAs were found to dominate the sequencing space in all nine embryonic samples. In this study the contributions of the top seven miRNAs were found to total between 84% (HH5-7) and 65% (HH17-20 H) of the total normalised reads of all known animal miRNAs identified in the samples (**Figure 3.7B, D and F**). However, the identities of the highly expressed miRNAs differed from those published previously, with only 4 highly expressed miRNAs in common between the studies: miR-143, miR-22, miR-26 and miR-30. miR-208, highly detected in the adult mouse heart, has not been annotated in the chicken genome and no exact match to this miRNA was discovered by miRProf in the developing chick tissues. This miRNA is currently thought to be the only heart-specific microRNA<sup>93</sup>, and has been shown to play a role in cardiac stress response as well as in repressing the expression of skeletal muscle genes in the postnatal heart, rather than in the development of the heart<sup>6,94</sup>. In keeping with the current literature<sup>94</sup> that points to miR-208 being expressed after birth, miR-208 was not detected at any of the time points assayed.

In theory, the abundance of a specific microRNA in a sequencing library should be similar to the abundance of that microRNA in a biological sample<sup>95</sup>. However, proprietary Illumina adapter ligation has been shown to be biased<sup>64,96</sup>, and since only adapter ligated

molecules can be sequenced any bias in their ligation will give a skewed view of miRNA expression. The preference of truncated T4 RNA ligase 2 (Rnl2), used to ligate pre-adenylated adapters to the 3' ends of RNA molecules, for double stranded nucleotides upstream and single stranded nucleotides downstream of ligation sites is well known<sup>64,97</sup>. After ligation of the 3' adapter, a 5' adapter is added by T4 RNA ligase 1 (Rnl1), which favours single stranded substrates<sup>64,97</sup>. A recent literature search failed to reveal other published Illumina HiSeq profiles of miRNA expression in heart, but it is possible that the profiles of most frequently detected miRNAs obtained in the current study reflect an improvement in library construction protocol, compared to the profile obtained in adult mouse heart during the very early stages of Illumina GAI sequencing<sup>73</sup>, which has led to a decrease in sequencing bias, possibly by modest changes to adapter sequence or ligation conditions<sup>64,97</sup>.

### 3.3.6 Enrichment of predicted target genes of top expressed miRNAs

Further analysis of the top expressed miRNAs in the nine samples was done by splitting the samples into three groups: early (HH5-7, HH8-10, HH11-13), late heart (HH14-16 H, HH17-20 H, HH21-23 H) and late non-heart (HH14-16 NT, HH17-20 NT, HH21-23 NT). The majority of the top expressed miRNAs in samples corresponded to those expressed in the other members of the group as summarised in **Table 3.4**, **Table 3.5** and **Table 3.6**. In order to assess whether the miRNAs unique to samples when compared with the other members of the group had different or similar roles to the common miRNAs, predicted target genes of the unique miRNAs were evaluated. It was shown that the vast majority of predicted target genes of the unique miRNAs were also predicted target genes of the common miRNAs, suggesting that the top expressed miRNAs in each group act together to affect essential targets. It is likely that these miRNAs play many essential roles during these stages of development, and further bioinformatic analysis of the shared targets might give a plausible starting point in unravelling their many functions.

Three miRNAs were consistently expressed at very high levels in all nine samples assayed: miR-26, miR-30 and miR-92 (**Figure 3.7G, H**). Components of the miR-17-92 cluster have been shown to be essential for cardiomyocyte proliferation in adult mice<sup>98</sup>, miR-26

has important functions as a regulator of potassium current changes in mice<sup>99</sup> and hypertrophy<sup>100</sup>, and miR-30 is pro-angiogenic in zebrafish embryos<sup>101</sup> whereas miR-17 and miR-20 repress angiogenesis<sup>102</sup>. An analysis of the predicted target genes of these three miRNAs showed that more than 40% of the predicted targets of miR-26 and miR-92 are also predicted target genes of the others (**Figure 3.9**). However, only 17% of the predicted target genes of miR-30 are also targeted by either or both of miR-26 and miR-92. Analysis of the shared predicted target genes revealed a large number of significant, though broad, GO terms including cell differentiation (GO:0030154), nervous system development (GO:0007399) and cardiovascular system development (GO:0072358), confirming the many and varied essential roles these miRNAs are known to play during development and in the adult organism.

The degree of overlap of all the predicted targets of miR-26, miR-30 and miR-92 was maintained when only genes known to be expressed in human heart tissue were considered (**Figure 3.9B**, **Figure 3.10B**). This strongly suggests that miR-26 and miR-92 are working in concert to target a core set of genes. A FunDo analysis of the overlapping genes known to be expressed in the human heart (**Table Apx A.3**, **Table Apx A.4**) revealed 25 genes significantly associated with disease (**2.8.6**, **Table Apx A.5**). Two of these diseases affect the heart (**Table 3.8**). Atherosclerosis, the hardening of the arteries, can lead to further pathology and death in diseased hearts<sup>103</sup>.

**Table 3.8: Overlapping predicted target genes, known to be expressed in the heart, of at least two of miR-26, miR-30, and miR-92 are significantly associated with disease of the heart, amongst other diseases.**

Disease	Overlapping predicted target genes of miR-26, miR-30 and miR-92
Atherosclerosis	PDE4D (miR-26 & miR-30) ATXN1 (miR-30 & miR-92) CACNA1C (miR-26, miR-30 & miR-92) BDNF (miR-10 & miR-30)
Heart failure	ADRB1 (mir-30 & miR-92) LOX (miR-26 & miR-30)

This analysis has identified five genes known to be expressed in the heart and predicted to be targeted by at least two of miR-26, miR-30 and miR-92, suggesting that the regulation of these five genes may be essential for normal heart function.

miR-10 is initially not found in the top 10 miRNAs by normalised read count, but dominates from HH14-16, especially in non-heart samples (**Figure 3.7C, E and H**). Although it is not in the top 10 most highly expressed miRNAs in the early stages, it is still highly expressed at around position 15 in these samples. The predicted target genes of miR-10 not shared with the three common top expressed miRNAs (miR-26, miR-30 and miR-92) return only three significant GO terms, all related to nerve development: glossopharyngeal nerve development (GO:0021563), glossopharyngeal nerve morphogenesis (GO:0021615), and nervous system development (GO:0007399). The non-heart NT samples do contain the developing neural tube, but also contain many other cell types including the migrating cardiac neural crest, that will eventually form part of the heart, and cervical somites, that will develop into skeletal muscle. Therefore it is possible that in addition to its known roles in vasculogenesis and angiogenesis<sup>91,92</sup>, miR-10 may function to counteract factors promoting neural development in non-neural cells.

The extent to which predicted targets of miR-10 overlapped with predicted targets of miR-26, miR-30 or miR-92 known to be expressed in the heart was very low (less than 3%) (**Figure 3.10B**), similar to the low values of overlap seen when all the predicted target genes were compared (**Figure 3.9B**). Interestingly, only a single gene associated with a disease in the heart was a predicted target of miR-10: BDNF, which was also found to be targeted by miR-30 (**Table 3.8**).

### **3.3.7 Unique predicted target genes of top expressed miRNAs**

Finally, all the predicted target genes unique to each of the three groups of samples (early, late heart and late non-heart) were examined (**Table 3.7**). The unique predicted target genes of the early samples (HH5-7, HH8-10, HH11-13) were found to be significantly associated with signalling by TGF $\beta$  (REAC:170834). The TGF $\beta$  cytokines are involved in early heart development, during the formation of the endocardial cushions, in vasculogenesis, and in the development of the myocardium in ventricles<sup>104</sup>. In the older heart group (HH14-16 H, HH17-20 H, HH21-23 H) predicted target genes unique to this group were significantly associated with the GO term for congenital malformation of the

great arteries (HP:0011603), suggesting that miRNAs unique to this group function to ensure normal development of the great arteries and providing a shorter list of potentially very important miRNA:target gene interactions to evaluate in the context of normal heart development. The older non-heart group (HH14-16 NT, HH17-20, HH21-23 NT) had unique predicted target genes significantly associated with the positive regulation of cardiac muscle hypertrophy (GO:0010613). It is possible that this analysis reduced the thousands of predicted target genes of the top expressed miRNAs to a few interactions that could be essential for tissue development at these stages, simplifying future experimental investigations.

The most abundant miRNAs identified in nine chick embryo samples play essential roles during early development, and an attempt has been made to focus on some possibly novel functions by the bioinformatic analysis of the sequencing data obtained from these samples.

## 4 Evaluation of miRNAs important during heart development

### 4.1 Introduction

Following on from the summary of the sequencing strategy employed in this project (3.1) and the bioinformatic analysis of the known animal miRNAs detected in the samples assayed by NGS (3.2), this chapter focuses on bioinformatic analyses of miRNAs that are known to play roles in heart development. Adapter sequences and non-mapping reads were removed by Simon Moxon (2.8.1), but all subsequent analyses performed in this chapter were performed by the author.

#### 4.1.1 miRNAs are essential for heart development and function

The vital role of microRNAs during heart development and in heart function was illustrated by two studies that generated the conditional knockout of *Dicer* after deletion by the Cre recombinase enzyme under the transcriptional control of two heart-specific promoters<sup>4</sup>. *Dicer* has been conditionally deleted in the heart by placing *Cre* under the control of an early heart-specific promoter, *α-Myosin Heavy Chain (αMHC)*, which is expressed in mice from as early as E14.5<sup>15</sup>. The researchers observed postnatal lethality as well as dilated cardiomyopathy (DCM) that resembled the most common form of cardiomyopathy associated with human heart failure. The early conditional deletion of *Dicer* in mice by an *Nkx2.5*-promoter-driven Cre, active from E8.5, lead to embryonic lethality in mutant mice; developmental heart defects were observed on autopsy<sup>5</sup>. Furthermore, *Dicer* levels were observed to be much reduced in human patients with hearts failing from dilated cardiomyopathy, compared to patients with non-failing hearts<sup>4,15</sup>. Treatment of patients with failing hearts by fitting Left Ventricular Assist Devices (LVADs) resulted in *Dicer* levels being returned to levels similar to those detected in non-failing hearts<sup>15</sup>. Microarray results from several studies of hypertrophic or failing hearts show dysregulation of an assortment of microRNAs<sup>6</sup>.

Deletion of the *Egfl7* gene in mice resulted in animals with vascular malformations<sup>105</sup>. However, current evidence points to these abnormalities being a result of the deletion of miR-126, which is hosted by the *Egfl7* gene<sup>105</sup>. This suggests that other disease phenotypes observed by targeted deletion of coding genes in model organisms may be a result of the inadvertent removal of their hosted miRNAs.

#### **4.1.2 miRNAs influence proliferation and differentiation**

Several miRNAs have been shown to affect proliferation and differentiation in embryonic and adult hearts. Both miR-1 and members of the miR-17-92 cluster promote differentiation of cardiomyocytes<sup>6,106-108</sup>. However, miR-133 inhibits differentiation and stimulates proliferation of myoblasts<sup>4</sup>. The miR-15 family were shown to work in concert to repress several cell cycle regulators which lead to suppression of proliferation of cardiomyocytes<sup>106</sup>, and miR-101 inhibits the proliferation of cardiac fibroblasts<sup>109</sup>. miR-21, miR-24 and miR-499 have been shown to inhibit apoptosis, whereas the miR-15 family are linked to the induction of apoptosis<sup>106</sup>.

#### **4.1.3 MicroRNAs involved in vasculogenesis and angiogenesis**

miRNAs are essential for the proper development of the blood vessels. In vascular smooth muscle cells (VSMCs) miR-143 and miR-145 reduce proliferation and promote differentiation<sup>106</sup>, in contrast to miR-221 and miR-222, that induce proliferation at the expense of differentiation<sup>74</sup>. miR-18a, miR-19 and miR-30 are pro-angiogenic, whereas the miR-17-92 cluster is associated with reduced angiogenesis<sup>101,106</sup>. miR-126 is essential for blood vessel development in mice and zebrafish by directly targeting sprouty-related protein EVH1 domain containing 1 (SPRED1) and phosphoinositol-3 kinase regulatory subunit 2 (PIK3R2), repressors of the vascular endothelial growth factor (VEGF) pathway<sup>74</sup>. However, miR-126 has also been shown to negatively affect proliferation and to increase apoptosis of endothelial cells<sup>74</sup>.

#### 4.1.4 miRNAs involved in patterning the heart

miR-126 is required, in the outflow tract and in the endocardial cushions, for valve formation by repressing the inhibitors of the VEGF signalling pathway<sup>110</sup>. The expression of miR-138 during early cardiac looping is essential for patterning in the zebrafish heart<sup>111</sup>.

#### 4.1.5 Other miRNAs

*miR-208* is located in introns of the  $\alpha$ MHC gene, *Myh6*, and the  $\beta$ MHC gene, *Myh7*. Although it is expressed exclusively in the heart and slow skeletal muscle<sup>93</sup>, deletion of this miRNA during development did not result in any apparent developmental defects<sup>94</sup>. miR-208 is currently thought to play a role in cardiac stress response as well as in repressing the expression of skeletal muscle genes in the heart, rather than in the development of the heart<sup>6,94</sup>. The induction of hypertrophy in diabetic mouse hearts resulted in the downregulation of miR-130<sup>79</sup> (**Figure 4.1A**). Normal zebrafish heart development has been shown to depend on miR-138 expression in ventricles<sup>106,111</sup>, which is transcribed from intergenic regions of chromosome 2 and 11 in the chick (**Figure 4.1B**). It has been suggested that *miR-499*, located in an intron of *Myh7B* (**Figure 4.1C**), is also cardiac and slow skeletal muscle-specific due to its co-expression with *Myh7b*<sup>6,93</sup>.

<b>A</b>			
miR-130a-3p	Chr 15	Intergenic	CAGUGCAAUAUAAAAAGGGCAU
miR-130b-3p	Chr 15	Intergenic	CAGUGCAAUA <b>AUG</b> AAAGGGCGU
miR-130c-3p	Chr 19	Intron 1, SKA2	CAGUGCAAUGUAAAAAGGGCAU

<b>B</b>			
miR-138-5p	Chr 2	Intergenic	AGCUGGUGUUGUGAAUC
miR-138-3p	Chr 11	Intergenic	AGCUGGUGUUGUGAAUC

<b>C</b>			
miR-499-5p	Chr 20	Intron 11, MYH7B	UUAAGACUUGUAGUGAUGUUUAG

**Figure 4.1: The mature sequence and genome location of miR-130 (A), miR-138 (B), and miR-499 (C) in the chick.**

In the mouse, overexpression of miR-24 resulted in embryonic lethality whereas overexpression of miR-214 did not cause any defects<sup>6</sup>. miR-195 is upregulated in hypertrophic human hearts and in mouse models of cardiac hypertrophy, and its overexpression *in vitro* in rat cardiomyocytes resulted in hypertrophic growth<sup>6</sup>. Hypertrophy was also induced within weeks of birth in mice overexpressing miR-195 in cardiac muscle; eventually the elevated levels of miR-195 resulted in dilated cardiomyopathy and heart failure. miR-21 levels were increased after chemical induction of hypertrophy *in vitro* and following surgical intervention leading to hypertrophy *in vivo*<sup>106</sup>. The miR-144 / 451 cluster has been shown to increase survival of cardiomyocytes *in vitro* after chemical treatment to mirror ischemic damage, although the direct targets of these miRNAs have yet to be identified in this context<sup>112</sup>. Furthermore, miR-451 is known to promote erythroid maturation by targeting the 3' UTR of GATA-binding protein 2 (GATA2) in Zebrafish<sup>113</sup>.

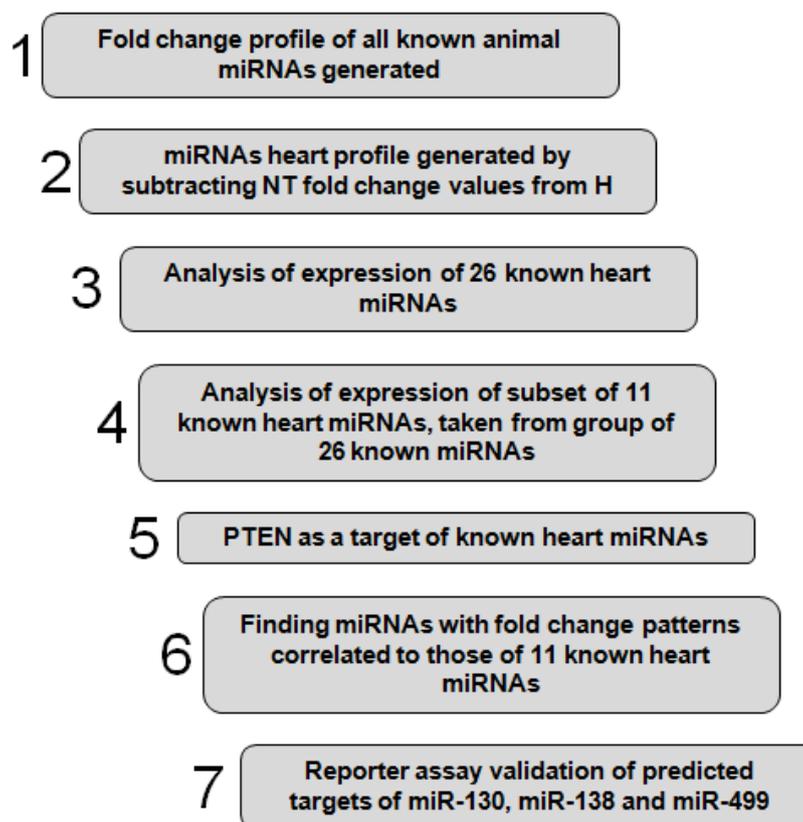
MicroRNAs clearly have many and diverse roles both during development and after birth, in many organs including the heart. The effects of microRNA dysregulation observed during *in vitro* experiments are often not seen *in vivo*, and differences between transient and permanent knockdown of microRNAs *in vivo* also exist<sup>114</sup>. The chick remains an excellent model species for embryonic experiments and despite a large number of microRNAs being expressed during chick embryogenesis<sup>115</sup>, the functions of the vast majority of microRNAs during development have yet to be established.

A miRNA profile specific to heart tissue samples (HH14-16, HH17-20, HH21-23) was generated by subtraction of the contributions of miRNAs in NT samples from their contributions in Heart samples (4.2.2). Furthermore, this chapter serves to highlight the expression in these samples of known miRNAs identified as important during heart development and function (Table 4.4). The fold change patterns of all detected miRNAs over their levels at HH5-7 were determined and miRNAs that were greatly upregulated during development were identified (4.2.1). Following a review of current literature 26 miRNAs known to play roles in heart development and function were identified. The fold change profiles of these 26 miRNAs were determined (4.2.3). A further analysis of the 11 most interesting heart miRNAs identified in this literature search exposed significant

correlation between these 11 miRNAs and several miRNAs not currently known to function in heart development (**4.2.6**).

## 4.2 Results

In this chapter bioinformatic analyses were focused towards miRNAs that are known to have roles in the development or function of the heart. An overview of these analyses is presented in **Figure 4.2**, as they are reported and discussed in each section below.



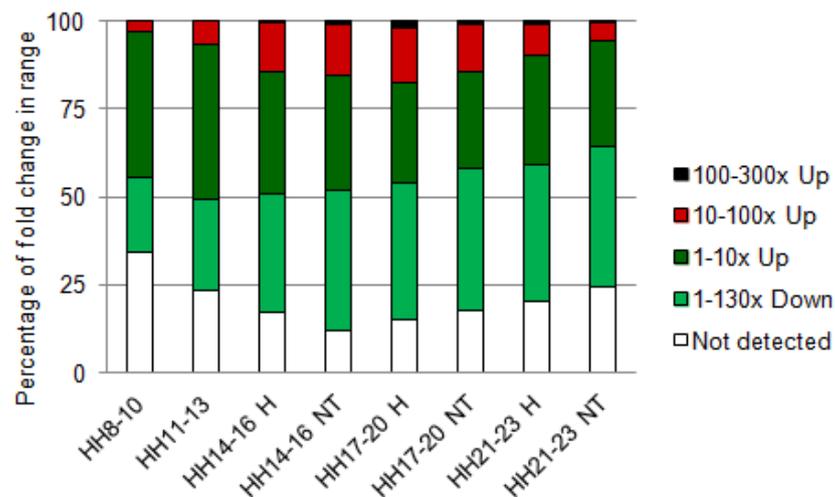
**Figure 4.2: Overview of analyses performed in Chapter 4.** Several analyses were performed in Chapter 4 ranging from generation of a fold change profile for all known miRNAs (1) to target validation of three selection miRNAs (7).

### 4.2.1 Few miRNAs have large increases in expression

In order to identify miRNAs that undergo large changes in expression during the developmental time period assayed and to create a profile of the pattern of change in the levels of miRNA expression, the fold change of known animal miRNAs in normalised reads per million was determined relative to the least developed sample, HH5-7. This allowed identification of the extent to which the largest group of miRNAs changed across all the samples, establishing that most miRNAs do not have large changes in expression

(**Figure 4.3**). In addition, it allowed the identification of the few miRNAs that did show large spikes in expression, pointing to their massive upregulation at specific stages of development (**Table 4.1**).

Samples and miRNAs were clustered based on miRNA fold change using custom R scripts written by the author (**Figure Apx A.2**). A breakdown of miRNAs by their fold change at each stage (**Figure 4.3**) shows that most of the 423 known animal miRNAs detected are expressed at a similar or lower level compared to that established at HH5-7, ranging from 49% (HH11-13) to 64% (HH21-23 NT). The remaining miRNAs show a modest upregulation over HH5-7 levels, with from 3% (HH8-10) to 16% (HH17-20 H) of miRNAs having a 1 to 10 fold increase in expression. Fewer than 2% of miRNAs have the most extreme increase of between 100 and 300 fold at some point during the developmental time series assayed. This shows that most miRNAs are expressed at lower or nearly constant levels as development progresses, with very few miRNAs dramatically upregulated.

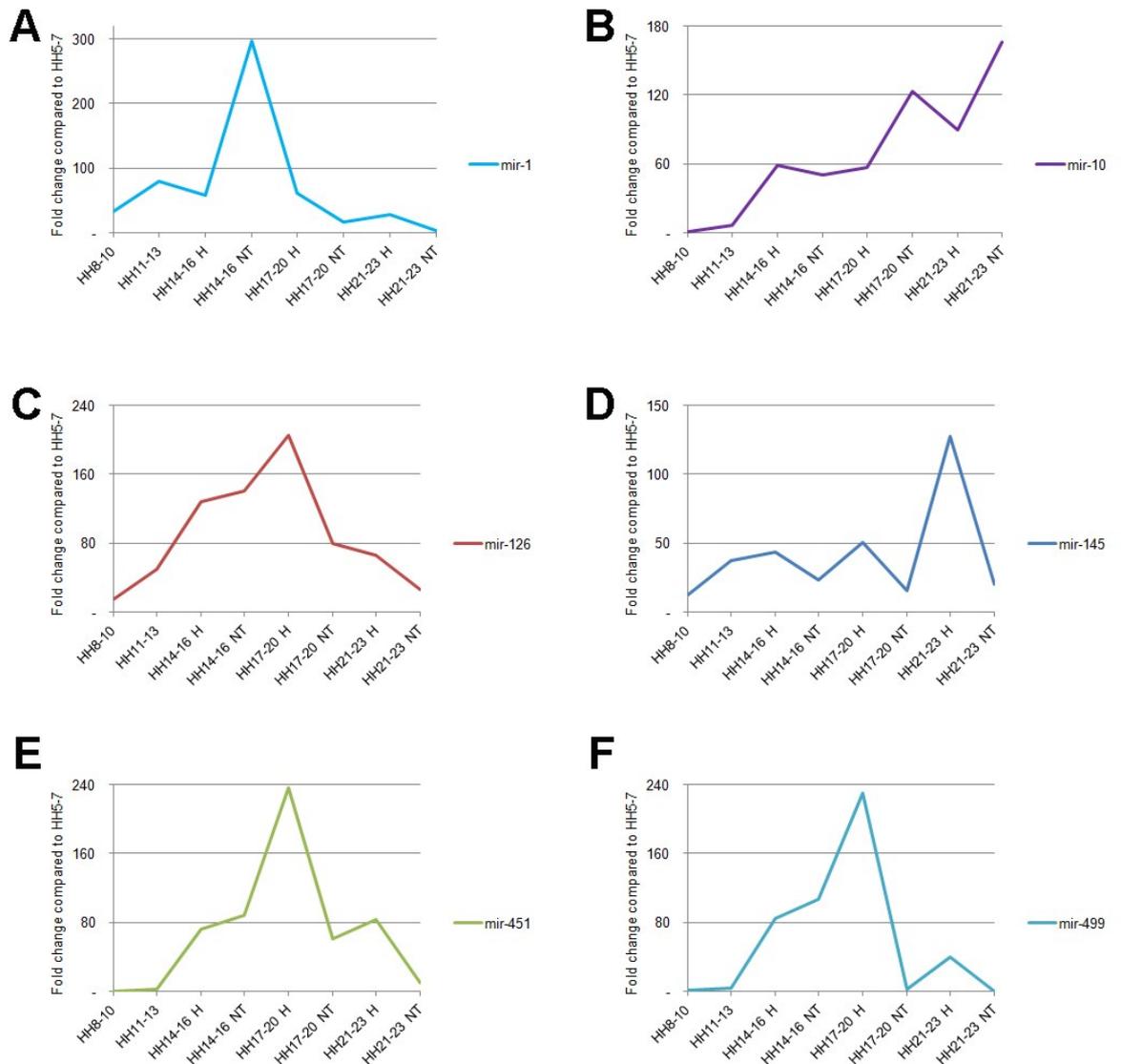


**Figure 4.3: The fold change of normalised reads per million for each known animal miRNA detected in the nine samples compared to reads at HH5-7.** Only miRNAs detected in more than one sample were considered. The percentage of miRNAs detected broken down by their fold change over HH5-7 levels. Many miRNAs were found to decrease between 2 and 130 fold from their HH5-7 levels. Most miRNAs that were upregulated were less than 10 fold increased over their HH5-7 levels. Very few (14) miRNAs were found to have levels 100 fold increased over their HH5-7 levels. White: Not detected. Light green: 1 to 130fold down. Dark green: 1 to 10 fold increased. Red: 10 to 100 fold increased. Black: 100 to 300 fold increased.

**Table 4.1: A few miRNAs have levels more than 100 times increased over their levels at HH5-7:** Fourteen miRNAs were found to have an increase in normalised read numbers more than 100 times that detected at HH5-7, in at least one of the nine samples assayed. Six of these miRNAs have known roles in heart development and function (underlined).

miRNAs with a fold change larger than 100 in at least one sample				
<u>miR-1</u>	<u>miR-126</u>	miR-187	<u>miR-451</u>	miR-9
<u>miR-10</u>	<u>miR-145</u>	miR-199	miR-490	miR-99
miR-100	miR-1451	miR-449	<u>miR-499</u>	

Only 14 miRNAs were found to have fold change values of more than 100 over their HH5-7 values (**Figure 4.3, Table 4.1**). Six of these miRNAs are known to play roles in heart development and function<sup>74,92,106,110,116</sup>, but have fold change patterns very different from one another (**Figure 4.4**). Three of these miRNAs (miR-126, miR-451 and miR-499) show a peak in upregulation in the HH17-20 heart tissue sample, whereas only miR-145 showed a large spike in expression in the heart at HH21-23.



**Figure 4.4:** The fold change patterns of six of the 14 miRNAs with a larger than 100 fold increase over their reads detected at HH5-7 in at least one sample. These miRNAs have all been implicated in heart development and function. **A:** miR-1<sup>106</sup>. **B:** miR-10<sup>92</sup>. **C:** miR-126<sup>74</sup>. **D:** miR-145<sup>106</sup>. **E:** miR-451<sup>116</sup>. **F:** miR-499<sup>106</sup>.

As discussed in 4.3.1 the timing of the upregulation of these miRNAs suggests that they may function during some of the important milestones of cardiovascular development, such as valve formation (HH17-20).

#### 4.2.2 Creating a profile of miRNAs in Heart by subtraction of NT sample values

A profile of miRNAs in the later heart samples (HH14-16 H, HH17-20 H, HH21-23 H) was generated by subtracting the fold change values of each miRNA in NT samples from

the corresponding Heart samples. This analysis filtered out miRNAs involved in processes common to both Heart and NT samples and generated a profile of miRNAs that may be essential for heart development.

**Table 4.2: Breakdown of the profile generated when subtracting miRNA fold change values in NT from those of the corresponding Heart samples.** The bulk of miRNAs in the HH14-16, HH17-20 and HH21-23 samples were found to have 1 to 10 fold higher fold change values in Heart samples than in NT samples, as shown in row three.

Change	% miRNAs	% miRNAs	% miRNAs
	HH14-16	HH17-20	HH21-23
>10x higher in Heart	1	11	8
1 – 10x higher in Heart	42	43	51
Heart			
Unchanged	15	13	11
1 – 10x lower in Heart	30	20	12
>10x lower in Heart	3	4	2
Undetected	8	9	16

More than 40% of the 292 known miRNAs detected in at least one of the HH14-16, HH17-20 and HH21-23 time points had fold change values 1 to 10 fold higher in the Heart sample than in the corresponding NT sample (**Table 4.2**). Very few (between 1% and 11%) of miRNAs were found to be more than 10 fold higher in Heart samples (**Table Apx A.6**). Similarly, less than 5% of miRNAs had fold change levels more than 10 fold lower in Heart samples. As **Table Apx A.6** shows, 10 of 27 miRNAs known to be important for heart development and function were identified as being more than 10 fold increased in these Heart samples. Of the remaining 28 miRNAs, 20 are conserved in humans and may have as yet undiscovered functions during heart development.

An increase in miRNA expression in Heart samples suggests that such miRNAs may be involved in repressing targets to shape the forming heart. In addition, profiles that resemble the H – NT profiles of known heart important miRNAs may identify miRNAs essential for heart development during these later stages. Therefore an analysis of this expression profile was conducted to identify miRNAs that have patterns of H – NT expression similar

to known important heart miRNAs in an attempt to identify as yet undiscovered miRNAs that have roles in heart development. A literature search revealed 26 miRNAs that have roles in heart development and function<sup>74,101,106,109,110</sup>. A correlation between the H – NT profile of each of these 26 heart miRNAs and the remaining miRNAs detected in the HH14-16, HH17-20 and HH21-23 samples was performed (**Table 4.3**). In addition miR-130, which was consistently identified as very highly expressed in later Heart samples (**Figure 3.7C**), was correlated against the remaining miRNAs identified in this screen. miR-92 did not have any significant correlations, and is therefore excluded from the table.

**Table 4.3: Breakdown of numbers of miRNAs that correlate to 26 known heart miRNAs based on a H – NT fold change profile established for the HH14-16, HH17-20 and HH21-23 samples.** Some of the heart miRNAs correlated significantly to other heart miRNAs, as shown in column three. The numbers of miRNAs that correlate to each heart miRNA that are also conserved in humans are shown in column four. Significant correlations were determined by two-tailed P values in GraphPad prism, using a 95% confidence interval.

Heart miRNAs	Total correlated miRNAs	Correlated heart miRNAs	Correlated miRNAs conserved in humans
miR-1	29	4	8
miR-10	1	0	0
miR-101	14	1	4
miR-125	2	1	1
miR-126	14	2	9
miR-128	14	0	4
miR-130	15	0	3
miR-133	19	2	12
miR-138	4	0	1
miR-143	17	4	7
miR-145	1	0	1
miR-15	3	0	1
miR-16	15	1	7
miR-17	8	1	6
miR-18	15	0	3
miR-181	6	2	3
miR-19	9	3	5
miR-20	14	0	3
miR-21	7	1	7
miR-221	13	1	7
miR-222	9	3	5
miR-24	7	1	3
miR-30	1	0	0
miR-31	6	0	3

miR-363	26	4	8
miR-499	8	1	5

A summary of the numbers of correlating miRNAs, of correlating heart miRNAs, and of correlating miRNAs that are conserved in humans is shown in **Table 4.3**. Detailed lists of the correlating miRNAs can be found in **Table Apx A.7**, **Table Apx A.8**, and **Table Apx A.9**.

This analysis has identified miRNAs conserved in humans that correlate by H – NT profile to known heart miRNAs, highlighting networks of miRNAs that may share similar functions during heart development.

#### **4.2.3 miRNAs with a role in heart development and function were detected**

A literature review identified 26 miRNAs known to play various roles in heart development and function<sup>74,101,106,109,110</sup> (**Table 4.4**). A hierarchical clustering of the fold change values of the 26 miRNAs was performed using a custom R script written by the author in order to identify similarities between the expression patterns of these heart important miRNAs. This served to identify the closely related expression patterns of miR-126 and miR-499, well known and relative unstudied miRNAs respectively. Further examination of the biological function of miR-499 identified its regulation of an activator of miR-126 transcription (**4.3.1**).

**Table 4.4: 26 miRNAs identified from a literature search have roles in heart development and function.**

Chosen miR	Role in heart
miR-1	Inhibits proliferation and promotes differentiation of cardiomyocytes <sup>106</sup> .
miR-10	Positively affects vasculogenesis and angiogenesis by its effects on VEGF signalling <sup>91,92</sup> .
miR-15	Inhibition of cardiomyocyte proliferation by inhibition of several cell cycle regulators <sup>106</sup> .
miR-16	Inhibition of cardiomyocyte proliferation by inhibition of several cell cycle regulators <sup>106</sup> .
miR-17	Deletion of the miR-17-92 cluster leads to reduced cardiomyocyte proliferation in mice <sup>98</sup> .
miR-18	Deletion of the miR-17-92 cluster leads to reduced cardiomyocyte proliferation in mice <sup>98</sup> .
miR-19	Deletion of the miR-17-92 cluster leads to reduced cardiomyocyte proliferation in mice <sup>98</sup> .
miR-20	Deletion of the miR-17-92 cluster leads to reduced cardiomyocyte proliferation in mice <sup>98</sup> .
miR-21	Inhibition of apoptosis of cardiomyocytes <sup>106</sup> .
miR-24	Inhibition of apoptosis of cardiomyocytes <sup>106</sup> .
miR-30	Ectopic expression of miR-30 led to excessive angiogenesis in zebrafish embryos <sup>101</sup> .
miR-31	Negatively modulates cardiac epithelial-to-mesenchymal transition by targeting <i>Isl1</i> <sup>117</sup> .
miR-92	Deletion of the miR-17-92 cluster leads to reduced cardiomyocyte proliferation in mice <sup>98</sup> .
miR-101	Inhibits proliferation of cardiac fibroblasts in adult mice <sup>109</sup> .
miR-125	Prevents apoptotic signalling and protects the myocardium of adult mice after injury <sup>118</sup> .
miR-126	Essential for valve formation <sup>110</sup> and angiogenesis <sup>74</sup> .
miR-128	Targets <i>Isl1</i> <sup>119</sup> which is essential for cardiac specification and vasculogenesis <sup>120</sup> , as well as cardiac EMT <sup>117</sup> .
miR-133	Inhibits the proliferation of cardiomyocytes <sup>106</sup> .
miR-138	Required for regulation of patterning in the looping zebrafish heart <sup>111</sup> .
miR-143	Reduces proliferation and promotes differentiation of VSMCs <sup>106</sup> , along with miR-145.
miR-145	Reduces proliferation and promotes differentiation of VSMCs <sup>106</sup> , along with miR-143.
miR-181	Expression is induced during myogenesis in human heart and skeletal muscle <sup>121</sup> .
miR-363	Possible role in negative regulation of angiogenesis <sup>122</sup> .
miR-499	Decreases apoptosis in cardiomyocytes <sup>106</sup> .

The resulting dendrogram (**Figure 4.5**) was coloured to show the grouping of highly correlated miRNAs. Most miRNAs cluster closely, confirming their very similar patterns of expression in these samples.

miR-10 and miR-125, both of the miR-10 family, did not correlate well by expression pattern (**Figure 4.5**) and are located on different chromosomes (Chr 1 and Chr7 respectively). Despite the sequence similarities of these miRNAs across their entire hairpins, they have only three matching bases in their seed regions. In addition, their known functions in heart development differ greatly (**Table 4.4**). These data confirm that

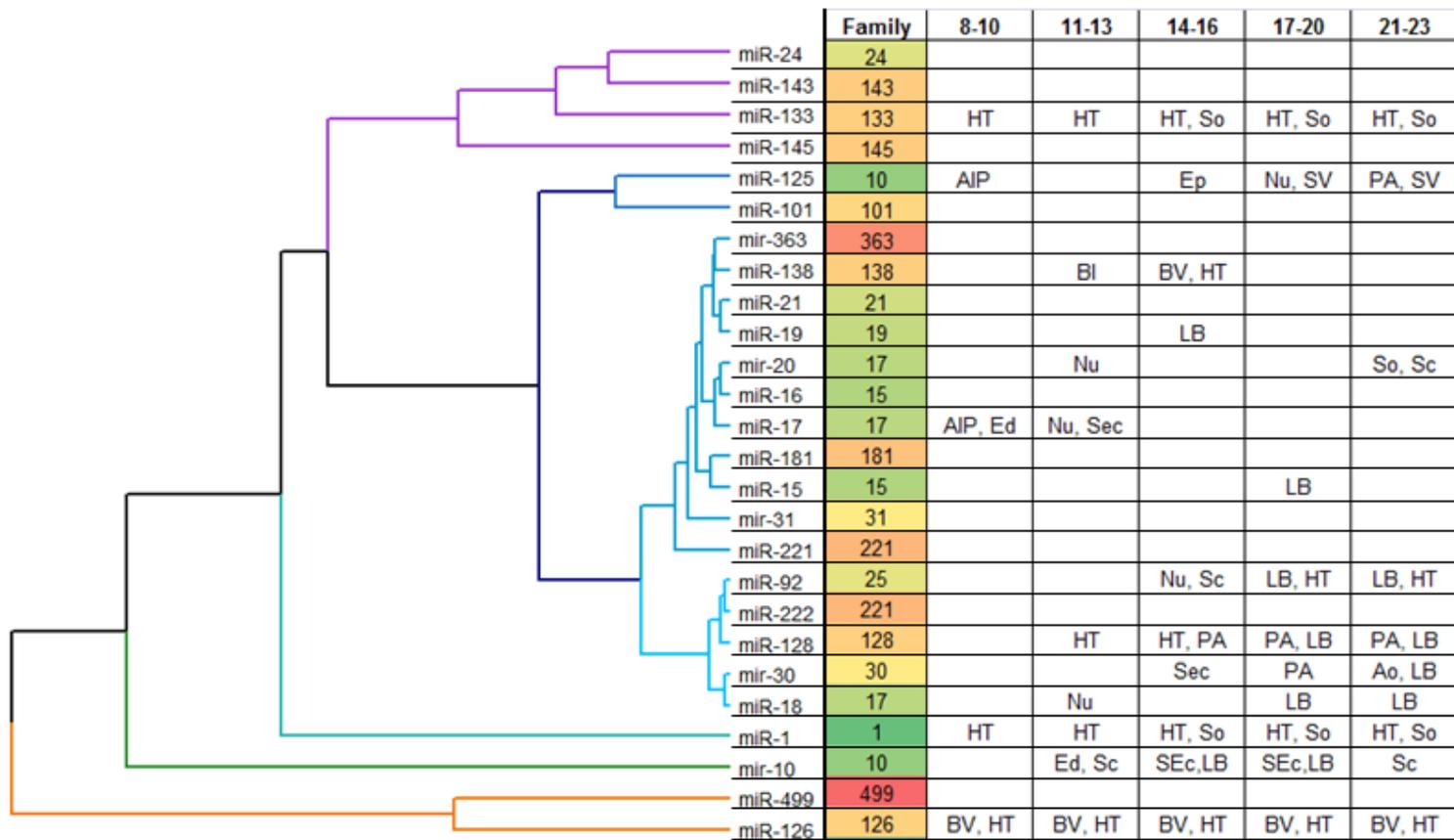
although miR-10 and miR-125 may have had a common ancestor, they appear not to be co-regulated in these tissues.

miR-143 and miR-145 correlated closely by fold change expression pattern (**Figure 4.5**). miR-143 has not yet been annotated in the chicken genome, but in the human genome is located less than 2 Kbp from miR-145 in an intergenic region of Chr5. In this analysis their expression patterns were found to correlate well which indicates that these miRNAs may be co-regulated in the chicken, as they are known to be in mice and humans<sup>106</sup>.

miR-363 has not been annotated in the chicken genome but has been identified in this screen. In humans, miR-363 is clustered on ChrX with miR-18, miR-19, miR-20 and miR-92. The closely related expression patterns of miR-363, miR-19, miR-20 and miR-17 observed here indicates that miR-363 and these members of the miR-17-92 cluster may be regulated by the same factors in these samples (**Figure 4.5**).

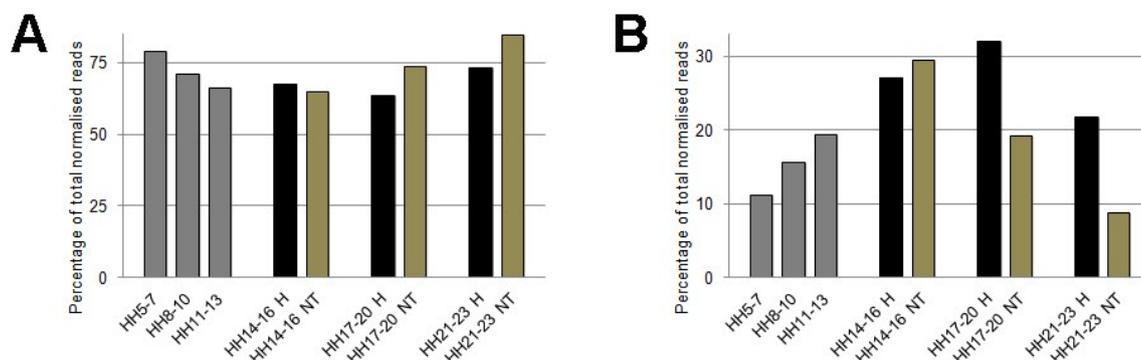
Two members of the miR-221 family, miR-221 and miR-222, are clustered together on Chr1 and here had similar patterns of expression (**Figure 4.5**). These miRNAs are known to be highly expressed together in endothelial cells and therefore their correlated expression patterns in the developing heart confirms previous findings<sup>106</sup>.

The clustering of miR-499 and miR-126 together (**Figure 4.5**) was noteworthy and, as is discussed in **4.3.1**, suggests that miR-499 may regulate miR-126 by targeting its transcriptional activator, ETS1.



**Figure 4.5 miRNAs known to be involved in heart development and function from current literature cluster well based on their fold change patterns across all samples.** A coloured dendrogram of 26 miRNAs known to play a role in heart development and function, clustered based on the fold change in their normalised read values over those recorded in the HH5-7 sample. Clustering was done using hierarchical clustering function in the statistical package R. Column two of the grid shows which miRNA family each miRNA belongs to. The LNA WISH expression of each miRNA, as annotated on the Geisha ISH web resource, is indicated for each time point. HT: Heart tube. So: Somites. Sc: Spinal cord. Ed: Endoderm. SEc: Surface Ectoderm. LB: Limb Buds. AIP: Anterior Intestinal Portal. Nu: Neural Tube. Ao: Aorta. PA: Pharyngeal Arches and Clefts. Ep: Epicardium. SV: Sinus Venosus. BV: Blood Vessels. BI: Blood Islands.

To establish whether the 26 published heart miRNAs (**Table 4.4**) show average increased expression in heart samples, the percentage of the total known miRNA normalised reads taken up by these 26 miRNAs were examined across all the samples (**Figure 4.6**).



**Figure 4.6: The percentage of the total normalised reads due to miRNAs with known roles in the heart. A:** The portion of normalised reads from annotated animal miRNAs taken up by 26 known heart miRNAs. **B:** The portion of normalised reads from annotated animal miRNAs taken up by 24 known heart miRNAs, after removal of miR-10 and miR-92 reads.

The 26 miRNAs known to function in the heart represented between 64% (HH17-20 H) and 79% (HH21-23 NT) of the total reads from all the known miRNAs recorded in the samples (**Figure 4.6A**), and do not appear to be expressed at a higher level in the heart tissue samples. However, this domination was largely due to the contributions of two miRNAs: miR-92 and miR-10, which comprised more than two thirds of the reads recorded in HH5-7 and HH21-23 NT samples respectively. Removal of the reads due to these two miRNAs highlighted the increased expression of the remaining 24 miRNAs in the heart samples over that detected in the matched non-heart samples (**Figure 4.6B**).

**Table 4.5: Thirteen miRNAs were found to have slightly increased reads in the HH14-16 NT sample over that detected in the HH14-16 H sample.** Increases ranged from 103% (miR-126) to 489% (miR-1) of the HH14-16 H reads.

miRNAs with slightly increased reads in HH14-16 NT over HH14-16 H				
miR-1	miR-133	miR-17	miR-222	miR-499
miR-126	miR-143	miR-181	miR-24	
miR-128	miR-16	miR-21	miR-31	

The reversal of this trend in HH14-16 samples is not due to the effects of any one miRNA, but rather from the cumulative effect of 13 out of 24 miRNAs having slightly increased fold change numbers in the non-heart sample (HH14-16 NT) over the heart sample (HH14-16 H) (**Table 4.5**). The migration of cardiac neural crest is known to depend on miRNA expression and miR-21 and miR-181 are essential for this migration<sup>123</sup>. Furthermore, miR-143 is expressed in migrating cardiac neural crest cells that will eventually form vascular smooth muscle cells and controls their differentiation into smooth muscle<sup>124</sup>. It is possible that the increased read numbers of these miRNAs reflect their expression in the cardiac neural crest, which migrates through the dissected area in the HH14-16 NT sample prior to their arrival in the outflow tract at HH17 (**Table 3.1**), which would implicate the remaining miRNAs in **Table 4.5** as being expressed in cardiac neural crest and possibly having important roles during this migration.

