The role of microRNA-31 in cardiac fibrosis associated with atrial fibrillation



A thesis submitted for the degree of Doctor of Philosophy

Michaelmas term 2020

Jasha Winston Trompf

Lincoln College

Radcliffe Department of Medicine, University of Oxford

Declaration

I declare that the work presented in this thesis is my own unless otherwise stated.

Acknowledgements

My supervisors Svetlana Reilly and Craig Lygate made an enormous effort to guide me through this project while also giving me space to pursue other ideas and interests. Their mentoring proved to be invaluable and I could not have received better supervision. Thank you Svetlana and Craig.

I would also like to thank Keith Channon and his lab for always being approachable and for excellent coaching in experimental design, and I thank Ricardo Carnicer for providing clear oversight and purpose to the project. Thank you Peter Hinton for your essential role in managing this project. I also thank members of the Reilly lab, Lucia Moreira, Neelam Mehta and Kathryn Aguilar-Agon for helping me get through the final stages of this work.

Table of Contents

Declaration ..... **i**

Acknowledgements ..... **ii**

Table of Contents ..... **iii**

Figures ..... **vii**

Tables ..... **xi**

List of Abbreviations ..... **xii**

Abstract ..... **xvi**

Chapter 1 Introduction **1**

1.1 Clinical definition and epidemiology of atrial fibrillation **2**

1.2 Aetiology of atrial fibrillation **3**

1.2.1 Atrial cellular electrophysiology and electrical remodelling 4

1.2.2 Structural remodelling 6

1.3 Cardiac fibroblasts and extracellular matrix **8**

1.3.1 Cell signalling leading to cardiac fibrosis 9

1.3.2 Markers for healthy and disease state cardiac fibroblasts 12

1.4 The role of microRNAs in structural remodelling **15**

1.5 miR31 in atrial fibrillation **18**

1.5.1 Transcriptional regulation and genesis of microRNA31 19

1.6 Project aims and hypotheses **22**

Chapter 2 Methods and Materials **23**

2.1 Gene cloning and vector constructs **23**

2.1.1 Cloning and purification of ETS1 expression vector 23

2.1.2 Identification and validation of putative ETS1 binding sites within the wild type hsa-LOC554202 promoter 26

2.1.3 Generating wild type hsa-LOC554202 promoter *Firefly* Luciferase and mutagenesis constructs....28

2.2 Tissue culture experiments....30

2.2.1 Receipt, expansion and storage of hAFs....30

2.2.2 Preparing hAFs for experimentation....31

2.2.3 Lipofection of hAFs with synthetic miR31 inhibitor and mimic for subsequent RNA and protein analyses....32

2.2.4 BrdU cell proliferation assay in hAFs....32

2.2.5 Picro Sirius red assay for collagen detection in hAFs....33

2.2.6 Knockdown of ETS1 with locked nucleic acid in hAFs....34

2.2.7 Overexpression of ETS1 expression vector in hAFs....34

2.2.8 Culture of HEK293 cells....34

2.2.9 Dual-Glo firefly luciferase reporter assay in HEK293 cells....34

2.3 Immunofluorescence and confocal microscopy....35

2.4 Flow cytometry....36

2.5 *In vivo* experiments with mice....37

2.5.1 Genotyping....37

2.5.1.1 *Collagen1a2-Cre-ER(T)* transgene genotyping....37

2.5.1.2 Floxed pre-miR31-Inv transgene genotyping....39

2.5.1.3 Breeding strategies....41

2.5.2 Tamoxifen preparation and administration....41

2.5.3 Checking for genomic inversion of the *pre-miR31-Inv* transgene....42

2.5.4 Assessment of murine cardiac function with echocardiography....43

2.5.4.1 Anaesthesia....43

2.5.4.2 Echocardiography....43

2.5.5 Harvesting mouse tissue....44

2.5.6 Hydroxyproline assay....44

2.6 RNA extraction, cDNA synthesis and qPCR 45

2.7 RNA sequencing 47

2.8 Protein extraction and immunoblotting 48

2.9 Patient Cohorts and atrial fibroblast isolation 49

2.10 Experimental design and statistics 50

2.11 Materials and reagents 52

2.12 Primers 57

2.13 Contributions from other lab members 59

Chapter 3 Results: Investigating the effects of miR31 manipulation in hAFs 60

3.1 Flow cytometry and immunofluorescence validate hAFs 62

3.2 Validation of tools for manipulation of miR31 in hAFs 69

3.3 Gain- and loss-of miR31 alters the poly-A tail transcriptome of hAFs .. 71

3.4 Gain- and loss-of miR31 alters mRNA abundance of fibrotic genes 76

3.4.1 Validation of RNA-seq results using RT-qPCR 85

3.5 MiR-31 abundance effects fibrotic phenotype in hAFs *in vitro* 90

3.5.1 Assessing cell proliferation 90

3.5.2 Assessing collagen synthesis 94

3.6 Discussion and plans for future experiments 97

Chapter 4 Results: Investigating transcriptional regulation of miR31 by ETS1 in human atrial fibroblasts in vitro 102

4.1 The effects of loss- and gain-of ETS1 on miR31 transcripts in hAFs ... 106

4.2 Mutagenesis of ETS1 binding sites at the LOC554202 promoter 117

4.3 Discussion and plans for future experiments 124

Chapter 5 Results: Validation of a novel transgenic mouse and consequences of in vivo up-regulation of miR31 in fibroblasts 127

5.1 *Col1a2-Cre-ER(T)* litters show Mendelian inheritance of transgene 133

5.1.1 Optimisation of genotyping suggests GelRed retards DNA migration 135

5.2 *Pre-miR31-Inv* litters show Mendelian inheritance of transgene 138

5.3 Crossing *Col1a2-Cre-ER(T)* *with* *pre-miR31-Inv* transgenic mice generates tamoxifen-inducible fibroblast-specific miR31 mice 140

5.4 Dose-dependent response of miR31 in murine atria by varying tamoxifen treatment regimens 143

5.5 Tamoxifen induces inversion of the *pre-miR31-Inv* transgene at the *Gt(ROSA26)Sor* locus 145

5.6 Increased miR31 expression is genotype specific 150

5.7 Evaluating the effects of miR31 overexpression on cardiac fibrosis and function in pre-miR31-Inv transgenic mice 153

5.7.1 Validation of echocardiography image analysis against training data. 154

5.7.2 Mice carrying *Col1a2-CreER(T)* or *pre-miR31-Inv* do not show changes in systolic heart function before tamoxifen administration 159

5.7.3 Tamoxifen affects systolic heart function independent of genotype ... 161

5.7.4 Increased miR31 expression does not affect systolic heart function ... 165

5.7.5 Increased miR31 alters mRNA and protein abundance of selected fibrogenesis related genes in mouse atria 169

5.7.6 Increased miR31 abundance results in elevated total atrial collagen... 181

5.8 Discussion and plans for future experiments 182

Chapter 6 Discussion 184

6.1 Conclusion 191

References 192

Appendices 215

1. hAF donor 1 certification 215

2. hAF donor 2 certification 217

3. RNA-seq quality control 219

4. 477 predicted mRNA targets of miR31 by TargetScan7.2 226

Figures

Figure 2.1 Workflow used to generate ETS1 expression vector 25
Figure 2.2 ETS1 DNA-binding motifs 27
Figure 2.3 Workflow used to generate WT LOC554202 promoter Firefly Luciferase vector and mutagenesis constructs 29
Figure 2.4 Workflow diagram showing receipt, expansion and storage of hAFs ... 30
Figure 2.5 Workflow diagram showing different experiments performed in hAFs 31
Figure 3.1 miR31 abundance in fibroblasts derived from the right atrial appendage from patient donors 60
Figure 3.2 Immunostaining validates presence of intracellular proteins found in cardiac fibroblasts 63
Figure 3.3 Flow cytometry gating strategy used on negative control hAF cell population (i.e. unstained cells) 65
Figure 3.4 hAFs stained with anti-CD45, -CD31, -CD90, -PDGFRα and –VAP1 antibodies. 67
Figure 3.5 miR31 expression in hAFs following enrichment or depletion of miR31 with small oligonucleotides 70
Figure 3.6 RT-qPCR measurements of miR31 expression in hAFs following enrichment or depletion of miR31 with small oligonucleotides 73
Figure 3.7 116 significantly altered cell signalling networks determined with Ingenuity Pathway Analysis in miR31 mimic sequencing dataset 82
Figure 3.8 Significantly altered cell signalling networks determined with Ingenuity Pathway Analysis in miR31 inhibitor sequencing dataset 84
Figure 3.9 Venn diagrams showing the number of genes significantly changed in miR31 mimic and inhibitor datasets 85
Figure 3.10 QPCR of putative direct miR31 targets in hAFs validates RNA-seq 88
Figure 3.11 Donor 1 hAF cell proliferation following enrichment or depletion of miR31 with small oligonucleotides 91
Figure 3.12 miR31 expression in hAF donor 2 (18TL312486) following enrichment or depletion of miR31 with small oligonucleotides 92
Figure 3.13 Cell proliferation in donor 2 hAFs (18TL312486) following enrichment or depletion of miR31 with small oligonucleotides 93

Figure 3.14 The effects of TGFβ1 on cell proliferation in donor 2 hAFs 94
Figure 3.15 hAF cell collagen production following enrichment or depletion of miR31 with small oligonucleotides 95
Figure 3.16 hAF cell collagen production following enrichment or depletion of miR31 with small oligonucleotides 96
Figure 4.1 pri-miR31 abundance in human right atrial appendage 102
Figure 4.2 5'-3' 1061 bp segment upstream of exon-1 of the long non-coding RNA LOC554202 with eleven identified putative ETS1 binding sites 104
Figure 4.3 Immunoblotting of ETS1 in human RAA protein lysates 105
Figure 4.4 Dose and time-course experiments of ETS1 expression in primary hAFs following lipofection with anti-ETS1 LNA 107
Figure 4.5 ETS1 depletion alters LOC554202 transcription in hAFs 108
Figure 4.6 ETS1 depletion alters miR31 abundance in hAFs 109
Figure 4.7 NotI/KpnI digested and gel purified pcDNA3.1(+) 110
Figure 4.8 Colony PCR of DH10-beta cells following transformation with ETS1 ORF ligated into pcDNA3.1(+) vector 111
Figure 4.9 Dose experiment of ETS1 expression vector in hAF cells 113
Figure 4.10 ETS1 enrichment alters miR31 abundance in hAFs 114
Figure 4.11 Visualising ETS1 in primary hAF cells by immunofluorescence 115
Figure 4.12 Trialling lower doses of ETS1 expression vector in hAF cells 116
Figure 4.13 Mutagenesis constructs of WT LOC554202 promoter Firefly Luciferase vector 118
Figure 4.14 qPCR analysis and immunoblotting of time-course and dose experiment of ETS1 expression vector in HEK293 cells 120
Figure 4.15 Visualising expression and localisation of ETS1 in HEK293 cells by immunofluorescence 121
Figure 4.16 Luciferase assay in HEK293 cells 123
Figure 5.1 Principle of Cre-recombinase dependant inversion switch at the targeted Gt(ROSA26)Sor locus 131
Figure 5.2 Expression of miR31 following Cre-dependant inversion at the targeted Gt(ROSA26)Sor locus 132
Figure 5.3 Col1A2-Cre-ER(T) example litter genotyping 134
Figure 5.4 GelRed retards DNA migration in genotyping gels 137

Figure 5.5 Example gels from various breeding strategies for pre-miR31-Inv transgenic mice colony 139
Figure 5.6 Crossing Col1a2-Cre-ER(T)$^{(Hemi)}$ with pre-miR31-Inv$^{(+/-)}$ strains produce pups carrying both transgenes 142
Figure 5.7 Pilot experiment showing miR31 expression in different tissues following varying intraperitoneal tamoxifen regimens in Col1a2-Cre-ER(T)$^{(Hemi)}$ / pre-miR31-Inv$^{(+/-)}$ mice 144
Figure 5.8 Genomic DNA PCR analysed for inversion, recombination and trapping at the 5' end of the pre-miR31-Inv allele at Gt(ROSA26)Sor locus.... 147
Figure 5.9 Inversion, recombination and trapping at 3' end of pre-miR31-Inv 149
Figure 5.10 Checking for unwanted inversion pre-tamoxifen treatment 149
Figure 5.11 Work flow diagram of tamoxifen administration in mice 150
Figure 5.12 miR31 fold change in various tissues of interest 151
Figure 5.13 eGFP mRNA transcripts are expressed in pre-miR31-Inv $^{(+/-)}$ / Cre $^{(Hemi)}$ post-tamoxifen 152
Figure 5.14 Workflow diagram showing timings of echocardiography measurements, tamoxifen administration and tissue harvest 153
Figure 5.15 Example murine short-axis echocardiograms 154
Figure 5.16 Comparison of expert and student analyses of echocardiographic images visualised using correlation and Bland-Altman plots 157
Figure 5.17 Echocardiographic data re-analysed following refinement of measurement technique 159
Figure 5.18 Echocardiographic parameters pre-tamoxifen 161
Figure 5.19 Before-after plots showing changes in mouse weight during tamoxifen treatment protocol 162
Figure 5.20 Before-after plots showing the effects of tamoxifen on echocardiographic readouts within genotype groups. 165
Figure 5.21 Echocardiographic parameters post-tamoxifen 166
Figure 5.22 Raw qPCR values of housekeeping gene transcripts in mouse atria 170
Figure 5.23 Mouse atrial mRNA and protein changes of genes predicted as direct miR31 targets from hAF RNA-seq and Targetscan7.2 173
Figure 5.24 Mouse atrial mRNA and protein changes of genes that decreased in hAF RNA-seq following increased miR31 but didn't necessarily change after miR31 depletion 174

Figure 5.25 Mouse atrial mRNA and protein changes of genes that increased in hAF RNA-seq following increased miR31 but didn't necessarily change after miR31 depletion.....175
Figure 5.26 mRNA fold changes in mouse atria of candidate pro-fibrotic genes177
Figure 5.27 Protein changes in mouse atria of candidate pro-fibrotic genes.....178
Figure 5.28 LV and atria collagen content 181

Tables

Table 2.1 ETS1 ORF PCR reaction mix and thermocycling 24

Table 2.2 Col1a2-Cre-ER(T) genotyping reaction mix and thermocycling 38

Table 2.3 pre-miR31 transgene allele genotyping conditions for 5' and 3' ends .. 40

Table 3.1 Flow cytometry event count with gating in negative control hAFs..... 66

Table 3.2 Flow cytometry event count following gating in hAFs probed for surface markers CD45, CD31, CD90, PDGFRα and VAP1 68

Table 3.3 RNA Nanodrop absorbance readings from all technical replicates in miR31 enrichment and depletion experimental repeats..... 72

Table 3.4 Stratification of sequencing data..... 75

Table 3.5 Manual screen for pro-fibrotic genes of interest..... 76

Table 3.6 Ingenuity pathways with relevance to cardiovascular fibrosis with significantly altered genes..... 84

Table 3.7 Genes significantly changed in both sequencing datasets. Underlined genes are predicted as direct miR31 targets by TargetsScan7.2 86

Table 5.1 Col1a2-Cre-ER(T) colony maintenance genotyping tally 134

Table 5.2 pre-miR31-Inv colony maintenance genotyping tally 140

Table 5.3 pre-miR31-Inv X Col1a2-Cre-ER(T) genotyping tally 142

Table 5.4 Echocardiographic readouts of all mice used in this study expressed as mean ± SD..... 168

Table 5.5 Summary of all genes explored in mouse atria and relevant hAF RNA-seq data..... 179

List of Abbreviations

ACTA2 aka αSMA	Actin alpha 2 aka Alpha smooth muscle actin
ACTB	Actin beta
ACTN2	Actinin alpha 2
AF	Atrial fibrillation
AngII	Angiotensin II
APD	Action potential duration
AT1R	Angiotensin II type 1 receptor
BAP1	BRAC1 associated protein 1
BMP	Bone morphogenetic protein
BrdU	Bromodeoxyuridine
BSA	Bovine serum albumin
CaMKII	Calcium/calmodulin dependent kinase II
ChIP	Chromatin immunoprecipitation
CIP	Calf intestinal alkaline phosphatase
CMV	Human cytomegalovirus immediate early promoter
CRP	C reactive protein
CTGF	Connective tissue growth factor
DAPI	4′,6-diamidino-2-phenylindole
DDR2	Discoidin domain receptor 2
ECL	Enhanced chemiluminescence
ECM	Extracellular matrix
ENCODE	Encyclopaedia of DNA Elements
EMSY	The childhood moniker of Emma Whitley
ERK	Extracellular regulated kinase
ETS1	Erythroblastosis transcription factor 1
FBS	Foetal bovine serum

FGF1	Fibroblast Growth Factor 1
FGM-3	Fibroblast growth medium-3
FLNA	Filamin A
FN1	Fibronectin
FSC-A	Forward scatter area
FSP1	Fibroblast specific protein 1
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
GATA5	GATA binding protein 5
Gt(ROSA26)Sor	Gene trap reverse orientation splice acceptor 26, Philippe Soriano
hAF	Human atrial fibroblast
HEK293	Human embryonic kidney cells 293
HEM / Hemi	Hemizygous
HG-miR31	non-coding splicing product from LOC554202
HIF1α	Hypoxia inducible factor 1 alpha
HPRT1	Hypoxanthine phosphoribosyl transferase 1
I_{CaL}	Ca^{2+} current, L-type channel
I_{K1}	K^+ current, inward rectifier
I_{KR}	K^+ current, rapid delayed rectifier
I_{KS}	K^+ current, slow delayed rectifier
I_{KTO}	K^+ current, transient outward
I_{KUR}	K^+ current, ultra-rapid delayed rectifier
I_{Na}	Na^+ current, voltage gated
ITGB3	Integrin beta 3
JAK	Janus kinase
JMJD8	Jumonji Domain Containing 8
JNK	c-Jun N-terminal kinase
KLF5	Kruppel like zinc finger transcription factor five
KDM5B	Lysine demethylase 5B

K-S	Kolmogorov–Smirnov statistical test
LAP	Latency associated propeptide
LB broth	Luria Bertani broth
LNA	Locked nucleic acid
LOC554202	Long noncoding RNA 554202
LOX	Lysyl oxidase
LTBP1	latent TGFβ binding protein 1
MAP	Mean arterial pressure
MAPK	Mitogen activated protein kinase
MCS	Multiple cloning site
miR31	microRNA-31-5p
MMP	Matrix metalloproteinase
NF-κB	Nuclear factor kappa-light-chain-enhancer of activated B cells
p38	Sub family of MAPKs
pcDNA3.1(+)_ETS1_WT	ETS1 expression vector
PDGFA	Platelet derived growth factor subunit A
PDGFB	Platelet-derived growth factor subunit B
PDGFRα	PDGF receptor alpha
pGL4[*Luc4*]	Encodes the luciferase reporter gene luc2 (*Photinus pyralis*)
PKN2	Protein Kinase N2
POSTN	Periostin
PPIA	Peptidylprolyl isomerase A
Pre-miR31	precursor-microRNA-31
PSR	Picro sirius red
RISC	RNA induced silencing complex
RNU58A	RNA, U58A small nuclear
RS18	Ribosomal protein 18S
RyR	Ryanodine receptor

SCA1	Stem cell antigen 1
SMAD	Mothers against decapentaplegic homolog
SPRY1	Sprouty 1 RTK signalling antagonist
SSC-A	Side scatter Area
STAT	Signal transducers and activators of transcription
STK40	Serine/threonine kinase 40
TAC	Transverse aortic constriction
TAK1	TGFβ activated kinase 1
TBP	TATA box binding protein
TBX18	T-box 18
TCF21	transcription factor 21
TGFβ	Transforming growth factor β
Thy1	Thymus-cell antigen 1 (CD90)
TIE2	Tunica interna endothelial cell kinase 2
TIMP	Tissue inhibitor of metalloproteinase
TNFα	Tumour necrosis factor alpha
TNS1	Tensin 1
TSS	Transcription start site
TβRI	TGFβ receptor type I
TβRII	TGFβ receptor type II
U6	RNA, small nucleolar U6
VAP1	Vascular adhesion protein
VEGFB	Vascular Endothelial Growth Factor B Subunit
VEGFR1	Vascular Endothelial Growth Factor receptor 1
WT1	Wilms' tumour suppressor gene 1

Abstract

Cardiac fibrosis is a histological hallmark of atrial arrhythmogenic structural remodelling characterised by increased deposition and accumulation of fibrillar collagen. Despite recent advances in understanding cardiac fibrogenesis, there is a paucity of effective anti-fibrotic therapies. Given the robust effect that single microRNAs can have on gene transcript abundance, these small oligonucleotides have been recognised as potential targets for therapeutic manipulation in a wide range of disease states. Our group has recently shown that increased miR31 expression in cardiomyocytes contributes to electrical remodelling begetting atrial fibrillation and our more recent unpublished data show that miR31 is also upregulated in human atrial fibroblasts derived from patients with chronic AF compared to sinus rhythm controls. Given that miR31 has been shown to drive pro-fibrotic phenotype changes in fibroblasts in other disease models, including ischaemic heart disease, we hypothesised that the observed increase in atrial fibroblast miR31 expression in chronic AF was contributing to fibrotic changes seen in this disease state. Consequently, I conducted experiments whereby miR31 abundance was either enriched or depleted in primary human atrial fibroblasts obtained from healthy patient donors to observe whether manipulation of miR31 *in vitro* would cause pro-fibrotic gene expression- and phenotype-changes. Results did show significant changes in mRNA of key pro-fibrotic signalling genes however inter-donor variability made it difficult to draw robust conclusions on the role of miR31 in fibroblast phenotype. Additional *in vitro* experiments in human atrial fibroblasts were conducted to investigate whether the transcription factor ETS1 is involved in regulating miR31 expression and results did show increased miR31 abundance following ETS1 depletion. Further, the consequence of miR31 up-regulation on fibrogenesis *in vivo* was explored in a novel, tamoxifen inducible, global fibroblast-specific miR31 transgenic mouse strain. Results of these mouse experiments demonstrated that increased fibroblast production of miR31 *in vivo* leads to similar pro-fibrotic gene expression and protein changes observed in human atrial fibroblasts thereby corroborating the separate experimental systems used. Overall, the results of this project are the preliminary evidence that miR31 is involved in the pro-fibrotic response in human atrial fibroblasts, as well as in the intact mouse, and provide the foundation upon which further experiments can be performed and justified.

Chapter 1

Introduction

The rhythm of cardiac muscle contraction is maintained intrinsically by specialised myocytes called pacemaker cells that are capable of spontaneous cycling of their membrane potential. The sinoatrial node is the primary pacemaker of the heart and generates electrical impulses that propagate uniformly through the atria, the atrioventricular node and His-Purkinje system to the ventricles (Monfredi et al. 2010). The organised spread of these electrical signals is essential for coordinated contraction and relaxation of the atria and ventricles which ensures efficient coupling between these separate chambers. Perturbations in conduction through these tightly regulated circuits can lead to irregular and/or incomplete cardiac contractions known as arrhythmias (Chen et al. 2017). Atrial fibrillation is the most common arrhythmia and is thought to occur when electrical impulses or triggers, which arise ectopically around the atria and surrounding structures, initiate disorganised conduction patterns in vulnerable tissue substrates (Lip et al. 2016). The genesis of these tissue substrates involves a complex interplay between different cell types, including resident myocytes and fibroblasts, and investigations demonstrate fibrotic structural changes characterised by increased accumulation of fibrillar collagen (Yue et al. 2011). This chapter covers what is understood about cardiac fibrosis and the development of atrial fibrillation with particular emphasis on microRNA-31-5p (miR31) and also specifies gaps in knowledge that need further scientific investigation. Project aims and hypotheses are then stated as well as experimental approaches used.

1.1 Clinical definition and epidemiology of atrial fibrillation

Atrial fibrillation (AF) is an arrhythmia characterised by temporally- and spatially-irregular electrical activation of the atrial chambers resulting in incomplete and rapid atrial contraction (400-600 bpm) (Prystowsky et al. 2017). The perturbed electrical activity of the atria is detectable through standard surface ECG as multiple, low amplitude, and abnormally shaped P waves, often referred to as fibrillatory waves or oscillations (Park et al. 2019). The abnormal activation of the atria propagates electrical impulses through the atrioventricular node and ventricles leading to disorganised ventricular depolarisations (Piccini et al. 2017). This too is detectable with ECG as irregularly-irregular QRS complexes and leads to compromised atria-ventricle coupling thereby impeding ventricular filling and cardiac output (Michaud et al. 2018). Clinically AF is classified as paroxysmal if it self-terminates in <7 days, and persistent if it lasts >7 days. AF that continues for 1 year is defined as longstanding persistent, and if patients continue to show repeated ECG that is not amenable to treatment after one year, AF is defined as permanent or chronic (Heijman et al. 2014, Lip et al. 2016). Most epidemiological studies and clinical data are derived from patients with permanent AF (Kirchhof et al. 2016) and therefore the disease burden described hereafter likely underestimates the true prevalence and clinical concern of AF.

AF is the most common sustained arrhythmia and is estimated to affect 1 – 1.5% of the general population worldwide (Chugh et al. 2014, Lip et al. 2012). Age is the salient risk factor for this disorder and in the UK, AF affects 2% of the population aged over 45 years and 4% aged over 65 years (Davis et al. 2012). The prevalence of AF is likely to triple within the next two-three decades (Go et al. 2001, Miyasaka et al. 2006) which is attributed to an ageing population as well as the growing prevalence of other AF risk factors (Andrade et al. 2014, Borschel et al. 2019). AF is associated with increased morbidity and mortality and patients show a five-fold increased risk of stroke (Healey et al. 2012, Wolf et al. 1991) as well as a three-fold increased risk of heart failure (Wang et al. 2003). The absence of organised atrial contraction during AF results in cardiac haemodynamic disturbances thereby promoting formation of thromboemboli which then elevates patient risk for thrombolytic events (Ding et al. 2020). Consequently, the presence of AF is estimated to account for 25-30% of all thromboembolic strokes (January et al. 2014).

The irregularly-irregular ventricular rate in AF also reduces heart contractile efficiency and causes atrioventricular dyssynchrony which can lead to de novo heart failure (Ling et al. 2016). Further, incident AF is associated with a 1.5-fold increase in mortality in men and a two-fold increase in women, even in asymptomatic cases of this disease (Andersson et al. 2013, Benjamin et al. 1998, Stewart et al. 2002). Thus, it is unsurprising that AF presents a significant financial burden for health care, in the UK alone it is estimated that clinical care for patients with AF accounts for 1-2% of the National Health Service budget, with this projected to increase to 4% in the next two decades as the prevalence of AF rises (Burdett et al. 2020).

1.2 Aetiology of atrial fibrillation

AF is the clinical manifestation of underlying disease processes that affect the physiological conduction pathways which propagate electrical impulses from the SA node through the atria and the rest of the myocardium (Perna et al. 2012). The organised spread of these electrical signals requires healthy pacemaker cells and cardiomyocytes arranged in insulated circuits of extra cellular matrix which are predominately maintained by cardiac fibroblasts (Heijman et al. 2014). Reversible causes of AF directly impede the normal cycling of myocyte and pacemaker cell membrane potentials and include: thyrotoxicosis (Franklyn et al. 2012), carbon monoxide poisoning (Dallas et al. 2012), anaphylaxis (Kounis et al. 2020) and excessive alcohol consumption (Voskoboinik et al. 2021). Furthermore, cardiac or thoracic surgery can cause AF (Greenberg et al. 2017), as well as infections that cause septic shock, pericarditis and/or myocarditis (Balik et al. 2017, Maharaj et al. 2018, Peretto et al. 2019).

Sustained AF is more complex and involves three components: initiation of the arrhythmia (clinically observed as paroxysmal AF), arrhythmia maintenance (observed as persistent and longstanding persistent AF), and progression toward longer-lasting (permanent/chronic AF) (Heijman et al. 2014). All these components require both electrical and structural remodelling of the atria to generate ‘vulnerable substrate’ which is tissue capable of maintaining fibrillation following stimulation from ectopic triggers derived from the atria itself and/or the pulmonary veins (Andrade et al. 2014, Kawai et al. 2019). Vulnerable substrate formation involves multiple cell types

including resident myocytes, fibroblasts, endothelial cells and circulating inflammatory cells, which are themselves sensitive to a wide range of contributing factors. Consequently, sustained AF is a disorder that is caused by the augmentation of multiple factors including genetic pre-disposition (Christophersen et al. 2016), other heart diseases (Odutayo et al. 2016) and/or altered neuro-hormonal factors (Chen et al. 2014). Diseases that are strongly associated with sustained AF include valvular heart disease, coronary artery disease and myocardial infarction, chronic heart failure, sinoatrial disease and cardiomyopathy and hypertension (Odutayo et al. 2016). Other diseases associated with AF include diabetes, chronic kidney disease and obesity (Ferreira et al. 2015). When sustained AF occurs in the absence of other disease states it is then termed as lone AF (Tello-Montoliu et al. 2015).

1.2.1 Atrial cellular electrophysiology and electrical remodelling

The morphology of myocyte action potentials depend predominately on the biophysical properties of voltage-gated ion channels expressed in the outer-cell and sarcoplasmic membranes (Bartos et al. 2015). Under normal conditions the successive opening, inactivation, and closure of these channels gives rise to the five distinct phases of the myocyte action potential (phase 0, 1, 2, 3 and 4). Phase zero is the initial rapid depolarisation caused by influx of Na^+ through voltage-gated Na^+ channels (I_{Na}) and phase 1 is the small downward deflection (early repolarisation) mediated by transient outward K^+ (I_{KTO}) (Bohnen et al. 2015) and the *ultra-rapid* delayed rectifier K^+ channels (I_{KUR}) (Ravens et al. 2010). Phase 2 is the plateau phase, which is maintained by *inward* Ca^{2+} through L-type channels (I_{CaL}) and *outward* K^+ through the *rapid* delayed rectifier channels (I_{KR}). Phase 3 is repolarization, which occurs with sustained outward K^+ through *slow* delayed rectifier channels (I_{KS}) and brings myocytes to rest phase (phase 4) (Grant 2009, Trenor et al. 2017). Myocyte action potentials cause contraction by excitation-contraction coupling whereby Ca^{2+} influx through L-type channels during the plateau phase stimulates sarcoplasmic ryanodine receptors (RyR) to open, causing a large release of Ca^{2+} (Ca^{2+}-induced Ca^{2+}-release) which drives Ca^{2+}-troponin binding and myocyte contraction (Mayourian et al. 2018).

The hallmark feature of electrical remodelling in AF patients when compared to sinus rhythm controls is a *shortening* of the myocyte action potential duration (APD) and hence the atrial effective refractory period (Lip et al. 2016). The observed reduction in atrial effective refractory period leads to AF by facilitating conduction circuit re-entry. Circuit re-entry is the collection of premature action potentials in myocyte conduits and can occur as either circular or spiral electrical activity (Chen et al. 2017) (Liu et al. 2020). Conversely, AF can also be associated with a *lengthening* of the myocyte APD which serves to either postpone the beginning of the next action potential (delayed after-depolarisations), or cause quick subsequent low-amplitude action potentials (early after-depolarisations) (Heijman et al. 2014). Either consequence of increased APD, as well as decreased APD, ultimately perturbs the uniform spread of current across the atria thereby promoting atrial fibrillation (Avula et al. 2019, Martins et al. 2014).

The changes in APD (both shortening and lengthening), as well as other deviations from the standard action potential waveform, are largely the result of expression- and molecular-changes in ion channels. In AF with decreased APD, there is a significant reduction in the density of I_{CaL} channels thereby reducing Ca^{2+} influx which shortens the plateau phase duration (Denham et al. 2018). Further, familial atrial fibrillation is associated with gain-of-function mutations of I_{KR} (Hong et al. 2005) and I_{KS} channels which drive repolarization thus also shortening the action potential plateau phase (Lundby et al. 2008, Nielsen et al. 2014, Olesen et al. 2012). Conversely, in AF with increased APD, there is significant downregulation of I_{KTO} which results is an increased plateau voltage and markedly delayed repolarization which together extend APD (Ni et al. 2019). AF can also be driven by increased activity in inward rectifier K^+ channels (I_{K1}) (Tsai et al. 2008), which ordinarily set the resting membrane potential in myocytes, limit the outward K^+ current during the positive phase of the action potential, and then contribute to phase 3 repolarization (Dhamoon et al. 2005). The increased activity of I_{K1} channels is supported by increased mRNA and protein abundance of Kir2.1 which is the α-subunit of the channel (Gaborit et al. 2005) but how increased I_{K1} affects APD in AF is debated. Computer simulations of 2-dimensional human atrial cell networks did however show that increased I_{K1} shortens the ADP thereby facilitating circuit re-entry and fibrillation (Pandit et al. 2005).

1.2.2 Structural remodelling

Structural remodelling is the other key contributing factor to vulnerable substrate formation and refers to the collective changes that physically impede uniform action potential propagation by conferring increased resistance and current block (Nattel 2016). Progressive accumulation of fibrotic tissue in the atria is a major component of structural remodelling and is caused by excessive formation and redistribution of extracellular matrix (ECM), predominately fibrillar collagen, either in regions of cardiomyocyte loss (reparative fibrosis) or in areas not affected by injury (reactive fibrosis) (Yue et al. 2011). Aberrant ECM deposition interrupts myocyte bundle continuity leading to conduction slowing which initiates and maintains circuit re-entry. As a further part of the pro-fibrotic response, resident cardiac fibroblasts begin to proliferate and migrate into interstitial niches where they interact with cardiomyocytes electrically through newly synthesised connexin gap junctions (Burstein et al. 2009, van der Velden et al. 2000). As fibroblasts are inexcitable they act as current sinks which confer increased resistance and blockage to circuit passage (Dzeshka et al. 2015).

Cardiac fibrosis is present in almost all forms of cardiovascular disease including chronic heart failure, valvular heart disease, hypertension, myocardial infarction and atrial fibrillation (Murtha et al. 2017). In the majority of patients with AF, atrial fibrosis occurs secondary to other underlying cardiovascular disease (e.g hypertension) that have worsened the atrial substrate necessary for AF development (Burstein et al. 2008). However a minority of patients develop AF without the presence of pro-fibrotic co-morbidities or underlying cardiovascular disease (i.e. lone AF). In patients with lone AF, early histological studies on atrial septum biopsies were the first to show increased cardiac fibrosis in this disease state (Frustaci et al. 1991, Frustaci et al. 1997, Levy 1997). Later, MRI investigations also confirmed cardiac fibrosis in a different cohort of patients with lone AF (Mahnkopf et al. 2010) and post-mortem studies confirmed that fibrosis was associated with the duration of years lived with AF specifically and not due to differing patient age (Platonov et al. 2011). Further, in experimental models where AF is induced with implanted pacemakers or trans-oesophageal burst pacing in healthy animals, fibrosis is observable soon after AF induction (Everett et al. 2006, Lin et al. 2007, Pan et al. 2007).

Together these studies suggest that atrial fibrosis and atrial fibrillation are intrinsically linked, independent of other co-morbidities. However this relationship is not well understood and consequently the available treatments for AF are limited to restoring and maintaining normal sinus rhythm with antiarrhythmic drugs and/or catheter ablation and are not focused on resolving cardiac fibrosis specifically (Calvo et al. 2018). For example, K^+ channel-blocking drugs are routinely used to treat AF but can show non-selective effect on both atrial and ventricular electrical properties thereby causing significant side effects (Ravens et al. 2017). A better understanding of the mechanisms that control the progress from atrial fibrosis to AF and vice versa would provide important insights into the development of this disorder as well as provide novel drug targets that act as common drivers for both electrical and structural myocardial remodelling. (Lip et al. 2016).

1.3 Cardiac fibroblasts and extracellular matrix

Cardiac fibroblast is a non-specific term used to describe all non-excitable interstitial spindle-shaped cells capable of producing extracellular matrix in the heart and are frequently sub-characterised by their location (e.g. ventricular, atrial, septal and valvular fibroblasts) (Tallquist 2020). Histological and flow cytometry based investigations in mice suggest that fibroblasts make up ~10 % of the total cell number in the adult heart (Pinto et al. 2016) and lineage tracing experiments indicate that cardiac fibroblasts form and disperse from the epicardium (Smith et al. 2011, Acharya et al. 2012, Asli et al. 2014), and to a lesser extent from endocardial endothelial cells undergoing epithelial-to-mesenchymal transition (EMT), in the developing embryo (Kisanuki et al. 2001, Ali et al. 2014).

Cardiac fibroblasts maintain myocardial ECM which is all cell derived non-living material and is sub-characterised into 'fibrous structures' and 'ground-substance.' Fibrous structures are defined as either *collagenous*, which are made up of different collagen sister proteins, or *elastic*, which are made predominately of elastin and fibrillin. Ground-substance consists mainly of water - which acts as a medium through which signalling and matricellular proteins and ions can diffuse – as well as a wide range of hydrophilic glyco-conjugated proteins that create an amorphous gel-like meshwork that encapsulates resident cells (Omelyanenko et al. 2014, Rienks et al. 2014). In the myocardium, ECM is predominately made up of collagen type I fibrils (encoded by Col1a1 and Col1a2 sister genes) and collagen type III fibrils (encoded solely by the Col3a1 gene) (Bashey et al. 1992, Ivey et al. 2016). Collagen fibrils are enzymatically cross-linked by lysyl oxidase (LOX) in the extracellular space to form wide tough fibres, and in myocardial tissue, these fibres are arranged into a large interconnected honeycomb like structure which gives tensile integrity to the heart as it sustains repeated expansion-contraction cycles (Stephenson et al. 2016). Myocardial ground substance contains versican, aggrecan and other proteoglycans, as well as fibronectin and laminin glycoproteins, which support both fibroblasts and myocytes by bridging transmembrane integrins to extracellular collagen scaffolds (Bashey et al. 1992, Bagchi et al. 2016, Mohammadzadeh et al. 2019).

1.3.1 Cell signalling leading to cardiac fibrosis

Cardiac fibrosis is characterised by increased production and decreased degradation of cardiac ECM and is driven by a wide range of circulating and locally synthesized pro-fibrotic factors as well as mechanical stimuli. Fibroblasts are the primary cell type responsible for integrating different pro-fibrotic stimuli and demonstrate significant changes in cell behaviour including proliferation, migration and differentiation into a secretory, but non-proliferative, phenotype called myofibroblasts that sustain ECM production (Bochaton-Piallat et al. 2016). A recent study showed that myofibroblasts specifically, not resident myocytes or endothelial cells, are the primary source of collagen production in mouse hearts acutely and chronically stressed with transverse aortic constriction (TAC) (Khalil et al. 2019). These changes in cardiac fibroblast behaviour and phenotype are caused by increased activity in major pro-fibrotic signalling pathways, importantly canonical and non-canonical transforming growth factor beta (TGFβ) signalling (Molkentin et al. 2017) and angiotensin II (AngII) signalling (Nattel et al. 2020).

In physiological conditions, TGFβ is secreted by cardiac fibroblasts in an inactive complex consisting of TGFβ and latency associated propeptide (LAP) which is anchored to latent TGFβ binding protein 1 (LTBP1) in the ECM. Changes in mechanical forces (for example, in hypertension and valvular disease) can release active TGFβ from the latent complex, which even in small quantities, is capable of driving a pro-fibrotic response following binding to its transmembrane type I and type II receptors (TβRI and TβRII) (Annes et al. 2003, Herum et al. 2017). Active TGFβ can also be derived from non-cardiac sources including circulating plasmids and megakaryocyte-specific TGFβ knockout in mice can partially protect against cardiac fibrosis induction in different disease models (Meyer et al. 2012, Laurence et al. 2017).

Binding of TGFβ to TβRI and TβRII facilitates phosphorylation of the cytoplasmic serine/threonine domain of TβRI which in turn phosphorylates mothers-against decapentaplegic homolog 2/3 (SMAD2/3) that bind cytoplasmic SMAD4 which together translocate to the nucleus and drive gene expression (referred to as canonical TGFβ signalling). Gene expression changes following canonical TGFβ activity include TGFβ and its receptors (i.e. TGFβ induced TGFβ signalling), type I and III collagen, lysyl oxidase and fibronectin as well as matricellular proteins including thrombospondin, periostin and connective tissue growth factor (CTGF) (Khalil et al. 2017, Yousefi et al. 2021). Thrombospondin and CTGF bind latent-TGFβ in the ECM to facilitate interaction with TβRI/II receptors thereby reinforcing further TGFβ signalling (Rienks et al. 2014, Dorn et al. 2018). Periostin is a ligand for cell integrins and mediates cell-matrix and intercellular interactions to facilitate fibroblast motility (Walker et al. 2016, Landry et al. 2017). Canonical TGFβ activity also stimulates expression of cell-matrix adhesion proteins including integrins, fibronectin and laminin to further aid in fibroblast motility (Schroer et al. 2015). As an additional layer of regulation, canonical TGFβ activity increases ECM turnover by invoking expression changes in matrix metalloproteinase (MMPs) and their tissue inhibitors (TIMPS) (Nattel 2017).

Binding of TGFβ to TβRI/II also simulates non-canonical TGFβ signalling which activates mitogen activated protein kinases (MAPKs) including: the extracellular-signal regulated kinase 1 and 2 (ERK1/2), c-Jun N-terminal kinases (JNKs), the p38 isoforms (α, β, γ, and δ) and TGFβ activated kinase 1 (TAK1) (Dzeshka et al. 2015). Increased activity of MAPKs is widely studied in cancer cell lines and facilitate cell proliferation (Guo et al. 2020, Maik-Rachline et al. 2019), metabolic reprogramming (Papa et al. 2019) and cell differentiation (Semba et al. 2020). In myocardial fibrosis, pharmacological interventions that reduce p38 and ERK blunt TGFβ-dependent cardiac remodelling in animal models of myocardial infarction (Kompa et al. 2008, See et al. 2004). Furthermore, cardiac overexpression of active TAK1 and ERK is implicated in myocardial fibrosis in pressure overload mouse models (Zhang et al. 2000, Lim et al. 2021).

In addition to TGFβ signalling, increased binding of AngII to its type I receptor (AT1R) drives pro-fibrotic activity in cardiac fibroblasts (Nattel et al. 2020). AngII can be delivered from circulating blood plasma which is increased in a number of disease states associated with atrial fibrillation including hypertension, chronic kidney disease and diabetes (Laffer et al. 2020, Oparil et al. 2018). Active AngII is also derived from local cardiac sources, including myocytes and fibroblasts, which increase expression of AngII and AT1R following increased mechanical strain (Burstein et al. 2008). Binding of AngII to AT1R leads to increased kinase activity of ERK1/2 and JNKs which in turn activate Kruppel like zinc finger transcription factor five (KLF5) (Forrester et al. 2018). KLF5 enhances canonical TGFβ signalling by interacting with the SMAD2/3/4 complex and consequently induces further ECM and TGFβ production to enhance the fibrotic response (Takeda et al. 2010, Hara et al. 2017). Canonical TGFβ signalling also upregulates AngII and AT1R and this leads to reciprocal positive feedback between these major pro-fibrotic pathways (Rosenkranz 2004, Tadevosyan et al. 2017).

Furthermore, KLF5 activates the expression of platelet derived growth factor A (PDGFA). Again through autocrine and paracrine signalling, PDGFA binds to its membrane receptors to stimulate Janus kinase (JAK) and its activation products: signal transducers and activators of transcription (STATs), which further promote expression of pro-fibrotic factors including endothelin-1 and TGFβ (Chen et al. 2017, Nattel et al. 2020).

1.3.2 Markers for healthy and disease state cardiac fibroblasts

'Fibroblast' traditionally defined all 'fibre' producing cells which adhere to surfaces when cultured *in vitro*, but unlike epithelial cells, are spindle shaped and not restricted to forming monolayers (Carrel et al. 1926). Keeping to this definition, fibroblasts include a range of cell types such as interstitial fibroblasts, pericytes, fibrocytes, myofibroblasts, osteoblasts etc. (LeBleu et al. 2020). The use of molecular biomarkers is a more specific way of defining fibroblast and sub-types as it differentiates cells based on mRNA and/or protein abundance, however using this approach in experimentation can by definition exclude fibroblast types that have unknown and unique expression profiles. For example, single cell RNA-seq can be used to sub-categorise fibroblasts, however this technique relies on pre-defined markers to sort and select cells and thus excludes ill-defined fibroblast types from characterisation (Svensson et al. 2018).

With this limitation in mind, one of the most utilised fibroblast biomarkers is the expression of type I collagen fibrils which are comprised of both COL1A1 and COL1A2 protein. Transgenic mice that harbour a random genomic insertion of the green florescent protein (GFP) cassette cloned to the Col1a1 promoter, have been generated and used to visualise fibroblasts in developing and adult tissues including bone, liver and heart (Kalajzic et al. 2002, Yata et al. 2003). Similarly, Cre driver mouse lines that carry constructs of Cre-recombinase cloned to the Col1a1 and/or Col1a2 promoter have been generated and are frequently used to manipulate expression of floxed targets in fibroblasts in a wide range of organs (Kim et al. 2004, Zheng et al. 2002).

Owing to the non-selective expression of collagen type I across multiple fibroblast sub-types and organ systems, the utility of other biomarkers for identifying both embryonic and adult cardiac fibroblast have been explored. For example, adult cardiac fibroblasts express surface membrane PDGF receptor alpha (PDGFRα) and lineage tracing experiments in mice demonstrated that selective deletion of PDGFRα in embryonic epicardial cells results in failure of cardiac fibroblasts to form and disperse from these cells (Smith et al. 2011). PDGFRα expression is retained in healthy adult cardiac fibroblasts at low levels and shows significant upregulation in cardiac disease states (Chong et al. 2013). Similarly, the selective deletion of transcription factor 21 (TCF21) also results in failure of cardiac fibroblast formation from epicardial cells, TCF21 expression is also retained at low levels in normal adult cardiac fibroblasts and can increase after injury (Acharya et al. 2012, Braitsch et al. 2013). Developing cardiac fibroblasts of epicardial origin also express discoidin domain receptor 2 (DDR2), which binds extracellular collagen and other ECM protein but DDR2 expression is limited to a only a small subset of adult cardiac fibroblasts (Morales et al. 2005, Banerjee et al. 2007, Ali et al. 2014). In addition to the epicardial cell population, a unique population of endocardial endothelial cells that express tunica interna endothelial cell kinase 2 (TIE2) give rise to a smaller population of cardiac fibroblasts in the developing embryo, but TIE2 expression is reduced in these cells in extra-uterine life (Kisanuki et al. 2001, Ali et al. 2014).

The membrane glycoprotein thymus-cell antigen 1 (Thy1, CD90) is another marker that has been used to identify and sort healthy adult cardiac fibroblasts *in vitro* (Ali et al. 2014), and disease state cardiac fibroblasts *in vivo* can be identified by a reduction of Thy1 surface expression (Hudon-David et al. 2007, Li et al. 2020). In addition, vimentin is an intermediate filament protein that has been used as a biomarker in both *in vitro* and *in vivo* studies since it is expressed in cardiac fibroblasts in both the healthy and injured heart (Moore-Morris et al. 2014).

In the case of myofibroblast identification, the expression of alpha smooth muscle actin (αSMA) is the most widely used biomarker. αSMA is contractile and ordinarily expressed in vascular smooth muscle cells but is significantly upregulated in myofibroblasts and contributes to scar contracture seen in fibrosis (Ribatti et al. 2019, Shinde et al. 2017). Furthermore, periostin shows high abundance in embryonic cardiac fibroblasts, no expression in adult cardiac fibroblasts, and high expression in myofibroblasts and is therefore also used to mark these cells (Conway et al. 2008) (Landry et al. 2017). In addition, one study showed that vascular adhesion protein (VAP1) could be used to distinguish dermal fibroblasts from myofibroblasts (Hsia et al. 2016). Other markers that have been explored and used include fibroblast specific protein 1 (FSP1), stem cell antigen 1 (SCA1), Wilms' tumour suppressor gene 1 (WT1) and t-box 18 (TBX18) (Fu et al. 2020).

All known fibroblast markers, including all those mentioned in this section, show varied expression patterns across different tissues and cells and consequently there are no markers that exclusively identify fibroblasts and/or their subtypes. As a consequence, *in vitro* and *in vivo* experiments using fibroblasts will utilise multiple different markers for validation. Similarly all developmental and/or disease models using fibroblast drivers are forced to accept the caveat that systemic, and not organ-specific, fibroblast activity will occur (Tarbit et al. 2019).

1.4 The role of microRNAs in structural remodelling

MicroRNAs are small (19-25 nucleotides) endogenous non-coding RNAs that regulate gene expression post-transcriptionally by mediating sequence-specific mRNA degradation and/or translational repression (Ambros et al. 2003, Lagos-Quintana et al. 2001). This activity is facilitated by sequence complementarity between mRNA target sites and the microRNA seed sequence (7 nucleotides close to the 5' end of the 19-25 nucleotide microRNA) (Bartel 2009). Hundreds of miRNAs have been identified in humans (Bartel 2004) and bioinformatic analyses suggest that more than 60% of mRNA transcripts carry microRNA binding sites and that these predominate in the mRNA 3' untranslated region (3' UTR) (Agarwal et al. 2015, Bartel 2009). Analysis of cell-specific gene expression atlases demonstrated that the presence of all 3' UTR target regions in the total mRNA pool was inversely correlated with the abundance of microRNAs capable of binding those same target regions (Farh et al. 2005). This finding was among the first to suggest that (a): the 3′ UTR is linked with tissue-specific expression patterns of microRNAs in humans and therefore microRNAs likely have robust effects on their transcript targets and (b): genes that are necessary for any given cell function are under evolutionary pressure to avoid site matching with microRNAs. Consequent to these early studies, changes in microRNA cellular abundance, as well as deviations in circulating plasma and/or plasmid vesicle microRNA abundance, have been implicated in a wide range of physiological and disease states including cardiovascular disease and cancer. Therefore circulating microRNAs are postulated for use as biomarkers for disease, while specific tissue microRNAs are postulated as targets for therapeutic manipulation (van den Berg et al. 2017).

Multiple studies have demonstrated a link between altered microRNA signatures in atrial tissue (Wang et al. 2015), blood plasma (Liu et al. 2012, McManus et al. 2014) and platelets in patients with AF compared to sinus rhythm controls (Goren et al. 2014). Furthermore, different experimental models of cardiac fibrosis show that microRNAs directly regulate mRNA abundance of genes involved in the pro-fibrotic response. For example, miR27b has two predicted binding sites in the 3'UTR of TβRI, and viral delivery of miR27b to the pericardial space blunted atrial fibrosis induced with an AngII osmotic minipump in mice (Wang et al. 2018). Similarly, viral delivery of miR30c attenuates fibrosis by directly targeting TβRII in an abdominal aortic constriction (AAC) model for atrial fibrosis in rats (Xu et al. 2018).