#### 4.2.4 11 Known heart miRNAs correlate well by fold change

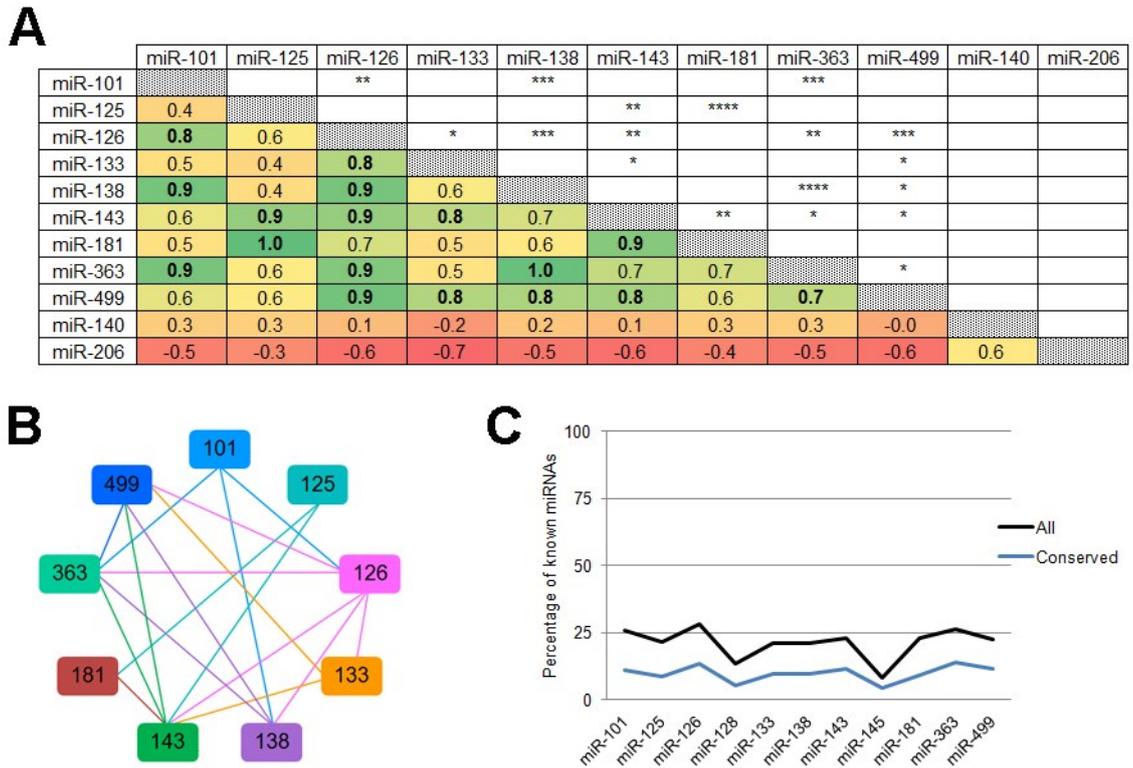
Although all 26 known miRNAs are worthy of further study, analysis of such a large group can quickly become cumbersome. The list of 26 known miRNAs evaluated in **Figure 4.5** and **Figure 4.6** was narrowed down to eleven miRNAs (**Table 4.6**) as described below.

**Table 4.6: Eleven miRNAs were selected for further investigation based on a list of 26 miRNAs known to have roles in heart development and function.**

Chosen miR	Role in heart
miR-101	Inhibits proliferation of cardiac fibroblasts in adult mice <sup>109</sup> .
miR-125	Prevents apoptotic signalling and protects the myocardium of adult mice after injury <sup>118</sup> .
miR-126	Essential for valve formation <sup>110</sup> and angiogenesis <sup>74</sup> .
miR-128	Targets Isl1 <sup>119</sup> which is essential for cardiac specification and vasculogenesis <sup>120</sup> , as well as cardiac EMT <sup>117</sup> .
miR-133	Inhibits the proliferation of cardiomyocytes <sup>106</sup> .
miR-138	Required for regulation of patterning in the looping zebrafish heart <sup>111</sup> .
miR-143	Reduces proliferation and promotes differentiation of VSMCs <sup>106</sup> , along with miR-145.
miR-145	Reduces proliferation and promotes differentiation of VSMCs <sup>106</sup> , along with miR-143.
miR-181	Expression is induced during myogenesis in human heart and skeletal muscle <sup>121</sup> .
miR-363	Possible role in negative regulation of angiogenesis <sup>122</sup> .
miR-499	Decreases apoptosis in cardiomyocytes <sup>106</sup> .

miR-101, miR-125 and miR-181 have not been studied during embryonic development, but their effects in adult mice suggested that they may be important during heart development<sup>109,118,121</sup>. miR-126 was selected as it is essential for the formation of the heart<sup>74,110</sup>, and because it clustered by expression pattern to miR-499. miR-499 is not well studied<sup>106</sup>, but is located within an intron of and co-expressed with Myh7b<sup>125</sup>, which is essential for normal heart development<sup>126</sup>. Both miR-126 and miR-499 were selected to further elucidate miR-499 roles during heart development. miR-128, miR-133 and miR-145 were selected as they have several interesting roles during heart development<sup>106,117,119,120</sup>. miR-138 has not been studied in the chick or well studied in general and is essential for looping of the zebrafish heart<sup>111</sup>, and its rapid upregulation from the onset of looping (HH11-13, **Figure 4.9H**) was thought to be noteworthy. miR-143 was selected because although it is not annotated in the chicken genome, its expression pattern correlated highly with that of miR-145 which suggests that these miRNAs may function in the chick as they do in mice<sup>106</sup>, during VSMC proliferation and differentiation. In addition, miR-130 was consistently detected in the top 10 most highly expressed miRNAs in early (HH5-7, HH8-10, HH11-13) and later heart (HH14-16 H, HH17-20 H, HH21-23 H) samples (**Figure 3.7A, C**). As it had not been studied in the context of heart development, miR-130 was selected based on its expression in these samples.

The level of correlation between these 11 miRNAs was assessed based on their fold change in normalised reads over that recorded at HH5-7 (**Figure 4.7**). Some of these chosen miRNAs, such as miR-126, have well characterised roles in heart development, whereas others, such as miR-181, are only suspected of functioning in the growing heart (**Table 4.6**).



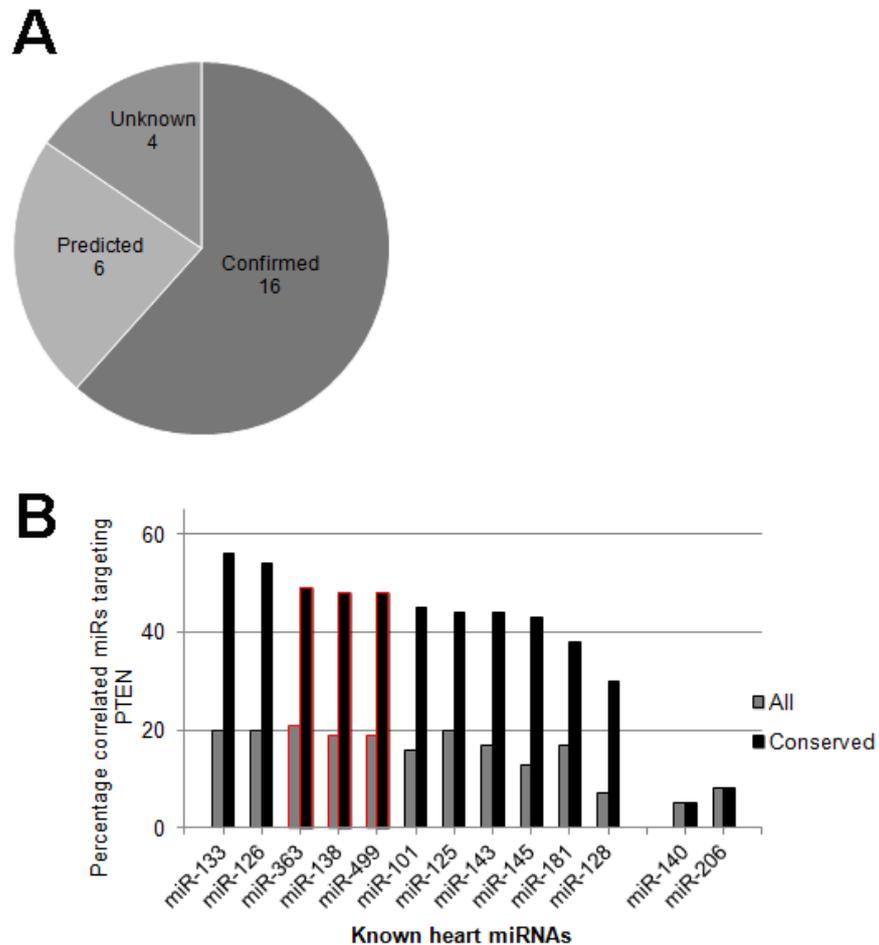
**Figure 4.7: An assessment of the extent to which 11 miRNAs with known roles in heart development and function correlate based on their fold change patterns across all samples.** In order to determine whether the fold change patterns of 11 miRNAs with known roles in heart development and function (**Table 4.6**) were similar, the correlation of each pairing was assessed. 9 of the 11 were found to have significantly similar changes in their read numbers over that detected at HH5-7, when considered over all the remaining 8 samples. **A:** A grid displaying the correlation coefficients and the statistical significance of the correlations. Correlation coefficients are coloured from low to high in red to green. The top triangle of the grid shows the level of significance of the correlations.  $0.05 \geq P > 0.01$  : \*;  $0.01 \geq P > 0.001$  : \*\*;  $0.001 \geq P > 0.0001$  : \*\*\*;  $P \leq 0.0001$  : \*\*\*\*. Significance was only observed in pairings with correlation coefficients larger than 0.737. **B:** Schematic diagram indicating the significant correlations between the 11 of 26 miRNAs with known roles in heart development. **C:** An indication of the number of known miRNAs that were found to correlate with the 11 chosen miRNAs, showing that less than 30% of the 412 remaining miRNAs correlated with these chosen miRNAs. The numbers are reduced even further when only miRNAs also conserved in humans are considered.

Statistically significant correlations were found to exist between 9 of the 11 miRNAs chosen for further investigation when examined by fold change in normalised reads (**Figure 4.7A** and **B**). miR-126 correlated well with six of the 11 chosen miRNAs (miR-133, miR-138, miR-143, miR-181, miR-363 and miR-499). However, two of the chosen miRNAs were found to not be statistically significantly correlated to any of the others: miR-128 and miR-145. Control correlations were done with skeletal muscle-specific miR-206<sup>127</sup>, and chondrogenic miR-140<sup>128</sup>. The fold change patterns of the control miRNAs did not significantly correlate with the patterns of the 11 chosen heart miRNAs (**Figure 4.7A**). Despite high levels of correlation between the 9 heart miRNAs observed in **Figure 4.7A**,

these correlations were found to be specific to heart miRNAs and not reflected in the overall fold change patterns of all the miRNAs detected in the screen. The 9 miRNAs with significant correlations to each other were found have fold change patterns significantly similar to less than 30% of the 424 miRNAs expressed in at least two samples (**Figure 4.7C**). Only a small number of the correlating miRNAs, ranging from 19 (miR-145) to 58 (miR-363), were found to be conserved in humans. This analysis suggests that the correlations observed may reflect a concerted effect of these 9 miRNAs on heart development, and has uncovered a number of conserved miRNAs that also have very similar patterns of expression to these heart important miRNAs. It is conceivable that these miRNAs have similar targets or pathways in common, and many of the 26 miRNAs involved in heart development and function are known or predicted to target a single gene, PTEN, as is discussed in **4.2.5**.

#### **4.2.5 PTEN is an enriched target of most of the miRNAs with functions in the heart**

The phosphatase and tensin homolog gene, PTEN, is known to be essential for normal myocardial development in certain parts of the developing heart<sup>129,130</sup>. The 3' UTR of PTEN has 68 target sites for miRNAs conserved among vertebrates, 2 of which have been confirmed experimentally (miR-130, miR-182). Despite not being predicted to target this gene by TargetScan<sup>131</sup>, more than half (16) of the 26 heart miRNAs shown in **Figure 4.5** have been experimentally confirmed to inhibit PTEN<sup>132</sup>, and a further 6 are conserved predicted targeting miRNAs of PTEN (**Figure 4.8A**). Furthermore, 8 of the 11 selected heart miRNAs are known to target PTEN, and only 2 are not predicted to target this gene (miR-138, miR-499).



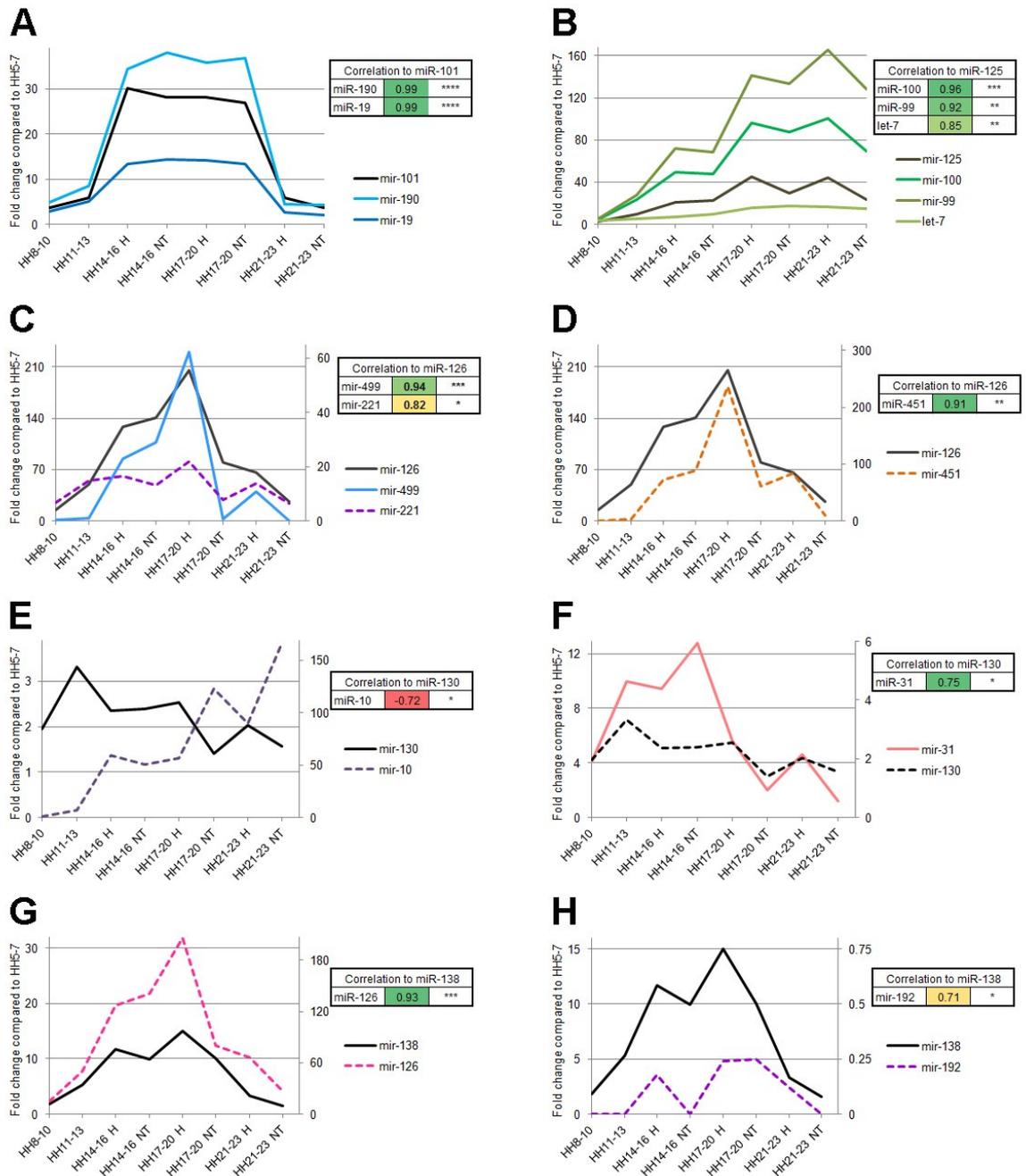
**Figure 4.8: PTEN is an enriched target of miRNAs known to affect heart development and function, as well as of their correlated known animal miRNAs.** **A:** Breakdown of the PTEN-targeting status of 26 miRNAs described in **Figure 4.5**. 16 are experimentally confirmed to target PTEN and 6 are predicted to target this gene. **B:** Around 20% of miRNAs with significantly correlated fold change patterns to 11 known heart miRNAs are predicted or confirmed to target PTEN. This number increases to between 38% and 56% when only miRNAs conserved in humans are considered. The bars outlined in red are miRNAs correlated to three miRNAs not experimentally shown to target PTEN to date.

About 20% of the miRNAs significantly correlated to the 11 chosen heart miRNAs are also either predicted or confirmed to target PTEN. This number increased to between 30% (miR-128) and 56% (miR-133) when only miRNAs conserved in humans were considered (**Figure 4.8B**). PTEN is not a predicted target of either the chondrogenic miR-140 or the skeletal muscle-specific miR-206, and is a predicted target of only 5% and 8% of their significantly correlated miRNAs respectively. This analysis strongly suggests that one of the major functions of the heart miRNAs examined, and of most of their conserved correlated miRNAs, is to target PTEN. Furthermore, a “miRNA sponge” pseudogene of *PTEN*, *PTEN1*, contains multiple binding sites for miRNAs that also target the 3' UTR of

*PTEN*<sup>133</sup>, in what appears to be a mechanism to further modulate the translation of this gene by decreasing the size of the miRNA pool targeting *PTEN* at any one time.

#### **4.2.6 Selected correlated miRNA fold change patterns**

In this section the expression patterns of miRNAs found to correlate well to some of the 11 chosen heart miRNAs (**Table 4.6, Figure 4.9**) were examined, along with some of the implications of these correlations. To determine whether the correlated miRNAs target similar pathways, GO terms analysis of their predicted and confirmed target genes was performed.



**Figure 4.9: Normalised reads fold change over levels recorded at HH5-7 for miRNAs that correlate well with selected miRNAs with known roles in heart development and function. Dotted lines were used for profiles plotted on the secondary axis. A:** miR-101 with miR-190 and 19. **B:** miR-125 with miR-100, miR-99 and let-7. **C:** miR-126 and miR-499 with miR-221. **D:** miR-126 with miR-451. **E:** miR-130 with miR-10. **F:** miR-130 with miR-31. **G:** miR-138 with miR-126. **H:** miR-138 with miR-192.

miR-101, miR-19 and miR-190 have very similar fold change patterns across the stages of chick development assayed (**Figure 4.9A**). As both miR-101 and miR-19 are known to function in the forming heart<sup>98,109</sup>, the significant correlation of miR-190 expression may

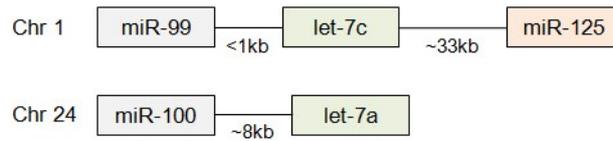
indicate a similar role for this miRNA. The predicted and confirmed targets of miR-101, miR-190 and miR-19 were found to be significantly enriched for genes involved in the development of the nervous system (**Table 4.7**).

**Table 4.7: Significantly enriched common GO terms of the predicted and confirmed target genes of miR-101, miR-190 and miR-19.**

Significant GO term	P value
Nervous system development (GO:0007399)	$0.4 \times 10^{-3}$
Neurogenesis (GO:0022008)	$0.3 \times 10^{-1}$
Neuron differentiation (GO:0030182)	$0.2 \times 10^{-1}$
Generation of neurons (GO:0048699)	$0.5 \times 10^{-1}$

The observed enrichment of targets involved in neurogenesis and neuron differentiation suggests that miR-101, miR-19 and miR-190 may ensure that heart tissue remains committed to the cardiac lineage (**4.3.6**).

miR-100, let-7 and miR-125 were found to have very similar changes in expression (**Figure 4.9B**). miR-99, a relatively unstudied miRNA, also correlated well with these miRNAs (**Figure 4.9B**). miR-100 and miR-99 are both members of the miR-99 family and have sequences that differ by only one nucleotide (**Figure 4.10B**). The genomic locations of miR-100, let-7, miR-125 and miR-99 are shown in **Figure 4.10A**. miR-100 expression results in the repression of the adult cardiac genes  $\alpha$ -myosin heavy chain ( $\alpha$ MHC) and sarco/endoplasmic reticulum  $\text{Ca}^{2+}$  - ATPase (SERCA2a) during development<sup>134</sup>. The importance of miRNA-mediated downregulation of *SERCA2a* was recently confirmed when miR-25 was shown to inhibit the translation of this gene in humans<sup>135</sup>. However, miR-25 is not known in the chick and has not been identified in any of the samples in this screen. miR-125, miR-100 and let-7 have significantly correlated patterns of expression, suggesting their co-regulation despite miR-100 being located on a separate chromosome. A lack of predicted targets of miR-99 and miR-100 proved a frustrating barrier to further GO terms analysis of these correlated miRNAs.

**A****B**

miR-99a-5p	AACCCGUAGAUCCGAUCUUGUG
miR-100-5p	AACCCGUAGAUCCGAAUCUUGUG
let-7a-2-3p	CUGUACAACCUCUAGCUUUC
let7-c-5p	UGAGGUAGUAGGUUGUAUGGUU
miR-125b-5p	UCCUGAGACCCUAAACUUGUGA

**Figure 4.10: miR-100, let-7, miR-125 and miR-99 are clustered in the chicken genome. A:** A schematic of the genomic locations of miR-100, let-7, miR-125 and miR-99. **B:** The mature sequences of miR-100, let-7a, let-7c, miR-99 and miR-125 found in the chick.

miR-126, essential for angiogenesis and valve formation<sup>74,110</sup>, has a fold change profile very similar to that of miR-499 (**Figure 4.9C**) and miR-451 (**Figure 4.9D**). Target prediction programmes return very few predicted targets for miR-126, but 131 genes have been shown to be affected by this miRNA<sup>132</sup>. Only 16 of the 86 predicted target genes of miR-451 are shared with the list of known miR-126 target genes. However, a comparison found that 82% of the GO terms of the predicted targets of miR-451 were common to the terms from confirmed targets of miR-126, suggesting that despite having few target genes in common these miRNAs may be targeting similar pathways during these developmental stages.

The fold change patterns of miR-130 and miR-31 were found to be significantly correlated (**Figure 4.9F**). miR-31 is known to target *Isl1*, a transcription factor that is essential during heart development<sup>117,120,136</sup>. In addition, miR-130 represses the translation of a reporter construct containing the 3' UTR of LEF1, a known transcriptional activator of *Isl1* (**Figure 4.13C**). Together these findings implicate miR-130 as a regulator of heart development, as discussed further in 4.3.6.

A significant negative correlation between miR-10 and miR-130 was observed (**Figure 4.9E**). Not many of the confirmed or predicted targets of miR-130 and miR-10 are shared (**Table 4.8**), and only 25% of the 100 most significant GO terms associated with combined confirmed and predicted targets of miR-130 and miR-10 are common to both miRNAs.

**Table 4.8: Few of the predicted or confirmed target genes of miR-130 are also predicted or confirmed targets of miR-10.**

miRNA	Common confirmed targets	Common predicted targets
miR-130	18.5% (13/70)	3.9% (18/462)
miR-10	16.0% (13/81)	20.6 (18/87)

However, too little is currently known about the functions of miR-130 to allow further analysis of this finding (**4.3.6**).

A significant correlation between the fold change pattern of miR-138 and miR-126, as well as miR-138 and miR-192 was observed (**Figure 4.9G, H**). Cardiac patterning in zebrafish requires the expression of miR-138<sup>111</sup>. Furthermore miR-138 translationally repressed the expression of a reporter construct containing the 3' UTR of ZEB2, an activator of PTEN, a known target of miR-192 (**Figure 4.17**). Repression of ZEB2 leads to growth and angiogenesis, as is the effect of miR-126 on its target PIK3R2 in the same pathway (**Figure 4.17**). The correlated expression patterns of miR-138, miR-126 and miR-192 may reflect their targeting of the PTEN pathway in the growing heart (**4.3.5**).

Several novel roles for miRNAs correlated to known heart miRNAs were suggested in this section, as is discussed further in sections **4.3.5** and **4.3.6**.

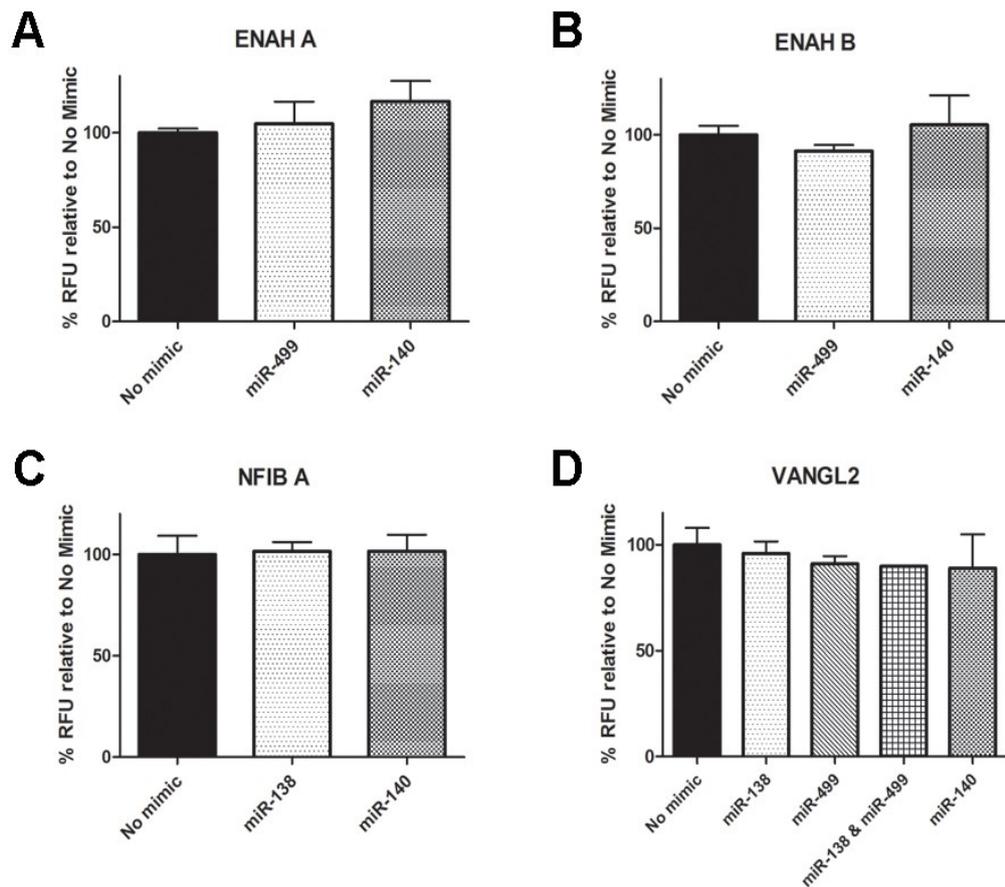
#### **4.2.7 Validating targets of miR-130, miR-138 and miR-499**

Three miRNAs and 15 of their target genes were chosen for further investigation. miR-138 is required for the regulation of patterning in the looping zebrafish heart<sup>111</sup>, and miR-499 is associated with decreased apoptosis in cardiomyocytes<sup>106</sup>. miR-130 was selected based on its fold change profile as a more speculative miRNA with possible roles in heart

development. Validation was done by luciferase assays of modified pGL3 reporter constructs<sup>47</sup> containing the cloned 3' UTRs of the target genes, amplified from chick RNA.

The 3' UTR of the longest transcript of *Enabled homolog (Drosophila) (ENAH)* contains two conserved predicted miR-499 binding sites. *ENAH* is known to have an essential role in intercalated disk function at the interface between cardiac myocytes<sup>137</sup>. This gene has nine transcripts, three of which result in protein products<sup>138</sup>. The remaining two transcripts, both much shorter, have identical 3' UTRs that also match to a downstream portion of the 3' UTR of the long transcript. This shorter 3' UTR contains only a single predicted miR-499 target site. miR-499 was found not to target any of these predicted sites (**Figure 4.11A, B**).

The *nuclear factor I B* gene, *NFIB*, is predicted to have miR-130 and miR-138 target sites. *NFIB* is thought to direct tissue-specific gene expression during development and is expressed in muscle during mouse development and in adult mouse hearts<sup>139</sup>. Although the annotated 3' UTR of *NFIB* in the chicken genome is 1,800bp long, TargetScan predicted miR-130 and miR-138 target sites up to 3,000bp further downstream of the end of the annotated 3' UTR region. Two reporter constructs were prepared with different portions of this long UTR. Construct A contained the first 1,700bp of the 3' UTR of *NFIB*, with two predicted target sites for miR-138. The second construct, B, contained sequence from an unannotated region 3,600bp to 5,000bp from the start of the 3' UTR, as well as single miR-130 and miR-138 predicted target sites. Neither of the predicted miR-138 target sites in construct A was found to reduce the levels of luciferase (**Figure 4.11C**). Both miR-130 and miR-138 were found to target construct B (**Figure 4.13D**). Time constraints lead to only a single mutant construct being made, with the miR-130 predicted target site mutated. The reporter assay showed that the action of miR-130 on the predicted miR-130 target site is responsible for the inhibitory effect of this miRNA on *NFIB* 3'UTR sensor expression.

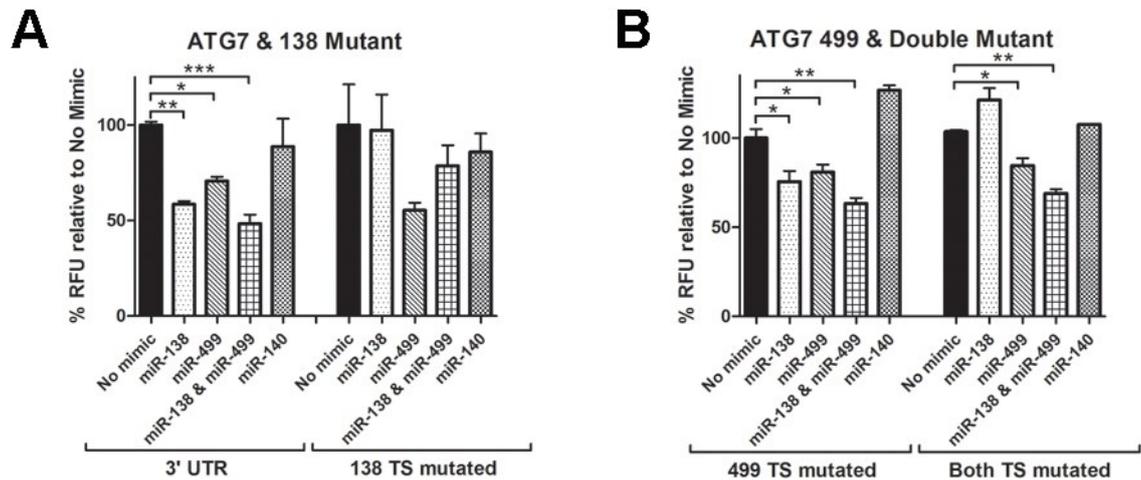


**Figure 4.11: The predicted target sites in the 3' UTRs of three genes were shown to not be true target sites by luciferase reporter assays. A,B:** Neither miR-499 nor the control, miR-140, repressed the translation of a luciferase reporter constructs containing the 3' UTRs of the two splice variants of ENAH. **C:** The first 1,700nts of the NFIB 3' UTR contained two miR-138 target sites, neither of which were responsive to miR-138 mimic treatment. **D:** VANGL2 did not respond to miR-138 or miR-499. One way ANOVA with Tukey post test statistical analyses were performed on the results of each construct.

The *VANGL planar cell polarity protein 2 (VANGL)* mRNA is predicted to contain target sites for miR-138 and miR-499. The *VANGL* gene encodes Vangl2, a transmembrane protein that is a component of the Wnt/PCP signalling pathway essential during development<sup>140</sup>. PCP signalling is required for OFT formation and cardiac looping, although it is dispensable in the CNC<sup>141</sup>. Neither miR-138 nor miR-499 was found to target *VANGL2* (Figure 4.11D).

TargetScan predicted miR-138 and miR-499 target sites within the *Autophagy related 7 (ATG7)* mRNA. *ATG7* induces basal autophagy without detrimental effects on cell survival, and reduces misfolded protein content in protein misfolding-stressed cells<sup>142</sup>. The 3' UTR of *ATG7* in the chick has not been annotated and therefore the 2,000bp following

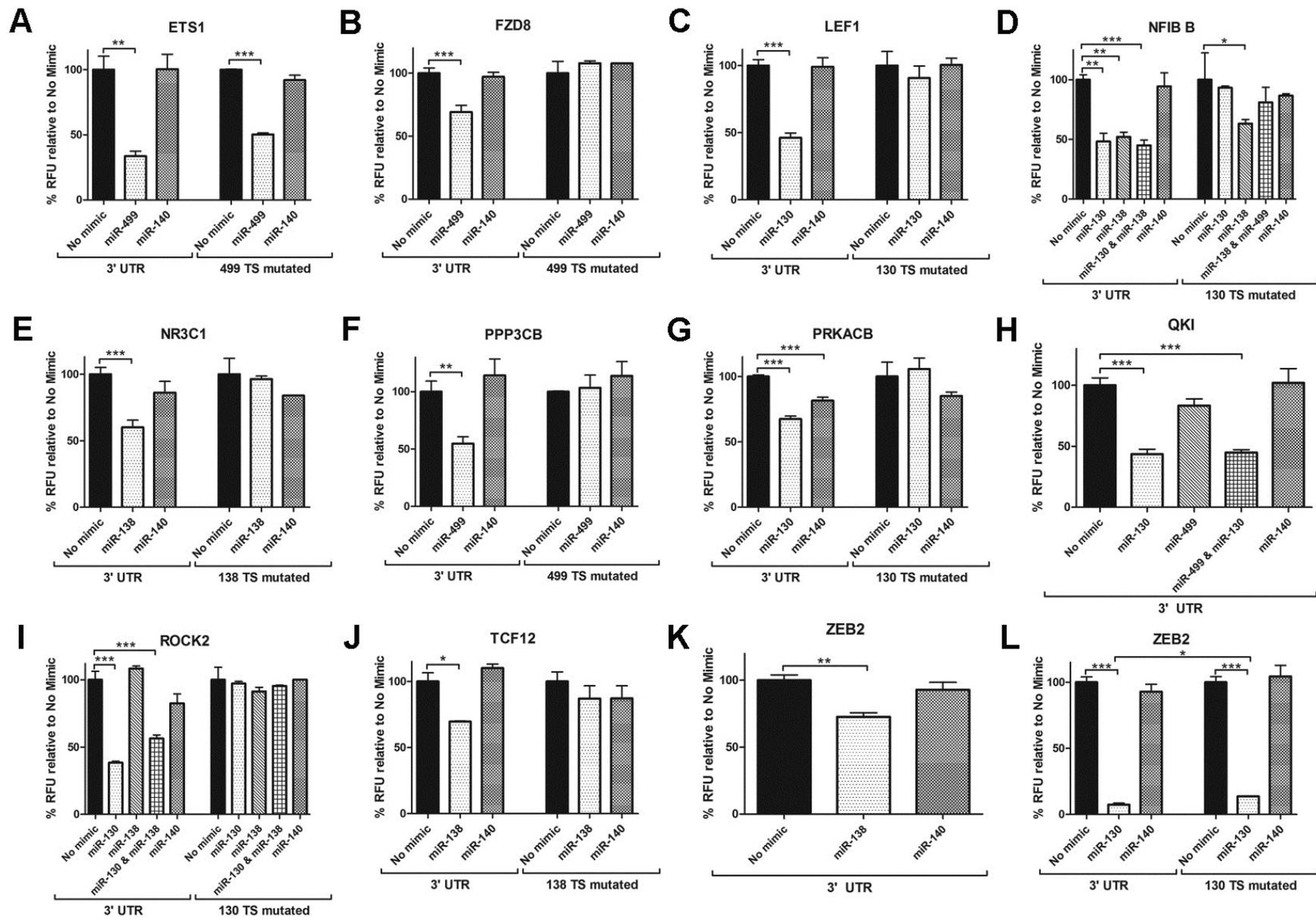
the coding region of this gene were cloned into a luciferase miRNA-sensing construct<sup>47</sup> and assayed with the two predicted targeting miRNAs, miR-138 and miR-499, either individually or together. miR-138 reduced luciferase levels by around 50%, but not when the predicted miR-138 target site was mutated (**Figure 4.12A**). The effects of miR-499 (**Figure 4.12A**) were found to not be due to the predicted target site (**Figure 4.12B**).



**Figure 4.12: The effects of two miRNAs, miR-138 and miR-499, on luciferase reporter construct containing the ATG7 3' UTR. A:** The unmodified 3' UTR and the 3' UTR with the miR-138 target site mutated were assayed with the microRNA mimics indicated below the graph. **B:** The miR-499 site and both the miR-499 and miR-138 sites were mutated and assayed with miRNA mimics as shown. One way ANOVA with Tukey post test statistical analyses were performed on the results of each construct.  $0.05 \geq P > 0.01$  : \*;  $0.01 \geq P > 0.001$  : \*\*;  $0.001 \geq P > 0.0001$  : \*\*\*;  $P \leq 0.0001$  : \*\*\*\*.

Both splice variants of the v-ets avian erythroblastosis virus E26 oncogene homolog 1 gene (*ETS1*) contain single predicted target sites for miR-499 in their 3' UTRs<sup>138</sup>. *ETS1* is a transcription factor that regulates *Gata4* and is briefly active in multiple cardiac lineages early in cardiac development but restricts to the endocardium where it remains active through cardiogenesis<sup>143</sup>. A reporter assay showed that miR-499 translationally represses *ETS1* (**Figure 4.13A**), although mutation of two nucleotides in the seed region of the target site did not abrogate the repression by miR-499. The online miRNA target prediction tool RNAHybrid<sup>144</sup> predicted a second, more energetically favourable target site for miR-499 located roughly 100bp upstream of the TargetScan predicted target site (**Figure 4.14**). This predicted target site contains G:U wobble base pairings in the seed region, that have been demonstrated not to interfere with miRNA binding<sup>145</sup>. The translational repression of *ETS1* by miR-499 may be a result of interaction with this site, although additional reporter assays

on a mutant construct with the extended predicted site mutated would allow further elucidation of the binding of miR-499 to this 3' UTR.



**Figure 4.13: Reporter assay results of eleven genes found to contain miRNA target sites in their 3' UTRs.** In each case the reporter construct is shown on the left (3' UTR) and the construct with a mutated target site, where available, is shown on the right. The miRNA mimics used are indicated below each bar. 130 TS mutated: miR-130 target site was mutated. 138 TS mutated: miR-138 target site was mutated. 499 TS mutated: miR-499 target site was mutated. **A:** ETS1. **B:** FZD8. **C:** LEF1. **D:** NFIB 3' UTR from 3,600 – 5,000bp. The miR-130 target site was mutated. **E:** NR3C1. **F:** PPP3CB. **G:** PRKACB. **H:** QKI. **I:** ROCK2. The miR-130 target site was mutated. **J:** TCF12. **K:** ZEB2 assayed with miR-138. **L:** ZEB2 assayed with miR-130. One way ANOVA with Tukey post test statistical analyses were performed on the results of each construct.  $0.05 \geq P > 0.01$  : \*;  $0.01 \geq P > 0.001$  : \*\*;  $0.001 \geq P > 0.0001$  : \*\*\*;  $P \leq 0.0001$  : \*\*\*\*.

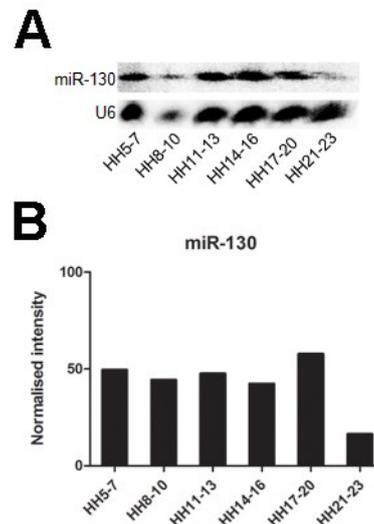


**Figure 4.14: The 3' UTR of ETS1 has a second predicted target site for miR-499.** **A:** The online miRNA target site prediction tool RNAHybrid<sup>144</sup> predicted a second, more energetically favourable target site for miR-499 in the 3' UTR of ETS1, with wobble base pairing. **B:** The 3' UTR of ETS1 aligned with the miR-499 target site predicted by TargetScan indicated by the black arrow. A construct containing the 3' UTR of ETS1 with the third and fifth nucleotide of the target site mutated did not significantly increase reporter levels ( ). Half of the seed region of miR-499 is involved in the RNAHybrid predicted target site.

TargetScan predicted a target site for miR-499 within *Frizzled8* (*FZD8*) mRNA. *FZD8* is a receptor in the non-canonical Wnt signalling pathway<sup>7</sup>. Although *FZD8* does not currently have an annotated 3' UTR<sup>138</sup>, a predicted target site for miR-499 can be found approximately 700bp from the end of the coding region of this gene. This target site was responsible for the inhibition of the translation of a luciferase reporter construct containing the 1,000bp sequence from the end of the coding region of *FZD8* (**Figure 4.13B**).

The 3' UTR of Lymphoid enhancer binding factor 1 (*LEF1*) contains a single miR-130 putative target site, as predicted by TargetScan. *LEF1*, a component of the canonical Wnt signalling pathway, is required for the formation of major blood vessels in *X.tropicalis*<sup>146</sup>.

The 3' UTR of *LEF1* contains a single miR-130 predicted target site, which was found to induce translational repression in a luciferase reporter construct containing the entire annotated 3' UTR of *LEF1* (**Figure 4.13C**). miR-130 has not been detected in the heart by LNA ISH from HH19<sup>147</sup>, but northern blot analysis of the developing chick heart showed expression of miR-130 from HH5 to HH23 (**Figure 4.15**).



**Figure 4.15: Northern blot analysis of miR-130 expression in the developing chick heart.** **A:** Northern blot for miR-130 and U6 on RNA purified from HH5-7, HH8-10, HH11-13, HH14-16, HH17-20 and HH21-23 cardiogenic or heart tissue. A lower amount of HH8-10 RNA was loaded onto the polyacrylamide gel. **B:** Quantification of miR-130 band intensity relative to U6 band intensity performed with ImageJ software<sup>148</sup> showed relatively constant levels of miR-130 until HH21-23, where much less miR-130 was detected in the embryonic chick heart.

NR3C1, the nuclear receptor subfamily 3 group C member 1, is predicted to have a miR-138 target site in its 3' UTR. NR3C1 is a glucocorticoid receptor<sup>138</sup>. The normal development of embryonic organs is fundamentally dependent on glucocorticoids, but adult cardiovascular function is negatively impacted by exceptionally elevated levels of glucocorticoid exposure during development<sup>149</sup>. The annotated 3' UTR of *NR3C1* in ENSEMBL 71 is only 207bp long, but a target site for miR-138 is predicted approximately 2,000bp downstream of the end of the annotated 3' UTR. A luciferase reporter construct containing the sequence from 1,300 – 2,500bp from the start of the annotated *NR3C1* 3' UTR is repressed by miR-138 specifically at the predicted target site (**Figure 4.13E**).

The *PPP3CB* gene encodes calcineurin A,  $\beta$  isozyme, which is predicted to have miR-499 target sites in both its splice variants. Calcineurin plays a crucial role in the development of myocardial hypertrophy<sup>150</sup>. The 3' UTRs of both *PPP3CB* splice variants contain a miR-499 target site in the 500bp their 3' UTRs have in common. This portion of the 3' UTR was cloned into a luciferase reporter construct, which was targeted by miR-499 at the predicted target site (**Figure 4.13F**).

A miR-130 target site is predicted by TargetScan in the 3' UTR of the *PRKACB* gene. cAMP-dependent protein kinase catalytic subunit beta of Protein Kinase A (PKA) is formed from *PRKACB*. The signalling molecule cAMP activates PKA which then phosphorylates target proteins. miR-130 inhibited a luciferase reporter construct containing the first 1,200bp of the 3' UTR of *PRKACB* (**Figure 4.13G**).

The QKI KH domain containing RNA binding protein (QKI), predicted to have two miR-130 and two miR-499 target sites by TargetScan, is essential for blood vessel development but not for heart muscle differentiation or function<sup>151</sup>. None of the splice variants of this gene have an annotated 3' UTR at this time. A construct was prepared containing sequence from approximately 1,700bp to 2,300bp after the end of the coding sequence. This construct was predicted to contain two miR-130 and two miR-499 target sites. miR-130 was found to repress the translation of *QKI* (**Figure 4.13H**). However, due to time constraints a mutant target site of this reporter construct was not prepared and therefore it not certain whether the predicted target site of miR-130 is required for the translational repression observed. miR-499 was found not to target *QKI*.

*ROCK2*, the Rho-associated coiled-coil containing protein kinase 2 gene, is predicted by TargetScan to have binding sites for miR-130 and miR-138. ROCK2 negatively affects the canonical Wnt signalling pathway by binding to and ensuring the degradation of TGF $\beta$  type I receptors<sup>152</sup>. ROCK2 has been shown to prevent mesoderm induction by ectopic Nodal signals in zebrafish<sup>152</sup>, and has many other known functions including controlling VSMC contraction<sup>153</sup> and inducing endothelial cell differentiation and migration during endocardial cushion development<sup>154</sup>. miR-130, but not miR-138, was found to repress a reporter construct containing the 50bp 3' UTR of *ROCK2* along with 950bp of downstream

sequence (**Figure 4.13I**). The repression was shown to require miR-130 binding to a predicted target site 780bp downstream of the annotated end of the *ROCK2* 3' UTR.