CTGF and TβR1 harbour binding sites for miR133a, and both genes showed increased mRNA and protein abundance in rat and mouse hearts following reduction of miR133a expression in hypertension disease models (Duisters et al. 2009). In addition, atrial miR133a was downregulated in canines following nicotine-induced atrial remodelling (Shan et al. 2009). Again in canines, when chronic heart failure is induced by ventricular tachypacing, atrial miR29b expression significantly decreases with a corresponding increase of its target genes COL1A1, COL3A1, and fibrillin. The same study showed that miR29b knockdown in mice enhances COL1A1 gene expression and cardiac collagen content, and that both plasma and atrial miR29b content was decreased in humans with chronic heart failure and/or atrial fibrillation (Dawson et al. 2013). The Framingham study also confirmed a reduction of plasma miR29b in patients with incident AF (McManus et al. 2014).

Sprouty 1 RTK signalling antagonist (SPRY1) has protective effects as it inhibits ERK1/2, the activity of which is increased in pro-fibrotic signalling (Thum et al. 2008). SPRY1 is a direct target of miR21 and in ischemic rat models, the upregulation of miR21 was associated with reduced SPRY1 which facilitated pro-fibrotic signalling (Cardin et al. 2012). This role is supported by the fact there is increased miR21 in the left and right atrial appendage derived from patients with AF-associated valvular heart disease (Adam et al. 2012, Cooley et al. 2012). Similarly SMAD7 has protective effects as it inhibits SMAD2/3, the activity of which is increased in fibrosis (Nattel 2017). SMAD7 is also targeted by miR21 and consequently increased miR21, with downregulated SMAD7, is associated with myocardial fibrosis following atrial pacing in rabbits (He et al. 2016). However one study demonstrated that miR21 knockout in mice, and/or inhibition with locked nucleic acid (LNA) anti-miRs, showed no change in cardiac fibrosis in response to TAC or induced myocardial infarction (Patrick et al. 2010).

1.5 miR31 in atrial fibrillation

miR31 is a pleiotropically acting microRNA that is broadly conserved in mammals and most vertebrates including zebrafish. The human ortholog of mature miR31 was first cloned and sequenced from Henrietta Lacks (HeLa) cells screened for novel short RNAs (Lagos-Quintana et al. 2001) after the original identification of miR31 in *D. melanogaster* embryo lysate (Elbashir et al. 2001). Subsequent large scale sequencing of more than one hundred human small-RNA libraries from nine different organ systems confirmed miR31 abundance in heart tissue, connective tissue, the lungs, liver and kidneys (Landgraf et al. 2007).

Our group has recently shown that miR31 abundance is increased in atrial myocytes from patients with persistent AF and that miR31 begets electrical remodelling by depleting dystrophin and neuronal nitric oxide synthase (Reilly et al. 2016). Subsequent to this observation, we measured miR31 abundance in fibroblasts derived from patients with persistent AF and saw a similar increase when compared to sinus rhythm controls (*see results figure 3.1*). Consequently the literature was queried for a role for miR31 in atrial fibrillation-associated fibrotic signalling and/or fibrosis associated with any other disease state. The Framingham study showed no change in plasma miR31-5p in AF patients, however the functionally less active antisense sequence of miR31-5p (i.e. miR31-3p) was significantly reduced in AF patient plasma (McManus et al. 2014). One study showed that miR31 was among the most highly expressed microRNAs in rat hearts following induced myocardial infarction with coronary artery ligation, and that post-ischemic administration of miR31 inhibitor helped preserve cardiac structure and function (Martinez et al. 2017). Another study investigating hypertrophic scar formation, a disease characterised by excessive ECM deposition, identified miR31 among the most differentially expressed microRNAs in disease-state fibroblasts compared to controls, and that miR31 positively regulated the hypoxia inducible factor 1 alpha (HIF1α) pathway by targeting HIF1α inhibitors (Wang et al. 2017). Later studies demonstrated a similar role for miR31 in keloid scarring (Zhang et al. 2017), and in TGFβ-induced liver fibrosis (Hu et al. 2015).

Given the role of miR31 in other disease models of fibrosis, as well as the paucity of information linking miR31 to structural remodelling in AF, we devised an experimental strategy to investigate the role of miR31 in human atrial fibroblasts *in vitro* and also *in vivo* by creating a novel transgenic mouse strain capable of Cre-dependent overexpression of miR31 in fibroblasts (*see forthcoming project aims and hypotheses*). Our lab has very recently shown that global disruption of calcitonin receptor (CALCR) signalling in mice causes atrial fibrosis and that atrial-specific knockdown of calcitonin promotes atrial fibrosis and atrial fibrillation (Moreira et al. 2020). CALCR mRNA carries predicted binding sites for miR31 in its 3'UTR, and should CALCR mRNA and protein prove to be reduced following upregulation of miR31 in fibroblasts, this finding will help us understand the contribution of miR31 to this novel pro-fibrotic signalling mechanism. This would promote miR31 as a therapeutic target in AF, and since our lab has also demonstrated that increased miR31 promotes electrical remodelling in myocytes (Reilly et al. 2016), targeting miR31 for degradation could protect against both electrical and structural remodelling seen in AF.

1.5.1 Transcriptional regulation and genesis of microRNA31

Canonically, microRNAs originate from long primary transcripts, generated by Pol-II and -III, that are processed in the nucleus by the RNase-III enzyme Drosha to generate ~70 nucleotides hairpin precursor microRNAs (pre-miRs) (Ozsolak et al. 2008). Pre-miRs are exported to the cytoplasm and cleaved by another RNase-III enzyme Dicer to produce the mature miR and antisense passenger microRNA star (miR*). The mature microRNA is loaded onto the RNA-induced silencing complex (RISC) for mRNA targeting (Ambros et al. 2003, Bartel 2004). For a significant minority of microRNAs, including miR31, their coding sequences are nested in introns of transcripts destined for splicing. As a result, the initial processing of these transcripts to generate pre-miRs is thought to use Drosha in consonance with the spliceosome or simply the spliceosome on its own (Okamura et al. 2007, Ruby et al. 2007). Regardless, the following miR31 predecessor transcripts can be considered when investigating how levels of miR31 are perturbed: one, the initial Pol-II/III generated primary transcript (pri-miR31); two and three, the Drosha/spliceosome generated non-coding mature mRNA (HG-miR31) and precursor-miRNA (pre-miR31); and four and five, the Dicer generated miRNAs (miR31 and miR31*).

Marson *et al.* were the first to predict a putative transcription start site (TSS) for miR31 (Marson et al. 2008), later it was recognised that this TSS belonged to the previously described long noncoding RNA LOC554202 (Corcoran et al. 2009) and that miR31 sits in the first intron of this transcript (Augoff et al. 2012). LOC554202 is located at 9p21.3 which is a loci of interest in cancer since it is adjacent to clusters of the cyclin-dependent kinase inhibitor 2 (CDKN2) and interferon (IFN) families (Laurila et al. 2013). Consequently, studies investigating the epigenetic and transcriptional regulation of microRNA31 have predominately been in cancer cell lines. For example, LOC554202 was downregulated by hypermethylation of the CpG island overlaying its transcription start site in breast cancer cells (Augoff et al. 2012 16). In addition, the polycomb-repressive complex supresses LOC554202 by catalysing histone H3K27 methylation in prostate (Zhang et al. 2014) and chordoma cell lines (Ma et al. 2017). Another study showed that B-cell lymphoma 6 (BCL6) transcription factor binds upstream of the LOC554202 TSS and represses its expression through de-methylation of H3K4me1/3 and methylation of H3K27 in follicular helper T-cells (Ripamonti et al. 2017).

Again in a breast cancer cell line, one study showed that association of erythroblastosis transcription factor 1 (ETS1) with the LOC554202 promoter reduced miR31 abundance. The authors elucidated and proposed a mechanism whereby ETS1 recruits EMSY[1], that subsequently employs the lysine demethylase 5B (KDM5B) to demethylate H3K4me3 thereby silencing the LOC554202 promoter (Vire et al. 2014). This was supported by the observation that ETS1 chromatin immunoprecipitation sequencing (ChIP-seq) shows binding peaks at the LOC554202 promoter in A549 cells, a lung epithelial cancer line (Luo et al. 2020). ETS1 is inhibited by phosphorylation at residues Ser282 and 285 by the calcium/ calmodulin-dependent kinase II (CaMKII) (Fleischman et al. 1995, Vetter et al. 2005), the activity of which is known to be increased in AF and to play an important role in its pathogenesis (Heijman et al. 2014, Purohit et al. 2013). We subsequently queried the abundance of ETS1 protein in the right atrial appendage (RAA) in patients with chronic AF and found a reduction in ETS1

[1] EMSY is not an abbreviation of a larger protein name. The protein was named EMSY which is the childhood nickname of Emma Whitley (formerly Emma Hughes-Davies) who is a breast cancer nurse. The nurses' brother, Luke Hughes-Davies, who studies breast cancer, discovered the protein EMSY and named it after her as the first letters of the protein's amino acid sequence reads 'SISTER' (Hughes-Davies et al. 2003).

when compared to controls (*see results figure 4.3*). Following this observation, we speculated that the observed increase in miR31 abundance in atrial fibroblasts derived from AF patients was the result of ETS1 inhibition and/or reduction of presence at the LOC554202 promoter.

As a final note on miR31 expression, it is conceivable that miR31 is transcribed independently of LOC554202, from its own unique promoter, as this is estimated to occur in more than 25% of all intragenic miRs (Ozsolak et al. 2008, Corcoran et al. 2009). However, ChIP-on-chip (i.e. ChIP followed by the use of DNA microarray cards to identify DNA crosslinked to the immunoprecipitated protein) experiments in lung epithelial cells, showed Pol-II signal abundance closest to miR31 at the LOC554202 TSS (Corcoran et al. 2009). In addition H3K4me3, a TSS marker, was overrepresented at the LOC554202 TSS in human ES cells (Marson et al. 2008), B cells (Barski et al. 2007) and T cells (Guenther et al. 2007) but not any closer to the miR31 coding sequence reiterating that miR31 is likely transcribed from LOC554202. Of further interest, like miR31, the coding sequence of interferon epsilon (IFNE) is also located within the first intron of LOC554202, and sits downstream of miR31. There have been no studies looking at the relationship between IFNE and miR31 expression, nonetheless the close proximity of these sequences is intriguing given that interferon expression is sensitive to a wide range of stimuli and feedback loops, however this is speculative and was deemed of lesser importance to this project.

1.6 Project aims and hypotheses

Given the observed increase in miR31 abundance in fibroblasts derived from patients with chronic AF, as well as our group's previous publication showing that increased miR31 is associated with electrical remodelling in cardiomyocytes, we asked whether increased miR31 acts as a common driver for both, electrical, and structural remodelling in chronic AF. We have also observed a reduction in ETS1 in RAA protein derived from chronic AF patients, and given the documented role of ETS1 in recruiting inhibitory transcription factors thereby blocking miR31 expression, we asked whether the observed reduction in RAA ETS1 was facilitating increased miR31 abundance seen in cardiac fibroblasts. To answer these questions the project was split into three experimental aims:

Aim 1: To investigate the effects of miR31 manipulation on pro-fibrotic gene expression and phenotype in healthy human atrial fibroblasts *in vitro*

Aim 2: To investigate the role of ETS1 on miR31 abundance in healthy human atrial fibroblasts *in vitro*

Aim 3: To validate a novel transgenic mouse and investigate the functional consequences of fibroblast specific miR31 up-regulation *in vivo*

The overarching hypothesis of this project is that increased miR31 expression in atrial fibroblasts has an intrinsic role in AF-associated fibrosis, and that a reduction of ETS1 occupancy at the LOC554202 promoter, relieves inhibition of LOC554202 transcription, resulting in increased miR31 abundance in these cells.

Chapter 2

Methods and Materials

2.1 Gene cloning and vector constructs

2.1.1 Cloning and purification of ETS1 expression vector

pcDNA3.1(+) expression vector was obtained as a gift from my co-supervisor, Prof Craig Lygate, WTCHG, Oxford. One µg pcDNA3.1(+) was digested with Not1 (NEB: R0189S) and Kpn1 (NEB: R0142S) restriction enzymes for two hours in CutSmart buffer (NEB: B7204) at 37 °C, then treated with calf intestinal alkaline phosphatase (CIP) (NEB: M0290) for 45 mins at 37 °C to stabilise overhanging 5´ and 3´ ends. Digested products were then electrophoresed in 1% agarose TBE gel and purified using GeneClean-III kit whereby excised gel bands were heated at 55 °C to solubilise DNA in TBE buffered NaI solution before labelling with silica in glass-milk slurry (MP-Biomedicals: 111001600). Silica-labelled DNA was pelleted by centrifugation, the supernatant discarded, and then washed three times in Wash Buffer with intermittent spins to pellet and remove supernatant. DNA was then liberated from glass-milk by adding nuclease-free H_2O, the silica was then pelleted and supernatant containing DNA retrieved.

In parallel the ETS1 ORF was PCR amplified from pDONR223_ETS1_WT (Addgene: 82118) using PfuUltra HF DNA Polymerase (Agilent: 600380) with forward cloning primer, integrated with a Kpn1 docking and cleavage site, and a reverse cloning primer, integrated with a Not1 docking and cleavage site. The PCR reaction and thermocycling conditions are shown in *Table 2.1*.

hsa-ETS1 ORF

Forward: (TTTTTT)(AAGCTT)ATGAAGGCGGCCGTCGAT

Reverse: (AAAAAA)(GCGGCCGC)CTACTCGTCGGCATCTGGCTT

Table 2.1 ETS1 ORF PCR reaction mix and thermocycling

Component	**Volume (uL)**	**Final conc./amount**
PfuUltra HF Polymerase (2.5 U/µL)	1	2.5 U
10x PfuUltra HF Buffer	5 uL	1x
dNTPs (25 mM each dNTP)	0.5	0.25 mM
Forward primer (10 uM)	1	0.5 uM
Reverse primer (10 uM)	1	0.5 uM
Template DNA (~30 ng/uL)	1	30 ng
Nuclease free water	40.5	
Total	50	
Thermocycling step	**Temperature**	**Time**
Initial denaturation	95 °C	1 min
Denaturation	95 °C	50 sec
Primer annealing	60 °C	50 sec
Extension	68 °C	5 min

} 18 cycles

The PCR amplified ETS1 ORF was then ligated into the Not1/Kpn1 digested pcDNA3.1(+) using T4 DNA ligase (NEB: M0202S) by incubation at 4 °C for 48 hrs in ligation buffer (NEB: B0202S) to create pcDNA3.1(+)_ETS1_WT expression vector (frequently referred to as ETS1 expression vector throughout this document). The molar ratio of ETS1 ORF / pcDNA3.1(+) was 5 : 1 in this reaction. *Figure 2.1* depicts the workflow used to generate ETS1 expression vector.

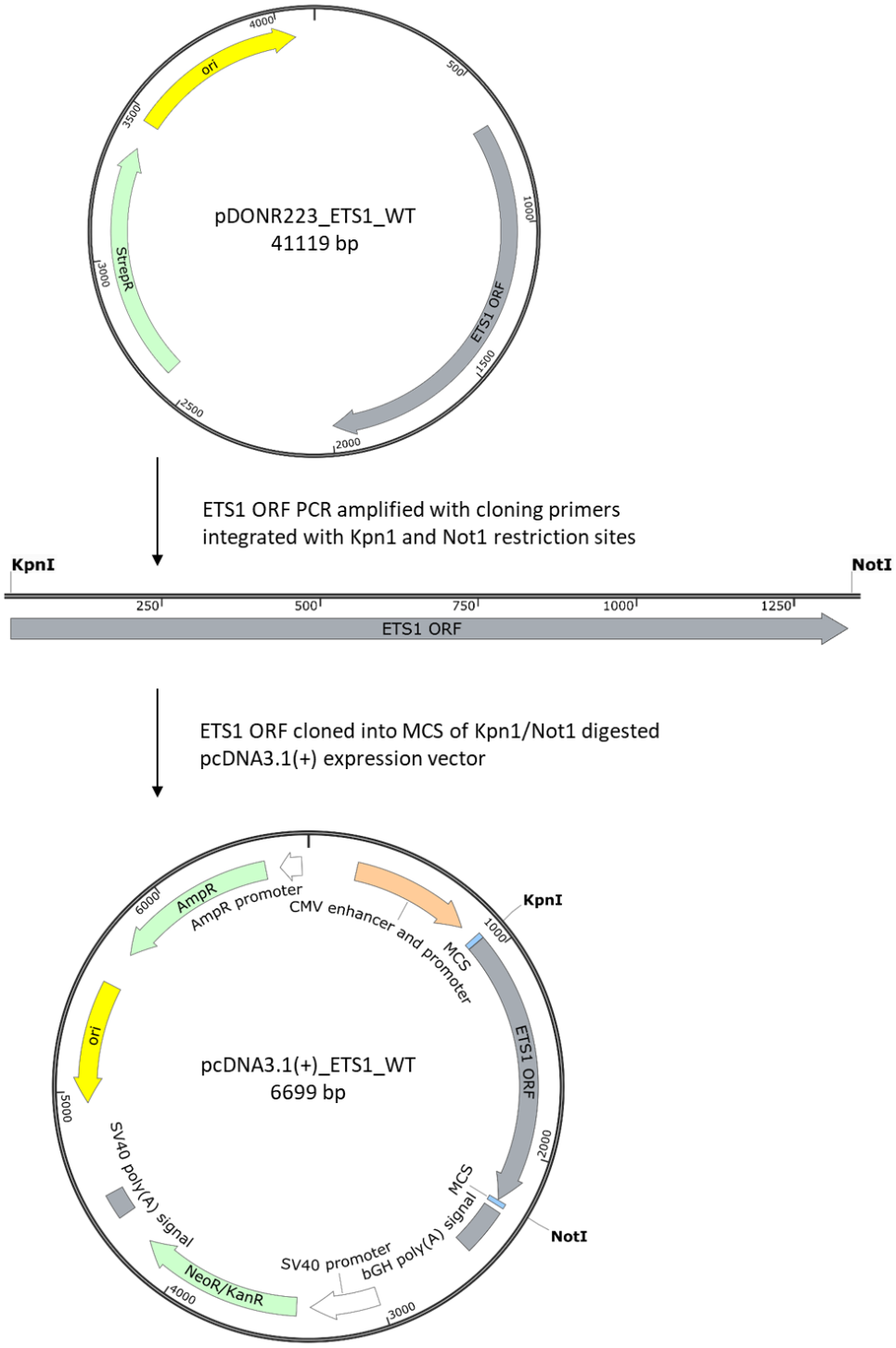


Figure 2.1 Workflow used to generate ETS1 expression vector

To make a stock of ETS1 expression vector, 1 mL of cryopreserved DH10-beta cells in Luria-Bertani (LB) broth were transformed by heat-shocking at 37 °C for 5 minutes with 1 µg vector before chilling on ice for 1 minute. Cells were then cultured on ampicillin (100 µg/mL) LB-agar selection plates and five single colonies picked and grown overnight in 5 mL starter culture. Simultaneously the five colonies were assessed for transformation by colony PCR using the same ETS1 PCR amplification protocol above. Starter cultures selected for expansion were incubated in 400 mL LB broth, also with ampicillin (100 µg/mL), overnight at 37 °C. Cells were pelleted using an ultra-centrifuge, fitted with a JLA-10.500 rotor, at 6000 g for 6 minutes and plasmids were then purified by Maxiprep (Qiagen: 12163) according to manufacturer's instructions. Construct sequences were then validated with sequencing at Source Bioscience.

2.1.2 Identification and validation of putative ETS1 binding sites within the wild type hsa-LOC554202 promoter

The 1061 base pair sequence immediately upstream of the first exon of the LOC554202 was obtained from UCSC genome browser (hg38, chr9: 21559699 – 21560759, negative strand) and was analysed *in-silico* for transcription factor binding sites. Initially JASPAR (2018 release) (www.jaspar.genereg.net) was used using the Scan feature, set to 90% relative profile score threshold, that analysed the 1061 bp sequence for ETS1 binding sites using two separate DNA motifs (matrices ID: MA0098.1 and MA0098.3) (*figure 2.2A*). Subsequently, the same stretch of DNA was analysed using the TRANSFAC (release 2018) (www.genexplain.com/transfac) MATCH tool that scanned for ETS1 binding sites using the one available binding motif (matrices ID: V$ETS1_02) (*figure 2.2B*). For further validation, the Encyclopaedia of DNA Elements (ENCODE) (www.encodeproject.org) was queried for ETS1 ChIP-seq data by entering 'ETS1'in their search engine. ETS1 ChIP-seq data was then visualised on the UCSC genome browser (hg38), by using the upload ENCODE track feature, to identify previously validated ETS1 binding peaks at the LOC554202 promoter.

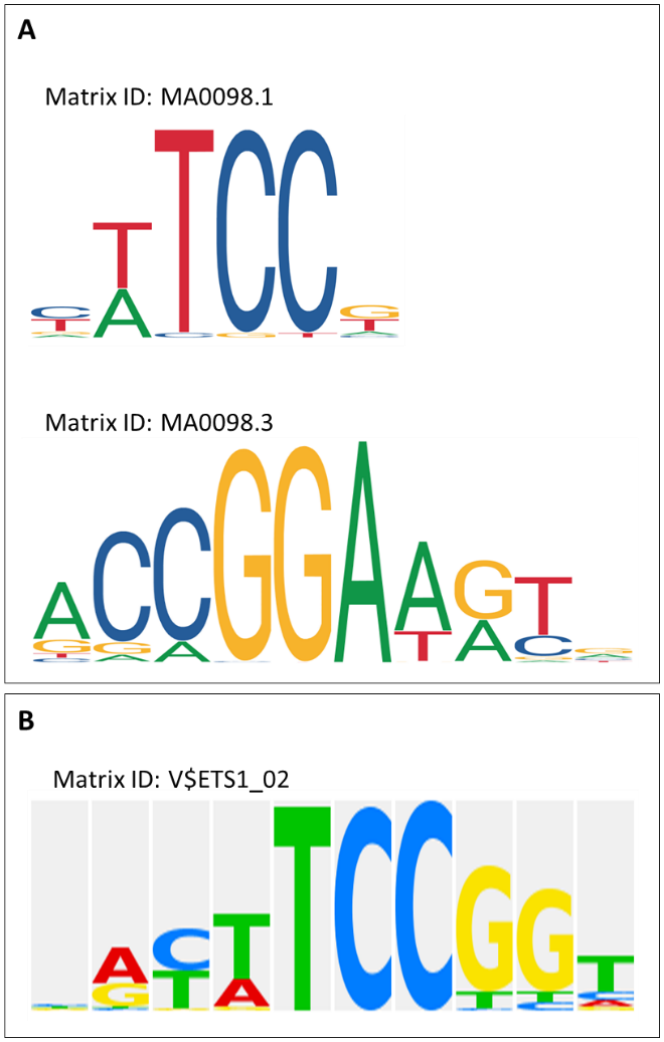


Figure 2.2 ***ETS1 DNA-binding motifs***

A: the two different ETS1 DNA binding motifs used by JASPAR (2018 release) to identify ETS1 binding sites at the LOC554202 promoter.
B: the TRANSFAC (release 2018) ETS1 DNA binding motif used to identify ETS1 binding sites at the LOC554202 promoter.

2.1.3 Generating wild type hsa-LOC554202 promoter *Firefly* Luciferase and mutagenesis constructs

Once putative ETS1 binding sites were identified upstream of the first exon of LOC554202, this 1061 bp segment of DNA was PCR amplified from human genomic DNA (Promega: G1521A) using PfuUltra HF DNA Polymerase (*identical protocol in table 2.1 above*) with forward primer integrated with a Xhol docking and cleavage site, and a reverse cloning primer integrated with a HindIII docking and cleavage site.

hsa-LOC554202 promoter

Forward: (TTTTTT)(CTCGAG)TGGGCTTGCGGAGGGCAT
Reverse: (AAAAAA)(AAGCTT)GCGAGCCCCTGGCGGCGC

The PCR amplified LOC554202 promoter was then ligated into the Xhol/HindIII digested luciferase reporter plasmid pGL4.10[*Luc2*] (Promega: E6651), using T4 DNA ligase (NEB: M0202S) incubated at 4 °C for 48 hours in ligation buffer to create pGL4.10[*Luc2*]_LOC554202_Promoter_WT vector (frequently referred to as WT LOC554202 promoter *Firefly* Luciferase vector in this document). To make a stock of this vector, protocol was identical to ETS1 vector (*final paragraph of section 2.1.1*) whereby DH10-beta cells were transformed with the ligated product, grown in LB-broth with ampicillin, and plasmids were purified by MaxiPrep (Qiagen: 12163).

Wild type LOC554202 promoter *Firefly* Luciferase vector was then used as a template for generating eleven unique mutagenesis constructs using the QuickChange Site-Directed Mutagenesis kit (Agilent: 200521). For each of the eleven constructs, 10 ng of pGL4.10[*Luc2*]_LOC554202_WT was PCR amplified using PfuUltra HF DNA Polymerase (*identical protocol in table 2.1.1 above*) and mutagenesis primers (*see section 2.12 for all primer sequences*). PCR amplified products were then treated with Dpn I restriction enzyme at 37 °C for 1 hour to digest the original non-mutated DNA. Stocks of these mutagenesis constructs were then generated similar to ETS1 expression vector (*final paragraph of section 2.11*) and mutagenesis was confirmed by sequencing of both the forward and reverse strand using the above XhoI and HindIII cloning primers as sequencing primers. Sequencing was performed by Source Bioscience. *Figure 2.3* depicts the workflow used to generate wild type LOC554202 promoter *Firefly* Luciferase vector and mutagenesis constructs.

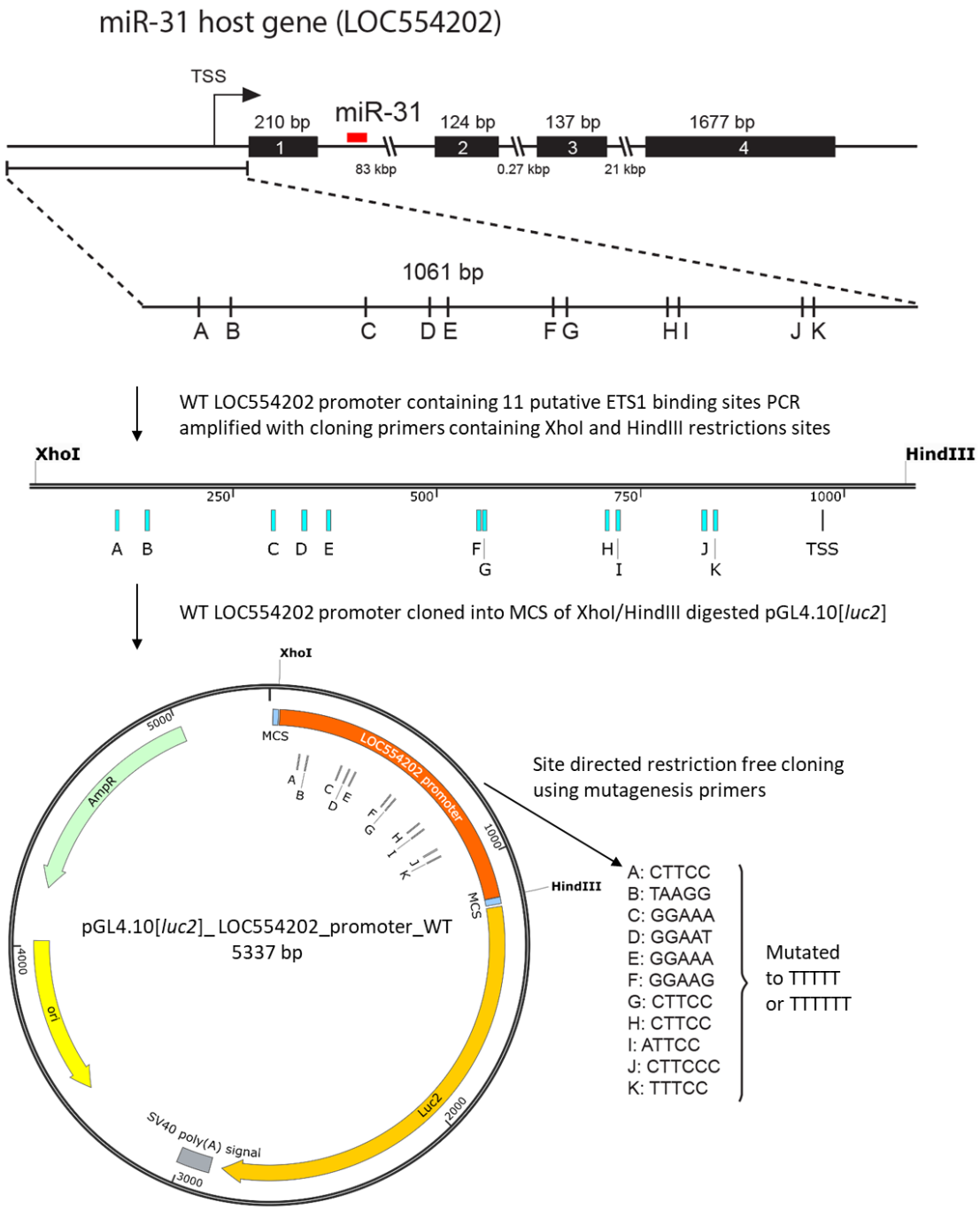


Figure 2.3 ***Workflow used to generate WT LOC554202 promoter Firefly Luciferase vector and mutagenesis constructs***

2.2 Tissue culture experiments

2.2.1 Receipt, expansion and storage of hAFs

Normal primary human atrial fibroblasts (hAFs) (Lonza: CC-2903) from donor 1 (Lot: 0000634284) and donor 2 (Lot: 18TL312486) (see appendices for donor certificates), were received on their third passage on dry ice in 1 mL cryogenic vials (~0.82 x10^6 cells / 1 mL) which were then stored in liquid nitrogen vapour phase at -170 °C. To increase cell amount and develop a bank for experimentation, hAFs were thawed and seeded in a T75 containing 10 mL Fibroblast Growth Medium-3 (FGM-3) (Lonza: CC-4526) which contains a proprietary concentration or recombinant human insulin, fibroblast growth factor-β, and gentamicin sulphate amphotericin-β supplemented with 10% foetal bovine serum (FBS) (Thermo: 10270098). hAFs were then expanded with two more passages (~1 passage/ 48 hrs), before trypsinisation and storage again at -170 °C in 10% dimethylsulfoxide (DMSO) (*figure 2.4*). For all tissue culture procedures with hAFs, media was changed every 24 hours, cells were passaged at confluence and the humidified cell incubator was set at 37 °C and 5% CO_2.

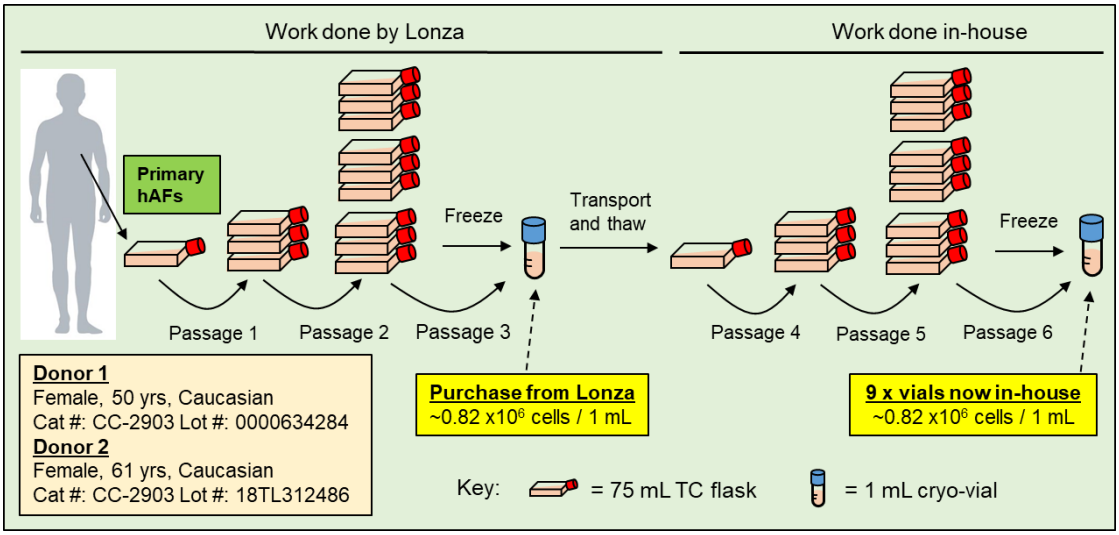


Figure 2.4 Workflow diagram showing receipt, expansion and storage of hAFs

2.2.2 Preparing hAFs for experimentation

Stored hAFs that were now on their sixth passage, were thawed and seeded (~0.82 x10^6 cells) in T75 flasks before a media change at 24 hours. At 48 hours cells were then trypsinized and seeded in either 6-, 24- or 96-well plates depending on their destined experimental treatment (*figure 2.5*).

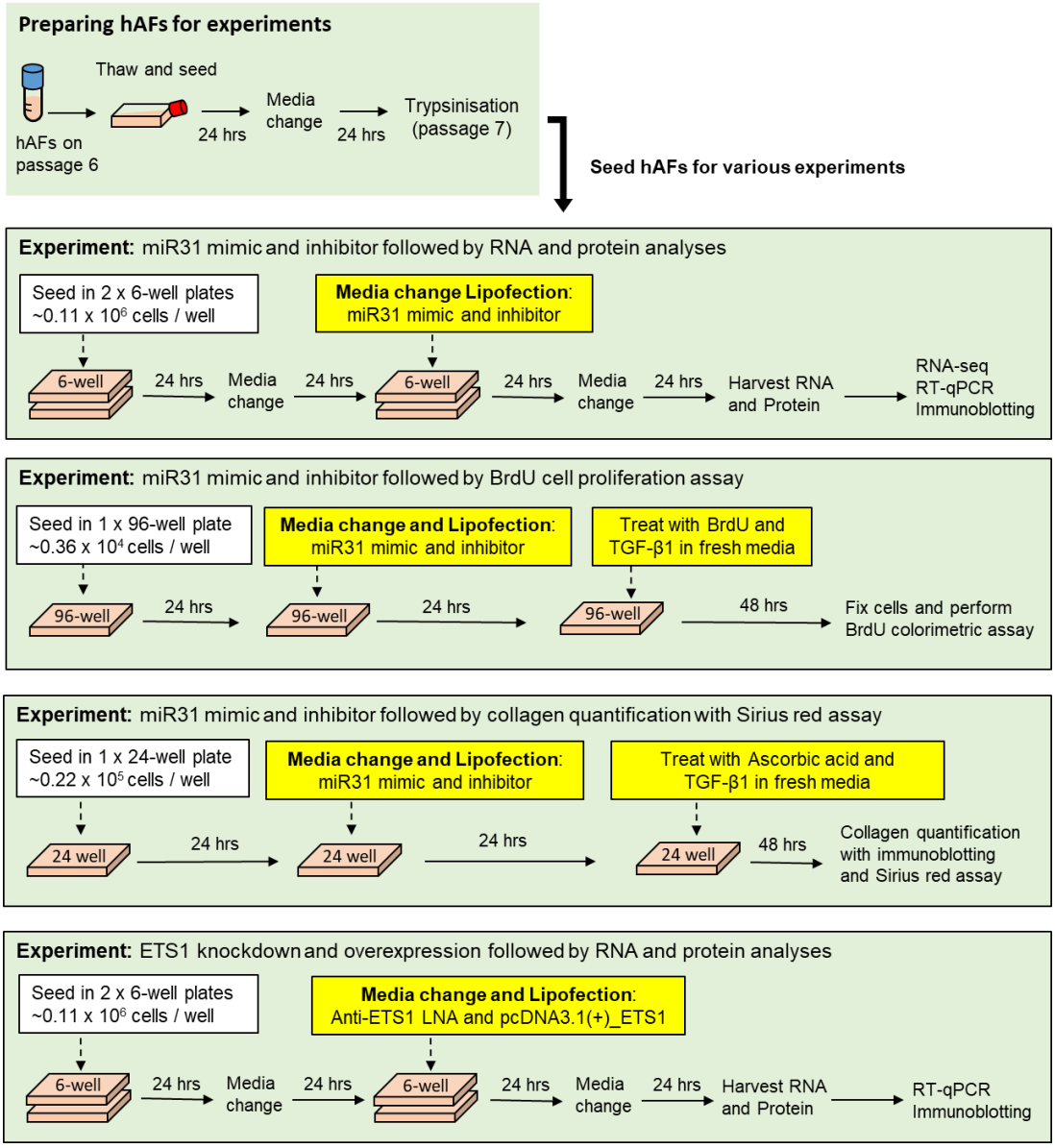


Figure 2.5 Workflow diagram showing different experiments performed with hAFs

2.2.3 Lipofection of hAFs with synthetic miR31 inhibitor and mimic for subsequent RNA and protein analyses

For overexpression and inhibition of miR31 experiments in hAFs that were intended for RNA and/or protein analyses, cells were seeded in 6-well plates (~0.11 x 10^6 cells/well) with a media change 24 hours later. Following another 24 hours, media was replaced with 1.3 mL antibiotic-free media and then hAFs were lipofected using 2.5 uL RNAiMax (ThermoFisher: 13778030) with 75 pmol of either miR31 inhibitor, inhibitor negative control, mimic or mimic negative control (Dharmacon: IH-300507-06, IN-001005-01, C-300507-05 and CN-001000-01) diluted in 300 uL Opti-MEM medium (ThermoFisher: 31985062). Media was changed 24 hours later and cells were harvested at 48 hours for RNA (*see section 2.6*) and protein (*see section 2.8*) analyses.

2.2.4 BrdU cell proliferation assay in hAFs

hAFs were seeded in a 96-well plate (~0.36 x 10^4 cells/ well) and the following day media was replaced with 40 uL antibiotic free media before cells were lipofected using 0.08 uL of RNAiMax with 2.5 pmol of either miR31-mimic, mimic-control, miR31-inhibitor or inhibitor-control (same oligonucleotides as above) diluted in 10 uL Opti-MEM. Cells were left to incubate for 24 hours, media was then replaced with 100 uL media containing bromodeoxyuridine (BrdU) (0.2 uL BrdU/ mL media) (Sigma-Aldrich: JA1595) and recombinant human TGFβ1 (5 ng / mL media) (BioLegend: 580702) and left for a further 48 hours. BrdU incorporation into DNA was then measured using BrdU-ELISA kit (Sigma-Aldrich: QIA58) whereby cells were first fixed with 4% PFA for 30 minutes, washed with PBS and probed for BrdU with 100 uL of anti-BrdU antibody (kit component: JA1599) for 1 hour at room temperature. Wells were then washed three times with 100 uL Wash Buffer and then 100 uL of Peroxidase Goat Anti-Mouse IgG HRP conjugated antibody (kit Component: JA1618) was applied before another three washes with Wash Buffer. 100 uL of HRP Substrate was then added for 15 minutes and the reaction finalized by then adding 100 uL Stop Solution. Absorbance was measured using a spectrophotometric plate reader at 500 nm (FLUOstar Omega).

2.2.5 Picro Sirius red assay for collagen detection in hAFs

hAFs were seeded in a 24-well plate (~0.22 x 10^5 cells/ well) and the following day media was replaced with 240 uL antibiotic free media before cells were lipofected using 0.5 uL of RNAiMax with 15 pmol of either miR31-mimic, mimic-control, miR31-inhibitor or inhibitor-control diluted in 60 uL Opti-MEM. Cells were left to incubate for 24 hours. Media was then replaced with 300 uL FGM-3 containing ascorbic acid (100 pmol / mL media) and TGFβ1 (5 ng / mL media). Cells were left to incubate for 48 hours, the media was then retrieved, spun at 10 000 x g for 1 minute to pellet unwanted cell debris and transferred to clean microcentrifuge tubes.

To then concentrate collagen from the retrieved media samples, collagen was precipitated by incubating media at 4 °C for 24 hours with concentrating solution (Chondrex: 90626) (250 uL / 1 mL media) which was then centrifuged at 10 000 x g for three minutes to obtain a collagen pellet that was then resuspended in 100 uL of 0.05 M acetic acid. Collagen concentration was then estimated using a Sirius Red assay collagen detection kit (Chondrex: 9062) whereby 500 uL of Sirius Red solution was added to samples which were then spun at 10 000 x g to pellet bound collagen. Supernatant was removed and the pellet washed with Wash Solution before another spin and then resuspended in Extraction Buffer. Sample absorbance was measured spectrophotometrically at 510 nM in a 96-well plate and concentration determined using a standard curve generated by making serial dilutions (250 ug/ mL - 8 ug/ ml) of bovine Type-1 collagen standard that were subject to the same precipitation protocol.

2.2.6 Knockdown of ETS1 with locked nucleic acid in hAFs

RNA-mediated interference was used to reduce ETS1 expression in hAF cells using Lipofectamine RNAiMax (ThermoFisher: 13778030) and Opti-MEM medium with 50-200 nM control Locked nucleic acid (LNA) (Qiagen: LG00000002) or anti-ETS1 LNA (Qiagen: 688573-1). hAF ells were seeded in 6-well plates (~0.11 x 10^6 cells/well) with a media change 24 hours later. Following another 24 hours, media was replaced with 1.3 mL antibiotic-free media and then in technical triplicates, hAFs were lipofected using 2.5 uL RNAiMax with 50-200 pmol of anti-ETS1 LNA or negative control diluted in 300 uL Opti-MEM medium. Media was changed 24 hours later, and cells were harvested at 48 hours for RNA and protein.

2.2.7 Overexpression of ETS1 expression vector in hAFs

For ETS1 overexpression in hAFs, procedure was identical to ETS1 knockdown except hAFs were lipofected using 2.5 uL Lipofectamine-2000 (ThermoFisher-11668019) with either 25-400 ng ETS1 expression vector (pcDNA3.1(+)_ETS1_WT) or empty pcDNA3.1(+) control vector diluted in 300 uL Opti-MEM medium.

2.2.8 Culture of HEK293 cells

HEK293 cells were a gift from our collaborator's lab (Prof Charles Redwood, RDM, Oxford) and were maintained in Dulbecco's modified Eagle medium (DMEM) (Thermo: 11960044), containing Penicillin-Streptomycin (Sigma: P4333) and supplemented with 10% FBS. Cells were cultured at 37 °C in 5% CO_2 in a humidified incubator, passaged indefinitely and media was always changed 24 hours after passaging and every 48 hours thereafter. For ETS1 overexpression in HEK293 cells protocol was identical to hAF cells (*see section 2.2.7 above*).

2.2.9 Dual-Glo firefly luciferase reporter assay in HEK293 cells

HEK293 cells were seeded in 24 well tissue culture plates (~3.3x10^5 cells/well) before a media change at 24 hours. Following another 24 hours, media was replaced with 275 µL antibiotic free 10% FBS DMEM media and lipofected using 0.5 µL Lipofectamine with 50 ng WT LOC554202 promoter *Firefly* Luciferase vector (pGL4.10[*Luc2*]_LOC554202_Promoter_WT) diluted in 25 µL Opti-MEM. To normalise for variation in transfection efficiency, cells were also transfected with 50 ng

Renilla Luciferase vector (pGL4.74) at the same time. Further, as HEK293 cells do not express endogenous ETS1, cells were transfected with 100 ng ETS1 expression vector. Dual-Luciferase Reporter assay (Promega: E2920) was then performed 48 hours after transfection by first removing media and then applying 60 µL Dual-Glo Luciferase Assay Reagent to each well for 10 minutes. *Firefly* luminescence was then measured in a spectrophotometric plate reader (FLUOstar Omega) and then 60 µL Dual-Glo Stop-and-Glo reagent was applied to each well for another 10 minutes before measuring *Renilla* luminescence.

2.3 Immunofluorescence and confocal microscopy

Immunofluorescence was performed on hAF and HEK293 cells. Cells were cultured on sterilized coverslips coated in mouse laminin (Corning: 354232), fixed with ice-cold methanol-acetone (1:1) for 5 minutes at room temperature, permeabilized by three lots of 5 minute incubation periods in PBS, and blocked in PBS containing 3% BSA. Samples were then incubated for 1 hour with primary antibodies in PBS containing 3% BSA and then washed three times for 5 minutes in PBS before incubation for 1 hour at room temperature with secondary antibody. Primary anti-ETS1 (1:300, Cell Signalling-14069) was used and was probed with secondary antibody Alexa Fluor-488 (1:300, ThermoFisher: A11070). Primary anti-FLMNA (1:200, Merck MAB1680 C) was probed with secondary antibody Alexa Fluor-568 (1:300, ThermoFisher: A11019) and primary anti-vimentin (1:200, Novus Biologicals, NBP1-97672) was probed with Alexa Fluor-488 (1:300 ThermoFisher: A11017). Cell nuclei were counter stained using TO-PRO (1:500, ThermoFisher: T3605) which was included in with the secondary antibody solution. Coverslips were then rinsed with PBS and mounted using SlowFade Gold Mountant (ThermoFisher: S36936) on glass slides, before examination with confocal microscopy (Leica: DM6000) using green, red, and far-red channels. Images were processed with ImageJ (version 1.52a).

2.4 Flow cytometry

Flow cytometry was performed on hAF cells. One million hAF cells were suspended in 1 mL Flow Cytometry Staining Buffer (ThermoFisher: 00-4222-26) and then incubated on ice for 30 minutes with 10 µg of each fluorochrome-conjugated antibody together with 1 ug of DAPI (Roche: 10236276001). Antibodies used were: Anti-CD45, PerCP-Cyanine5.5 conjugated (ThermoFisher: 45-0451-82), Anti-CD31, PE-Cyanine7 conjugated (ThermoFisher: 25-0319-41), Anti-CD90, PE conjugated (R&D Systems: FAB2067P), Anti-PDGFRα, APC conjugated (R&D Systems: FAB1264A) and Anti-VAP1, Alexa Fluor-488 conjugated (R&D Systems: IC39571G). Given the use of multiple fluorochrome-conjugated primary antibodies, and thus the potential for overlapping fluorophore emissions, we used UltraComp compensation beads (ThermoFisher: 01-2222) to generate a compensation matrix. To prepare compensation beads, a single drop of beads was added to separate microcentrifuge tubes for each fluorochrome-conjugated antibody used. One ug of each antibody was then added to their respective tubes and mixed by flicking before being incubated for 15 minutes at 4 °C in the dark. 2 mL of Flow Cytometry Staining Buffer was then added to each tube which were then mixed and centrifuged at 500 x g for 3 minutes. The supernatant was then decanted and the precipitate resuspended in 0.5 mL of Flow Cytometry Staining Buffer. Compensation beads and labelled hAF cell sample preparations were then passed through a Becton Dickinson Aria III Fusion sorter and data was analysed with BD FACSDiva software (v8.0.1). Gating strategy is shown and described in results (*section 3.1*, *figure 3.3*).

2.5 *In vivo* experiments with mice

All animal studies were conducted in accordance with the Animals (Scientific Procedures) Act 1986 (HMSO, UK) under Project Licence 30/3374 and Personal Licence IFDDF67FB which were issued and approved by the Home Office.

2.5.1 Genotyping

DNA was extracted from ear notch biopsies using *PureLink* Genomic DNA isolation kit (Thermo: K182002) according to manufacturer's instructions. Briefly, biopsies were digested in 180 µL of DNA Lysis buffer with 20 µL Proteinase-K solution for 12-24 hours at 55 °C with occasional vortexing. Homogenates were centrifuged for 1 min at 10,000 x g to pellet unwanted tissue debris and the supernatants recovered, treated with RNase and mixed with 200 µL of Binding buffer, before applying to *PureLink* Spin Columns. Columns were spun briefly and washed with Wash Buffers1 and 2 with intermediate spins. DNA was eluted in 40 µL Elution Buffer (10 mM Tris-HCl, pH 9.0, 0.1 mM EDTA) and stored at -20 °C.

2.5.1.1 ***Collagen1a2-Cre-ER(T)*** **transgene genotyping**

PCR genotyping conditions and primer design for *Col1a2-Cre-ER(T)* transgenic mice was previously optimised and published (Zheng et al. 2002) and was performed as a duplex reaction with primers amplifying both the wild type *Hprt1* allele, as an internal positive control, and the randomly integrated *Col1a2-Cre-ER(T)* transgenic allele (*see chapter 5 results figure 5.3 for transgene schematic and primer binding sites*). The *Col1a2-Cre-ER(T)* is a hemizygous strain therefore checking for zygosity was not done in these mice.

Col1a2-Cre-ER(T) **(product size: 186 bp)**

Forward (Col1a2): CAGGAGGTTTCGACTAAGTTGG

Reverse (Cre): CATGTCCATCAGGTTCTTGC

Wild type *Hprt1* third intron (product size: 314 bp)

Forward: TGAAAAAGTCCACTAATTAAAACCA

Reverse: CTAACAACCCTTTCTCTCAAGGT

PCR was performed in 96-well reaction plates (ThermoFisher: 4346907) using 2x BioMix red master mix (Bioline: Bio-25006) with primer pairs, water and $MgCl_2$. Reaction mix and thermocycling conditions are given in *Table 2.2*.

Table 2.2 Col1a2-Cre-ER(T) genotyping reaction mix and thermocycling conditions

Component		**Volume (uL)**	**Final conc./amount**
2x BioMix Red		10	1x
$MgCl_2$ (50 mM)		0.25	0.625 mM
Forward primer (10 uM)		0.5	0.25 uM
Reverse primer (10 uM)		0.5	0.25 uM
Template DNA (~30 ng/uL)		1	30 ng
Nuclease free water		7.75	
Total		20	
Thermocycling step	**Temp. (°C)**	**Time**	**Number of Cycles**
Initial Denaturation	94	10 mins	1
Touchdown	94	20 sec	10 (with -0.5 °C per cycle decrease)
	65	15 sec	
	68	10 sec	
Denaturation	94	15 sec	28
Annealing	60	15 sec	
Extension	72	10 sec	
Final extension	72	2 min	
Hold	4	Indefinite	

2.5.1.2 **Floxed pre-miR31-Inv transgene genotyping**

Genotyping PCRs were performed as separate reactions for the 5' and 3' ends for the floxed pre-miR31-Inv transgenic allele which is inserted in the first intron of the *ROSA26* locus (*see chapter 5 results figure 5.1 for transgene schematic and primer binding sites*). To query mouse zygosity, the presence of wild type *ROSA26* first intron was always checked, this was also performed as a separate reaction (i.e. not as duplex).

5'end floxed *pre-miR31-Inv* allele (product size: 602 bp)

Forward (CAG): CAGCCATTGCCTTTTATGGT

Reverse (Neomycin): GTTGTGCCCAGTCATAGCCGAATAG

3'end floxed *pre-miR31-Inv* allele (product size: 1257 bp)

Forward (GFP): CGGGAGAAATGGATATGAAGTACTGGGC

Reverse (ROSA26, remaining first intron): ACACGCTGAACTTGTGGCCG

Wild type *ROSA26* first intron (product size: 603 bp)

Forward: AAAGTCGCTCTGAGTTGTTAT

Reverse: GGAGCGGGAGAAATGGATATG

For the *pre-miR31-Inv* transgene PCR was performed using Immolase Taq Polymerase (Bioline: 21046) in a reaction mix including betaine (Sigma: BO300), dNTPs (Bioline: BIO-39044), immolase buffer, primer pairs, $MgCl_2$ and water. Reaction mix and thermocycling conditions are given in *Table 2.3* below. For the wild type *ROSA26,* reaction mix and thermocycling was identical to *Table 2.2* above.

Table 2.3 pre-miR31-Inv transgenic allele genotyping reaction mix and thermocycling conditions for 5' and 3' ends

Component		**Volume (µL)**	**Final conc./amount**
10x Immolase Buffer		2	1x
$MgCl_2$ (50 mM)		0.6	1.5 mM
Primer 1 (10 uM)		1	0.5 uM
Primer 2 (10 uM)		1	0.5 uM
dNTPs (100 mM)		0.3	1.5 mM
Immolse Taq Polymerase		0.2	
Betaine (5 M)		4	1 M
Nuclease free water to final volume		8.9	
DNA (~30 ng/µL)		2	60 ng
Total		20	
5' end Thermocycling (Hygro-30)	**Temp. (°C)**	**Time**	**Number of Cycles**
Initial Denaturation	95	10 min	1
Denaturation	94	1 min	
Annealing	57	1 min	35
Extension	72	30 sec	
Final Extension	72	10 min	1
Hold	4	Indefinite	
3' end Thermocycling (Hygro-90)	**Temp. (°C)**	**Time**	**Number of Cycles**
Initial Denaturation	95	10 min	1
Denaturation	94	1 min	
Annealing	57	1 min	35
Extension	72	90 sec	
Final Extension	72	10 min	1
Hold	4	Indefinite	

Amplified PCR products were then mixed with 6 x DNA loading buffer (NEB: B7025S) and loaded onto 1.5% agarose TBE gel containing 0.002% ethidium bromide and electrophoresed at 120 V for 35 min and visualized by UV transluminescence using ChemiDoc imaging software (version 5.1). No template control (NTC) was used to monitor contamination and primer-dimer formation. Invitrogen 1 kb plus (10787018) was used to mark band size.

2.5.1.3 **Breeding strategies**

The *Col1a2-Cre-RT(T)* mouse line exists as hemizygous animals. Our first breeding pair was purchased from JAX lab (B6.Cg-Tg(Col1a2-cre/ERT,-ALPP)7Cpd/J) and included a WT male and hemizygous female. As this strain was generated by pro-nuclear injection with random integration of the transgene, no hemizygous mice were crossed together in order to ensure gene dosage was controlled. Colonies were maintained by crossing *either* male or female hemizygous genotypes with wild types. Pre-miR31-Inv transgenic mice were developed by the Davies lab, Oxford (WTCHG, Transgenic core facility) and colonies were initially maintained by crossing *either* male or female heterozygous mice with wild types. To improve efficiency, heterozygous males and females were frequently paired. In addition, homozygous males proved to be viable studs and were therefore often mated with heterozygous or wild type females. To create *Col1a2-Cre-ER(T) : pre-miR31-Inv* progeny, heterozygous *pre-miR31-Inv* mice were crossed with hemizygous *Col1a2-Cre-ER(T)* mice.

2.5.2 Tamoxifen preparation and administration

Tamoxifen in corn oil solution was prepared by first dissolving 80 mg tamoxifen (Sigma: T5648) in 750 uL sterile 100% ethanol (passed through 0.22 µm syringe filters) heated at 50 °C for 5 minutes with occasional inversion to aid solubilisation. 3.25 mL corn oil was then added to dissolved tamoxifen to make 4 mL stock (20 mg tamoxifen/ mL) and 1 mL aliquots were stored at -20 C for future use. Considering tamoxifen's photosensitivity, this preparation was made in opaque tubes and in the dark where possible.

Intra-peritoneal administration of tamoxifen was optimised using a time-course experiment whereby mice at eight weeks of age were injected either with plain corn oil (Sigma: C8267) or with 2 mg tamoxifen/ 30 g mouse (Sigma: T5648) for either four or seven consecutive days. This regiment was chosen considering a published protocol that suggested an optimal dose of tamoxifen for Cre activity (Pillai et al. 2017). Tamoxifen was administered using 26G needle and 1 mL syringes and following optimisation, regimen was kept to seven consecutive days with the same dose.

2.5.3 Checking for genomic inversion of the *pre-miR31-Inv* transgene.

Following tamoxifen administration, the *pre-miR31-Inv* transgene is inverted in fibroblast cells. This inversion and recombination event was checked using gDNA extracted from ear tissue obtained after the two-week tamoxifen washout period (*see figure 5.8 in Chapter 5 results for experimental work flow diagram*). Inversion at the 5' and 3' ends was confirmed using the Hygro-90 and Hygro-30 protocols respectively (these protocols are shown in *Table 2.3* above).

5' end of inverted *pre-miR31-Inv* allele (product sizes: 1221 and 1143 bp)

Forward (CAG): CAGCCATTGCCTTTTATGGT

Reverse (GFP) : CGGGAGAAATGGATATGAAGTACTGGGC

3' end of inverted *pre-miR31-Inv* allele (product sizes: 716 and 638 bp)

Forward (Neo): GTTGTGCCCAGTCATAGCCGAATAG

Reverse (ROSA26): ACACGCTGAACTTGTGGCCG

2.5.4 Assessment of murine cardiac function with echocardiography

2.5.4.1 **Anaesthesia**

General anaesthesia in mice was induced with a 4% isoflurane/oxygen mixture using a Benson anaesthetic machine and chamber. Absence of righting and toe-pinch reflex with depressed respiratory rate were used to confirm adequate depth of anaesthesia. Mice were weighed and then fixed with tape in the supine position on a 37 °C heated homeothermic stage that also monitors ECG using in-build leads. Ointment was applied to the eyes to prevent drying (Novartis: Viscotears Liquid Gel). Chest fur was removed with an electric shaver and a layer of acoustic coupling gel was applied. Throughout the imaging protocol anaesthesia was maintained with varying concentrations of isoflurane administered through a nosecone and HR rate was monitored using ECG. Approximately 1.2% isoflurane induced heart rates appropriate for image acquisition (400-450 bpm), however isoflurane concentrations were carefully titrated to elevate or reduce heart rate when needed.

2.5.4.2 **Echocardiography**

Echocardiographic images were acquired with the VisualSonics Vevo 2100 system fitted with an ultra-high frequency MS-550D MicroScan transducer (22-55 MHz). Parasternal short and long axes images were obtained at mid-papillary level, mice were then recovered from anaesthesia in a heated incubator. Image analysis was blinded between control and treatment groups and performed on Vevo Lab, Edition 1.7.1, Imaging System (VisualSonics, Toronto, Canada). Short axis B-mode images from 3-5 cardiac cycles were averaged and used to estimate left ventricular *epicardial* end-diastolic area (LV-Epi-EDA) and *endocardial* end-diastolic (LV-Endo-EDA) and end-systolic area (LV-Endo-ESA). Alone, LV-Endo-EDA measurements are used as a surrogate for left ventricular end-diastolic *volume* (LVEDV) and hence dilation (Lindsey et al. 2018). Further, myocardial cross-sectional area (myoCSA) was calculated as a measure of hypertrophy (as it is a surrogate for LV mass), as well as fractional area change (FAC %), as a measure of ejection fraction (Collins et al. 2003):

$$MyoCsa\ (mm^2) = \text{LV-Epi-EDA} - \text{LV-Endo-EDA}$$

$$FAC\ \% = \frac{\text{LV-Endo-EDA} - \text{LV-Endo-ESA}}{\text{LV-Endo-EDA}} \times 100$$

2.5.5 Harvesting mouse tissue

Seven days following cessation of tamoxifen mice were culled by Schedule 1 cervical dislocation, the chest cavity was then opened to expose the heart which was then excised and washed with ice-cold PBS. The left and right ventricle, and both atria together, were then dissected and placed in individual cryo-tubes and snap frozen in liquid nitrogen in cryogenic vials then stored at -80 °C until required.

2.5.6 Hydroxyproline assay

Tissue was kept on dry ice and pulverised to powder using a spatula and divided into aliquots of 10 mg and placed in glass screw-thread vials (ThermoFisher: B7999-1). 100 uL of distilled water and of HCL (10 M) was added to the vials which were then incubated at 120 ℃ for 24 hours to hydrolyse the samples. 10 uL of each hydrolysed tissue sample was then plated in duplicates onto a 96-well plate and hydroxyproline was then modified by adding 100 uL chloramine T solution, followed by 100 uL DMAB solution (Chondrex: 6017). The plate was incubated at 60 ℃ for 30 minutes which results in a coloured product and absorption was measured at 570 nm using a plate reader. The concentration of hydroxyproline was then calculated from a standard curve from serially diluted (400 – 6.3 ug/mL) pure hydroxyproline standard subject to the same protocol.

2.6 RNA extraction, cDNA synthesis and qPCR

Total RNA was extracted from mouse tissues, as well as hAF- and HEK293-cell cultures using *mirVana* miRNA Isolation Kit (Thermo: AM1560). For mouse tissue RNA extraction, samples were lysed in 300 uL Lysis/Binding Buffer containing β-Mercaptoethanol using pre-cooled stainless-steel beads vibrating at 50 Hz in a tissue pulveriser (Qiagen: TissueLyser LT). For hAF and HEK293 cells cultured in 6-well plates, cells were lysed in each well using 500 uL of the same Lysis/Binding Buffer and mechanically homogenised with a pipette tip. Lysed samples were then transferred to micro centrifuge tubes for vortexing briefly and treated with 0.1 volumes of homogenate additive. Total RNA was then isolated by organic extraction whereby Acid-Phenol:Chloroform was added to lysates (1:1) which were then vortexed and centrifuged to retrieve the upper aqueous phase containing RNA. RNA was then precipitated using 1.25 volumes of room temperature 100% ethanol and applied to miRNA Isolation Kit Filter Cartridges that were then washed with Buffers 1 and 2 with intervening spins. RNA was eluted in 40 uL of nuclease-free water heated to 95 °C and were then stored at -80 °C.

To generate a library of cDNA from RNA transcripts containing *polyA* tails, first-strand cDNA was synthesised using QuantiTect Reverse Transcriptase kit (Qiagen: 205313). 400 ng of extracted total RNA made up in 12 uL nuclease free water was initially treated with 2 uL of 7x Wipeout Buffer for 2 minutes at 42 °C to digest any residual genomic DNA. The resultant 14 uL RNA solution was then combined with 8 uL of master mix which contains propriety quantities of: Quantiscript Reverse Transcriptase, reverse transcriptase buffer, dNTPs and oligo-dT primers. The final 20 uL reaction mix was then incubated at 42 °C for 15 minutes for cDNA synthesis by reverse transcription followed by heating at 95 °C for 3 minutes for enzyme deactivation.

For MicroRNA transcripts, cDNA was made using TaqMan MicroRNA Reverse Transcription Kit (Thermo: 4366596). 150 ng of extracted total RNA was mixed with propriety quantities of microRNA specific stem-loop primers, MultiScribe Reverse Transcriptase, RNase Inhibitor, buffer and dNTPs to a final volume of 15 uL. The reaction mix was then incubated at 16 °C for 30 minutes to facilitate primer hairpin folding, then 42 °C for 30 minutes for primer annealing and reverse transcription, and then at 85 °C for 5 minutes for enzyme deactivation.

For qPCR, protocol was identical for cDNA derived from polyA mRNA and for cDNA derived from microRNA transcripts. Two uL of cDNA[2] was added to 3 uL TaqMan Fast Advanced Master Mix (Thermo: 4444557) including housekeeping and gene-of-interest (GOI) reporter probes (*see section 2.11 for all probes used*) to make up a final volume of 5 uL. Thermocycling was then performed in a 96-well plate using QuantStudio 7 Flex Real Time PCR System (ThermoFisher: 4485701). The temperature profile included initial denaturation (20 secs at 95 °C) and 40 cycles of denaturation (1 sec at 95 °C) with annealing/elongation (20 secs at 60 °C). Each reaction was performed in duplicates and data recorded with QuantStudio 7 Software (version 1.2). The amount of GOI target, normalized to the endogenous reference gene, was calculated using the delta Ct method (2^ΔCt) where ΔCt is the difference in PCR cycle number between the GOI and reference gene at some pre-defined fluorescence threshold in the exponential phase of template amplification. This calculation was performed under the assumption that in the exponential phase of PCR amplification, copy number increased two-fold per cycle for both the GOI and reference gene. To then determine the fold-change of gene transcript following an experimental intervention, the amount of target, normalized to the endogenous reference was then normalized further to the average of control group values (2^ΔΔCt) (Schmittgen et al. 2008).

[2] 2uL of normal cDNA contains ~20 ng input RNA converted to cDNA. Conversely, 2uL of microRNA cDNA contains ~10 ng input RNA converted to cDNA.

2.7 RNA sequencing

Total RNA was extracted from hAF cell cultures using *mirVana* miRNA Isolation Kit (Thermo: AM1560) as above and eluted in 40 uL nuclease-free water. Samples were then treated with RNase-Free DNase Set (Qiagen: 79254) whereby 5 µL of Buffer RDD was added to each sample followed by 1.25 µL of DNase I stock solution (2.7 Kunitz units / uL). RNA clean-up was then performed by precipitating RNA with 1.25 volumes 100% ethanol which was then applied to fresh *mirVana* miRNA Isolation Kit columns (Thermo: AM1560) that were then washed as above before eluting clean RNA. Sample replicates with 260/230 and 260/280 NanoDrop spectrophotometric absorbance ratios of ~2 were selected.

750 ng of RNA per sample was made up in 30 µL nuclease-free water which was then used for RNA-seq transcriptional profiling which was performed as a service at the genomic core facility of the Wellcome Centre for Human Genetics (Oxford). Total RNA quantity and integrity were assessed there using Quant-IT RiboGreen RNA Assay Kit (ThermoFisher: R11490) and Tapestation system (Agilent). Samples were then normalised to 100 ng. RNA was enriched for protein encoding mRNA transcripts by using poly-adenosine oligonucleotides conjugated to magnetic beads with NEBNext Magnetic Isolation Module (E7490). Generation of double stranded cDNA and library construction were performed with NEBNext Ultra II Directional RNA Library Prep Kit for Illumina (E7760L) with their own adapters and barcode tags (dual indexing, based on DOI: 10.1186/1472-6750-13-104). The concentrations used to generate the multiplex pool were determined by Picogreen. The final size distribution of the pool was determined using a Tapestation (Agilent), and quantification was determined by Qubit (Thermo) before sequencing on an Illumina NovaSeq6000 as 150bp paired end.

Sequencing data was then aligned to hg38 and quality checked (see appendices for quality control data). Differential gene expression was then analysed using edgeR statistical package (Bioconductor release 3.12) which generated a list of gene sequencing read counts, p-value, adjusted p-value (p-value corrected for multiple tests), likelihood ratio (LR) and fold change relative to control. Gene lists were then analysed using Ingenuity pathways analysis (IPA) (Qiagen version 42012434). The miR31 mimic vs. control gene list was analysed by first applying a cut-off of adj. $p < 0.00000001$. Conversely, adj. $p < 0.05$ was used for inhibitor vs. control list.

2.8 Protein extraction and immunoblotting

For cultured hAF and HEK293 cells, cells were lysed with 350 µL of ice-cold protein lysis buffer (Sigma: R0278) containing tablet protease (Roche: 04693159001) and phosphatase inhibitors (Roche: 4906845001), and homogenised mechanically on plate using the back-end of a pipette tip before being transferred to micro centrifuge tubes. For mouse and human tissue protein extraction, tissue was lysed in the same buffer (300 µL buffer per 5 mg tissue) using pre-cooled stainless-steel beads vibrating at 50 Hz in a tissue pulveriser (Qiagen: TissueLyser LT). Cell lysate and tissue lysate samples were then sonicated for 30 secs three times using a table-top spectacle-cleaner (Bransonic: B200). Protein concentration was estimated using Pierce Protein Assay Kit (ThermoFisher: 23227) where initially a stock working reagent was made by combining 50-parts solution A with 1-part solution B. 20-parts of working solution were then added to 5-parts of serial dilutions (2 - 0.125 mg/µl) of bovine serum albumin (BSA) as standard, and to protein samples. Concentrations were then determined spectrophotometrically at 562 nM in a 96-well plate (FLUOstar Omega). Protein samples were stored temporarily at -20 °C and long term at -80 °C.

Cell lysates were loaded into NuPAGE 4-12% Bis-Tris Protein Gels (ThermoFisher: NP0335) (10 ug/well), electrophoresed in Nupage MOPS SDS running buffer (ThermoFisher: NP0001) and transferred in transfer buffer: (ThermoFisher: NP0006) onto a nitrocellulose membrane (ThermoFisher: 88018) activated briefly in 100% methanol. The membrane was then washed in PBST (0.1%) for 10 minutes before blocking with 5% non-fat milk in PBST (0.1%) overnight at 4 °C. Primary and secondary horse radish peroxidase (HRP)-conjugated antibodies were applied in 5% non-fat milk in PBST (0.1%) for 1 hour before washing three times for 10 minutes in PBST (0.1%) after each antibody.

Anti-ETS1 (1:1000, Cell Signalling: 14069), anti-COL1 (1:1000, Merck: 234167), anti-STK40 (1:1000, Novus: NBP1-33391), anti-COL3A1 (1:1000, Novus: NBP2-15946), anti-TNS1 (1:1000, Novus: NBP1-84129), anti-BAP1 (1:1000, Cell Signalling: 13187S), anti-αSMA (1:1000, Sigma: A5228), anti-FN1 (1:1000, Sigma: F3648), anti-CALCR (1:1000, Bio-Rad: AHP635) and anti-FGF1 (Abcam: EPR19989) primary antibodies were used and were probed for with anti-IgG HRP-conjugated secondary antibody (1:2500, Promega: W4011). HRP conjugated anti-GAPDH served as loading control (1:2500 Sigma: G9295). Subsequently enhanced chemiluminescence (ECL) SuperSignal West Dura (ThermoFisher: 34075) was used as substrate for HRP and membranes were then developed using X-ray films and analysed with ImageJ software (version 1.52a). For stripping, nitrocellulose membranes were submerged in Restore Plus Striping buffer (ThermoFisher: 46430) for 10 minutes and then washed briefly with distilled tap water and then in PBST (0.1%) for 10 minutes on an orbital shaker.

2.9 Patient Cohorts and atrial fibroblast isolation

Studies involving human participants were approved by the local Research Ethics Committee (South Central - Berkshire B Research Ethics Committee, UK; ref: 07/Q1607/38). Patients undergoing cardiac surgery for either coronary artery bypass grafting or valve repair/replacement in the John Radcliffe Hospital, Oxford, were invited to take part in the study. Following informed written consent, the right atrial appendage was obtained during cardiac surgery from anaesthetised patients at the *beginning* of the surgical procedure. Samples were rinsed of blood and then for fibroblast isolation from the right atrial appendage, tissue was cut into 2–3 mm^3 pieces and digested using 4 mg/ml collagenase II and 0.0625% trypsin in DMEM. Cells were then washed two times with PBS and plated in 6-well plates in 10% FBS FBM-3 medium (Lonza: CC-3131). Media was renewed every 2-3 days, cells were passaged at ~80–90% confluence and cultured in a humidified cell incubator set at 37 °C and 5% CO_2.

2.10 Experimental design and statistics

All statistical analyses were performed using GraphPad Prsim (version 7.05) except for RNA-seq data (*see section 2.7 above for RNA-seq analysis).* Experiments involving purchased hAFs (Lonza: CC-2903) were performed on cells from single donors, therefore individual data points are not biological replicates, but instead represent experimental repeats that were performed on different weeks. Data derived from these experiments were checked for normality using the Kolmogorov-Smirnov (KS) statistical test. Parametric data was then analysed using a two-tailed *paired* Student's *t*-test while non-parametric data were compared using a paired Wilcoxon test. Conversely, other experiments involved the purification of hAFs from separate patient donors (sinus rhythm controls vs. arrhythmia disease group) and in such cases, each data point represented a biological replicate and groups were compared using an *unpaired* Student's *t*-test as data was parametric.

For experiments involving mice, each data point represents an individual mouse. For echocardiography, all data was analysed blinded to mouse genotypes that were decoded following completion of measurements. To correct for an imbalance in gender distribution, which is a principal driver of weight distribution, genotype groups contained equal number of males and females. The distribution of all measured heart dimensions in mice were first checked for normality using the KS test. Means of normally distributed data were compared to *WT/WT* genotype control using one-way ANOVA with Dunnett's correction for multiple comparison of means. Non-parametric data were compared to *WT/WT* genotype group using Kruskal-Wallis test with Dunn's correction for multiple comparison of means. In the case of paired echocardiographic data, groups were compared using a paired Student's *t*-test or with a paired Wilcoxon test if data was non-parametric. Conversely, for gene expression studies, immunoblot and hydroxyproline assay in mice atria, an *unpaired* Student's *t*-test was used to compare the *miR31/Cre* group to *WT/Cre* control group when data was parametric, while the Mann–Whitney U test was used for non-parametric data.

For echocardiography image analysis training, the difference between the expert's and student's analysis, together with the average of those same measurements was determined using the Bland-Altman statistical test. Data points determined to be outside the 95% limits of agreement (±2 SD from the mean of differences) of the Bland-Altman plot were then reanalysed with additional coaching from the expert.

2.11 Materials and reagents

Antibodies

Name	Supplier	Identifier
Primary antibodies		
Anti-ETS1 (Monoclonal rabbit anti-human)	Cell Signalling	14069
Anti-FLMNA (Monoclonal mouse anti-human)	Merck	MAB1680-C
Anti-COL1 (Polyclonal mouse anti-rabbit)	Merck	234167
Anti-STK40 (Polyclonal rabbit anti-human)	Novus Biologicals	NBP1-33391
Anti-COL3A1 (Polyclonal rabbit anti-human)	Novus Biologicals	NBP2-15946
Anti-TNS1 (Polyclonal rabbit anti-human)	Novus Biologicals	NBP1-84129
Anti-BAP1 (Monoclonal rabbit anti-human)	Cell Signalling	13187S
Anti-αSMA (Monoclonal mouse anti-human)	Sigma	A5228
Anti-FN1 (Polyclonal rabbit anti-human)	Sigma	F3648
Anti-CALCR (Polyclonal rabbit anti-rat)	Bio-Rad	AHP635
Anti-FGF1 (Monoclonal rabbit antibody)	Abcam	EPR19989
Anti-Vimentin (Monoclonal mouse anti-human)	Novus Biologicals	NBP1-97672
Anti-VAP1, Alexa Fluor-488 conjugated (Monoclonal mouse anti-human)	R&D Systems	IC39571G
Anti-GAPDH, HRP-conjugated (Monoclonal mouse anti-human)	Sigma	G9295

Anti-CD45, PerCP-Cyanine5.5-conjugated (Monoclonal rat anti-mouse)	ThermoFisher	45-0451-82
Anti-CD31, PE-Cyanine7-conjugated (Monoclonal mouse anti-human)	ThermoFisher	25-0319-41
Anti-CD90, PE-conjugated (Monoclonal mouse anti-human)	R&D Systems	FAB2067P
Anti-PDGFRα, APC-conjugated (Monoclonal mouse anti-human)	R&D Systems	FAB1264A
Secondary antibodies		
HRP-conjugated Anti-IgG secondary antibody (Anti rabbit (H+L) IgG, unspecified source)	Promega	W4011
Alexa Fluor-488 secondary antibody (Polyclonal goat anti-rabbit (H+L) IgG)	ThermoFisher	A-11070
Alexa Fluor-488 secondary antibody (Polyclonal goat anti-mouse (H+L) IgG)	ThermoFisher	A-11017
Alexa Fluor-568 secondary antibody (Polyclonal goat anti-mouse (H+L) IgG)	ThermoFisher	A-11019

Commercial assays and kits

Name	Supplier	Identifier
TaqMan Gene Expression Master Mix	ThermoFisher	4369016
Pierce™ BCA Protein Assay Kit	ThermoFisher	23225
mirVana miRNA Isolation Kit	ThermoFisher	AM1561
RNase-Free DNase Set	Qiagen	79254
Hydroxyproline Assay Kit	Chondrex Inc	6017
Sirius Red Collagen Detection Kit	Chondrex Inc	9062
BrdU Cell Proliferation Assay	Calbiochem	QIA58

Machines and software

Name	Supplier
GraphPad Prism version 7.0a	GraphPad Software Inc.
Vevo Lab, Edition 1.7.1	VisualSonics Inc
Real-Time PCR Software V1.2	ThermoFisher
QuantStudio™ 7 Flex (qPCR Machine)	ThermoFisher

Small interfering oligonucleotides

Name	Sequence	Supplier	Identifier
miR31 mimic	Not supplied	Dharmacon	C-300507-05
Control mimic	Not supplied	Dharmacon	CN-001000-01
miR31 inhibitor	Not supplied	Dharmacon	IH-300507-06
Control inhibitor	Not supplied	Dharmacon	IN-001005-01
Anti-ETS1 LNA	(G*T*G*G*A*C*T*T*A*G*A*G*A*G*A*A)	Qiagen	LG00190128
Control LNA	(A*A*C*A*C*G*T*C*T*A*T*A*C*G*C)	Qiagen	LG00000002

*****phosphorothioate backbone modification

Human TaqMan gene expression assays

Gene symbol	Identifier
ACTB	Hs01060665_g1
BAP1	Hs01109276_g1
ETS1	Hs00428293_m1
FLNA	Hs00924645_m1
HG-miR31	Hs01007339_g1
HPRT1	Hs02800695_m1
miR31	02279
MPRS18A	Hs01026310_m1
PKN2	Hs00178944_m1
PPIA	Hs04194521_s1
pri-miR31	Hs03302684_pri
RNU58A	001207
STK40	Hs008894269_m1
TBP	Hs00427620_m1
TGFβ1	Hs00998133_m1
TGFβ2	Hs00234244_m1
TNS1	Hs00917032_m1
UCN2	Hs00264218_s1
U6, small nucleolar RNA	001973

Mouse TaqMan gene expression assays

Gene symbol	Identifier
Acta2 (actin)	Mm01546133_m1
Actn2	Mm00473657_m1
Bap1	Mm00550845_m1
Calcr	Mm00432282_m1
Col1a1	Mm00801666_g1
Col3a1	Mm01254476_m1
Ctgf	Mm01192933_g1
Fgf1	Mm00438906_m1
Fn1	Mm01256744_m1
Gapdh	Mm99999915_g1
Hprt1	Mm03024075_m1
Itgb3	Mm00443980_m1
Mmp2	Mm00439498_m1
Mmp9	Mm00442991_m1
Mprs18a	Mm00452788_m1
Pdgfb	Mm00440677_m1
Pkn2	Mm00618304_m1
Postn	Mm01284919_m1
Ppia	Mm02342430_g1
Stk40	Mm00512134_m1
Tbp	Mm01277042_m1
Tgf-β1	Mm01178820_m1
Tgf-β2	Mm00436955_m1
Timp4	Mm01184417_m1
Tns1	Mm00452886_m1
Vegf-b	Mm00442102_m1

2.12 Primers

Cloning primers

Name	Sequence (5'-3')
hsa-ETS1 forward cloning primer with Kpn1 docking and cleavage site	(TTTTTT)(AAGCTT)ATGAAGGCGGCCGTCGAT
hsa-ETS1 reverse cloning primer with Not1 docking and cleavage site	(AAAAAA)(GCGGCCGC)CTACTCGTCGGCATCTGGCTT
hsa-LOC554202 forward cloning primer XhoI docking and cleavage site	(TTTTTT)(CTCGAG)TGGGCTTGCGGAGGGCAT
hsa-LOC554202 reverse cloning primer with HindIII docking and cleavage site	(AAAAAA)(AAGCTT)GCGAGCCCCTGGCGGCGC

hsa-LOC544202 promoter mutagenesis primers

Name	Sequence (5'-3')
miR-31_A_25_F	TCAGTATTTA(TTTTT)AGAGATGGGC
miR-31_A_25_R	GCCCATCTCT(AAAAA)TAAATACTGA
miR-31_A_31_F	CCCTCAGTATTTA(TTTTT)AGAGATGGGCAAC
miR-31_A_31_R	GTTGCCCATCTCT(AAAAA)TAAATACTGAGGG
miR-31_A_25_F	TCAGTATTTA(TTTTT)AGAGATGGGC
miR-31_A_25_R	GCCCATCTCT(AAAAA)TAAATACTGA
miR-31_A_31_F	CCCTCAGTATTTA(TTTTT)AGAGATGGGCAAC
miR-31_A_31_R	GTTGCCCATCTCT(AAAAA)TAAATACTGAGGG
mIR-31_B_25_F	TCTTCATGTT(TTTTT)AATTCAAACC
miR-31_B_25_R	GGTTTGAATT(AAAAA)AACATGAAGA
miR-31_B_31_F	TAGTCTTCATGTT(TTTTT)AATTCAAACCCAA
miR-31_B_31_R	TTGGGTTTGAATT(AAAAA)AACATGAAGACTA
miR-31_C_25_F	AAGGGGCAGT(TTTTT)GGTTCAGAGC
miR-31_C_25_R	GCTCTGAACCA(AAAAA)CTGCCCCTT
miR-31_C_31_F	CTGAAGGGGCAGT(TTTTT)GGTTCAGAGCACC
miR-31_C_31_R	GGTGCTCTGAACCA(AAAAA)CTGCCCCTTCAG
miR-31_D_25_F	CAAAGAGGGC(TTTTT)CAAATCTGGG
miR-31_D_25_R	CCCAGATTTG(AAAAA)GCCCTCTTTG

miR-31_D_31_F	TCCCAAAGAGGGC(TTTTT)CAAATCTGGGCGC
miR-31_D_31_R	GCGCCCAGATTTG(AAAAA)GCCCTCTTTGGGA
miR-31_E_25_F	GTGGGTACCC(TTTTT)AAATTCTTGG
miR-31_E_25_R	CCAAGAATTT(AAAAA)GGGTACCCAC
miR-31_E_31_F	CAGGTGGGTACCC(TTTTT)AAATTCTTGGTGG
miR-31_E_31_R	CCACCAAGAATTT(AAAAA)GGGTACCCACCTG
miR-31_F_25_F	GTCCGACCAA(TTTT)TACTTCCGCT
miR-31_F_25_R	AGCGGAAGT(AAAAA)ATTGGTCGGAC
miR-31_F_31_F	GGGGTCCGACCAA(TTTTT)TACTTCCGCTGTT
miR-31_F_31_R	AACAGCGGAAGT(AAAAA)ATTGGTCGGACCCC
miR-31_G_25_F	CAAGGAAGTA(TTTTT)GCTGTTCAAT
miR-31_G_25_R	ATTGAACAGC(AAAAA)TACTTCCTTG
miR-31_G_31_F	GACCAAGGAAGTA(TTTTT)GCTGTTCAATTTG
miR-31_G_31_R	CAAATTGAACAGC(AAAAA)TACTTCCTTGGTC
miR-31_H_25_F	GTGGCCGCCG(TTTTT)CCCACAAAAT
miR-31_H_25_R	ATTTTGTGGG(AAAAA)CGGCGGCCAC
miR-31_H_31_F	GTAGTGGCCGCCG(TTTTT)CCCACAAAATTCC
miR-31_H_31_R	GGAATTTTGTGGG(AAAAA)CGGCGGCCACTAC
miR-31_I_25_F	CCCCCACAAA(TTTTT)GTCACTCGGG
miR-31_I_25_R	CCCGAGTGAC(AAAAA)TTTGTGGGGG
miR-31_I_31_F	CTTCCCCCACAAA(TTTTT)GTCACTCGGGAGC
miR-31_I_31_R	GCTCCCGAGTGAC(AAAAA)TTTGTGGGGGAAG
miR-31_J_25_F	CCAGGCGCCT(TTTTTT)TCCCCCTCC
miR-31_J_25_R	GGAGGGGGA(AAAAAA)AGGCGCCTGG
miR-31_J_31_F	CGTCCAGGCGCCT(TTTTTT)TCCCCCTCCTTT
miR-31_J_31_R	AAAGGAGGGGGA(AAAAAA)AGGCGCCTGGACG
miR-31_K_25_F	CTCCCCCTCC(TTTTT)GCGGCGGTCG
miR-31_K_25_R	CGACCGCCGC(AAAAA)GGAGGGGGAG
miR-31_K_31_F	TCCCTCCCCCTCC(TTTTT)GCGGCGGTCGCGG
miR-31_K_31_R	CCGCGACCGCCGC(AAAAA)GGAGGGGGAGGGA

Mouse genotyping primers

Mouse strain/ allele	Sequence (5'-3')	Product size
Col1a2-Cre-ER(T)		
Col1a2-F	F: CAGGAGGTTTCGACTAAGTTGG	186 bp
Cre-ER(T)-R	R: CATGTCCATCAGGTTCTTGC	
5' miR-31-Inv		
Cag-F	F: CAGCCATTGCCTTTTATGGT	602 bp
Neo-R	R: GTTGTGCCCAGTCATAGCCGAATAG	
3' miR-31-inv		
eGFP-R	F: ACACGCTGAACTTGTGGCCG	1257 bp
ROSA-R	R: CGGGAGAAATGGATATGAAGTACTGGGC	
WT ROSA26	F: AAAGTCGCTCTGAGTTGTTAT	603 bp
	R: GGAGCGGGAGAAATGGATATG	
eGFP	F1: CTTGTACAGCTCGTCCATGC	229 bp
	R2: GGTGAACTTCAAGATCCGCC	
	F2: TCCTTGAAGTCGATGCCCTT	167 bp
	R2: CACATGAAGCAGCACGACTT	
WT Hprt1	F: TGAAAAAGTCCACTAATTAAAACCA	314 bp
	R: CTAACAACCCTTTCTCTCAAGGT	

2.13 Contributions from other lab members

Lucia Moreira performed immunofluorescence hAF stain of vimentin and filamin A. Neelam Mehta performed the FACS experiment. Kathryn Aguilar-Agon performed the hydroxyproline assay and all immunoblotting in Section 5.7.5.

Chapter 3

Results: Investigating the effects of miR31 manipulation in human atrial fibroblasts

Our group has previously shown that atrial myocytes derived from patients with chronic AF have increased miR31 abundance and that miR31 begets electrical remodelling in these cells by depleting dystrophin and nNOS (Reilly et al. 2016). After this study, we queried whether miR31 expression in fibroblasts derived from atrial tissue from patients with AF showed a similar change in abundance. For this pilot study, the right atrial appendage (RAA) was excised from patient donors with chronic AF (cAF) or with sinus rhythm (SR), fibroblasts were then isolated from the RAA and RNA was then extracted from fibroblasts to query miR31 expression by RT-qPCR. Results indicated a significant increase in miR31 in patients with chronic AF compared to patients with SR (*figure 3.1*).

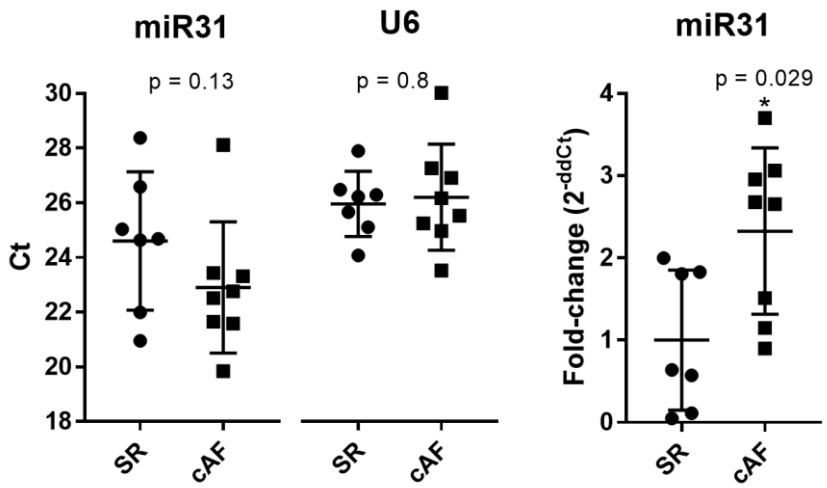


Figure 3.1 ***miR31 abundance in fibroblasts derived from the right atrial appendage from patient donors***

RNA was extracted from fibroblasts derived from the right atrial appendage from patients with SR and cAF to measure miR31 expression by RT-qPCR. Kolmogorov–Smirnov (K-S) test was used to determine that raw Ct values were parametric and that fold change values were non-parametric. Means of normally distributed data were compared with unpaired Student's *t*-test while non-parametric data were compared using Mann–Whitney U test. SR patients n = 7, cAF patients n = 8. Error bars depict mean ± SD.

Upon inspection of the above RT-qPCR data, it was noted that the abundance of miR31 and U6 microRNAs showed marked variation within groups. I considered whether this was due to technical variability introduced when performing the RT-qPCR methodology, however subsequent figures showing RT-qPCR results for miR31 (*see figures 3.5 and 3.6 below*), which made use of cultured hAFs from a single primary donor, rather than multiple donors, showed tight error bars indicating excellent repeatability. This suggested that technical variability for the RT-qPCR methodology must be low in my hands and that the extent of variability observed in *figure 3.1* above was likely the result of other factors. For example, the natural biological variation seen between human subjects as well as variation caused by sex, age, and other confounding variables such as co-morbidities, smoking status and/or alcohol consumption, medication use and surgical handling of the retrieved tissues. These factors could not be corrected for in the experimental design due to the low sample number because of limited access to donor patient tissue. In due course however, the power of the experiment may be improved by increasing the sample size as the availability of tissue increases.

Despite the variation of miR31 in human donor patients as shown in *figure 3.1* above, we became interested in whether miR31 is involved more broadly in myocardial fibrogenesis or specifically in the aetiology or maintenance of AF-associated fibrosis. Consequently, in this part of the project experiments were conducted on primary human atrial fibroblasts (hAFs) cultured *in vitro* to query whether miR31 would facilitate pro-fibrotic gene expression changes and phenotype in this model. Due to the low availability of donor tissue all experiments were performed as technical replicates, not biological replicates, on cells that were expanded up from a single donor, donor 1 (CC-2903: 0000634284). Following depletion of cells from donor 1, some experiments were proceeded with cells from a second donor, donor 2 (CC-2903: 18TL312486), and again, these experiments were performed as technical, not biological replicates. These donors had no previous history of atrial fibrillation or myocardial fibrosis (*see appendices for cell donor details and certification*) and the donor used for each experiment is specified in the figure legends throughout this document.

In my experimental design, miR31 was targeted either for degradation using a reverse complementary hairpin RNA oligonucleotide or overexpressed in hAFs using a double-stranded oligonucleotide with identical sequence to the endogenous miR31. To gain an unbiased view of how these experimental interventions might affect underpinning gene expression changes facilitating fibrogenesis in hAFs, the poly-A tail transcriptome was sequenced by the genomic core facility at the Wellcome Trust Centre (Oxford). Sequence data was then validated by me with RT-qPCR. In tandem hAF fibrotic phenotype was queried by assessing cell proliferation, using a BrdU ELISA assay (Abcam: 126556), and collagen production, using Sirius Red Collagen detection kit (Chondrex: 9062). Given that these experiments were performed as technical replicates on hAFs from single donors, and not as biological replicates, it was not informative at this stage of the project to perform power calculations to gauge the number of biological replicates needed to reach statistical significance. In future, as more samples are acquired from new patient donors, power calculations will be incorporated into the study in order to improve the validity of these experiments.

3.1 Flow cytometry and immunofluorescence validate human atrial fibroblasts

Currently there are no intracellular or surface markers that can be used to discern fibroblast type however all *known* fibroblasts express Collagen type-I and therefore this has been routinely used to validate fibroblasts *in vitro* (Stellato et al. 2019). Indeed our *in vitro* model uses human atrial fibroblasts which are first characterised by the supplier Lonza, by positive immunofluorescent staining for the αI procollagen chain (encoded by Col1a1), and negative for von Willebrand factor VIII related antigen which is used to mark endothelial cells. In addition to this validation, we sought to perform further fibroblast authentication in-house by assessing the presence of intra-cellular markers Filamin A and Vimentin which have been shown previously to be expressed in cardiac fibroblasts (Moore-Morris et al. 2014, Moreira et al. 2020). Immunostaining results confirmed cytoplasmic abundance of both Filamin A and Vimentin in hAFs (*figure 3.2*).

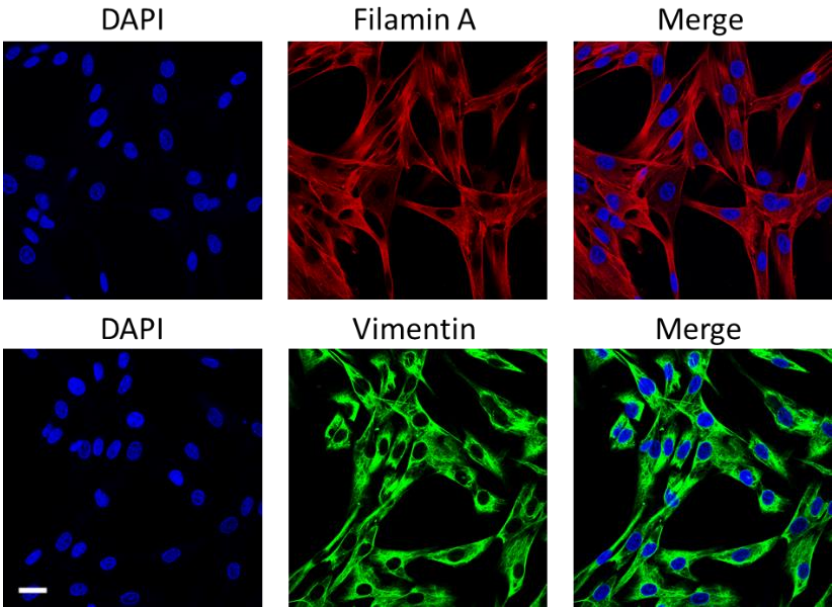


Figure 3.2 ***Immunostaining validates presence of intracellular proteins found in cardiac fibroblasts***

hAF cells on passage 7 were fixed with methanol: acetone (1:1) and incubated with To-Pro for nuclear staining (blue), anti-filamin A (red) and anti-vimentin (green) antibodies for cytoskeletal probing. Alexa Fluor (Invitrogen) secondary antibodies were used and imaging performed with a Leica DM 6000 confocal microscope. hAF donor 1 was used for this experiment. Scale bar is 50 µm.

Although Filamin A and Vimentin are abundant in cardiac fibroblasts, these are non-specific markers as Filamin A shows ubiquitous expression while Vimentin is expressed in most cells derived from the primitive mesenchyme (Mendez et al. 2010). Therefore for additional in-house confirmation of hAFs we queried the presence of cell membrane protein THY1 (i.e. CD90) which has been used previously to identify and sort cardiac fibroblasts *in vitro* (Ali et al. 2014). Cardiac fibroblasts also express low levels of PDGFRα which is significantly upregulated in disease state myofibroblasts (Chong et al. 2013, Ivey et al. 2016). Thus to demonstrate that these cells were normal and not showing myofibroblast-like properties, we also queried the expression of PDGFRα. In addition, we queried the presence of VAP1, which is ordinarily used as an endothelial cell surface marker, however one study did show increased VAP1 expression could be used to distinguish myofibroblasts from dermal fibroblasts (Hsia et al. 2016).

For this experiment cultured fibroblasts were harvested and incubated in suspension with anti-CD90, -PDGFRα and -VAP1 antibodies. As leukocytes also express CD90 and endothelial cells also express VAP1 and low amounts of PDGFRα, in addition to the above antibodies, fibroblasts in suspension were also incubated with antibodies against surface markers CD45, which marks leukocytes, and CD31, which marks endothelial cells. Fibroblasts do not express CD45 and CD31, therefore CD31-, CD45-, and CD90+ cells identified with flow cytometry would be confidently recognised as fibroblasts. In addition to the cell population incubated with the above antibodies, a negative control group of hAFs were utilised in this experiment that were not incubated with any antibodies.

For data processing, gating was first performed on scatter and fluorescence intensity plots derived from negative control hAF cells to exclude cell debris, cell doublets and dead cells while selecting for CD31-CD45-CD90+PDGFRα+VAP1+ cells (*figure 3.3*). *Table 3.1* provides the breakdown of total events enclosed by each gating strategy used on hAF negative control cells.

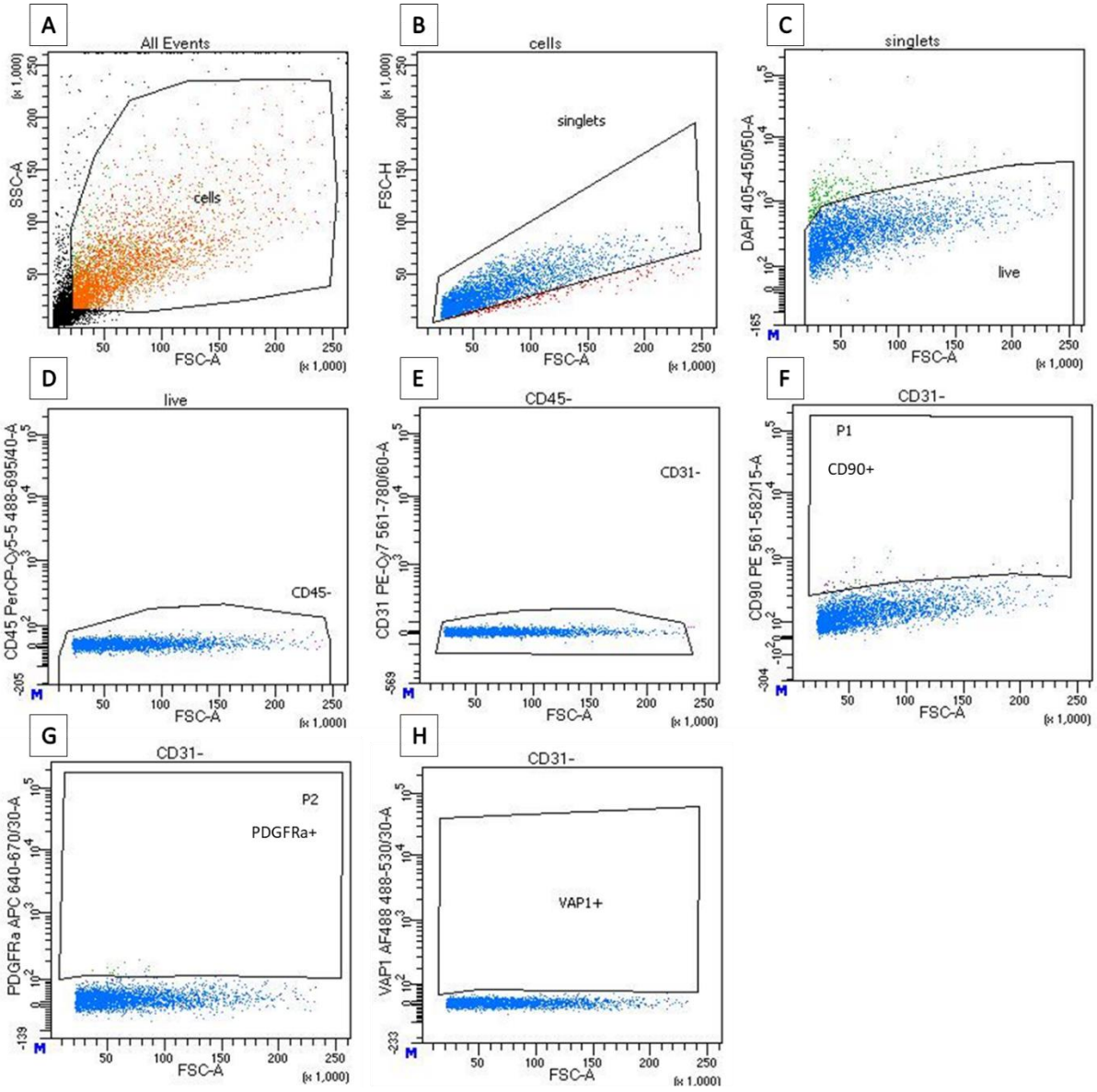


Figure 3.3 ***Flow cytometry gating strategy used on negative control hAF cell population (i.e. unstained cells)***

A: a total of one million hAFs on passage 7 from donor 1 were used for this experiment. Gating was first performed in side scatter area (SSC-A) against forward scatter area (FSC-A) plots to exclude debris.
B: gating was then performed in forward scatter height (FSC-H) against FSC-A plots to exclude cell doublets.
C: live cells were then selected as DAPI-negative in FSC-H against log DAPI plots
D-H: gating was then performed on live cells to select CD31-, CD45-, CD90+, PDGFRα+, and VAP1+ cells

Table 3.1 ***Flow cytometry event count following gating in negative control hAFs***

Population	#Events	%Parent	%Total
All Events	11,427	####	100.0
cells	3,969	34.7	34.7
singlets	3,686	92.9	32.3
live	3,437	93.2	30.1
CD45-	3,437	100.0	30.1
CD31-	3,432	99.9	30.0
P1	44	1.3	0.4
P2	15	0.4	0.1
NOT(P1)	3,388	98.7	29.6
VAP1+	0	0.0	0.0

Table 3.1 gives a breakdown of the number of events counted within each separate gating perimeter. The one million hAF cells used in this experiment were not incubated with antibodies and produced a total of 11427 events (first row). Following gating to exclude debris, 3969 cells were counted (second row). Of these cells, the exclusion of cell doublets gave 3686 single cells (third row) and the exclusion of dead cells, which are DAPI positive, gave 3437 live cells (fourth row). Gating was then performed to select cells negative for CD45 (fifth row) and also negative for CD31 (sixth row). Among the CD31 negative population, gating was performed to select cells positive for CD90 (seventh row), PDGFRα (eighth row) and VAP1 (tenth row). Row nine is extra information showing the number of cells that are negative for CD45 and CD31 as well as CD90.