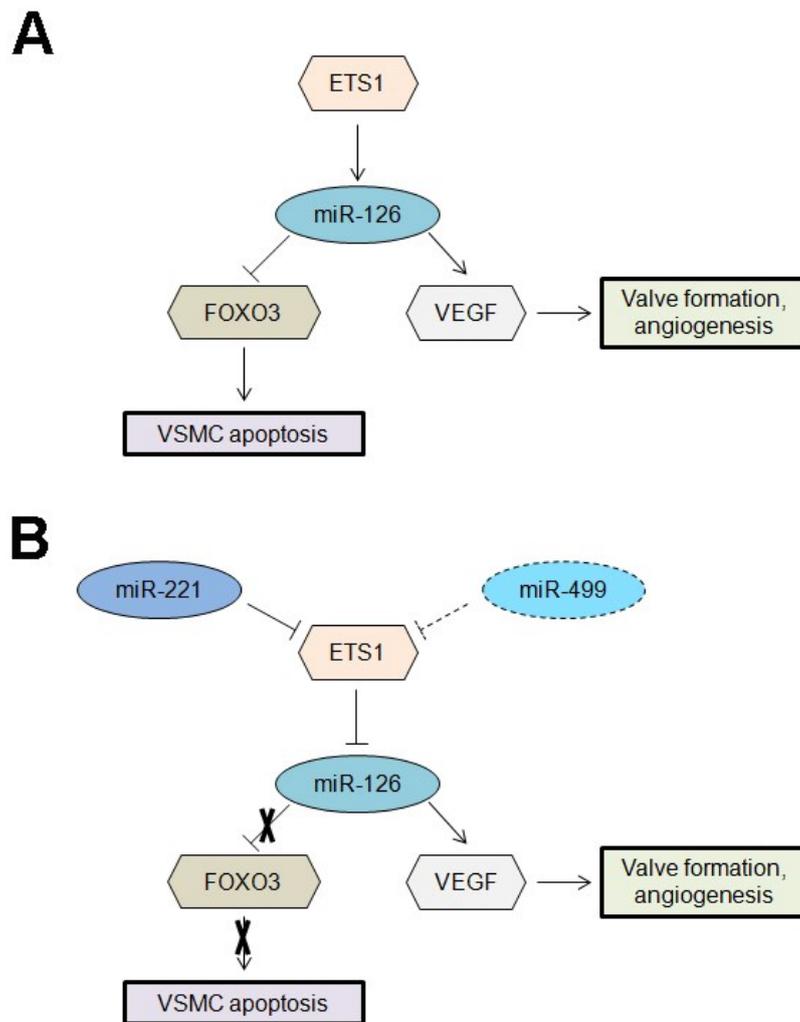
miR-138 was found to repress a reporter construct containing the 3' UTR of the *Transcription factor 12 (TCF12)* gene via a predicted target site (**Figure 4.13J**). Canonical Wnt signalling induces the transcription of target genes by the interaction of  $\beta$ -Catenin with the TCF/LEF family of transcription factors<sup>7</sup>.

A reporter construct containing the 3' UTR of Zinc finger E-box binding homeobox 2 (*ZEB2*) gene was inhibited on transfection with miR-138 (**Figure 4.13K**). Unfortunately a mutant construct was not prepared and therefore it is unknown whether this inhibition can be disrupted by mutation to the predicted miRNA binding site. The same 3' UTR construct was repressed to less than 10% of the control level by miR-130 (**Figure 4.13L**), and although mutation of the predicted miR-130 target site did statistically reduce the repression, the difference is only 6% and therefore supplementary binding of miR-130 to this 3' UTR is likely.

## 4.3 Discussion

### 4.3.1 Few miRNAs have large increases in expression

Although 423 known animal miRNAs annotated in miRBase release 20 were detected in more than one of nine developing chick samples, most had normalised read values equal to or less than that recorded in the HH5-7 sample (**Figure 4.3**). All of the remaining miRNAs were found to be upregulated compared to their HH5-7 levels. Most of these show an increase of between 1 and 10 fold, and a few increased between 100 and 300 fold over HH5-7 in at least one sample (**Table 4.1**). Six of the fourteen miRNAs upregulated between 100 and 300 fold (miR-1, miR-10, miR-126, miR-145, miR-451, miR-499) have known roles in heart development and function<sup>74,92,106,110,116</sup> (**Table 4.6**), but have varied fold change patterns (**Figure 4.4**). miR-126, miR-451 and miR-499 (**Figure 4.4C, E, F**) all have large increases in expression levels in the HH17-20 Heart sample. Valve formation is known to start at this stage as a result of the cardiac neural crest entering the looped heart, and therefore the concomitant increased expression of these miRNAs at this stage may indicate functions for these miRNAs during this process. Indeed, miR-126 has been reported to be required for valve formation by its positive effects on Vascular Endothelial Growth Factor (VEGF)<sup>110</sup>, supporting this hypothesis. During valve development VEGF signalling is essential in endocardial cells at the outflow tract cushions but not in endocardial cells at the atrioventricular cushions<sup>110</sup>. ETS1 is a reported activator of miR-126 expression<sup>74</sup> (**Figure 4.16A**), and has also been found to be a target of miR-499 in this study (**Figure 4.13A**), which suggests that miR-499 might function to limit VEGF signalling during valve formation by indirectly reducing the levels of miR-126 in the atrioventricular cushions.



**Figure 4.16: miR-499 is implicated to regulate miR-126 by targeting ETS1.** miR-126 is essential for valve formation and angiogenesis, but also increases VSMC apoptosis by repressing FOXO3 (**A**). **B**: miR-221 targets ETS1, the transcriptional activator of miR-126, relieving some of the pro-apoptotic effects of miR-126. miR-499 has been found to repress ETS (**Figure 4.13A**), and may function in a similar way to miR-221 in VSMCs.

The fold change patterns of miR-126 and miR-499 have been found to be highly significantly correlated (**Figure 4.7A**, **Figure 4.9C**), as have the patterns of miR-126 and miR-221(**Figure 4.9C**). miR-221 is known to target ETS1 and has been shown to increase SMC proliferation<sup>155</sup>, in contrast to the pro-apoptotic effects of miR-126 on SMCs<sup>156</sup> (**Figure 4.16B**). Inhibition of ETS1 by miR-499 may influence miR-126 levels, and as a result valve formation, in a similar way to the interplay between miR-126 and miR-221 in SMCs.

### 4.3.2 Creating a profile of miRNAs in Heart by subtraction of NT sample values

The subtraction of fold change values of miRNAs in the NT samples at later stages (HH14-16, HH17-20, HH21-23) from their values in the corresponding Heart samples resulted in the generation of a profile of miRNA expression in the heart over these stages of development (4.2.2). Only 38 of all the known animal miRNAs detected at these time points were found to be more than 10 fold higher in Heart samples than in NT samples (Table Apx A.6), including 10 miRNAs known to be important for heart development and function (Table 4.4). The upregulation of the remaining 28 miRNAs in Table Apx A.6 in Heart over NT samples suggests that some of these miRNAs may also have important functions in the foetal or adult heart.

miRNAs with fold change profiles, after subtraction of NT reads, that correlate significantly with those of 26 heart miRNAs selected from literature (Table 4.4) and miR-130 were identified (Table 4.3, Table Apx A.7, Table Apx A.8, and Table Apx A.9). This analysis revealed networks of miRNAs that have similar NT – H fold change profiles to known heart miRNAs and provides an insight into groups of miRNAs that may function in concert to regulate heart development and function.

### 4.3.3 miRNAs with a role in heart development and function were detected

Clustering 26 miRNAs known to have roles in heart development and function by their fold change in normalised reads over levels at HH5-7 showed that most have broadly similar expression patterns in the samples assayed (Figure 4.5). miR-1 and miR-10 did not correlate well with any of the others, and are both very highly upregulated in non-heart samples (Figure 4.4A and B). As discussed above, miR-126 and miR-499 cluster together away from the remaining miRNAs. miR-143 and miR-133 were found to be highly correlated and have been shown to have negative effects on the proliferation of VSMCs and cardiomyocytes respectively<sup>106</sup>. Furthermore, members of the miR-17-92 cluster (miR-17, miR-18, miR-19, miR-20 and miR-92) were shown to correlate closely, as did members of the miR-15 family (miR-15 and miR-16).

The 26 miRNAs with reported functions in heart development and function were found to represent between 65% (HH14-16 H) and 85% (HH21-23 NT) of the total reads from 423 known miRNAs identified in more than one of the nine chick samples (**Figure 4.6A**). However, this was mostly due to the effects of two highly expressed miRNAs with many known functions during development: miR-10<sup>157</sup> and miR-92<sup>158</sup>. Once the reads from these two miRNAs were removed, increased expression of the remaining 24 miRNAs in heart samples over non-heart samples emerged from HH17-20 (**Figure 4.6B**). This enrichment coincides with valve formation and septation, which takes place in the chick heart at HH17-20 and HH21-23 respectively. It is therefore conceivable that these 24 miRNAs may contribute to cardiac patterning at these stages.

Further bioinformatic analysis of the targets of these miRNAs might reveal whether they share targets in the pathways that control these essential events, giving a starting point for further elucidation of the regulation of heart patterning. Eleven of these 24 miRNAs were selected for such further analysis, based on criteria shown in **Table 4.6**, and the results are discussed in section **4.3.4**.

#### **4.3.4 11 Known heart miRNAs correlate well by fold change**

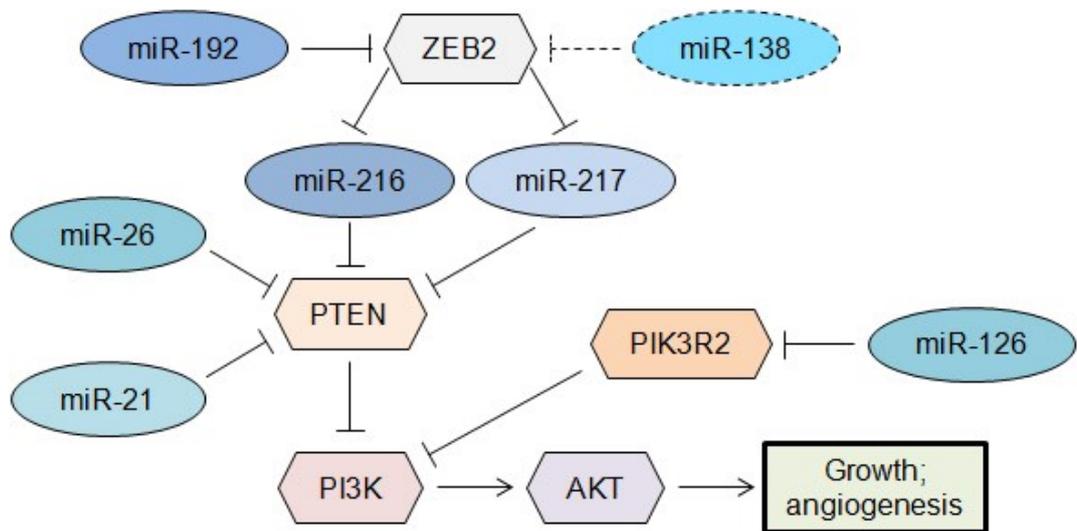
Even though nine of eleven chosen heart miRNAs were found to significantly correlate with at least one of the others (**Figure 4.7A and B**), correlations were generally rare with less than a third of the remaining 412 known animal miRNAs expressed in more than one sample correlating with these eleven miRNAs (**Figure 4.7C**). Many of the correlating miRNAs were found to be poorly conserved (**Figure 4.7C**), leaving only 68 correlating miRNAs that are also conserved in humans (**Table Apx A.10, Table Apx A.11**). As the 11 heart miRNAs (**Table 4.6**) themselves are highly conserved, this small subset of conserved miRNAs that correlate to them may have important functions similar to the 11 chosen heart miRNAs that have yet to be elucidated. Further analysis of predicted targets and expression patterns of these correlated miRNAs should give a good starting point in estimating whether these miRNAs are likely to have important biological functions in common with their correlating heart miRNAs.

#### 4.3.5 PTEN is an enriched target of most of the miRNAs with functions in the heart

Deletion of the phosphatase and tensin homolog gene, *PTEN*, in a cardiomyocyte-specific manner resulted in hypertrophic growth and reduction in autophagy in mice<sup>129</sup>. However, its expression in the endothelium, the layer of cells on the inside of blood and lymphatic vessels, is essential for normal development<sup>130</sup>. Many of the 26 miRNAs with reported roles in heart development and function (**Table 4.4**), have *PTEN* as a confirmed or predicted target (**Figure 4.8A**). The majority (8) of the eleven selected heart miRNAs have been experimentally proven to target *PTEN*, and one more is predicted to target it. Surprisingly around 20% of all the sequenced miRNAs significantly correlated to the 11 heart miRNAs are also either predicted or confirmed to target *PTEN* (**Figure 4.8B**). Furthermore, a larger enrichment of *PTEN*-targeting was observed when only miRNAs conserved in humans were considered. An effort was made to determine whether the enrichment of correlated *PTEN*-targeting miRNAs to the known heart miRNAs is unique by assessing the chondrogenic miR-140 and the skeletal muscle-specific miR-206, along with their correlated miRNAs. Neither miR-140 nor miR-206 have been shown or predicted to target *PTEN*, and less than 10% of their correlated miRNAs were *PTEN*-targeting miRNAs. A more comprehensive bioinformatic comparison with many more miRNAs lacking functions in the heart would more fully test this observation. Although the inhibition of this gene is essential for growth in many tissues<sup>159</sup>, a large number of miRNAs with expression patterns similar to those with essential roles during cardiac development are known or predicted to inhibit the translation of *PTEN*, suggesting that a large network of miRNAs with similar expression patterns in chick heart development allow growth by repressing *PTEN*. A pseudogene of *PTEN*, *PTEN1*, has recently been discovered that has multiple binding sites of miRNAs that also target *PTEN* in its 3' UTR<sup>133</sup>. This has led to the speculation that pseudogenes may function as lures for miRNAs, leaving the target less affected. This discovery highlights the very complex nature of mRNA regulation by miRNAs, and further supports the study of these ubiquitous interactions.

*PTEN* negatively regulates the powerful PI3K-AKT cascade to limit growth and angiogenesis, and it is modulated by miR-21<sup>160</sup> (**Figure 4.17**). miR-126 is essential for

angiogenesis in the developing heart<sup>74</sup>, and it has been shown to target PI3-kinase regulatory subunit- $\beta$  (PIK3R2), a negative regulator of PI3K and therefore of growth and angiogenesis. Interestingly, miR-126 and 54% of its conserved correlated miRNAs are predicted or confirmed to target PTEN (**Figure 4.8B**). Attenuation of the negative regulation of the PI3K-AKT cascade allows growth and angiogenesis to take place without the risk that a more permanent relief of downregulation would bring. It appears that this repressor of growth and angiogenesis is a target of a large number of miRNAs with many different expression profiles in the heart, in keeping with the role of miRNAs in “fine tuning” the expression of essential targets.



**Figure 4.17: The attenuation of PTEN, a negative regulator of growth and angiogenesis, by miRNAs.** PTEN prevents the powerful PI3K-AKT signalling cascade from causing senescence or cancer by directly targeting PI3K, thus reducing growth and angiogenesis. Many miRNAs, such as miR-21, miR-26, miR-216 and miR-217 are known to directly target PTEN and reduce its negative effects on growth. Other miRNAs have been shown to induce growth after indirect reduction in PTEN levels by inhibiting a repressor of miRNAs that directly target PTEN (ZEB2, miR-192), or by repressing a negative regulator of growth that targets PI3K downstream of PTEN (miR-126, PIK3R2). miR-138 has been found to target ZEB2 *in vitro* in this study (**Figure 4.13K**).

Zinc Finger E-box binding homeobox 2 (ZEB2) is a DNA-binding repressor of transcription known to interact with activated SMADs, factors that induce transcription as a result of TGF- $\beta$  signalling, which are essential during development<sup>161</sup>. ZEB2 is known to repress the transcription of the PTEN-targeting miR-216 and miR-217 (**Figure 4.17**). However, TGF- $\beta$  induces the transcription of miR-192 which represses ZEB2 and thus allows the inhibition of PTEN by miR-216 and miR-217<sup>160</sup>. miR-138 has been shown to target ZEB2 *in vitro* (**Figure 4.13K**), and has a fold change pattern significantly correlated

to that of miR-192 (**Figure 4.9H**). This suggests that miR-138 might, along with miR-192, allow the inhibition of PTEN by targeting ZEB2 (**Figure 4.17**). Both miR-216 and miR-217 were detected in all nine samples assayed in this study, with average normalised reads around 70 and 20 reads per million respectively. miR-216 and miR-217 have not been studied during chick development, and this analysis provides a starting point for further elucidation of their potentially important roles during heart development.

#### 4.3.6 Selected correlated miRNA fold change patterns

The fold change patterns of miR-101, miR-19 and miR-190 were found to be significantly correlated (**Figure 4.9A**). miR-101 is known to inhibit the proliferation of cardiac fibroblasts<sup>109</sup>, whereas miR-19 not only induces cardiomyocyte proliferation<sup>98</sup> but also decreases neural stem cell differentiation<sup>162</sup>. This suggests that miR-101 and miR-19 may function to expand the cardiomyocyte population and ensure commitment to the cardiac lineage in these cells. Furthermore, common GO terms significantly associated with the predicted and confirmed target genes these three miRNAs were associated with nervous system development, neurogenesis, neuron differentiation, and generation of neurons (**Table 4.7**). The extent to which miR-101, miR-19 and miR-190 ensure commitment to the cardiac fate may be elucidated by their inhibition in the developing heart.

miR-100 ensures the proper induction of the foetal cardiac gene programme by its indirect repression of adult cardiac genes<sup>134</sup>. miR-100, let-7 and miR-125 are clustered within 10Kbp of each other on the same chromosome and are co-expressed in most animals. The human genome contains three such clusters, indicating possible essential roles for all these miRNAs<sup>163</sup>. In the chicken genome, miR-125 is still clustered with let-7 but miR-100 is present on a separate chromosome, also clustered with let-7 (**Figure 4.10A**). Interestingly, despite being located on separate chromosomes, the expression change profiles of miR-100, let-7 and miR-125 are still very highly correlated (**Figure 4.9B**). The conserved miRNA miR-99, a member of the same family as miR-100, is clustered with let-7c and miR-125 in the chick (**Figure 4.10A**). Although this miRNA is not part of the miR-100 / let-7 / miR-125 cluster in humans, it is located very close to let-7c on another chromosome (Chr 21). The only GO term significantly associated with all four miRNAs is very general

(GO:0009893, positive regulation of a metabolic process), but this is mainly due to a dearth of predicted targets for miR-99 and miR-100. The conservation of all four miRNAs suggests that, along with let-7, they may have important functions in higher vertebrates and possibly in embryonic heart development.

The possible implications of the correlation between the expression patterns of miR-126 and miR-499 (**Figure 4.9C**) were discussed above (**4.3.1**).

miR-126 negatively regulates the repressors of the VEGF signalling pathway and as a result is required for the development of heart valves and blood vessels<sup>74,110</sup>. miR-126 and miR-451 were found to have highly correlated patterns of expression (**Figure 4.9D**). miR-451 is a known inducer of erythroid maturation<sup>112</sup>. In the chick valve formation begins around HH17, and the rapid increase in miR-126 expression in the HH17-20 sample confirms its role in this process (**Figure 4.9D**). Very few of the confirmed targets of miR-126 were also predicted targets of miR-451, but the majority of GO terms significantly associated with the predicted targets of miR-451 were also associated with the confirmed targets of miR-126. This suggests that miR-126 and miR-451 may have similar effects during development by targeting similar pathways, even if they do not target the same genes. It would be interesting to determine whether miR-451 is required for valve development or angiogenesis, and whether inhibition of this miRNA recapitulates the phenotype observed when miR-126 is absent.

A significant negative correlation between the fold change patterns of miR-130 and miR-10 was observed (**Figure 4.9E**). miR-10 is known to have many functions during development<sup>157</sup>, and is pro-angiogenic as a result of its effects on the VEGF signalling network<sup>92</sup>. However, miR-10 levels were shown to increase rapidly when human embryonic stem cells were induced to form neuronal cells following treatment with retinoic acid<sup>164</sup>, and miR-10 was found to occupy a larger portion of the total normalised reads in non-heart NT samples than in heart samples (**Figure 3.7D** and **F**). Furthermore, very few of the targets (**Table 4.8**) and only a quarter of the 100 most significant GO terms associated with the predicted and confirmed targets of miR-130 and miR-10 were common to both miRNAs. miR-130 may have effects opposite to miR-10 in the developing embryo,

but not enough information is currently available on miR-130 functions *in vivo* to substantiate this.

The LIM homeobox transcription factor Islet-1 (Isl1) is required for cardiac specification<sup>120</sup>, cardiomyocyte differentiation<sup>136</sup>, vasculogenesis<sup>120</sup> and cardiac EMT<sup>117</sup>. Isl1 is targeted by miR-31, which results in decreased cardiac fibrogenic EMT<sup>117</sup>. The expression of *Isl1* has recently been shown to be promoted by Wnt signalling via a novel TCF/LEF1 binding site, which was identified during early cardiomyocyte differentiation<sup>165</sup>. *Isl1* is expressed in the chick cardiac crescent and heart tube (HH6-11), but its expression becomes limited to the second heart field in the pharyngeal arches later on (HH14-24)<sup>17,147</sup>. However, miR-130 and miR-31 were found to have very similar changes in expression (**Figure 4.9F**), and miR-130 targets LEF1 *in vitro* (**Figure 4.13C**). These results suggest several possible roles for miR-130 in heart development, and further studies are required to assess whether the inhibition of miR-130 has functional effects on heart development, possibly resulting from its effects on the *Isl1* transcriptional activator, LEF1. Interestingly, miR-138 was found to target TCF12 (**Figure 4.13J**), and is known to be essential for normal cardiac looping<sup>111</sup>. An investigation of the combinatorial effect of dysregulation of miR-130 and miR-138 on *Isl1* and therefore on the embryonic chick heart may yield interesting results.

miR-138 and miR-126 as well as miR-138 and miR-192 were found to have significantly similar fold change profiles (**Figure 4.9G, H**). The possible related effects of miR-138, miR-192 and miR-126 were discussed in 4.3.5 (**Figure 4.17**).

#### 4.3.7 Validating targets of miR-130, miR-138 and miR-499

Three miRNAs and 15 of their predicted target genes were selected for further investigation. Luciferase reporter construct assays were used to determine the effects of miR-138, known to be essential for patterning in zebrafish during heart looping, and miR-499, associated with decreased cardiomyocyte apoptosis. miR-130 was assessed in a similar way, and was selected based on its presence in the top detected miRNAs in Heart but not NT samples (**Figure 3.7**).

miR-138 and miR-499 have been shown to target the 3' UTR of *ATG7* (**Figure 4.12**) to a moderate degree. In zebrafish, the loss of *ATG7* leads to defective heart looping and malformations in heart structures, amongst other cardiac effects<sup>166</sup>. Although autophagy is essential for the development of many heart structures, it is possible that its downregulation by these miRNAs is required in other tissues during cardiogenesis, or that miR-138 and miR-499 function to ensure *ATG7* levels stay within a certain range. The knockdown observed as a result of translational repression brought on by treatment with miR-138 and miR-499 was not very large. This could also suggest that the targeting of *ATG7* by these miRNAs is simply a coincidence, with limited functional effects. However, both target sites are broadly conserved among vertebrates, although only the miR-138 site is retained in humans. Such conservation is often an indication of a functional role for the targeting miRNAs, which may be confirmed by overexpression of *ATG7* and inhibition of these miRNAs *in vivo*.

miR-130 and miR-138 were found to target a region of sequence approximately 2,000bp downstream of the annotated end of the *NFIB* 3' UTR, and the predicted miR-130 target site was experimentally validated (**Figure 4.13D**). Expression of *NFIB* has only been detected in adult mouse hearts and is thought to direct tissue specific gene expression during development<sup>139</sup>, suggesting that these miRNAs may ensure tissue identity during development by translationally repressing this gene. Therefore an upregulation of *NFIB* as a result of miR-130 and miR-138 inhibition may result in a loss of commitment to a cardiac fate in developing embryos.

*ETS1*, a transcription factor that regulates *Gata4*<sup>143</sup> and activates the transcription of miR-126<sup>74</sup>, is a target of miR-499 *in vitro* (**Figure 4.13A**). However, the mutation of two nucleotides in the predicted target site did not remove the repression. This could be due to the presence of another non-canonical target site within the 3' UTR being responsible for most of the inhibition observed, or by miR-499 binding proceeding despite the mutation of two nucleotides in the predicted binding site. As miR-126 is intimately involved in angiogenesis and valve formation, regulation of its activator by miR-499 adds a layer of fine control over these essential processes (**4.3.1**).

Both canonical and non-canonical Wnt signalling is essential at different stages and in different tissues during heart development<sup>7</sup>. FZD8 is a receptor for the non-canonical WNT-5A<sup>167</sup>, and was repressed by the action of miR-499 on the target site in its 3' UTR (**Figure 4.13B**). Non-canonical Wnt signalling, via WNT5A, is required in cardiac neural crest cells that have migrated into, but not in those still migrating towards, the outflow tract for normal septation to proceed. Furthermore, canonical Wnt signalling must be high in the endocardial cushions for valve formation to occur normally between HH17 and HH20. The levels of miR-499 detected at HH17-20 in the heart sample show a large increase during valve formation. It is therefore possible that miR-499 ensures domination of canonical Wnt signalling in the endocardial cushions by targeting *FZD8*, as well as preventing premature non-canonical Wnt signalling in migrating cardiac neural crest cells. If so, these actions would assure both normal septation and valve formation in the developing heart, which may be prevented following inhibition of miR-499 *in vivo*.

miR-130 levels have been shown to increase during this phase of heart development (**Figure 4.9E**), and miR-130 targets *LEF1* *in vitro* (**Figure 4.13C**). Although *LEF1* is required for the formation of major blood vessels in the heart<sup>146</sup>, non-canonical Wnt signalling must be high for normal heart looping to occur<sup>7</sup>. miR-130 appears to be virtually absent from the heart from HH21 (**Figure 4.15**), concomitant with increased canonical Wnt signalling in the endocardium which is required to support endocardial cell proliferation<sup>7</sup>. This implies that the inhibitory effect of miR-130 on canonical Wnt signalling is necessary during earlier development, but not during later proliferation. Artificially increasing the levels of miR-130 in the heart at HH21-23 may prevent this endocardial proliferation *in vivo*.

A miR-138 target site approximately 2,000bp downstream of the end of the annotated 3' UTR of *NR3C1* is responsive to miR-138 (**Figure 4.13E**). The function of the adult heart is negatively affected by elevated levels of glucocorticoids during development<sup>149</sup>. As *NR3C1* is a glucocorticoid receptor, it is conceivable that miR-138 attenuates the levels of glucocorticoids cells are exposed to during development, thus preventing poor cardiovascular health in adults. Testing the validity of this conclusion would involve *in*

*vivo* studies on the effects in the adult of reduction of embryonic miR-138 levels, and would therefore be quite challenging.

Calcineurin is known to activate glycogen synthase kinase-3 $\beta$  (GSK-3 $\beta$ ), a molecule intimately involved in canonical Wnt signalling<sup>168</sup>. PPP3CB, the  $\beta$  isozyme of calcineurin A, was translationally inhibited by miR-499 *in vitro* at a predicted target site (**Figure 4.13G**). Therefore, translational inhibition of PPP3CB may result in reduced canonical Wnt signalling as a result of miR-499 action. miR-499 expression peaks at HH17-20 in the heart (**Figure 4.9C**), a time point which coincides with valve formation. At this time canonical Wnt signalling is repressed in cells undergoing epithelial to mesenchymal transition (EMT) in the endocardial cushions as part of valve formation<sup>7</sup>. Once the EMT is complete canonical Wnt signalling is again upregulated in these cells to induce their proliferation. miR-499 may attenuate canonical Wnt signalling in the endocardial cushions by repressing *PPP3CB*, a known activator of GSK-3 $\beta$ , thereby allowing EMT of endocardial cells as part of valve development in the chick. However, whether this action of miR-499 on *PPP3CB* is biologically relevant is yet to be determined.

miR-130 was shown to target *PRKACB* via a predicted target site in its 3' UTR (**Figure 4.13G**). *PRKACB* is an essential part of PKA, which phosphorylates target proteins. Heart failure results when calcium release channels on the sarcoplasmic reticulum of cardiac muscle become sensitised to calcium-induced activation following hyperphosphorylation of the ryanodine receptor 2 (RyR2) by Protein Kinase A (PKA)<sup>169</sup>. The targeting of *PRKACB* by miR-130 indicates that miR-130 may reduce the PKA phosphorylation of RYR2 and therefore protect heart muscle from the negative effects of hyperphosphorylation.

Both miR-130 and miR-499 were predicted to target *QKI*. Although *QKI* is known to play a role during blood vessel development, cardiac function and muscle differentiation remains normal when *QKI* is knocked out<sup>151</sup>. This gene appears to be essential for yolk sac vascular remodelling and vasculogenesis in the developing chick embryo.

**Table 4.9: Many miRNAs with known roles in heart development and function have conserved predicted target sites in the 3' UTR of *QKI*.**

miRNAs with conserved predicted target sites in the 3' UTR of <i>QKI</i>			
miR-1	miR-20	miR-125	miR-181
miR-10	miR-24	miR-128	miR-221
miR-15	miR-30	miR-133	miR-222
miR-16	miR-31	miR-138	miR-499
miR-17	miR-92	miR-143	
miR-18	miR-101	miR-145	

miR-130, but not miR-499, was found to repress the translation of a *QKI* reporter plasmid (**Figure 4.13H**). Interestingly, this gene contains conserved predicted target sites for 22 of the 26 known heart miRNAs examined previously (**Figure 4.5, Table 4.9**). Only four miRNAs did not have predicted target sites in this gene: miR-126, also known to be essential for blood vessel development<sup>74</sup>, and three others, two of which (miR-19, miR-21) have fold change patterns significantly correlated to that of miR-126. Although some of the predicted target sites are likely to not be true target sites, such an enrichment of heart miRNA target sites may indicate an essential need for the repression of *QKI* during normal heart development. It is worth noting that 17 splice variants are annotated for this gene, with 15 encoding *QKI* proteins of various sizes. Despite the large number of splice variants, the 3' UTR of *QKI* is not currently annotated in the latest chicken genebuild (Galgal4<sup>138</sup>). It would be very interesting to determine whether the splice variants have unique 3' UTRs with different predicted miRNA target sites.

A luciferase reporter containing the entire 3' UTR of *ROCK2*, along with a 950bp stretch of sequence downstream of the end of the annotated 3' UTR, was inhibited by the action of miR-130 on a predicted target site (**Figure 4.13I**). The *ROCK2* gene is known to play many essential roles during development, ranging from inhibition of mesoderm induction<sup>152</sup> to VSMC contraction<sup>153</sup> and endothelial cell differentiation and migration in endocardial cushions<sup>154</sup>. It is difficult to suggest what function the inhibition of *ROCK2* by miR-130 is likely to have, and further validation and examination of this interaction *in vivo* may yield some interesting results.

TCF12 expression is repressed by miR-138 (**Figure 4.13J**) by specifically binding a predicted target site in its 3' UTR. Transcription factor 12 (TCF12) interacts with  $\beta$ -Catenin to activate transcription as a result of canonical Wnt signalling. However, *TCF12* has two splice variants that generate proteins of different lengths<sup>138</sup>. The shorter protein coding transcript has an annotated 3' UTR 2,300bp long, which contains the miR-138 target site that has been validated *in vitro*. The longer protein coding transcript has an annotated 3' UTR only 11bp long, and does not contain any predicted target sites. Although it is possible that this difference in UTR may be due to incomplete annotation of the newest chicken genebuild, it may be that the second splice variant of this gene escapes regulation by miR-130 as a result of a truncated 3' UTR. The TCF family of transcription factors are essential during development<sup>7</sup> and any investigation into the dysregulation of TCF12 is likely to be complex.

The possible role of miR-138 in development as a result of targeting *ZEB2* has been discussed (**4.3.5**). miR-130 was shown to greatly inhibit the translation of a reporter construct containing the 3' UTR of *ZEB2* (**Figure 4.13K**). Although miR-130 does not have a significantly correlated fold change pattern to miR-192, as is the case for miR-138 (**Figure 4.9H**), this miRNA may also be essential for PTEN inhibition leading to growth and angiogenesis (**Figure 4.17**). If both miR-130 and miR-138 are found to target *PTEN in vivo*, it would be interesting to observe the effect on growth and angiogenesis following inhibition of both miRNAs at the same time.

Several reporter constructs were found to be translationally repressed by one or more of three selected miRNAs. The analysis of this repression in the context of heart development and in light of the expression profile of each of the repressing miRNAs has revealed many potentially important interactions, as discussed above and in **7.3.3**.

## 5 Novel miRNAs

### 5.1 Introduction

In this chapter 3 of 14 putative miRNAs, predicted by Simon Moxon (2.8.4), were validated in chicken embryonic tissue. Target sites of the novel miRNAs were predicted by searching for 6nt sequences matching the seed region of the novel miRNAs in any of the annotated 3' UTRs of the chicken genome (Galgal4), using custom R scripts written by the author (2.8.5). A GO terms analysis was performed on the predicted target genes of the novel miRNAs (2.9.2, Table 5.4).

Profiling miRNA expression by NGS has the potential of discovering novel miRNAs, which cannot easily be detected by microarray or qPCR methods. Since the advent of small RNA deep sequencing the number of miRNAs submitted to miRBase has increased very rapidly and attempts are being made to provide a measure of confidence in the miRNAs annotated in this repository<sup>170</sup>.

Although 734 precursors and 996 mature miRNAs have been identified in the chick to date, 1,872 precursors and 2,578 mature human miRNA sequences are currently annotated in miRBase release 20<sup>41,170-173</sup>. This study identified 31 miRNAs known in other animals that were detected in the chick samples assayed and that map to the chicken genome (Table 3.3), but that are not registered as existing in the chicken. Furthermore, more than half of the reads from the miRNA-sized fraction of the HH5-7 sample did not correspond to any known animal miRNAs (Figure 3.5A).

The miRNA profile of the developing chick has been assessed by NGS<sup>72,77,174-176</sup>. However, an essential step during library preparation for sequencing has been shown to favour certain miRNA sequences<sup>64,96</sup> leading to their over-representation in sequencing results and perhaps causing novel miRNAs to be overlooked. Based on the normalised reads profile of the known miRNAs identified in this study, it is possible that changes in the latest proprietary Illumina workflow may have reduced this bias, which may allow the

identification of miRNAs not previously identified. To determine whether any novel miRNAs were expressed during the stages of development investigated, bioinformatic predictions of possible miRNAs were complemented with experimental validation of the predicted structures.

## 5.2 Results

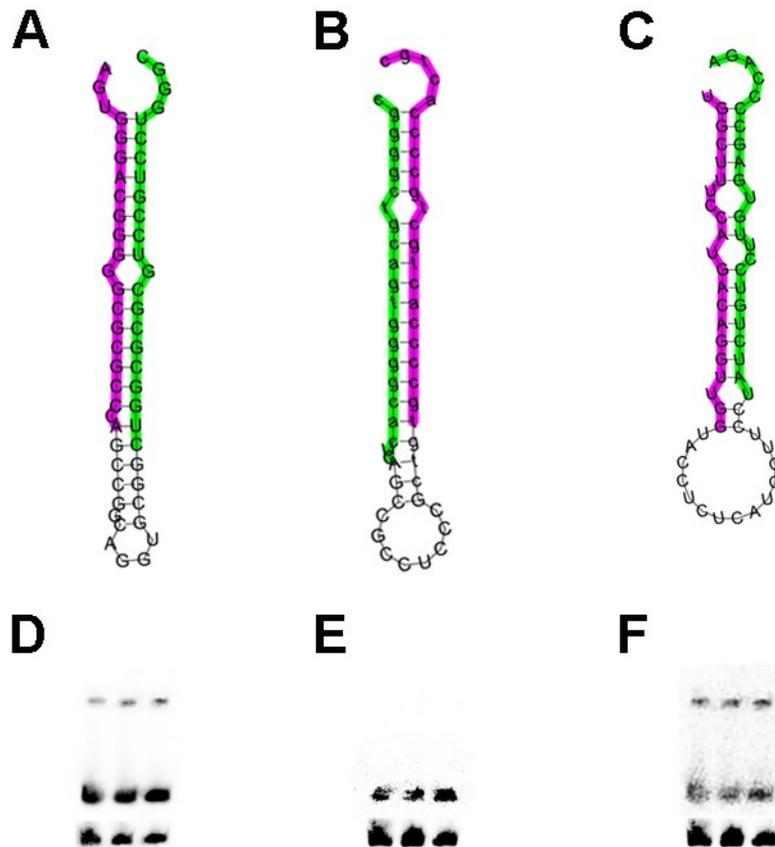
The UEA Small RNA workbench tool miRCat<sup>42</sup> was used in collaboration with Simon Moxon to identify possible novel miRNAs sequenced in any of the nine chick samples in this study.

**Table 5.1: The genomic locations and mature sequences of the fourteen potential miRNAs predicted by Simon Moxon using miRCat<sup>42</sup>.**

Name	Location	Mature sequence
Putative 1	chr1/13939419-13939524(+)	CCGAGCGCTGCCCGGGCCGTGG
Putative 2	chr1/56587108-56587198(-)	TTAGAGGACAATGGCCAATACT
Putative 3	chr1/67350434-67350518(-)	TTGAATGCAGGAGCTCCTCCT
Putative 4	chr10/14163398-14163493(-)	CTTGGCTGCTGAGGGACTCTCC
Putative 5	chr13/16456080-16456166(-)	TCATCAGAGAGGCTGGGCTCTGT
Putative 6	chr14/8505895-8505989(+)	TTTCTGGTACAGTGAAGTCCGAGC
Putative 7	chr15/4870073-4870156(-)	GGGGATGTAGCTCAGTGGTAGA
Putative 8	chr19/147190-147267(-)	TGCGCTGCGCTCCCTGCACAG
Putative 9	chr2/63171551-63171630(+)	TTGTGCCTGCATAAACTGACT
Putative 10	chr4/8488401-8488480(-)	TATCTGTCTTGTGAGCCCCAGA
Putative 11	chr4_JH375168_random/82858-82942(-)	TTCAGGGTGCTGCAGAGAGCA
Putative 12	chr5/17763791-17763866(+)	CTGGCGCGCTCCGTCTGGGC
Putative 13	chrZ/70567939-70568039(-)	ATGCAGAAGTGCACGGAAACAGCT
Putative 14	chrUn_JH376230/1786-1869(-)	CGGGGCTGCAGTGGGGCACTC

Fourteen putative novel miRNAs were identified by miRCat (**Table 5.1**) based on annotated miRNAs listed in miRBase release 19<sup>41,171-173</sup>. miRCat predicts sRNAs from NGS data in part by assessing the adjusted minimum free energy of the target read, along with that of the surrounding sequence in the chicken genome, folded with RNAfold. Pooled whole chick embryo RNA was probed for the mature miRNA sequence, and three potential miRNAs were confirmed: Putative 14 (**Figure 5.1A, D**), Putative 10 (**Figure 5.1B, E**), and Putative 12 (**Figure 5.1C, F**).

However, miRBase release 20<sup>41,170-173</sup> was found to contain an annotated miRNA matching perfectly to the sequence of Putative 10 (**Figure 5.2B**). This miRNA, miR-1462-3p, is only known in the chick and is present in intron 1 of the *Dachshund2* (*DACH2*) gene on chromosome 4 (**Figure 5.5**)<sup>138</sup>.



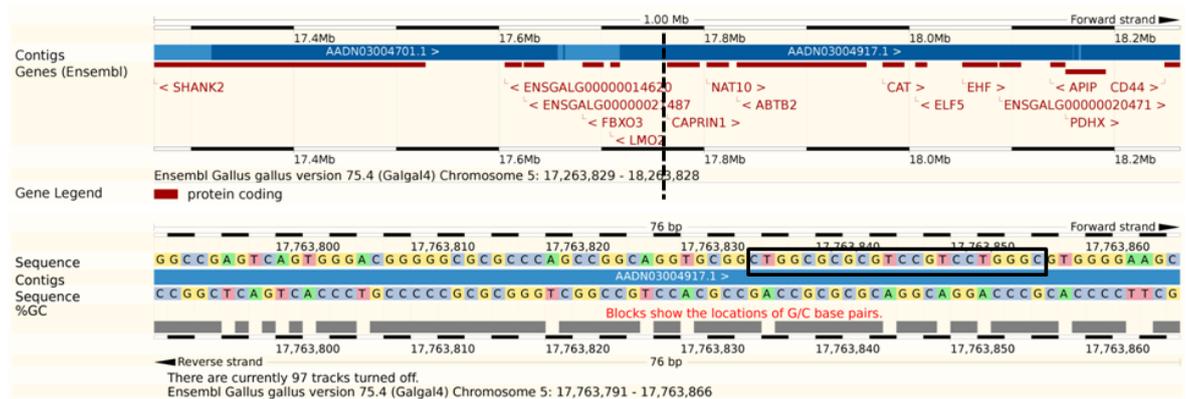
**Figure 5.1: Novel chick miRNAs identified from sequencing data.** The predicted secondary structure, generated with miRCat<sup>42</sup> by Simon Moxon, of Putative 12 (A), Putative 14 (B), and Putative 10 (C). All three miRNAs were confirmed by Northern blot. **D:** Both the pre-miRNA and mature miRNA of Putative 12 can be observed. **E:** Putative 14 was detected in chick samples. **F:** Putative 10 pre-miRNA and mature miRNA was detected (N=3). A U6 loading control band is shown at the bottom of each membrane. Membrane F was probed for Putative 10 after probing for Putative 14 and stripping. The same loading control image is shown for these blots.



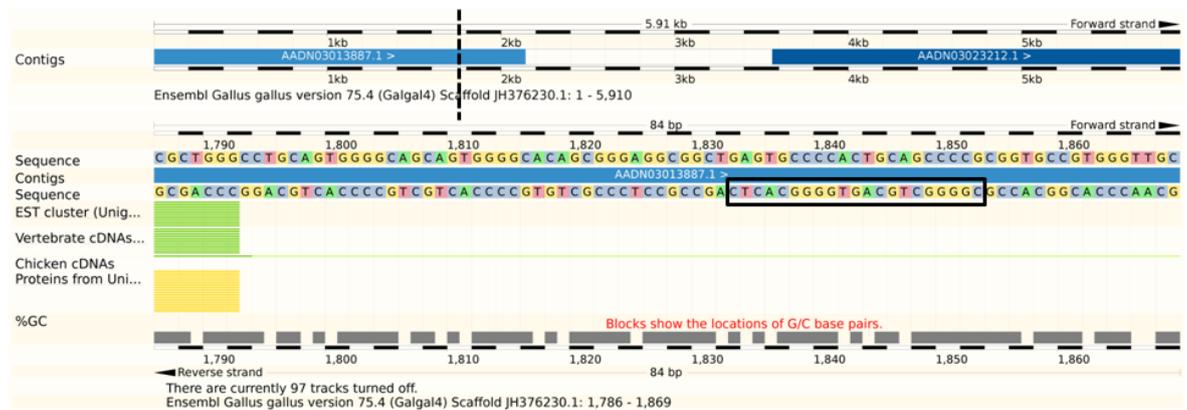
**Figure 5.2: Assessment to determine whether novel miRNAs have been annotated in miRBase.** miRBase Blast<sup>41,170-173</sup> results comparing the putative miRNAs to all known miRNAs. **A:** Putative 14 has limited sequence homology with hsa-miR-1202, but has a large e-value number and is therefore unlikely to be

a match. **B:** Putative 10 matches perfectly to gga-miR-1462-3p. This miRNA is only known in chick in miRBase release 20.

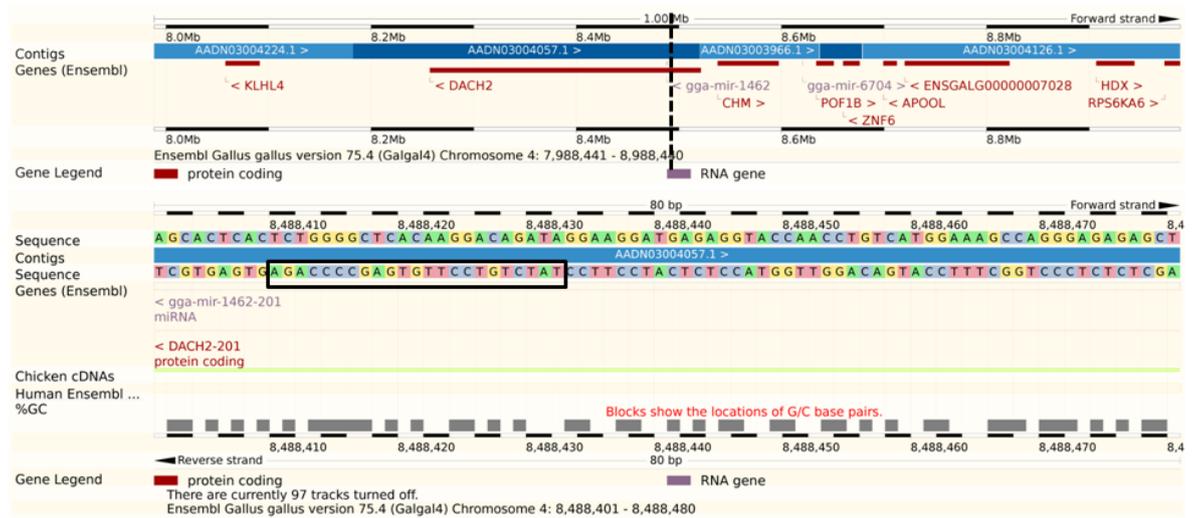
Putative 14 has very limited sequence homology with miR-1202 known in humans (**Figure 5.2**), and was identified from the JH376230 scaffold of the current chicken genebuild (**Figure 5.4**)<sup>138</sup>. Finally, sequences matching Putative 12 were not found in miRBase release 20. This miRNA is positioned 50bp upstream of the coding region of the *cell cycle associated protein 1 (CAPRIN1)* gene in the chicken genome (**Figure 5.3**).



**Figure 5.3: Genomic location of Putative 12.** Putative 12 is located 50bp upstream from the start of the coding region of cell cycle associated protein 1 (CAPRIN1) on chromosome 5<sup>138</sup>. The location is indicated in the top panel by a dotted black line and the mature miRNA sequence is highlighted in a black box in the bottom panel.



**Figure 5.4: Genomic location of Putative 14.** Putative 14 was identified from the JH376230 scaffold of the Galgal4 chicken genebuild<sup>138</sup>, indicated by the dotted line. The mature miRNA sequence is highlighted by a black box.