Once gating dimensions were established in the negative control hAF cell population, the identical strategy was then used for hAF cells treated with surface marker antibodies (*figure 3.4*). *Table 3.2* provides a breakdown of total events enclosed by each gating strategy used on hAF cells and indicated that >95% of the cell population did not express leukocyte marker CD45 or endothelial cell marker CD31. Within this subset, 65.6% of the cells expressed fibroblast marker CD90, 36.8% expressed VAP1 and only 0.9% expressed PDGFRα.

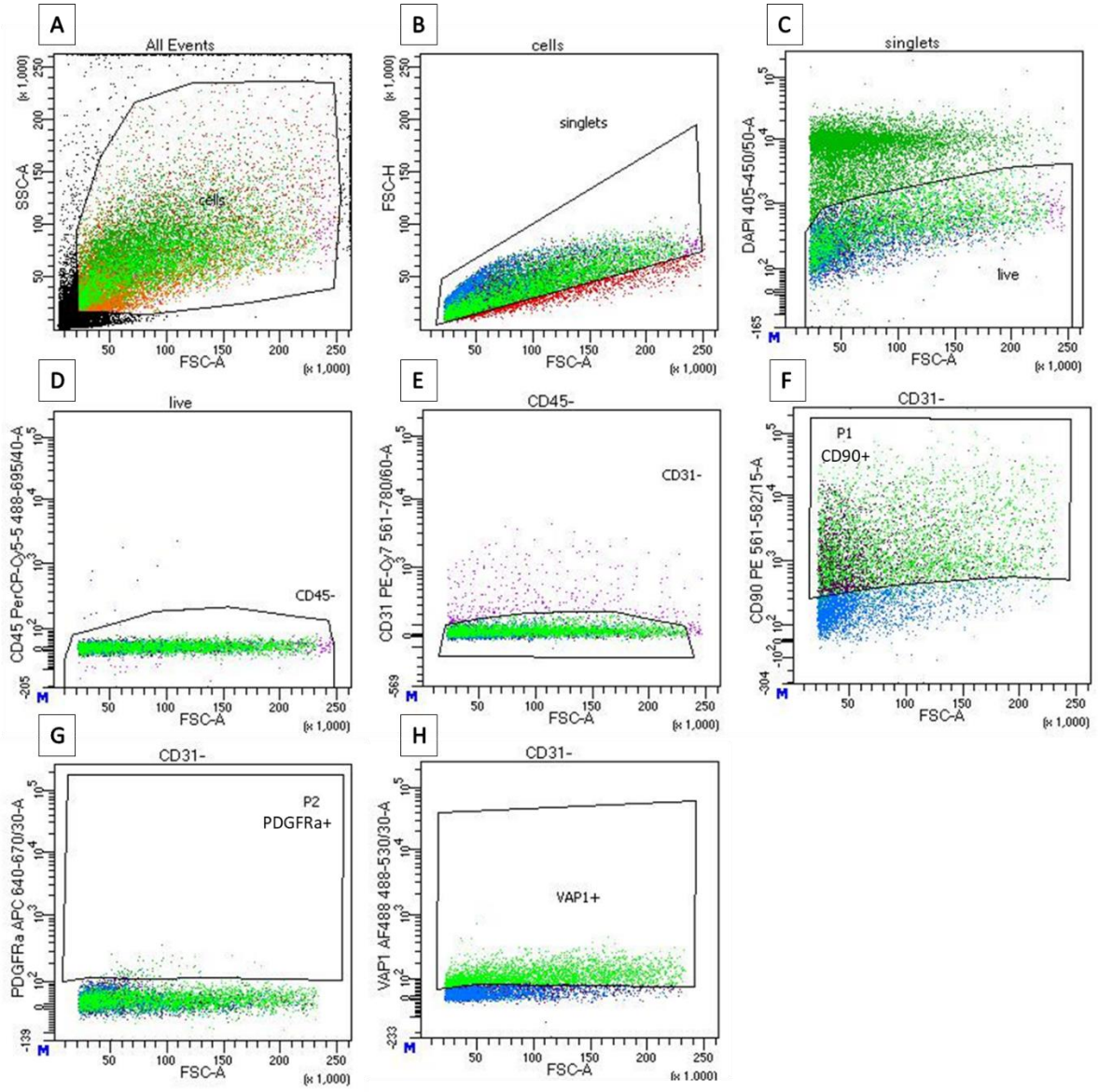


Figure 3.4 ***hAFs stained with anti-CD45, -CD31, -CD90, -PDGFRα and –VAP1 antibodies.***

Gating established on the hAF negative control cells was then applied to the data plots derived from hAFs incubated with antibodies CD45, CD31, CD90, PDGFRα and VAP1.
A: gating was performed first in side scatter area (SSC-A) against forward scatter area (FSC-A) plots to exclude debris.
B: gating was then performed in forward scatter height (FSC-H) against FSC-A plots to exclude cell doublets.
C: live cells were then selected as DAPI-negative in FSC-H against log DAPI plots
D-H: gating was then performed on live cells to select CD31-, CD45-, CD90+, PDGFRα+, and VAP1+ cells

Table 3.2 ***Flow cytometry event count following gating in hAFs probed for surface markers CD45, CD31, CD90, PDGFRα and VAP1***

Population	#Events	%Parent	%Total
All Events	52,575	####	100.0
cells	19,114	36.4	36.4
singlets	17,324	90.6	33.0
live	8,433	48.7	16.0
CD45-	8,422	99.9	16.0
CD31-	8,050	95.6	15.3
P1	5,279	65.6	10.0
P2	75	0.9	0.1
NOT(P1)	2,771	34.4	5.3
VAP1+	2,961	36.8	5.6

Table 3.2 gives a breakdown of the number of events counted within each separate gating perimeter. The one million hAF cells used in this experiment were treated with antibodies against surface markers CD45, CD31, CD90, PDGFRα and VAP1 and produced a total of 52575 events (first row). Following gating to exclude debris, 19114 cells were counted (second row). Of these cells, the exclusion of cell doublets gave 17324 single cells (third row) and the exclusion of dead cells, which are DAPI positive, gave 8433 live cells (fourth row). Gating was then performed to select cells negative for CD45 (fifth row) and also negative for CD31 (sixth row). Among the CD31 negative population, gating was performed to select cells positive for CD90 (seventh row), PDGFRα (eighth row) and VAP1 (tenth row). Row nine is extra information showing the number of cells that are negative for CD45 and CD31 as well as CD90.

These in-house fibroblast authentication experiments have indicated that hAFs purchased from Lonza express Vimentin and Filamin A and that the majority of these cells express fibroblasts surface marker CD90. A large minority also express VAP1, however cells expressing PDGFRα, which is upregulated in myofibroblasts, were negligible. Given that there are no fibroblast-specific markers, or markers to distinguish between fibroblast sup-type specifically, the results from these two in-house experiments, together with authentication from Lonza that showed abundance of collagen type 1, were deemed sufficient validation of hAF cells to commence experiments.

3.2 Validation of tools for manipulation of miR31 in hAFs

To reduce miR31 abundance in hAFs, miR31 was targeted for degradation using a reverse complementary hairpin RNA oligonucleotide, with a proprietary backbone modification for increased stability (Dharmacon IH-300507-06). For a negative control I used a reverse complementary hairpin RNA oligonucleotide targeting C. *elegans* miR67 (Dharmacon: IN-001005-01) as this microRNA has no known sequence homology to any endogenous mammalian microRNAs (Griffiths-Jones et al. 2006). Conversely, for ectopic overexpression of miR31 in primary hAFs, a double-stranded oligonucleotide with identical sequence to endogenous miR31 was used (Dharmacon: C-300507-05). Similarly, for a negative control, I used double-stranded oligonucleotide mimic of C. *elegans* miR67 (Dharmacon: CN-001000-01).

To validate these tools for manipulation of miR31 in hAFs, I conducted a pilot experiment whereby hAFs cultured in six well plates were treated with 75 pmol of miR31 mimic or inhibitor and their respective controls. MiR31 and small nuclear RNA U58A (RNU58A) cDNA was then made and RT-qPCR used to measure miR31 abundance. Raw Ct values indicated an increase in miR31 with miR31 mimic treatment and reduction with miR31 inhibitor treatment. RNU58A remained stable and could therefore be used as reference gene in RT-qPCR gene expression calculations. Following normalisation, miR31 showed a ~400-fold increase in the mimic group and a ~10-fold decrease in the inhibitor group, when compared to their controls (*figure 3.5*). These values indicated successful manipulation of miR31 in hAFs and therefore I proceeded with further experiments to investigate the consequences of miR31 perturbation on hAF cell gene expression changes and phenotype.

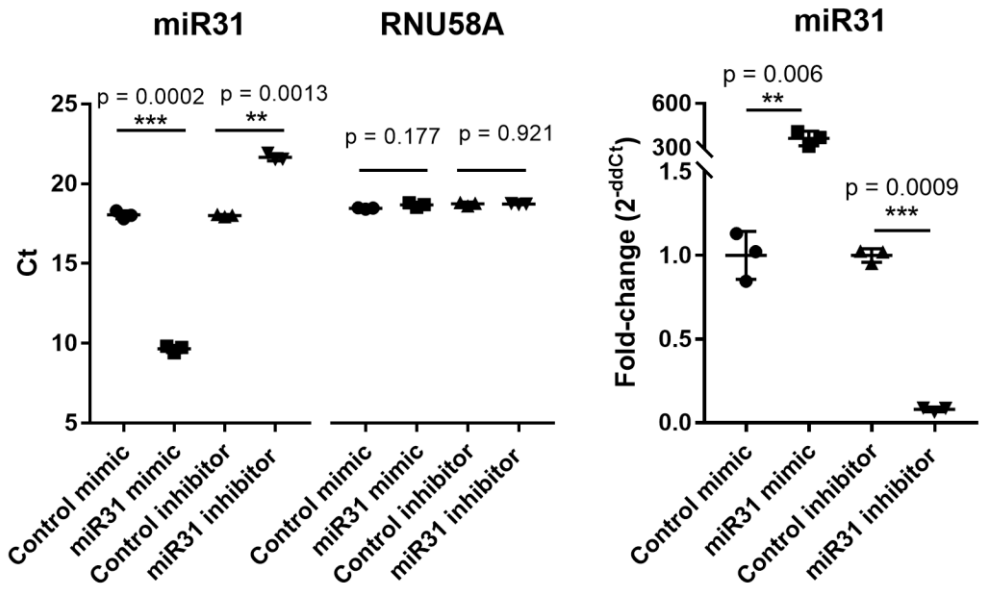


Figure 3.5 ***miR31 expression in hAFs following enrichment or depletion of miR31 with small oligonucleotides***

hAFs on passage 7 from donor 1 were lipofected in technical triplicates using six well plates using RNAiMax with 75 pmol of miR31 inhibitor (Dharmacon IH 300507 06) or control inhibitor which targets C. *elegans* miR67 (Dharmacon: IN 001005 01). Similarly, hAFs were treated with 75 pmol miR31 mimic (Dharmacon: C-300507-05) or mimic control identical to C. *elegans* miR67 (Dharmacon: CN-001000-01). Forty-eight hours following lipofection RNA was harvested for RT-qPCR, RNU58A was used as a housekeeping gene for normalization. Paired Student's *t*-test was used for statistical analysis as data was shown to be parametric with Kolmogorov–Smirnov test. Results are expressed as fold change of control (either mimic control or inhibitor control). Data is presented as mean ±SD.

3.3 Gain- and loss-of miR31 significantly alters the poly-A tail transcriptome of human atrial fibroblasts

Given our interest in the role of miR31 in human atrial fibroblasts and whether miR31 abundance affects change in the fibrotic phenotype of these cells, I first sought to take an unbiased view on how miR31 alters the transcriptome in these cells. Single microRNAs can bind hundreds of targets and through this action can drive significant change in cell behaviour and phenotype. Gaining a better understanding of which gene transcripts are significantly altered following enrichment or depletion of miR31 would reveal cell pathways driving changes in hAF cell phenotypes caused by miR31 dysregulation. Rather than looking at different predicted miR31 targets individually, I chose to sequence the entire poly-A tail transcriptome to gain an unbiased view of changes in the RNA portion ordinarily translated into protein.

For the RNA-sequencing experiment, miR31 was enriched or depleted in cultured hAFs on passage 7 from donor 1 by lipofection with miR31 mimic and inhibitor oligonucleotides. To acquire RNA of sufficient amount and purity, as well as an adequate number of samples for sequencing, this experiment was performed in technical triplicates for each treatment group and then repeated four times. Total RNA was then purified and absorbance at 230, 260 and 280 nM was measured to estimate concentration and purity. Our RNA extraction protocol uses cell lysis buffer containing guanidine thiocyanate which absorbs strongly at 230 nm, followed by organic extraction using phenol-chloroform which absorbs strongly at 280 nm. RNA absorbs strongly at 260 nm thus low 260/230 and 260/280 ratios were used to reveal contamination from guanidine thiocyanate and phenol-chloroform respectively. Poly-A tail library preparation for RNA sequencing requires 260/280 and 260/230 ratios of ~2 and therefore replicates from experimental repeats used for sequencing were selected on this basis (*Table 3.3* highlighted samples).

Table 3.3 ***RNA Nanodrop absorbance readings from all technical replicates in miR31 enrichment and depletion experimental repeats***

miR31 mimic and control treated hAF RNA samples					
Sample name	ng/ul	A260	A280	260/280	260/230
Expt1_Rep1 control_mimic	63.45	1.586	0.761	2.09	1.92
Expt1_Rep2 control_mimic	62.26	1.556	0.747	2.08	2
Expt1_Rep3 control_mimic	63.84	1.596	0.769	2.07	1.89
Expt1_Rep1 miR31_mimic	49.75	1.244	0.605	2.05	1.98
Expt1_Rep2 miR31_mimic	44.21	1.105	0.554	1.99	1.73
Expt1_Rep3 miR31_mimic	66.76	1.669	0.813	2.05	1.08
Expt2_Rep1 control_mimic	123	3.075	1.477	2.08	2.04
Expt2_Rep2 control_mimic	85.75	2.144	1.043	2.06	1.98
Expt2_Rep3 control_mimic	91.88	2.297	1.105	2.08	1.92
Expt2_Rep1 miR31_mimic	104	2.599	1.243	2.09	1.49
Expt2_Rep2 miR31_mimic	86.34	2.158	1.022	2.11	1.96
Expt2_Rep3 miR31_mimic	54.78	1.37	0.655	2.09	1.54
Expt3_Rep1 control_mimic	132.4	3.31	1.595	2.08	1.72
Expt3_Rep2 control_mimic	123.9	3.097	1.507	2.06	2.03
Expt3_Rep3 control_mimic	104.1	2.601	1.269	2.05	1.94
Expt3_Rep1 miR31_mimic	109.6	2.741	1.33	2.06	1.55
Expt3_Rep2 miR31_mimic	107.2	2.679	1.281	2.09	1.96
Expt3_Rep3 miR31_mimic	119.8	2.995	1.441	2.08	2.01
Expt4_Rep1 control_mimic	74.7	1.867	0.905	2.06	1.94
Expt4_Rep2 control_mimic	86.18	2.154	1.033	2.08	1.64
Expt4_Rep3 control_mimic	80.43	2.011	0.974	2.06	1.55
Expt4_Rep1 miR31_mimic	64.74	1.619	0.77	2.1	1.89
Expt4_Rep2 miR31_mimic	73.11	1.828	0.888	2.06	1.87
Expt4_Rep3 miR31_mimic	84.4	2.11	1.005	2.1	2.01

miR31 inhibitor and control treated hAF RNA samples					
Sample name	ng/ul	A260	A280	260/280	260/230
Expt1_Rep1 control_inhibitor	16.77	0.42	0.19	2.24	0.22
Expt1_Rep2 control_inhibitor	33.18	0.83	0.43	1.95	1.76
Expt1_Rep3 control_inhibitor	28.1	0.7	0.34	2.05	1.54
Expt1_Rep1 miR31_inhibitor	52.95	1.32	0.66	2.02	1.85
Expt1_Rep2 miR31_inhibitor	44.27	1.11	0.55	2.02	1.94
Expt1_Rep3 miR31_inhibitor	43.58	1.09	0.53	2.07	1.83
Expt2_Rep1 control_inhibitor	100.4	2.51	1.18	2.12	1.22
Expt2_Rep2 control_inhibitor	80.58	2.02	0.96	2.1	1.88
Expt2_Rep3 control_inhibitor	87.73	2.19	1.05	2.1	1.9
Expt2_Rep1 miR31_inhibitor	96.28	2.41	1.13	2.14	1.64
Expt2_Rep2 miR31_inhibitor	83.05	2.08	0.97	2.14	1.17
Expt2_Rep3 miR31_inhibitor	83.52	2.09	1	2.09	1.91
Expt3_Rep1 control_inhibitor	54.69	1.37	0.66	2.06	1.9
Expt3_Rep2 control_inhibitor	58.75	1.47	0.74	1.98	1.48
Expt3_Rep3 control_inhibitor	47.52	1.19	0.58	2.06	1.81
Expt3_Rep1 miR31_inhibitor	39.6	0.99	0.49	2.02	1.52
Expt3_Rep2 miR31_inhibitor	44.2	1.11	0.55	2	1.85
Expt3_Rep3 miR31_inhibitor	47.58	1.19	0.58	2.06	1.96
Expt4_Rep1 control_inhibitor	96.88	2.42	1.17	2.07	1.95
Expt4_Rep2 control_inhibitor	100.2	2.5	1.2	2.09	1.89
Expt4_Rep3 control_inhibitor	108.9	2.72	1.3	2.1	1.9
Expt4_Rep1 miR31_inhibitor	85.81	2.15	1.03	2.08	1.68
Expt4_Rep2 miR31_inhibitor	101.3	2.53	1.21	2.09	1.86
Expt4_Rep3 miR31_inhibitor	81.96	2.05	0.98	2.09	1.96

Having selected the technical replicates from my four experimental repeats to be used for transcriptomic analysis, I first used RT-qPCR to confirm miR31 enrichment or depletion in these samples. Like the initial pilot experiment (*figure 3.5 above*) raw Ct values indicated an increase in miR31 abundance with miR31 mimic and a reduction with miR31 inhibitor in all experimental repeats. RNU58A remained stable and could be used as housekeeping gene and following normalisation miR31 showed a ~400-fold increase above control in the miR31 mimic group, with a ~10-fold reduction in the miR31 inhibitor group (*figure 3.6*).

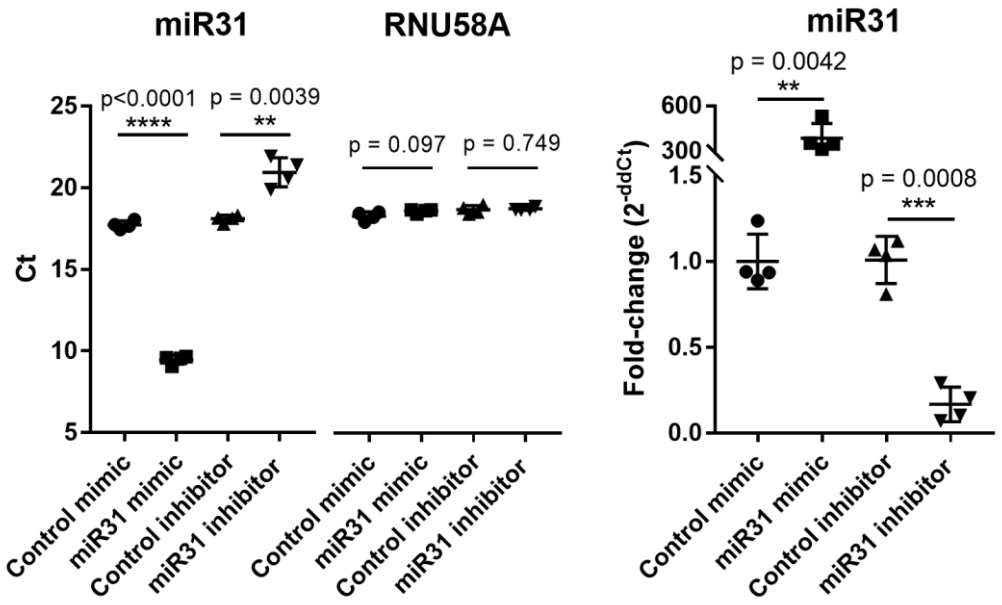


Figure 3.6 ***RT-qPCR measurements of miR31 expression in hAFs following enrichment or depletion of miR31 with small oligonucleotides from four experimental repeats***

Donor 1 hAFs on passage 7 were cultured in six-well plates and lipofected using RNAiMax in technical triplicates with 75 pmol of miR31 or inhibitor oligonucleotides or respective controls. Forty-eight hours following lipofection RNA was harvested. The experiment was repeated four times and replicates of sufficient purity were selected to check miR31 abundance before sequencing. RNU58A used as housekeeping for normalization in RT-qPCR and results expressed as fold-change of control. Paired Student's *t*-test was used for statistical analysis as data was shown to be parametric with Kolmogorov–Smirnov test. N = 4 in each group.

Following confirmation of depletion or enrichment of miR31 in hAF RNA, samples were sequenced as a service at Oxford's Wellcome Trust Centre genomic core facility. Sequencing data was deemed of sufficient quality for analysis by their in-house quality control measures (*see appendices for sequencing quality control*). Sequencing read counts were then analysed using edgeR statistical package (Bioconductor release 3.12) which compared miR31 mimic treated hAFs to mimic control (four in each group) as well as compared miR31 inhibitor to inhibitor control (also four in each group). This generated a comprehensive list of gene sequencing read count, p-value, adjusted p-value (p-value corrected for multiple statistical tests), likelihood ratio (LR) and fold change relative to control (*gene lists can be supplied upon request*).

Analysis of the miR31 mimic vs. control sequencing read count showed that 4824 genes were significantly altered (adjusted p value < 0.05). Conversely the miR31 inhibitor vs. control analysis showed 1144 genes significantly changed (adjusted p value < 0.05). The large difference in these numbers (4824 vs 1144) could be explained by the fact that miR31 mimic caused supra-physiological levels of miR31 (~400-fold greater than control), whereas miR31 inhibitor reduced levels of endogenous miR31 (~10-fold less than control). Therefore off-target secondary effects would be more likely in the miR31 mimic dataset and this could lead to greater transcript abundance changes.

Given that the function of microRNAs is to target mRNA transcripts for degradation, the expected result of increasing miR31 would be a net reduction in primary mRNA targets. Conversely, a reduction in miR31 would presumably cause a net increase in mRNA primary targets. To determine whether miR31 was behaving congruently with these expectations in my experimental system, the sequencing data was stratified to query the breakdown of the up/down regulated genes from the full list of significantly changed genes (i.e. all genes with adj p < 0.05) as well as the top 500, 200, 100, 50 and 25 significantly changed genes. In the miR31 mimic dataset there was an even split of up/down regulated genes when the whole gene list was considered and a consistent trend in the top 25, 50, 100, 200 and 500 genes that they decreased, not increased, when compared to control (*Table 3.4 upper half*). This suggested that miR31 mimic was having a significant off-target effect but also indicated that direct effects could be elucidated by applying more stringent p-value cut-offs to the sequencing dataset. Again, this may be explained when considering that miR31 mimic resulted in very high levels of miR31 and therefore surplus miR31 was able to elicit large secondary effect. Indeed to correct for this large secondary effect, I applied a stricter adjusted p-value when analysing the gene list using Ingenuity pathways analysis (IPA) (Qiagen version 42012434) (*next section*).

In the case of the miR31 inhibitor sequencing data, again there was a relatively stable split of up/down regulated genes in the total gene list (i.e. all genes with adj p < 0.05). However in the top 25, 50, 100, 200 and 500 genes that significantly changed, there was a bias towards transcripts *decreasing*, rather than increasing in abundance (*Table 3.4 lower half*). This was unexpected but could suggest that the primary effects of miR31 depletion were having immediate and significant knock-on effects. For

example, DICER1 mRNA abundance was shown to increase by ~100% (adjusted $p < 0.0001$) in the miR31 inhibitor dataset as well as decrease by ~25% (adjusted $p = 0.0303$) in the miR31 mimic dataset. DICER1 is also a bioinformatically predicted direct miR31 target (Targetscan7.2). Given that the role of DICER1 is to process pre-microRNAs to generate microRNAs which then proceed to target a vast array of mRNA transcripts in the RISC complex, it would be expected that changes in DICER1 transcript abundance would have significant knock-on effects and that this would affect the sequencing data.

Table 3.4 Stratification of sequencing data

miR31 mimic vs. control mimic sequencing data			
	Number of Genes	**Up-regulated (%)**	**Down-regulated (%)**
All altered genes	4824	2464 (51%)	2360 (49%)
Top 500	500	194 (39%)	306 (61%)
Top 200	200	63 (31%)	137 (69%)
Top 100	100	28 (28%)	72 (72%)
Top 50	50	13 (26%)	37 (74%)
Top 25	25	7 (28%)	18 (72%)
miR31 inhibitor vs. control inhibitor sequencing data			
	Number of Genes	**Up-regulated (%)**	**Down-regulated (%)**
All altered genes	1144	515 (45%)	629 (55%)
Top 500	500	202 (40%)	298 (60%)
Top 200	200	85 (43%)	115 (57%)
Top 100	100	44 (44%)	56 (56%)
Top 50	50	19 (38%)	31 (62%)
Top 25	25	10 (40%)	15 (60%)

3.4 Gain- and loss-of miR31 significantly alters mRNA abundance of pro-fibrotic genes

As an initial approach to query the effects of miR31 abundance on pro-fibrotic genes and pathways, I manually screened the significantly changed genes (adj. $p < 0.05$) in the mimic and inhibitor RNA-seq data sets using the ctrl-F feature in excel. I looked for names, and/or ENSEMBL IDs, of genes that were of interest given their function or use as markers in fibroblasts and myofibroblasts. Results of this initial screen are presented in *Table 3.5*.

Table 3.5 Manual screen for pro-fibrotic genes of interest

Gene	miR31 mimic hAF RNA-seq	miR31 inhib hAF RNA-seq	Gene	miR31 mimic hAF RNA-seq	miR31 inhib hAF RNA-seq
TGFβ1	—	—	Thy1	—	—
TGFβ2	↑	—	VAP1	—	—
TβRI	↑	—	αSMA	—	—
TβRII	—	—	Vimentin	—	—
AngII	—	—	FLNA	—	—
AT1R	—	↓	TCF21	—	—
PDGF subunit A	↓		DDR2	—	—
PDGF subunit B	—	—	TIE2	—	↑
PDGFRα	—	—	FSP1	—	—
CTGF	—	—	SCA1	↑*	—
COL1A1	—	—	WT1	—	—
COL1A2	—	—	TBX18	↓	—
COL3A1	↑	↑	SMAD 2	—	—
MMP2	—	—	SMAD 3	—	—
MMP9	—	—	SMAD 4	—	—
TIMP4	—	—	SMAD 7	↑	
LOX	—	—	ETS1	—	—
Fibronectin	—	—	IFNE	↑	—
Fibrillin 1	—	—	CALCR	—	—
Periostin	—	—	MKI67	↓	—

Continued.

Gene	miR31 mimic hAF RNA-seq	miR31 inhib hAF RNA-seq	Gene	miR31 mimic hAF RNA-seq	miR31 inhib hAF RNA-seq
COL4A1	—	↑	AGTRAP	↑	—
COL4A5	↑	—	PDGFC	↓	—
COL5A1	—	↑	PDGFRL	↓	—
COL5A3	—	↑	FGF1	↓	—
COL8A1	↑	—	FGF5	↓	↓
COL8A2	↓	—	FGF22	↓	—
COL11A1	↓	—	FGFR1	↓	—
COL14A1	↑	—	FGFR3	↓	
LOXL1	↓	—	THRα	—	↑
LOXL4	↓	—	FBN2	↓	—
MMP24	↑	—			
MMP25	↑	—			

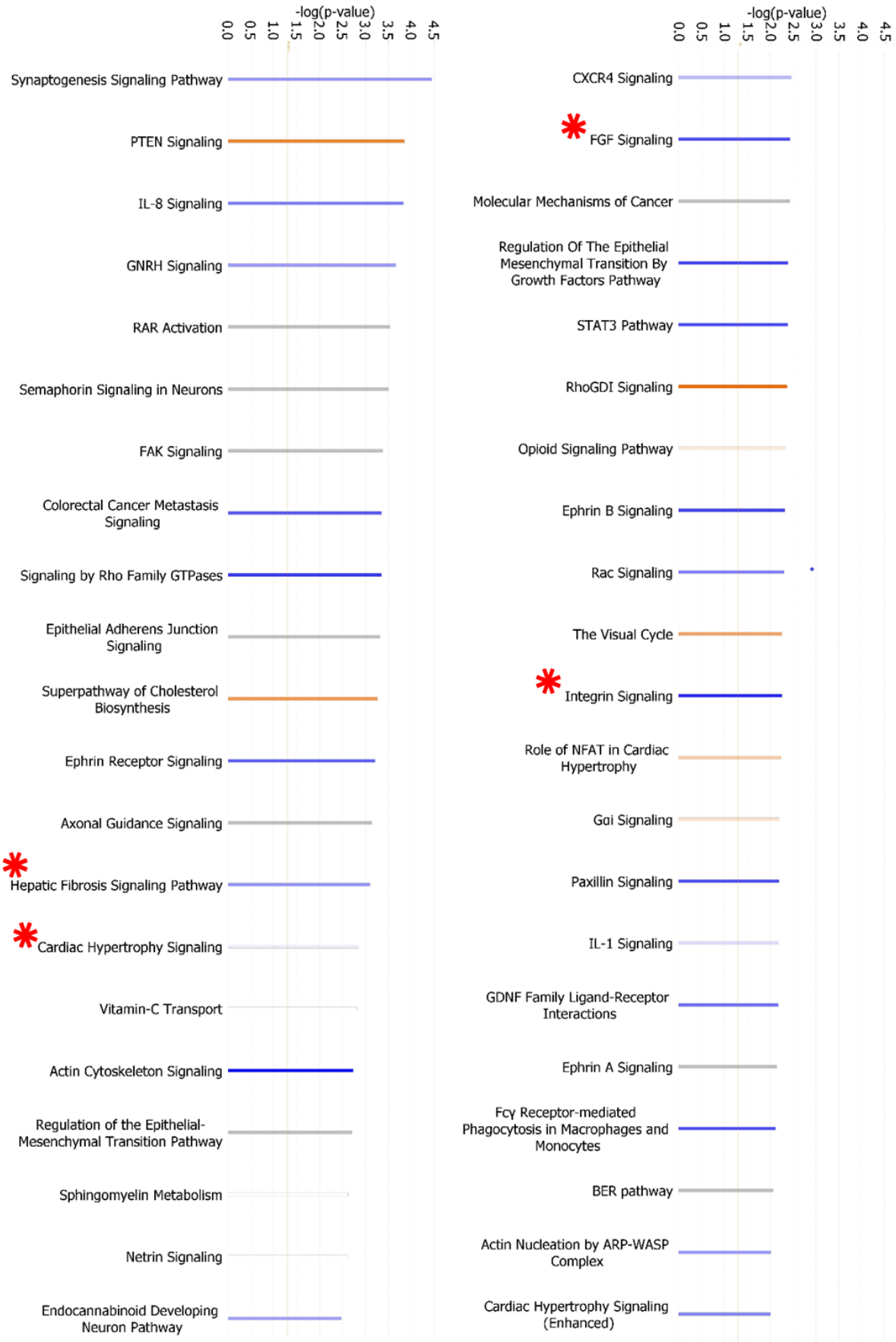
Significant increase (↑), significant decrease (↓), no change (—)

* SCA1 it is found as Ataxin 1 (ATXN1) in the raw gene list

Lysyl oxidase like 1 (LOXL1), thyroid hormone receptor alpha, (THRα) angiotensin II receptor associated protein (AGTRAP), Fibrillin2 (FBN2), platelet derived growth factor receptor like (PDGFRL)

To then gain a broader view of the transcriptomic changes, I analysed the gene lists using Ingenuity pathways analysis (IPA) (Qiagen version 42012434). IPA assesses the cellular signalling networks likely impacted with given gene expression changes. IPA is optimised for lists of 300-1200 genes, therefore for the miR31 mimic vs control data set I applied a stringent cut-off of adjusted p-value (adjusted $p < 0.00000001$), and a fold-change cut-off of ±3x, which gave a list of 861 genes (496 upregulated and 365 down regulated). IPA looks through the submitted list for genes represented in its own pre-defined pathway gene list and this comparison generates a p-value for how significantly enriched each pathway category is. The software also considers the fold-change as well as direction of change of each gene from the gene list and this information is used to class pathways as activated or repressed according to the known function of genes within a given pathway.

IPA analysis of the miR31 mimic dataset indicated 116 pathways significantly altered ($p < 0.05$) (*figure 3.7*). Of the 116 pathways I recognised seven pathways of potential importance to cardiac fibrosis including FGF, TGFβ and ERK/MAPK signalling, cardiac hypertrophy, hepatic fibrosis, renin-angiotensin signalling and integrin signalling. It is important to note that I was initially blinded to the direction of change of each of these pathways which meant that my attention was drawn to them purely based on their apparent association with AF and not because I could use their direction of change to support my hypotheses. This said, after the direction of change was decoded, it revealed that all of the seven aforementioned pathways showed *increased* enrichment following increased miR31 abundance. This supported our hypothesis that elevated miR31 drives pro-fibrotic change in hAFs.


-log(p-value)
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
Synaptogenesis Signaling Pathway
PTEN Signaling
IL-8 Signaling
GNRH Signaling
RAR Activation
Semaphorin Signaling in Neurons
FAK Signaling
Colorectal Cancer Metastasis Signaling
Signaling by Rho Family GTPases
Epithelial Adherens Junction Signaling
Superpathway of Cholesterol Biosynthesis
Ephrin Receptor Signaling
Axonal Guidance Signaling
Hepatic Fibrosis Signaling Pathway
Cardiac Hypertrophy Signaling
Vitamin-C Transport
Actin Cytoskeleton Signaling
Regulation of the Epithelial-Mesenchymal Transition Pathway
Sphingomyelin Metabolism
Netrin Signaling
Endocannabinoid Developing Neuron Pathway
-log(p-value)
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
CXCR4 Signaling
FGF Signaling
Molecular Mechanisms of Cancer
Regulation Of The Epithelial Mesenchymal Transition By Growth Factors Pathway
STAT3 Pathway
RhoGDI Signaling
Opioid Signaling Pathway
Ephrin B Signaling
Rac Signaling
The Visual Cycle
Integrin Signaling
Role of NFAT in Cardiac Hypertrophy
Gαi Signaling
Paxillin Signaling
IL-1 Signaling
GDNF Family Ligand-Receptor Interactions
Ephrin A Signaling
Fcγ Receptor-mediated Phagocytosis in Macrophages and Monocytes
BER pathway
Actin Nucleation by ARP-WASP Complex
Cardiac Hypertrophy Signaling (Enhanced)


-log(p-value)
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
fMLP Signaling in Neutrophils
Germ Cell-Sertoli Cell Junction Signaling
Cholesterol Biosynthesis I
Cholesterol Biosynthesis II (via 24,25-dihydrolanosterol)
Cholesterol Biosynthesis III (via Desmosterol)
Tec Kinase Signaling
Wnt/β-catenin Signaling
Renin-Angiotensin Signaling
Regulation Of The Epithelial Mesenchymal Transition In Development Pathway
Clathrin-mediated Endocytosis Signaling
Role of Macrophages, Fibroblasts and Endothelial Cells in Rheumatoid Arthritis
SAPK/JNK Signaling
HGF Signaling
Gαq Signaling
Necroptosis Signaling Pathway
Mevalonate Pathway I
Regulation of Actin-based Motility by Rho
Sumoylation Pathway
Semaphorin Neuronal Repulsive Signaling Pathway
IL-15 Production
Agrin Interactions at Neuromuscular Junction
-log(p-value)
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
Estrogen-mediated S-phase Entry
Reelin Signaling in Neurons
Role of Osteoblasts, Osteoclasts and Chondrocytes in Rheumatoid Arthritis
Osteoarthritis Pathway
Gαs Signaling
Selenocysteine Biosynthesis II (Archaea and Eukaryotes)
Glioblastoma Multiforme Signaling
Regulation of Cellular Mechanics by Calpain Protease
Pancreatic Adenocarcinoma Signaling
Endocannabinoid Neuronal Synapse Pathway
White Adipose Tissue Browning Pathway
Dermatan Sulfate Biosynthesis
Thrombin Signaling
Retinol Biosynthesis
α-Adrenergic Signaling
Mitochondrial Dysfunction
TGF-β Signaling
G-Protein Coupled Receptor Signaling
Adipogenesis pathway
Superpathway of Geranylgeranyldiphosphate Biosynthesis I (via Mevalonate)
Role of Tissue Factor in Cancer


-log(p-value)
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
Sertoli Cell-Sertoli Cell Junction Signaling
PKCθ Signaling in T Lymphocytes
Cdc42 Signaling
PEDF Signaling
Apelin Adipocyte Signaling Pathway
Cholecystokinin/Gastrin-mediated Signaling
Ovarian Cancer Signaling
CREB Signaling in Neurons
Virus Entry via Endocytic Pathways
Estrogen Receptor Signaling
Retinoate Biosynthesis I
G Beta Gamma Signaling
Mouse Embryonic Stem Cell Pluripotency
Myc Mediated Apoptosis Signaling
ERK/MAPK Signaling
Amyloid Processing
Sperm Motility
Phospholipase C Signaling
Glutamine Biosynthesis I
Adenine and Adenosine Salvage VI
Aryl Hydrocarbon Receptor Signaling
-log(p-value)
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
Endocannabinoid Cancer Inhibition Pathway
UVC-Induced MAPK Signaling
Sirtuin Signaling Pathway
Phosphatidylethanolamine Biosynthesis II
PAK Signaling
B Cell Receptor Signaling
cAMP-mediated signaling
Breast Cancer Regulation by Stathmin1
Notch Signaling
Small Cell Lung Cancer Signaling
Regulation of IL-2 Expression in Activated and Anergic T Lymphocytes

Figure 3.7 ***116 significantly altered cell signalling networks determined with Ingenuity Pathway Analysis in miR31 mimic sequencing dataset***

Adjusted p-value cut-off of < 0.00000001, and a fold-change cut-off of ±3x, was applied to the total list of significantly changed genes in the miR31 mimic dataset. This gave a list of 861 genes that was then analysed using IPA which generated a list of 116 significantly altered pathways ($p < 0.05$). Blue intensity indicate extent to which pathways were activated, red-orange intensity indicate extent of pathway repression. Grey indicate pathways without strong directional change. Red stars indicate pathways of potential importance to cardiac fibrosis.

For analysis of the miR31 inhibitor vs control dataset, I applied an adjusted p-value cut-off of < 0.05, and a fold-change cut-off of ±3x, which gave a list of 1144 genes (515 upregulated and 629 downregulated). IPA results indicated 52 pathways significantly altered (*figure 3.8*) of which five pathways appeared important including BMP, PDGF and TGF-β signalling, hepatic fibrosis and integrin signalling. Following decoding of their directional change, all these pathways showed *reduced* activity with miR31 inhibition, which again supported our hypothesis.

Following this observation I generated a list of the composite genes that had been identified within each pathway of interest (*Table 3.6*). This gene list would serve as a useful reference point for future studies wanting to validate the role of these pathways in functional assays and proteomic investigations studying fibrogenesis in hAFs.


-log(p-value)
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
Axonal Guidance Signaling
Adipogenesis pathway
Semaphorin Neuronal Repulsive Signaling Pathway
Hepatic Fibrosis / Hepatic Stellate Cell Activation
Semaphorin Signaling in Neurons
Chondroitin Sulfate Degradation (Metazoa)
IL-8 Signaling
Dermatan Sulfate Degradation (Metazoa)
Hepatic Fibrosis Signaling Pathway
ILK Signaling
Clathrin-mediated Endocytosis Signaling
PDGF Signaling
Pentose Phosphate Pathway (Oxidative Branch)
Iron homeostasis signaling pathway
Oleate Biosynthesis II (Animals)
Ferroptosis Signaling Pathway
ErbB2-ErbB3 Signaling
Synaptogenesis Signaling Pathway
Osteoarthritis Pathway
IL-4 Signaling
BMP signaling pathway
-log(p-value)
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
Th2 Pathway
Actin Cytoskeleton Signaling
ERK5 Signaling
Insulin Receptor Signaling
RAN Signaling
γ-linolenate Biosynthesis II (Animals)
Acute Myeloid Leukemia Signaling
Glioma Invasiveness Signaling
Oncostatin M Signaling
Tumor Microenvironment Pathway
Integrin Signaling
IL-15 Signaling
Angiopoietin Signaling
Macropinocytosis Signaling
Superoxide Radicals Degradation
IL-7 Signaling Pathway
TGF-β Signaling
Circadian Rhythm Signaling
IL-9 Signaling
NGF Signaling
JAK/Stat Signaling

Figure 3.8 ***Significantly altered cell signalling networks determined with Ingenuity Pathway Analysis in miR31 inhibitor sequencing dataset***

Adjusted p-value cut-off of <0.05, and a fold-change cut-off of ±3x, was applied to the total list of significantly changed genes in the miR31 inhibitor dataset. This gave a list of 1144 genes that was then analysed using Ingenuity pathway analysis which generated a list of 42 significantly altered pathways ($p < 0.05$). Blue intensity indicate extent to which pathways were activated, red-orange intensity indicate extent of pathway repression. Grey indicate pathways without strong directional change. Red stars indicate pathways of potential importance to cardiac fibrosis.

Table 3.6 ***Ingenuity pathways with relevance to cardiovascular fibrosis with significantly altered genes***

miR31 mimic vs. control mimic sequencing data	
Pathway	**Genes**
FGF signalling	FGF1, FGF5, FGFR1, FGFR3, ITPR1, MAPKAPK2, MET, RAC1, SOS2,
Cardiac hypertrophy	ADCY3, ADCY4, ADRA1B, ADRB2, CDC42, CHP1, GNB1L, GNG2, IRS1, MAP3K12, MAPK9, MAPKAPK2, NFATC4, OPN1SW, PRKAR2A, RAC1, RAC2, RRAS2, SRF
Hepatic fibrosis	ACVR2B, ARAF, AXIN1, BCL2, CDC42, DVL3, FGFR1, FOXO4, FZD1, IL1RAP, IRAK1, ITGB3, APK9, NOX4, OPN1SW, PRKAR2A, RAC1, RAC2, RRAS2, SNAI1, SOS2, TCF3, TCF7L1, TLR4, TNFRSF1B, VEGFB
Integrin signalling	ASAP1, CAPN5, CAPNS1, CDC42, DOCK1, ILK, ITGB3, PAK4, RAC1, RAC2, RRAS2, SOS2, SRC, TLN2, TSPAN6, WIPF1
TGFβ signalling	ACVR2B, BCL2, BMPR1A, CDC42, MAPK9, RRAS2, SKI, SOS2
Renin-angiotensin signalling	ADCY3, ADCY4, ITPR1, MAPK9, PAK4, PRKAR2A, PTGER2, RAC1, RRAS2, SOS2
ERK/MAPK signalling	ARAF, DOCK1, PAK4, PRKAR2A, RAC1, RAC2, RRAS2, SOS2, SRC, SRF, TLN2, YWHAH
miR31 inhibitor vs. control inhibitor sequencing data	
Pathway	**Genes**
BMP signalling	BMP1, BMP6, BMPR1B, FST, PRKAR2A, RAP2B, RRAS, SMAD6
Integrin signalling	CAPN5, GSN, ITGA11, ITGB3, MYL9, MYLK2, NCK2, PFN2, PIK3C2B, RAP2B, RHOD, RHOG, RND3, RRAS, TLN2
PDGF signalling	ABL2, EIF2AK2, INPP5K, JAK3, PDGFC, PIK3C2B, RAP2B, RRAS, TYK2
TGFβ signalling	ACVR2A, ACVR2B, BMPR1B, IRF7, RAP2B, RRAS, SMAD6, SMURF2
Hepatic fibrosis	ACVR2A, ACVR2B, AGTR1, CACNA1A, COL3A1, COL5A3, CREB5, FOXO4, FZD5, ITGB3, LEPR, MYL9, MYLK2, NOX4, PDCD4, PDGFC, PIK3C2B, RKAR2A, RAP2B, RHOD, RHOG, RND3, RPS6KB2, RRAS, TLR4, CAM1, VEGFB

3.4.1 Validation of RNA-seq results using RT-qPCR

Having shown changes in a wide a range of signalling pathways following manipulation of miR31, I then sought to validate my RNA-seq data by showing congruent gene expression changes in candidate genes of interest in the same RNA samples using RT-qPCR. To focus validation on a select few genes, I chose to look at genes that were significantly changed, but in opposing directions, in both miR31 mimic and inhibitor experiments. Such genes were selected as by definition they showed the most reproducible behaviour with manipulation of miR31 in hAFs. Genes up-regulated in the miR31 inhibitor experiment and down-regulated in the miR31 mimic experiment were of particular interest as this list would presumably contain many direct miR31 targets. Conversely, genes down-regulated in the miR31 inhibitor experiment and up-regulated in the miR31 mimic experiment would likely contain genes that are reproducibly sensitive to the *secondary* effects of miR31 manipulation. When looking for genes that changed in opposing directions in both miR31 mimic and inhibitor datasets, I found 161 genes that simultaneously decreased with miR31 mimic, and increased with miR31 inhibitor (*figure 3.9A*). In addition, 157 genes increased and reduced following miR31 enrichment and depletion respectively (*figure 3.9B*).

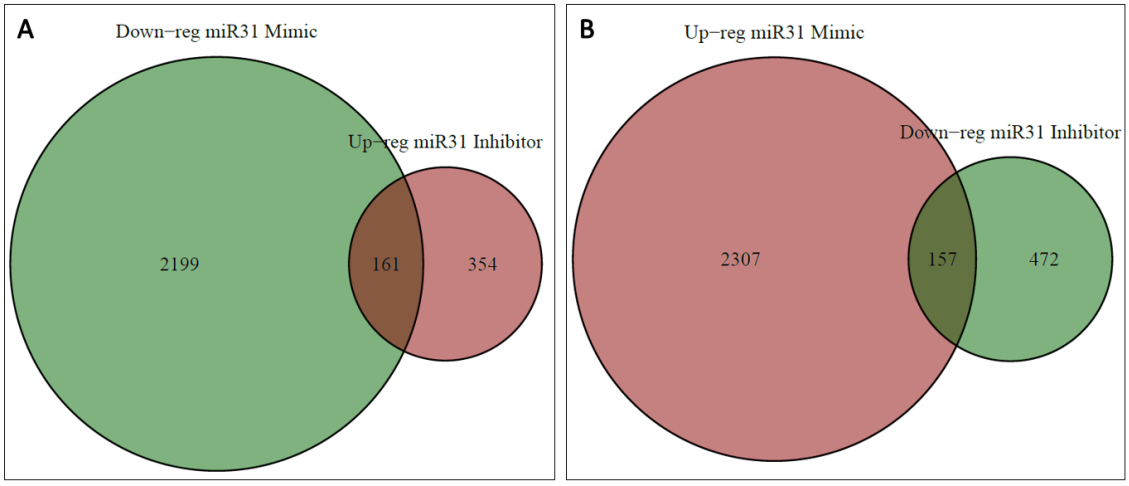


Figure 3.9 ***Venn diagrams showing the number of genes significantly changed in miR31 mimic and inhibitor datasets***

A: 161 genes simultaneously decreased with miR31 mimic, and increased with miR31 inhibitor treatments.
B: 157 genes simultaneously increased with miR31 mimic, and decreased with miR31 inhibitor treatments.