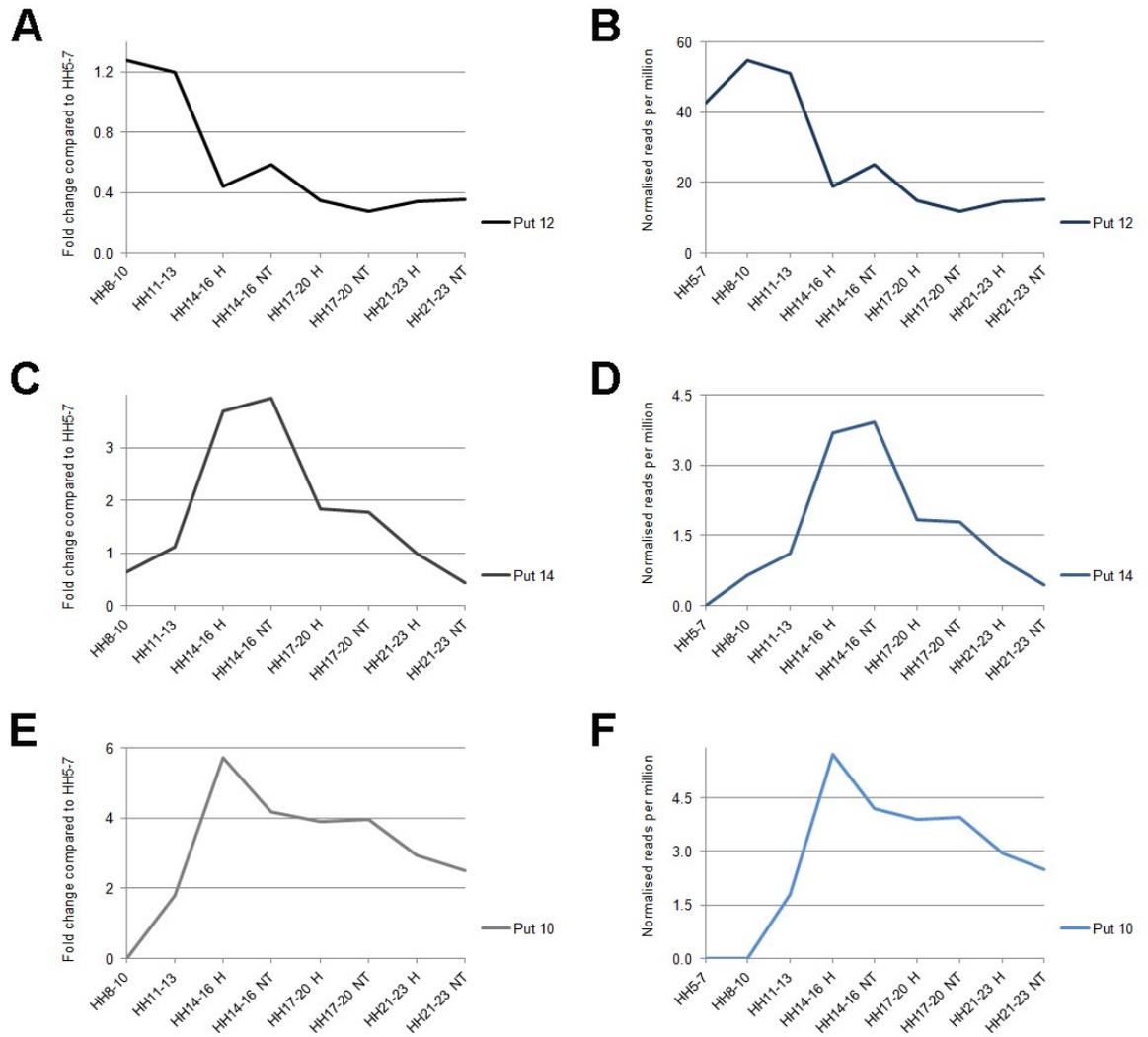
**Figure 5.5: Genomic location of Putative 10.** Putative 10 is located on chromosome 4 in intron 1 of the *DACH2* (*Dachshund2*) gene as indicated by the dotted line<sup>138</sup>. The sequence of the mature miRNA is shown outlined in black.

Putative 12, Putative 14 and Putative 10 have distinctive expression patterns in the embryonic samples sequenced (Figure 5.6). Putative 12 is downregulated as embryogenesis proceeds (Figure 5.6A, B), whereas both Putative 14 (Figure 5.6C, D) and Putative 10 (Figure 5.6E, F) have sharp increases in normalised reads in the HH14-16 sample that gradually decreases.

**Table 5.2: The seed 6-mer sequences of putative miRNAs used to check for targets in all chicken 3' UTRs.** The seed region, nt 2-7, of the predicted mature miRNA sequence was used to generate a 6-mer predicted target site which was then checked against all annotated chicken 3' UTRs for each of the putative miRNAs confirmed by northern blot.

Putative miRNA	Seed 6-mer	6-mer target site
Putative 12	TGGCGC	GCGCCA
Putative 14	GGGGCT	AGCCCC
Putative 10	ATCTGT	ACAGAT

A custom R<sup>78</sup> script was used to identify 6nt target sites, complementary to the seed sequence of each putative miRNA (Table 5.2), in all the annotated 3' UTRs of the chicken genome<sup>177</sup>.



**Figure 5.6: Expression of three predicted novel miRNAs in nine chick embryo NGS samples.** The fold change in normalised reads over that detected for each putative miRNA in the HH5-7 sample and the normalised reads per million for Putative 12 (A, B), Putative 14 (C,D), and Putative 10 (E,F).

**Table 5.3: Many possible target sites in a large number of chicken genes with human orthologues were identified for the three predicted miRNAs confirmed in chick.**

Putative miRNA	Target sites	Genes	Human orthologues
Putative 12	189	170	155
Putative 14	1,088	952	838
Putative 10	1,821	1,588	1,484

This method identified a large number of possible target sites for each putative miRNA in the 3' UTRs of many chicken genes and their human orthologues (Table 5.3).

**Table 5.4: Significant GO terms associated with the predicted target genes of three putative miRNAs confirmed in the developing chick.**

Putative 12	Putative 14	Putative 10
Cytoplasm (GO:0005737)	Cell proliferation (GO:0008283)	Generation of neurons (GO:0048699)
	Growth (GO:0040007)	Nervous system development (GO:0007399)
	Developmental growth (GO:0048589)	Neurogenesis (GO:0022008)

The GO terms significantly associated with the predicted targets of Putative 12, Putative 14, and Putative 10 ranged from genes involved with proliferation to components of the cytoplasm (**Table 5.4**). As the custom method of target identification is limited to perfect complementarity with the seed region, targets for Putative 10, known as miR-1462, GO terms analysis was performed on targets predicted by TargetScan<sup>131</sup>.

The GO terms analysis of Putative 12 returned a very general term, Cytoplasm, as the only significant association of the 179 predicted target genes of this miRNA. Repeated analysis with slightly different settings did not return any other significant GO term associations. Lowering the stringency of the analysis of the predicted target genes to also include non-significant results showed only six possible associations, all with very general headings such as “Positive regulation of cell development”, “Hip dysplasia”, and “Cell cycle checkpoints”. The limited nature of basing target predictions solely on seed sequence similarity may have led to the inclusion of genes that are not targets, which may be obscuring associations of actual targets.

### 5.3 Discussion

Three of fourteen predicted novel miRNAs were confirmed as true miRNAs by northern blot on RNA from chick embryos (**Figure 5.1**). However, the prediction was based on known miRNAs annotated in miRBase release 19, and one of the confirmed novel miRNAs, Putative 10 has since been found to match perfectly to miR-1462-3p (**Figure 5.2B**) annotated in miRBase release 20. The genomic location of Putative 10 identified by miRCat (**Table 5.1**) was also found to contain an annotated copy of miR-1462 in ENSEMBL (**Figure 5.5**). Therefore Putative 10 appears to be the known chicken miRNA miR-1462. This miRNA was detected at quite high levels in three biological replicates of RNA from whole chick embryos (**Figure 5.1F**). However, **Figure 5.6F** shows that the normalised reads of Putative 10 in the sequenced samples were very low. Putative 10 is located in intron 1 of *DACH2* on chromosome 4 (**Figure 5.5**). Developing somites express *DACH2*, which functions with *EYA2* to control the differentiation of skeletal muscle and is expressed in the chick during the developmental stages assayed in the northern blots<sup>178</sup>. Many of the genes predicted by TargetScan to be targets of miR-1462 had significant association with generation of neurons, nervous system development, and neurogenesis GO terms (**Table 5.4**). Additionally, the predicted conserved targets of miR-1462 showed a significant association with the known targets of miR-135. miR-135, via its target Smad5, inhibits osteogenesis induced by BMP2 in C2C12 cells<sup>179</sup>. Therefore, miR-1462 may be functioning in concert with miR-135 to ensure commitment to a skeletal muscle fate.

Putative 12 is located 50bp upstream of *cell cycle associated protein 1 (CAPRINI)* (**Figure 5.3**), and miRBase release 20 does not contain any sequences matching it. Its expression is high in the early stages assayed (HH5-13) but drops sharply from HH14 onwards (**Figure 5.6A, B**). This may indicate that Putative 12 has functions during the very early stages of heart development. Although 155 human genes were identified that corresponded to the 170 chicken genes containing 189 predicted target sites of Putative 12, these genes were associated with a single GO term: cytoplasm (**Table 5.4**). More than half of the predicted targets (54%) of Putative 12 were associated with this GO term. However, this term is very broad and further GO terms analyses did not yield any other significant associations of the

predicted target genes. The lack of significant associations may be due to inclusion of a large number of genes that are not targets of this miRNA which could be obscuring any significant associations of real targets of this miRNA.

Putative 14 is present on the JH376230 scaffold and although it has some homology with miR-1202 known in humans, it is doubtful that these miRNAs are identical. The expression of this miRNA increases sharply in both heart and non-heart tissue in the HH14-16 sample (**Figure 5.6C, D**). Putative 14 is predicted to target genes associated with cell proliferation, growth and developmental growth (**Figure 5.4**). Interestingly, a significant association was found between the predicted targets of this miRNA and confirmed targets of three miRNAs known to inhibit growth and proliferation. miR-198 reduces proliferation in lung cancer cells by targeting FGFR1<sup>180</sup>, miR-593 is thought to reduce proliferation of esophageal cancer by its action on Polo-kinase 1 (PLK1)<sup>181</sup>, and miR-129 inhibited the growth and migration of thyroid cancer cells<sup>182</sup>. Putative 14 may have similar effects on proliferation, especially at the HH14-16 time point which showed a rapid increase in Putative 14 expression.

Whole chick RNA was probed for all fourteen predicted novel miRNAs by Northern blot, but bands were only detected for the three putative miRNAs shown in this chapter. The normalised read counts for all eleven undetected miRNAs were quite low, ranging from 3 to 188 reads per million and it is possible that these miRNAs do exist but were below the limits of detection of the Northern blots performed. Putative 4 and Putative 10 both have fewer than 5 reads per million at their peak and both were detected by Northern blot. However, whole embryo RNA was used to perform the Northern blots and therefore if a putative miRNA is expressed at low levels in a specific tissue, such as the heart, its signal may be undetectable without isolating RNA from just the heart for Northern blots.

## 6 The role of miR-1 during early chick heart development

### 6.1 Introduction

In this chapter the biological effects of miR-1 inhibition, as achieved by antagomiR cardiac microinjection at HH14, were examined. miR-1 was found to translationally inhibit *FZD7* and *RARβ* *in vivo*, and its dysregulation at HH14 resulted in developmental delay of embryos.

The cardiac and skeletal muscle-specific<sup>4,5,183,184</sup> miR-1 is indispensable during heart development<sup>5,6,107,108,183,185</sup>. miR-1 inhibits the Wnt signalling receptor *FZD7* and the Retinoic Acid (RA) signalling receptor *RARβ*<sup>8</sup>. Both Wnt and RA signalling are essential for cardiogenesis at several stages of development<sup>7,34</sup>, and the effects of miR-1 dysregulation with respect to these signalling pathways were examined.

#### 6.1.1 miR-1

miR-1 has been found to be specifically expressed in skeletal and cardiac muscle, in a manner that is evolutionarily conserved<sup>4,5,73,183,184</sup>. The muscle-specificity of miR-1 expression has been shown by the repression of murine miR-1 sensor constructs in mouse hearts, by the *in vivo* expression of miR-1 LacZ promoter fusion constructs in differentiating cardiac and skeletal muscle precursors, as well as by the detection of miR-1 expression in Zebrafish and *D. melanogaster* muscle by LNA-probe *in situ* hybridization<sup>108,184</sup>. This microRNA is transcriptionally regulated by MyoD<sup>5</sup>, Myf5 and myogenin, myogenic regulatory factors (MRFs)<sup>186</sup>, by members of the myocyte enhancer factor 2 (MEF2) transcription factor family, as well as by serum response factor (SRF)<sup>5,183,184</sup>. Texel sheep, a very muscular breed popular for its meat, have a G-to-A mutation in the 3' UTR of their myostatin mRNA which allows the closely related miR-1 and miR-206 mature microRNAs to bind and translationally repress this gene, causing the heavy muscularity the breed is known for<sup>6</sup>. Additionally, miR-1 has been associated with

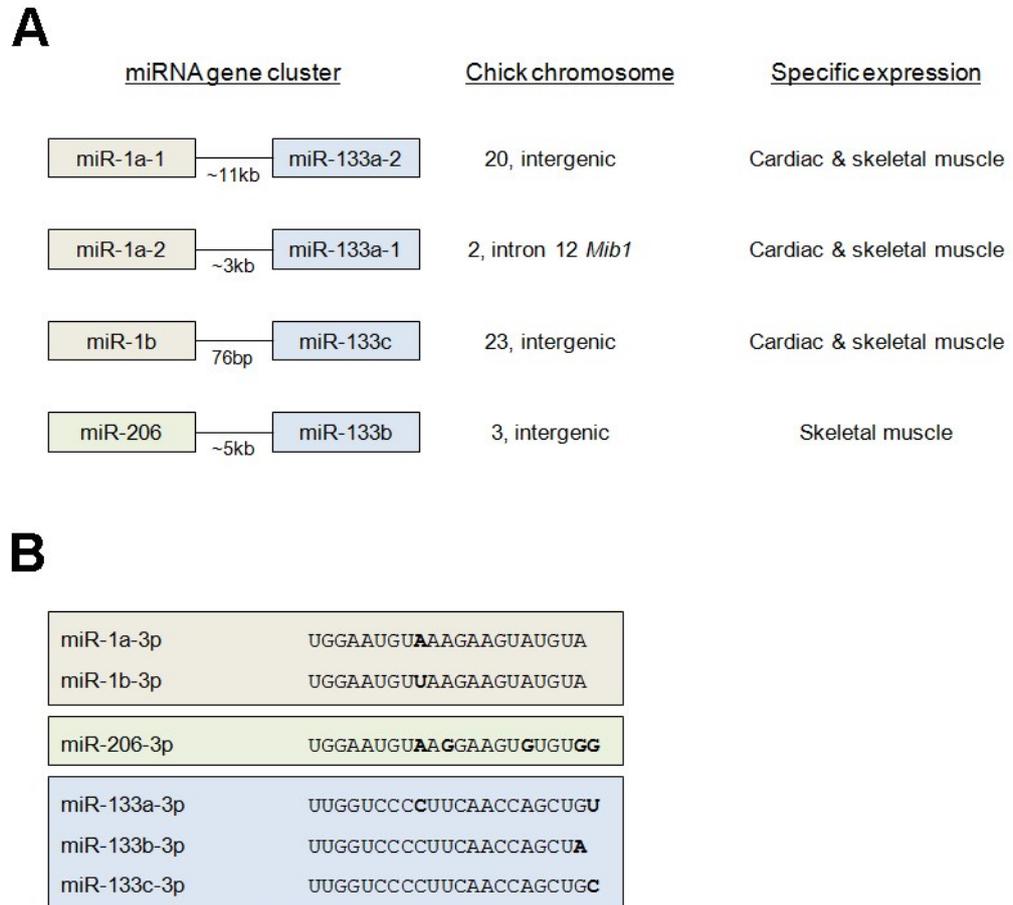
myogenesis and with the maintenance of differentiation of muscle cells<sup>107,108,185</sup>, as well as with the regulation of key components of the cardiac conduction system<sup>6</sup>.

miR-1 expression in human heart and skeletal muscle was confirmed by Northern blot assays<sup>184</sup>. miR-1 has been implicated in cardiac dysfunction with miR-1 expression levels decreased in diseased human hearts and in three cardiac hypertrophy models<sup>6,184</sup>, and significantly elevated in samples from humans with coronary artery disease (CAD)<sup>185</sup>. This microRNA is evolutionarily conserved with a single *miR-1* gene being present in both the worm and fly genomes and with two *miR-1* loci, *miR-1-1* and *miR-1-2*, present in the zebrafish, mouse and human genomes<sup>184</sup>. miR-1-1 and miR-1-2 have identical mature sequences<sup>114</sup>. The most recent release of the miRBase microRNA database<sup>41,170,171,173</sup>, release 20, contains three annotated copies of the *miR-1* gene in the chicken genome.

The *miR-1a-1 / miR-133a-2* cluster is located in an intergenic region on chromosome 20, whilst *miR-1a-2 / miR-133a-1* is found on chromosome 2 in an antisense direction within intron 12 of the *Mib1* gene (**Figure 6.1A**). A third copy of the *miR-1* gene originates from the *miR-1b / miR-133c* cluster within intron 1 of the putative *ISG12-1* gene on chromosome 23 (**Figure 6.1A**). The mature sequences of miR-1a-1 and miR-1a-2 are identical, whereas miR-1b differs from the miR-1a mature sequence by one nucleotide (**Figure 6.1B**). *miR-1a* expression is present in the chick embryo heart tube by Hamburger-Hamilton stage (HH) 11 and in the myotome of the somite from HH14<sup>115</sup>. The Geisha chicken embryo gene expression database<sup>17,147</sup> indicates that miR-1b is expressed in the heart from HH13 until at least HH27.

In *D. melanogaster* SRF is known to regulate the expression of miR-1 in cardiac muscle and MEF2 also plays a role in the regulation of muscle-specific miR-1 expression<sup>183</sup>. In mice the *miR-1-1 / miR-133a-2* cluster is found on chromosome two in a non-coding region<sup>4</sup>, whilst the *miR-1-2 / miR-133a-1* cluster is positioned in the antisense direction in an intron of the ubiquitously expressed *Mib1* gene, which is involved in the Notch signalling pathway<sup>5</sup>. Both of these clusters have been shown to be expressed specifically in skeletal and cardiac muscle<sup>6,114,184</sup>, with expression starting in the ventricular myocardium and the interventricular septum from E8.5 and continuing until adulthood<sup>114</sup>. SRF and

MEF2 cooperate with the MRFs myogenin and MyoD to induce the expression of miR1-1 and miR-1-2 in skeletal muscle<sup>4,6</sup>. MEF2 binds to the intronic enhancer within the cluster to turn on expression of *miR-1* in ventricular myocytes, whereas SRF has been shown to direct the expression of miR-1 in ventricular and atrial myocytes<sup>6</sup>.



**Figure 6.1: Chicken muscle-specific microRNAs. A:** The *miR-1a-1 / miR-133a-2* cluster is located in an intergenic region of chromosome 20, *miR-1a-2 / miR-133a-1* resides in intron 12 of the *Mib1* gene on chromosome 2, and *miR-1b / miR-133c* is found in an intergenic region of chromosome 23. The expression of these three clusters is restricted to cardiac and skeletal muscle. The skeletal muscle-specific *miR-206 / miR-133b* cluster is sited in an intergenic region of chromosome 3. **B:** Alignments of the mature microRNA sequences that result from the clusters shown in A, with letters in bold indicating sequence differences. The skeletal muscle-specific miR-206 is very similar to miR-1.

A 4.6 kilobase (kb) fragment from the genomic region around *miR-1-1* and *miR-1-2*, conserved between human and mouse, directed lacZ expression in transgenic mouse hearts<sup>108</sup>. This transgene was expressed in the heart from E8.5 and in the myotome from E11.5, confirming the cardiac and skeletal muscle-specificity of miR-1 expression<sup>108</sup>. After cardiac looping and as cardiomyocyte differentiation progressed, the expression of the

transgene became more robust with the strongest expression seen in the less proliferative inner curvature of the looping heart tube and atria<sup>108</sup>.

The early expression of miR-1 in embryonic stem cells (ESCs) forced their differentiation toward a cardiac fate<sup>6</sup>. The overexpression of *miR-1* and the closely related *miR-206* in cultured C2C12 cells advanced myogenic differentiation whereas the knockdown of these microRNAs prevented myogenic differentiation<sup>4</sup>. The expression of non-muscle genes is repressed in a microarray of HeLa cells transfected with human miR-1, which suggests that miR-1 may be responsible for the maintenance of muscle cell identity<sup>184</sup>. Furthermore, skeletal myoblasts were forced to differentiate into skeletal muscle cells by the overexpression of *miR-1 in vitro*<sup>5</sup>, and hypertrophic growth of cardiomyocytes was inhibited by miR-1 *in vitro*<sup>6</sup>. Histone deacetylase 4 (HDAC4) is a signal-dependent enzyme that causes the repression of MEF2 and leads to the repression of muscle differentiation<sup>93</sup>. In C2C12 cells miR-1 was shown to repress the translation of HDAC4, leaving MEF2 free to induce myoblast differentiation in conjunction with the MRF MyoD<sup>107</sup>.

The injection of miR-1 into *Xenopus laevis* embryos resulted in increased differentiation and decreased proliferation of muscle cells, as well as anomalous heart development<sup>6,107</sup> similar to that seen when *miR-1* was overexpressed in *D. melanogaster* embryos<sup>183</sup>. When miR-1, controlled by a  $\beta$ MHC promoter which gives high levels of miR-1 expression by E9.0, was overexpressed in developing mouse hearts, lethality resulted during mid-embryogenesis at E13.5, possibly due to heart failure as a result of the observed deficiency in cardiomyocytes<sup>108</sup>. It has been suggested that this deficiency of cardiomyocytes is due to miR-1 targeting the mRNA of Hand2<sup>108</sup>, a bHLH transcription factor that regulates the expansion of the ventricles of the embryonic heart<sup>187</sup>. Postnatal miR-1 overexpression confirmed that miR-1 translationally represses Hand2, as significantly reduced Hand2 protein was detected in transgenic mouse hearts when compared with non-transgenic littermate hearts whilst Hand2 mRNA levels remained constant<sup>108</sup>.

Groups of undifferentiated cells along with a general loss of differentiation in some muscle was observed when *miR-1* was deleted from the *D. melanogaster* genome<sup>184</sup>. Deletion of

*miR-1* resulted in a third of mutants dying at embryonic stages, a third dying at hatching and the final third dying at larval stages after displaying reduced mobility<sup>183</sup>. Defects in cardiac and somatic muscle patterning were seen in the most severely affected embryos that died during embryogenesis, and surviving embryos had a loss of cardiac cells in the dorsal vessel<sup>183</sup>.

Deletion of *miR-1-2* did not result in the upregulation of *miR-1-1* expression and it was reported that about half of *miR-1-2*<sup>-/-</sup> mutants die shortly after birth due to ventricular septal defects (VSD)<sup>5</sup>. No difference in Hand2 mRNA levels was detected at E15.5 between surviving *miR-1-2*<sup>-/-</sup> mutants and their wildtype littermates, but Hand2 protein levels were four times higher in the *miR-1-2*<sup>-/-</sup> mutant animals. As a narrow range of Hand2 concentrations are required for normal cardiomyocyte development<sup>187</sup>, the dysregulation observed in *miR-1-2*<sup>-/-</sup> mutants strongly imply that this microRNA plays an essential role during heart development by fine-tuning the protein levels of Hand2. In addition, the walls of the hearts of the surviving *miR-1-2*<sup>-/-</sup> mutants appear thickened, with many myocytes apparently undergoing nuclear division<sup>5</sup>. Although mouse cardiomyocytes do undergo nuclear and cellular division in the first two weeks of life before permanently exiting the cell cycle, a significant increase in mitotic myocytes was found in *miR-1-2*<sup>-/-</sup> mutants compared to their wildtype littermates<sup>5</sup>. In fact, mitotic myocytes were still found in some 3-month-old mutant animals<sup>5</sup>. Together with the overexpression studies, these results lead to the suggestion that miR-1 has an essential role in inhibiting cell proliferation during and following embryonic heart development<sup>5,6</sup>.

Of the surviving *miR-1-2*<sup>-/-</sup> mutants, about 15% died of heart and ventricular dysfunction by 3 months of age and despite the rest of the animals appearing normal many suffered sudden death, leading the researchers to suspect cardiac electrophysiologic defects to be the cause<sup>5</sup>. Adult mutant mice were found to have significantly lower heart rates than their wildtype littermates and had two-fold higher *Irx5* mRNA and 5-fold higher *Irx5* protein levels than wildtype hearts. *Irx5* is a transcription factor that represses a potassium channel *Kcnd2* and thus regulates cardiac repolarisation. In accordance with the observed increase in *Irx5* mRNA and protein levels, *miR-1-2*<sup>-/-</sup> hearts had lower levels of *Kcnd2* mRNA.

These results suggest that miR-1 is involved in the regulation of the cardiac electrical system by targeting *Irx5*<sup>5</sup>.

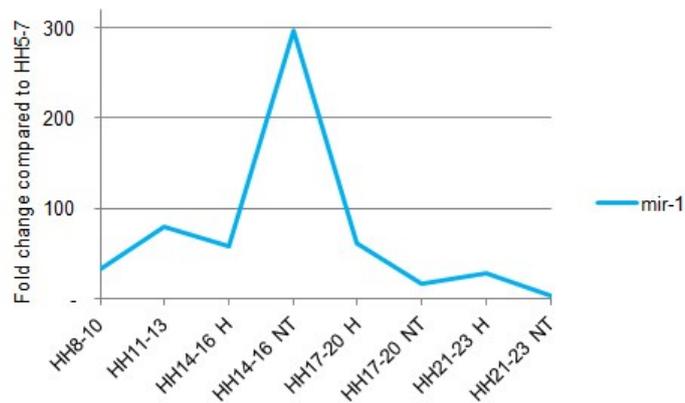
Rats subjected to experimental myocardial infarction (MI), the formation of necrosis caused by obstruction of blood circulation, were found to have an increase of about 2.6 fold of *miR-1* expression within the ischemic zone but no increase in the non-ischemic zone<sup>185</sup>. Ischemic arrhythmias were increased when miR-1 RNA was injected into the infarcted myocardium but were suppressed significantly by the injection of antagomiR-1 into this region. Further investigation revealed that miR-1 translationally downregulates the potassium channel protein Kir2.1 as well as the cardiac gap junction channel protein connexin 43 (Cx43), leading to increased arrhythmias in ischemic hearts<sup>185</sup>. When antagomiR-1 and siRNA against Cx43 and Kir2.1 mRNA were co-injected, significant arrhythmias were induced in ischemic hearts, confirming that these arrhythmias were due to the downregulation of Cx43 and Kir2.1 and thus indicating that miR-1 possibly has a pathological role in ischemic hearts<sup>185</sup>.

To date miR-1 is intimately involved in both the embryonic development and postnatal function of cardiac cells with its effects ranging from differentiation of cardiomyocytes to control of components of the conduction system. Microarray assays of adult hearts of *miR-1-2* knockout mice showed significant upregulation of 45 and downregulation of 25 protein-coding genes<sup>5</sup>, suggesting that many important targets of this microRNA still remain unstudied. Previous work done in the Münsterberg laboratory illustrated the importance of miR-1 function during somite myogenesis<sup>127</sup>, and miR-1 was shown to target the Wnt receptor Frizzled 7 (FZD7) as well as Retinoic Acid Receptor  $\beta$  (RAR $\beta$ ) *in vitro*<sup>8</sup>.

Like miR-1, miR-133 is known to be expressed specifically in skeletal and cardiac muscle under the transcriptional regulation of MyoD, MEF2 and SRF<sup>5,33,183,184</sup>, but unlike miR-1 this microRNA inhibits the differentiation and stimulates the proliferation of myoblasts<sup>4</sup>. Along with miR-1, miR-133 has been found to be involved in the regulation of components of the cardiac conduction system and to increase arrhythmias<sup>6</sup>, possibly by the repression of a potassium channel protein, KCNQ1<sup>188</sup>.

*In vitro* miR-133 repressed the differentiation and advanced the proliferation of C2C12 cells by the repression of its targets SRF and polypyrimidine tract-binding protein (nPTB)<sup>5,6</sup>, and inhibited ESC differentiation towards cardiomyocytes<sup>6</sup>. Additionally, miR-133 has been found to repress the translation of the cardiac potassium channel protein ether-a-go-go-related gene (ERG), resulting in delayed myocyte repolarization, when it was introduced into cardiomyocytes *in vitro*. The normal hypertrophic response of both neonatal and adult cardiomyocytes to agonist stimulation was inhibited by the overexpression of miR-133 by an adenoviral vector<sup>6</sup>. Despite the proximity of the microRNA genes and their overlapping expression patterns, evidence has been found that indicates miR-1, miR-133 and miR-206 can be expressed as independent transcriptional units<sup>6</sup>.

In this study miR-1 was detected in all nine samples assayed (**Figure 6.2**), and was found to have a large spike in expression in the HH14-16 NT sample. This large difference between the HH14-16 H and HH14-16 NT samples may have been due to contamination of the HH14-16 NT sample with material from a developing somite, known to express miR-1 HH14<sup>115</sup>.



**Figure 6.2: The fold change pattern of miR-1 over its levels in the HH5-7 sample.** miR-1 was found to have a very large fold change increase in the HH14-16 NT sample.

### 6.1.2 Frizzled 7

The Frizzled (FZD) family of 7-transmembrane receptors can bind members of the cysteine-rich, glycosylated Wnt-protein family on the extracellular surface of cells<sup>26</sup>. The Wnt signalling pathway is highly conserved and is involved in a large number of biological processes during embryogenesis and in adults<sup>25,26</sup>. FZD receptors have been shown to be involved in the regulation of the cytoskeleton, cell proliferation, apoptosis induction and, in *D. melanogaster*, receiving polarity signals and transmitting these signals to adjacent cells<sup>26</sup>.

After myocardial infarction in rats, myofibroblasts expressed FZD2 whilst migrating into the infarct area as well as during their proliferation following migration<sup>26</sup>. Deletion of the *Dvll* gene in mice resulted in infarct rupture much more often than in wildtype mice;  $\beta$ -catenin could not be detected in the hearts of mice that died following infarct rupture whereas wildtype hearts had high levels of  $\beta$ -catenin<sup>26</sup> showing that Wnt-Fzd signalling is essential for wound healing in the heart. Even though the loss of  $\beta$ -catenin abated myocyte-myocyte interactions by weakening adherens junctions, healthy hearts did not appear affected and only once myocardial infarction occurred was the structural integrity of the ventricular wall threatened. This suggests that the Wnt-Frizzled signalling cascade is essential for healing following a myocardial infarction. Frizzled-associated Wnt signalling has been implicated as being involved in angiogenesis, a process which is upregulated in ischemic heart disease and results in blood flow being reinstated to the wounded tissue<sup>26</sup>. Following induced infarction in rat hearts, the cell adhesion complexes of newly formed endothelial cells were disrupted, and concurrently  $\beta$ -catenin was found to be translocated to the cytoplasm from the plasma membrane<sup>26</sup>, indicating that Wnt-Fzd signalling is involved in the process of vascularisation of the diseased heart.

Foetal, neonatal and overloaded rat hearts have been shown to have higher expression of FZD receptors compared to healthy adult hearts<sup>189</sup>. Secreted Frizzled-related proteins (sFRPs) antagonise the Wnt / FZD signalling pathway by binding to and therefore inhibiting the normal action of either Wnt or FZD<sup>189</sup>. Human adult ventricular cardiomyocytes have been found to express sFRPs in the ventricular myocardium of both

healthy and diseased hearts<sup>189</sup>. Wnt3A and Wnt8 activity in the cardiogenic mesoderm of *X. laevis* is inhibited by the action of the Wnt antagonists Dickkopf-related protein 1 (Dkk-1) and Crescent<sup>27</sup>. These findings lead to speculation that Wnt antagonism, in part, initiates Nkx2.5 expression, later maintained by the growth factor Bone Morphogenetic Protein (BMP), to initiate cardiogenesis<sup>27</sup>. Similarly, in HH8-9 chick embryo explants, Wnt-3a and Wnt-1 signals from the neural tube and BMP antagonists noggin and chordin secreted by the notochord were found to inhibit cardiogenesis in the anterior paraxial mesoderm<sup>28</sup>. When the anterior paraxial mesoderm of chick embryos was exposed to both BMP-4 and a Wnt antagonist, the heart was found to become enlarged as a result of the formation of a larger pool of cardiac myocyte precursor cells<sup>28</sup>. The migration of mesodermal cells into the heart was also stimulated by Wnt antagonists and BMP molecules<sup>28</sup>. Cardiogenesis was induced when GSK3 $\beta$  was ectopically expressed in *X. laevis* ventral mesoderm, indicating that heart formation is a result of the disruption of Wnt signalling via the canonical pathway<sup>27</sup>.

A member of the FZD family, FZD7, is known to function as part of the canonical Wnt signalling pathway<sup>25</sup>. Whole-mount ISH (WISH) on developing mouse embryos showed expression of FZD7 to be ubiquitous and comparable to receptors involved in anterior-posterior patterning in the gastrulating embryo<sup>25</sup>, and in *X. laevis* FZD7 has been implicated in neural crest development by its interaction with Wnt signals<sup>190</sup>. The Geisha chicken embryo gene expression database<sup>17</sup> shows FZD7 expression in chick from HH3, in the early mesoderm and primitive streak, up to at least HH10, in the paraxial mesoderm and neural tube. In the frog FZD7 is expressed in migrating cardiac progenitor cells (CPCs) and in the pericardium and myocardium of the developing heart<sup>189</sup>.

Previously, miR-1 expression was reported as skeletal and cardiac muscle-specific<sup>4,5,183,184</sup> and essential for both the development and function of the heart<sup>5,6,107,108,183,185</sup>. *In vitro* assays on C2C12 cells have confirmed that miR-1 can suppress sensor constructs containing the 3' UTR of FZD7, at the level of translation<sup>8</sup>. The inhibition of canonical Wnt signalling by Wnt antagonists, along with BMP secretion, is essential for cardiogenesis during development<sup>26-28</sup>. The translational suppression of FZD7, a transmembrane receptor that binds Wnt<sup>25</sup>, by miR-1 could possibly ensure that FZD7 receptors are absent and therefore cannot bind stray Wnt signals, and so ensure the commitment of the miR-1-expressing cell to the cardiac lineage. In adult hearts the Wnt-FZD cascade is essential for healing following myocardial infarction in mice<sup>26</sup>, and miR-1 expression is reported to be upregulated in the ischemic zone following myocardial infarction<sup>185</sup>. It is possible that the upregulation of miR-1 following an infarction results in the translational suppression of its target FZD7, which leads to a disruption in myocyte-myocyte interactions and a further weakening of the heart.

### 6.1.3 Retinoic acid receptor $\beta$

Retinoic acid (RA), a small lipophilic molecule, directly activates transcription when bound to one of three retinoic acid receptors (RAR $\alpha$ , RAR $\beta$  or RAR $\gamma$ ). In this signalling pathway RA is derived from retinol via retinaldehyde by the action of retinol dehydrogenases and retinaldehyde dehydrogenases. The RARs form heterodimers with retinoid X receptors (RXRs) and act in the nucleus regardless of whether RA is bound<sup>191</sup>. If RA is not bound to the RAR/RXR heterodimer, this complex binds co-repressor complexes which repress transcription by keeping the chromatin structure of the target DNA in a stabilised state<sup>34</sup>. However, when the agonist binds to the RAR/RXR heterodimer, a conformational change leads to the release of the co-repressor complexes allowing co-activator proteins to become attached, which results in transcriptional activation of the gene<sup>34,191</sup>. RA signalling has pervasive roles during development, ranging from early dorsoventral axis patterning to the morphogenesis and differentiation of the heart<sup>34</sup>, confirmed by the number of abnormalities that resulted when rat and avian embryos were deprived of dietary vitamin A<sup>120</sup>.

Embryonic stem cells exposed to RA are induced to differentiate into their previously specified lineages<sup>34,192</sup>. *In vivo* experiments on chick embryos have shown that ectopic RA causes a large number of heart anomalies, especially when embryos are treated during early development (HH3 – HH6)<sup>193</sup>. Local application of RA to the cardiogenic mesoderm inhibited the formation of a fused cardiac crescent and lead to cardia bifida, similar to the effects seen when the formation of the cardiac crescent is prevented by the removal of a piece of the precardiac mesoderm. A stronger phenotypic effect was seen in embryos treated with RA at HH5 than later stages, further indicating that excess RA causes cardia bifida by inhibiting the migration of cardiogenic mesoderm<sup>193</sup>.

Mouse mutants lacking either the *Rxra* or *Raldh2* genes died about midway though development and had impaired myocardial differentiation and growth<sup>34</sup>. The knockout of the genes encoding RAR $\alpha$  and RAR $\beta$  at the same time resulted in many abnormalities, including outflow tract and large vessel malformations, and mice died at birth. The heart, outflow tract and large vessels were also affected in mice with both the *Rarg* and *Rxrb* genes deleted; embryonic lethality was observed about halfway through gestation<sup>34</sup>.

The inhibition of RA signalling *in vivo* lead to an increase in the postnatal size of a subpopulation of mouse stem and progenitor cells that contains CPCs, a decrease in the number of differentiation markers expressed, as well as an increase in the expression of cardiac progenitor markers<sup>192</sup>. Cardiac Fgf8 is necessary for *Isl1* expression. As *Isl1* expression defines the second heart field, the inhibition of Fgf8 by RA at the posterior end of this region observed in mutant mice limited the size of the cardiogenic region<sup>120</sup>. These results suggest that RA negatively controls the size of the cardiac progenitor pool during early development in mice. Similarly, the bilateral heart fields expressing *Nkx2.5* and *cardiac myosin light chain 2* became enlarged as a result of the inhibition of RA signalling in zebrafish embryos<sup>34</sup>. Further work using zebrafish embryos indicated that lateral mesoderm will become specified to a cardiac fate in the absence of RA, providing extra support for the suggestion that RA signalling has an inhibitory effect on cardiac specification during early development<sup>34</sup>.

Epicardial progenitors migrate from the proepicardial organ (**Figure 1.2B**), over the outside of the myocardium and eventually cover the surface of the growing heart<sup>18</sup> (**Figure 1.2A**). The epicardium is responsible for inducing the underlying myocardium (**Figure 1.3**) to grow<sup>18</sup>, though RA signalling is not required during the initial formation and migration of the epicardium<sup>192</sup>. Experiments have shown that the growth-promoting effects of the epicardium depend on RA<sup>192</sup>, which has been proposed to act on liver cells and result in myocardial growth via *Epo* and *Igf2* expression<sup>120</sup>. RA stimulates the secretion of hepatic EPO, which in turn fuels IGF2 secretion when it becomes bound to epicardial receptors. IGF2 signalling in myocardium induces myocardial growth as a result of IGF2 secretion by the epicardium.

Cardiomyocyte differentiation was shown to be suppressed in RA deficient mice<sup>34,192</sup>; this phenotype could be rescued by exogenous addition of FGF<sup>192</sup>. FGF signalling, required for growth in the heart by recruiting cells to the arterial pole and known to induce differentiation of cardiomyocytes, is dependent on RA signalling in murine epicardial and proliferating myocardial cells<sup>34,192</sup>. In addition, the migration of neural crest cells is disturbed when RA is absent<sup>191</sup>, resulting in a loss of outflow tract septation<sup>34</sup>. It has been suggested that a concentration gradient of RA is responsible for the development of the segmental units of the hindbrain, possibly due to FGF repressing a RA-metabolizing enzyme, Cytochrome P450 26 A1 (CYP26A1), along its own concentration gradient<sup>34</sup>.

In zebrafish embryos miR-138 is expressed in the ventricular chamber during early development and in the outflow tract during later development<sup>111</sup>. Chemical inhibition of miR-138 caused looping defects, cardiac dysfunction and reduced cardiomyocyte differentiation in the vast majority of treated embryos, similar to the effects seen when RA signalling was induced in the heart. The expression of miR-138 is reported to be essential for normal cardiogenesis during the early stages of cardiac looping<sup>111</sup>. During RA signalling retinoic acid dehydrogenase (*Raldh2*) produces activated RA, and in the zebrafish embryo *Raldh2* is expressed only in the atrioventricular canal of the embryonic heart. The gene encoding *Raldh2*, *aldh1a2*, was found to have a conserved miR-138 binding site, and miR-138 inhibition resulted in upregulated *Raldh2* expression in ventricular cardiomyocytes as well as cardiac abnormalities<sup>111</sup>. In addition, miR-138 was

found to repress the ventricle-specific protein versican. Versican is known to induce valve formation and its gene, *cspg2*, is activated by RA signalling. These results suggest that miR-138 plays a role in cardiac development by inhibiting the RA signalling pathway in zebrafish ventricular cardiomyocytes<sup>111</sup>.

Retinoic acid receptor  $\beta$  is one of several predicted targets of miR-1, and *in vitro* experiments have shown that miR-1 can repress RAR $\beta$  at the transcript level<sup>8</sup>. The size of the CPC pool is limited by RA<sup>34,120,192</sup>, and RA signalling is required for FGF-induced myocardial differentiation *in vivo* as well as for myocardial growth stimulated by epicardium<sup>192</sup>.

Given the importance of the known roles of miR-1 during heart development, the effects of miR-1 inhibition on embryonic development were studied along with the consequences of miR-1 inhibition on the translation of *FZD7* and *RAR $\beta$* , known targets of this miRNA *in vitro*. Although miR-1 is reasonably well studied, it is very unlikely that most of its roles have been discovered and its regulation of *FZD7* or *RAR $\beta$*  has not been elucidated elsewhere.

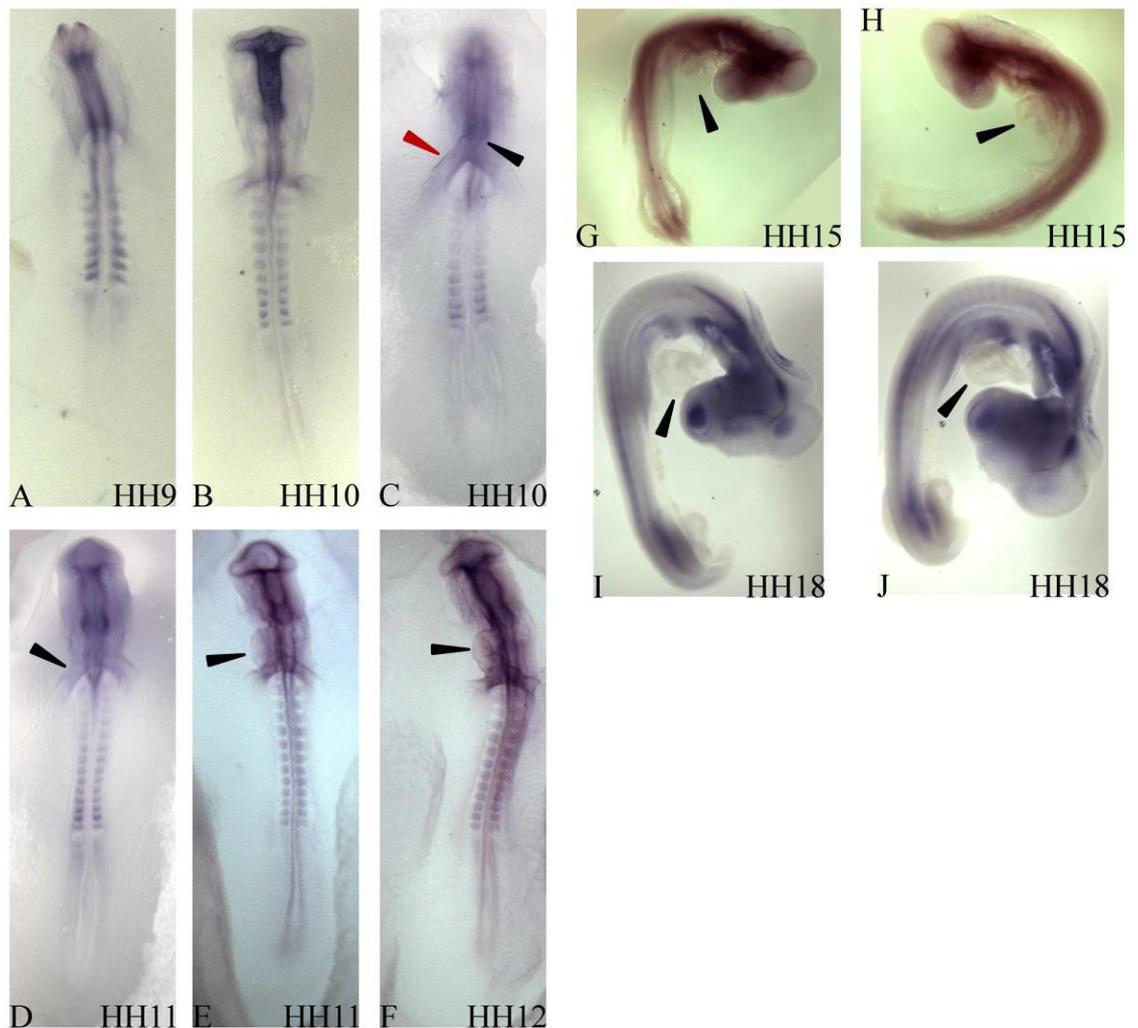
## 6.2 Results

### 6.2.1 FZD7 and RAR $\beta$ mRNA is expressed in the developing chick heart

In order to establish an expression profile of *FZD7* and *RAR $\beta$*  in the chicken embryo, WISH assays were performed on embryos ranging from HH4 to HH18.

*FZD7* is expressed in primitive streak and the area pellucida of HH4 and HH4+ embryos (**Figure Apx D.1A, B**), as well as in the cardiogenic and paraxial mesoderm at HH8 (**Figure Apx A.1D**). *FZD7* is still expressed in the cardiogenic mesoderm at HH9 (**Figure 6.3A**), and in the heart tube at HH10 (**Figure 6.3B, C**), HH11 (**Figure 6.3D, E**), and HH12 (**Figure 6.3F**). By HH15 (**Figure 6.3G, H**) and HH18 (**Figure 6.3I, J**) however, *FZD7* has been transcriptionally downregulated in the heart.

The expression profile from HH4 to HH10 established in **Figure Apx A.1** and **Figure 6.3** confirms the ISH analysis of *FZD7* expression published on the Geisha web resource. In addition, the expression of *FZD7* was established up to HH18 and *FZD7* was found to be expressed in cardiogenic and heart tissue from HH4 up to HH14.