To narrow validation of RNA-seq on a select few genes, I then cross-checked the list of 161 genes that were repressed with miR31 mimic and increased with miR31 depletion against a list of 477 genes bioinformatically predicted as direct miR31 targets by Targetscan (Release 7.2). This comparison indicated twenty-two genes in common which are underlined in *Table 3.7.*

Table 3.7 ***Genes significantly changed in both sequencing datasets. Underlined genes are predicted as direct miR31 targets by TargetsScan7.2***

157 genes that both ***Increased* with miR31 mimic and** ***Decreased* with miR31 inhibitor**
AMMECR1, DCP2, YTHDF2, SREK1IP1, CA2, F2RL2, TXNRD1, EAF1, OSGIN1, CYP4V2, CFHR1, NMRAL2P, THUMPD1, ASF1A, CFAP300, GPR162, HMOX1, OGA, BCL2L2, TDG, TYW3, RAB39B, STK35, PDIK1L, EPHA7, PKIA, YY2, ZC3H6, WASF3, HDHD2, SDSL, TMEM170A, KPNA5, THNSL1, RIT1, NDUFA13, MAP3K13, AC007796.1, PFN2, C5ORF15, PPM1K, CDNF, B4GALNT1, RP11-473I1.9, RBKS, RP11-473I1.10, CREBRF, PGAM5, ZFHX2-AS1, PER1, ZNF791, BCAN, ZNF697, PDP1, ZNF736, FOPNL, LINC02762, MAFG, IFITM1, DPM3, MIR22HG, ZNF682, NEAT1, LUC7L, ZNF117, PLS3-AS1, RP11-111M22.3, ZEB1-AS1, EXOC3-AS1, MX2, STAT5B, MOSMO, RP11-288L9.4, TIPARP, F3, OVGP1, OASL, RSAD2, DNM3OS, LINC00475, MIR29B2CHG, MTF1, IFIH1, RP11-262H14.4, RP1-39G22.7, LENG8-AS1, LINC01003, LONRF2, ZNF529-AS1, DNAJC3-DT, C18ORF25, ZNF326, MYCL, RFLNB, SGPL1, SRPK3, CNOT6LP1, CKMT2-AS1, RP11-227H15.4, C15ORF40, MIRLET7I, C20ORF204, HERC6, SYT5, RP3-428L16.2, SCAMP1-AS1, NECAB2, VAMP2, RP11-45A16.4, BPG299F13.17, KPTN, SLC3A2, IFI44L, IRF7, ODF3B, HOXB2, CMPK2, ASH1L-AS1, DCP1A, ZNF680, RP11-694I15.7, ARRDC1-AS1, SH3BGR, SMTNL1, CTD-2017D11.1, CEBPG, UBAC2-AS1, RP11-104G3.2, ZNF429, PCGF1, TTC32, TRIM22, ISG15, TRIM52, SLC16A1, RP11-158H5.7, ZNF486, TMEM147-AS1, PAXIP1-AS1, MAP2, ELOVL6, ZNF737, GRPEL2, ZNF91, ZNF626, ARMT1, YTHDC2, FBXO30, CTD-2561J22.2, ZNF606, BST2, FBLL1, ANKRD12, IRF2BP2, FENDRR, ADAR, SAMD9
161 genes that both **D*ecreased* with miR31 mimic and** ***Increased* with miR31 inhibitor**
STK40, KCTD20, DICER1, SORL1, TP53INP2, PLAGL2, RP4-607J23.2, GYPC, ACAA2, ARHGEF40, IGF2BP2, ITGB3, MAP3K3, RPS6KB2, PI16, HIST1H2BK, TLR4, EIF4EBP1, ACVRL1, TMEM109, B4GALT5, PLEKHA1, FOXN3, NAGPA, CXCL12, NYNRIN, ITPKB, GPR137, PACS2, ZNF815P, CPA4, NSMF, C17orf103, RP11-21L23.2, FAM222B, NFE2L3, BAP1, DUSP22, SH3GL1P3, TTC28, TUBB2B, ABR, CAP2, SMAD6, MAMDC2, SLC29A1, ZNF236-DT, TSPEAR-AS1, PREB, SULF1, ZDHHC3, UCN2, PTGIS, FAM217B, CHST2, STRA6, RAB5B, ECE1, ACSS2, TNS1, LLGL1, AP1M1, CCNJL, PODXL, PKN2, TMEM104, AC093616.4, FAM214B, STK38, TSPAN9, MYH1, CSRP1, BTRC, B3GALT2, GPR157, GPX8, ABLIM1, IGF2BP3, TBC1D13, ENG, LARGE1, LRRC28, CAPN5, FBLIM1, CD248, ZDHHC16, TTC13, SPATC1L, VEGFB, OSBPL5, IVNS1ABP, STARD3, SNPH, GXYLT1, NCK2, CDH11, MINAR1, RNF216P1, NOTCH3, SH3RF2, CHMP1A, SYNC, TSPAN11, TP73-AS1, ABCD1, TMEM168, INPP5K, HYAL2, FAM160B2, PCYOX1, BTN3A2, PAPOLA, FBLN1, UQCRC1, CTD-2044J15.2, CTC-205M6.5, ZFHX2, KANK1, LRRC8A, TLN2, EXTL3, SFTA1P, ASCC2, ANK1, RP11-288H12.4, SLC43A2, JMJD8, ZC3HAV1L, TRAK2, LEPROTL1, SOGA1, IDH2, PCDHGB7, TRAPPC12-AS1, PRKAR2A, CLCN3, ADCY9, FARP1, TNFRSF10C, TM9SF4, NTAN1, CPPED1, SVEP1, RHOG, ZFYVE26, ZDHHC7, RASSF7, NRG3, ACP7, SFXN3, PHETA2, LMOD1, LITAF, CTD-2196E14.9, THBS2, DAB2, PDPN, CCN3, HYAL3, NUPR1, CABIN1

I then conducted a literature search querying a role in fibrogenesis of the twenty-two genes that had simultaneously decreased with miR31 mimic, increased with miR31 inhibitor, and that were predicted miR31 direct targets. Following the literature search, three genes were of interest including BRAC1-associated protein 1 (BAP1), Protein Kinase N2 (PKN2) and Tensin 1 (TNS1). BAP1 is a deubiquitinating enzyme that has been shown to control chromatin accessibility of key cytoskeleton genes thereby promoting cell motility in liver organoids which was of interest given the increase in fibroblast motility observed in pro-fibrotic fibroblasts (Artegiani et al. 2019). TNS1 is a cell membrane protein that binds fibronectin in the ECM and one study showed TNS1 was essential for myofibroblast differentiation and ECM production in pulmonary fibroblasts (Bernau et al. 2017). PKN2 is an effector protein of Rho GTPases in a range of cellular pathways including cytoskeletal organisation (Vincent et al. 1997), cell cycle (Schmidt et al. 2007) and cell migration (Lachmann et al. 2011) and a recent study showed that cardiomyocyte specific deletion of PKN2 results in cardiac dysfunction and fibrosis in mice (Sakaguchi et al. 2019). Further, Serine/Threonine Kinase 40 (STK40) was among the most significantly changed genes in both the miR31 inhibitor (adj. $p = 2.65 \times 10^{-35}$) and mimic (adj. $p = 3.37 \times 10^{-50}$) sequencing datasets and is a previously validated miR31 target (Yu et al. 2018), therefore despite no documented role of STK40 in fibrosis, I used STK40 as a positive control in RNA-seq validation.

Before validating RNA-seq with RT-qPCR, I first questioned which genes could serve as reference genes for normalisation in this assay. To answer this I screened my RNA-seq datasets to determine which common housekeeping genes had, or had not, changed following miR31 manipulation. Housekeeping genes that had changed and were therefore not selected included: GAPDH, RS18, FLMNA, FLMNB, HBMS, IPO8, PGK, POLR2A, TFRC, UBS and YWHAZ. Conversely, actin beta (ACTB), hypoxanthine phosphoribosyl transferase 1 (HPRT1) and TATA box binding protein (TBP) were housekeeping genes that did not appear to change with miR31 manipulation and therefore the expression of these genes was determined using RT-qPCR. Results indicated that ACTB, HPRT1 and TBP remained stable and I therefore used the geometric mean of these genes for normalisation in RT-qPCR (*figure 3.10 upper graphs*). STK40, BAP1, TNS1 and PKN2 mRNA levels were then checked and results indicated a congruent and statistically significant change in STK40, BAP1 and PKN2 in both the miR31 mimic and inhibitor experiments when compared to

control (*figure 3.10 lower graphs*). TNS1 did follow the predicted trend, however statistical analysis suggested that this was not significant which likely reflects the difference of error introduced by the separate methodologies of gene expression measurements.

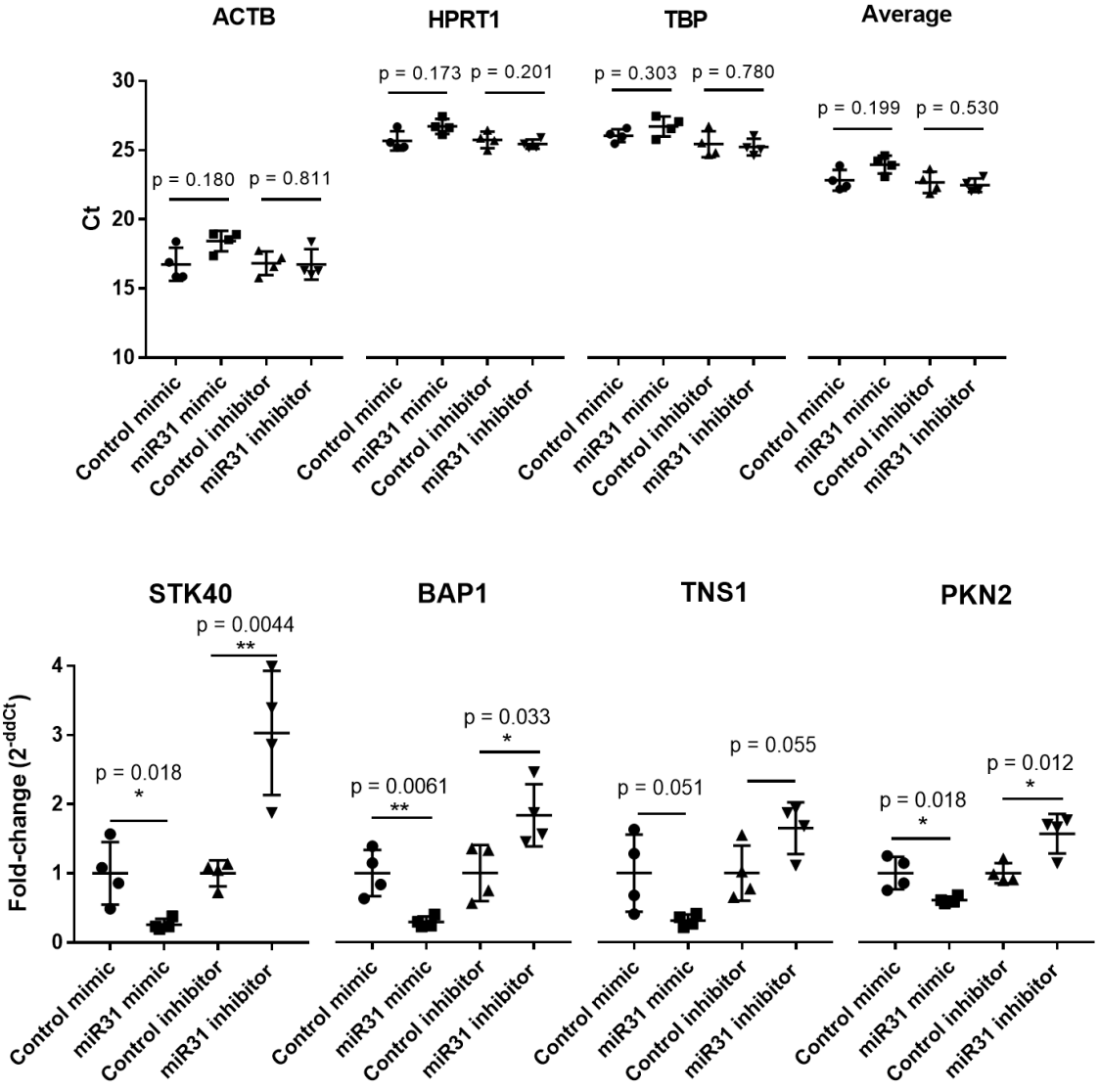


Figure 3.10 ***RT-qPCR of putative direct miR31 targets in hAFs validates RNA-seq datasets***

The same RNA samples that were used for sequencing were assessed by RT-qPCR for changes in direct miR31 targets predicted by TargetScan7.2. Paired Student's *t*-test used for statistical analysis as data was shown to be parametric with Kolmogorov–Smirnov test. Results are expressed as fold change of control (either mimic control or inhibitor control).

In experiments so far I have investigated the effects of miR31 abundance on the transcriptomic profile of hAFs *in vitro*. To achieve this I first validated my hAF cell model by showing abundance of key intracellular and surface markers and then confirmed enrichment and depletion of miR31 in these cells following treatment with miR31 mimic and inhibitor respectively. Subsequent to these interventions the poly-A tail transcriptome was sequenced to attain an unbiased view of miR31 gene regulation and these results were analysed using Ingenuity Pathway Analysis software. Results indicated that changing miR31 abundance significantly affects transcript abundance of genes involved in a wide range of cell signalling pathways including BMP, PDGF, FGF, TGFβ1 and ERK/MAPK signalling, cardiac hypertrophy, hepatic fibrosis and integrin signalling. These results will help explain the role of miR31 in fibrosis in hAFs when considered with proteomic measurements and functional assays. Following this observation, I then validated the RNA-seq data with RT-qPCR by focusing on genes that had shown a reduction with miR31 enrichment, an increase with miR31 depletion, and that were also predicted miR31 targets that had been shown previously to be important in fibrosis.

3.5 MiR-31 abundance effects fibrotic phenotype in human atrial fibroblasts *in vitro*

It is unknown whether miR31 has a necessary or sufficient role in driving pro-fibrotic phenotype changes in hAFs. In this part of the project I sought to determine whether manipulating miR31 abundance would affect hAF cell proliferation and collagen production which are hallmark phenotypic features of pro-fibrotic fibroblasts.

3.5.1 Assessing cell proliferation

To determine whether miR31 abundance effects hAF cell proliferation I used BrdU ELISA assay that measures BrdU, a synthetic nucleotide, incorporation into DNA as a surrogate for changes in cell number. A pilot experiment was performed on hAFs from donor 1 cultured in a single 96 well plate whereby in eight technical replicates, miR31 was silenced or overexpressed with 2.5 pmol miR31 inhibitor or mimic. Results of this initial pilot experiment indicated that increasing miR31 abundance did increase hAF cell proliferation while reducing miR31 had an opposing effect (*figure 3.11*).

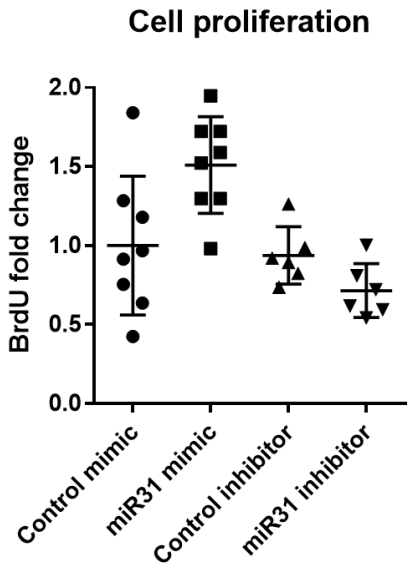


Figure 3.11 ***Donor 1 hAF cell proliferation following enrichment or depletion of miR31 with small oligonucleotides***

hAFs on passage 7 were cultured in a single 96-well plate and lipofected using RNAiMax in eight technical replicates on the same day with 2.5 pmol of miR31 mimic and inhibitor oligonucleotides with respective controls. Twenty-four hours later cells were treated with BrdU and following another 48 hours cell proliferation was estimated using BrdU ELISA assay (Abcam: 126556) and plate spectrophotometer. hAF donor 1 (0000634284) was used for this experiment. Error bars depict mean ± SD however no statistical comparisons were made as this experiment represents data from a single experimental repeat performed using a 96 well plate.

I then sought to repeat the above experiment, however at this stage of the project cells from donor 1 (0000634284) had been depleted so experiments were proceeded with donor 2 (18TL312486). These cells had also been validated previously by the supplier Lonza by positive immunofluorescent staining for collagen type I, and negative for von Willebrand factor VIII related antigen (*see appendices for certification*). Given that Lonza certification of donor 1 was shown to be valid by our in-house authentication (*section 3.1*), we trusted Lonza certification of donor 2 and did not perform additional validation of donor 2 cells in-house. Before proceeding with functional assays in donor 2, I sought to show that miR31 abundance could be controlled with miR31 mimic and inhibitor in donor 2 specifically. Results indicated a significant increase in miR31 in mimic treated cells ($p = 0.035$) and a reduction of miR31 in inhibitor treated cells ($p = 0.005$) (*figure 3.12*).

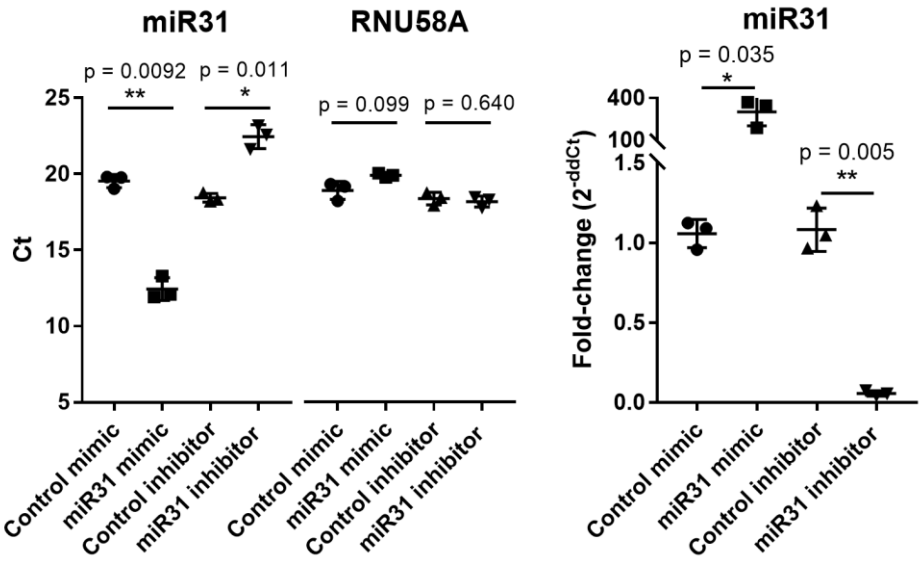


Figure 3.12 ***miR31 expression in hAF donor 2 (18TL312486) following enrichment or depletion of miR31 with small oligonucleotides***

hAFs from donor 2 on passage 7 were cultured in a single 24 well plate and lipofected in technical triplicates using RNAiMax with 15 pmol of miR31 mimic or inhibitor and their respective controls. Forty-eight hours following lipofection RNA was harvested for RT-qPCR, RNU58A was used as housekeeping gene for normalization. The experiment was repeated three times and graph data points are the average from each experimental repeat. Paired Student's *t*-test was used for statistical analysis. Results are expressed as fold-change of control. Error bars depict mean ± SD.

Having shown successful manipulation of miR31 in hAFs from donor 2, cell proliferation was then measured in these cells using BrdU ELISA assay (Abcam: 126556). HAFs were cultured in a single 96 well plate, and in four technical replicates miR31 was silenced with 2.5 pmol of inhibitor and overexpressed with 2.5 pmol of mimic. This experiment was then repeated three times. Results indicated that manipulating miR31 abundance had no effect on cell proliferation of hAFs from donor 2 (*figure 3.13*).

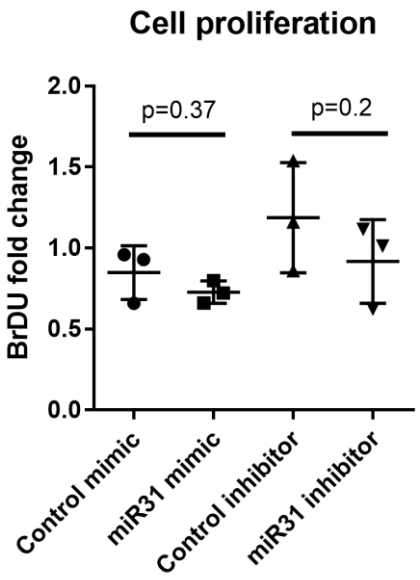


Figure 3.13 ***Cell proliferation in donor 2 hAFs (18TL312486) following enrichment or depletion of miR31 with small oligonucleotides***

hAFs on passage 7 were cultured in a single 96-well plate and lipofected, using RNAiMax, in four technical replicates with miR31 mimic and inhibitor oligonucleotides. Twenty-four hours later cells were treated with BrdU and following another 48 hours cell proliferation was estimated using BrdU ELISA assay (Abcam: 126556). The experiment was repeated three times and graph data points are the average from each experimental repeat. Paired Student's *t*-test was used for statistical analysis. Values are reported as fold change of untreated cells. Error bars depict mean ± SD.

TGFβ1 is a potent growth factor previously shown to stimulate cardiac fibroblast cell proliferation *in vitro* (Sigel et al. 1996). Given that hAFs from donor 1 showed changes in proliferation following miR31 manipulation, but this result was not reproducible in hAFs from donor 2, I queried whether TGFβ1 would cause proliferation in these cells. This experiment would inform me whether hAFs from donor 2 were not responding to miR31 as expected due to a ceiling effect in proliferation being reached in these cells. Indeed when donor 2 hAFs were cultured in a 96 well plate and treated with 0.5 ng of TGFβ1 following lipofection with 2.5 pmol of miR31 mimic, inhibitor or their respective controls, there was no observed effect on proliferation (*figure 3.14*). This result suggested that hAFs from donor 2 had reached a ceiling effect in their capacity to proliferate in response to TGFβ1, a known stimulant of proliferation, and therefore drawing conclusions on the effects of miR31 manipulation in hAF cell proliferation from donor 2 would not be viable.

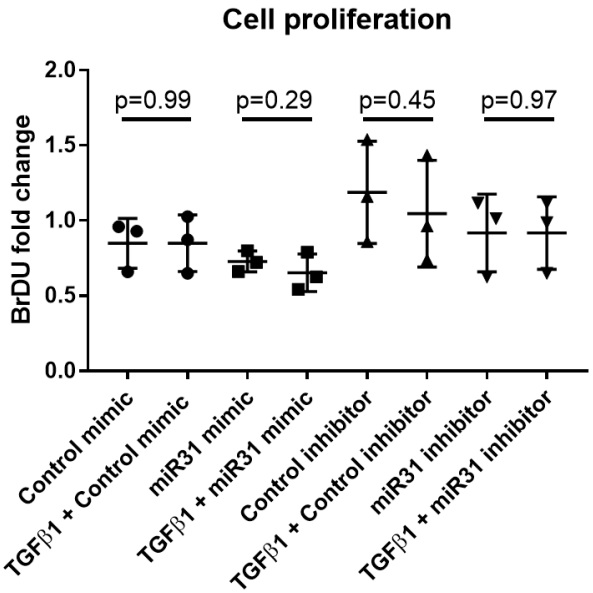


Figure 3.14 ***The effects of TGFβ1 on cell proliferation in donor 2 hAFs***

Donor 2 hAFs were cultured in a 96 well plate and lipofected with 2.5 pmol of miR31 mimic and inhibitor oligonucleotides. Twenty-four hours later cells were treated with BrdU and 0.5 ng TGFβ1 and following another 48 hours cell proliferation was estimated using BrdU ELISA assay. The experiment was repeated three times and graph data points are the average from each experimental repeat. Paired Student's *t*-test used for statistical analysis. Values are reported as fold change of untreated cells. Error bars depict mean ± SD.

3.5.2 Assessing collagen synthesis

Further to the above cell proliferation assay, experiments were also performed on hAFs from donor 2 to query collagen production following miR31 enrichment or depletion using Sirius Red Collagen detection kit (Chondrex: 9062). Results of this experiment showed extreme variability in the control groups which indicated that the experimental technique had not been adequately optimised by myself (*figure 3.15*). See *section 3.6* below for thoughts on optimisation strategies.

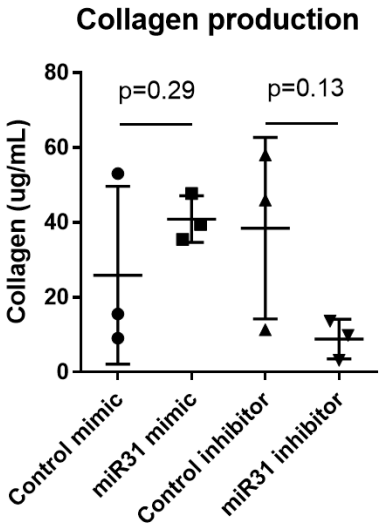


Figure 3.15 ***hAF cell collagen production following enrichment or depletion of miR31 with small oligonucleotides***

hAFs on passage 7 were cultured in a single 24-well plate and lipofected with miR31 mimic and inhibitor oligonucleotides. Twenty-four hours later cells were treated with ascorbic acid and 1.5 ng TGFβ1, and following another 48 hours collagen production was estimated using Sirius Red Collagen detection kit (Chondrex: 9062). hAF donor 2 was used for this experiment. The experiment was repeated three times and graph data points are the average from each experimental repeat. Paired Student's *t*-test was used for statistical analysis. Error bars depict mean ± SD.

In addition to looking at the effects of miR31, I also queried whether TGFβ1 would cause collagen production in donor 2 hAFs. TGFβ1 is a known stimulant of collagen synthesis in fibroblasts (Butt et al. 1995) and therefore the addition of TGFβ1 in the experimental design would serve as a positive control. Results indicated that in the presence of TGFβ1, a significant increase in collagen production was observed with miR31 mimic ($p = 0.0016$) and therefore donor 2 cells were likely sensitive to TGFβ1 however, again, the extreme variability within experimental groups made the data largely uninterpretable.

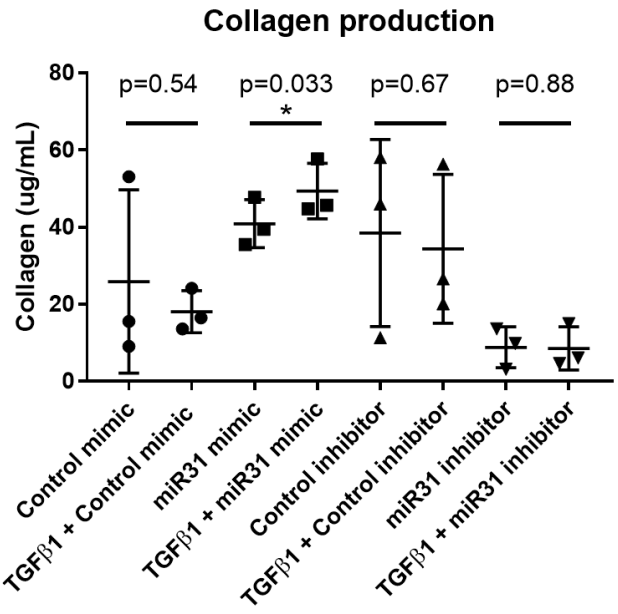


Figure 3.16 ***hAF cell collagen production following enrichment or depletion of miR31 with small oligonucleotides***

hAFs on passage 7 were cultured in a single 24-well plate and lipofected with miR31 mimic and inhibitor oligonucleotides. Twenty-four hours later cells were treated with ascorbic acid and 1.5 ng TGFβ1, and following another 48 hours collagen production was estimated using Sirius Red Collagen detection kit (Chondrex: 9062). hAF donor 2 was used for this experiment. The experiment was repeated three times and graph data points are the average from each experimental repeat. Paired Student's *t*-test was used to compare groups. Error bars depict mean ± SD.

3.6 Discussion and plans for future experiments

In this chapter I explored the effects of depleting and enriching miR31 abundance in primary hAFs with a particular interest in whether miR31 would facilitate pro-fibrotic gene expression or phenotype changes in these cells. Due to closure of the University buildings during the Covid-19 pandemic my experiments were necessarily cut short and further work is required to comprehensively address the aims of this part of the project. The forthcoming paragraphs give an outline of experiments performed while highlighting the limitations of my results and provide details of future experiments that ideally would have been done to complete this chapter.

Experiments were conducted on primary hAFs and at the onset I validated my experimental system by demonstrating intracellular expression of Vimentin and surface expression of THY1 (i.e. CD90). I also demonstrated negligible surface expression of PDGFRα which is significantly upregulated in disease state myofibroblasts indicating that these cells were appropriate for experimentation (Chong et al. 2013, Ivey et al. 2016). This validation part of the chapter could have been improved by also querying the level of expression of αSMA which is the most widely used biomarker for identifying the differentiation of fibroblasts to myofibroblasts. Cardiac fibroblasts are known to differentiate in culture with increased passage number and attachment to plastic culture plates (Santiago et al. 2010, Landry et al. 2019). HAFs were received on passage three which I then expanded and cryopreserved at passage six to be later used for experimentation (*see section 2.2.1 receipt, expansion and storage of hAFs*). Given that these cells were subject to a series of passages before commencing experimentation, it would have been informative to gauge the expression of αSMA using qPCR and immunoblotting at different passage numbers. Such data could have been used as a point of reference to determine whether my hAF population had already reached an activated state and could be used to explain any discrepancies observed in experiments that I later performed. Despite this limitation, I performed all experiments on cells on identical passage number (passage 7) with careful attention to plating technique to try and normalise for the effects of passaging.

Following validation of hAFs by looking at expression markers I then commenced experiments and looked at the effects of enriching and depleting miR31 in these cells. I queried changes in the transcriptome with RNA-seq as well as changes in cell proliferation and collagen production. In addition to each individual experiment having its own limitations, which are discussed shortly, all experiments in this chapter shared the overarching caveat that they were performed as technical replicates, not biological replicates and were therefore pilot studies. To improve the validity of these pilot studies, I chose to perform each technical replicate on separate weeks on freshly thawed cells as opposed to simply performing multiple technical replicates in a single tissue culture plate. Taking this approach would at least reflect the variability associated with the entire longitudinal experimental protocol and not just measurement variability from a single experimental run. This increases confidence in my findings, but future experiments should still be performed using biological replicates to reflect biological variability and improve the validity of these experiments.

Another overarching caveat to this chapter was that validation of hAFs from donor 1 was more comprehensive than donor 2, but despite this, the bulk of the BrdU proliferation and Sirius red collagen detection experiments were performed on donor 2 cells. In addition, I included TGFβ1 treated cells as a positive control when performing experiments with donor 2 cells however this was not used earlier in the chapter when using donor 1 cells. Consequently the study could have been improved by validating both donor hAFs, as well as future hAF samples, for intracellular vimentin and surface expression of THY1 while also measuring αSMA expression to determine the extent of fibroblast to myofibroblast trans-differentiation in each donor. In addition, I would use a TGFβ1 treated hAF cell population as a positive control group in all my experiments to first show a robust pro-fibrotic response with the particular donor population in hand. This would help elucidate whether or not a ceiling effect had been reached in these cells and whether it would be practical to then look for additive effects of miR31.

The RNA-seq experiment was performed as technical replicates repeated on four different days and using this dataset I was able to narrow down a gene list of putative direct miR31 targets of potential importance to fibrosis which included BAP1, PKN2 and TNS1 (Bernau et al. 2017, Artegiani et al. 2019, Sakaguchi et al. 2019). I then performed validation experiments for the RNA-seq data by showing congruent changes in BAP1, PKN2 and TNS1 using qPCR however protein changes in these genes still need to be confirmed with immunoblotting. In addition, follow up experiments are needed to determine whether these proteins are necessary and/or sufficient for pro-fibrotic miR31 signalling in hAFs. For example, BAP1, PKN2 and TNS1 could, in succession, be targeted for knockdown in hAFs, using gene silencing oligonucleotides, to determine whether the absence of any of these proteins stimulate hAFs to proliferate, produce collagen or differentiate. Conversely, gain of function experiments with expression vectors of these gens could be used to determine if increased abundance of these proteins blunts pro-fibrotic miR31 signalling. Collectively these experiments would help elucidate whether BAP1, PKN2 and TNS1 are essential downstream mediators of miR31 fibrotic signalling in hAFs.

With regards to the BrdU cell proliferation assay, this assay was first performed on donor 1 cells in a single 96 well plate and initial results suggested that miR31 increased hAF proliferation. However when this assay was proceeded with donor 2 cells, and the experiment repeated on three separate weeks, results showed high variability between experimental runs making it difficult to detect the effects of miR31 on proliferation. Furthermore, when TGFβ1 was used as a positive control in this experiment, results were similarly variable and as a consequence, further optimisation is needed to reduce the variability between experimental repeats. One approach could be to monitor proliferation over a shorter time period during the 48 hour protocol, for instance every 4-6 hours as opposed to every 16-24 hours, this would help reduce the chance of cells overcrowding and clustering which may be contributing to signal noise when BrdU intensity is measured.

BrdU signal represents changes in the amount of newly synthesised DNA, which is the biochemical signature of cellular replication, however it is plausible that apparent changes observed with BrdU labelling are due to other biological changes as a result of miR31 mimic/inhibitor expression. For example, newly generated hAF daughter cells may have improved or impaired survival and/or plate attachment characteristics. As a consequence, BrdU signal would not necessarily reflect the total number of new daughter cells, but rather the remaining population of daughter cells that endured the length of the protocol. In light of this caveat, the BrdU experiment would have been improved by normalising BrdU signal to the total number of cells in respective wells. The normalization step would help to remove sample-to-sample variability caused by factors other than that being tested in the experiment, namely proliferation. Total cell number could be estimated by DAPI staining hAFs following completion of the BrdU-ELISA assay and then counting the total number of stained nuclei with ImageJ.

In addition, future experiments could use a TUNEL assay to look specifically at the effects of miR31 mimic/inhibitor on hAF cell survival in order to better gauge the validity of the BrdU assay results. Alternatively, other methods of assessing proliferation can be used in addition to the BrdU assay. For example, Ki67 is a nuclear protein expressed during all phases of the cell cycle, but is absent in quiescent cells, and is commonly used to visualise cell proliferation. Ki67 immunofluorescence staining could be used to detect hAFs undergoing the cell-cycle, and like the BrdU assay, this value could then be normalised to total cell number of hAFs stained blue with DAPI.

Similar to the BrdU experiment, the Sirius red collagen detection assay was repeated on three separate weeks and showed high variability between experimental runs and requires optimisation. The collagen detection assay uses Sirius Red to bind and precipitate collagen however the small quantities of collagen generated *in vitro* can make it difficult to visualise, wash and re-suspend which may explain the high variability within groups. This experiment may therefore be improved by increasing the tissue culture plate size used, and hence cell number, from which the produced collagen is retrieved thereby increasing the visibility of the collagen pellet during purification.

Furthermore, the variability with collagen measurements could be due to differences in cell numbers in each experimental run, therefore the collagen content in future experiments could be normalised to total cell number which could be determined with DAPI staining hAFs after collagen retrieval from the well media. Alternatively collagen content could be normalised to total protein content which can also be extracted after the collagen retrieval step. Furthermore, activated hAFs produce more collagen and display increased proliferation rates compared to inactive hAFs and thus, as mentioned earlier in this section, all future experiments would ideally first measure the expression of αSMA using qPCR and immunoblotting to query the extent of hAF activation before proceeding with functional assays.

Chapter 4

Results: Investigating transcriptional regulation of miR31 by ETS1 in human atrial fibroblasts in vitro

We had shown previously that patients with chronic AF had increased miR31 abundance in atrial fibroblasts derived from the RAA when compared to sinus rhythm donors (*see figure 3.1 above*) and following this observation we speculated on the upstream regulation of miR31. MicroRNAs originate from long primary transcripts that undergo a series of processing steps to generate mature microRNAs, therefore their biogenesis can be affected at many levels (*see introduction section 1.5.1*). Given that we had RNA samples extracted from atrial fibroblasts of patients with AF, it was feasible to then explore the same RNA for changes in pri-miR31. This information would be a more direct indication as to whether LOC554202, the long non-coding RNA which encodes the miR31 sequence, was showing changes in transcriptional activity. Results of this enquiry showed a borderline increase in pri-miR31 abundance in patients with chronic AF compared to sinus rhythm donors which indicated that increased LOC554202 transcription may be causing increased miR31 abundance (*figure 4.1*).

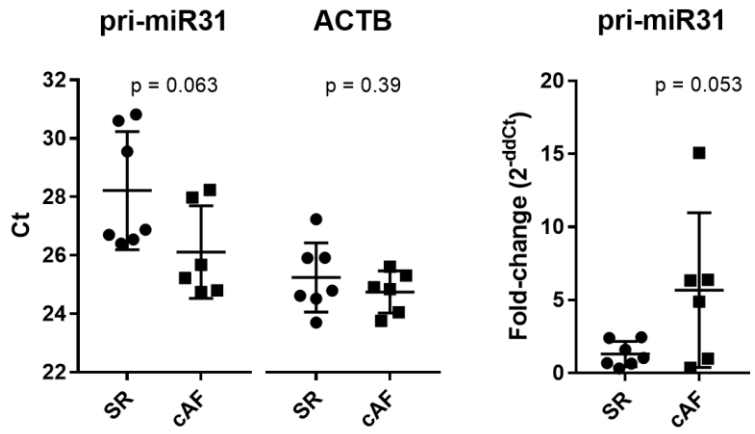
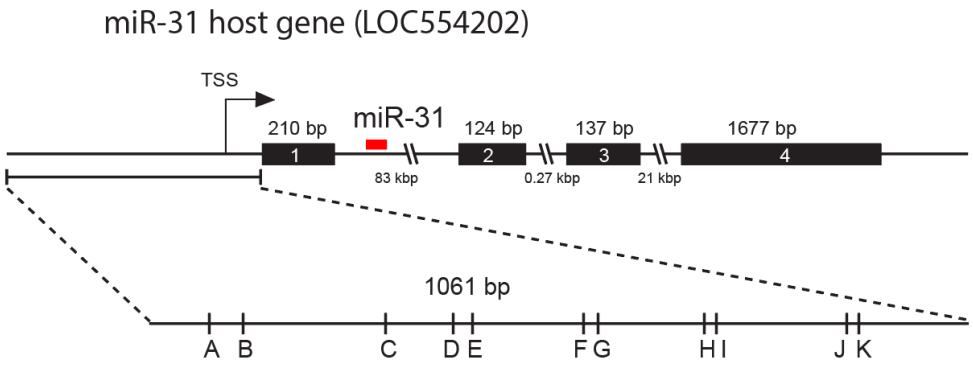


Figure 4.1 ***pri-miR31 abundance in human right atrial appendage***

RNA was extracted from fibroblasts derived from the right atrial appendage from patients with SR and cAF to measure pri-miR31 expression by qPCR. Graphs depict raw Ct values for pri-miR31 and ACTB, as well as pri-miR31 normalised to ACTB and presented as a fold-change of SR control. K-S statistical test was used to determine that data was parametric and therefore groups were compared using an unpaired Student's *t*-test. SR patients n = 7, cAF patients n = 6. Error bars depict mean ± SD. Actin beta (ACTB).

Further to the above result, I conducted an *in-silico* analysis using JASPAR (2018 release) of a ~1 kbp segment immediately upstream of exon-1 of LOC554202. This approach identified eleven putative ETS1 binding sites which were denoted as A-K. ETS1 is a transcription factor shown to bind the LOC554202 promoter, to favour transcriptional repression by employing EMSY and KDM5B, and has potential roles in atrial fibrillation and fibrosis (*see introduction section 1.5.1).* For further investigation into the identified ETS1 binding sties at the LOC554202 promoter, I performed a similar *in-silico* analysis on the same ~1 kbp segment using TRANSFAC (2019 release) which confirmed ETS1 binding sites E and G. The ENCODE 3 database was also queried for ETS1 ChIP-seq data at the LOC554202 promoter which showed an ETS1 binding peak over sites H and I in A549 cells, a lung epithelial cancer cell line. Lastly, a literature search of ETS1 involvement in miR31 transcriptional regulation indicated ETS1 binding at sites G and K by ChIP-qPCR in MCF-7 cells, a breast cancer cell line (Vire et al. 2014). All aforementioned ETS1 binding sites at the LOC554202 promoter and their methods of identification are depicted below in *figure 4.2*.



5’-TGGGCTTGCGGAGGGCATAGTTGAGTAGGACAGCCAGAACTTCTGGCTAGGTCATCGGGTCGGA

TTACCCTGCCATGTACCCCTCAGTATTTA**CTTCC**AGAGATGGGCAACTTACTATAGTCTTCATGTT**TAA**
(Site A) **(B)**
GGAATTCAAACCCAAATGAAAAGTGTTGTGAAAGCATAAATATCTCTGCTTGTATACTACAGTACTGA

ATATGAAACCACATATTGCAATAAAAATAACTAACGATGGCCCATGCATTTCACAAGGCGCTCTCATA

CACCTGAAGGGGCAGT**GGAAA**GGTTCAGAGCACCCAGGCTGTCCCAAAGAGGGC**GGAAT**CAAATCT
(C) **(D)**
GGGCGCAGGTGGGTACCC***GGAAA***AAATTCTTGGTGGAGATGGCATCTGAGAGAGGCGTCGAAGGA
(E)
CACGGAGGGTTGAGGATAGGAGAGAGGCCACTCCAGATGGAAAAACACCCTCGTTTTCCAAATGCC

AGTTATACCACGGGGCCTGTGATAGCCGCCACGTTTCTAGAAGACCCCTAGAAGAGCAGGGGTCCGA

CCAA**GGAAG**TA***CTTCC***GCTGTTCAATTTGCTAAGTTTGCATTCGGAGAGTACAAGGGCGAGCTCTGG
(F) **(G)**
CTCTTGCAGGTGGACTCCCTCTCCCTTAGCTCTGACTGCGAGAAGCTCCCACCCGCAAAGGGCTACTG

TGCCAAAAGCGAGCGTAGTGGCCGCCG**CTTCC**CCCACAAA**ATTCC**GTCACTCGGGAGCTGCGTTGCA
(H) **(I)**
GTGGAGAAACCTGGGTCCCGTCCGGGCCCAAGGGAGGAGGCTCAGCACGGCCGCTCCCGCAGCCCC

GTCCAGGCGCCT**CTTCCC**TCCCCCTCC**TTTCC**GCGGCGGTCGCGGCGGCGGCTGGAGCGGGAGCCGG
(J) **(K)**
GGGCGGCGAGGGGCGGAGCCGGGGCGGAGCCACGGCGCTGGCGGCCCCGGCGGCGGCGGCAGGT

GTGGATGGGGCGAAGGTGCGGGAACGTCC|ACTCCCGCGCGCCCCTCTCGGGGACCAGGGCGGCCT

CCCAGGAGGAGCTTGGTGAGCAGCTGCGACCTGTGCATAACTTGGGGCGCCGCCAGGGGCTCGC-3’

Figure 4.2 ***5’-3’ 1061 bp segment upstream of exon-1 of the long non-coding RNA LOC554202 with eleven identified putative ETS1 binding sites***

1061 bp (hg38, chr9: 21559699 – 21560759, negative strand) immediately upstream of exon 1 of LOC554202 was analysed in JASPAR which identified eleven putative ETS1 binding sites (denoted A-K and in bold). Analysis in TRANSFAC validated two sites (E and G, italicized) and the ENCODE-3 database indicated ChIP-seq ETS1 binding at two sites in A549 cells (H and I, highlighted blue). ETS1 binding by ChIP-qPCR in MCF-7 cells (G and K highlighted green) (Vire et al. 2014). Sites on the positive strand are indicated by a double underline (6 sites), while sites on the negative strand are indicated by a single underline (5 sites). LOC554202 TSS indicated by small vertical line highlighted yellow.

Following the above *in-silico* analysis, literature search, and querying available databases, we hypothesised that increased miR31 abundance in atrial fibroblasts was a result of reduced ETS1 binding at the LOC554202 promoter. Consequently, the abundance of ETS1 in the right atrial appendage (RAA) from patients with sinus rhythm and chronic AF was estimated with immunoblotting. Results did indicate a borderline reduction in RAA ETS1 protein abundance from patients with chronic AF when compared to sinus rhythm donors (*figure 4.3*).

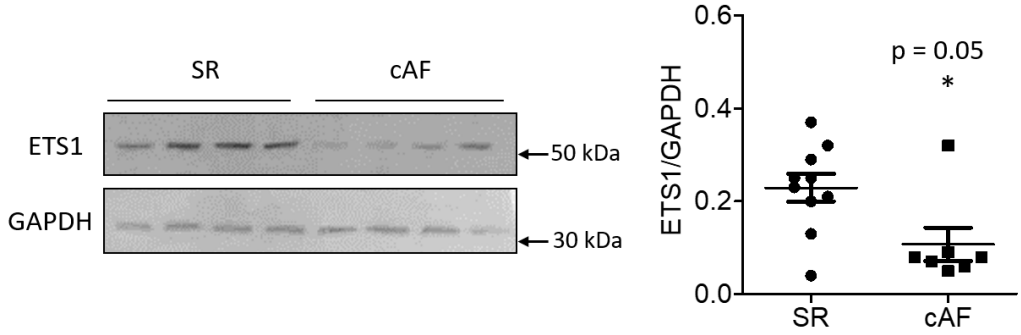


Figure 4.3 ***Immunoblotting of ETS1 in human right atrial appendage protein lysates***

Protein lysate was extracted from the right atrial appendage from donors with sinus rhythm (SR) and chronic atrial fibrillation (cAF). Band intensities were measured using ChemiDoc imaging software (version 5.1) and ETS1 measurements were normalised to GAPDH which was used as loading control. Kolmogorov–Smirnov test was used to determine that data was non-parametric and therefore groups were compared using Mann–Whitney U test. SR patients n = 10, cAF patients n = 7. Error bars depict mean ± SD.

Given the observed increase in both miR31 and pri-miR31 in atrial fibroblasts from patients with cAF compared to sinus rhythm controls, as well as the reduction in ETS1 protein in RAA lysates, experiments were devised to study the effects of ETS1 on LOC554202 transcription *in vitro* in hAFs. ETS1 abundance was either depleted by lipofection with commercially available antisense-ETS1 locked nucleic acids, or enriched with an ETS1 expression vector. Following these experimental interventions miR31 and its predecessor transcripts were measured. In addition to these experiments, I sought to use a luciferase-based reporter assay in hAFs to test whether mutating the eleven identified ETS1 binding sites (A-K) would alter LOC554202 transcription. However, I chose to first optimise this experimental approach in HEK293 cells, a model routinely used in such assays because of their reliability for transfection and ease of obtention.

4.1 The effects of loss- and gain-of ETS1 on miR31 and its predecessor transcripts in human atrial fibroblasts

I first asked whether ETS1 expression accounts for changes in miR31, within the normal physiological range of ETS1 abundance in primary hAFs and devised an ETS1 knockdown strategy using anti-ETS1 locked nucleic acids (LNA). To determine adequate LNA dose for ETS1 knockdown, donor 1 hAFs were cultured in 6-well plates and in single wells were treated with either 50, 100 or 200 pmol of anti-ETS1 LNA or control LNA and harvested 48 hours later for RNA analysis by qPCR. Results indicated significant reduction in ETS1 mRNA at all doses of anti-ETS1 LNA (*figure 4.4A*). To better gauge the duration of ETS1 knockdown, a time course experiment was then performed whereby RNA and protein were harvested 48, 72- and 96-hours post transfection with 100 pmol anti-ETS1 LNA. qPCR and immunoblotting indicated sustained reduction of ETS1 mRNA (*figure 4.4B*) and protein at all time-points (*figure 4.4C*). Because of these experiments, subsequent ETS1 knockdown experiments in hAFs were performed using 100 pmol anti-ETS1 LNA with harvest at 48 hours post transfection.

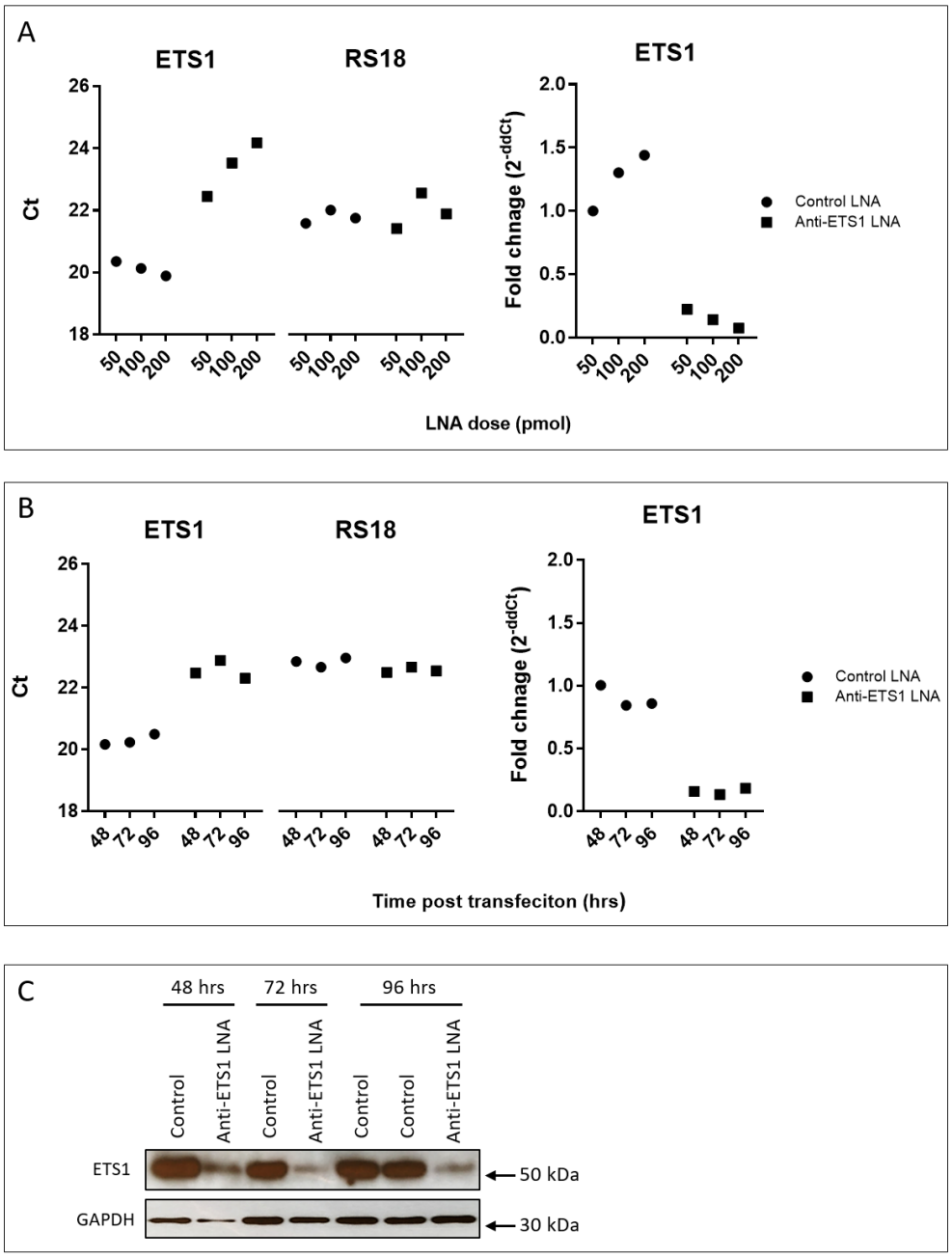


Figure 4.4 ***Dose and time-course experiments of ETS1 expression in primary hAFs following lipofection with anti-ETS1 LNA***

All experiments used donor 1 hAFs and were performed once in a single six well plate, therefore no statistical comparisons were performed on the data.

A: hAF cells were transfected with 50-200 pmol anti-ETS1 LNA, RNA was harvested 48 hours later for qPCR. Graphs depict raw Ct values as well as ETS1 normalised to RS18 and presented as fold-change of control.
B: hAF cells were transfected with 100 pmol anti-ETS1 LNA and RNA was harvested either 48, 72 or 96 hours later to measure ETS1 and RS18 expression using qPCR.
C: hAF cells were transfected with 100 pmol anti-ETS1 LNA. Protein was harvested either 48, 72 or 96 hours later for immunoblotting experiments to measure expression of ETS1 as well as GAPDH which was used as a loading control.

ETS1 knockdown in hAFs was then performed six times on separate weeks with each experimental repeat performed on hAFs on identical passage and from donor 1. RNA was then extracted and qPCR used to measure gene expression changes. Raw Ct values showed that ETS1 was significantly depleted and indicated that ribosomal protein S18 (RS18) remained stable and could be used as a housekeeping gene (*figure 4.5A*). Further, pri-miR31, a surrogate measure for changes in LOC554202 transcription did show a borderline significant reduction (p = 0.052) in its raw Ct value, and HG-miR31, a product of splicing pri-miR31, followed a similar trend (p = 0.069) (*figure 4.5A*). After normalizing Ct values to RS18, both pri-miR31 and HG-miR31 showed a slight increase however this was not statistically significant (*figure 4.5B*).

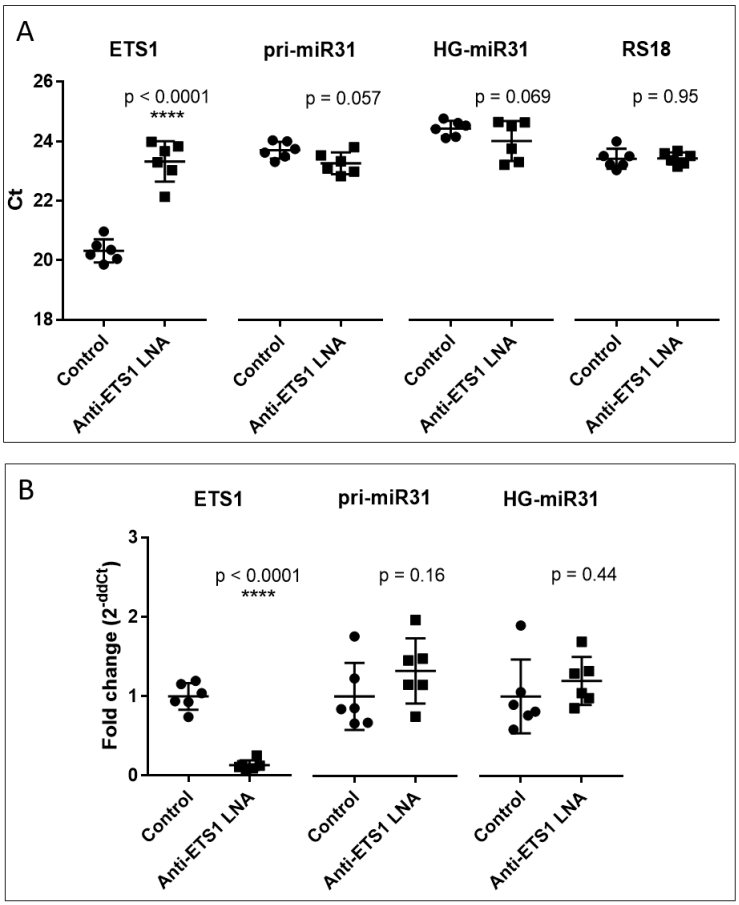


Figure 4.5 ***ETS1 depletion alters LOC554202 transcription in hAFs***

hAF total RNA was extracted 48 hours following ETS1 depletion with 100 nM anti-ETS1 LNA for qPCR experiments. Paired Student's *t*-test used to compare group means as data was shown to be parametric with K-S test. Experiment performed on hAFs from donor 1 and repeated on six different days. Error bars depict mean ± SD.

A: graphs of raw qPCR Ct values for ETS1, pri-mi31, HG-miR31 and RS18
B: ETS1, pri-mi31, HG-miR31 Ct values normalised to RS18 and presented as fold-change of control.

I then checked miR31 and RNU58A abundance. Raw Ct values indicated a statistically significant change in miR31 and that RNU58A was very variable but did not change on average following ETS1 depletion. Following normalisation to RNU58A, miR31 was shown to increase however this was not statistically significant (p = 0.11) (*figure 4.6*).

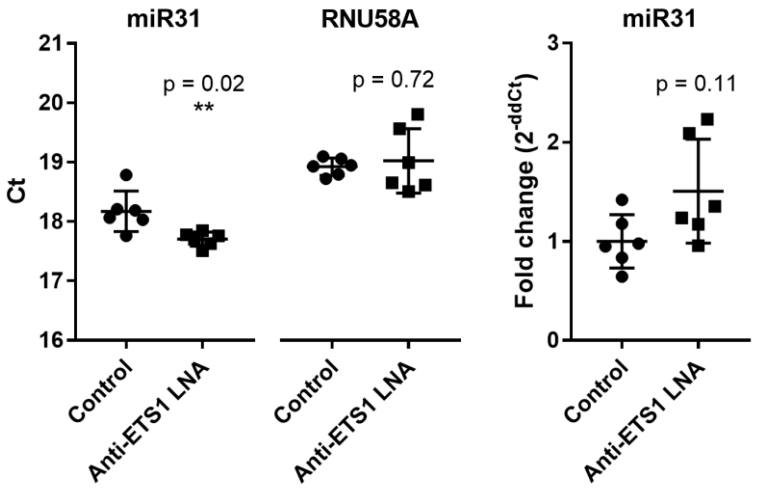


Figure 4.6 ***ETS1 depletion alters miR31 abundance in hAFs***

hAF total RNA was extracted 48 hours after ETS1 knockdown. Graphs depict raw qPCR Ct values for miR31 and RNU58A as well as miR31 normalised to RNU58A and presented as fold-change of control. Paired Student's *t*-test was used to compare means as data was shown to be parametric with K-S test. The experiment was performed on hAFs from donor 1 and repeated on six different days. Error bars depict mean ± SD.

In tandem to the above experiments, I asked whether enrichment of ETS1 in hAF cells would affect miR31 abundance. To query this I created an ETS1 expression vector to be used to increase ETS1 abundance in hAFs. The 5428 bp pcDNA3.1(+) vector was digested with Not1 and Kpn1 restriction enzymes which liberated a 58 bp segment of the multiple cloning site (MCS) while linearizing the remaining 5370 bp vector backbone. These digested products were then purified using gel-electrophoresis and an extraction kit (*figure 4.7*).

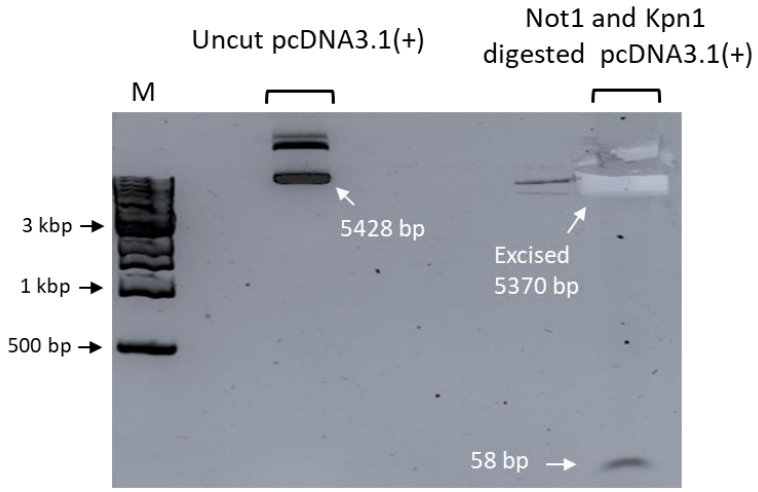


Figure 4.7 ***NotI/KpnI digested and gel purified pcDNA3.1(+)***

pcDNA3.1(+) expression plasmid was digested with KpnI and NotI restriction enzymes and then electrophoresed in a 1% agarose TBE gel containing 0.002% ethidium bromide for 90 minutes at 80 V. Digested band was then excised and purified with a kit (MP-Biomedicals: 111001600). NEB GeneRuler 1 kb used for DNA Ladder (M).

The ETS1 1326 bp ORF was then PCR amplified from purchased vector pDONR223_ETS1_WT (Addgene: 82118) using a KpnI forward cloning primer and NotI reverse cloning primer, which was then ligated into digested pcDNA3.1(+) expression plasmid. DH10-beta cells were then transformed with the resultant plasmid and five single colonies were checked using colony PCR amplifying the ETS1 ORF. Results indicated only colony one was carrying the expected plasmid (*figure 4.8*), it was therefore selected for expansion, plasmid purification and then validated with sequencing. The resultant ETS1 expression vector was named pcDNA3.1(+)_ETS1_WT and was used for ETS1 overexpression studies in both hAF and HEK293 cells .

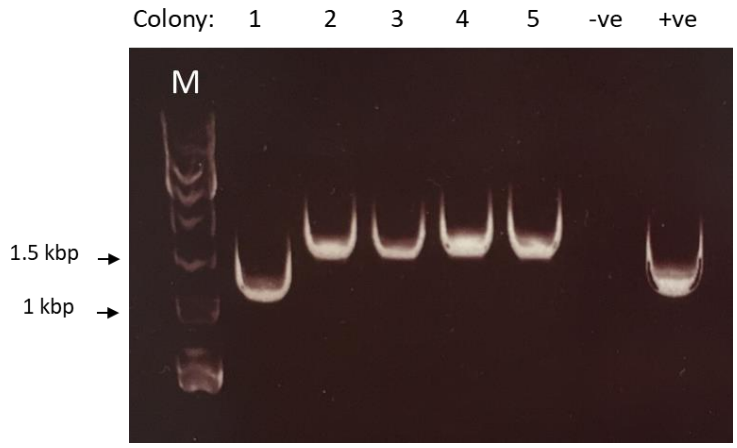


Figure 4.8 ***Colony PCR of DH10-beta cells following transformation with ETS1 ORF ligated into pcDNA3.1(+) vector***

DH10-beta cells were transformed with pcDNA3.1(+) vector containing ETS1 ORF and colony PCR performed to check for successful ligation and transformation. TBE gel containing 0.002% ethidium bromide electrophoresed for 30 minutes at 120 V. Negative control is identical PCR reaction without any colony added. Positive control is the PCR amplified ETS1 ORF from pDONR223_ETS1_WT (Addgene: 82118). Marker (M): NEB GeneRuler 1 kb used for DNA Ladder.

To determine an adequate dose of ETS1 vector for overexpression of ETS1 in hAFs, an initial small-scale experiment was conducted whereby hAFs, cultured in 6-well plates, where treated in technical duplicates with either 100, 200 or 400 ng of pcDNA3.1(+)_ETS1_WT or with empty pcDNA3.1(+) as a control. RNA and protein was harvested 48 hours post transfection and qPCR and immunoblotting were used as experimental readouts (48 hours selected to match my ETS1 knockdown experiments). QPCR showed an increase in ETS1 mRNA transcript abundance in all wells transfected with pcDNA3.1(+)_ETS1_WT (*figure 4.9A*) and several genes were then queried as potential reference genes for normalisation. ACTA2, FLNA GAPDH showed high variability between treatment groups, however RS18 remained stable and was therefore nominated to be used as a housekeeping gene (*figure 4.9A*). After normalisation to RS18, ETS1 mRNA change was recorded to surpass 1000-fold at the lowest dose of pcDNA3.1(+)_ETS1_WT expression vector (*figure 4.9B*).

In addition, ETS1 protein abundance in all treatment conditions was shown to increase by immunoblotting while GAPDH remained constant and therefore could serve as an adequate loading control (*figure 4.9C*). In order to demonstrate that control hAFs also expressed endogenous ETS1, *figure 4.9C* also includes a micrograph of the X-ray film exposed for a longer time thereby revealing ETS1 in the control lane. It was noted that transfection of hAFs with ETS1 expression vector resulted in the detection of two bands, one of identical molecular weight to endogenous ETS1 (also seen in *figure 4.12* below), and a second heavier band. I queried the literature for studies that had used the same antibody (Cell Signalling: 14069), and depending on the cell line used, some studies showed immunoblots with a single ETS1 band (Sobral et al. 2020; Zhao et al. 2021), while others showed two bands similar to my results but did not comment on what the heavier band was (Chen et al. 2016; Lei et al. 2019). ETS1 activity has been shown to be inhibited by phosphorylation at residues Ser282 and Ser285 (Fleischman et al. 1995, Vetter et al. 2005), as well as by acetylation at lysine residues Lys8 and Lys18 (Czuwara-Ladykowska et al. 2002; Chen et al. 2017). These studies used anti-ETS1 antibodies that detected these specific post-translational modifications, and their results indicated that the molecular weight of phosphorylated and acetylated ETS1 were similar to the molecular weight of the heavier band in my experiments.

In light of these previous findings, there is good reason to consider that the heavier band in my ETS1 overexpression experiments is either phosphorylated or acetylated ETS1. Consequently, further immunoblotting experiments in hAFs using antibodies specific to ETS1 post-translational modifications would be needed to discern specifically what the modification is. The ETS1 overexpression experiment posed another query as to why the heavier ETS1 band was only detected in hAFs treated with ETS1 vector, but not in the control population of hAFs. One explanation may be that the supra-physiological expression of ETS1 in hAFs treated with ETS1 vector initiated endogenous feedback loops resulting in phosphorylation and/or acetylation of ETS1. Regardless of the identity of the post-translational modification in the heavier band in the ETS1 immunoblots, we chose to proceed with experiments because there was clear increased expression in the un-modified ETS1 protein which was the main objective at this stage of the project.

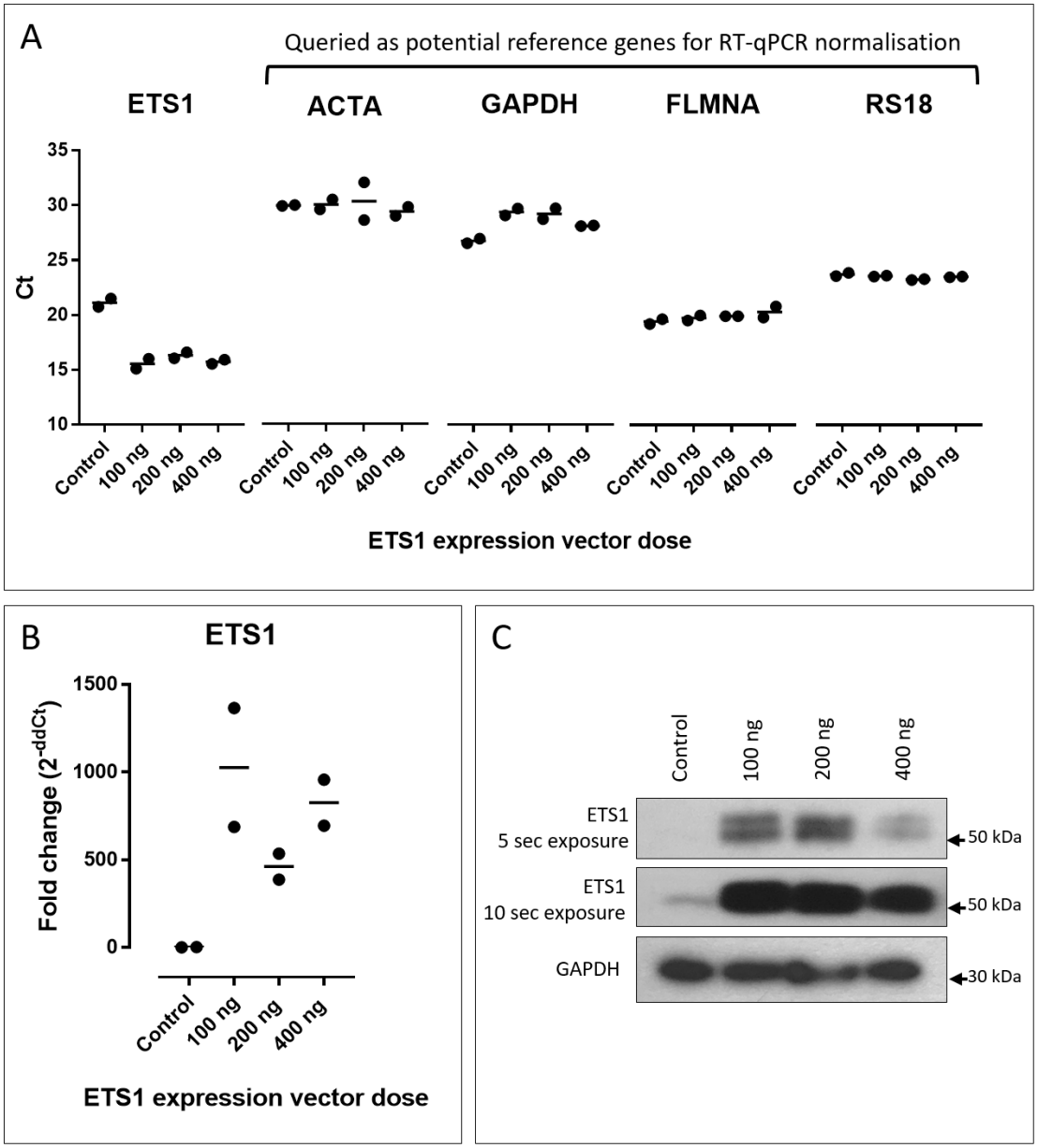


Figure 4.9 ***Dose experiment of pcDNA3.1(+)_ETS1_WT expression in hAF cells***

hAF total RNA and protein was extracted 48 hours after transfection using lipofectamine with 100 ng empty pcDNA3.1(+) (control) or increasing doses (100-400 ng) pcDNA3.1(+)_ETS_WT. The experiment was performed as technical duplicates and therefore no statistical comparisons were performed on the data.

A: graphs of raw qPCR Ct values for ETS1, ACTA2, GAPDH, FLNA and RS18.
B: ETS1 Ct values normalised to RS18 and presented as fold-change of control
C: western blot analysis of hAF protein extracts to measure ETS1 expression with GAPDH serving as a loading control.

After determining a dose and time-course of pcDNA3.1(+)_ETS1_WT expression vector that increased ETS1 mRNA and protein abundance, I repeated the experiment four times and investigated any effect on miR31 abundance. Results indicated a significant, albeit highly variable, increase in ETS1 mRNA abundance (*figure 4.10A*) and raw Ct values indicated an increase in miR31 abundance while RNU58A remained relatively stable and could therefore serve as a housekeeping gene in RT-qPCR (*figure 4.10B*). Following normalisation, miR31 showed a close to significant increase in the ETS1 vector treated hAF cells (*figure 4.10 B*).

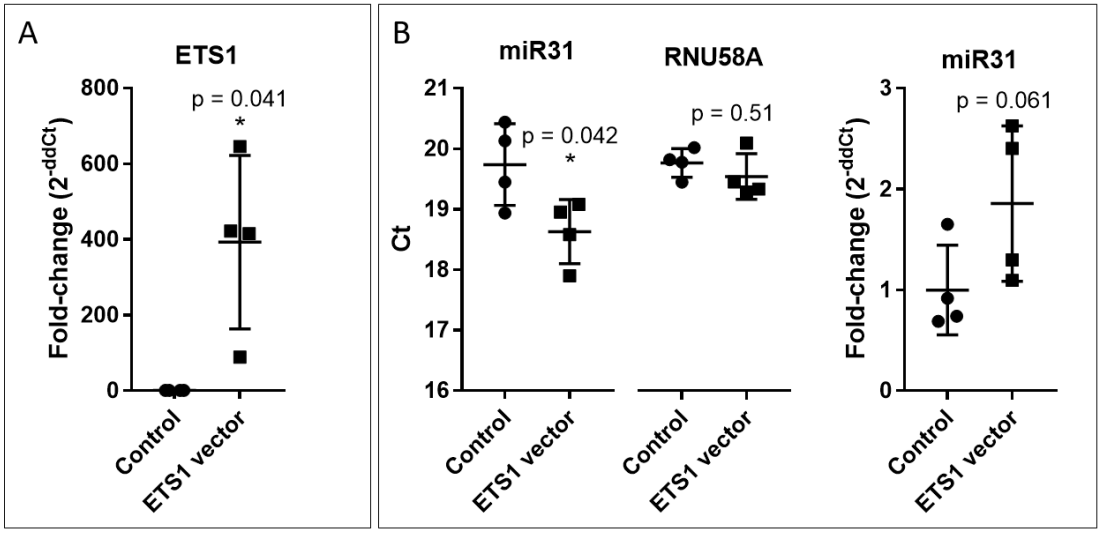


Figure 4.10 ***ETS1 enrichment alters miR31 abundance in hAFs***

hAF total RNA was extracted 48 hours after transfection using lipofectamine with 100 ng ETS1 expression vector or empty pcDNA3.1(+) as a control. The experiment was repeated four times and each data point is from one experimental repeat. Paired Student's *t*-test used to compare means as data was shown to be parametric with K-S statistical test. Error bars depict mean ± SD.

A: qPCR data of ETS1 expressed as fold-change of control using RS18 as housekeeping for normalisation.
B: raw Ct values for miR31 and RNU58A, as well as miR31 normalised to RNU58A and presented as fold-change of control.

Given the observed *increase* in miR31 abundance following transfection with ETS1 vector was incongruent with our expectation that miR31 would decrease, I queried whether ETS1 was locating to the nucleus where it could elicit effect on LOC554202 transcription. Immunofluorescence did show nuclear abundance however results indicated that only approximately 10% of all cells were highly abundant in ETS1 and that this was present in the nucleus (*figure 4.11*).

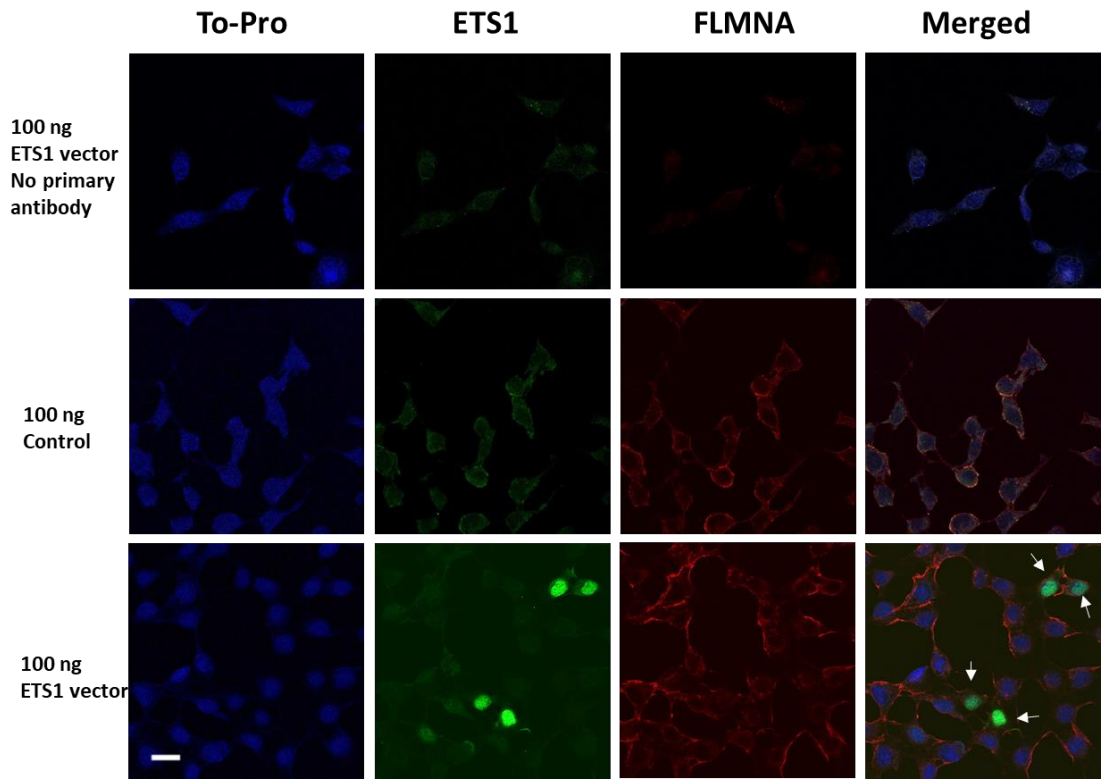


Figure 4.11 ***Visualising ETS1 in primary hAF cells by immunofluorescence***

Primary hAFs were transfected with 100 ng ETS1expression vector or 100 ng empty pcDNA3.1 vector (negative control) and fixed 48 hrs post transfection. To-Pro was used for nuclear staining (blue), ETS1 (green), filamin-A (FLMNA) for cytoskeleton (red). White arrows indicate nuclear ETS1. Performed on hAFs from donor 1. Scale bar is 50 µm.

The former experiments showed that the ETS1 expression vector was entering only a small portion of transfected hAFs, and in that small portion ETS1 was being expressed supra-physiologically such that when harvesting RNA from the total population of hAFs, ETS1 mRNA and miR31 abundance was significantly altered and highly variable. Further, miR31 expression was shown to increase, not decrease, following ETS1 expression. Following these observations I sought to bring the experimental system closer to an endogenous state and queried whether a lower dose of pcDNA3.1(+)_ETS1_WT would have a more predictable effect on miR31. Consequently, hAF cells were cultured in 6-well plates and in technical triplicates treated with 25, 50 and 100 ng pcDNA3.1(+)_ETS1_WT and harvested 48 hours later. Immunoblotting again showed increased ETS1 expression in all doses of pcDNA3.1(+)_ETS1_WT vector (*figure 4.12A*), and qPCR indicated a ~20-fold increase in ETS1 mRNA (normalised to RS18) at the lowest dose (25 ng) (*figure 4.12B*).

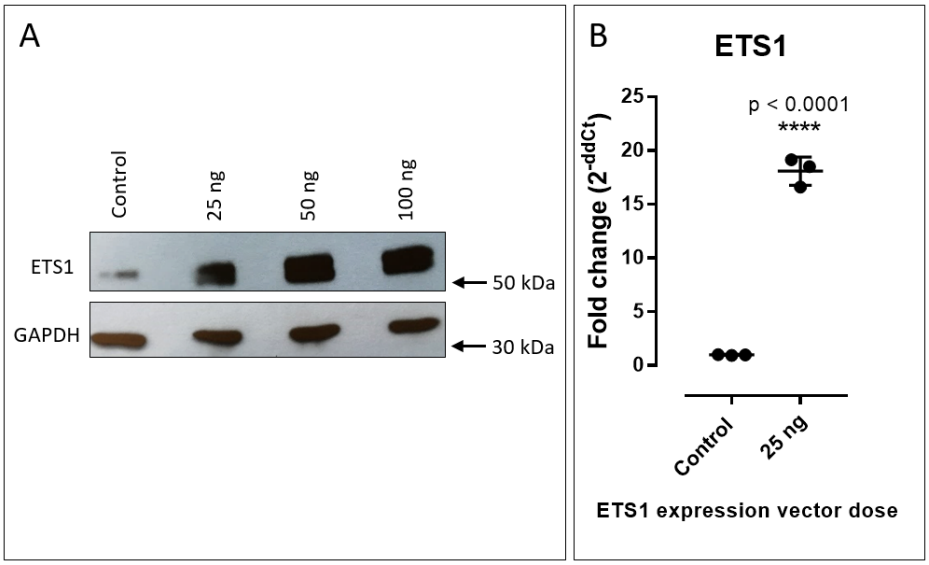


Figure 4.12 ***Trialling lower doses of pcDNA3.1(+)_ETS1_WT expression in hAF cells***

hAF total RNA and protein extracted 48 hours after transfection using lipofectamine with 25 ng empty pcDNA3.1(+) (control) or increasing doses (25-100 ng) pcDNA3.1(+)_ETS_WT.

A: Western blot analysis of hAF protein extracts measuring ETS1 expression with GAPDH serving as a loading control.
B: qPCR data showing ETS1 Ct values normalised to RS18 and presented as fold-change of control. Paired Student's *t*-test used to compare means as data was shown to be parametric with K-S statistical test. Error bars depict mean ± SD.

Although ETS1 expression was still reaching significantly higher than endogenous levels, it would be important to investigate further whether improving transfection efficiency with 25 ng pcDNA3.1(+)_ETS1_WT would show homogenous ETS1 expression in hAFs. In addition, it would be important to determine whether 25 ng, or indeed lower amounts, of ETS1 vector would lead to a predicted *decrease* in miR31 abundance. In the interest of time however, I could not pursue the ETS1 overexpression experiments further to answer these questions.

4.2 Mutagenesis of ETS1 binding sites at the LOC554202 promoter

Alongside the above ETS1 experiments in hAFs, I sought to quantify whether ETS1 affects LOC554202 transcription *in vitro* using a luciferase based assay in HEK293 cells. In this assay, HEK293 cells were transfected with a mutagenesis construct carrying the LOC554202 promoter with one of the ETS1 binding sites mutated to retard ETS1 binding, and cloned upstream of *Firefly* Luciferase gene cDNA in a reporter plasmid. HEK293 cells do not express endogenous ETS1, therefore to control ETS1 abundance, cells were co-transfected with ETS1 expression plasmid and luminescence readouts were taken to quantify the effect of ETS1 on LOC554202 expression. To first generate the mutagenesis constructs, the 1061 bp stretch of LOC554202 promoter was PCR amplified from human gDNA (Promega: G1521A) and cloned into a Firefly pGL4.10[*Luc2*] reporter plasmid. This construct was then used as a template in restriction-free cloning using eleven unique pairs of mutagenesis primers to target and mutate the ETS1 binding sites (A-K). The resultant eleven constructs were sequenced and mutations confirmed by comparison to WT (*figure 4.13*).

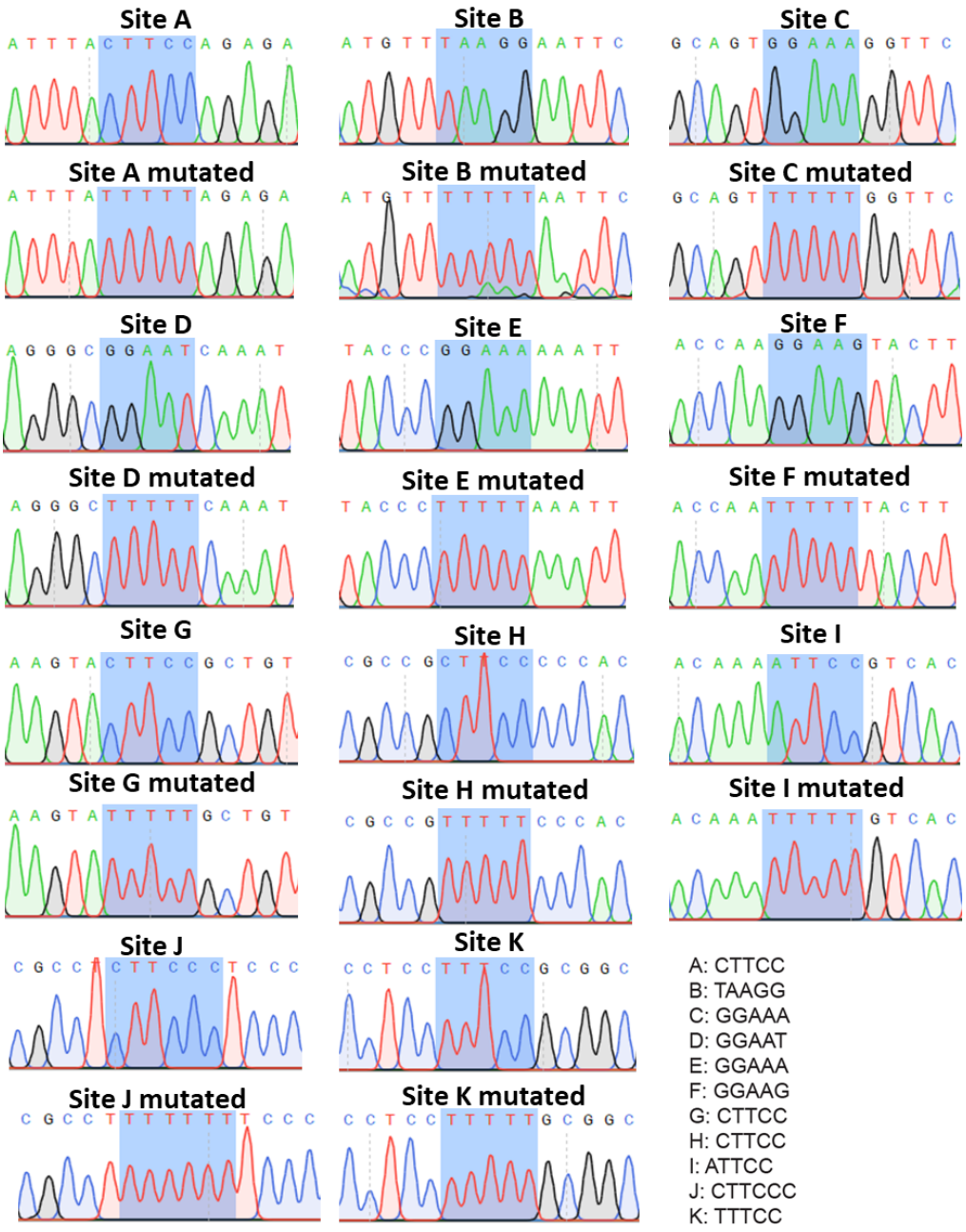


Figure 4.13 ***Mutagenesis constructs of WT LOC554202 promoter Firefly Luciferase vector***

The WT LOC554202 promoter *Firefly* Luciferase vector was used as template to generate eleven mutagenesis constructs that were validated by sequencing of both the forward and reverse strand using the XhoI and HindIII cloning primers respectively as sequencing primers. Results show eleven sites A-K on the forward strand in their original WT form and in mutant form.

For optimizing ETS1 expression in HEK293 cells, a dose and time-course experiment was conducted whereby, in technical duplicates, cells cultured in 12-well plates were lipofected with 100, 200 or 400 ng of pcDNA3.1(+)_ETS1_WT expression vector, or with 100 ng empty pcDNA3.1(+) vector as control. HEK293 cells were harvested either 24 or 48 hours post-transfection for mRNA and protein analysis. qPCR results showed an increase in ETS1 mRNA transcript abundance in all wells transfected with pcDNA3.1(+)_ETS1_WT (*figure 4.14A*) and RS18 mRNA abundance remained stable and could therefore be used to as a reference gene for normalisation in future qPCR analyses in HEK293 cells (*figure 4.14A*). After normalisation, ETS1 mRNA change was recorded to surpass 1000-fold at even the lowest dose of pcDNA3.1(+)_ETS1_WT expression vector (*figure 4.14B*). In addition, ETS1 protein abundance in all treatment conditions was shown to increase by immunoblotting. GAPDH remained constant and therefore could serve as an adequate loading control (*figure 4.14C*). As a consequence of this experiment, 100 ng of pcDNA3.1(+)_ETS1_WT expression vector with harvest at 48 hours was determined as suitable starting point for future ETS1 expression experiments in HEK293 cells. Forty-eight hour harvest was also used in my hAF cell experiments therefore this time-course was chosen for consistency between experimental models.

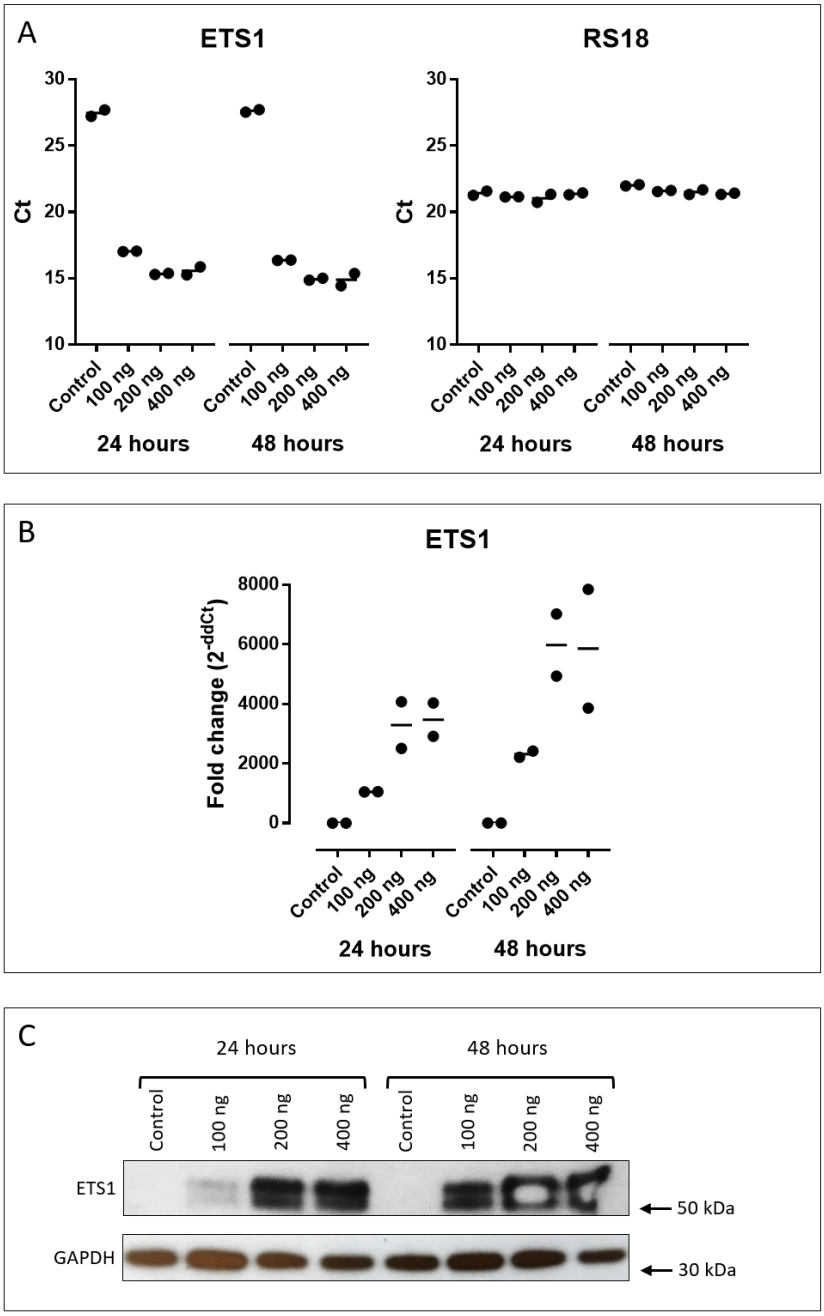


Figure 4.14 ***qPCR analysis and immunoblotting of time-course and dose experiment of ETS1 expression vector in HEK293 cells***

HEK293 total RNA and protein extracted 24 and 48 hours after transfection using lipofectamine with empty pcDNA3.1(+) (control) or increasing doses (100-400 ng) ETS1 expression vector.

A: raw ETS1 and RS18 qPCR Ct values at different doses of ETS1 expression vector 24 and 48 hour time points after transfection
B: ETS1 Ct values normalised to RS18 and presented as fold-change of control.
C: western blot analysis of HEK293 protein extracts to measure ETS1 expression with GAPDH serving as a loading control.

Having shown ETS1 mRNA and protein abundance in HEK293 cells I then asked whether ETS1 was indeed translocating to the nucleus where it could elicit transcriptional effect on our mutant LOC554202 vector constructs. For this I used immunofluorescence and confocal microscopy to visualize ETS1 location in HEK293 cells 48 hours following transfection with ETS1 expression vector. Image analysis did indicate a high abundance of ETS1 in the nucleus, however this was only observable in approximately 10% of the HEK293 cell population (*figure 4.15*). This experiment therefore indicated that our ETS1 expression vector was acting as intended but the transfection protocol could be optimised to improve efficiency and facilitate more homogeneous ETS1 expression.

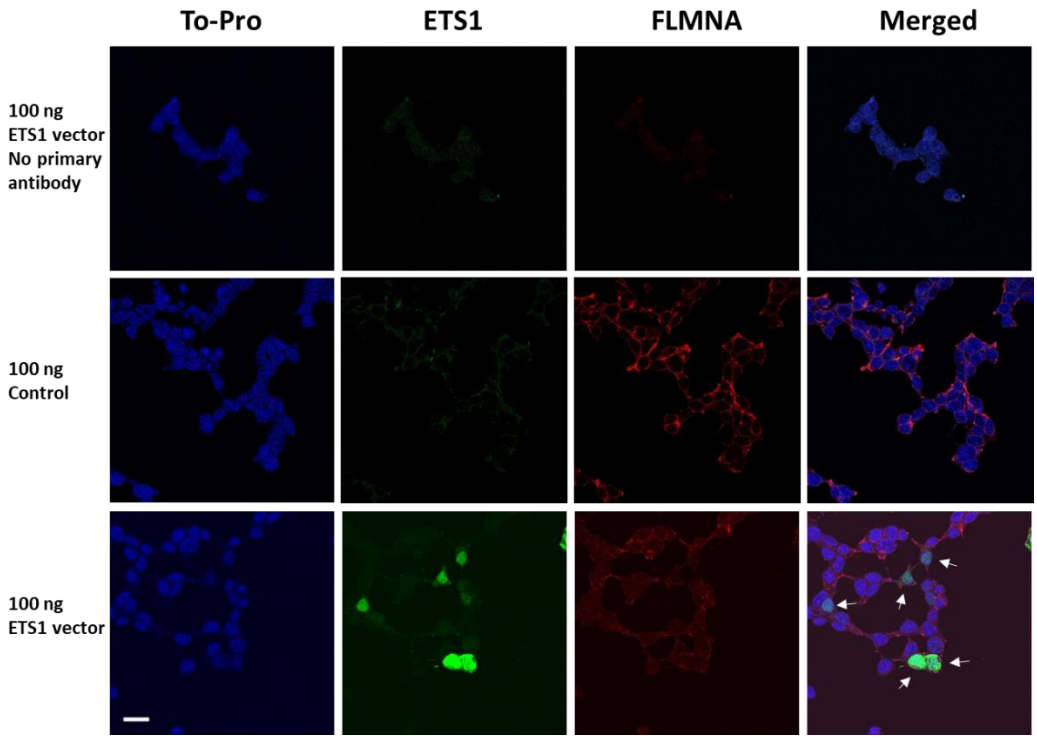


Figure 4.15 ***Visualising expression and localisation of ETS1 in HEK293 cells by immunofluorescence***

HEK293 cells were transfected with 100 ng pcDNA3.1_ETS1_WT expression vector or 100 ng empty pcDNA3.1 vector (negative control) and fixed in methanol-acetone (1:1) 48 hrs post transfection. To-Pro was used for nuclear staining (blue), ETS1 (green), FLMNA for cytoskeleton (red). White arrows indicate nuclear localisation of overexpressed ETS1. Scale bar is 50 µm.

Having demonstrated ETS1 nuclear expression in HEK293 cells following lipofection with 100 ng of ETS1 expression vector, I then queried whether ETS1 would affect the expression of WT LOC554202 promoter *Firefly* Luciferase vector. For this pilot experiment I chose not to include the mutagenesis constructs as I was simply testing whether the luciferase assay would work with the basic constructs used. HEK293 cells were cultured in a 24-well plate and lipofected with 100 ng ETS1 vector together with 50 ng of WT LOC554202 promoter *Firefly* Luciferase vector and 50 ng of *Renilla* Luciferase control vector which would act as an internal control to demonstrate equal transfection efficiency between wells. Luminescence readings were then taken following the addition of *Firefly* Luciferase substrate and then following addition of *Renilla* Luciferase substrate.

Results of this experiment showed that the presence of the WT LOC554202 promoter *Firefly* Luciferase vector caused an increase in luminescence and that the subsequent addition of ETS1 marginally lowered the observed increase (*figure 4.16 upper graph*). However given that luminosity derived from the *Renilla* Luciferase vector was negligible compared to background levels detected in vehicle control treated cells (*figure 4.16 lower graph*), this experiment needs further optimising to demonstrate equal transfection efficiency between control and treatment groups. One approach could be to first increase the amount of *Renilla* Luciferase vector used and hence the luminescence signal as this pilot experiment showed near-zero luminescence with the amount of vector used.

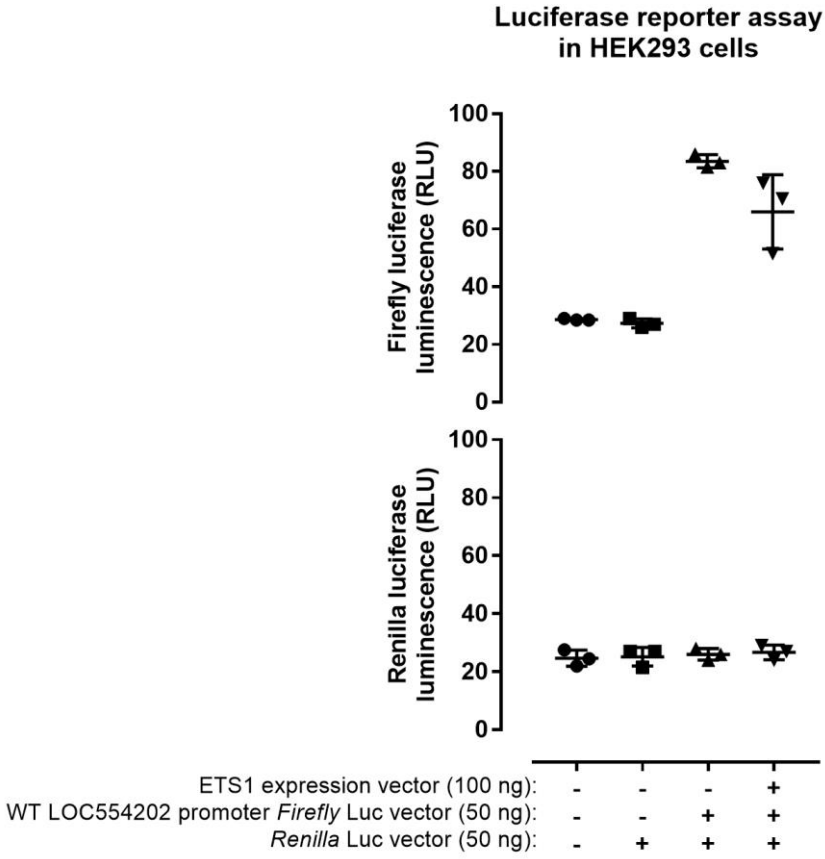


Figure 4.16 ***Luciferase assay in HEK293 cells***

HEK293 cells cultured in a 24-well plate were lipofected using Lipofectamine with varying combinations of 100 ng of ETS1 expression vector, 50 ng of WT LOC554202 promoter *Firefly* Luciferase vector and 50 ng *Renilla* Luciferase vector used as transfection control. Vehicle control was also used (far left group) that was subject to the same lipofection protocol but without the presence of vectors. The experiment was performed as triplicates in a single 96 well plate, statistical tests were not performed as these data were used to simply demonstrate the negligible luminescence derived from the *Renilla* Luciferase construct. RLU stands for Relative Light Units and is the raw output from the plate reader when luminescence was measured following the sequential addition of *Firefly* and *Renilla* luciferase substrate.

4.3 Discussion and plans for future experiments

In this chapter I conducted ETS1 knockdown and overexpression studies in hAFs to understand the relationship between ETS1 abundance and miR31 expression in these cells. I also commenced experiments to study the effects of mutating ETS1 binding sites at the LOC554202 promoter on transcription of this allele by using a luciferase based reporter assay in HEK293 cells. Like the previous chapter, closure of the University buildings during the pandemic meant that experiments were cut short and further work is required to address the aims of this part of the project. This section outlines the limitations of my results and provides details of future work that ideally would have been done to complete this chapter.

I first depleted ETS1 in hAFs and repeated this experiment on six different days, and following ETS1 depletion there was a consistent and significant increase in miR31 when looking at the raw qPCR Ct values however this significant effect was lost after normalising miR31 signal to RNU58A. Upon inspection of the raw RNU58A data from this experiment, it was noticed that there was high variability primarily within the group treated with anit-ETS1 LNA (*see figure 4.6*). Given that experiments using hAFs in the previous chapter showed tight error bars for RNU58A (*see figure 3.5*) indicating that technical variability for the qPCR methodology was low in my hands, the extent of variability of RNU58A observed in the ETS1 knockdown experiment was likely a real biological result of depleting ETS1 in hAFs. Therefore this experiment would have been improved by trialling differing microRNA housekeeping genes to which miR31 data could then be normalised to in the ETS1 experiments. For example, small nuclear RNA U6 (snRNA U6) is a commonly used housekeeping gene for microRNA studies and could be measured in hAFs, alternative options that have been trialled in human heart tissue and could be tested in hAFs are small nucleolar RNA C/D Box 48 (SNORD48), as well as ribosomal subunit 5 (5S) (Mase et al. 2017). Let-7i and miR16 are also frequently used and could also be measured in hAFs (Wang et al. 2018; Donati et al. 2019).

Following ETS1 knockdown experiments, I then went on to look at the effects of increasing ETS1 abundance on miR31 expression by transfecting hAFs with ETS1 expression vector using lipofectamine. Immunoblotting and qPCR data showed that lipofection with 100 ng of ETS1 vector caused supra-physiological abundance of ETS1 in hAFs (*see figure 4.9*), however when immunofluorescence staining was performed it was noted that only a small portion of transfected hAFs (~10%) showed ETS1 expression (*see figure 4.11*). Together these experiments suggested that ETS1 vector was working as intended to increase ETS1 abundance however optimisation was needed to bring ETS1 expression closer to endogenous levels as well as achieve more homogenous expression of ETS1 across the entire hAF population. One approach could be to trial increased quantities of lipofectamine reagent to improve transfection, alternative approaches could test different plating densities of hAFs for optimal confluency at the time of transfection. In addition, different methods of transfection could be tried, for instance nucleofection. Our lab has access to a Nucleofector 4D-unit (Lonza) and I commenced ETS1 transfection experiments using this machine however electroporation of hAFs often resulted in significant cell death and failure of hAFs to re-attach to tissue culture plates (data not recorded). Future experiments could try increasing the cell number used to offset cell death or try coating tissue culture plates with laminin to improve hAF attachment following nucleofection.

Alongside ETS1 experiments in hAFs, I sought to use a luciferase reporter assay in HEK293 to study the effects of mutating ETS1 binding sites at the LOC554202 promoter. An initial experiment was conducted whereby cells were transfected with WT LOC554202 *Firefly* Luciferase construct and this showed good luminescence following addition of luciferase substrate (*figure 4.16*). In the experimental design cells were co-transfected with *Renilla* luciferase vector (pGL4.74) as an internal control to normalise for any variation in transfection efficiency between wells and treatment groups. However results indicated that luminosity derived from the *Renilla* Luciferase vector was negligible compared to background levels and consequently this experiment needs optimising. Different approaches could include trialling increased amounts of lipofectamine and/or *Renilla* Luciferase vector, other ways could be to trial nucleofection to improve transfection efficiency, or trial different seeding densities for optimal confluency at transfection.

Given that *Renilla* luminescence signal was essentially zero with 30 ng of vector used (the same amount used as *Firefly* vector), it is more plausible that the issue was with expression of the vector within HEK293 cells and not with delivery of the construct, as one would expect at least some luminescence even if the transfection efficiency was low. The *Renilla* luciferase vector (pGL4.74) used contains the promoter sequence of the human simplex virus thymidine kinase gene (HSV-TK) to drive transcription of *Renilla* luciferase. The HSV-TK promoter has shown variable activity between different cell lines (Shifera et al. 2009), therefore one approach could be to trial other commercially available *Renilla* luciferase vector constructs that utilise the human cytomegalovirus (CMV) promoter (Promega: E2261) or the simian vacuolating virus 40 (SV40) promoter (Promega: E2231) which may prove to work more effectively in HEK293 cells. Alternatively the luciferase reporter assay could be trialled in other readily available cell lines, for instance HeLa cells.