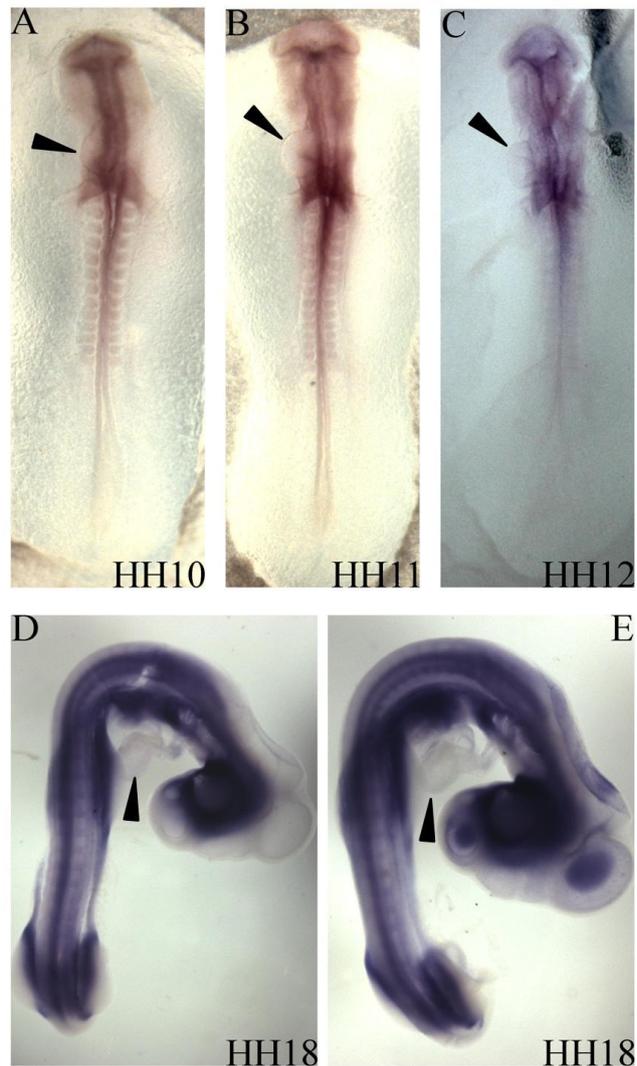


**Figure 6.3: FZD7 expression in the chick during later stages of development.** FZD7 mRNA WISH on chick embryos. At HH9 FZD7 is expressed in the cardiogenic and paraxial mesoderm (**A**). **B, C:** Dorsal (**B**) and ventral (**C**) views of a HH10 embryo shows FZD7 expression in the neural tube, neural plate, paraxial mesoderm and in the heart tube (black arrow). The omphalomesenteric vein (red arrow, **C**) also expresses FZD7 mRNA. **D, E, F:** FZD7 is expressed in the cardiogenic mesoderm, heart tube (black arrows, **E, F**), omphalomesenteric vein (**D**), paraxial mesoderm and neural tube at HH11 and HH12. **G, H, I, J:** FZD7 expression is not detected in the heart at HH15 (black arrow, **G, H**) and HH18 (black arrow, **I, J**).

A very slight *RAR $\beta$*  mRNA signal seen in HH10 (**Figure 6.4A**), HH11 (**Figure 6.4B**), and HH12 (**Figure 6.4C**) heart tubes. By HH18 *RAR $\beta$*  expression appears completely downregulated (**Figure 6.4D, E**).

The expression of *RAR $\beta$*  at HH10 seen in **Figure 6.4A, B** and **C** mirrors the profile of *RAR $\beta$*  published in the GEISHA web resource. Information on *RAR $\beta$*  expression between

HH11 and HH18 is not available from the GEISHA web resource. RAR $\beta$  was here found to be low at HH11 and absent by HH18.



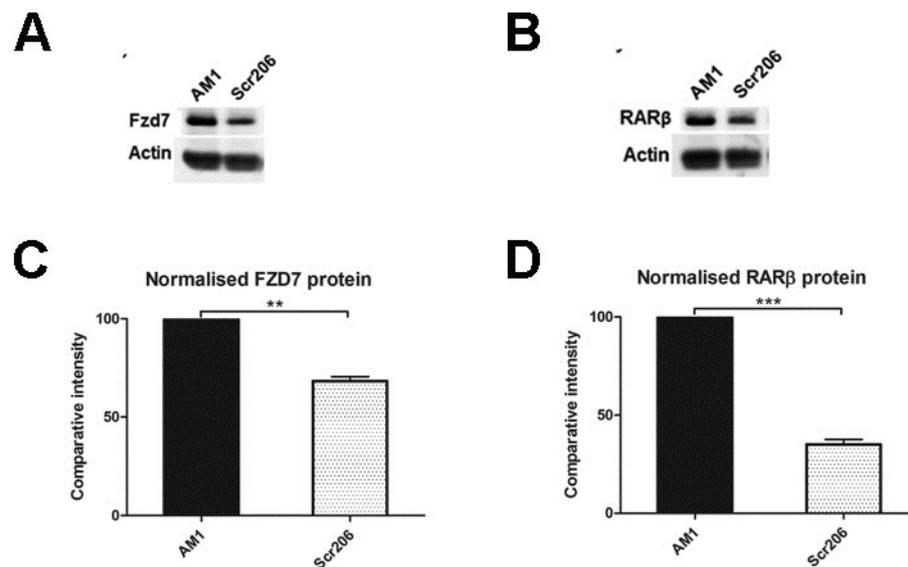
**Figure 6.4: RAR $\beta$  expression in the chick during early stages of development.** WISH for RAR $\beta$  mRNA on chick embryos. **A, B, C:** Ventral view of HH10 (**A**), HH11 (**B**) and HH12 (**C**) embryos with very slight RAR $\beta$  expression seen in the heart tube (black arrows). **D, E:** By HH18 RAR $\beta$  expression has been completely downregulated in the heart (black arrows).

### 6.2.2 miR-1 reduces FZD7 and RAR $\beta$ protein levels in the heart

2'-O-methyl modified cholesterol conjugated and fluorescein isothiocyanate (FITC)-labeled antisense RNAs are known to reduce the levels of endogenous miRNAs and very specifically act on just their complementary miRNA<sup>194,195</sup>. The conjugation of cholesterol

to these small molecules allows direct movement into cells without the need for damaging electroporation. AntagomiR-1 (AM1) inhibits miR-1 expression, whereas an antagomiR with a scrambled sequence complementary to miR-206 (Scr206) does not affect miR-1 levels *in ovo*<sup>127</sup>.

To assess whether miR-1 translationally inhibits FZD7 and RAR $\beta$ , AM1 and a control antagomiR containing a scrambled miR-206 sequence (Scr206) were microinjected into the hearts of HH14 embryos *in ovo*. Western blot analysis of injected heart tissue showed an approximate 38% reduction in FZD7 levels and a 65% drop in RAR $\beta$  levels (**Figure 6.5**). Three biological replicate assays were performed, a representative of which is shown in **Figure 6.5A** and **B**. Each biological assay loaded 40 $\mu$ g of protein per lane (**2.11**), isolated from 24 injected hearts. Quantifications of bands for **Figure 6.5C** and **D** were performed as described in **2.13**.



**Figure 6.5: miR-1 translationally inhibits FZD7 and RAR $\beta$  in the developing chick heart.** Inhibition of miR-1 by antagomiR-1 (AM1) microinjection into HH14 embryonic hearts resulted in increased FZD7 (**A**) and RAR $\beta$  (**B**) protein levels compared to the control. An antagomiR with a scrambled miR-206 sequence (Scr206) was injected as a negative control and actin was used as a loading control. **C,D:** ImageJ<sup>148</sup> quantification (**2.13**) of FZD7 and RAR $\beta$  Western blot bands, normalised to actin and then to AM1, shows a statistically significant increase in protein levels when miR-1 was inhibited with antagomiR-1. N=3, paired t-test. P=0.0002 (**C**) and P=0.0005 (**D**).

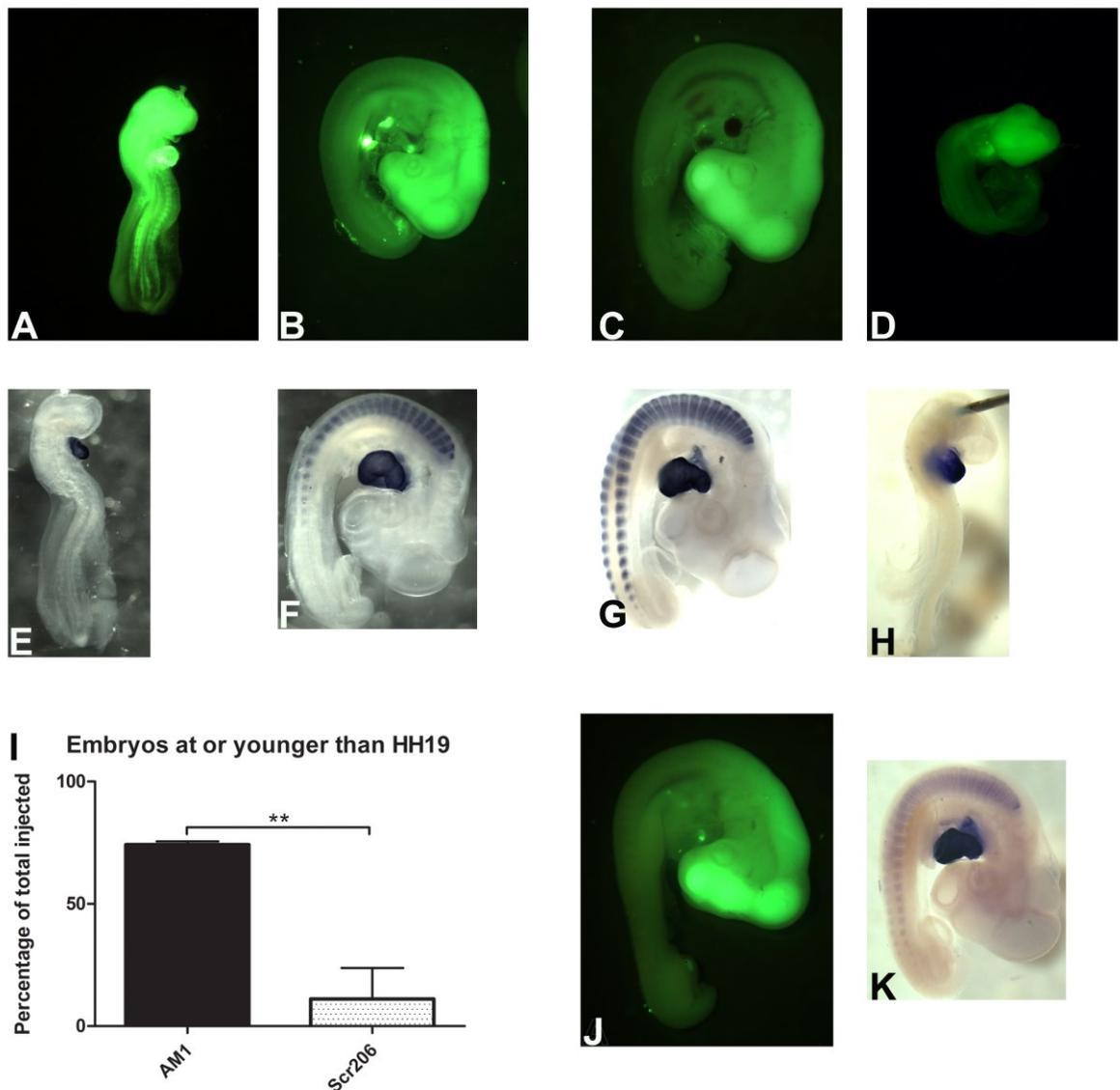
### 6.2.3 Reduction of miR-1 impairs embryonic growth

AntagomiRs are FITC-labelled and therefore embryos successfully injected *in ovo* can be identified by their fluorescence upon harvesting. As expected, injection of antagomiRs into the beating chicken heart resulted in widespread distribution throughout the embryo (**Figure 6.6A, B, C, D, J**). AM1 injection is known to result in miR-1 inhibition whereas Scr206 does not affect miR-1 levels *in ovo*<sup>127</sup>.

During preliminary investigations inhibition of miR-1 by AM1 at HH14 resulted in a striking developmental delay phenotype, discussed further below, compared to the effects of miR-1 inhibition at HH11 (**Figure Apx D.2**) or HH20 (**Figure Apx D.2, Figure Apx D.3, Figure Apx D.4**). Therefore investigations into the effects of miR-1 dysregulation at HH14 were prioritised over further studies at HH11 and HH20.

Embryos were microinjected at HH14 *in ovo*, after being allowed to cool to room temperature to slow the heart rate and minimize damage to the heart during injection, and left to develop for 24 hours, which should have resulted in embryos around developmental stage HH20. The AM1 and Scr206 embryos from three different experiments were staged by counting somites and divided into two groups: those that had reached stage HH20 and older, and those that were at or younger than stage HH19. An average of 75% AM1 injected embryos were found to be younger than HH19, whereas only 9% of Scr206 injected embryos had not reached HH20 or older at the time of harvest. These results were consistent across the three different experiments, and the portion of delayed embryos after AM1 injection was significantly different from that after Scr206 injection (**Figure 6.6I**).

Whole mount *in situ* hybridisation (WISH) for myosin heavy chain 15 (*Myh15*) mRNA was performed on all injected embryos to visualise the heart (**Figure 6.6E, F, G, H, K, Figure Apx D.3, Figure Apx D.5**). *Myh15* is expressed in the heart tube from HH9 and in the somites from HH16<sup>17,147</sup>.



**Figure 6.6 Inhibition of miR-1 by AM1 at HH14 resulted in significant developmental delay not seen upon injection with a control antagomiR (Scr206) nor following inhibition of miR-206 by AM206.** Representative embryos injected with AM1 (A, B, E, F) or Scr206 (C, D, G, H). Most (74%) of embryos injected with AM1 were found to be developmentally delayed (I). 26% of embryos injected with AM1, in contrast with 91% of embryos injected with Scr206, did not show developmental delay. Inhibition of miR-206, a miRNA very closely related in sequence to miR-1, by AM206 at HH14 did not result in the developmental delay observed on inhibition of miR-1 (J, K). The expected widespread distribution of the FITC-labelled antagomiRs following successful cardiac microinjection (A, B, C, D and E) with the corresponding *Myh15* mRNA expression pattern (E, F, G, H and K) is shown in representative embryos. I: Comparison of the fraction embryos found to be at or younger than HH19 after AM1 or Scr206 microinjection at HH14. Inhibition of miR-1 with AM1 resulted in an average of 74% embryos younger than HH19, compared with 9% younger than HH19 after Scr206 injections. N=3, n=10, P=0.001.

#### 6.2.4 Reduction of growth is not seen when miR-206 is inhibited

miR-1 and miR-206 share several targets, such as FZD7 and RAR $\beta$ <sup>8</sup>, which may be affected when AM1 is circulated throughout the embryo. miR-206 was inhibited by injection with AM206 in order to determine whether the observed developmental delay on inhibition of miR-1 at HH14 (**Figure 6.6I**) may have been due to dysregulation of target genes common to miR-1 and miR-206 outside the heart.

The beating heart is essential for development and transfer of nutrients throughout the developing embryo. Injection into the heart results in widespread distribution of the injected antagomiR, as is apparent by FITC fluorescence (**Figure 6.6A, B, C, D, J, Figure Apx D.2** and **Figure Apx D.4**). This resembles the distribution of antagomiR-206 injected into the vitelline vein, which resulted in greatly reduced miR-206 levels in the somites of chick embryos at 24 and 28 hours post injection<sup>196</sup>.

miR-1 and miR-206 have very closely related sequences that differ by only 5 nucleotides in the mature miRNA, and have identical seed sequences (**Figure 6.1A**). miR-1 is expressed specifically in cardiac and skeletal muscle, whereas miR-206 is restricted to skeletal muscle only (**Figure 6.1A**). miR-1 is expressed in the heart from HH11 and in somites from HH14<sup>115</sup>, and miR-206 can be detected in somites from HH14<sup>197</sup>. AntagomiR-206 (AM206) has a sequence very similar to that of AM1, and in the chick embryo AM1 completely inhibits miR-1 but does not affect miR-206 levels<sup>127</sup>. Therefore, to assess whether the developmental delay observed as a result of miR-1 inhibition could have been due to the dysregulation of target genes common to miR-1 and miR-206 outside the heart, miR-206 was inhibited by injection with AM206 (**Figure 6.6J, K, Figure Apx D.6, Figure Apx D.7**).

All AM206 injected embryos had reached at least HH20 at harvest 24 hours after microinjection (**Figure 6.6J, K** and **Figure Apx D.7**). This confirms that the effect observed when miR-1 was inhibited from HH14 was not due to the inhibition of targets shared by miR-1 and miR-206 outside the heart.

In this chapter miR-1 was confirmed to target both FZD7 and RAR $\beta$  *in vivo*, and a developmental delay phenotype was observed following inhibition of miR-1 at HH14 by microinjection into the heart *in ovo*. The possible reasons for this delay are discussed further in **6.3**.

### 6.3 Discussion

miR-1 is known to target the canonical Wnt receptor FZD7 and the retinoic acid receptor RAR $\beta$  *in vitro*<sup>8</sup>. FZD7 is expressed in the heart of chick embryos until around HH15 (**Figure 6.3G, H**), and RAR $\beta$  is absent from the heart by HH18 (**Figure 6.3I, J**). AntagomiR-1 inhibition of miR-1 in developing chicken heart resulted in an average of 38% and 65% reduction of FZD7 and RAR $\beta$  protein levels respectively (**Figure 6.5**).

Microinjection of cholesterol conjugated and FITC labelled antagomiRs into the embryonic heart at HH14 resulted in widespread distribution of the antagomiR throughout the embryo (**Figure 6.6, Figure Apx D.5, Figure Apx D.7 and Figure Apx D.9**). When HH14 embryos were injected with a known specific inhibitor of miR-1, AM1, 74% of embryos were found to be developmentally delayed compared with the 9% of control (AM206) injected embryos (**Figure 6.6I**). Most (91%) of the control injected embryos had reached at least HH20, in keeping with the expected development of embryos in 24 hours after HH14.

miR-1 and miR-206 have identical seed sequences and are very similar in sequence overall, with only 5 nucleotides difference between them (**Figure 6.1B**). Therefore it is not surprising that these miRNAs target many of the same genes, including both FZD7 and RAR $\beta$ <sup>8</sup>. miR-1 is expressed in skeletal and cardiac muscle, whereas miR-206 is absent from the heart (**Figure 6.1B**). To determine whether upregulation of target genes expressed in skeletal muscle that are targeted by both miR-1 and miR-206 was responsible for the developmental delay observed, miR-206 was inhibited by antagomiR-206 injection into the heart at HH14 (**Figure 6.6J, K**). After 24 hours embryos were harvested and found to have reached the expected HH20 developmental stage, indicating that the developmental delay observed when miR-1 was inhibited was not a result of upregulation of target genes in skeletal muscle that are also targeted by miR-206. Due to the essential nature of the beating heart during embryonic development and the widespread distribution of any material injected into the heart throughout the embryo, it is difficult to say whether the developmental delay observed when miR-1 was inhibited was due to heart-specific upregulation of its target genes or whether other genes normally suppressed by miR-1 in

skeletal muscle caused this effect. However, direct injection of AM1 into somites at this stage did not result in an obvious effect<sup>127</sup>.

Heart development was observed to be consistent with the overall developmental stage achieved by delayed AM1-injected (**Figure 6.6A, E**), unaffected AM1 injected (**Figure 6.6B, F**), and control injected (**Figure 6.6C, G**) embryos. Canonical Wnt signalling is suppressed in cells undergoing EMT<sup>7</sup>, and non-canonical Wnt signalling is thought to be essential for epicardial EMT<sup>198</sup>. Epicardial cells undergo EMT and migrate over the myocardium at around HH14, forming the epicardium<sup>18</sup>. Growth in the heart at around HH12 is mostly from migration, but the forming epicardium induces ventricular myocardial cells to become proliferative at HH14 to HH16<sup>18</sup>. As both canonical and non-canonical Wnt receptors compete for the same ligand, it is feasible that the downregulation of canonical Wnt receptors would be essential during times when and in cells where non-canonical signalling is essential. Therefore, the developmental delay that resulted from miR-1 inhibition at HH14 may be due to the loss of FZD7 inhibition allowing this receptor to compete for Wnt ligands and therefore reducing the non-canonical signalling essential for epicardial EMT and as a result, reducing myocardial growth normally expected at HH14 to HH16. Interestingly, preliminary results following antagomiR-1 injections at HH11 and HH20 (**Figure Apx D.2, Figure Apx D.3, Figure Apx D.4**) did not reflect the developmental delay observed when miR-1 was inhibited at HH14. This further suggests that the delay is a result of a specific event around HH14, and may provide further evidence for the initial suggestion that the downregulation of FZD7 is required for epicardial EMT. The targeting of FZD7 by miR-1 is essential for cardiomyocyte commitment<sup>199</sup>, confirming the importance of the interaction between miR-1 and this target. FZD7 overexpression in the HH14 embryonic heart would reveal the effects of complete loss of repression of this gene on heart development. Loss of FZD7 downregulation, such as was achieved in this study by AM1 injection, may affect known markers of epicardial EMT<sup>200</sup> such as *platelet-derived growth factor receptor  $\alpha$*  (*PDGFRA*), *sex determining region Y box 9* (*SOX9*), *fibroblast growth factor 1/2/7* (*FGF1/2/7*), and *Wilms tumor 1* transcription factor (*WT1*).

However, miR-1 is also predicted to target the *Snail family zinc finger 2 (SNAI2)* transcription factor<sup>131</sup>. SNAI2 is a repressor of *E-cadherin* transcription, which must be downregulated for EMT to proceed<sup>201</sup>. SNAI2 is expressed in the primitive streak and mesoderm of the chick embryo in cells undergoing EMT during gastrulation. However, SNAI2 downregulation did not result in increased levels of E-cadherin or in a discernible loss of EMT in gastrulating embryos<sup>201</sup>, which suggests that EMT during gastrulation does not depend on E-cadherin downregulation. In the chick SNAI2 is expressed in the migrating cardiac neural crest at HH15<sup>17,147</sup> and in the ventricle at HH21<sup>202</sup>, and is known to be required for EMT in the chick heart<sup>203</sup>. It would be interesting to determine whether miR-1 targets SNAI2 *in vivo* and to complement FZD7 investigations by overexpressing SNAI2 at HH14 to examine whether miR-1 ensures commitment to a cardiac fate as migrating cells move into the heart by repressing both these targets. At HH22-HH23 apoptosis is increased in the ventricle around the junction with the OFT<sup>18</sup>. The epicardium has been implicated as being the source of factors that induce apoptosis in the myocardium<sup>18</sup>, and SNAI2 has been shown to have anti-apoptotic activity in haematopoietic cells<sup>203</sup>. It may be that its expression in tissue lining the ventricular cavity serves to counteract the epicardium-induced apoptosis in this tissue, and that miR-1 ensures apoptosis in the rest of the ventricle by decreasing the activity of *SNAI2* by transcriptionally repressing this gene.

Epicardial-induced growth is known to depend on RA signalling<sup>192</sup> and miR-1 has here been shown to reduce the expression of RAR $\beta$  (**Figure 6.5**). The genetic knockout of the RAR $\beta$  gene in mice did not result in heart developmental defects and embryos were found to be viable<sup>34</sup>, whereas RAR $\alpha$ :RAR $\beta$  null mutants died shortly after birth and displayed outflow tract and large vessel malformations<sup>34</sup>. Furthermore, the cardiac neural crest cells that migrate into the heart and contribute to septation of the outflow tract require RA signalling for their migration<sup>34</sup>. It is possible that miR-1 aids outflow tract septation by repressing RA signalling in these cells as they enter the heart and further dysregulation of miR-1 and overexpression of RAR $\beta$  in the heart between HH18-HH20 would test this hypothesis. The differences in the observed phenotypes between RAR $\beta$  and RAR $\alpha$ :RAR $\beta$  genetic knockout mice, along with the evidence showing RA signalling is required for cardiac neural crest cell migration and resulting outflow tract formation, suggests however

that it may be RAR $\alpha$  that is essential for this aspect of cardiogenesis. It would be interesting to determine whether RAR $\beta$  competes with RAR $\alpha$ , and whether abrogation of the effect of AM1 injections at HH14 results by simultaneous reduction of RAR $\beta$  protein levels with a dominant negative form of RAR $\beta$ . If these receptors do compete for RA ligand and this competition results in the observed effect of miR-1 dysregulation, then a similar effect should be seen when RAR $\alpha$  is inhibited with a dominant negative form of the RAR $\alpha$  protein by electroporation to the heart.

Furthermore, miR-1 has been shown to target the HAND2 transcription factor which is required for cardiomyocyte expansion in the ventricles<sup>108</sup>. When Hand2 was inhibited by antisense oligonucleotide injections at HH8 developmental arrest followed 18 hours after treatment<sup>204</sup>, and overexpression of miR-1 at E9 in the developing mouse heart also lead to developmental arrest (at E13.5)<sup>108</sup>.

Investigating the *in vivo* functions of miRNAs during development is extremely challenging and yet can give invaluable information regarding the roles of a miRNA in a complex biological system with multiple targets and many different tissue types. miR-1 is known to be essential for cardiogenesis and it is likely to have many different roles during development dependent on the specific spatiotemporal effects that result from its functional interactions with its many targets, and further study of this fascinating miRNA is definitely required. Although this chapter raises more questions than it answers, it does give a glimpse into the specific effects of miR-1 dysregulation at HH14 and has led to several suggestions for further work.

## 7 General discussion and concluding remarks

### 7.1 Study aims

The aims of this study were:

1. To generate a miRNA expression profile of the developing chicken heart from embryonic stage HH5 to stage HH23.
2. To explore the functions of miR-1 during chicken embryonic heart development.

A NGS miRNA expression profile was generated for 9 samples from 6 time points (HH5-7, HH8-10, HH11-13, HH14-16, HH17-20, HH21-23). Bioinformatic analyses of these data revealed 31 known animal miRNAs not currently annotated in the chicken genome (7.2.1), established that miR-130 is highly expressed at early stages and in later stages in heart tissue samples (7.2.2), and showed that three miRNAs (miR-26, miR-30, miR-92) were highly expressed in all samples (7.2.3).

A fold change profile of all identified miRNAs over their levels at HH5-7 was generated, which revealed 14 miRNAs with increases of expression more than 100 fold over their HH5-7 levels (7.3.1). Of these 14 miRNAs, 6 are known to be involved in heart development and function. Subtraction of the fold change values of each miRNA in NT samples from their values in matched Heart samples generated a Heart miRNA expression profile that revealed 38 miRNAs detected at 10 fold higher levels in Heart samples than in NT samples (7.3.2). Additionally, 13 reporter constructs were validated to be repressed by at least one of miR-130, miR-138, or miR-499 (7.3.3). The most notable of these included a regulator (ETS1) of miR-126 expression. miR-126 is essential for valve formation and angiogenesis. In addition, miR-138 was found to regulate ZEB2 which affects growth and angiogenesis by its effects on the PTEN signalling cascade. Two predicted novel miRNAs were confirmed in chicken embryonic tissue (7.4).

miR-1 was shown to target both *FZD7* and *RARβ* *in vivo* (7.5). Finally, miR-1 inhibition by antagomiR cardiac injections *in ovo* at HH14 was shown to result in developmental delay of 75% of injected embryos (7.5).

## 7.2 miRNA expression profile

The miRNA expression profile of areas around the forming heart of the chicken embryo was established by NGS for nine samples ranging from HH5, when cardiac progenitor cells are specified, to HH23, during ventricular septation.

### 7.2.1 Known animal miRNAs not annotated in the chicken genome were identified by NGS

Analysis of all the sequences detected within the miRNA-sized fraction revealed the presence of 31 miRNAs that map to the chicken genome and are known in other animals but not known in the chick (Table Apx A.1). Three of these miRNAs (miR-143, miR-182, miR-363) are known to be involved in heart development and function<sup>75,88</sup>, and were found to be very highly expressed in the samples assayed (3.2.5). All 31 of these miRNAs will be submitted to miRBase for annotation of the chicken genome.

### 7.2.2 miR-130 is highly expressed in heart tissue samples

The normalised expression levels of miR-130 were very high in early samples (HH5-7, HH8-10, HH11-13) and in later heart but not NT samples (HH14-16, HH17-20, HH21-23) (Figure 3.7). Apart from its dysregulation in diabetic mouse hearts<sup>79</sup>, nothing is known of the roles of miR-130 in heart development and function. Interestingly, miR-31 and miR-130 have significantly correlated patterns of expression (Figure 4.9F). miR-31 is known to decrease cardiac fibrogenic EMT by the negative effects on its target *Isl1*<sup>117</sup>. *Isl1* expression during early cardiomyocyte differentiation has recently been shown to be promoted by Wnt signalling via a novel TCF/LEF1 binding site<sup>165</sup>. In this study miR-130 was found to reduce the expression of a reporter construct containing the 3' UTR of *LEF1*, and a reporter construct containing the 3' UTR of *TCF12* was inhibited by miR-138 (Figure

**4.13C and J).** The combined actions of miR-130 and miR-138 during cardiogenesis may result in the repression of *Isl1* by their effects on the translation of *LEF1* and *TCF12*.

### **7.2.3 miR-26, miR-30 and miR-92 were highly expressed across all stages**

All nine samples consistently showed very high levels of three miRNAs: miR-26, miR-30 and miR-92 (**Figure 3.7**). More than 40% of the predicted target genes of miR-26 and miR-92 were also shared by the other miRNAs, but a large number of the predicted targets of miR-30 were found to be unique (**Figure 3.9**). This high degree of predicted target overlap was maintained when only those predicted targets known to be expressed in the human heart were considered (**Figure 3.10, Table Apx A.3, Table Apx A.4**).

miR-10 was the most sequenced miRNA in the older samples and dominates especially in the non-heart samples (**Figure 3.7**). Although all four these highly expressed miRNAs share many predicted targets only three GO terms, all related to nerve development, were found to be significantly associated with the unique predicted targets of miR-10 (**Figure 3.9**). This suggests that miR-10 may suppress nerve development in non-neural tissues by counteracting factors promoting the formation of the neural tube, in addition to its known roles in angiogenesis miR-10<sup>92</sup>.

## **7.3 miRNAs important during heart development**

### **7.3.1 Finding highly expressed miRNAs by generation of a fold change profile**

The normalised reads recorded for most of the 423 known animal miRNAs detected in at least two samples were found to either stay the same as or fall below their levels recorded in the HH5-7 sample (**Figure 4.3**). Most of the miRNAs found to increase as development proceeded did so by between 1 and 10 fold over their recorded reads at the earliest time point, and only 14 miRNAs were found to increase between 100 and 300 fold over HH5-7 (**Table 4.1**). Six (miR-1, miR-10, miR-126, miR-145, miR-451, miR-499) of the 14 miRNAs found to increase by the largest margin have known roles in heart development

and function, which implies that the remaining 8 miRNAs that are upregulated more than 100 fold may also have important roles in the heart. Large increases in expression of three of these miRNAs (miR-126, miR-451, miR-499) were observed in the HH17-20 Heart sample (**Figure 4.4C, E, F**). Valve formation is known to start during these stages and the positive effects of miR-126 on VEGF are essential to this process<sup>110</sup>. Interestingly, miR-499 was found to inhibit a reporter construct containing the 3' UTR of ETS1, a known transcriptional activator of miR-126<sup>74</sup> (**Figure 4.13A**), discussed further in **7.3.3**.

### **7.3.2 Establishment of a heart miRNA expression profile, generated by subtracting miRNA fold change values in matched NT samples from Heart samples**

The subtraction of fold change values of known animal miRNAs detected in NT samples at the HH14-16, HH17-20 and HH21-23 time points from their values in matched Heart samples resulted in the generation of a fold change profile specific to the developing embryonic heart (**Table 4.2**). Fourteen miRNAs were found to be more than 100 fold increased in heart over their matched NT values, including 10 miRNAs known to have roles in the development or functioning of the heart (**Table Apx A.6**). Analysis of miRNAs with significantly correlated patterns of H – NT values over these three time points revealed possible networks of miRNAs that behave in a similar manner to known heart miRNAs (**Table 4.3**).

### **7.3.3 Validation of miRNA targets**

Several targets of miR-130, miR-138 and miR-499 were validated in this study following analysis of their effects on reporter constructs containing the 3' UTRs of several genes *in vitro*.

In this study miR-499 was found to target a reporter construct containing the 3' UTR of *ETS1* (**Figure 4.13A**) and the fold change patterns of miR-126, miR-499 and miR-221 were found to correlate closely (**Figure 4.7A, Figure 4.9C**). Recently, miR-221 was shown to modulate the pro-apoptotic effect miR-126 has on VSMCs by inhibiting the

transcriptional activator of miR-126, ETS1<sup>155</sup>. miR-499 may be functioning to attenuate the effects of miR-126 in a similar way to miR-221, as described in **Figure 4.16**.

miR-138 was found to inhibit the translation of a reporter construct containing the 3' UTR of *ZEB2*, which is known to inhibit two miRNAs (miR-216, miR-217) that target PTEN<sup>160</sup>. PTEN is an essential inhibitor of the PI3K-AKT cascade, and is known to have negative effects on growth and angiogenesis<sup>160</sup>. *ZEB2* is also known to be targeted by miR-192<sup>160</sup>, which was found to have a significantly correlated fold change pattern to that of miR-138. PTEN is a predicted and confirmed target not only of most of 11 heart miRNAs evaluated (**Table 4.6, Table 4.8A**), but also of many of the miRNAs significantly correlated to each of these miRNAs by fold change pattern (**Table 4.8B**). This suggests that miR-138 might, along with miR-192, allow the inhibition of PTEN by targeting *ZEB2* (**Figure 4.17**).

Several predicted target genes with many and varied roles during development were validated as targets of three miRNAs, miR-130, miR-138 and miR-499, *in vitro*. This provides a starting point for further investigations into the essential functions of miRNAs during development.

## 7.4 Novel miRNAs

Three of thirteen predicted novel miRNAs were identified as a result of this screen, although one of these has since been annotated in miRBase as miR-1462-3p. The predicted target genes of miR-1462 were significantly associated with neurogenesis, generation of neurons and with nervous system development, as well as with the known targets of miR-135. miR-135 is known to inhibit osteogenesis in C2C12 cells<sup>179</sup>. The dysregulation of miR-1462 and miR-135 during the development of skeletal muscle may indicate an essential role for these miRNAs. Although the predicted target genes of Putative 12 did not have a significant association with any specific cellular processes, its expression is high during the early stages of development and further investigation of its expression in the developing chick may narrow down its potential roles during chick cardiogenesis, if any. The predicted target genes of Putative 14 were found to be significantly similar to the

known targets of three miRNAs that negatively affect growth and proliferation: miR-198, miR-593, and mir-129<sup>181,182,185</sup>.

## 7.5 miR-1 affects growth during early chick heart development

miR-1 was found to translationally inhibit the canonical Wnt receptor FZD7 as well as the retinoid acid receptor RAR $\beta$  *in vivo* (**Figure 6.5**). Inhibition of miR-1 at HH14 by microinjection with a cholesterol conjugated antagomiR directly into the heart resulted in a significant number of embryos being developmentally delayed (**Figure 6.6I**). The developmental delay was not seen in control injections with antagomiR-206 or scrambled 206 (**Figure 6.6I, J, K**). The hearts of AM1 injected embryos were found to be at a developmental stage consistent with their overall development, suggesting that miR-1 inhibition may have resulted in the loss of growth-stimulating epicardium due to a lack of epicardial EMT at this stage of development. This effect may have been due to increased levels of the canonical Wnt receptor FZD7, normally repressed by miR-1, competing with non-canonical Wnt receptors for ligand. Further examination of the observed effects of miR-1 inhibition could be made by evaluating epicardial EMT in treated and control treated embryos, and by overexpressing FZD7 in the HH14 heart.

However, miR-1 was also shown to repress RAR $\beta$  in the developing chick heart and outflow tract septation is known to depend upon RA signalling in migrating cardiac neural crest cells<sup>34</sup>. Genetic knockouts of RAR $\alpha$  and RAR $\beta$  suggest that this RA signalling proceeds mainly through RAR $\alpha$ . Repression of RAR $\beta$  levels in these cells may reduce the competition for RA ligand, allowing miR-1 to aid outflow tract septation by its repression of RAR $\beta$ .

Although novel miRNAs and several previously unknown interactions of known miRNAs and mRNAs were elucidated in this study, much work can still be done to further examine the many and varied roles of these miRNAs during cardiogenesis. Furthermore, some of the roles of the extremely interesting miR-1 are well known, but many of its functions are likely still unexplored and merit further investigation. This profiling study may serve as a

guide for future studies of miRNAs involved in heart development and function, and several suggestions for future work have been made.

## 8 References

1. Buckingham, M., Meilhac, S. & Zaffran, S. Building the mammalian heart from two sources of myocardial cells. *Nat Rev Genet* **6**, 826-835 (2005).
2. Small, E.M. & Olson, E.N. Pervasive roles of microRNAs in cardiovascular biology. *Nature* **469**, 336-342 (2011).
3. Srivastava, D. Making or breaking the heart: from lineage determination to morphogenesis. *Cell* **126**, 1037-1048 (2006).
4. Townley-Tilson, W.H., Callis, T.E. & Wang, D. MicroRNAs 1, 133, and 206: critical factors of skeletal and cardiac muscle development, function, and disease. *Int J Biochem Cell Biol* **42**, 1252-1255 (2010).
5. Zhao, Y., *et al.* Dysregulation of cardiogenesis, cardiac conduction, and cell cycle in mice lacking miRNA-1-2. *Cell* **129**, 303-317 (2007).
6. Chen, J.F., Callis, T.E. & Wang, D.Z. microRNAs and muscle disorders. *J Cell Sci* **122**, 13-20 (2009).
7. Gessert, S. & Kuhl, M. The multiple phases and faces of wnt signaling during cardiac differentiation and development. *Circ Res* **107**, 186-199 (2010).
8. Goljanek-Whysall, K., *et al.* Regulation of multiple target genes by miR-1 and miR-206 is pivotal for C2C12 myoblast differentiation. *J Cell Sci* **125**, 3590-3600 (2012).
9. Ruijtenbeek, K., De Mey, J.G. & Blanco, C.E. The chicken embryo in developmental physiology of the cardiovascular system: a traditional model with new possibilities. *Am J Physiol Regul Integr Comp Physiol* **283**, R549-550; author reply R550-541 (2002).
10. Hamburger, V. & Hamilton, H.L. A series of normal stages in the development of the chick embryo. 1951. *Dev Dyn* **195**, 231-272 (1992).
11. Lee, R.C., Feinbaum, R.L. & Ambros, V. The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* **75**, 843-854 (1993).
12. Reinhart, B.J., *et al.* The 21-nucleotide *let-7* RNA regulates developmental timing in *Caenorhabditis elegans*. *Nature* **403**, 901-906 (2000).
13. Wightman, B., Ha, I. & Ruvkun, G. Posttranscriptional regulation of the heterochronic gene *lin-14* by *lin-4* mediates temporal pattern formation in *C. elegans*. *Cell* **75**, 855-862 (1993).
14. Bernstein, E., *et al.* Dicer is essential for mouse development. *Nat Genet* **35**, 215-217 (2003).
15. Chen, J.F., *et al.* Targeted deletion of Dicer in the heart leads to dilated cardiomyopathy and heart failure. *Proc Natl Acad Sci U S A* **105**, 2111-2116 (2008).
16. Bartel, D.P. MicroRNAs: target recognition and regulatory functions. *Cell* **136**, 215-233 (2009).
17. Darnell, D.K., *et al.* GEISHA: an in situ hybridization gene expression resource for the chicken embryo. *Cytogenet Genome Res* **117**, 30-35 (2007).
18. Martinsen, B.J. Reference guide to the stages of chick heart embryology. *Dev Dyn* **233**, 1217-1237 (2005).
19. Abu-Issa, R. & Kirby, M.L. Heart field: from mesoderm to heart tube. *Annu Rev Cell Dev Biol* **23**, 45-68 (2007).
20. Abu-Issa, R. & Kirby, M.L. Patterning of the heart field in the chick. *Dev Biol* **319**, 223-233 (2008).
21. Brand, T. Heart development: molecular insights into cardiac specification and early morphogenesis. *Dev Biol* **258**, 1-19 (2003).
22. Evans, S.M. Myocardial lineage development. *Circ. Res.* **107**, 1428-1444 (2010).
23. Yamagishi, H., *et al.* Molecular embryology for an understanding of congenital heart diseases. *Anat Sci Int* **84**, 88-94 (2009).
24. Bittel, D.C., Kibiryeva, N., Marshall, J.A. & O'Brien, J.E. MicroRNA-421 Dysregulation is Associated with Tetralogy of Fallot. *Cells* **3**, 713-723 (2014).
25. Kemp, C.R., *et al.* Expression of Frizzled5, Frizzled7, and Frizzled10 during early mouse development and interactions with canonical Wnt signaling. *Dev Dyn* **236**, 2011-2019 (2007).

26. van Gijn, M.E., Daemen, M.J., Smits, J.F. & Blankesteyn, W.M. The wnt-frizzled cascade in cardiovascular disease. *Cardiovasc Res* **55**, 16-24 (2002).
27. Schneider, V.A. & Mercola, M. Wnt antagonism initiates cardiogenesis in *Xenopus laevis*. *Genes Dev* **15**, 304-315 (2001).
28. Tzahor, E. & Lassar, A.B. Wnt signals from the neural tube block ectopic cardiogenesis. *Genes Dev* **15**, 255-260 (2001).
29. Wodarz, A. & Nusse, R. Mechanisms of Wnt signaling in development. *Annu Rev Cell Dev Biol* **14**, 59-88 (1998).
30. van Amerongen, R. & Nusse, R. Towards an integrated view of Wnt signaling in development. *Development* **136**, 3205-3214 (2009).
31. Mayor, R. & Theveneau, E. The role of the non-canonical Wnt-planar cell polarity pathway in neural crest migration. *Biochem J* **457**, 19-26 (2014).
32. Wagner, M. & Siddiqui, M.A. Signal transduction in early heart development (I): cardiogenic induction and heart tube formation. *Exp Biol Med (Maywood)* **232**, 852-865 (2007).
33. Malizia, A.P. & Wang, D.Z. MicroRNAs in cardiomyocyte development. *Wiley Interdiscip Rev Syst Biol Med* **3**, 183-190 (2011).
34. Niederreither, K. & Dolle, P. Retinoic acid in development: towards an integrated view. *Nat Rev Genet* **9**, 541-553 (2008).
35. Krol, J., Loedige, I. & Filipowicz, W. The widespread regulation of microRNA biogenesis, function and decay. *Nat Rev Genet* **11**, 597-610 (2010).
36. Filipowicz, W., Bhattacharyya, S.N. & Sonenberg, N. Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? *Nat Rev Genet* **9**, 102-114 (2008).
37. Gu, S., Jin, L., Zhang, F., Sarnow, P. & Kay, M.A. Biological basis for restriction of microRNA targets to the 3' untranslated region in mammalian mRNAs. *Nat Struct Mol Biol* **16**, 144-150 (2009).
38. Grimson, A., *et al.* MicroRNA targeting specificity in mammals: determinants beyond seed pairing. *Mol Cell* **27**, 91-105 (2007).
39. Prüfer, K., *et al.* PatMaN: rapid alignment of short sequences to large databases. *Bioinformatics* **24**, 1530-1531 (2008).
40. Meyer, L.R., *et al.* The UCSC Genome Browser database: extensions and updates 2013. *Nucleic Acids Res* **41**, D64-69 (2013).
41. Kozomara, A. & Griffiths-Jones, S. miRBase: integrating microRNA annotation and deep-sequencing data. *Nucleic Acids Res* **39**, D152-157 (2011).
42. Stocks, M.B., *et al.* The UEA sRNA workbench: a suite of tools for analysing and visualizing next generation sequencing microRNA and small RNA datasets. *Bioinformatics* **28**, 2059-2061 (2012).
43. Flatow, J.H., M. Feng, G. Du, P. Kibbe, W.A. Lin, S.M. FunDO: Exploring Genes Using Functional Disease Ontology Annotations. <http://projects.bioinformatics.northwestern.edu/fundo> (2009).
44. Osborne, J.D., *et al.* Annotating the human genome with Disease Ontology. *BMC Genomics* **10 Suppl 1**, S6 (2009).
45. Du, P., *et al.* From disease ontology to disease-ontology lite: statistical methods to adapt a general-purpose ontology for the test of gene-ontology associations. *Bioinformatics* **25**, i63-68 (2009).
46. Falcon, S. & Gentleman, R. Using GOstats to test gene lists for GO term association. *Bioinformatics* **23**, 257-258 (2007).
47. Tuddenham, L., *et al.* The cartilage specific microRNA-140 targets histone deacetylase 4 in mouse cells. *FEBS Lett* **580**, 4214-4217 (2006).
48. Rasband, W.S. ImageJ. U. S. National Institutes of Health, Bethesda, Maryland, USA [http://imagej.nih.gov/ij/\(1997-2014\)](http://imagej.nih.gov/ij/(1997-2014)).
49. Abramoff, M.D., Magelhaes, P.J. & Ram, S.J. Image Processing with ImageJ. *Biophotonics International* **11**, 36-42 (2004).
50. Metzker, M.L. Sequencing technologies - the next generation. *Nat Rev Genet* **11**, 31-46 (2010).
51. Ansorge, W.J. Next-generation DNA sequencing techniques. *N Biotechnol* **25**, 195-203 (2009).

52. Quail, M.A., *et al.* A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. *BMC Genomics* **13**, 341 (2012).
53. LifeTechnologies. 5500 Series Genetic Analysis Systems. <[http://tools.lifetechnologies.com/content/sfs/brochures/cms\\_096631.pdf](http://tools.lifetechnologies.com/content/sfs/brochures/cms_096631.pdf)> (Accessed 19 February 2014).
54. Roche. GS FLX+ System. <[http://454.com/downloads/GSFLXApplicationFlyer\\_FINALv2.pdf](http://454.com/downloads/GSFLXApplicationFlyer_FINALv2.pdf)> (Accessed 18 February 2014).
55. Loman, N.J. Next Generation Genomics: World Map of High-throughput Sequencers. <<http://omicsmaps.com/stats>> (Accessed 18 February 2014).
56. Zhang, J., Chiodini, R., Badr, A. & Zhang, G. The impact of next-generation sequencing on genomics. *J Genet Genomics* **38**, 95-109 (2011).
57. Junemann, S., *et al.* Updating benchtop sequencing performance comparison. *Nat Biotechnol* **31**, 294-296 (2013).
58. Loman, N.J., *et al.* Performance comparison of benchtop high-throughput sequencing platforms. *Nat Biotechnol* **30**, 434-439 (2012).
59. Saunders, C.J., *et al.* Rapid whole-genome sequencing for genetic disease diagnosis in neonatal intensive care units. *Sci Transl Med* **4**, 154ra135 (2012).
60. Yin, J.Q., Zhao, R.C. & Morris, K.V. Profiling microRNA expression with microarrays. *Trends Biotechnol* **26**, 70-76 (2008).
61. Git, A., *et al.* Systematic comparison of microarray profiling, real-time PCR, and next-generation sequencing technologies for measuring differential microRNA expression. *RNA* **16**, 991-1006 (2010).
62. Chen, C., *et al.* Real-time quantification of microRNAs by stem-loop RT-PCR. *Nucleic Acids Res* **33**, e179 (2005).
63. Balcells, I., Cirera, S. & Busk, P.K. Specific and sensitive quantitative RT-PCR of miRNAs with DNA primers. *BMC Biotechnol* **11**, 70 (2011).
64. van Dijk, E.L., Jaszczyszyn, Y. & Thermes, C. Library preparation methods for next-generation sequencing: Tone down the bias. *Exp Cell Res* **322**, 12-20 (2014).
65. Craig, D.W., *et al.* Identification of genetic variants using bar-coded multiplexed sequencing. *Nat Methods* **5**, 887-893 (2008).
66. Smith, A.M., *et al.* Highly-multiplexed barcode sequencing: an efficient method for parallel analysis of pooled samples. *Nucleic Acids Res* **38**, e142 (2010).
67. Hicks, J.A., Tembhurne, P. & Liu, H.C. MicroRNA expression in chicken embryos. *Poult Sci* **87**, 2335-2343 (2008).
68. Li, T., Wu, R., Zhang, Y. & Zhu, D. A systematic analysis of the skeletal muscle miRNA transcriptome of chicken varieties with divergent skeletal muscle growth identifies novel miRNAs and differentially expressed miRNAs. *BMC Genomics* **12**, 186 (2011).
69. Hicks, J.A., Trakooljul, N. & Liu, H.C. Discovery of chicken microRNAs associated with lipogenesis and cell proliferation. *Physiol Genomics* (2010).
70. Yao, J., Wang, Y., Wang, W., Wang, N. & Li, H. Solexa sequencing analysis of chicken pre-adipocyte microRNAs. *Biosci Biotechnol Biochem* **75**, 54-61 (2011).
71. Burnside, J., *et al.* Deep sequencing of chicken microRNAs. *BMC Genomics* **9**, 185 (2008).
72. Rathjen, T., *et al.* High throughput sequencing of microRNAs in chicken somites. *FEBS Lett* **583**, 1422-1426 (2009).
73. Rao, P.K., *et al.* Loss of cardiac microRNA-mediated regulation leads to dilated cardiomyopathy and heart failure. *Circ Res* **105**, 585-594 (2009).
74. Wei, Y., Schober, A. & Weber, C. Pathogenic arterial remodeling: the good and bad of microRNAs. *Am J Physiol Heart Circ Physiol* **304**, H1050-1059 (2013).
75. Deacon, D.C., *et al.* The miR-143-adducin3 pathway is essential for cardiac chamber morphogenesis. *Development* **137**, 1887-1896 (2010).
76. Ishizu, H., Siomi, H. & Siomi, M.C. Biology of PIWI-interacting RNAs: new insights into biogenesis and function inside and outside of germlines. *Genes Dev* **26**, 2361-2373 (2012).
77. Shao, P., *et al.* Drastic expression change of transposon-derived piRNA-like RNAs and microRNAs in early stages of chicken embryos implies a role in gastrulation. *RNA Biol* **9**, 212-227 (2012).

78. R Development Core Team (2012). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org/>. (2012).
79. Chavali, V., Tyagi, S.C. & Mishra, P.K. Differential expression of dicer, miRNAs, and inflammatory markers in diabetic Ins2<sup>+/-</sup> Akita hearts. *Cell Biochem Biophys* **68**, 25-35 (2014).
80. Pan, J.B., *et al.* PaGenBase: a pattern gene database for the global and dynamic understanding of gene function. *PLoS One* **8**, e80747 (2013).
81. Reimand, J., Arak, T. & Vilo, J. g:Profiler--a web server for functional interpretation of gene lists (2011 update). *Nucleic Acids Res* **39**, W307-315 (2011).
82. Bisping, E., *et al.* Transcription factor GATA4 is activated but not required for insulin-like growth factor 1 (IGF1)-induced cardiac hypertrophy. *J Biol Chem* **287**, 9827-9834 (2012).
83. Yu, H., Waddell, J.N., Kuang, S. & Bidwell, C.A. Park7 expression influences myotube size and myosin expression in muscle. *PLoS One* **9**, e92030 (2014).
84. Li, J., *et al.* The mAkapbeta scaffold regulates cardiac myocyte hypertrophy via recruitment of activated calcineurin. *J Mol Cell Cardiol* **48**, 387-394 (2010).
85. Zhang, W., *et al.* Ca<sup>2+</sup>/calmodulin-dependent protein kinase I $\delta$  orchestrates G-protein-coupled receptor and electric field stimulation-induced cardiomyocyte hypertrophy. *Clin Exp Pharmacol Physiol* **37**, 795-802 (2010).
86. Fredj, S., *et al.* Role of interleukin-6 in cardiomyocyte/cardiac fibroblast interactions during myocyte hypertrophy and fibroblast proliferation. *J Cell Physiol* **204**, 428-436 (2005).
87. Castro, M.G., *et al.* Screening of the endothelin1 gene (EDN1) in a cohort of patients with essential left ventricular hypertrophy. *Ann Hum Genet* **71**, 601-610 (2007).
88. Taurino, C., *et al.* Gene expression profiling in whole blood of patients with coronary artery disease. *Clin Sci (Lond)* **119**, 335-343 (2010).
89. Drake, C.J. Embryonic and adult vasculogenesis. *Birth Defects Res C Embryo Today* **69**, 73-82 (2003).
90. James, J.M., Gewolb, C. & Bautch, V.L. Neurovascular development uses VEGF-A signaling to regulate blood vessel ingression into the neural tube. *Development* **136**, 833-841 (2009).
91. Jin, P., *et al.* Differentiation of two types of mobilized peripheral blood stem cells by microRNA and cDNA expression analysis. *J Transl Med* **6**, 39 (2008).
92. Hassel, D., *et al.* MicroRNA-10 regulates the angiogenic behavior of zebrafish and human endothelial cells by promoting vascular endothelial growth factor signaling. *Circ Res* **111**, 1421-1433 (2012).
93. van Rooij, E., Liu, N. & Olson, E.N. MicroRNAs flex their muscles. *Trends Genet.* **24**, 159-166 (2008).
94. van Rooij, E., *et al.* Control of stress-dependent cardiac growth and gene expression by a microRNA. *Science* **316**, 575-579 (2007).
95. Tian, G., *et al.* Sequencing bias: comparison of different protocols of microRNA library construction. *BMC Biotechnol* **10**, 64 (2010).
96. Linsen, S.E., *et al.* Limitations and possibilities of small RNA digital gene expression profiling. *Nat Methods* **6**, 474-476 (2009).
97. Sorefan, K., *et al.* Reducing ligation bias of small RNAs in libraries for next generation sequencing. *Silence* **3**, 4 (2012).
98. Chen, J., *et al.* mir-17-92 cluster is required for and sufficient to induce cardiomyocyte proliferation in postnatal and adult hearts. *Circ Res* **112**, 1557-1566 (2013).
99. Luo, X., *et al.* MicroRNA-26 governs profibrillatory inward-rectifier potassium current changes in atrial fibrillation. *J Clin Invest* **123**, 1939-1951 (2013).
100. Zhang, Z.H., *et al.* MicroRNA-26 was decreased in rat cardiac hypertrophy model and may be a promising therapeutic target. *J Cardiovasc Pharmacol* **62**, 312-319 (2013).
101. Bridge, G., *et al.* The microRNA-30 family targets DLL4 to modulate endothelial cell behavior during angiogenesis. *Blood* **120**, 5063-5072 (2012).
102. Doebele, C., *et al.* Members of the microRNA-17-92 cluster exhibit a cell-intrinsic antiangiogenic function in endothelial cells. *Blood* **115**, 4944-4950 (2010).
103. Moura, J., Borsheim, E. & Carvalho, E. The Role of MicroRNAs in Diabetic Complications-Special Emphasis on Wound Healing. *Genes (Basel)* **5**, 926-956 (2014).

104. Azhar, M., *et al.* Transforming growth factor beta in cardiovascular development and function. *Cytokine Growth Factor Rev* **14**, 391-407 (2003).
105. Liu, N. & Olson, E.N. MicroRNA regulatory networks in cardiovascular development. *Dev Cell* **18**, 510-525 (2010).
106. Chen, J. & Wang, D.Z. microRNAs in cardiovascular development. *J Mol Cell Cardiol* **52**, 949-957 (2012).
107. Chen, J.F., *et al.* The role of microRNA-1 and microRNA-133 in skeletal muscle proliferation and differentiation. *Nat Genet* **38**, 228-233 (2006).
108. Zhao, Y., Samal, E. & Srivastava, D. Serum response factor regulates a muscle-specific microRNA that targets Hand2 during cardiogenesis. *Nature* **436**, 214-220 (2005).
109. Pan, Z., *et al.* MicroRNA-101 inhibited postinfarct cardiac fibrosis and improved left ventricular compliance via the FBJ osteosarcoma oncogene/transforming growth factor-beta1 pathway. *Circulation* **126**, 840-850 (2012).
110. Stankunas, K., Ma, G.K., Kuhnert, F.J., Kuo, C.J. & Chang, C.P. VEGF signaling has distinct spatiotemporal roles during heart valve development. *Dev Biol* **347**, 325-336 (2010).
111. Morton, S.U., *et al.* microRNA-138 modulates cardiac patterning during embryonic development. *Proc Natl Acad Sci U S A* **105**, 17830-17835 (2008).
112. Zhang, X., *et al.* Synergistic effects of the GATA-4-mediated miR-144/451 cluster in protection against simulated ischemia/reperfusion-induced cardiomyocyte death. *J Mol Cell Cardiol* **49**, 841-850 (2010).
113. Pase, L., *et al.* miR-451 regulates zebrafish erythroid maturation in vivo via its target gata2. *Blood* **113**, 1794-1804 (2009).
114. Liu, N., *et al.* microRNA-133a regulates cardiomyocyte proliferation and suppresses smooth muscle gene expression in the heart. *Genes Dev* **22**, 3242-3254 (2008).
115. Darnell, D.K., *et al.* MicroRNA expression during chick embryo development. *Dev Dyn* **235**, 3156-3165 (2006).
116. Chen, H., *et al.* Micro-RNA-195 and -451 regulate the LKB1/AMPK signaling axis by targeting MO25. *PLoS One* **7**, e41574 (2012).
117. Bronnum, H., *et al.* Islet-1 is a dual regulator of fibrogenic epithelial-to-mesenchymal transition in epicardial mesothelial cells. *Exp Cell Res* **319**, 424-435 (2013).
118. Wang, X., *et al.* MicroRNA-125b protects against myocardial ischaemia/reperfusion injury via targeting p53-mediated apoptotic signalling and TRAF6. *Cardiovasc Res* (2014).
119. Witman, N., Heigwer, J., Thaler, B., Lui, W.O. & Morrison, J.I. miR-128 regulates non-myocyte hyperplasia, deposition of extracellular matrix and Islet1 expression during newt cardiac regeneration. *Dev Biol* **383**, 253-263 (2013).
120. Brade, T., *et al.* Retinoic acid stimulates myocardial expansion by induction of hepatic erythropoietin which activates epicardial Igf2. *Development* **138**, 139-148 (2011).
121. Gambardella, S., *et al.* Overexpression of microRNA-206 in the skeletal muscle from myotonic dystrophy type 1 patients. *J Transl Med* **8**, 48 (2010).
122. Wang, X.H., *et al.* MicroRNA-320 expression in myocardial microvascular endothelial cells and its relationship with insulin-like growth factor-1 in type 2 diabetic rats. *Clin Exp Pharmacol Physiol* **36**, 181-188 (2009).
123. Huang, Z.P., *et al.* Loss of microRNAs in neural crest leads to cardiovascular syndromes resembling human congenital heart defects. *Arterioscler Thromb Vasc Biol* **30**, 2575-2586 (2010).
124. Cordes, K.R., *et al.* miR-145 and miR-143 regulate smooth muscle cell fate and plasticity. *Nature* **460**, 705-710 (2009).
125. Yeung, F., Chung, E., Guess, M.G., Bell, M.L. & Leinwand, L.A. Myh7b/miR-499 gene expression is transcriptionally regulated by MRFs and Eos. *Nucleic Acids Res* **40**, 7303-7318 (2012).
126. Dellefave, L. & McNally, E.M. Sarcomere mutations in cardiomyopathy, noncompaction, and the developing heart. *Circulation* **117**, 2847-2849 (2008).
127. Goljanek-Whysall, K., *et al.* MicroRNA regulation of the paired-box transcription factor Pax3 confers robustness to developmental timing of myogenesis. *Proc Natl Acad Sci U S A* **108**, 11936-11941 (2011).

128. Swingler, T.E., *et al.* The expression and function of microRNAs in chondrogenesis and osteoarthritis. *Arthritis Rheum* **64**, 1909-1919 (2012).
129. Xu, X., Roe, N.D., Weiser-Evans, M.C. & Ren, J. Inhibition of Mammalian Target of Rapamycin With Rapamycin Reverses Hypertrophic Cardiomyopathy in Mice With Cardiomyocyte-Specific Knockout of PTEN. *Hypertension* **63**, 729-739 (2014).
130. Bhattacharya, S., Macdonald, S.T. & Farthing, C.R. Molecular mechanisms controlling the coupled development of myocardium and coronary vasculature. *Clin Sci (Lond)* **111**, 35-46 (2006).
131. Lewis, B.P., Burge, C.B. & Bartel, D.P. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* **120**, 15-20 (2005).
132. Dweep, H., Sticht, C., Pandey, P. & Gretz, N. miRWalk--database: prediction of possible miRNA binding sites by "walking" the genes of three genomes. *J Biomed Inform* **44**, 839-847 (2011).
133. Poliseno, L., *et al.* A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. *Nature* **465**, 1033-1038 (2010).
134. Sucharov, C., Bristow, M.R. & Port, J.D. miRNA expression in the failing human heart: functional correlates. *J Mol Cell Cardiol* **45**, 185-192 (2008).
135. Wahlquist, C., *et al.* Inhibition of miR-25 improves cardiac contractility in the failing heart. *Nature* **508**, 531-535 (2014).
136. Fonoudi, H., *et al.* ISL1 protein transduction promotes cardiomyocyte differentiation from human embryonic stem cells. *PLoS One* **8**, e55577 (2013).
137. Eigenthaler, M., *et al.* Disruption of cardiac Ena-VASP protein localization in intercalated disks causes dilated cardiomyopathy. *Am J Physiol Heart Circ Physiol* **285**, H2471-2481 (2003).
138. Flicek, P., *et al.* Ensembl 2013. *Nucleic Acids Res* **41**, D48-55 (2013).
139. Chaudhry, A.Z., Lyons, G.E. & Gronostajski, R.M. Expression patterns of the four nuclear factor I genes during mouse embryogenesis indicate a potential role in development. *Dev Dyn* **208**, 313-325 (1997).
140. Glasco, D.M., *et al.* The mouse Wnt/PCP protein Vangl2 is necessary for migration of facial branchiomotor neurons, and functions independently of Dishevelled. *Dev Biol* **369**, 211-222 (2012).
141. Sinha, T., Wang, B., Evans, S., Wynshaw-Boris, A. & Wang, J. Disheveled mediated planar cell polarity signaling is required in the second heart field lineage for outflow tract morphogenesis. *Dev Biol* **370**, 135-144 (2012).
142. Pattison, J.S., Osinska, H. & Robbins, J. Atg7 induces basal autophagy and rescues autophagic deficiency in CryABR120G cardiomyocytes. *Circ Res* **109**, 151-160 (2011).
143. Schachterle, W., Rojas, A., Xu, S.M. & Black, B.L. ETS-dependent regulation of a distal Gata4 cardiac enhancer. *Dev Biol* **361**, 439-449 (2012).
144. Rehmsmeier, M., Steffen, P., Hochsmann, M. & Giegerich, R. Fast and effective prediction of microRNA/target duplexes. *RNA* **10**, 1507-1517 (2004).
145. Didiano, D. & Hobert, O. Perfect seed pairing is not a generally reliable predictor for miRNA-target interactions. *Nat Struct Mol Biol* **13**, 849-851 (2006).
146. Roel, G., Gent, Y.Y., Peterson-Maduro, J., Verbeek, F.J. & Destree, O. Lef1 plays a role in patterning the mesoderm and ectoderm in *Xenopus tropicalis*. *Int J Dev Biol* **53**, 81-89 (2009).
147. Bell, G.W., Yatskievych, T.A. & Antin, P.B. GEISHA, a whole-mount in situ hybridization gene expression screen in chicken embryos. *Dev Dyn* **229**, 677-687 (2004).
148. Schneider, C.A., Rasband, W.S. & Eliceiri, K.W. NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* **9**, 671-675 (2012).
149. Rog-Zielinska, E.A., *et al.* Glucocorticoid receptor is required for foetal heart maturation. *Hum Mol Genet* **22**, 3269-3282 (2013).
150. Panther, F., Williams, T. & Ritter, O. Inhibition of the calcineurin-NFAT signalling cascade in the treatment of heart failure. *Recent Pat Cardiovasc Drug Discov* **4**, 180-186 (2009).
151. Noveroske, J.K., *et al.* Quaking is essential for blood vessel development. *Genesis* **32**, 218-230 (2002).

152. Zhang, Y., *et al.* Rock2 controls TGFbeta signaling and inhibits mesoderm induction in zebrafish embryos. *J Cell Sci* **122**, 2197-2207 (2009).
153. Wang, Y., *et al.* ROCK isoform regulation of myosin phosphatase and contractility in vascular smooth muscle cells. *Circ Res* **104**, 531-540 (2009).
154. Zhao, Z. & Rivkees, S.A. Rho-associated kinases play a role in endocardial cell differentiation and migration. *Dev Biol* **275**, 183-191 (2004).
155. Zhu, N., *et al.* Endothelial enriched microRNAs regulate angiotensin II-induced endothelial inflammation and migration. *Atherosclerosis* **215**, 286-293 (2011).
156. Zhou, J., *et al.* Regulation of vascular smooth muscle cell turnover by endothelial cell-secreted microRNA-126: role of shear stress. *Circ Res* **113**, 40-51 (2013).
157. Lund, A.H. miR-10 in development and cancer. *Cell Death Differ* **17**, 209-214 (2010).
158. Li, N., Wei, C., Olena, A.F. & Patton, J.G. Regulation of endoderm formation and left-right asymmetry by miR-92 during early zebrafish development. *Development* **138**, 1817-1826 (2011).
159. Knobbe, C.B., Lapin, V., Suzuki, A. & Mak, T.W. The roles of PTEN in development, physiology and tumorigenesis in mouse models: a tissue-by-tissue survey. *Oncogene* **27**, 5398-5415 (2008).
160. Inui, M., Martello, G. & Piccolo, S. MicroRNA control of signal transduction. *Nat Rev Mol Cell Biol* **11**, 252-263 (2010).
161. Gaarenstroom, T. & Hill, C.S. TGF-beta signaling to chromatin: How Smads regulate transcription during self-renewal and differentiation. *Semin Cell Dev Biol* (2014).
162. Jovicic, A., *et al.* Comprehensive expression analyses of neural cell-type-specific miRNAs identify new determinants of the specification and maintenance of neuronal phenotypes. *J Neurosci* **33**, 5127-5137 (2013).
163. Sokol, N.S. Small temporal RNAs in animal development. *Curr Opin Genet Dev* **22**, 368-373 (2012).
164. Parsons, X.H. MicroRNA Profiling Reveals Distinct Mechanisms Governing Cardiac and Neural Lineage-Specification of Pluripotent Human Embryonic Stem Cells. *J Stem Cell Res Ther* **2**(2012).
165. Lu, H., *et al.* Wnt-promoted Isl1 expression through a novel TCF/LEF1 binding site and H3K9 acetylation in early stages of cardiomyocyte differentiation of P19CL6 cells. *Mol Cell Biochem* (2014).
166. Lee, E., *et al.* Autophagy is essential for cardiac morphogenesis during vertebrate development. *Autophagy* **10**, 572-587 (2014).
167. Kumawat, K., *et al.* Noncanonical WNT-5A signaling regulates TGF-beta-induced extracellular matrix production by airway smooth muscle cells. *FASEB J* **27**, 1631-1643 (2013).
168. Kim, Y., *et al.* Calcineurin dephosphorylates glycogen synthase kinase-3 beta at serine-9 in neuroblast-derived cells. *J Neurochem* **111**, 344-354 (2009).
169. Marx, S.O., *et al.* PKA phosphorylation dissociates FKBP12.6 from the calcium release channel (ryanodine receptor): defective regulation in failing hearts. *Cell* **101**, 365-376 (2000).
170. Kozomara, A. & Griffiths-Jones, S. miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res* **42**, D68-73 (2014).
171. Griffiths-Jones, S. The microRNA Registry. *Nucleic Acids Res* **32**, D109-111 (2004).
172. Griffiths-Jones, S., Grocock, R.J., van Dongen, S., Bateman, A. & Enright, A.J. miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Res* **34**, D140-144 (2006).
173. Griffiths-Jones, S., Saini, H.K., van Dongen, S. & Enright, A.J. miRBase: tools for microRNA genomics. *Nucleic Acids Res* **36**, D154-158 (2008).
174. Cutting, A.D., *et al.* The potential role of microRNAs in regulating gonadal sex differentiation in the chicken embryo. *Chromosome Res* **20**, 201-213 (2012).
175. Hicks, J.A., Trakooljul, N. & Liu, H.C. Discovery of chicken microRNAs associated with lipogenesis and cell proliferation. *Physiol Genomics* **41**, 185-193 (2010).
176. Glazov, E.A., *et al.* A microRNA catalog of the developing chicken embryo identified by a deep sequencing approach. *Genome Res* **18**, 957-964 (2008).

177. Kasprzyk, A. BioMart: driving a paradigm change in biological data management. *Database (Oxford)* **2011**, bar049 (2011).
178. Heanue, T.A., *et al.* Synergistic regulation of vertebrate muscle development by Dach2, Eya2, and Six1, homologs of genes required for Drosophila eye formation. *Genes Dev* **13**, 3231-3243 (1999).
179. Li, Z., *et al.* A microRNA signature for a BMP2-induced osteoblast lineage commitment program. *Proc Natl Acad Sci U S A* **105**, 13906-13911 (2008).
180. Yang, J., Zhao, H., Xin, Y. & Fan, L. MicroRNA-198 inhibits proliferation and induces apoptosis of lung cancer cells via targeting FGFR1. *J Cell Biochem* **115**, 987-995 (2014).
181. Ito, T., *et al.* Polo-like kinase 1 regulates cell proliferation and is targeted by miR-593\* in esophageal cancer. *Int J Cancer* **129**, 2134-2146 (2011).
182. Duan, L., Hao, X., Liu, Z., Zhang, Y. & Zhang, G. MiR-129-5p is down-regulated and involved in the growth, apoptosis and migration of medullary thyroid carcinoma cells through targeting RET. *FEBS Lett* (2014).
183. Kwon, C., Han, Z., Olson, E.N. & Srivastava, D. MicroRNA1 influences cardiac differentiation in Drosophila and regulates Notch signaling. *Proc Natl Acad Sci U S A* **102**, 18986-18991 (2005).
184. Sokol, N.S. & Ambros, V. Mesodermally expressed Drosophila microRNA-1 is regulated by Twist and is required in muscles during larval growth. *Genes Dev* **19**, 2343-2354 (2005).
185. Yang, B., *et al.* The muscle-specific microRNA miR-1 regulates cardiac arrhythmogenic potential by targeting GJA1 and KCNJ2. *Nat Med* **13**, 486-491 (2007).
186. Sweetman, D., *et al.* Specific requirements of MRFs for the expression of muscle specific microRNAs, miR-1, miR-206 and miR-133. *Dev Biol* **321**, 491-499 (2008).
187. McFadden, D.G., *et al.* The Hand1 and Hand2 transcription factors regulate expansion of the embryonic cardiac ventricles in a gene dosage-dependent manner. *Development* **132**, 189-201 (2005).
188. Xiao, J., *et al.* MicroRNA miR-133 represses HERG K<sup>+</sup> channel expression contributing to QT prolongation in diabetic hearts. *J Biol Chem* **282**, 12363-12367 (2007).
189. Schumann, H., Holtz, J., Zerkowski, H.R. & Hatzfeld, M. Expression of secreted frizzled related proteins 3 and 4 in human ventricular myocardium correlates with apoptosis related gene expression. *Cardiovasc Res* **45**, 720-728 (2000).
190. Abu-Elmagd, M., Garcia-Morales, C. & Wheeler, G.N. Frizzled7 mediates canonical Wnt signaling in neural crest induction. *Dev Biol* **298**, 285-298 (2006).
191. Dolle, P., Fraulob, V., Gallego-Llamas, J., Vermot, J. & Niederreither, K. Fate of retinoic acid-activated embryonic cell lineages. *Dev Dyn* **239**, 3260-3274 (2010).
192. Lin, S.C., *et al.* Endogenous retinoic acid regulates cardiac progenitor differentiation. *Proc Natl Acad Sci U S A* **107**, 9234-9239 (2010).
193. Osmond, M.K., Butler, A.J., Voon, F.C. & Bellairs, R. The effects of retinoic acid on heart formation in the early chick embryo. *Development* **113**, 1405-1417 (1991).
194. Krutzfeldt, J., *et al.* Silencing of microRNAs in vivo with 'antagomirs'. *Nature* **438**, 685-689 (2005).
195. Meister, G., Landthaler, M., Dorsett, Y. & Tuschl, T. Sequence-specific inhibition of microRNA- and siRNA-induced RNA silencing. *RNA* **10**, 544-550 (2004).
196. McGlinn, E., *et al.* In ovo application of antagomiRs indicates a role for miR-196 in patterning the chick axial skeleton through Hox gene regulation. *Proc Natl Acad Sci U S A* **106**, 18610-18615 (2009).
197. Sweetman, D., *et al.* FGF-4 signaling is involved in mir-206 expression in developing somites of chicken embryos. *Dev Dyn* **235**, 2185-2191 (2006).
198. von Gise, A., *et al.* WT1 regulates epicardial epithelial to mesenchymal transition through beta-catenin and retinoic acid signaling pathways. *Dev Biol* **356**, 421-431 (2011).
199. Lu, T.Y., *et al.* Overexpression of microRNA-1 promotes cardiomyocyte commitment from human cardiovascular progenitors via suppressing WNT and FGF signaling pathways. *J Mol Cell Cardiol* **63**, 146-154 (2013).
200. von Gise, A. & Pu, W.T. Endocardial and epicardial epithelial to mesenchymal transitions in heart development and disease. *Circ Res* **110**, 1628-1645 (2012).
201. Hardy, K.M., Yatskevych, T.A., Konieczka, J., Bobbs, A.S. & Antin, P.B. FGF signalling through RAS/MAPK and PI3K pathways regulates cell movement and gene expression in

- the chicken primitive streak without affecting E-cadherin expression. *BMC Dev Biol* **11**, 20 (2011).
202. Marin, F. & Nieto, M.A. Expression of chicken slug and snail in mesenchymal components of the developing central nervous system. *Dev Dyn* **230**, 144-148 (2004).
203. De Craene, B., van Roy, F. & Berx, G. Unraveling signalling cascades for the Snail family of transcription factors. *Cell Signal* **17**, 535-547 (2005).
204. Srivastava, D., Cserjesi, P. & Olson, E.N. A subclass of bHLH proteins required for cardiac morphogenesis. *Science* **270**, 1995-1999 (1995).

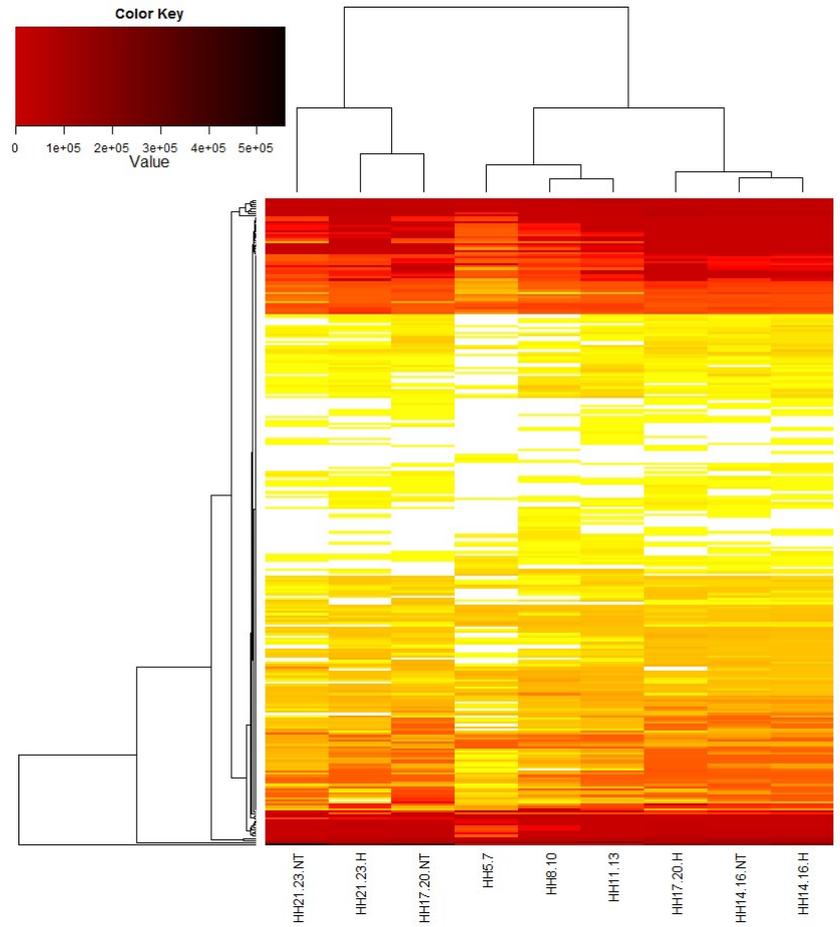
## Appendix A

**Table Apx A.1: Animal miRNAs sequenced from chicken tissue but not yet formally annotated as chicken miRNAs in miRBase.**

miRNA	miRNA	miRNA	miRNA
miR-7	miR-210	miR-2478	miR-3618
miR-72	miR-338	miR-2976	miR-3620
miR-96	miR-363	miR-2984	miR-4454
miR-141	miR-716	miR-2985	miR-5100
miR-143	miR-1277	miR-3074	miR-6240
miR-145	miR-1386	miR-3120	miR-7565
miR-182	miR-1388	miR-3591	miR-7641
miR-192	miR-2424	miR-3604	

**Table Apx A.2: miRNAs detected in a single sample only.**

miRNA	Sample	Raw reads	miRNA	Sample	Raw reads	miRNA	Sample	Raw reads
miR-6545	8-10	3	miR-7476	14-16 NT	1	miR-1579	17-20 NT	1
miR-6603	8-10	3	miR-1603	17-20 H	2	miR-1601	17-20 NT	1
miR-6576	8-10	1	miR-1688	17-20 H	2	miR-6647	17-20 NT	1
miR-1567	11-13	1	miR-3620	17-20 H	2	miR-6712	17-20 NT	1
miR-6581	11-13	1	miR-1277	17-20 H	1	miR-6594	21-23 H	2
miR-1554	14-16 H	1	miR-1605	17-20 H	1	miR-7473	21-23 H	2
miR-1634	14-16 H	1	miR-1642	17-20 H	1	miR-1735	21-23 H	1
miR-1713	14-16 H	1	miR-1648	17-20 H	1	miR-1759	21-23 H	1
miR-1741	14-16 H	1	miR-1742	17-20 H	1	miR-6640	21-23 H	1
miR-7456	14-16 H	1	miR-1760	17-20 H	1	miR-1754	21-23 NT	3
miR-1683	14-16 NT	5	miR-1813	17-20 H	1	miR-1691	21-23 NT	2
miR-1768	14-16 NT	3	miR-3591	17-20 H	1	miR-3074	21-23 NT	2
miR-6629	14-16 NT	3	miR-6567	17-20 H	1	miR-6708	21-23 NT	2
miR-1574	14-16 NT	1	miR-6571	17-20 H	1	miR-7565	21-23 NT	2
miR-1588	14-16 NT	1	miR-6650	17-20 H	1	miR-1582	21-23 NT	1
miR-1675	14-16 NT	1	miR-6706	17-20 H	1	miR-1631	21-23 NT	1
miR-1718	14-16 NT	1	miR-6695	17-20 NT	4	miR-1637	21-23 NT	1
miR-3618	14-16 NT	1	miR-6600	17-20 NT	3	miR-1656	21-23 NT	1
miR-6672	14-16 NT	1	miR-1592	17-20 NT	2	miR-1762	21-23 NT	1
miR-6681	14-16 NT	1	miR-2984	17-20 NT	2	miR-2976	21-23 NT	1
miR-6685	14-16 NT	1	miR-1564	17-20 NT	1	miR-2985	21-23 NT	1



**Figure Apx A.1: Correlation of samples by known animal miRNA normalised reads.** A custom R script was used to generate a correlation heatmap with manually defined colour breaks. A five colour transition was used, ranging from white to black through yellow, orange and red. White: 0 to 0.001 reads per million. Yellow: 0.001 to 5 reads per million. Orange: 5 to 50 reads per million. Red: 50 to 1,000 reads per million. Black: 1,000 to 600,000 reads per million.

**Table Apx A.3: Overlapping predicted targets of miR-26, miR-30, miR-92 and miR-10 that are known to be expressed in the human heart.** This table contains the first 46 of 92 overlapping genes. **Table A4** contains genes 47 to 92.

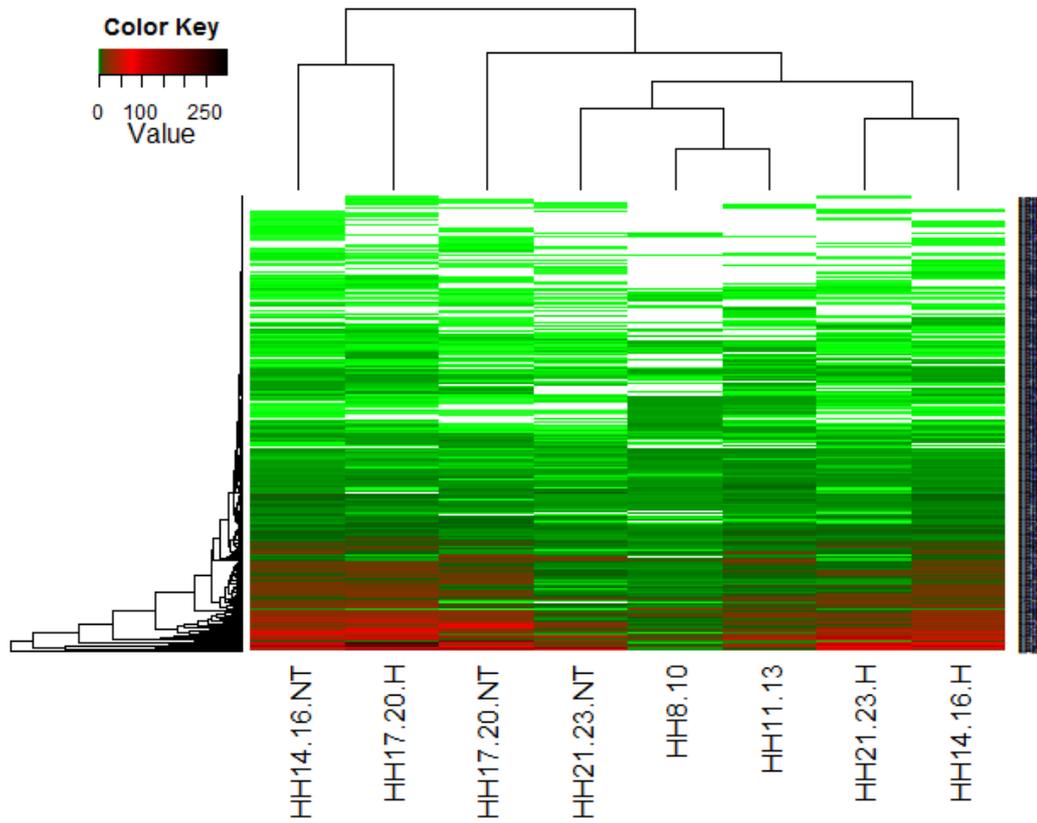
Gene Symbol	Overlapping predicted targets of				
	miR-26 & miR-30	miR-26 & miR-92	miR-30 & miR-92	miR-26 & miR-30 & miR-92	miR-10
ABL2	Y	Y	Y	Y	
BCL11A	Y	Y	Y	Y	
CACNA1C	Y	Y	Y	Y	
CEP350	Y	Y	Y	Y	
DENND1B	Y	Y	Y	Y	
DLG5	Y	Y	Y	Y	
FBXO28	Y	Y	Y	Y	
GALNT7	Y	Y	Y	Y	
HIPK1	Y	Y	Y	Y	
MTDH	Y	Y	Y	Y	
NOVA1	Y	Y	Y	Y	
TMCC1	Y	Y	Y	Y	
ZCCHC2	Y	Y	Y	Y	
ZFC3H1	Y	Y	Y	Y	
BRWD3	Y				
CACNB2	Y				
GDNF	Y				
KPNA6	Y				
LARP1	Y				
LIN28B	Y				
LOX	Y				
METAP2	Y				
PDCD10	Y				
PDE4D	Y				
PITX1	Y				
PLXNA2	Y				
RAP2C	Y				
SLC38A1	Y				
SLC38A2	Y				
SMAD1	Y				
SMARCA5	Y				
TNPO1	Y				
USP15	Y				
ZSWIM5	Y				
ARMC1		Y			
CASZ1		Y			
EIF4G2		Y			
MAN2A1		Y			
MITF		Y			
NAA15		Y			
PTEN		Y			
PTPRD		Y			
RBM47		Y			
RYR3		Y			
SSFA2		Y			
TET2		Y			Y

**Table Apx A.4: Overlapping predicted targets of miR-26, miR-30, miR-92 and miR-10 that are known to be expressed in the human heart.** This table contains genes 47 – 92; see **Table A3** for the first 46 overlapping genes.

Gene Symbol	Overlapping predicted targets of				
	miR-26 & miR-30	miR-26 & miR-92	miR-30 & miR-92	miR-26 & miR-30 & miR-92	miR-10
ZFH3		Y			
ZNF385B		Y			
ACTC1			Y		
ADRB1			Y		
ARID5B			Y		
ATP2A2			Y		
ATXN1			Y		
AZIN1			Y		
C20orf112			Y		
CFL2			Y		
CHD9			Y		
DCP1A			Y		
EIF2C3			Y		
EVI5			Y		
FAM160B1			Y		
FXR1			Y		
GOLGA1			Y		
GOLGA4			Y		
HECW1			Y		
HOXA11			Y		
LIMCH1			Y		
MIA3			Y		
MMD			Y		
MRPS25			Y		
NEDD4L			Y		
NUFIP2			Y		
PCDH10			Y		
PI4K2B			Y		
PPP1R12A			Y		
RAB23			Y		
RNF44			Y		
SETD5			Y		
SOBP			Y		Y
SSBP3			Y		
TMEM87A			Y		
TSC1			Y		
ZBTB34			Y		
ZNF521			Y		
BDNF					Y
CSMD1					Y
ESRRG					Y
GATA6					Y
GTF2H1					Y
KLF11					Y
NCOR2					Y
WDR26					Y

**Table Apx A.5: Diseases significantly associated with predicted overlapping target genes of miR-26, miR-30, and miR-92 that are also known to be expressed in the heart.** The number of genes found associated with each disease is shown in brackets after the disease name.

<b>Disease (no. genes)</b>	<b>P value</b>	<b>Disease (no. genes)</b>	<b>P value</b>
Autistic disorder (4)	<0.0001	Chronic obstructive airway disease (2)	0.014
Subarachnoid haemorrhage (2)	0.0003	<u>Heart failure (2)</u>	0.017
Huntington disease (2)	0.001	Dermatitis (2)	0.030
<u>Atherosclerosis (4)</u>	0.001	Leukemia (3)	0.034
Pancreas cancer (3)	0.001	Stomach cancer (2)	0.040
Drug abuse (3)	0.002	Neoplasm metastasis (2)	0.045
Retinal disease (2)	0.005	Diabetes mellitus (3)	0.048
Amyotrophic lateral sclerosis (2)	0.006		



**Figure Apx A.2: Correlation of samples by known animal miRNA fold change over HH5-7.** Only miRNAs detected in more than one sample were considered. A custom R script was used to construct a correlation heatmap with manually defined colour breaks. Correlations were done both by miRNA and by sample. A five colour transition was employed, ranging from white to black through green and red. The colour break ranges were defined by size of fold change as shown. White: 0, light green: 0.0077 to 1, dark green: 1 to 10, red: 10 to 100, black: 100 to 300.