Chapter 5

Results: Validation of a novel transgenic mouse and consequences of in vivo up-regulation of miR31 in fibroblasts

Having demonstrated a role of miR31 in facilitating pro-fibrotic gene expression changes, as well as the potential to drive pro-fibrotic phenotype in hAFs *in vitro*, in this part of the project I sought to investigate the function of miR31 in cardiac fibroblasts *in vivo* in mice. This approach was taken to verify the function of miR31 in more complex biological systems as well as uncover any new effects of changes in miR31 thereby revealing potential downstream biological targets for therapeutic manipulation. This approach would also provide us with a novel model in which miR31 can by studied in states known to induce cardiac fibrosis including aortic banding, or infusion of AngII or isoproterenol known to induce hypertension and cardiac fibrosis. Further, fibrosis associated with atrial fibrillation specifically can be studied by inducing AF with trans-oesophageal burst pacing in these mice. In addition, through this work we could generate a genetically modified mouse capable of overexpressing miR31, which would be of interest and use to the wider scientific community studying microRNAs *in vivo.*

To develop our experimental system such that miR31 expression could be controlled temporally, be ligand-dependant, and sustained in fibroblasts in mice we utilised the *Col1a2-Cre-ER(T)* transgenic mouse strain (Zheng et al. 2002) in combination with our novel *pre-miR31-Inv* transgenic mouse that was developed by the Davies lab, Oxford (WTCHG, Transgenic core facility). We chose to use the *Col1a2-Cre-ER(T)* transgenic mouse strain, as opposed to other promoter Cre driver mouse strains, based on two main considerations. Firstly, all available fibroblast Cre drivers are non-specific to fibroblast subtype and therefore show varied expression patterns across different tissues and stages of development which meant that, regardless of the mouse model used, the caveat would remain that systemic, and not cardiac specific, fibroblast Cre expression would likely occur. With this limitation in mind, the second consideration was to find a Cre driver mouse strain that had previously shown robust and reproducible expression in healthy adult mouse heart such that it may be used to target loxP flanked alleles.

We considered the periostin Cre drivers, namely *Postn-CreERT2* (Kaur et al. 2016) and *Postn-MerCreMer* (Kanisicak et al. 2016), however these were quickly excluded as periostin shows high abundance in embryonic cardiac fibroblasts, no expression in healthy adult cardiac fibroblasts, and then high expression in cardiac fibroblasts that have differentiated into myofibroblasts (Conway et al. 2008, Landry et al. 2017). Therefore periostin driver Cre models are well suited for targeting loxP flanked genes after the initiation of some heart disease or injury when cardiac fibroblasts are becoming activated, but are not well suited to targeting loxP flanked genes in healthy mouse hearts before the onset of disease. Given that we were interested in whether the induction of miR31 expression in the mouse heart would initiate and drive cardiac fibrosis, we chose not to use the periostin Cre driver models, and then considered other models.

Alternative Cre drivers were considered including *Tcf21-Cre* (Acharya et al. 2011), however previous reports have demonstrated that TCF21 expression shows high abundance in developing cardiac fibroblasts, low levels in normal adult cardiac fibroblasts, and variable expression after initiating cardiac disease (Acharya et al. 2012). For instance, TCF21 is increased in fibroblasts in hypertensive heart disease mouse models (Braitsch et al. 2013), but reduced in cardiac fibroblasts during myofibroblast differentiation following induced cardiac infarction (Fu et al. 2018). Consequently, the *Tcf21-Cre* mouse line may have been a viable option to target loxP flanked miR31 in normal adult cardiac fibroblasts, but we were unsure how it would behave in certain experiments where we may have had to induce heart disease and fibrosis, for instance with atrial burst pacing to cause arrhythmia. Consequently the *Col1a2-Cre-ER(T)* mouse line was then considered.

The *Col1a2-Cre-ER(T)* mice carry a chimeric gene, encoding the *Cre-ER(T)* fusion protein, under the control of a regulatory element from the endogenous mouse *Col1a2* gene previously demonstrated to be constitutively active and specific to fibroblasts in mouse (Armendariz-Borunda et al. 1992, Zheng et al. 2002). Several studies that have used the *Col1a2-Cre-ER(T)* mouse line have reported consistent expression of Cre with successful targeting of loxP flanked genes in both disease models and healthy mouse hearts. For example, *Col1a2-Cre-ER(T)* was used to knockout floxed p38a in healthy mouse hearts which then went on to show improved protection against myocardial injury with isoproterenol-induced heart failure (Li et al. 2017). Furthermore, *Col1a2-*

Cre-ER(T) mice were used to target and delete floxed interleukin-11 which resulted in reduced cardiac fibrosis compared to controls in an AngII model of hypertension (Schafer et al. 2017). Conversely, another study used *Col1a2-Cre-ER(T)* mice for successful knockout of floxed G protein–coupled receptor kinase 2 (GRK2) *after* the induction of ischemia/reperfusion injury, and demonstrated that GRK2 enhances fibrosis in this disease model (Woodall et al. 2016). Together these studies suggested that the *Col1a2-Cre-ER(T)* mouse line could be used in a range of experimental contexts and that Cre activity could be induced both before and after the use of different experimental interventions known to induce fibrosis. This meant that choosing *Col1a2-Cre-ER(T)* in our experiments would give us flexibility with when we desired to activate Cre and induce miR31 expression. Furthermore, given the novelty of the *pre-miR31-Inv* transgenic mouse, and that the inversion switch utilised (mechanism described in paragraphs below) had never been used before *in vivo*, it was important to choose a Cre driver line with robust and reproducible Cre expression in adult mouse heart.

Cre-ER(T) acts as a tamoxifen-dependent DNA recombinase targeting specific LoxP consensus sequences and this feature was used to induce miR31 overexpression in the novel *pre-miR31-Inv* transgenic mouse strain. Under normal conditions, the *pre-miR31-Inv* transgene sits inverted (off position) and downstream of a ubiquitous *CAG* promoter within the modified *Gt*(*ROSA26*)*Sor* locus. Recombinase activity by *Cre* results in flipping of the inverted *pre-miR31-Inv* sequence, thereby linking the transgene to the *CAG* promoter (*figure 5.1*).

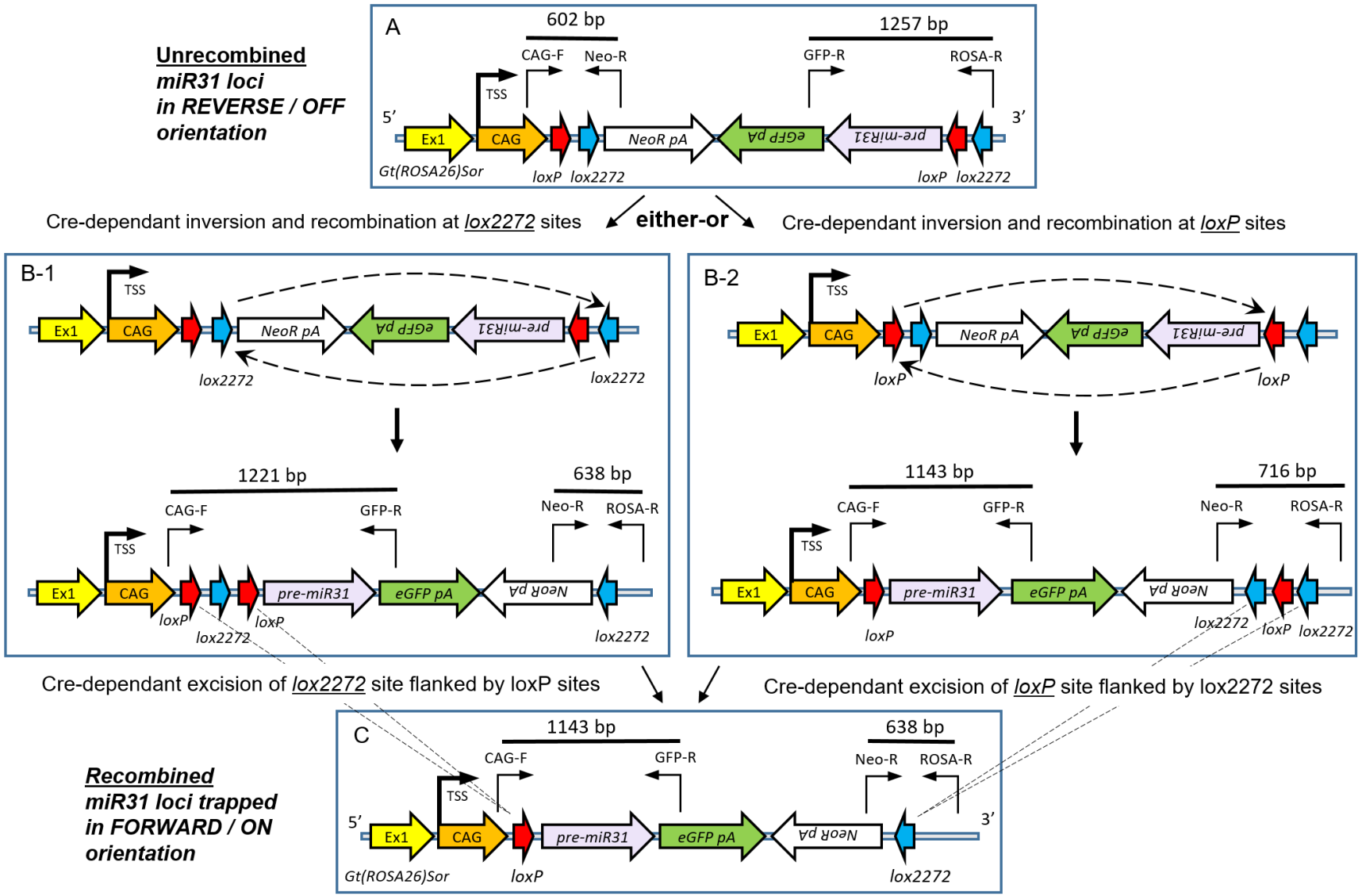

Unrecombined
miR31 loci
in REVERSE / OFF
orientation
A
602 bp
1257 bp
CAG-F
Neo-R
GFP-R
ROSA-R
TSS
5'
3'
Ex1
CAG
NeoR pA
eGFP pA
pre-miR31
Gt(ROSA26)Sor
loxP
lox2272
Cre-dependant inversion and recombination at lox2272 sites
either-or
Cre-dependant inversion and recombination at loxP sites
B-1
B-2
1221 bp
638 bp
1143 bp
716 bp
pre-miR31
eGFP pA
NeoR pA
Cre-dependant excision of lox2272 site flanked by loxP sites
Cre-dependant excision of loxP site flanked by lox2272 sites
C
Recombined
miR31 loci trapped
in FORWARD / ON
orientation

Figure 5.1 ***Principle of Cre-recombinase dependant inversion switch at the targeted Gt(ROSA26)Sor locus***

A: in the un-recombined state, the transgenic construct containing the *pre-miR31* (light pink arrow), *eGFP* with poly-A tail (green) and now redundant neomycin-resistance gene with poly-A tail (white), sits immediately downstream of the *Gt(ROSA26)Sor* Exon-1 (yellow) and *CAG* promoter (orange) and is flanked by two pairs of opposing loxP variants: *lox2272* (blue) and *loxP* (red).
B-1 and B-2: recombination occurs only between opposing pairs of variant loxP sites of the *same* type (loxP or lox2272) and causes the intervening sequence to be inverted and placed in the forward orientation with respect to the *CAG* promoter. Given there are two pairs of opposing loxP variants flanking the entire construct, *lox2272* (blue) and *loxP* (red), inversion can occur in either of two possible ways (B-1 or B-2).
C: following inversion, *Cre* excises the short sequence separating parallel loxP sites (again only of the same type) thereby trapping the entire construct in the forward orientation allowing for transcription.

Once the *pre-miR31-Inv* transgene is linked to the *CAG* promoter, transcription occurs to generate a primary mRNA transcript encoding *pre-miR31* and *eGFP*. Subsequent processing by the spliceosome and Drosha generates *pre-miR31* and *eGFP* mRNA. *Pre-miR31* is processed further by dicer to generate mature miR31 for functional studies while *eGFP* is used as a visual marker for transgene expression (*figure 5.2*).

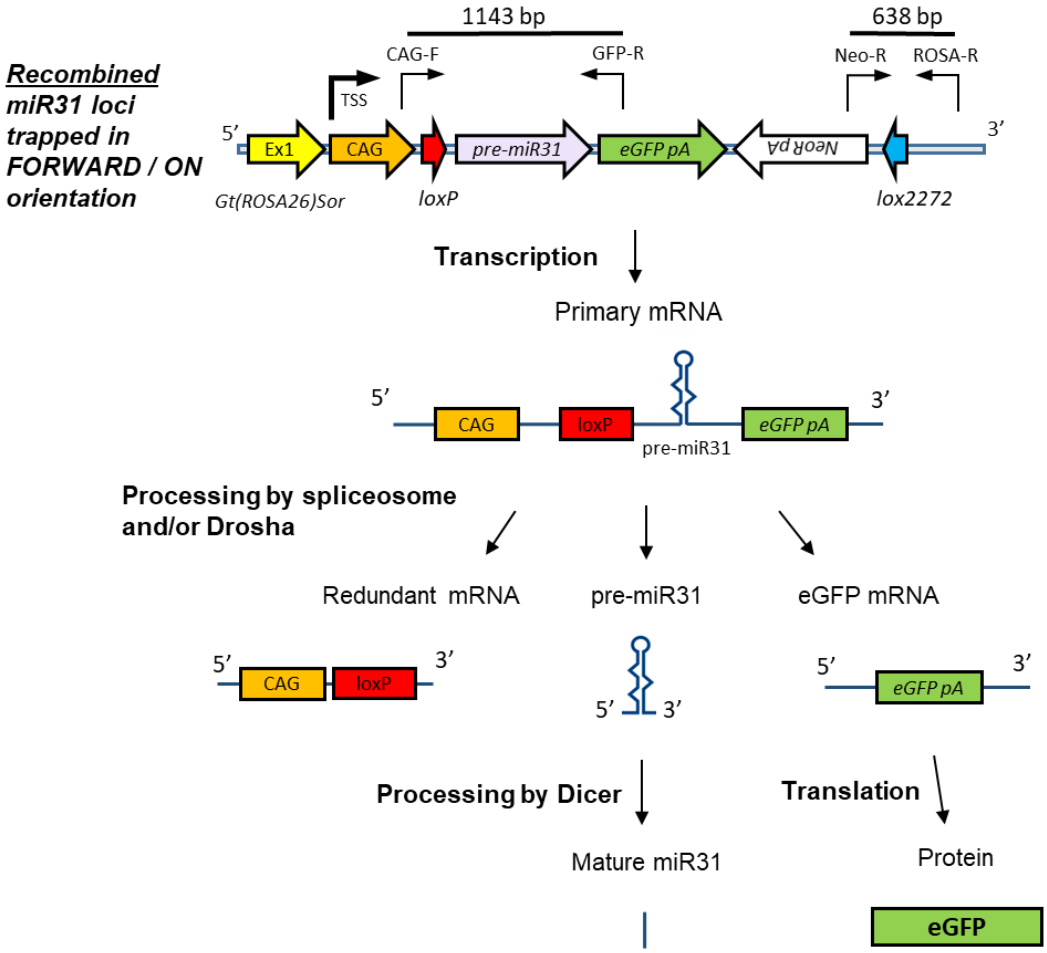


Figure 5.2 ***Expression of miR31 following Cre-dependant inversion at the targeted Gt(ROSA26)Sor locus***

In the recombined state the transgenic construct containing *pre-miR31* sits immediately downstream of the *CAG* promoter and transcription generates primary mRNA containing now redundant *CAG* and loxP, pre-miR31 and *eGFP* sequences. Processing by the spliceosome and/or Drosha liberates *eGFP* mRNA as well *as pre-miR31* which is processed further by Dicer to generate mature miR31.

5.1 *Col1a2-Cre-ER(T)* mice breeding strategies produce litters with Mendelian inheritance of transgene

The *Col1a2-Cre-ER(T)* transgenic mouse strain was generated by pro-nuclear injection which results in random integration of the transgenic allele by homologous recombination at varying copy number (Palmiter et al. 1986). Therefore *Col1a2-Cre-ER(T)* mouse colonies were maintained by crossing *hemizygous* with wild type mouse strains to control for Cre gene inheritance and expression level. Litter genotyping was performed as a multiplex with primers amplifying a 186 bp segment of the *Col1a2-Cre-ER(T)* transgene (*figure 5.3A*) as well as primers amplifying 314 bp of the third intron of the *Hprt1* gene as positive control (*figure 5.3B*). Breeding *Col1a2-Cre-ER(T)* hemizygous with wild types showed expected genotyping band sizes (*figure 5.3C*). In addition, I checked whether the inheritance of the *Col1a2-Cre-ER(T)* transgene was deviating from the expected Mendelian inheritance frequency by performing a Chi-square test. Results indicated that colony maintenance of the *Col1a2-Cre-ER(T)* transgenic mouse line did not significantly deviate ($p = 0.09$) from the expected 50:50 inheritance ratio (*table 5.1*).

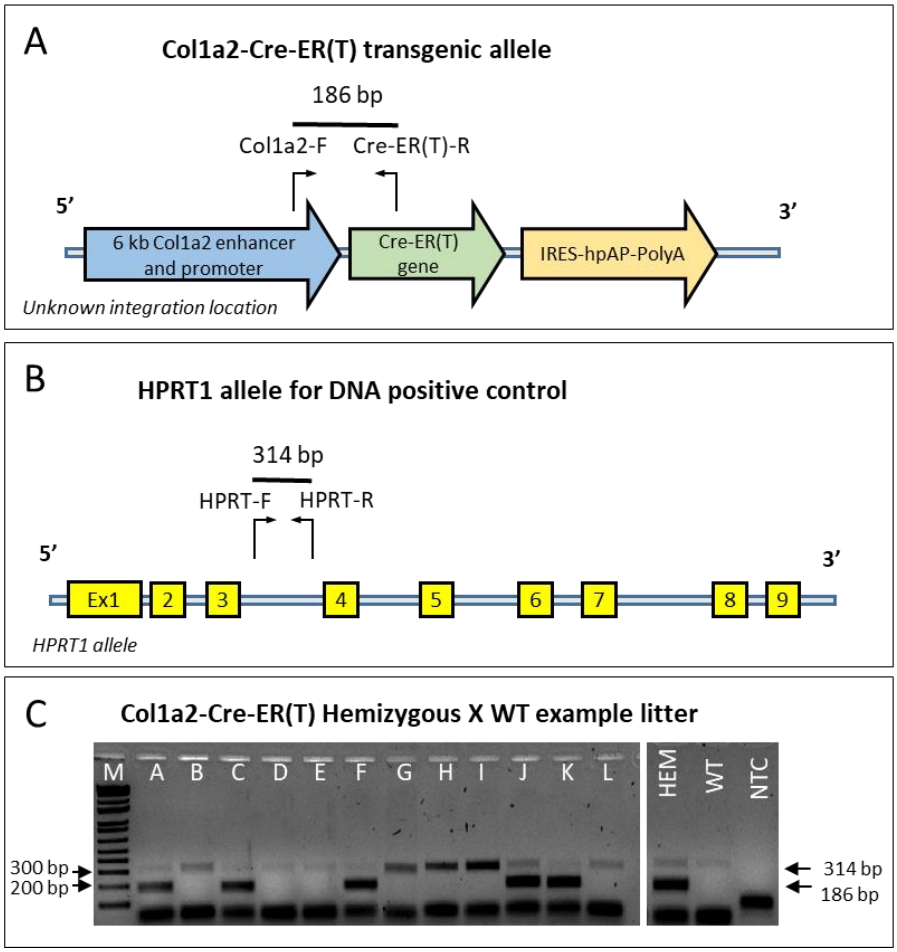


Figure 5.3 ***Col1A2-Cre-ER(T) example litter genotyping***

A: schematic of the randomly integrated *Col1a2-Cre-ER(T)* transgene showing forward and reverse primers (Col1a2-F and Cre-ER(T)-R) binding in the chimeric construct of the *Col1a2* promoter (pale blue) and adjacent *Cre-ER(T)* gene (pale green). The construct also includes the now-superfluous viral IRES sequence linked to the human placental alkaline-phosphatase reporter gene (hpAP) with PolyA site (pale yellow).
B: schematic of primer binding sites (HPRT-F and HPRT-R) specific to the 3rd intron of the *Hprt1* gene used as an internal positive control to validate presence of genomic DNA. Yellow boxes depict *Hprt1* gene exons.
C: littermate DNA (pups A-L) PCR amplified and electrophoresed in 1.5% agarose TBE gel containing 0.002% ethidium bromide for 35 minutes at 130 V. PCR was performed as multiplex with primers amplifying both the *Hprt1* allele (314 bp) and the *Col1A2-Cre-ER(T)* transgenic allele (186 bp). Hemizygous (HEM), wild type (WT) and no-template (NTC) controls were used. Invitorgen 1kb plus used for marker (M).

Table 5.1 ***Col1a2-Cre-ER(T) colony maintenance genotyping tally***

Breeding strategy	Breeding cages	Litters	Pup genotypes		Pup total	Chi-squared P
			WT	Cre		
Cre(Hemi) X WT	15	41	141 (44.7%)	114 (55.3%)	255	0.09

5.1.1 Optimisation of genotyping suggests GelRed retards DNA migration during electrophoresis

Genotyping *Col1a2-Cre-ER(T)* mice at the transgenic locus, and *Hprt1* allele as a positive control, initially showed issues whereby electrophoresed PCR products, expected to be the same size, ran at uneven sizes (*figure 5.4A*). I queried whether DNA migration was being retarded by the amount of DNA added to PCR reactions, the percentage of agarose in the agarose TBE gel, or the length of time the products were electrophoresed. These variables were tested by performing a series of PCRs with increasing amounts of template DNA that were then electrophoresed for 35 and 70 minutes on either 2 or 3% agarose TBE gels (*figure 5.4B*). Results indicated that band migration was affected by DNA amounts in higher, but not lower, percentage agarose gels and that time under electrophoresis did not account for retarding band migration. Therefore, low percentage agarose gels (1.5-2%) were used in future genotyping. To optimise further, I checked whether the use of GelRed would affect band migration when added to the agarose solution before pouring and setting, or when used to stain gels post electrophoresis (*figure 5.4C*). Gel images indicated GelRed was significantly altering band migration and therefore, in addition, all future genotyping was done using ethidium bromide which proved to resolve the issue of uneven band migration (*figure 5.4D).*

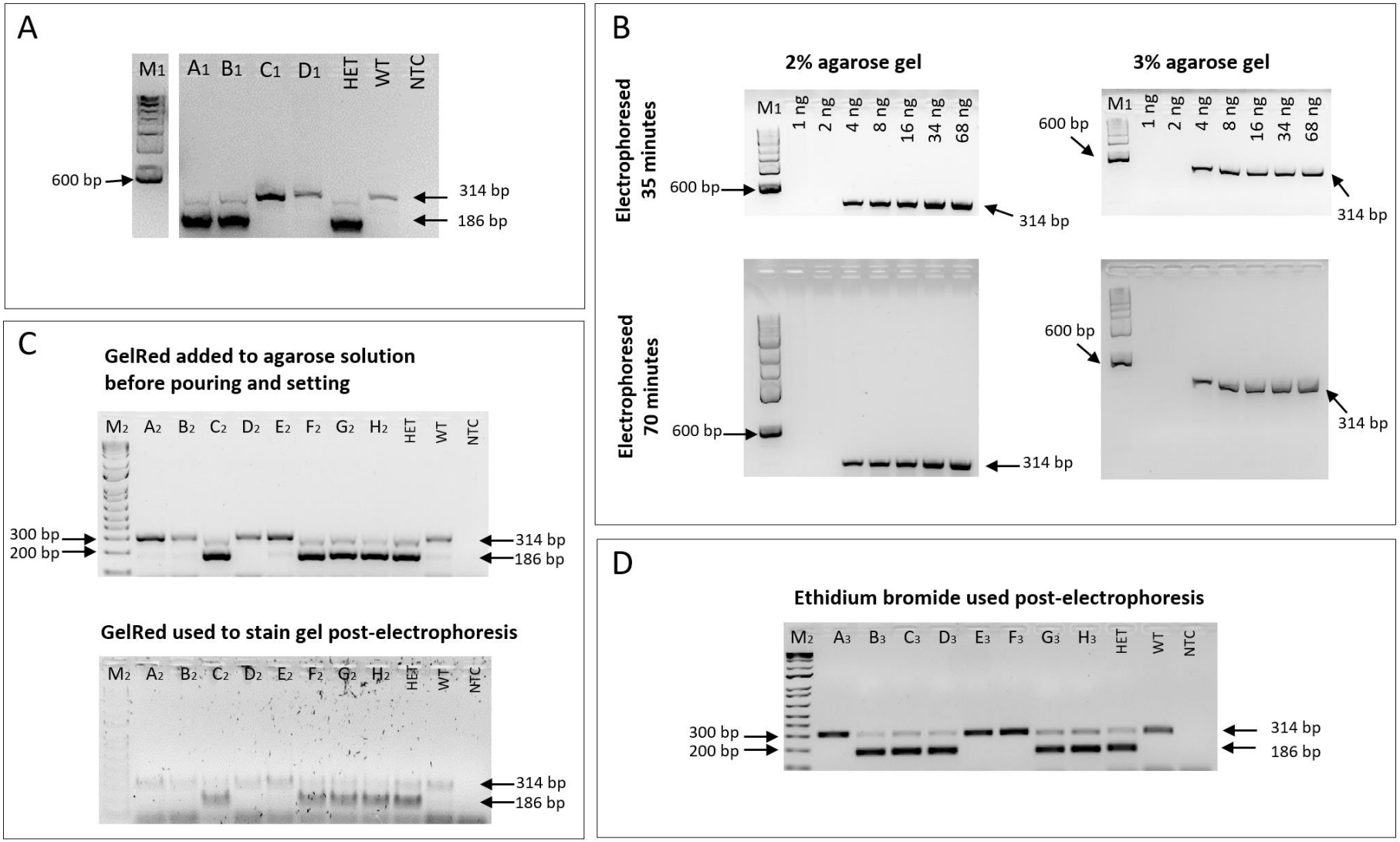

A
M1
A1 B1 C1 D1 HET WT NTC
600 bp
314 bp
186 bp
B
2% agarose gel
3% agarose gel
Electrophoresed 35 minutes
Electrophoresed 70 minutes
M1
1 ng 2 ng 4 ng 8 ng 16 ng 34 ng 68 ng
600 bp
314 bp
C
GelRed added to agarose solution before pouring and setting
M2 A2 B2 C2 D2 E2 F2 G2 H2 HET WT NTC
300 bp
200 bp
314 bp
186 bp
GelRed used to stain gel post-electrophoresis
D
Ethidium bromide used post-electrophoresis
M2 A3 B3 C3 D3 E3 F3 G3 H3 HET WT NTC

Figure 5.4 ***GelRed retards DNA migration in genotyping gels***

A: bands running at uneven sizes in an example litter (A1 - D1) were noticed when using 1.5% Agarose TBE gel run at 130 V for 35 minutes, with GelRed added to the agarose solution before pouring and setting. Heterozygous (HET), wild type (WT) and no-template (NTC) controls were used in all gels. Bioline HyperLadder 1 kb marker (M1) was bunching and predominately uninterpretable.
B: increasing amounts of DNA added to PCR reaction were loaded in two separate gels of different agarose concentration: 2% agarose TBE (left pane images) and 3% agarose TBE (right pane images), and images were taken after 35 minutes (top images) and 70 minutes (bottom images) of electrophoresis. Ran in TBE buffer, 130 Volts with GelRed added to the agarose solution before pouring and setting.
C: equal amounts (20 ng) of littermate DNA (A2 - H2) electrophoresed in 1.5% agarose TBE gel for 35 mins at 130 V. GelRed added to the agarose solution before pouring and setting (upper gel) or used to stain gel post-electrophoresis (lower gel). Invitrogen 1kb plus used for marker (M2).
D: equal amounts (20 ng) of littermate DNA (A3- H3) electrophoresed in 1.5% agarose TBE gel containing 0.002% ethidium bromide for 35 mins at 130 V.

5.2 *Pre-miR31-Inv* mice breeding strategies produce litters with Mendelian inheritance of transgene

Pre-miR31-Inv transgenic mice were generated using PhiC31 integrase that catalyses unidirectional recombination between specific sequences and is used for insertion of transgenic constructs into *defined* genomic locations (Dolatshad et al. 2015). Consequently, *pre-miR31-Inv* heterozygous mice carry a single copy of the *pre-miR31-Inv* transgene at the *Gt(ROSA26)Sor* locus which has benefits including that miR31 dosage is more predictable and closer to endogenous levels, and that dosage can be doubled in homozygous animals if needed. In addition, this approach avoids the random insertion of transgenes that can otherwise disrupt endogenous gene expression by causing unintended mutations. Genotyping of this mouse strain was performed using primers to amplify a 602 bp product at the 5' end of the *pre-miR31-Inv* transgene (*figure 5.5A*) and results showed correct genotyping band sizes (*figure 5.5A, B and C*). Various breeding strategies were trialled throughout the project and all produced Mendelian inheritance patterns (*table 5.2*).

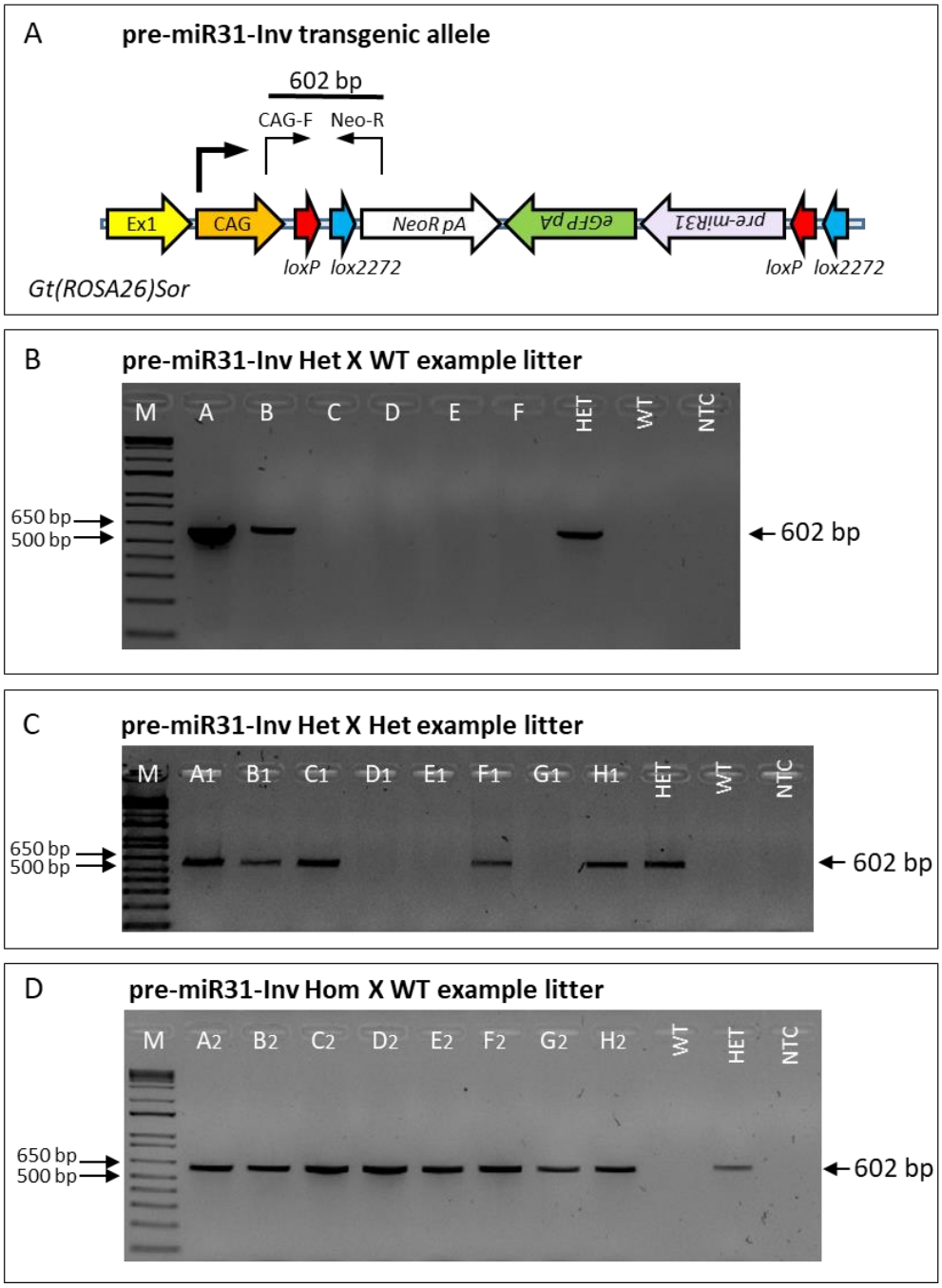


Figure 5.5 ***Example gels from various breeding strategies for pre-miR31-Inv transgenic mice colony***

A: schematic of the randomly *pre-miR31-Inv* transgene showing genotyping primers that sit in the CAG promoter (CAG-F) and neomycin-resistance gene (Neo-R) and produce a 602 bp product.
B-D: PCR products from different litters of varying breeding strategies were electrophoresed in 1.5% agarose TBE gel containing 0.002% ethidium bromide for 35 minutes at 130 V. Heterozygous (HET), wild type (WT) and no-template (NTC) controls were used in all gels and Invitrogen 1kb plus used for marker (M). Pups A - F from *pre-miR-31-Inv* heterozygous crossed with wild types (**B**), pups A_1 - H_1 from pre-miR-31-Inv heterozygous crossed with pre-miR-31-Inv heterozygous (**C**) and pups A_2 - H_2 from crossing *pre-miR-31-Inv* homozygous mice with wild types (**D**).

Table 5.2 pre-miR31-Inv colony maintenance genotyping tally

Breeding strategy	Breeding cages	Litters	Pup genotypes			Pup total	Chi-squared P
			WT	miR31(+/-)	miR31(+/+)		
miR31(+/-) X WT	3	8	24 (44.4%)	30 (55.6%)	NA	54	0.41
miR31(+/-) X miR31(+/-)	6	10	22 (33.8%)	27 (41.5%)	16 (24.6%)	65	0.23
miR31(+/+) X WT	2	7	NA	42 (100%)	NA	42	NA
miR31(+/+) X miR31(+/-)	2	2	NA	8 (47.1%)	9 (52.9%)	16	0.81
miR31(+/+) X miR31(+/+)	1	3	NA	NA	15 (100%)	15	NA

5.3 Crossing *Col1a2-Cre-ER(T) with pre-miR31-Inv* transgenic mice generates tamoxifen-inducible fibroblast-specific miR31 transgenic strains

To generate fibroblast specific *pre-miR31-Inv* transgenic mouse strains with tamoxifen inducible Cre, *Col1a2-Cre-ER(T)*$^{(Hemi)}$ mice were crossed with *pre-miR31-Inv*$^{(+/-)}$ mice. Results of an example litter are depicted below, again the third intron of the endogenous HPRT allele was used as a genotyping control and showed up positive in all littermates (*figure 5.6A*) while the *Col1a2-Cre-ER(T)* allele was inherited by approximately half the pups (*figure 5.6B*). Genotyping to query the inheritance of the *pre-miR31-Inv* transgene was performed at the 5' end and then at the 3' end for additional validation and indicated inheritance in three pups (*figure 5.6C*). To rule out the possibility of some unknown error in breeding, the first intron of the wild type ROSA locus was also PCR amplified and showed up positive in all littermates indicating that, as anticipated, no littermates at this stage were homozygous for the *pre-miR31-Inv* transgene in this litter and that breeding was therefore following the correct pattern (*figure 5.6D*). When the tally of all pup genotypes of this mouse colony was taken, results indicated Mendelian inheritance of both the *Col1a2-Cre-ER(T)* and *pre-miR31-Inv* transgenes (*table 5.3*).

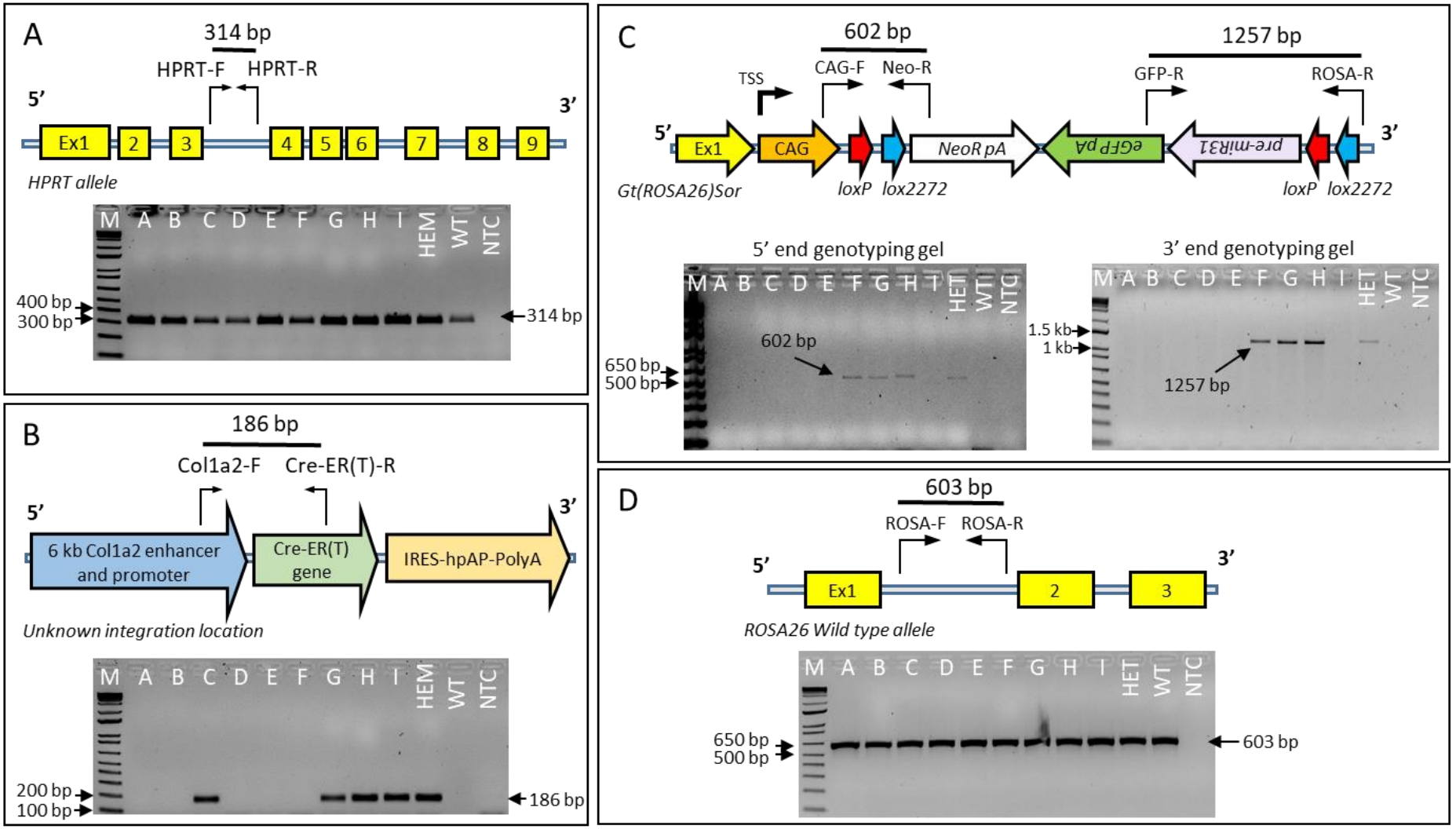

A
314 bp
HPRT-F
HPRT-R
5'
3'
Ex1
2
3
4
5
6
7
8
9
HPRT allele
M A B C D E F G H I HET WT NTC
400 bp
300 bp
314 bp
B
186 bp
Col1a2-F
Cre-ER(T)-R
5'
3'
6 kb Col1a2 enhancer and promoter
Cre-ER(T) gene
IRES-hpAP-PolyA
Unknown integration location
M A B C D E F G H I HET WT NTC
200 bp
100 bp
186 bp
C
602 bp
1257 bp
TSS
CAG-F
Neo-R
GFP-R
ROSA-R
5'
3'
Ex1
CAG
NeoR pA
eGFP pA
pre-miR31
Gt(ROSA26)Sor
loxP
lox2272
loxP
lox2272
5' end genotyping gel
3' end genotyping gel
M A B C D E F G H I HET WT NTC
602 bp
650 bp
500 bp
1.5 kb
1 kb
1257 bp
D
603 bp
ROSA-F
ROSA-R
5'
3'
Ex1
2
3
ROSA26 Wild type allele
M A B C D E F G H I HET WT NTC
650 bp
500 bp
603 bp

Figure 5.6 ***Crossing Col1a2-Cre-ER(T)$^{(Hemi)}$ with pre-miR31-Inv$^{(+/-)}$ strains produce pups carrying both transgenes***

A: schematic of internal positive control *Hprt1* primer binding sites (HPRT-F and HPRT-R) and PCR genotyping gel showing 314 bp products which validate presence of pup genomic DNA
B: schematic of primer binding sites (Col1a2-F and Cre-ER(T)-R) specific to the Col1a2 promoter and adjacent tamoxifen-inducible Cre-recombinase gene. Primers amplified a 186 bp product in littermates C, G, H and I and in HEM positive control which is shown in the PCR genotyping gel.
C: primers specific to the 5' end, Cag-F and Neo-R, and 3' end, GFP-R and ROSA-R, of the pre-miR-31-Inv allele were used and amplified a 602 and 1257 bp product respectively in littermates F, G, H and in HET positive control.
D: ROSA-F and ROSA-R, primers specific to the wild-type ROSA allele, were used to check zygosity, and amplified a 603 bp product as expected in all littermates.

Table 5.3 ***pre-miR31-Inv X Col1a2-Cre-ER(T) genotyping tally***

Breeding strategy	**Breeding cages**	**Litters**	**Pup genotypes**				**Pup total**	**Chi-squared P**
			WT/WT	miR31/WT	WT/Cre	miR31/Cre		
miR31(+/-) X Cre(Hemi)	10	35	83 (29.9%)	68 (24.4%)	64 (23.0%)	63 (22.7%)	278	0.30

5.4 Dose-dependent response of miR31 in murine atria by varying tamoxifen treatment regimens

With successful breeding of *Col1a2-Cre-ER(T)*$^{(Hemi)}$ with *pre-miR31-Inv*$^{(+/-)}$ mice, I then queried a suitable tamoxifen regimen for induction of miR31 in *Col1a2-Cre-ER(T)*$^{(Hemi)}$ */ pre-miR31-Inv*$^{(+/-)}$ mouse strain. Previous studies using the *Col1a2-Cre-ER(T)* mice utilised a variety of doses and time-courses of intraperitoneal tamoxifen to induce Cre activity including: 1 mg / day / 30 g for five days (Zheng et al. 2002); 1 mg / day / 30 g for 10 days (Pillai et al. 2017); 2.25 mg / day / 30 g for 7 days (Shi et al. 2020) and 3 mg / day / 30 g mouse for 5 days (Bageghni et al. 2018). Given the wide range of published regimens used, I chose to conduct a small pilot experiment to check whether a moderate dose of 2 mg / day / 30 g tamoxifen administered for four and seven days would induce miR31 expression. For this experiment nine mice, including both males and females, were treated with either plain corn oil (vehicle control) for four days or with tamoxifen-corn oil preparation for four or seven consecutive days. Mice were carefully monitored for behavioural changes and weighed every second day during the protocol and showed no adverse reaction or weight change with tamoxifen treatment.

Following cessation of tamoxifen, mice were sacrificed for tissue harvest and RNA was extracted from atria and left ventricle as well as outer-ear which would serve as a positive control as fibroblasts are the main cell type in ear tissue. RT-qPCR analysis indicated changes in miR31 abundance in all three tissues, atria, left-ventricle and ear, and that RNU58A remained relatively stable, particularly in the atria, and could therefore serve as a housekeeping gene for normalisation (*figure 5.7A*). MiR31 changes followed a dose response with positive correlation with the duration of tamoxifen treatment in all three tissues which indicated sensitivity of Cre to tamoxifen regimen (*figure 5.7B*). Importantly, the observed ~two-fold increase in *atrial* miR31, following a seven-day regimen of 2 mg / day / 30 g, was noted as similar to the change observed in human patients with chronic AF compared to sinus rhythm controls (*see above figure 3.1*). Therefore, to simulate perturbations in atrial miR31 relevant to AF, this dose and time course of tamoxifen treatment was used for all future experimentation in our mouse model.

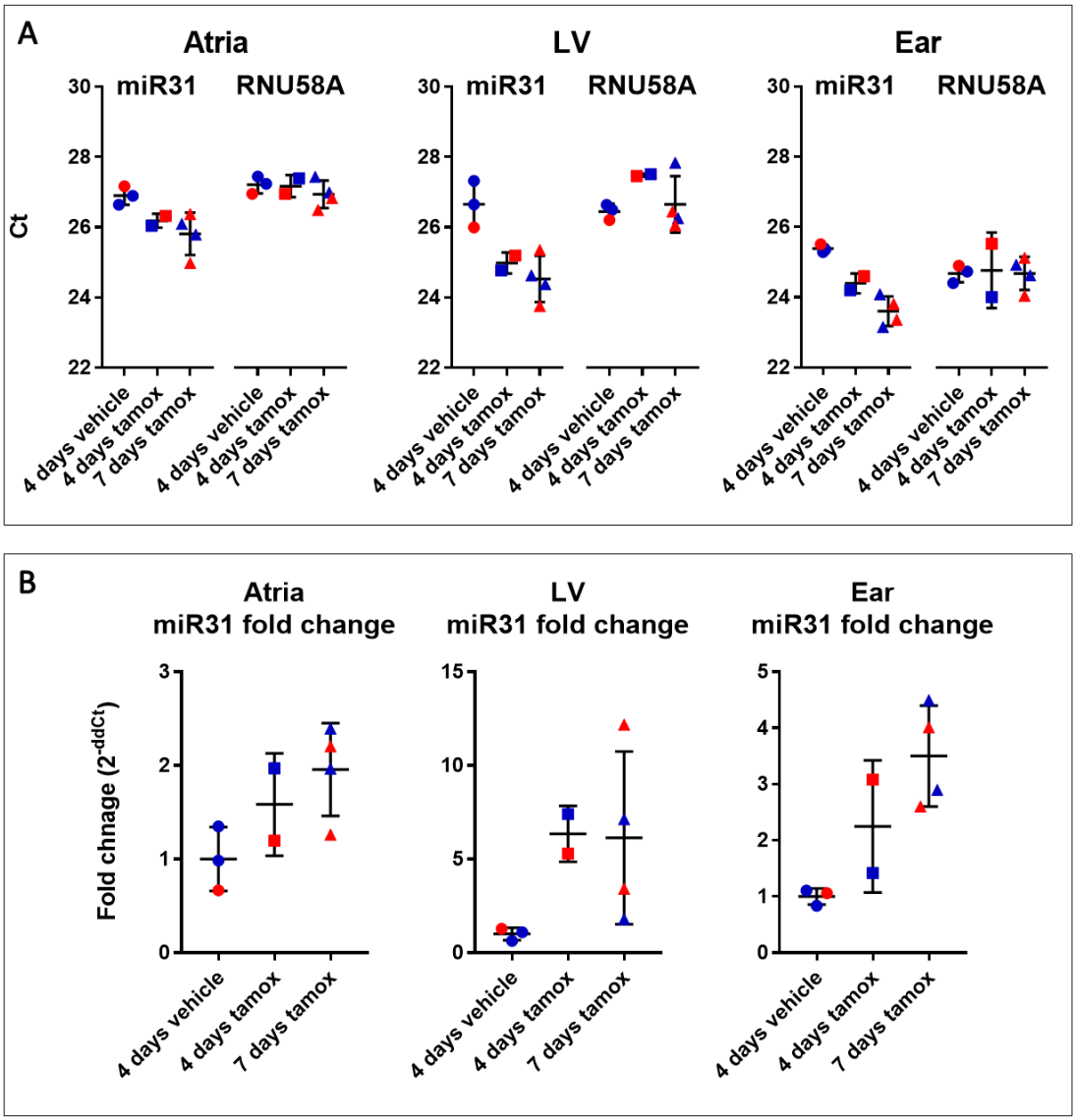


Figure 5.7 ***Pilot experiment showing miR31 expression in different tissues following varying intraperitoneal tamoxifen regimens in Col1a2-Cre-ER(T)$^{(Hemi)}$ / pre-miR31-Inv$^{(+/-)}$ mice***

Nine mice eight weeks of age including males and females carrying both the Cre and miR31 transgenes, were treated with intraperitoneal corn oil (vehicle control) or tamoxifen-corn oil preparation for four or seven consecutive days (2 mg / day / 30 g). Mice were then left for a week to clear tamoxifen before tissues including atria, LV and ear were harvested and cryopreserved at -80 °C. RNA was then extracted and analysed for changes in miR31 abundance using qPCR.

A: raw qPCR Ct values for miR31 and RNU58A from RNA derived from atria, LV and EAR
B: miR31 normalised to RNU58A and presented as fold-change of four-day vehicle control group. From left to right n = 3, 2 and 4. Blue colour indicate males, red indicate females.

5.5 Tamoxifen induces inversion of the *pre-miR31-Inv* transgene at the *Gt(ROSA26)Sor* locus

Results so far indicated that our newly generated miR31 transgenic mouse was performing as intended, which was that tamoxifen administration was inducing increased miR31 expression in mice carrying both the *Cre* and *pre-miR31-Inv* transgenes. Following this, I sought to specifically show inversion of the loxP flanked *pre-miR31-Inv* transgene following tamoxifen treatment. This observation would serve to further validate my experimental model as well as demonstrate that our Cre-recombinase dependent inversion switch was working as intended *in vivo*. For this experiment mice were genotyped pre-tamoxifen treatment (*figure 5.8A*) and then genomic DNA was extracted from ear tissue subsequent to the intraperitoneal tamoxifen treatment protocol (*figure 5.8B*). Ear genomic DNA was then PCR amplified to check for un-recombined pre-miR31-Inv, as well as inverted-recombined and trapped *pre-miR31-Inv* allele at the 5' end. Electrophoresed products indicated that irrespective of the presence of the *Cre* allele, all mice carrying the *pre-miR31-Inv allele* (mice F, G and H) retained some portion of ear tissue cells in which no recombination event occurred post tamoxifen treatment indicated by a 602 bp band (*figure 5.8C, upper left gel*). However, as expected only mice carrying both the *pre-miR31-Inv* and *Cre* alleles together (mice G and H) showed inversion-recombination and trapping events at the 5' end indicated by 1221 and 1143 bp bands respectively. To better visualize these Cre-dependent events, mouse samples G and H were re-run and electrophoresed further for improved resolution and showed expected band sizes (*figure 5.8C, bottom gel*).