**Table Apx A.6: miRNAs found to have fold change levels greater than 10 fold higher after subtracting miRNA reads detected in NT samples from that detected in Heart samples.** When fold change levels (determined by comparison with the HH5-7 sample) in NT samples from the HH14-16, HH17-20 and HH21-23 time points were subtracted from the values in the corresponding Heart samples, 38 miRNAs with levels more than 10 fold higher in Heart samples were identified. Some of these miRNAs, indicated in yellow, are known to have roles in heart development and function. miRNAs that are conserved in humans are indicated in bold and underlined.

miRNA	Fold increase in Heart samples		
	HH14-16	HH17-20	HH21-23
<b><u>mir-1</u></b>		26	15
<b><u>mir-10</u></b>	12		
<b><u>mir-100</u></b>			31
<b><u>mir-122</u></b>		11	21
<b><u>mir-125</u></b>		16	21
<b><u>mir-126</u></b>		125	40
<b><u>mir-133</u></b>		88	60
<b><u>mir-141</u></b>		27	
<b><u>mir-142</u></b>		50	28
<b><u>mir-143</u></b>		40	40
<b><u>mir-144</u></b>		22	12
<b><u>mir-145</u></b>	21	34	106
mir-1451	11	38	20
<b><u>mir-147</u></b>		18	
<b><u>mir-148</u></b>		20	
mir-1737		13	
mir-1773		37	29
mir-1788		15	
mir-1805		17	
<b><u>mir-187</u></b>		97	68
<b><u>mir-199</u></b>		16	
<b><u>mir-203</u></b>		22	
mir-2188		13	11
<b><u>mir-219</u></b>		18	
<b><u>mir-221</u></b>		14	
<b><u>mir-223</u></b>		25	
<b><u>mir-24</u></b>		27	
<b><u>mir-27</u></b>		14	11
<b><u>mir-301</u></b>		15	
<b><u>mir-375</u></b>		18	19
<b><u>mir-429</u></b>		11	
<b><u>mir-451</u></b>		109	46
<b><u>mir-454</u></b>		22	
<b><u>mir-490</u></b>		153	98
<b><u>mir-499</u></b>		226	39
mir-6631		11	16
mir-6694		89	36
<b><u>mir-99</u></b>		37	

**Table Apx A.7: miRNAs found to have significantly similar patterns of expression by fold change across the HH14-16, HH17-20 and HH21-23 time points, after subtracting miRNA reads detected in NT samples from that detected in H samples.  $0.05 \geq P > 0.01$  : \*;  $0.01 \geq P > 0.001$  : \*\*;  $0.001 \geq P > 0.0001$  : \*\*\*;  $P \leq 0.0001$  : \*\*\*\*. Heart miRNAs are shown in yellow. miRNAs conserved in *H. sapiens* are indicated in bold and underlined.**

Correlated to miR-1		Correlated to miR-101		Correlated to miR-130	
<b>miR-143</b>	*	miR-6582	****	<b>miR-33</b>	****
<b>miR-19</b>	*	miR-6689	****	miR-34	*
<b>miR-222</b>	*	<b>miR-125</b>		<b>miR-449</b>	*
<b>miR-363</b>	*	<b>miR-181</b>	*	miR-6549	****
miR-1619	*	miR-6631	*	miR-6555	****
miR-1651	*	<b>miR-126</b>		miR-6614	****
miR-1662	*	<b>miR-16</b>	*	miR-6630	****
miR-1663	*	<b>miR-221</b>	*	<b>miR-133</b>	
miR-1695	*	<b>miR-106</b>	**	<b>miR-101</b>	*
miR-1745	*	<b>miR-124</b>	*	<b>miR-21</b>	**
miR-1746	*	<b>miR-137</b>	*	miR-1388	*
miR-1751	*	miR-1451	**	<b>miR-144</b>	*
miR-1756	*	miR-1680	**	miR-1458	*
miR-1773	*	miR-1781	**	miR-1620	*
miR-1776	*	miR-1805	**	miR-1779	*
miR-1800	*	<b>miR-203</b>	**	<b>miR-187</b>	*
<b>miR-183</b>	*	<b>miR-215</b>	**	<b>miR-193</b>	*
miR-2188	*	<b>miR-217</b>	*	<b>miR-200</b>	**
<b>miR-22</b>	*	miR-3532	**	<b>miR-205</b>	**
<b>miR-338</b>	*	<b>miR-455</b>	*	miR-2129	*
miR-3525	*	<b>miR-128</b>		<b>miR-216</b>	*
<b>miR-489</b>	*	<b>miR-1306</b>	****	<b>miR-27</b>	*
miR-6550	*	miR-1456	****	<b>miR-3120</b>	*
miR-6561	*	miR-1668	****	<b>miR-3529</b>	**
miR-6604	*	miR-1689	****	<b>miR-490</b>	*
miR-6648	*	<b>miR-29</b>	****	miR-6582	*
miR-6660	*	<b>miR-365</b>	****	miR-6689	*
miR-6662	*	miR-5109	****	<b>miR-138</b>	
miR-72	*	miR-6240	****	miR-1454	****
<b>miR-10</b>		miR-6552	****	miR-1788	*
<b>miR-375</b>	*	miR-6586	****	miR-34	*
<b>miR-101</b>		miR-6612	****	<b>miR-449</b>	*
<b>miR-133</b>	*	miR-6709	****	<b>miR-143</b>	
miR-1388	****	miR-7	****	<b>miR-1</b>	*
<b>miR-142</b>	*	<b>miR-99</b>	*	<b>miR-19</b>	****
miR-1458	****	<b>miR-130</b>		<b>miR-222</b>	****
miR-1620	****	miR-1454	****	<b>miR-363</b>	****
miR-1731	*	miR-1650	****	miR-1619	****
miR-1779	****	miR-1690	****	miR-1663	****
<b>miR-187</b>	*	miR-1728	****	miR-1695	****
miR-2129	****	miR-1758	****	miR-1745	****
miR-456	*	miR-1788	*	miR-1746	****
miR-458	*	<b>miR-182</b>	****	miR-1751	****
<b>miR-490</b>	**	miR-2130	****	miR-1756	****

**Table Apx A.8: miRNAs found to have significantly similar patterns of expression by fold change across the HH14-16, HH17-20 and HH21-23 time points, after subtracting miRNA reads detected in NT samples from that detected in H samples.  $0.05 \geq P > 0.01$  : \*;  $0.01 \geq P > 0.001$  : \*\*;  $0.001 \geq P > 0.0001$  : \*\*\*;  $P \leq 0.0001$  : \*\*\*\*. Heart miRNAs are shown in yellow. miRNAs conserved in *H. sapiens* are indicated in bold and underlined.**

Correlated to Significance		Correlated to Significance		Correlated to Significance	
<b>miR-143</b>		<b>miR-18</b>		<b>miR-21</b>	
<u>miR-338</u>	****	<u>miR-182</u>	****	<u>miR-187</u>	*
<u>miR-375</u>	*	miR-2130	****	<u>miR-193</u>	*
<u>miR-489</u>	****	<u>miR-33</u>	****	<u>miR-216</u>	*
miR-6604	****	miR-34	*	<u>miR-27</u>	*
miR-6660	****	<u>miR-449</u>	*	<u>miR-3529</u>	****
miR-6662	****	miR-6549	****	<b>miR-221</b>	
<b>miR-145</b>		miR-6555	****	<u>miR-126</u>	*
<u>miR-100</u>	*	miR-6614	****	<u>miR-106</u>	*
<b>miR-15</b>		miR-6630	****	<u>miR-124</u>	**
miR-1597	*	<b>miR-181</b>		miR-1451	*
miR-1635	****	<u>miR-125</u>	*	miR-1680	*
<u>miR-199</u>	*	<u>miR-24</u>	*	miR-1781	*
<b>miR-16</b>		miR-1397	*	miR-1805	*
<u>miR-126</u>	*	<u>miR-140</u>	*	<u>miR-203</u>	*
<u>miR-106</u>	*	miR-1717	*	<u>miR-215</u>	*
<u>miR-135</u>	*	miR-6631	**	<u>miR-219</u>	*
<u>miR-137</u>	****	<b>miR-19</b>		miR-3532	*
miR-1451	*	<u>miR-1</u>	*	<u>miR-455</u>	**
miR-1559	*	<u>miR-143</u>	****	miR-9	*
miR-1680	*	<u>miR-363</u>	****	<b>miR-222</b>	
miR-1724	*	miR-1800	****	<u>miR-1</u>	*
miR-1781	*	<u>miR-338</u>	****	<u>miR-143</u>	****
miR-1805	*	miR-3525	****	<u>miR-363</u>	****
<u>miR-203</u>	*	<u>miR-375</u>	*	miR-1800	****
<u>miR-215</u>	*	miR-6550	****	<u>miR-338</u>	****
miR-3532	*	miR-6648	****	miR-3525	****
<u>miR-451</u>	*	<b>miR-20</b>		<u>miR-375</u>	*
miR-6694	*	miR-1650	****	miR-6550	****
<b>miR-17</b>		miR-1690	****	miR-6648	****
<u>miR-499</u>	*	miR-1728	****	<b>miR-24</b>	
<u>miR-148</u>	*	miR-1758	****	<u>miR-181</u>	*
<u>miR-153</u>	*	miR-1788	*	<u>miR-122</u>	*
miR-1720	**	<u>miR-182</u>	****	miR-1397	**
<u>miR-219</u>	*	miR-2130	****	<u>miR-140</u>	**
<u>miR-301</u>	*	<u>miR-33</u>	****	miR-1717	**
miR-9	*	miR-34	*	miR-1753	*
<u>miR-96</u>	*	<u>miR-449</u>	*	miR-6631	*
<b>miR-18</b>		miR-6549	****	<b>miR-30</b>	
miR-1454	****	miR-6555	****	miR-1625	*
miR-1650	****	miR-6614	****	<b>miR-31</b>	
miR-1690	****	miR-6630	****	miR-1773	*
miR-1728	****	<b>miR-21</b>		<u>miR-193</u>	*
miR-1758	****	<u>miR-133</u>	**	<u>miR-216</u>	*
miR-1788	*	<u>miR-144</u>	**	miR-2188	*

**Table Apx A.9: miRNAs found to have significantly similar patterns of expression by fold change across the HH14-16, HH17-20 and HH21-23 time points, after subtracting miRNA reads detected in NT samples from that detected in H samples.  $0.05 \geq P > 0.01$  : \*;  $0.01 \geq P > 0.001$  : \*\*;  $0.001 \geq P > 0.0001$  : \*\*\*;  $P \leq 0.0001$  : \*\*\*\*. Heart miRNAs are shown in yellow. miRNAs conserved in *H. sapiens* are indicated in bold and underlined.**

Correlated to miR-31 Significance		Correlated to miR-363 Significance	
<b>miR-31</b>		<b>miR-363</b>	
<u>miR-27</u>	*	<u>miR-22</u>	****
miR-72	*	miR-3525	****
<b>miR-363</b>		<u>miR-375</u>	*
<u>miR-1</u>	*	<u>miR-489</u>	****
<u>miR-143</u>	****	miR-6550	****
<u>miR-19</u>	****	miR-6561	****
<u>miR-222</u>	****	miR-6604	****
miR-1619	****	miR-6648	****
miR-1651	****	miR-6660	****
miR-1662	****	miR-6662	****
miR-1663	****	<b>miR-499</b>	
miR-1695	****	<u>miR-17</u>	*
miR-1745	****	<u>miR-124</u>	*
miR-1746	****	<u>miR-148</u>	*
miR-1751	****	miR-1720	*
miR-1756	****	miR-1729	*
miR-1776	****	<u>miR-219</u>	*
miR-1800	****	<u>miR-455</u>	*
<u>miR-183</u>	****	miR-9	**

**Table Apx A.10: Conserved miRNAs found to correlate significantly to 11 known heart miRNAs.** 68 miRNAs conserved in humans were found to correlate significantly by fold change pattern to 11 selected heart miRNAs. This table shows correlating miRNAs 1 to 34.

Correlating miRNA	Heart miRNAs							
	mir-125	mir-181						
let-7	mir-125	mir-181						
mir-100	mir-125	mir-143	mir-181					
mir-103	mir-125	mir-126	mir-133	mir-138	mir-143	mir-181	miR-363	miR-499
mir-106	mir-101	mir-126	mir-133	mir-138	miR-363	miR-499		
mir-107	mir-125	mir-126	mir-133	mir-138	mir-143	mir-181	miR-363	miR-499
mir-122	mir-133	mir-143	mir-145					
mir-141	mir-101	mir-126	mir-133	mir-138	mir-143	miR-363	miR-499	
mir-142	mir-101	mir-125	mir-126	mir-138	mir-143	mir-181	miR-363	miR-499
mir-144	mir-101	mir-126	mir-133	mir-138	mir-143	miR-363		
mir-146	mir-128							
mir-147	mir-101	mir-126	mir-138	mir-143	miR-363	miR-499		
mir-148	mir-101	mir-126	mir-138	mir-143	miR-363	miR-499		
mir-15	mir-125	mir-181						
mir-155	mir-128	mir-133						
mir-16	mir-101	mir-126	mir-138	miR-363	miR-499			
mir-17	mir-101	mir-126	mir-133	mir-138	miR-363	miR-499		
mir-18	mir-101	mir-128	mir-138					
mir-183	mir-125	mir-126	mir-138	mir-143	mir-181	miR-363	miR-499	
mir-184	mir-125	mir-145	mir-181					
mir-187	mir-125	mir-143	mir-145	mir-181				
mir-19	mir-101	mir-126	mir-138	miR-363				
mir-190	mir-101	mir-126	mir-138	miR-363				
mir-192	mir-138	miR-363						
mir-193	mir-125	mir-133	mir-143	mir-145	mir-181	miR-499		
mir-194	mir-125	mir-126	mir-143	mir-181	miR-363	miR-499		
mir-196	mir-133	miR-499						
mir-199	mir-125	mir-181	miR-363					
mir-20	mir-101	mir-126	mir-133	mir-138	miR-363	miR-499		
mir-200	mir-126	mir-128	mir-133	mir-138	miR-499			
mir-203	mir-101	mir-126	mir-133	mir-138	mir-143	miR-363	miR-499	
mir-204	mir-101	mir-125	mir-126	mir-138	mir-143	mir-181	miR-363	
mir-205	mir-125	mir-145	mir-181					
mir-21	mir-101	mir-126	mir-138	miR-363	miR-499			
mir-210	mir-130							

**Table Apx A.11: Conserved miRNAs found to correlate significantly to 11 known heart miRNAs.** 68 miRNAs conserved in humans were found to correlate significantly by fold change pattern to 11 selected heart miRNAs. This table shows correlating miRNAs 34 to 68.

Correlating miRNA	Heart miRNAs							
mir-211	mir-101	mir-125	mir-126	mir-138	mir-143	mir-181	miR-363	
mir-214	mir-125	mir-181						
mir-215	mir-101	mir-126	mir-138	mir-143	mir-181	miR-363	miR-499	
mir-218	mir-101	mir-138	miR-363					
mir-219	mir-101	mir-138	miR-363					
mir-22	mir-126	mir-128	mir-133					
mir-221	mir-126	mir-133	mir-138	mir-143	miR-499			
mir-223	mir-125	mir-126	mir-133	mir-138	mir-143	mir-181	miR-363	miR-499
mir-23	mir-145							
mir-24	mir-101	mir-126	mir-128	mir-133	mir-138	mir-143	miR-363	
mir-26	mir-101	mir-126	mir-133	mir-138	mir-143	miR-363		
mir-27	mir-101	mir-126	mir-138	mir-143	mir-181	miR-363	miR-499	
mir-29	mir-101	mir-126	mir-138	miR-363				
mir-30	mir-101	mir-126	mir-133	mir-138	mir-143	miR-363	miR-499	
mir-301	mir-101	mir-126	mir-138	miR-363	miR-499			
mir-302	mir-125	mir-130						
mir-3064	mir-101	mir-126	mir-128					
mir-31	mir-128	mir-130	mir-133					
mir-32	mir-101	mir-138	miR-363					
mir-33	mir-101	mir-138	miR-363					
mir-338	mir-101	mir-126	mir-138	miR-363				
mir-3529	mir-101	mir-126	mir-133	mir-138	miR-363	miR-499		
mir-3607	mir-125							
mir-367	mir-128							
mir-375	mir-133	mir-143	mir-145	miR-499				
mir-429	mir-101	mir-126	mir-133	mir-138	mir-143	miR-363	miR-499	
mir-451	mir-125	mir-126	mir-138	mir-143	mir-181	miR-363	miR-499	
mir-455	mir-101	mir-126	mir-133	mir-138	mir-143	miR-363	miR-499	
mir-490	mir-125	mir-126	mir-133	mir-143	mir-181	miR-499		
mir-5100	mir-128							
mir-551	mir-101	mir-126	mir-138	mir-143	mir-181	miR-363	miR-499	
mir-92	mir-125	mir-181						
mir-96	mir-101	mir-138	miR-363					
mir-99	mir-125	mir-181						

## Appendix B

### B.1 Oligonucleotide sequences

#### B.1.1 Illumina adapter sequences

RNA 3' adapter: 5'TGGAATTCTCGGGTGCCAAGG

Library	RNA 5' barcoded adapter
HH5-7	5'CAAGCAGAAGACGGCATAACGAGATCGTGATGTGACTGGAGTTCCTTGGCA CCCGAGAATTCCA
HH8-10	5'CAAGCAGAAGACGGCATAACGAGATACATCGGTGACTGGAGTTCCTTGGCA CCCGAGAATTCCA
HH11-13	5'CAAGCAGAAGACGGCATAACGAGATGCCTAAGTGACTGGAGTTCCTTGGCA CCCGAGAATTCCA
HH14-16 H	5'CAAGCAGAAGACGGCATAACGAGATTGGTCAGTGACTGGAGTTCCTTGGCA CCCGAGAATTCCA
HH14-16 NT	5'CAAGCAGAAGACGGCATAACGAGATCACTGTGTGACTGGAGTTCCTTGGCA CCCGAGAATTCCA
HH17-20 H	5'CAAGCAGAAGACGGCATAACGAGATATTGGCGTGACTGGAGTTCCTTGGCA CCCGAGAATTCCA
HH17-20 NT	5'CAAGCAGAAGACGGCATAACGAGATGATCTGGTGACTGGAGTTCCTTGGCA CCCGAGAATTCCA
HH21-23 H	5'CAAGCAGAAGACGGCATAACGAGATTCAAGTGTGACTGGAGTTCCTTGGCA CCCGAGAATTCCA
HH21-23 NT	5'CAAGCAGAAGACGGCATAACGAGATCTGATCGTGACTGGAGTTCCTTGGCA CCCGAGAATTCCA

## B.2 AntagomiR sequences

### AntagomiR-1

5' .F1.mU.\*mA.\*mC.mA.mU.mA.mC.mU.mU.mC.mU.mU.mU.mA.mC.mA.mU.mU  
 .\*mC.\*mC.\*mA.3'-Ch13'

### AntagomiR-206

5' .F1.mC.\*mC.\*mA.mC.mA.mC.mA.mC.mU.mU.mC.mC.mU.mU.mA.mC.mA.mU  
 .mU.\*mC.\*mC.\*mA.3'-Ch13'

### Scrambled-206

F1.mC\*.mA\*mU.mC.mC.mA.mU.mC.mA.mC.mU.mC.mA.mC.mU.mC.mC.mA.mU.  
 \*mC.\*mA.\*mU.3'-Ch1

## B.3 Primers for amplifying 3' UTR fragments of predicted target genes

Target gene	Forward primer	Reverse primer
ATG7	ATAGATCTTGAAGCTGTGTTGCTGCC TTT	ATGCTAGCTCAGAGAATCAAAGATT GCAGAA
ENAH A	ATCAGATCTGCCGATGTGACTCTTT AACTTG	TCGCTAGCCAAAAATGTCCACATGC ATAAA
ENAH B	ATAGATCTACCAATAGCGTTCAGCC AAT	ATGCTAGCTTTTGTGAACTGAACTT ACACACATC
ETS1	ATCAGATCTTGAAGAGCTTATGGGA CTTCATC	TCGCTAGCATCCGTTTGTTAATATT TTTCCTAGCC
FZD8	ATCAGATCTGGCAAGACCCTCGAGT CC	ATCGCTAGCCCCTCACTGGTGAGAA CCAC
LEF1	ATCAGATCTGCATGGTGGGAAAAGA AATG	ATCGCTAGCAATTCAAGTGCTGGGC TTTT
NFIB A	ATAGATCTAAAGAAGGAAGACGGGA	ATGCTAGCAAAGCACTACCTACAT

	TGG	AGGCAAAA
NFIB B	ATAGATCTGAGGGGAAGATGGATGGC TTT	ATGCTAGCAACAAAAGACCTTGACA GGATCA
NR3C1	TAAGATCTCCGACTGCTTCACCCTT TT	TAGCTAGCAGAGCCATTCCCAAGCA CT
PPP3CB	ATCAGATCTCCACGTCAACTTCTGG TTTACTC	TCGCTAGCCAGCCTCAGAAATTTTA AGACCA
PRKACB	ATAGATCTAGGAGTAAGAGGACAAG ATGACA	ATGCTAGCCCAGAATAAACTTGGC AATAACC
QKI	ATAGATCTCAGGTGGCTTATCTCCT ATGG	ATGCTAGCCAAAACCCAAAACCACA AACA
ROCK2	ATCAGATCTTTAAGGGCTAACCGCT TTCC	TCGCTAGCTTTTCATAGCACTTAAT GAGGGTTAAT
TCF12	ATCAGATCTGGTGACCTTTCTTCAC AAGGA	ATCGCTAGCAAAAAGATGGCACATT TATTGCT
VANGL2	ATCAGATCTGTCGTGCAGGTCCTG AG	ATCGCTAGCTGCAGTAAAACAACCC CAAA
ZEB2	ATCAGATCTCCTGCTTGAAAAACAC TGCT	ATCGCTAGCATGTAAGCAGGCTTGG GTGT

#### **B.4 Mutagenesis primers for generation of mutant reporter constructs with nucleotides in the predicted target site mutated**

<b>Primer</b>	<b>Sequence</b>
ATG7 138 Mut Fwd	CCTTAAAAGCAAGGATACCATGGGCTCCTATATTCCCATCATA AAAGAAGACTCACTGAA
ATG7 138 Mut Rev	TTCAGTGAGTCTTCTTTTATGATGGGAATATAGGAGCCCATGG TATCCTTGCTTTTAAGG
ATG7 499 Mut Fwd	GAAGTTAAGAGAATGTTCTACTATGTAGATCTAATTGCTTTGA GGATGAATCATTGATTC
ATG7 499 Mut Rev	GAATCAATGATTCATCCTCAAAGCAATTAGATCTACATAGTAG

	AACATTCTCTTAACTTC
ETS1 499 Mut Fwd	GCTTCTTTGGGTGAGATTCTGAGATCTAAGAAAGCCAGAAAA TGTAGCTCACTTTA
ETS1 499 Mut Rev	TAAAGTGAGCTACATTTTTCTGGCTTCTTAGATCTCAGAATC TCACCCAAAGAAGC
FZD8 499 Mut Fwd	TGCCTTATAACAGACTGAATCAAACAAAGCTTAATTATACAC CCCAGATAAATACGGGT
FZD8 499 Mut Rev	ACCCGTATTTATCTGGGGTGTATAATTAAGCTTTGTTTTGATT CAGTCTGTTATAAGGCA
LEF1 130 Mut Fwd	GCTTTTGTCTTCACTCCACCTCTGCCTAGATCTCTTAAGCCCT CGTCTCATCTGTGTCCA
LEF1 130 Mut Rev	TGGACACAGATGAGACGAGGGCTTAAGAGATCTAGGCAGAGGT GGAGTGAAGACAAAAGC
NFIB B 130 Mut Rev	AGGAGGCCAGAGACTGGATTTGGATCCTCAATGTTATCTTAAT CTTTTGACTGCTGAGC
NFIB B 130 Mut Fwd	GCTCAGCAGTCAAAGATTAAGATAACATTGAGGATCCAAATC CAGTCTCTGGCCTCCT
NFIB B 138 Mut Rev	CCTTGACAGGATCATATAAACTATATTGGATGTAGATCTTGGG ACCTTACTGCCAGCA
NFIB B 138 Mut Fwd	TGCTGGCAGTAAGGTCCCAAGATCTACATCCAATATAGTTTAT ATGATCCTGTCAAGG
NR3C1 138 Mut Rev	ACCTACGCTCTGTATTTACATAAAGCTTGTACCTGTATGCATA TTATTGCCCA
NR3C1 138 Mut Fwd	TGGGCAATAATATGCATACAGGTACAAGCTTTATGTAAATACA GAGCGTAGGT
PPP3CB 499 Mut Fwd	CATGAAATGTCTTTAAAACAATAGATTAGAATAAGCTTAAAAT TTCTGAGGCTGGCTAG
PPP3CB 499 Mut Rev	CTAGCCAGCCTCAGAAATTTAAGCTTATTCTAATCTATTGTT TTAAAGACATTTCATG
PRKACB 130 Mut Rev	GTCTCAGCTCCACCTCATCTGTTAACAGCTAGCAAAGATCCCA ATGTGAGCCCT
PRKACB 130 Mut	AGGGCTCACATTGGGATCTTTGCTAGCTGTAAACAGATGAGGT

Fwd	GGAGCTGAGAC
ROCK2 130 Mut Fwd	TGCTTTGTGGCAGAAAAATAGAGATCTATTTTTACACCTAGT AGGTATATCAATGTCAA
ROCK2 130 Mut Rev	TTGACATTGATATACCTACTAGGTGTAAAAATAGATCTCTATT TTTTCTGCCACAAAGCA
TCF12 138 Mut Fwd	GCTGATAAACCTTTCCAACCAGATCTAAGAAGTTCAAAAATT TCTGTCAATGTAACA
TCF12 138 Mut Rev	TGTTACATTGACAGAAAGTTTTTGAAGTTCTTAGATCTGGTTG GAAAGGTTTATCAGC
ZEB2 130 Mut Rev	TTTTAAGTTTGTAGTGATATATAAGTCGGATCCAATTCGTACA ATTTACTAGTTTGTG
ZEB2 130 Mut Fwd	CACAACTAGTAAATTGTACGAATTGGATCCGACTTATATATC ACTACAACTTAAAA
Amp GA 1 Fwd	CAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGA AGCT
Amp GA 1 Rev	GGAGCTGAATGAAGCCATACCAAACGACGAGCGTGA

## B.5 Northern blot oligos

Anti miR-130

TGCCCTTTTTTTGCCTG

Anti U6

GCTAATCTTCTCTGTATCGTTCC

Anti Putative 1

5' -CCACGGCCCGGGCAGCGCTCGG

Anti Putative 2

5' -AGTATTGGCCATTGTCCTCTAA

Anti Putative 3

5' -AGGAGGAGCTCCTGCATTCAA

Anti Putative 4

5' -GGAGAGTCCCTCAGCAGCCAAG

Anti Putative 5

5' -ACAGAGCCCAGCCTCTCTGATGA

Anti Putative 6

5' -GCTCGAGTCACTGTACCAGAAA

Anti Putative 7

5' -TCTACCACTGAGCTACATCCCC

Anti Putative 8

5' -CTGTGCAGGGAGCGCAGCGCA

Anti Putative 9

5' -AGTCAGTTTATGCAGGCACAA

Anti Putative 10

5' -TCTGGGGCTCACAAGGACAGATA

Anti Putative 11

5' -TGCTCTCTGCAGCACCCCTGAA

Anti Putative 12

5' -GCCCAGGACGGACGCGGCCAG

Anti Putative 13

5' -AGCTGTTTCCGTGCACTTCTGCAT

Anti Putative 14

5'-GAGTGCCCCACTGCAGCCCCG

## Appendix C

### C.1 Cloned 3' UTR sequences

The portion of 3' UTR of each target gene that was amplified is shown, with the primers used to amplify the 3' UTR shown underlined. In each case the position of the predicted miRNA target site is indicated.

#### C.1.1 ATG7 3' Region with miR-499 (green) and miR-138 (pink) target site

TGAACTGTGTTGCTGCCTTTTATCTGTTATTCGACAAGTTAATGATTGCTTACCCATTCTG  
ACTGCCATGAATATGAACAATTTAAATCACATACTGGCATGAATATCTTCACTGTAAAAGC  
AGTCATTTCTGGATATAACTCTTCGTATTTTCATTTTCATTATGTAAATAACCAAAGTAAT  
GTGGTTATGTTTAAATCTTATTCCTTTTTTTTTTCCCTCCATTTTCTTAATGAGACATACA  
TAAATTTGCTGCTTTAGTGAACCTTGGTATTTCAATTTGTTCCATTTTTTATTTGCTTGT  
AAGCAGACATGTCTAAGCAGCATCTCTAAATGCTTTTTCTTCATCTTTCCAACACTTGTGT  
ACCTTTACTTAGAGTTCATTATTAAGTTGAGAATATCTTCTGTTGGGAAATAAACTGCAT  
TCTTATTTACAGGAGAGCAGGAGTAGGTTAAAAACCTACACTTGCAGTGTATATCATGTTG  
CTTTCTATTTCAAAGTAAATATTTATAGTACAATTGATTCCCTTTTCTTTATTATGGTATG  
ATTGTTCTTACCTGTGTCAACGGAGAAATTATTCTCTGTAAAGATTTCATAATTTCAATTCT  
ATGGAAGAAATGTGGTTTGACTTATATTTACACAGGAAAGACCAAATGGATCTGTATCAT  
ACTCTCTGCTGAAGTTAAGAGAATGTTCTACTATGTAGTTAAATGCTTTGAGGATGAAT  
CATTGATTCACTAAAAGTTGTAGTTAAATGTATTACCAAATAAGAAATTGATTTTCTAA  
CTGTGCCATTCTTGGTTGCTGTTGTGTATTTCTTTACAGGTGTTAAACACCGATTCTGATT  
TTCATTTTGGCTGGTGTTTAAGTATTTCAATTTCTGCAGCATCAGGAATGGCTTCAGATAT  
TATGGCTGATACTGTGAGAATCTTTGGGTTCTCAAGGATTTGTAAAAATGGATTTGTAAAA  
TATCAAAGTAACTTACTACAAAACCTGAAAGCAAATCTATTGTTTTTTGTCTTGCAATAGGT  
TTTAGAAGTATAATTGAAATTCATTTTAAGGCTTTATATTGTGGAGATTATGGACTAACTT  
GCTCAGAAAAGAAGGCAGAATTGCTTTATTTGGCAACCAGTCTTTATTTCTATTATACATT  
ATGAAAATTAGTTGACCAACTAGTCAATCCAATTAATTTAGTTTGGTTTACTTCCTTTAAAC  
ATAAGCTCTGAATATCTGGTGCAGATAATTTACAGCTTCTCTCTTAGACCACATCAGAGG  
ATGTTCTGTCTGGGTCAAATTCACCTTAAAAGCAAGGATACCTGCAGCTCCTATATTCCCAT

CATAAAAGAAGACTCACTGAAACATGTTGAGGTTAGCATTAGTTGTTATGTCTATTTCTTT  
TTGTGATCACTAAAGTAGATTTAAGTTTTAAAGTCCTGTTTAGCTAACTTCATTGTAATAT  
ATATATCTGAATTAATATATTTTTTGATAGGAAACTGATCTTTTGGTGGTCAGGTAATGA  
TCAATAAGATTGATAACATGTGATATACATTTTTAAGTAGTTTATTAACATACAGCATTAC  
CTATGAGCTGTGTGATTTGCATGCTGTAGTTCTTCTGGAATCGCGTATCGGTAGTGGGTAG  
TTTAAAGGGAAAGACTCGCATTTTTATCTTTTGGAAATTGCTGAGTGTAAGTCTTCTGCA  
ATCTTTGATTCTCTGA

### **C.1.2 ENAH A 3' Region with miR-499 target site**

GCCGATGTGACTCTTTAACTTGGTATAAACTGACTACGTTTGTGAGCTGTTAGAAGAAAA  
TGGAAACAGACACTTGGGAAGGAGGGAGAAACATCATATTCTGAAAAGCTTCAGACGCATCA  
CTCTGGTGATATGCTATTTTCCCTCCTGGTTTGCAGCTTTTTTCTGACCTTTACAGCAGGG  
TGGAAAAGAGTCTTAAGGTTACTGTAATGATTGTGCAGAAATTCCTACACCTTTTCAGACAA  
GATCACAGTGCTTTGAAGTATACTCATGTCAAATAAAATTCACATGTTTCAGAATTTGT  
TGTTTTAAAAAAAAAAACAAACAGGGAAAAGCTTGGGAAGGCTTTTCAGAGAGGTTTTCTTT  
ATTAATGAAGGATTTAGCAAGTTATACTGTTGTACTTCAAATGTAACCTCAGGTTTGTCTTC  
CTATCATATCTGATGTTTCCATGAATAATTTAAAAGATCAGATGTGTGTAAGTTCAGTTCA  
CAAATATGAAACAATTAGAAATATGTACTTTTTAAAGGGCAATATTTGTAAGTGAATGACC  
TAGACAGTGTTACATGTCAAGATTTTGGGATTTTATGATAGGAAGCCTCTCTAGTTAGCC  
TTCATGCAATTTTATAGGAAAAAAAAAGAAAAGCAATTGCAGTCCAGAGTTTAAAGATGTAG  
ATGCCTTCTACATTGTTACAATGCTTTACCAAATCTAAGACTTCTACATAATGCGAATCA  
GCAGTCAAATGTAATCATTAGTATGTTGTAGATTTACGGTAGATTTTTGGGCTTTTTTTA  
GATTTATGCATGTGGACATTTTTG

### **C.1.3 ENAH B 3' Region with two miR-499 target sites**

ACCAATAGCGTTCAGCCAATGGCTGTTTCCACCTTCTACACTTCTCATGCTCAGTCTTGAA  
CAGAAGATTCAGATAAAAGAACGTTCACAATGTTTAGTTTTCAACAGAAGATACTGAGTGG  
TCTTCTGTAGTCATGGGAAATTATTTAGAAAGAGTATTTTCATGCATCCATGAATCAAGAAG  
TAAATGAACTACATGTAAGCTGTTACGGGTGTTTTTTTAAACACCTTAGTGTAGATTTCTT  
TTTACTTGTGGAATATACCAAAGCTGAGTATGAAAAAGGCTTTTTGCTGTGCTGATAAAAT

AGCAAATATCCGTCAGCTGATGTGGTAGAACGAGTTGACTGGTGTGGGGCAATAGTGCAGT  
GTCATCCCATATACTTCAGACACTAACTGTTTAATAAAAAGTTTTTTTTCTCTATTAATATA  
TTCCTACTACGCTGCTTACTCCAGCCGAATGGACAATATGATGCAAATAATGTGCAGCCCAGA  
GGCTCTCCTTAAACTAAAACCGAGTGAGCCCACACACATGCCACTTGATATCTCCAGCAAG  
TCCACCATTGAGATACGGGATGGAAAAATAAAACATTAATGAAGACTACAGAGGGGAGAAG  
GGCACTGCCTAGCTCAGAATTACACACGTCTTTAACCCTAGTGAAGCTACGTTATGTGTGAT  
TATTCACTTTTCTGTGACTGTTTTAGAAAGGCTGAGCTTTCTTTTGGGGGCTAGTTATGTATG  
CTGGCTTTCTTCCTTGTTTTAAAATGTATGATTTGTTTATTTCTCTTTCTCTGTAGCTATC  
AGGCAGGAACTGAGCAAGTCGAATACTGCATAGAAAGACTAGTACTAAGCCGATGTGACTC  
TTTAACTTGGTATAAAACTGACTACGTTTGTGAGCTGTTAGAAGAAAATGGAAACAGACAC  
TTGGAAGGAGGGAGAAACATCATATTCTGAAAAGCTTCAGACGCATCACTCTGGTGATATG  
CTATTTTCCCTCCTGGTTTGCAGCTTTTTTCTGACCTTTACAGCAGGGTGAAAAGAGGTCT  
TAAGGTTACTGTAATGATTGTGCAGAAATTCCTACACCTTTCAGACAAGATCACAGTGCTT  
TGAAGTATACTCATGTCAAAAATAAAATGCACATGTTTCAGAATTTGTTGTTTAAAAAAA  
AAACAAACAGGGAAAAGCTTGGGAAGGCTTTCAGAGAGGTTTTCTTTATTAATGAAGGAT  
TTAGCAAGTTATACTGTTGTACTTCAAATGTAACCTCAGGTTTGTCTTCCTATCATATCTGA  
TGTTTCCATGAATAAT ] TAAAAGATCAGATGTGTGTAAGTTCAGTTCACAAA

#### C.1.4 ETS1 3' Region with miR-499 target site

TGAAGAGCTTATGGGACTTCATCTTTTTTTTTTTTTTTTTTTTTTTTTTTGCTTTTTTTTTTCGTTAT  
TTTTTTGTTAGAGAGCAACTCAGATTTGATGCTTTGAGGATGTTGGGAGAGAGGGTGGAGG  
GCGAGTCAAACCTTCTGGTTGGAATAACCTCCAGAGTTTGGTTCAGAGAAGGATGAGCCTTT  
GTTTGCTGAGAGCTCCAGGACAGGCTTCTTTGGGTGAGATTCTGAGTTTAAGAAAGCCAG  
AAAAATGTAGCTCACTTTAAAAAGAAAAAAGAAAAAAGAAAAAAGCAAACAGGAAAA  
CCTCCCCGTCCCTCGCTGGCTAGGAAAAATATTAACAAACGGAT

### C.1.5 FZD8 3' Region with miR-499 target site

GGCAAGACCCTCGAGTCCTGGAGGGCCCTCTGCACCCGCTGCTGCTGGGCCAGCAAAGGTG  
CTGCCGTGGCCGGGGGTGCTGGAGCAGGAGCGGGCGGGCAGACCGCTATCGCTGCTGCTGG  
AGGACTTGGGGCTGGTGGAGGCGGCTCCCTGTACAGCGATGTCAGCACCGGCCTGACGTGG  
AGGTCTGGCACTGCCAGCTCCGTCTCCTACCCCAAGCAGATGCCTCTGTCTCAAGTGTGAG  
GAGGGGGAAAAGAGGCCATAGGTTAGGGAATGAGGGACACTGGCCTGCCTGAAATGCCCC  
TCACTGTTTGGTGCCATTAATGAACCACTAATGGTCCTTATTAGCCACAGCTTGTATAGAA  
CAATGCAGCTGGCAGAGGCCCCAGGCTCCAGCCCCCTCCTCCTTCCCCTCCATTCATATGC  
AGGGAGCATGTACTTCAAGGAGCCAGCATATTATTTTGGCTGGCGGAGGAAGGCAGGGGCTC  
ACCCGTGTCTTGCAGAGGGCAAGGCAGAGCGGCTCCTGATTTACTTTGGACTAACGGCAGC  
TCGTGGGAGGGAGGGTCTTCTCCCCATCCCAAAGCGGGAGTCTGCTGGAACTACAGGTT  
TTTGGGATATAGATGACCAGGCAAATGCCTTATAACAGACTGAATCAAACATGTCTTAAT  
TATACACCCAGATAAAATACGGGTTTCTACATTAAAGGATGTATTTATATAATTATTTGT  
TACCTTGTACAGATGTGTAAAATATGTATATATCCAAAGCTATACAGTCTGTAAATTTTTT  
TGTA AAAAAGTTTAGAGGCTACCCCTGTAAGAGCAAATATGAGTATTCTATTTTGTCAATA  
AAATGACTTTTGATAAATGACTTCTTCAAGCTTTTCCCCTACTCCTGCCCTGGTTACTGCA  
GCAGCTGTGTCCCTGGGTGGTTCTCACCAGTGAGGG

### C.1.6 LEF1 3' Region with miR-130 target site

GCATGGTGGGAAAAGAAATGCTTTCTCAACTTGCAAAGCAAAGGCGGCACACCAGGACCC  
CTGCTGGAGATGGAAGCTTGTTAAACCCTAGGTGTGTCCTTCATCCACACAGACCACCC  
AAGGAGTCTGTCTAGGATGTCATTCACCCTGATGTCACTGCTAAAGACACTGAATCATAAA  
GACAATCACTGCCAAAATCCTTCACTATCACGAGGTCCAGCCCTGATATAAATAGAGCGAA  
TAAGCTCCTGGGTCAGCAGAGCAGCTTTTGTCTTCACTCCACCTCTGCCTTGCACTCTTAA  
GCCCTCGTCTCATCTGTGTCCATGTTGAGTCAGCAAACAAAGTAGCGTGGCTGTCTCTGCC  
TGCTTCAGCCAACAGAGTAACTTAGAGCACCGCTGCCCATACAGCACATCGAGGAATGGAA  
TTTGCCCTGGACCAGGATGACACACTCAAACCTTCTTTTTATTTAATGGTGGCAGAGGATCA  
ATTTTGGTTTTAGCTTTTTTAAGAACTTGAATGCTTTTTGTATTTTAACTTTTAGTTTGT  
ATTTCTTGGTGTATATTTTACTAGGTATGAGCTTTTAAACAACCTGTCAAAAATCTGAACT

ACTTTTTAAATAATGGGGTTTATTTTCCTTTTTTTTTTTTTTTTTTTTTCTCTCCTGCCAGAG  
GAAAAGCCTGTATGGAAAAAGGTGGCTTTTTTTGTTAACTACTTGTAAACAAATAGTGGCC  
TCTCAGTCTATATAATACAGTGTTCATATAGAATAAATAAATGTAACAACAGTCAGCTGGT  
CAGTAGCCATGTCAGTCTTGCATCACCTCCTCTGTAAACATAAGCAATACATCCAGCAAAT  
CTGACTGCAGTGAGTTCTCCTTCTTACCCATCGCCCTTCTCCTTGTAAAAAGCCAGCACT  
TGAATT

### **C.1.7 NFIB A 3' Region with two miR-138 target sites**

AAAGAAGGAAGACGGGATGGGAGAAGAACTGAAGAACGAAGAATAAATAG**ACCAGCAATT**  
GCAGCACTTACAATCACTAATTCCTTATGGTCGAAACTGAAATGACATACAAAGGGTTCGA  
TGATATTTCACTGATGGGTAGAACGCAGCTCCTGCTAAGTAGCCTTTGTTATATGAAGTCC  
GCTGGGAAATAGATGTTCTGTCTCTGACAATATATTTAACTGACTTTCTAGATGCCTTAA  
TATTTGCATGATAAGCTAGTTTTATTGTTTTAGTATTCCTTGTGTTTACGCATGGGATCAC  
TATTCCTGGTTATCTCACAAAACAAAGGCTAGGCAGCAGCAGCAGATTTGCAGGAGGACAA  
GGGCTGTAAGCCAGCCGTTTTCTCAACACTCTGATTTCTAAATGTTTAAAAAAGAAAAAAG  
AAAAAAGCTCTGTAGCCTTTAGTTATCAGTATGGATTTATTAACCTTTGGCCCTTGA  
CTCAGCCTTTTCTAGTGTGTTACCCAGCTTGAATTTTCAATCAGAATGAACACACACGCT  
CTATGTAAGAGATGCCTCTATGTAGGTGAAAGTGTTATCTTTTCATGCAACAGTAAAAAA  
AAAAAAGAAAAAAGAAAAAACCATACAACTACAACTTTTTTACCCACTAAAGAAAC  
ACTTAGAGAGGCAAGTGCAATTTAAAAGTTGTATCAAAGGGTTCATCATTATAAGTCCTTC  
AGCCCTTGGACTCTAAATTGAGGGGATTAAGAAAGGAATTAATTTACTTTGAACAAAT  
TTATTTTTCCCCTCAGTTTTTGAGGGCATTAAAAGGCATTACATCAAGACAAATCATGTCC  
TTGAGAAAAATAAATCAATGAAAACACAGCACTTATGTTGGTTTAGCTGCTGCCTCCACAA  
AGGGGTAGAATTTATTTATTTAAAGTTACTGGTTGCATCAAGAACCATAGGGTTTATATT  
TTTCTGTATAATCTGCATCATAGAGACAAAGAGATAGGCAAATCCATGCCAAAAGGGCAA  
GCTTACTGTTTACAACTGGTAATACCAGGATGGGGATAAGATTTATTTAAATAATGCACT  
CTCTAAGTCTAATTACAGGGTGCTGAAAACCTTGCATGGTGCTTTCAGAAGTTAGCCTTGTT  
CAACAGATTTACATATGACAGAAATATAATGTGCATGGGCAGACATTTTGCCTCTGTGT  
CTCTAAAAAGTAATTAATTAATACTCTTTCCAGTAATCCCGATTTGCACGAAAATATA  
ATGTCCACATTACGTGCCTTGCCTTGAATCTAAAAAATGAAAAAACAACAAAAA

ATACAAAAAATCCACAAAAACAAACAAAAAATAATGACATCACTACACTTGTGTTTTGCTG  
CATTTATTATCATTTAAATCTTTACCATTTTTATGACAAAATATTTTGTACTCCAGATGGG  
AAAAAAGTGTGACATCATGGATTTTTTAGACAGTTATATCTTTATCTCACATTTATAAAG  
CATATCATGGCTGTGTATAATTGCCGCTTAAGAATTGTAATCGACCAGCAATATTTTCAGT  
ATTTTGGTGTTTTTTCTATTAACCTTTCATGTTTTTCATCTTCCAATTAATATTGGGGGGG  
GGAGGGGGGGAGGGGGTTCAAATTTATACGAATTATGCAATACCAAGTTTTGCCTATGTAG  
GTAGTGCTTTT

### C.1.8 NFIB B 3' Region with miR-130 (blue) and miR-138 (pink) target sites

GAGGGAAGATGGATGGCTTTATGCGAGCTTTTCTGTTATCTCTGACCTCTTGTATGAAAAA  
AAAAAAAAAAAAAAAAAAGAAGAGCTTGCAGATGTGATATTTTTAAAAGCAGATTCCTTCCA  
TGCGAATGAATTGGTTTTGTTAAGCAACTGTGGGTTTTTTCACACGTGCCTCTCTGATAGA  
AAAGTCTGGTTGGAAGAAAACACACTTTTCATTTATGTTTTGTTTGTGTTTTGTTTTGAAGC  
TCAGCAGTCAAAGATTAAGATAACATTTGAGAACTCAAATCCAGTCTCTGGCCTCCTCCTT  
CACCAGGAAGTGTATTCTGACTAAATCAAGCCACCCCGTTCTATGTGCACTGTTCTGCCTT  
TATTAAGTTGCTTTGCAGCCCTGGCCTCATGGCTGTGCCACTGAAGTGTGTTCCCCCTGGA  
GTGACATGAACGTTTGAGGCTTACTATGGAAAAAAAAAAAAAAAAAAAAAAAAAAGGCAGTGAA  
GGAGACTGTACATAAAGACATGGCAAAAATATTAATTATAGCAATATAGTTGTGGGGTGAT  
GTTTCAAGTGGGCAGCTCCATTGAAAATATGTGAATGGATTTGTGAAGCTGCAAGTAGCGAG  
AAGAAAGGAGGATTTTTCTGAATGAACCGCCTACCTTGTAGACAGTAATTTGTACACTGTAT  
AGTTTTGTAAAGATTTTTTTTTTTTTTAATTAATAACCTTGTTTGTAAAGCTAACTTTTT  
AACAAATTATAATGGAATATATGTCATTTCCATTTTTAAAGTAAACAAGAATATTCCTTGT  
TTGGAAGCTGGACTTGGGTTAAAACCTCCTCAGTTTCTTAATTTATGTGTTAAAAAAAAAAAA  
TAAATAAAAAATTTGTCCATGTTAGCAGATATTTACAGATTCCTGTGCTTGAAAAGCATA  
GGTTACTAATCCTTTAAAAATGTAATGGAGAAAAGTTATATTTTATGAAGGTTATTTTGT  
TGTATTTAGTATTGGAAAAGTTGGTTTTTTCAGAGCATTTCAAATGTTGGAAGCACCCTGT  
CTTTTTATTAGTATATATGGCCTTTAGCAAAAGTTTTTGTGATTGTTACATGATGGTATTT  
AAGATTAGGTTACAGAGCAATTCAGGATAGGCAGAGAATAAAAACAGTACTTATGTCTCA  
CATAGCTGTGTCCTCAGGGAGCAGAATTTTGGATTTGCAACTTGTAGCTTCATAAGGACTT  
AATGAAAGAGATTGCACAAGGACATACTCAGCTGTGACACCGGGGTGTGTTCTGTACCTAA

ATCTAAATTATATTTACTGTGTGAGGTTATGAAGGCTGCAGAAAGTTATCCCATATTCAGT  
GTACAGTATTCACTTTTAATGAAACAACCTCTTTAATGTTGCTGGCAGTAAGGTCCCAAGCA  
TGACATCCAATATAGTTTTATATGATCCTGTCAAGGTCTTTTGT

### C.1.9 NR3C1 3' Region with miR-138 target site

CCGACTGCTTCACCCTTTTCCCTCGGGAGCAGGAGAGCTCCGCTGCCGGCTTCGTGATCTTA  
GAGGTGTACGTCCACCCTTGTCTTTGTACGGCTGCATCCCTTCCCTATCGAGTACTAGCGT  
GCCATAAATCTGCCATAGGTCTCGTGAATTTCTCCTCGTTGCCATGGGTAGCTGTATTTTA  
AACGTTTAGAAGCTGTAGTAGCCTTTTCTACCTGCACCTTAACGATTTTCTTTATCTCGA  
AATTGAAATATTTACCAAGCCACTCTAAAATTTATTTGTACTGAAAGTGGCCCGATTATTT  
GTGTAATAACGATGAGAAAAATCATCTGGGAAGATACAAAACCTAATATATTTTTGTATTT  
AGTTATAGTTTTCCAATATGTATAATCGGTATATTCGCTGATCTGAGAAAAGAAGGGAAGG  
CTACTGCAGCTTTAAAAGCAATTTATTAATGACTGTAAAATAGCTTGTATAGTGTATAATA  
AGGATGATTTTTAGATGACAGATTGCTTTATCATGATACATGTAATTATATTTTTGTAGG  
GGTCACGGATATGCTGAATGGTAACCCGTGCCAGCCGTGCTTTGTCTGGAGGTAGTGCTGA  
AGGCACTGACATTCAGCATCCAAAGAAGTACTCTGTTAGAAAGAAAAACAAAAAACAAAA  
CCTGTGTTCAAGTGGCTGTTTTTAAGGTTGAATTTGCATTTCTCCTATCACTGTGTCAAAGC  
GCAAGTCGGACCTTGGATGGAACGCTCCCCTCCGTGCCTCCGCCTGAGCGGATGGGTTCCC  
TTCGTAGACAGCCGCACGAGCCGCAGCAAAGCAGCTGGAAACATTGGGCAATAATATGCA  
TACAGGTACCAGCAATATGTAAATACAGAGCGTAGGTTGTTCACAAACACACCAGTACCTTT  
CTGCCATGCAGCAGAAGTTCGTTTTCCGAGCGAGACAGGAAACTAAATTTTCATCCTTTCTT  
AAATCTTTTTGAATGTTATACCTATTAAAGCAAAGATACTATATTCAGTACTTTGGAGTA  
AAATATCATCTTTAATAAAAAAAAATACTCGTTCTTCTCTTTGAATAATCTGTACAAAA  
TACGTACGTTTCATGGCAGTTCAGGTGCCCTTCAGTGCTTGGGAATGGCTCT

### C.1.10 PPP3CB 3' Region with miR-499 target site

CCACGTCAACTTCTGGTTTACTCACTTATGTCTTGCCTGAGAATTCCTAGATTACTGCTTT  
TTGATGTGACAACAGACGAAATAGCCAGGGGGGTCAGCGAGGTGACTGCTTATCGGGAGGT

GGAACGCCATCTGCCTGACCCTGCTGAAGAACAGCGCCGTGCTTTGAGACAACAGCCCCCTC  
TCTTCACGCGGGCGACCGACCTCCTGGGAGCACCCCTTCGTGTATTCAGGCTGCTGCGACTGC  
AGAAGAGCTGGGAGGAAGGAAGCGGGCTGTTATCACTAAAGCACCCCTAGACTAGCATCTT  
ATGCCTGTTGTTGTCAACACTTGACGTTTACAAATGGGATTTTTCTTTTCGGTAGAGGTAGC  
TTGTTTGTGTTGTTGTTTTTTAAATCAGACTTTCAAATACGACCCTAAGAGCTAACTATTT  
ATTAATGGACATCGGTTTCATGAAATGTCTTTAAAACAATAGATTAGAATGGTCTTAAAT  
TTCTGAGGCTG

### C.1.11 PRKACB 3' Region with miR-130 target site

AGGAGTAAGAGGACAAGATGACATCAGGGCTCACATTGGGATCTTTGCACTCTGTAAACAG  
ATGAGGTGGAGCTGAGACCATCATATGTCAAAACAGTTACCTAGTTACTTCATTCCACAAA  
GCTGACTGAGGTCTTTATTGCCATCTTCCATGTGTGCGGTTTGCACCAACCCTTCTAACTA  
GGCACAATTAAGCAAGCACTGTTTGTGCAGTAACACAGAATAATAGACCACCTTCTACTT  
AACTTTGGTTTCTGTCGTTCTTTGTTTCTGTGTATTGTTTCATTTTTCCCCTTTAAAATGAA  
GTAGTTTTTTACCCAAGAATGCTCCATTTTTATTTGAATAATGTATTATTTGTGTGAGTGA  
AATGGAAGGTGTTGCGGCTAGTGTGATTTAGACTGAAGTGAAAGGTAAATGTTGCTTCTAG  
TCTGTGCCAGCCAATAATTCTTGTCTGTCTCGTGTCTGATGCTACAATTTATAATAATAAT  
TTTTTAGTAGCTCAGACTAAACCTAGCCAACATTTTTAGCGTTTTTTGAAGGTAGTATTTAT  
CGCCATATAAATGTCTGCTAAATTAACCTAAAAATGTTTCTTCCCTGGAGACTGACCAAAAT  
TCAGTAGAATCAATTGGATGGCTTAAGTAGGAAGACAGTAGCTAGACCCCGTGCCTTGGTT  
TTCTACATCCCTGCGATTTTCGTCACTGCCTCCTTCCCTTTGTATTTCCACCTAGGCAGTAG  
GTCAATTGCCATAGCCAGACTGTACCACGGAAAGCTTACAGTCTCAAGAGAGGGAGTAATT  
CCACCTGGTCTGGATTTTGGACCATATTTAACCAAGTCAACTTACCGGCATTTTCAAGTTGGC  
CATATGAAATGACTTTTAAGAACATTAACCCTAAACTCTTAAAGGAAGTTTTGGTCAGTAT  
TG TAGAGAATGACTCATCTTGTTTTTGGTTTTGTTTTTTTTTTTAAATGTTTTTTAGCCAATA  
TTTGGGCTCCCGAAGTATTCCTTGTTC AAGCCTTTTTTTGATTTATTTTTTTTTTAAAGGCA  
TCCTTTGAAAAAAAAGTAAATAGAAAATGATGTTAAAGCGAAAAACCACAAAAGTAAAT  
GTGCTGTGCCATTGGCTTGTGAAAAGCAAATGGTATAGCAAATGCAAATAATGTTCTAGA  
TGCATTGCTGGATGAACACGGTTATTTTTTTGTTAGCTGTTAAGAATAAGAGGTTATTGCCA  
AGTTTTATTCTG

### C.1.12 QKI 3' Region with miR-130 (blue) and miR-499 (green) target sites

CAGGTGGCTTATCTCCTATGGAGTAAGTCATGTATACCTCTATCTGTCTGCATGAAAACCA  
TGCTGAGATTTTCAGGTAAACATTGTTTTACTTTTCATCTAAACAACACCGGTGTT **GTCTTAA**  
GAATTGTACATGAAATAAACTTTACAACCTTTATTTTCATTTTATTCTGACTCCAAAATCCTA  
GATTATCCCCGGCCCGGCAGTTTCAACAGTTTTGCAAAAGGCTTTGAAAAGCATTGTTTCAG  
GAATGAAACAGATTCTTGAACATGTTAAAGTTATGTTGTAGAATTAATAATTCTAAGTATTT  
ATCATACTCATTGAGTTGTTCAATCACTGATACTTCCGTAGCTATGGCCATGAAAAACA  
ATGTCAGCAAACCGAAATATATATATTTTTTTTCAATGCAGAAATGTTTTTTTACGTGGTTAA  
ATGTTGCTAAAACCTATGTTTGCATTCCTGGGTCTAAAGTTTTTTTGTGTTGTTTTTGT  
TTGTTTTTTTTTTTTTTTAAATTCTAAATCGATTTTCAGATGACTTGGTAGTACTACTGCAT  
TCACTTTTTAAGTAGGTTTTTACTCTTGCATAGGTTTTTTTTTTAGAGATAAGCACTGCAT  
TTAAAGCAAGAATCCCGGAATGTGTAATGTTTTTTACTCTTTACAACCTAAGCATTTAGAA  
TTGGTGAAGGGTTGAAAAATGTAGCTCATTTTAACTAAGAATAAAACCCAGTTTATAAAT  
AATACAAAATTGAATTTGATGCATATAACAGAGAGACTGACTGATTTAATTGTAATAAACT  
GTAAATGGTGTGTCACAGTTGACAAACCTGTGTGTACTGTAGAAGGCCAGAGACATGGCAG  
AGGATGTATCACAGGTGTACTGGTATTTTTTAAAGGTAGCTTGCTTTGGTTATAAACTCCA  
CAAGAAATAATAATGCTCAGAATATTCTTGAATGTTATATCCAAACCAACTGTCACATAAT  
TTATATATAAATTTTAAATACAGTTTAAATCAAATTTGTAATCAGTATCAAATGCTTTAAGA  
GTCTGTGTTTCTAGTATACCTGTAAAGTGGTGAGATTGTGCCACATAAATATTTATTTTTTA  
GTTTTTGCTTATTTTTGAAACAGCCGTTTTAATTTAAACAAATCCCAGTTAGCATGATTCT  
GAAAAATATCAGGGAGCACGTGGTCTCATTATGCTACAGTAGTCATTTTTTTTTAAGTGTT  
TATGAACTGTATCCTAAATGATAGGCCATTTTTGTTATGTGAAGGTAAATCAACCCATGAG  
AATGATTGCTAAGTCGATGCTATCTGTGAATATTACACCTACTGGACTTCATTTGAATGAA  
ATGGCACAAATCTGACT **TTGCACT**GAATACCTTTCTCACTTTCATCTCCTCTGCCTTAGTC  
AAAATGGCAACAGTACAAAACTTTATCAACCTCAGAATTTATTAGCTATGATTAAGCTGT  
TGAATG **AGTCTTAA**AAAATAAAAATTAATAAAAAAAAAATCATACTACTGTTAAGTGGACCAA  
GTTTGGTGAAGGAGAATGTGAGAGAATGATTAAAGAAGGAACAGCTCAAGAACATTGGGAA  
TGATAACT **TTCCACTT**GAGAATTTTTTTTTTATGTTTTACTGTAATTTGTTTGTGGTTTTGG  
GTTTTG

### C.1.13 ROCK2 3' Region with miR-130 (blue) and miR-138 (pink) target sites

TTAAGGGCTAACCGCTTTCCATGAACACAGACACAACCTCAAGGTGATAATTTTCTTCCTGT  
TACAGCAAGACTGAAACATATCCCAAATATATTAATAATATATATTTTCCTGAGAGATTG  
TGCTATATATATCAGAGGTTTACATTTTTGCTGCAATATAAATTACTAATCTCTGAGGGTT  
TTCCTGTATTCCCCGAGTGACTGATCAAACGAGGAATGGCTGGTAAATAGTAAAAGGATA  
TGTTAGGTAACCTTTTAAGTTCTGTTCCACAAAAGCTTTCTCGGCAAGACACATACACTAC  
ATTTCATAAAAGGATTGAGAGGAATGAATATTTAAACTGAAGAAAGTGACTTCTCAGCTTT  
CATAAAGATGC **CACCAGCA**TGCCTGCGAGAACAGGAGGACGATCCATCATAGTGATATAGA  
CTGCATCTGTAAGAAGTGACCATTATACTGTGTATATATATTGTACTGTAATACTGCAAGA  
GGGTTTTAAAATCTTTCCACTTTATTTTTTTATGCTTATTAATCAGATACCGTTACTTAT  
TGCAGTTGCAACTATGCACTTGTATAAAGCCATAATGTTGAAGTTTATATCATTTCATACAT  
GTCTATCAAAGCTGCATGTCAGTGTTTCAGCAGCTAGTTCGAAGTATATGCTAGTTTCAGC  
TGCATCATCATGGGCAGTCCTGTTTTACCTGTCTAATTGCCTTAATTTAATTTTTTTTTT  
AGTTTTTTTTTGGCCATTCTCTGTATTTAAGGAAAAAAGTTTACCAGCTCTTAGGATGC  
AATTTTGCTTTGTGGCAGAAAAAATAG **TGCACTA**TTTTTACACCTAGTAGGTATATCAATG  
TCAACTTACTTTGAATGTATATCGACTGGTTTTGAGGGTGGAGAAGTGTTGGTTACAAAAA  
CAATGTCTGATAAGAATGGACTTACTAGACCATGTGCTTATACATGTAACTGCTCATCCAG  
CCCTTCTTACAAACCTCAATATAGTTATTGACTTATATTAACCCTCATTAAGTGCTATGAA  
AA

### C.1.14 TCF12 3' Region with miR-138 (pink) and miR-499 (green) target sites

GGTGACCTTTCTTCACAAGGACCACAGACAACCTAACATTATATGAAGACACAAACCTGACAG  
GAGAAAACACTTGAAGCAAGAAACCCAAATGCAATCTTATGATCAAAGCTACTGGTCAACA  
CCTGCATCAGGATTGAAGATAGAGGATCTTCAGATAGAATTTTCAGCCCATGAAATACTCTG  
AACATATCATTCTGTTGCAAGCAGTGTCGCTTCTGCACAATCAGAGACTGTTGTGATCT  
CTCCACTCATTGTGGAAGTTGCCTTGTGCCTAAACTGAATTGACAAATGCATTGTAACTAC  
AAATTTTATTTATTGTTATGGAAGTCTACATATAAAGGGAAAAGTTAACATGT  
TAAAAGCTGATACAATTTTCAGCTGGATGCCAGCATTTATTAAGCTGTTACATTCAGAGA  
ACAAAGCAGTGACAACCGTCGACCCGTAGCATTCCCAGCATACCTATTAGT **GTCTTAAAA**AA

AGGAAGGGAAAAGTCTTTTGTGTCCCTTTCCTCTCCCTTTTGCCATATGACTAGTGTTTTCC  
ACGCAATAGGAACTATTCCTTGGTATAGCTTTTTTTATGTTATTTTTGGTCTCTGGTAAT  
CTGAACAGTTATGGTCATAGTTACCCTTCCATACTGAGCTCCGTGTTTCAGGCCGGTAAGG  
GGCCTCCTGAGCAGGCTCCAGGCAGGAGCAGTACTGTGACCATAACATTTTCGTAGATTTAA  
AAATTGATTTTTATTTCCCTCAGACTTTCCTCAAATGTTTTTTTTTCTAATTATAAACAA  
AACCTTTTTAGGCTTCTAGGCTTCATAGTAAAGCTTGTAACGTGAAGTGTAATTGGGTGG  
GGAGAATCTACTTTGTTAGAGTTGCTTTGTTTTCCGAGCAGTAAGTACTACATATAGTACA  
TGTAAGTGTTAGCTGTATGTAAGCACAAAATGCATTAATAACAAAGGAGATTTTTTTTCA  
GGCTGTATTTTTGGTGAATAGTAAAATCCTAAATTCACCGTGGCAGGTAGTGTGATGACAT  
AGCCACCAGTAAGGTGGGCTTACTTACCTAATGGAATATAAGAGCAATGGTCACCAGTTTT  
TCCTTACTGTATCCATCAGTTATCTTGTCAATTTTTGTATGGCTTGTGAGTGGAGAAAAATC  
AGTGTGAATCAAATGCAAACAAATGACCCCTCCCAAGCCATATCAACATTAACAATTAAA  
TGTTTCTCCACCAGAACAGTGTCCATAAGTATTCCTTTGTCACAAATGTGTTGAGATGGTT  
AAAATCCCTTTGAACTTCCATGGAACTGTCTTCCACTGCAATGACCATGGGAGGAACTT  
CCTAAGCCAAGGTTGTCTGAGCTATTGTTTATATAGTGCATCTTTCATTACTGGTAATTGG  
TTCACCGTACAAAACAGGAACAAAACAGTACTTTTTATGAAGTCATTTAATTTGAAAGAAC  
AAAAGCTGGATCGTTAAAATAAACTAATGAATTTAAGCTCACGTTTGCAATGGTTCCTAAC  
AGAACAGTTTGTGAAAATAGTGGACTCATTTCTGTTACTTTAATTTGCATATACCATCAGA  
TAACACCTGATTATATCTGTGGCCTTGTCTTTCATTTTCAGTGTTTATTGGCTAAACAAATT  
TGTTGCTTATGACGTGTGAAAGTGATCACGTGCCACTGCCTGGCCTTTTTCCCTTCTAAGCT  
TGTTGTCTTTTTGGCTATATTAGACTTTGCAGTATGCCCAGAAGCTTTCCTTCATAAAATA  
GAAAGAAAAAACATTTGGCTTATTTTTCACTGTAGCTAGTCTTTTATAACAATAATCTTGT  
AAGAAAATTTCTTGAATTCTAAATATTACTCTTTCTAGATTTTTGAAATCGAAAAAGTTTT  
CAGTAAAAAGTTTCTTACTTTATTTTACTATATTAGGTAGTAAAAATGTAGGGTTATTTAC  
CATAACCTGTTCAATTAATATCAGAAATTTACAATAGCATTGTAAGACCATAGTAGGGTTCT  
AGCATAACCGTGTAGTACCTATGGAGTATTGTAAGAGCTAATTGTCCGAGATGAATTGCTTC  
TCATCTTGTTCTCCAGTTTCCATTGTTGGTTTATTGCAGATTTGTACCCTGTGTCAAATTT  
CAAGGTATTGCTGATAAACCTTTCCAACCAGCAAGCAAGAAGTTCAAAACTTTCTGTCAAT  
GTAACAGAAAACACTATGTATATAACATTTATGTAGCAATAAATGTGCCATCTTTTT

### C.1.15 VANGL2 3' Region with miR-138 (pink) and miR-499 (green) target sites

GTCGTGCAGGTCCCTGAGCACCAAGCCCCATCCCCGGCACGAGGCCGCTCGGTGCCCGTCC  
CCACGCTCCCACACCCAGCTCCTCGCAGCCATCCCCAGGAGGCACTGCCAGCACAGCACAG  
TGTCACAGCCGGACGGGCAGCATCTCAGTGCCATCCCCTCAAGGGCTGACCTTCGGCTGAG  
TTCTCTG**CACCAGCA**GCCTCGCTCGTGGCCAGGTCTGAAGCATCTCCTGCGTCTTCCTCTT  
CTGCTGCGACCCGCTGCGTCTCCG**GTCTTAA**AGCACCACCTGTGTCTGGACTGGTTGGTAA  
TAACGACACTTTGGGAGTGTTTTGCAAACCCCTGTGATCTGAGGCTGCTATGGAACCTTCC  
CTCCGCCCCACGGCCGGGGCAGACCCACAGGCACTGTCTGCTGCCATGGGAGACCCTCA  
GCTCCCTGCTGCTTTGCATGCAGAGCTCCTCACAGAAGCAGGAGGAGGTCTGATGATGAAG  
CTGAGAAACATCCCAGCGATGTGGGTGCCCAACCCACCCCCAGTCCCAGCACACCCCTC  
ACATAGGGCCCCGTTGCCCTATTCCAACGTGTGCTGGCTGCCAGAACCACAGAATCACTGA  
GATTGGAAGAGATCATTGAGTCCACCCATCTGTCATCGCCACCAACGGGGCTGGTGAGGGG  
CTGTGGGATTTGGGGTGCAGCACCCCGTGGCTTGCAGTGAGGCAGCCAAAGCACACCCAG  
GCCCTGGTGCTGGCTCCAAAGGGCTTCTTCATGGAGACCATGCTCTGTGGGCTCCTCAGCT  
CCTGTCCCATGGAGGTAGTGATGCAGTGACCCCCCAGGGGTGCAGAGCAGGGGGACCCC  
AACACTCAGGGCTGCAGCACAGCTTCTGCCTTCCCTCGGCTCTGTGAGGGCTGCCAATACC  
TCGTGGGGGTTCTGTATTTATCCATTTTTTTATGCTGGGGGTACCCCAAGCCGGGTTTG  
GGCTTCTTGGCCCTGAGAGCTGCCAAACCTTATCACTGTGACACACACATCCCCCCCCAC  
CCGAGGGGCTGTGGGGGCTTCCAGTAAGCACCAGTAGCAGAAAACAGATCCAAGTGGATC  
TCAAGGGCTTTCCAGCACAAAACGGGTGGGATGTAGTGTGTGTGTGGGGGGGGGGGACA  
CCGTCCCAAATTCCAGGGGTTTTTCTGGAAATCTCTTCTATTTTCCAGACTTTCCGTGGC  
TTTGTGACCCCGCTCAGCCCTGTGCAGCTTCCAGTCCCGTTTACACACGCAGCCCTTTGGG  
GCTGGTGCCGACAGGAGCCCAATTTGGGGTTGTTTTACTGCA

### C.1.16 ZEB2 3' Region with miR-130 (blue) and miR-138 (pink) target sites

CCTGCTTGAAAAACACTGCTGTGTTAAGCTGTTTCATGCACGTGCCTGATGCTTCCAGGAAG  
CCTGTAGAGATGGACTAAAGGGGCAGTTCAGCCAAGACAGAATTAGATGGAGTTTGAGCTG  
GGTATTGTTAAGAACTGCATTATGCAAAATTTTTGTACAGTGTTAAGGCCTAAAACTGTG  
TGTTTCCAGAGACTAAATCCTGTGTTTAAATAGCATTTATACTTTAAGCACAACTAGTAAAT  
TGTACGAA**TTGCACT**CGACTTATATATCACTACAACTTAAAAAAAAAAAAAAAAAGATATGTC

TAATTTATATTAATACATTTTAAAAAGGTGCCCGCACTACCATACATCAGTATTATTATTA  
TTATTATTCCTTTTTAATTTAATGTGCTCGCACTACAGTACATCAGTATTATGACTCCTCT  
GTACTTTCCTTTGCTATTCATACGTTTCCCAGTTTTTCAGCCTAAGTAACCACACAATTTTA  
GGCCTCGATTTTTATTTTTATTTTTCTGTGAAGGAAC TTGAAGTGATGCATGTGTGAATTTA  
AGATACCGAAGTCTTAAAGTGACCTGGACGTGAAGGAAAAAGTAAGATGAGAAATAAAGAA  
AGCCTTTGTAAGGTGGTTTTTAAAAGCCTTATATGCAAACCTTTTAATCTGTGTGTTTCTGC  
AAGTGCCATCCTTGTATAGTGTTAAGGGGGTAACGTGTGTTACCTTTG **CACCAGC** TTCAGT  
GTTAAGCTCACCTGTTCTTTGAAGCACCCATGTCAGTATTAGAAGAATAGGCAGCAGTTC  
CTTAGTTTACATATGTTTGTGCAATTTTTTTTCTGTACTTTTTTGTTCATTAATTTTGTGTCAG  
TATTACACCAAACCTTTTTTTTTTTTTTTTTTCCAACAAAAATTTTTTGCATTCATTTAAT  
TTTAGGTCAAATAACATTTTATTTATGTGGCTCATTTTATATTTCCCTAATTTTATTTATTT  
CATACTGTAGTGACAGTATTATAGTTCTTCAATATATAGATATATTTTAGTAAAAAAGGA  
ACATGACGTTGATCATTTGGGCAAATTTTACGTAAAGAGAAGAGCATTTATTGTGTTTTGG  
AACATTAATTGTGAGATGGGATTTTTCAATTTTATTATTTTATTTTTGTTTTTTCCAATTA  
CTGGAAATCCAAATTTGGGAACTTTTGATACGATCTTGTGAAAACACTGTATTTTCGACT  
GAAAATTCCACTTTCTTCATCTTGTTTTT

## C.2 Mutated target sites

The predicted target sites of genes were mutated as shown in red. In each case the predicted target site is shown above, with the mutant site below.

### C.2.1 ATG7 miR-138 mutant

A	C	C	T	G	C	A
C	C	A	T	G	G	

### C.2.2 ATG7 miR-499 mutant

G	T	A	G	T	T	A	A
		A	G	A	T	C	T

### C.2.3 ETS1 miR-499 mutant

A	G	T	T	T	T	A	A
A	G	A	T	C	T		

### C.2.4 FZD8 miR-499 mutant

	G	T	C	T	T	A	A
A	A	G	C	T	T		

### C.2.5 LEF1 miR-130 mutant

T	T	G	C	A	C	T
	A	G	A	T	C	T

### C.2.6 NFIB B miR-130 mutant

T	T	G	A	G	A	A	C	T	
		G	G	A	G	A	T	C	C

**C.2.7 NFIB B miR-138 mutant**

C	A	A	G	C	A		
A	A	A	G	A	T	C	T

**C.2.8 NR3C1 miR-138 mutant**

A	C	C	A	G	C	A	
A	A	A	G	C	T	T	

**C.2.9 PPP3CB miR-499 mutant**

	G	T	C	T	T	A	A
A	A	G	C	T	T		

**C.2.10 PRKACB miR-130 mutant**

T	T	G	C	A	C	T	
		G	C	T	A	G	C

**C.2.11 ROCK2 miR-130 mutant**

T	G	C	A	C	T	A	
A	G	A	T	C	T		

**C.2.12 TCF12 miR-138 mutant**

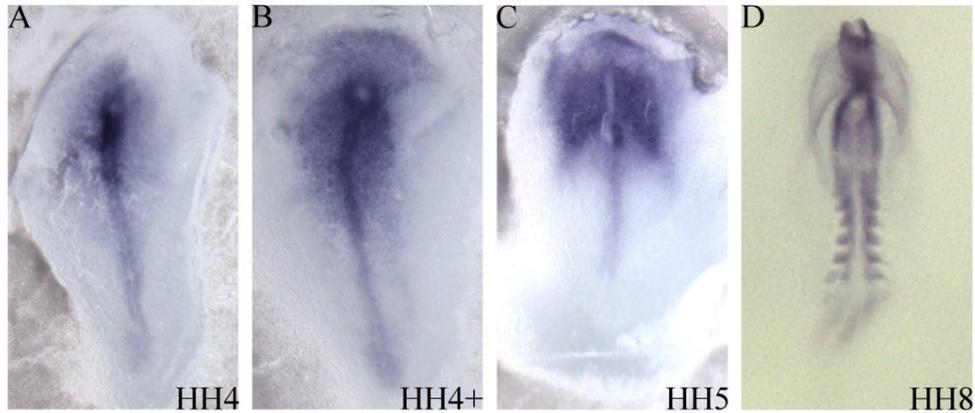
A	C	C	A	G	C	A		
			A	G	A	T	C	T

**C.2.13 ZEB2 miR-130 mutant**

T	T	G	C	A	C	T	
		G	G	A	T	C	C

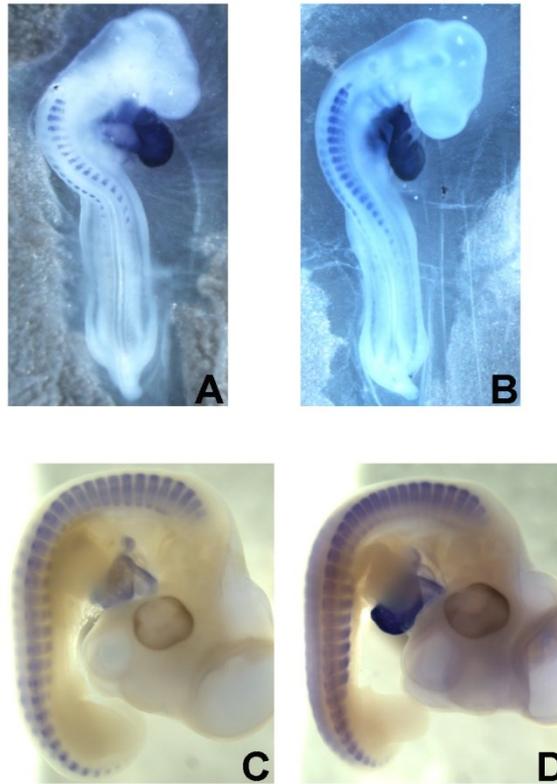
## Appendix D

### D.1 Expression of FZD7 mRNA during early stages of chicken development

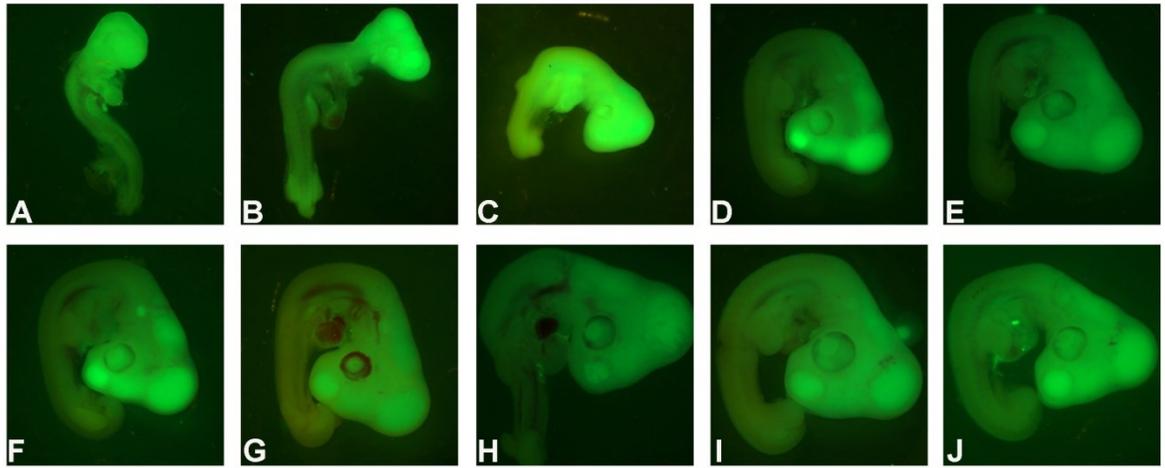


**Figure Apx D.1: FZD7 expression in the chick during early stages of development.** FZD7 mRNA whole mount in situ hybridization (WISH) of chick embryos at stage HH4 (A), HH4+ (B), HH5 (C) and HH8 (D). FZD7 is expressed in the primitive streak posterior to Hensen's node at HH4 and HH4+ (A, B), and also in the lateral plate mesoderm at HH5 (C). By HH8 FZD7 is expressed in the cardiogenic and paraxial mesoderm (D).

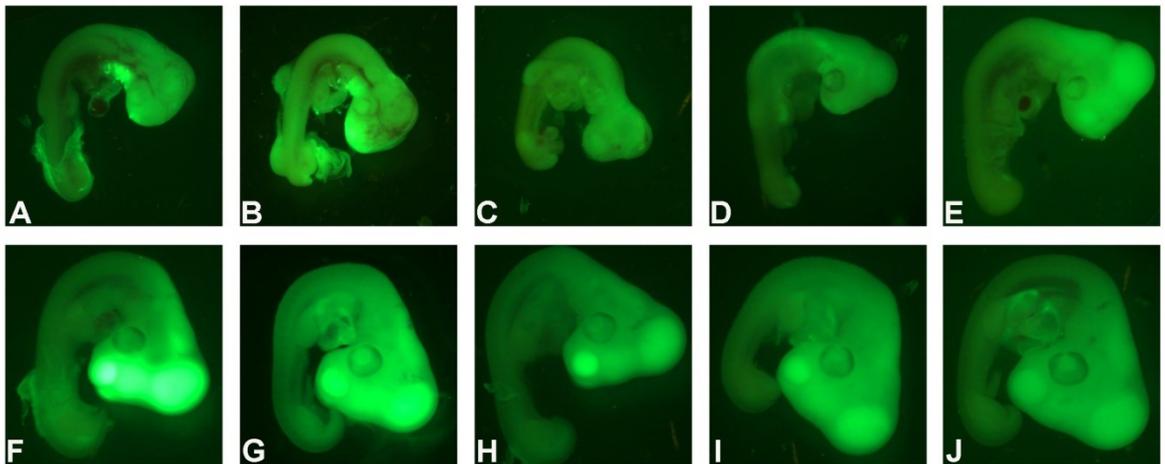
## D.2 Exploration of miR-1 dysregulation during heart development



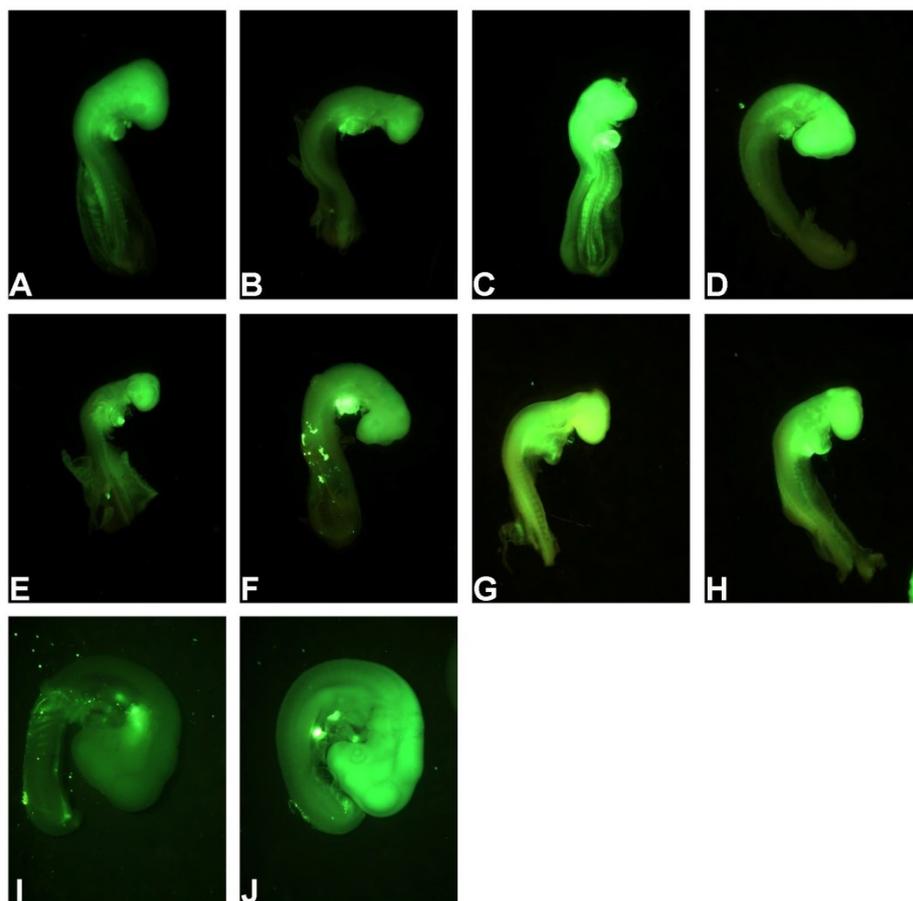
**Figure Apx D.2: AM1 and Scr206 injections into the HH11 and HH20 heart in ovo did not result in a developmental delay as observed when miR-1 was inhibited at HH14. A, B:** Microinjection into the HH11 chick heart with AM1 (A) or Scr206 control (B) did not appear to have a delaying effect on embryo maturation. However, due to the heart being positioned between the rest of the embryonic body and the yolk very few embryos were successfully injected with AM1 or Scr206 at this stage (N=3, n=3). **C, D:** Embryos injected with AM1 (C) or Scr206 (D) at HH20 developed at the same rate.



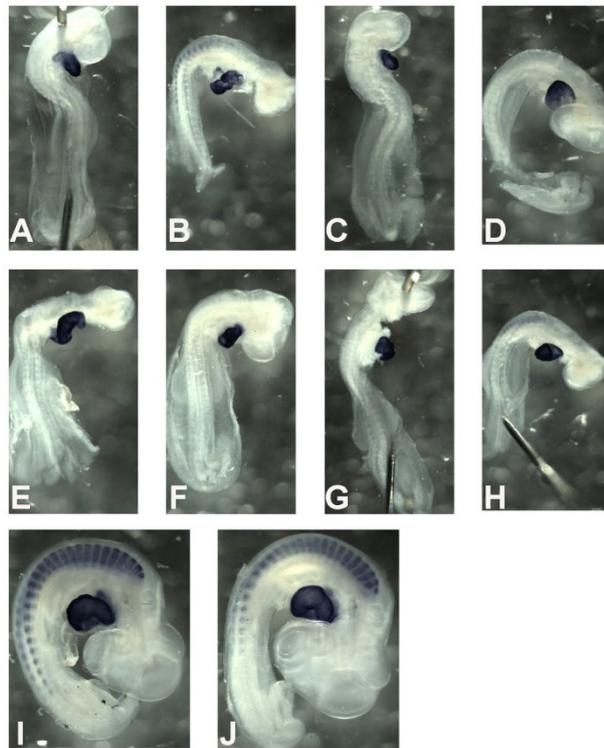
**Figure Apx D.3: AM1 cardiac microinjections at HH20 did not result in developmental delay.** AM1 microinjection into HH20 hearts *in ovo* resulted in embryos that all appeared about the same stage, but with 3 embryos (**A - C**) much smaller than the remaining seven embryos (**D - J**). All pictures show the entire field of view at 1.6x magnification and were taken immediately after harvest to confirm that antagomiR had successfully been injected into the embryonic heart. After injection embryos were left to develop for 24 hours.



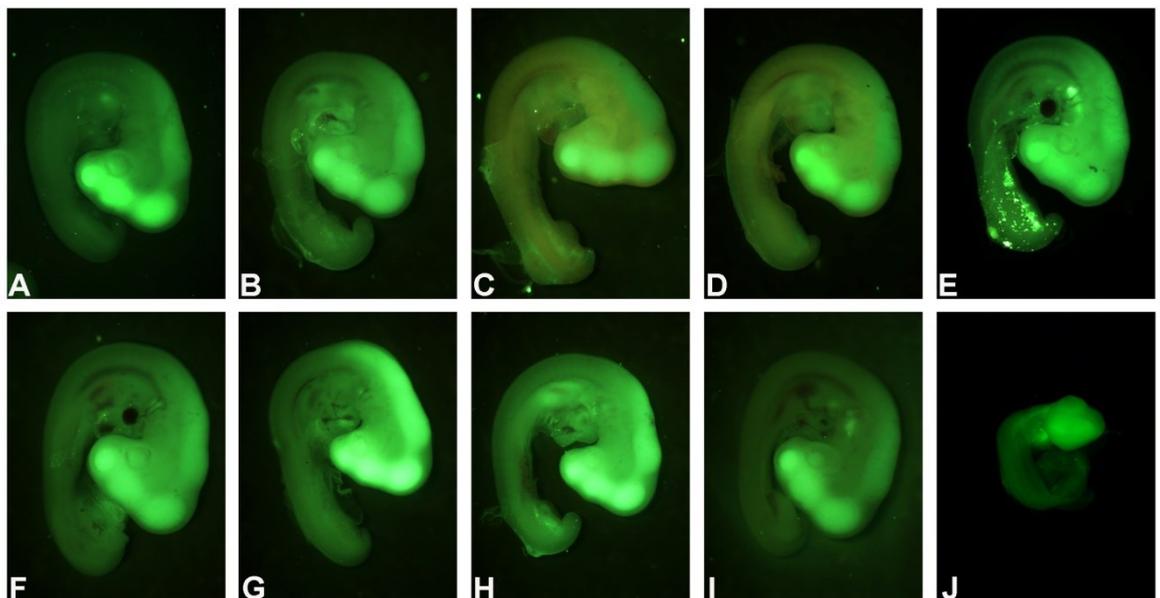
**Figure Apx D.4: Scr206 microinjections at HH20.** Three embryos (**A - C**) were found to be at a similar developmental stage to the majority (**D - J**) despite their smaller size, following Scr206 microinjections in HH20 hearts *in ovo*. Pictures were all taken at 1.6x magnification and the entire field of view is displayed. Embryos were left to develop for 24 hours after injection and the imaged immediately upon harvest to confirm successful injection of the FITC-labelled antagomiR.



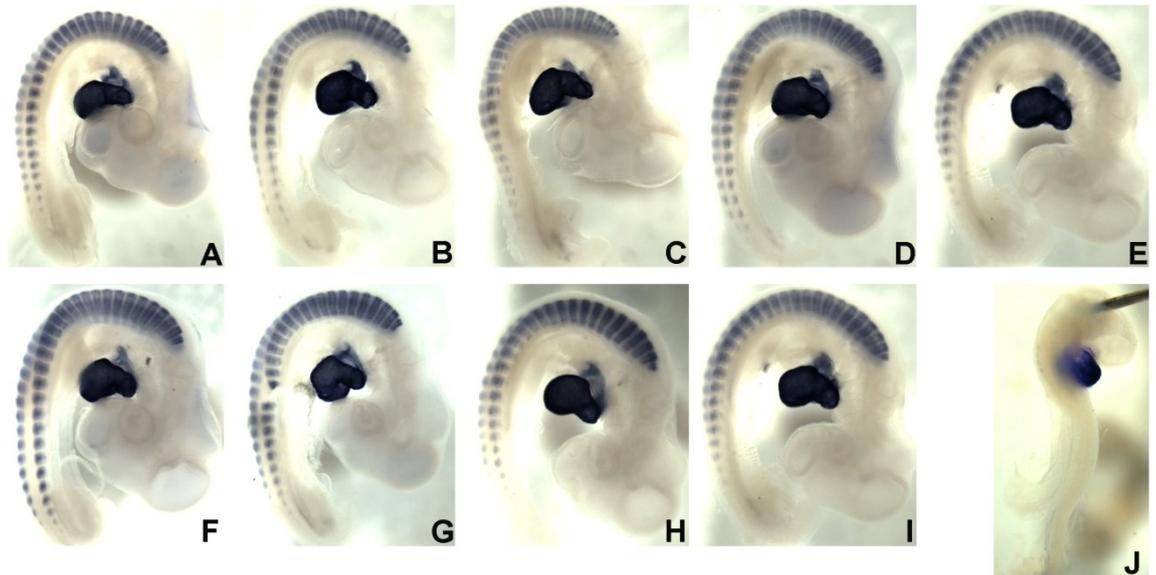
**Figure Apx D.5: AntagomiR-1 distribution *in ovo*.** AntagomiR-1 is conjugated to a FITC tag and successful microinjection into the beating heart *in ovo* results in the wide distribution of the miR-1 inhibitor. All pictures show the entire field photographed at 3.2x magnification and therefore size differences seen in this grid between smaller embryos (A – H) and larger embryos (I, J) is not due to magnification differences and reflect actual size differences between embryos.



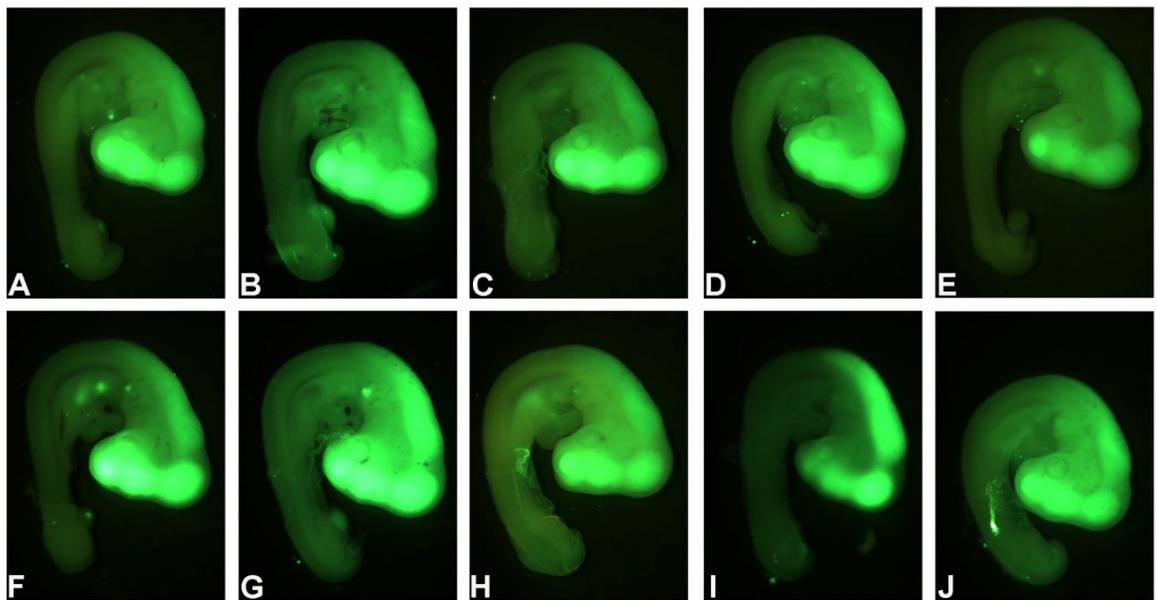
**Figure Apx D.6: Developmental delay was observed following inhibition of miR-1 at HH14.** AntagomiR-1 microinjected embryos (Figure Apx D.2) were probed for *Myh15* mRNA. The majority of miR-1-inhibited embryos were found to be at or younger than HH19. Embryos were injected *in ovo* at HH14 with 0.5mM antagomiR-1 and left to develop for 24 hours. Only two embryos (I, J) had reached the expected developmental stage (HH20) or older stages, with the remaining embryos developmentally delayed (A – H).



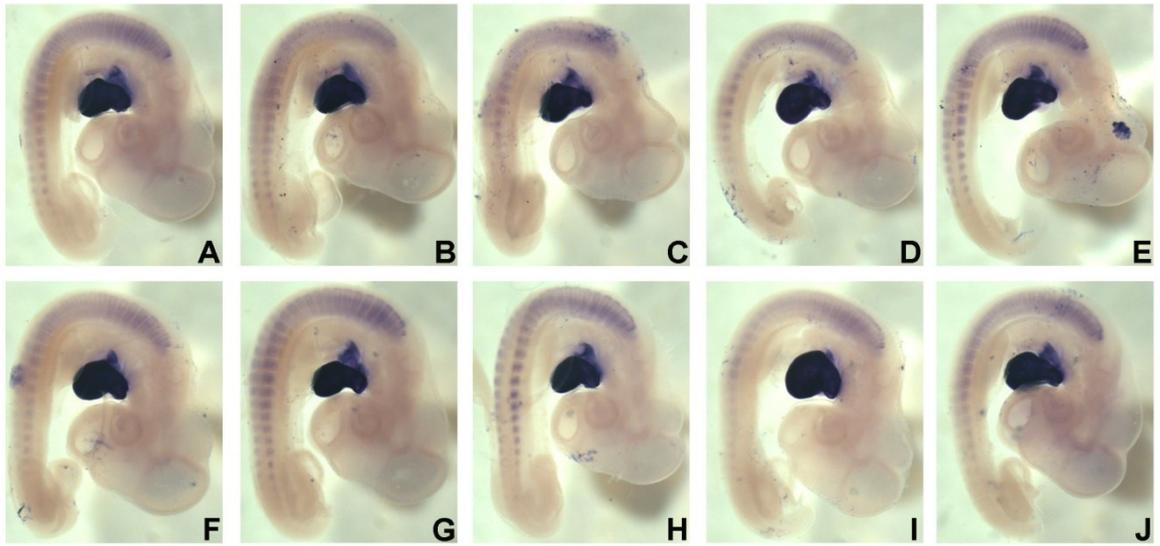
**Figure Apx D.7: Wide distribution of the control antagomiR, Scr206, was seen after *in ovo* microinjection into HH14 chick hearts.** The antagomiR is FITC labelled and embryos were imaged immediately after harvesting to confirm successful injection. All pictures show the entire field photographed at 3.2x magnification and almost all the embryos appear of a similar size (A – I) apart from one (J).



**Figure Apx D.8: Scr206 injected embryos did not display developmental delay.** *MyH15* WISH was performed on control (Scr206) microinjected embryos (**Figure Apx D.7**). Almost all the embryos were found to be at or older than HH20 (**A – I**), apart from one (**J**). *In ovo* injections with 0.5mM control antagomiR (Scr206) were performed at HH14 and embryos were left to develop for 24 hours.



**Figure Apx D.9: Microinjection with FITC labelled antagomiR-206 (AM206) into HH14 hearts *in ovo* resulted in a distribution of the antagomiR throughout the embryos.** Embryos were imaged as soon as they were harvested to confirm successful injection of AM206. The photographs were taken at 3.2x magnification and the entire field of view is displayed, giving an indication of relative embryo size. All embryos appear to be broadly similar in size (**A – J**).



**Figure Apx D.10: Inhibition of miR-206 at HH14 did not result in developmental delay.** AM206 microinjected embryos (Figure Apx D.9) were probed for MyH15 mRNA. The embryos were all found to have reached at least HH20 (A – J). The 0.5mM AM206 antagomiR was microinjected into the heart at HH14 and embryos were incubated for a further 24 hours following this treatment.

## Appendix E

**Table Apx E.1: Hamburger Hamilton stages of chicken embryo development.**

<b>HH stage</b>	<b>Age (hours)</b>	<b>HH stage</b>	<b>Age</b>
<b>1</b>	5 hrs	<b>15</b>	50 – 55 hrs
<b>2</b>	6 – 7 hrs	<b>16</b>	51 – 56 hrs
<b>3</b>	12 – 13 hrs	<b>17</b>	52 – 64 hrs
<b>4</b>	18 – 19 hrs	<b>18</b>	3 days
<b>5</b>	19 – 22 hrs	<b>19</b>	3 – 3.5 days
<b>6</b>	23 – 25 hrs	<b>20</b>	3 – 3.5 days
<b>7</b>	23 – 26 hrs	<b>21</b>	3.5 days
<b>8</b>	26 – 29 hrs	<b>22</b>	3.5 – 4 days
<b>9</b>	29 – 33 hrs	<b>23</b>	4 days
<b>10</b>	33 – 38 hrs	<b>24</b>	4.5 days
<b>11</b>	40 – 45 hrs	<b>25</b>	4.5 – 5 days
<b>12</b>	45 – 49 hrs	<b>26</b>	5 days
<b>13</b>	48 – 52 hrs	<b>27</b>	5 – 5.5 days
<b>14</b>	50 – 53 hrs	<b>28</b>	5.5 – 6 days