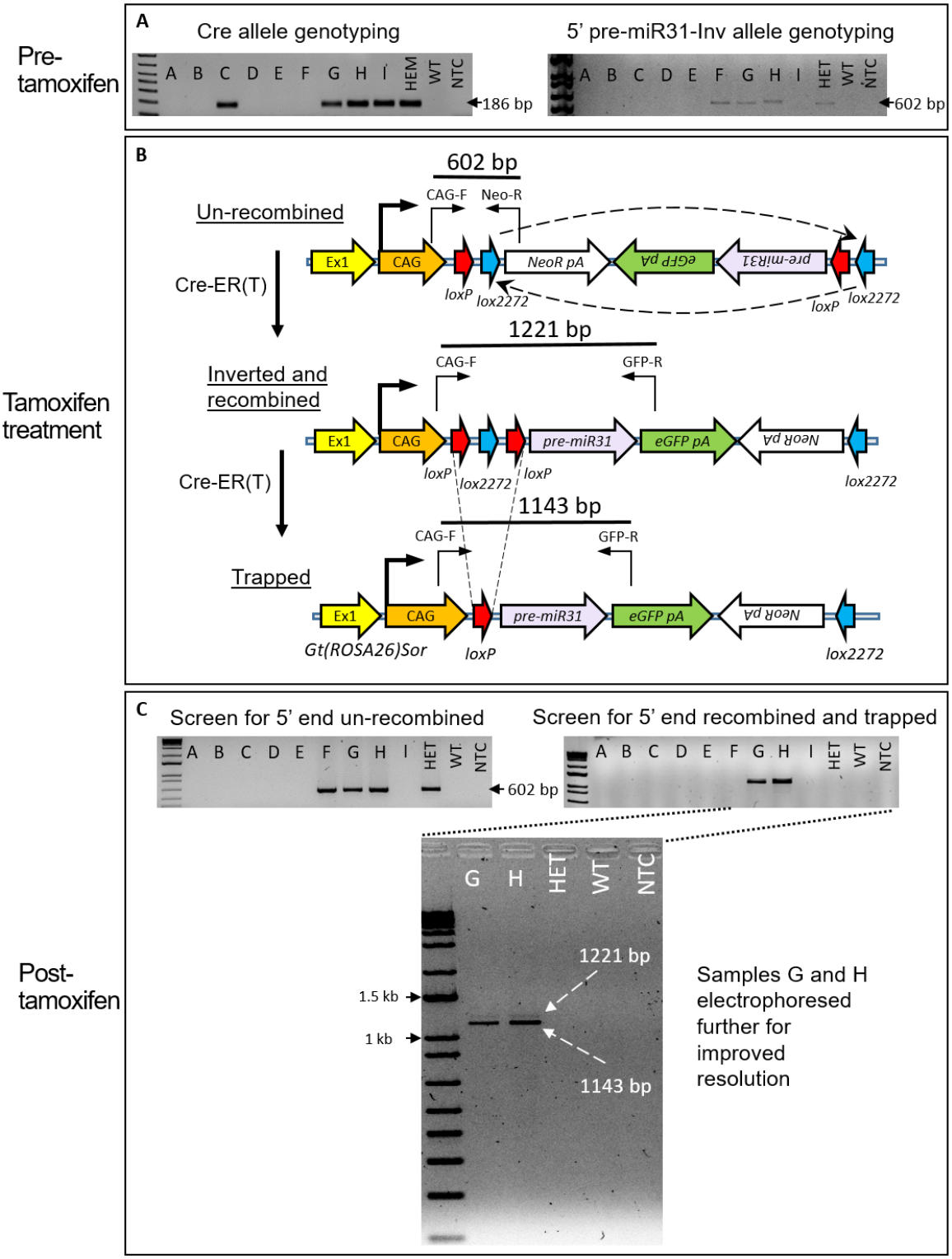

Pre-tamoxifen
A
Cre allele genotyping
A B C D E F G H I HEM WT NTC
186 bp
5' pre-miR31-Inv allele genotyping
A B C D E F G H I HET WT NTC
602 bp
Tamoxifen treatment
B
602 bp
Un-recombined
CAG-F
Neo-R
Ex1
CAG
NeoR pA
eGFP pA
pre-miR31
Cre-ER(T)
loxP
lox2272
loxP
lox2272
1221 bp
Inverted and recombined
CAG-F
GFP-R
Ex1
CAG
pre-miR31
eGFP pA
NeoR pA
loxP
lox2272
loxP
lox2272
Cre-ER(T)
1143 bp
Trapped
CAG-F
GFP-R
Ex1
CAG
pre-miR31
eGFP pA
NeoR pA
Gt(ROSA26)Sor
loxP
lox2272
Post-tamoxifen
C
Screen for 5' end un-recombined
A B C D E F G H I HET WT NTC
602 bp
Screen for 5' end recombined and trapped
A B C D E F G H I HET WT NTC
G H HET WT NTC
1221 bp
1.5 kb
1 kb
1143 bp
Samples G and H electrophoresed further for improved resolution

Figure 5.8 ***Genomic DNA PCR analysed for inversion, recombination and trapping at the 5' end of the pre-miR31-Inv allele at the targeted Gt(ROSA26)Sor locus***

A: gels depict littermates (A – I) genotyped for the Cre and pre-miR31-Inv transgenes pre-tamoxifen treatment.
B: schematic showing that Cre-dependent inversion-recombination and trapping at the 5' end of the pre-miR31-Inv transgene during tamoxifen administration is traceable using CAG-F and Neo-R primers for un-recombined transgene and CAG-F and GFP-R primers for inverted-recombined and trapped transgene.
C: left gel includes PCR products electrophoresed to screen for the un-inverted pre-miR31-Inv allele which showed a 602 bp product in littermates F, G, H and in HET positive control. Right gel includes PCR products to screen for inversion-recombination and trapping at the 5' end of the pre-miR31-Inv allele and shows 1221 and 1143 bp bands in littermates G and H. In the third gel samples G and H were re-run and electrophoresed for 30 minutes longer to better resolve and visualise band doublets to confirm two PCR products.

Following confirmation of inversion of the *pre-miR31-Inv* transgene at the 5' end, additional validation was performed at the 3' end. Genotyping pre-tamoxifen at the 3' end was in congruence with the above results showing that mice F, G and H carry the *pre-miR31-Inv* transgene (*figure 5.9A*) and that these same mice retained un-recombined transgene post-tamoxifen treatment (*figure 5.9C, left gel*). In addition, only mice carrying Cre, namely G and H, showed inversion-recombination and trapping at the 3' end (*figure 5.9C, right gel*).

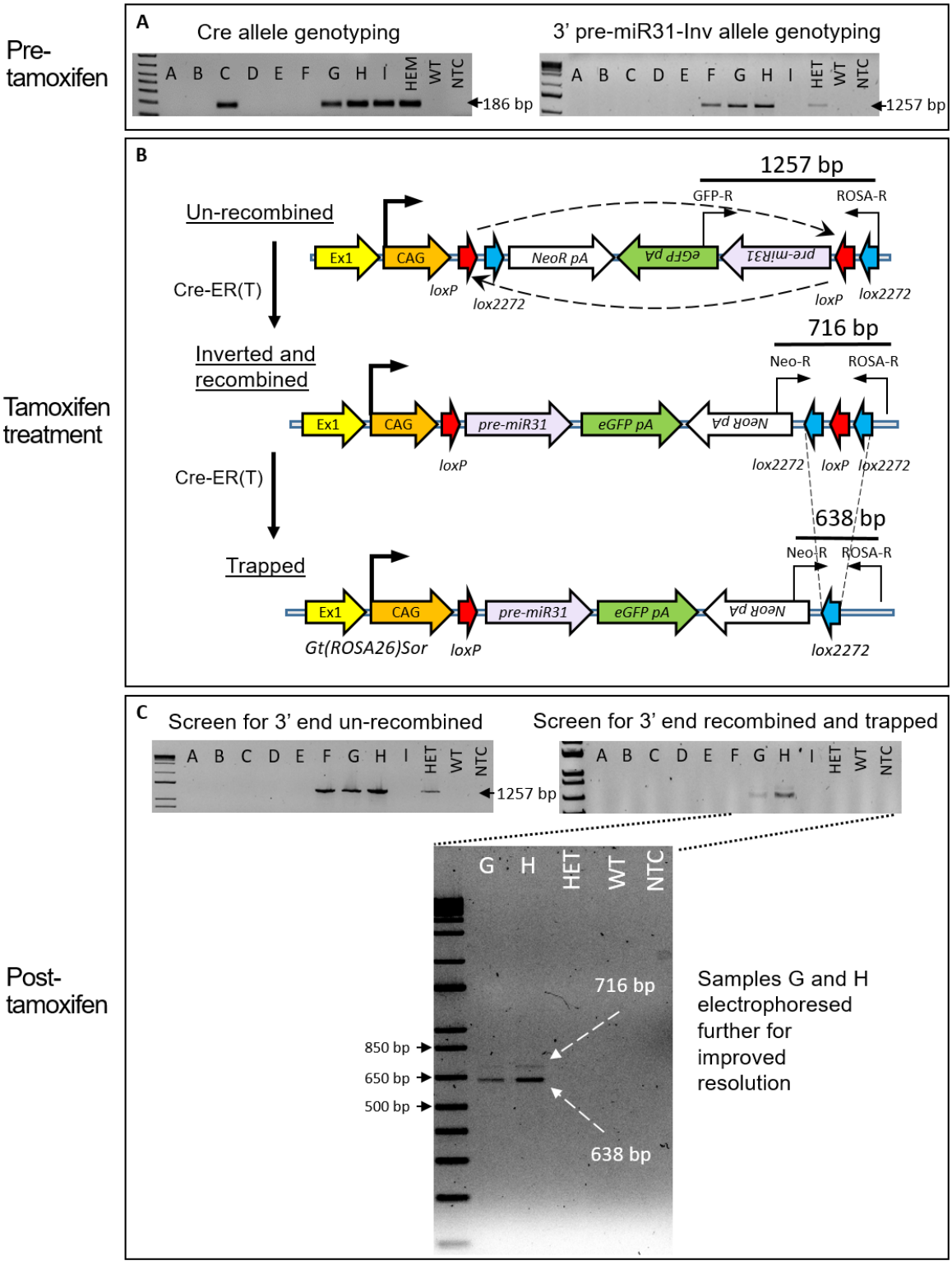

Pre-tamoxifen
A
Cre allele genotyping
A B C D E F G H I HEM WT NTC
186 bp
3' pre-miR31-Inv allele genotyping
A B C D E F G H I HET WT NTC
1257 bp
Tamoxifen treatment
B
1257 bp
GFP-R
ROSA-R
Un-recombined
Ex1
CAG
NeoR pA
eGFP pA
pre-miR31
loxP
lox2272
Cre-ER(T)
716 bp
Neo-R
ROSA-R
Inverted and recombined
pre-miR31
eGFP pA
NeoR pA
638 bp
Trapped
Gt(ROSA26)Sor
Post-tamoxifen
C
Screen for 3' end un-recombined
A B C D E F G H I HET WT NTC
1257 bp
Screen for 3' end recombined and trapped
A B C D E F G H I HET WT NTC
G H HET WT NTC
716 bp
638 bp
850 bp
650 bp
500 bp
Samples G and H electrophoresed further for improved resolution

Figure 5.9 ***Inversion, recombination and trapping at the 3' end of the pre-miR31-Inv allele***

A: the same littermates (A – I) genotyped for Cre and 3' pre-miR31-Inv transgenes pre-tamoxifen treatment.
B: schematic showing that Cre dependent changes are traceable at the 3' end
C: un-recombined indicated by a 1257 bp product in littermates F, G, H and in HET positive control in left gel. Inverted-recombined and trapped 3' pre-miR31-Inv allele indicated by 716 and 638 bp bands in littermates G and H in right gel. Samples G and H were re-run for 30 minutes longer to visualise band doublets in third gel.

As a final query, the same littermates were checked for inversion of the *pre-miR31-Inv* allele *pre*-tamoxifen treatment. This assay was used to uncover whether inversion of the *pre-miR31-Inv* allele, particularly in mice carrying *Cre*, was occurring independent of tamoxifen administration by some unknown mechanism. Results showed that no littermates carried an inverted *pre-miR31-Inv* allele pre-tamoxifen, indicating that inversion is indeed tamoxifen dependent (*figure 5.10*).



Figure 5.10 ***Checking for unwanted inversion pre-tamoxifen treatment***

The same littermates were genotyped for inversion of the *pre-miR31-Inv* allele pre-tamoxifen treatment. The positive control DNA carries the inverted *pre-miR31-Inv* and confirms PCR viability.

When considered together the experimental results so far have shown that our tamoxifen regimen increased miR31 abundance in mice carrying both the *Col1a2-Cre-ER(T)* and floxed *pre-miR31-Inv* transgenes and not in mice carrying either of the transgenes alone or in wild-types. Further, when ear gDNA from these mice was queried for Cre-dependent recombination events including inversion, recombination and trapping, at the floxed *pre-miR31-Inv* transgene, again this was only demonstrable in mice carrying both transgenes. This indicated that our inversion switch was working and was indeed leading to a strong overexpression of the miR31 transgene in our mouse model following tamoxifen treatment.

5.6 Increased miR31 expression is genotype specific

Subsequent to determining a suitable tamoxifen regimen correlative with increased miR31 expression in our pilot experiment, I sought to demonstrate that this observation was a result specific to mice carrying both the Cre and *pre-miR31-Inv* transgenes, and not simply a general and undocumented response of C57BL/6J mice to intraperitoneal tamoxifen. To show this I treated mice of all four genotype variations with the same tamoxifen regimen and compared miR31 expression levels (*figure 5.11*).

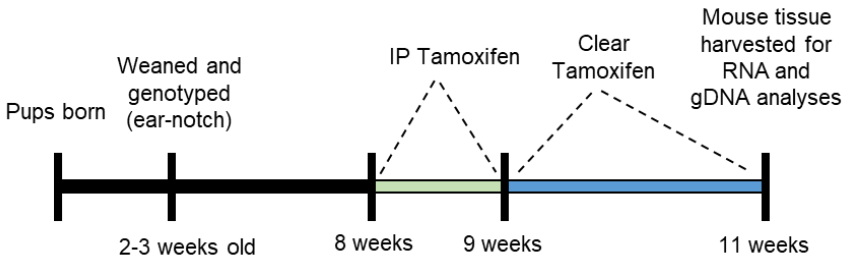


Figure 5.11 ***Work flow diagram of intraperitoneal tamoxifen administration in mice***

Mice were weaned and genotyped at two-three weeks of age. At eight weeks mice were treated with intraperitoneal tamoxifen for seven days and were then left to clear tamoxifen for two weeks before being sacrificed for tissue harvest.

Results indicated that in all tissues investigated, namely atria, LV and ear, that mice carrying both the *pre-miR31-Inv* and *Cre* alleles expressed significantly increased amounts of miR31 compared to the *WT/WT* group (*figure 5.12*). This observation indicated that the lox-p flanked *pre-miR31-Inv* transgene was likely undergoing a Cre-dependant recombination event leading to increased miR31 expression from the targeted allele and that our experimental system was working as designed. Despite this reassurance, this experiment also indicated that mice carrying the *pre-miR31-Inv* transgene alone (*miR31/WT*) were expressing increased abundance of miR31, albeit not statistically significant, when compared to *WT/WT* mice. The *pre-miR31-Inv* transgene is located at the *Gt(ROSA26)Sor* locus which is known to produce low level transcription off the reverse compliment strand, thus a plausible explanation of this observation is that miR31 is being expressed in small quantities independent of a Cre-recombination event.

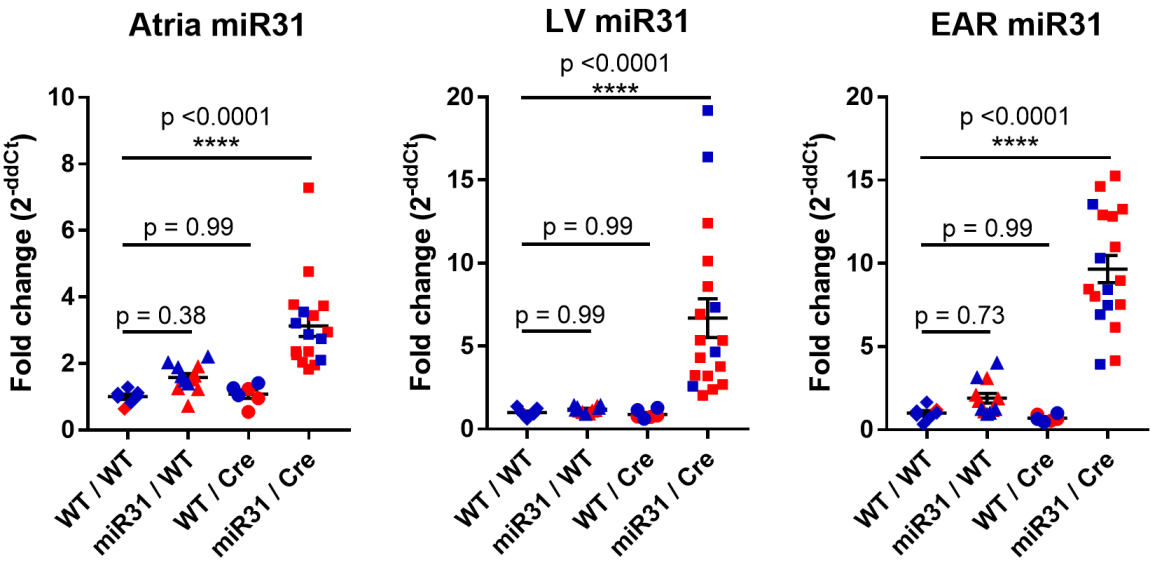


Figure 5.12 ***miR31 fold change in various tissues of interest is genotype specific***

Mice of all four genotype variations were treated with our intraperitoneal tamoxifen regimen. RNA was extracted and miR31 abundance measured using RT-qPCR. *RNU58A* was used as a housekeeping gene for normalisation and results are presented as fold change of *WT*/*WT*. For atria and ear, miR31 abundance showed normal distribution in each genotype group by K-S test so means were compared to *WT/WT* using one-way ANOVA with Dunnett's correction for multiple comparison. For LV, the miR31/Cre group showed non-parametric distribution of miR31 therefore all genotype groups were compared to *WT*/*WT* using Kruskal-Wallis test with Dunnett's correction for multiple comparison. N = 7, 12, 6, and 18 for genotype groups left to right. Blue colour indicate males, red indicate females. Data presented as mean ± SD.

In addition to showing increased miR31 expression in tissues of interest I queried whether *eGFP* mRNA was being expressed from the modified *Gt(ROSA26)Sor* locus in these mice as an *eGFP* cassette was included in the transgenic construct downstream of *pre-miR31-Inv* to visualise transgene expression (*see figure 5.1 above*). Ear RNA from *pre-miR31-Inv*$^{(+/-)}$ / *Cre*$^{(Hemi)}$ and from *WT/WT* control was assessed by one-step RT-PCR using two separate primer pairs at a range of primer annealing temperatures during thermocycling for optimisation. Electrophoresed PCR products indicated *eGFP* mRNA expression in *pre-miR31-Inv*$^{(+/-)}$ / *Cre*$^{(Hemi)}$ mice and not in *WT/WT* (*figure 5.13*). *eGFP* protein expression was not queried due to time constraints but would be useful in future investigations to visualise transgene expression. For example, *eGFP* expression could be used to estimate what portion of atrial fibroblasts in histological sections are expressing the transgene as microRNA *in situ* hybridisation protocols are

technically challenging (Nielsen 2012). In addition, future experiments will require the purification of atrial fibroblasts from *pre-miR31-Inv*$^{(+/-)}$ / *Cre*$^{(Hemi)}$ mouse atria and visualising *eGFP* expression could be used to confirm *in vitro* transgene expression using bench-top fluorescence microscopes.

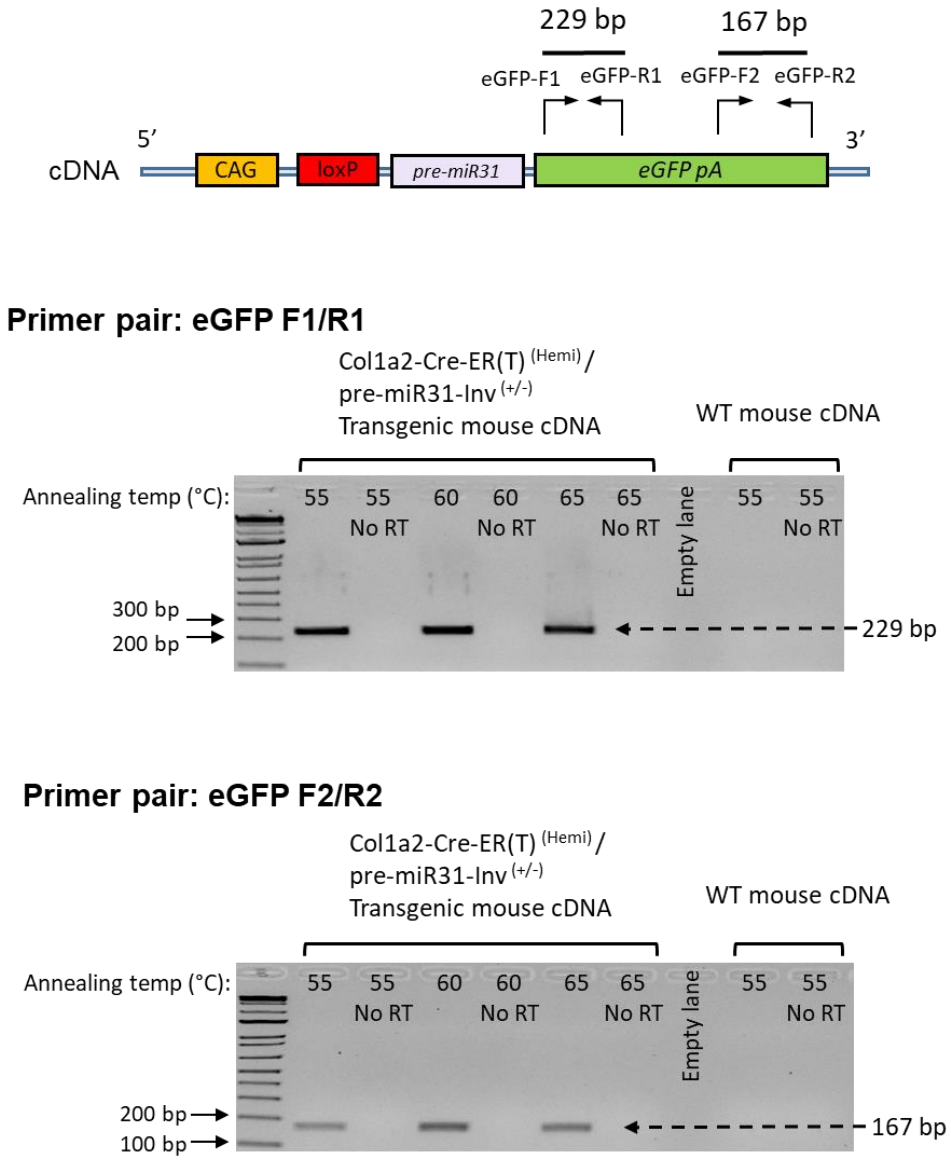


Figure 5.13 ***eGFP mRNA transcripts are expressed in pre-miR31-Inv (+/-) / Cre (Hemi) post-tamoxifen***

Ear RNA extracted post-tamoxifen was assessed by one-step RT-PCR for *eGFP* mRNA expression in miR31/Cre and *WT/WT* mice. Primer annealing temperatures 50, 55 and 60 °C were assessed during PCR as primers were newly designed. No reverse transcriptase (No-RT) was used for each PCR reaction as control to show PCR products were not amplicons from genomic DNA contamination.

5.7 Evaluating the effects of miR31 overexpression on cardiac fibrosis and function in pre-miR31-Inv transgenic mice

Having validated my in vivo experimental system I sought to then investigate the effects of fibroblast-specific increased miR31 abundance on cardiac fibrosis and function. For these experiments, I decided to first use echocardiography to screen for any overt and unintended changes in cardiac phenotype, specifically in left-ventricular chamber dimensions and wall thickness - which are surrogate measures for dilatation and hypertrophy respectively. Following echocardiography, mice were sacrificed and hearts were harvested to explore underlying molecular changes. RT-qPCR was used to assess gene expression profiles, immunoblotting for protein changes, and hydroxyproline assay for collagen production (*figure 5.14*).

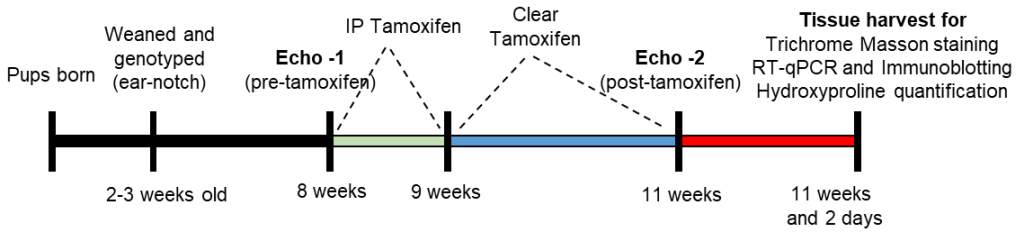


Figure 5.14 ***Workflow diagram showing timings of echocardiography measurements, intraperitoneal tamoxifen administration and tissue harvest***

Mice were weaned and genotyped at two-three weeks of age and then at eight weeks baseline echocardiography measurements were taken. Mice were then treated with intraperitoneal tamoxifen for seven days before being left to clear tamoxifen for two weeks. A second round of echocardiography measurements were then taken and following another two days mice were then culled for tissue harvest.

5.7.1 Validation of echocardiography image analysis against a training data set

In the following experiment echocardiography was used to visualise and generate images of murine cardiac tissues. During image analysis cardiac dimensions are measured by manually outlining structures using Vevo Lab Imaging System (Edition 1.7.1). I measured left ventricular *epicardial* end-diastolic area (LV-Epi-EDA) and *endocardial* end-diastolic (LV-Endo-EDA) and end-systolic area (LV-Endo-ESA) from a parasternal short-axis view at the level of the papillary muscles (*figure 5.15*).

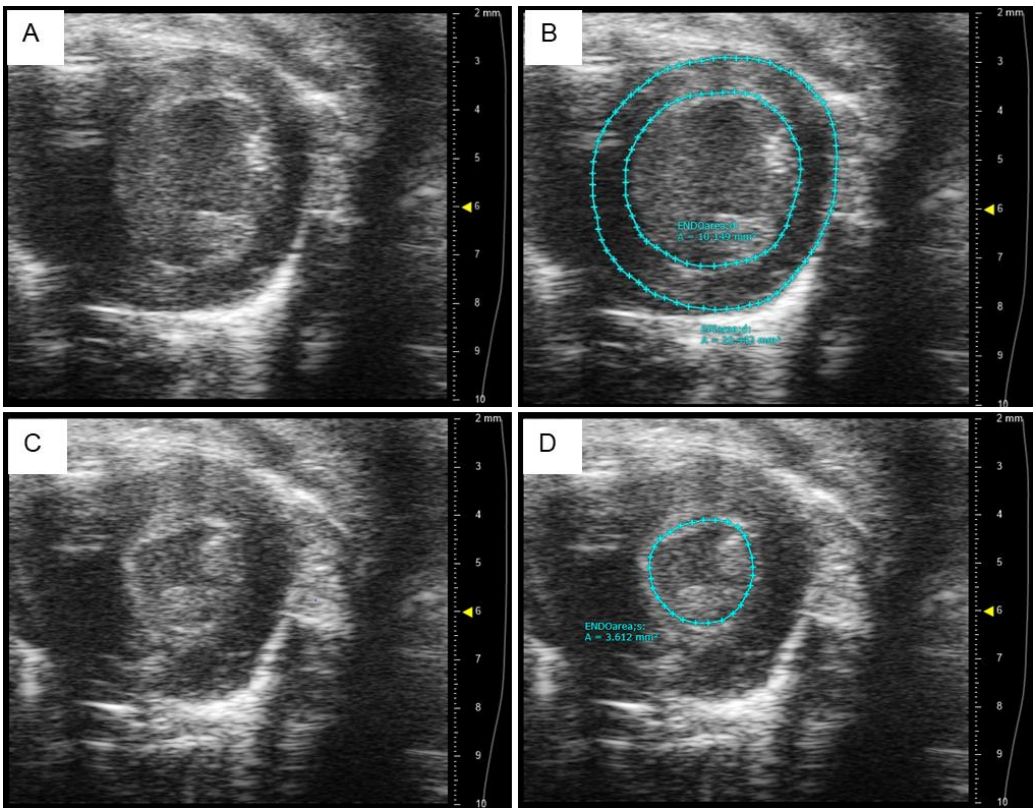
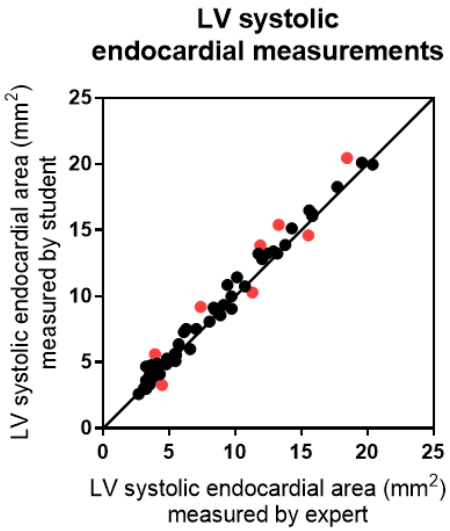
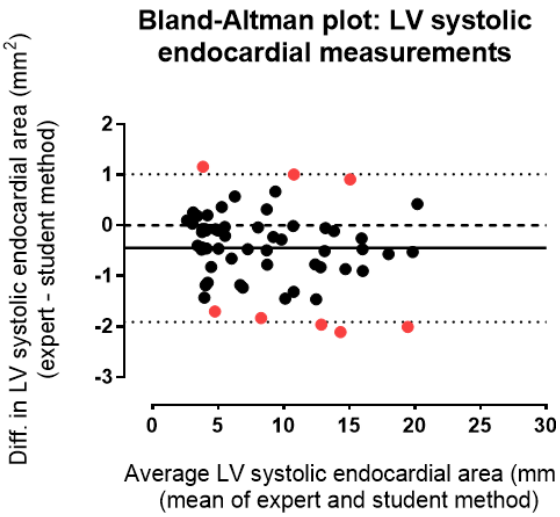


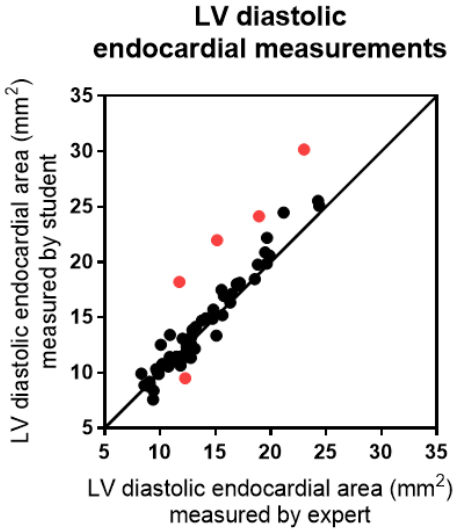
Figure 5.15 Example murine short-axis echocardiograms

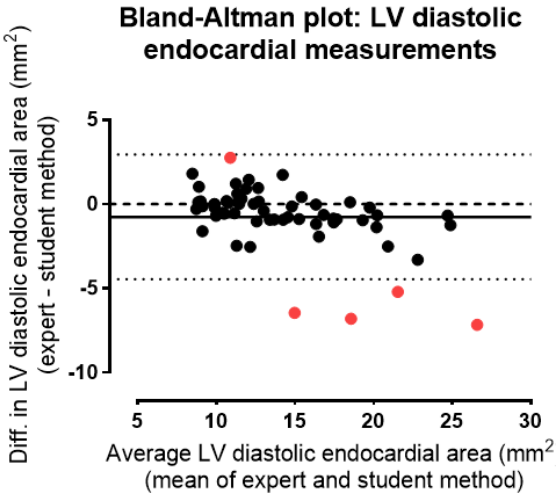
Parasternal short-axis views of murine LV at the level of the papillary muscles. Images A and B depict end-diastolic view with, and without, traced outlines of end-diastolic endocardial and epicardial areas. Image C and D depicts short-axis end-systolic view with, and without, traced outline of endocardial end-systolic area.

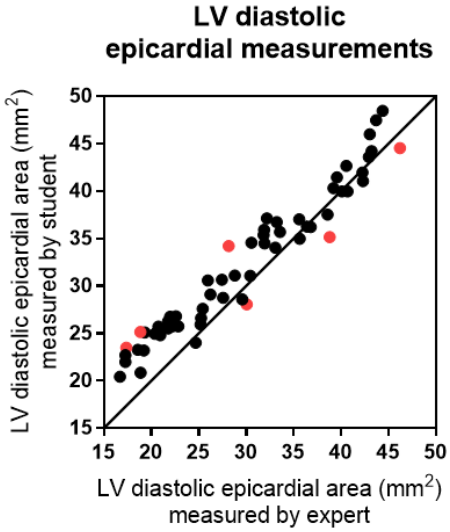
Technical variation is introduced when estimating heart dimensions from echocardiographic images as it requires manual outlining of perceived structures. To minimise error in my own experimental mice, I was first trained in echocardiographic image analysis by our local expert - Prof. Craig Lygate, using a previously generated data-set. To demonstrate competency and determine if our analyses differed, I plotted the results of my method against his on axes with equal scaling and a diagonal line of equality. These graphs showed high correlation between our measurements with a tendency of mine to overestimate end-diastolic endocardial and epicardial measurements (*figure 5.16, left side images*). I then used Bland-Altman analysis to query the agreement between the two datasets and more clearly determine systematic bias in my measurements. Bland-Altman plots depict the difference against the mean of two datasets and bias is observed by the mean of differences being non-zero. Outliers can also be identified as data points outside or close to the limits of agreement (±2 SD from the mean of differences). The Bland-Altman plots did show a small bias of mine to overestimate LV endocardial systolic and diastolic dimensions, and a larger bias towards overestimating LV diastolic dimensions (*figure 5.16, right side images*). Data points outside or close to the Bland-Altman limits of agreement were identified and following additional coaching and refinement of my measuring technique outliers were then re-analysed. Results following re-analysis indicated that outlying data points were brought closer to the mean of difference and this narrowed the standard deviation of the data thereby improving agreement between the two datasets (*figure 5.17*).


LV systolic endocardial measurements
LV systolic endocardial area (mm2) measured by student
LV systolic endocardial area (mm2) measured by expert
0
5
10
15
20
25


Bland-Altman plot: LV systolic endocardial measurements
Diff. in LV systolic endocardial area (mm2) (expert - student method)
Average LV systolic endocardial area (mm (mean of expert and student method)
2
1
0
-1
-2
-3
0
5
10
15
20
25
30


LV diastolic endocardial measurements
LV diastolic endocardial area (mm2) measured by student
LV diastolic endocardial area (mm2) measured by expert
5
10
15
20
25
30
35


Bland-Altman plot: LV diastolic endocardial measurements
Diff. in LV diastolic endocardial area (mm2) (expert - student method)
Average LV diastolic endocardial area (mm2 (mean of expert and student method)
5
0
-5
-10
5
10
15
20
25
30


LV diastolic epicardial measurements
LV diastolic epicardial area (mm2) measured by student
LV diastolic epicardial area (mm2) measured by expert
15
20
25
30
35
40
45
50

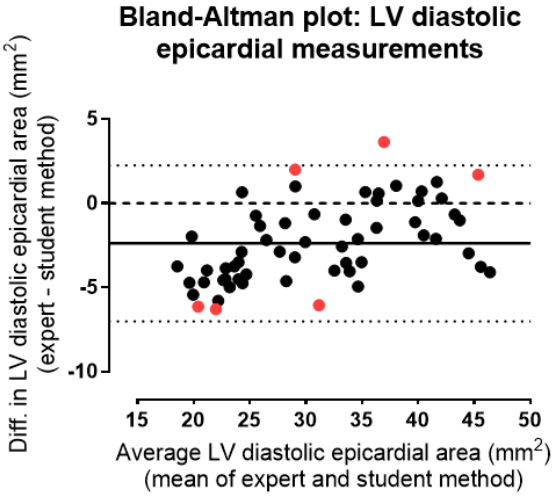
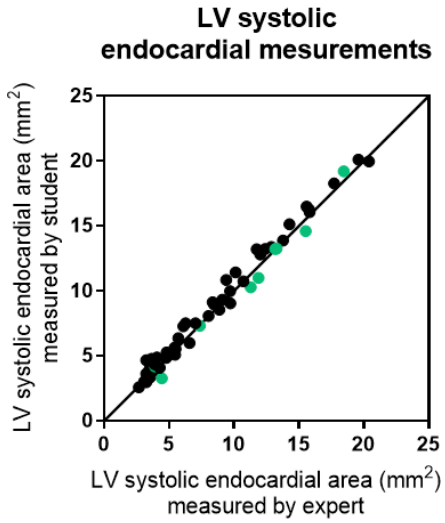
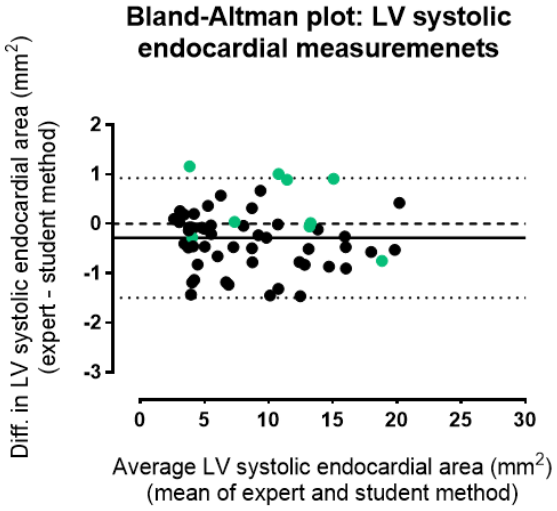
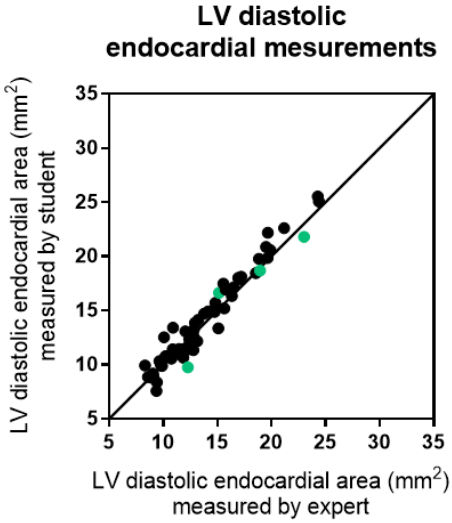

Bland-Altman plot: LV diastolic epicardial measurements
Diff. in LV diastolic epicardial area (mm2) (expert - student method)
Average LV diastolic epicardial area (mm2) (mean of expert and student method)
5
0
-5
-10
15
20
25
30
35
40
45
50

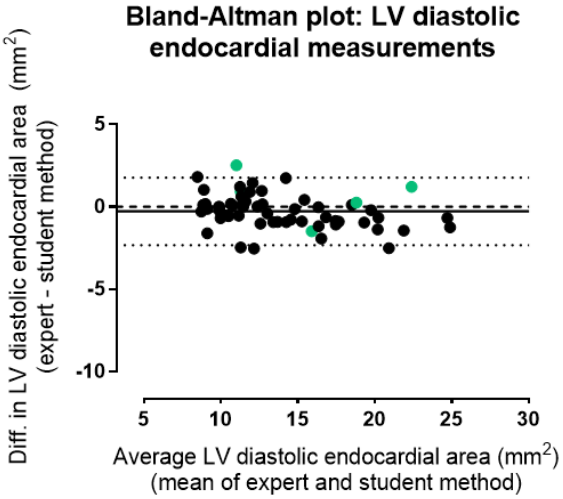
Figure 5.16 ***Comparison of expert and student analyses of echocardiographic images visualised using correlation and Bland-Altman plots***

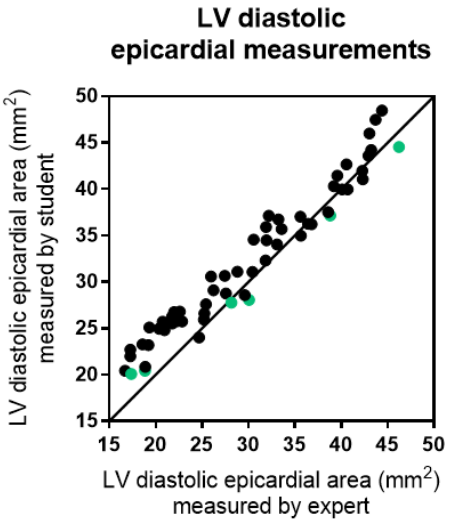
Echocardiographic images taken from fifty-nine mice were analysed independently by expert and student. Left side graphs have equal axes and diagonal line of equality to visualise obvious differences in subjective measurements. Right side graphs are Bland-Altman plots which identify data points close to, or outside (coloured red) two standard-deviations (limits of agreement: dotted line) from the mean of differences between the two datasets (solid horizontal line). Systematic bias is observed by the mean of differences being non-zero.


LV systolic endocardial mesurements
LV systolic endocardial area (mm2) measured by student
LV systolic endocardial area (mm2) measured by expert
0
5
10
15
20
25


Bland-Altman plot: LV systolic endocardial measuremenets
Diff. in LV systolic endocardial area (mm2) (expert - student method)
Average LV systolic endocardial area (mm2) (mean of expert and student method)
2
1
0
-1
-2
-3
0
5
10
15
20
25
30


LV diastolic endocardial mesurements
LV diastolic endocardial area (mm2) measured by student
LV diastolic endocardial area (mm2) measured by expert
5
10
15
20
25
30
35


Bland-Altman plot: LV diastolic endocardial measurements
Diff. in LV diastolic endocardial area (mm2) (expert - student method)
Average LV diastolic endocardial area (mm2) (mean of expert and student method)
5
0
-5
-10
5
10
15
20
25
30


LV diastolic epicardial measurements
LV diastolic epicardial area (mm2) measured by student
LV diastolic epicardial area (mm2) measured by expert
15
20
25
30
35
40
45
50

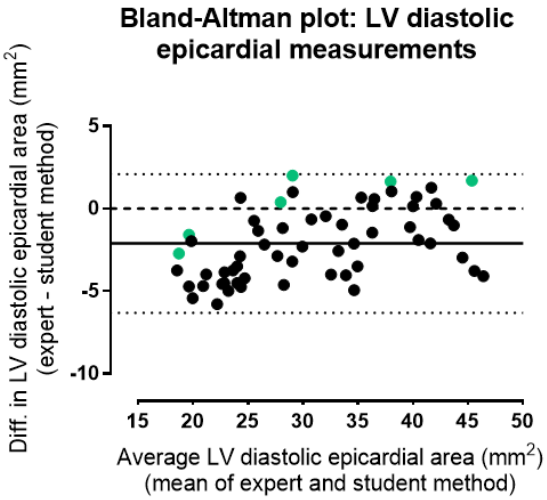

Bland-Altman plot: LV diastolic epicardial measurements
Diff. in LV diastolic epicardial area (mm2) (expert - student method)
Average LV diastolic epicardial area (mm2) (mean of expert and student method)
5
0
-5
-10
15
20
25
30
35
40
45
50

Figure 5.17 Echocardiographic data re-analysed following refinement of measurement technique

Echocardiographic images taken from fifty-nine mice were analysed independently by expert and student. Green dots represent data from images that were re-analysed following refinement of measurement technique. Left side graphs have equal axes and diagonal line of equality to visualise obvious differences in subjective measurements. Right side graphs are Bland-Altman plots which identify data points close to, or outside two standard-deviations (limits of agreement: dotted line) from the mean of differences between the two datasets (solid horizontal line). Systematic bias is observed by the mean of differences being non-zero

5.7.2 Mice carrying the *Col1a2-CreER(T)* or *pre-miR31-Inv* transgenes do not show changes in systolic heart function before tamoxifen administration

The former training exercise was used to demonstrate competency in echocardiographic image analysis and following this I sought to analyse my own data. In my experimental approach, I used the *Col1a2-CreER(T*) mouse strain to create a tamoxifen-inducible and fibroblast-specific system to overexpress miR31 in adult mice. There have been no previous studies investigating the adverse cardiovascular effects of simply carrying the *Col1a2-CreER(T*) transgene, and indeed no studies on the novel *pre-miR31-Inv* transgenic mouse either. Given that the aim of this *in vivo* work was to draw robust conclusions specifically on the effects of miR31 overexpression on heart phenotype, it was important to first query the potentially confounding effects caused by simply carrying the Cre and miR31 transgenes.

For this experiment, I first looked at the effect of genotype alone (i.e. before the use of tamoxifen) on heart rate, LV endocardial end diastolic area (LV-Endo-EDA), myocardial cross sectional area (MyoCSA) and fractional area change (FAC%). These parameters were chosen as they are widely used to assess systolic heart function and could therefore serve as metrics to screen for obvious changes in heart function (Lindsey et al. 2018). LV-Endo-EDA is a surrogate measure for left-ventricular dilatation, MyoCSA for cardiac hypertrophy and FAC% for ejection fraction. Diastolic images were not analysed since at this stage of the project I was simply screening for unintended consequences of genotype on cardiac function. In addition, given that pre-tamoxifen echocardiography data was taken in mice at 8 weeks of age where males

and females show significant weight differences, I chose to balance each genotype group with equal numbers of males in attempt to correct for obvious weight effects on heart dimensions and functional readouts. Results indicated that mice carrying either the miR31 or Cre transgene alone, or when carrying both transgenes together, did not show significant changes in body weight, heart rate, left-ventricular dilatation and hypertrophy or ejection fraction when compared to wild-type controls (*figure 5.18*).

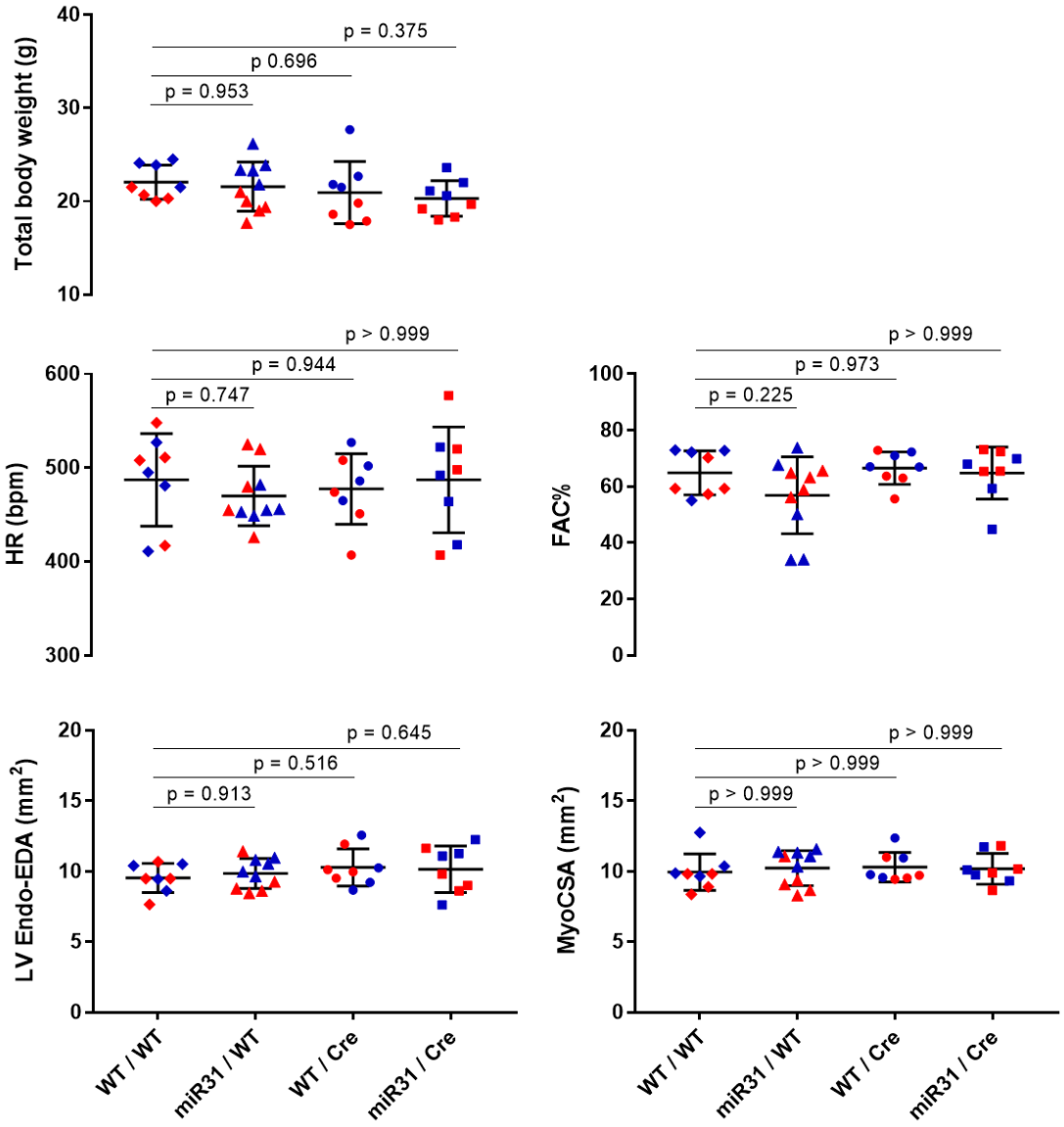


Figure 5.18 Echocardiographic parameters pre-tamoxifen

Figure shows the effects of genotype on various experimental readouts pre-tamoxifen administration in mice 8 weeks of age. From left to right genotype group n = 8, 10, 8 and 8. Each genotype group contains equal number of males and females to balance for gender effects. Red symbols represent females, blue represent males. K-S statistical test determined that all data was normally distributed except for MyoCSA in the WT/Cre group. Means of normally distributed data was compared to WT/WT control using one-way ANOVA with Dunnett's correction for multiple comparison of means. Means of non-parametric data were compared to WT/WT genotype group using Kruskal-Wallis test with Dunn's correction for multiple comparison of means. Low HR suggests excessive cardiac depression due to general anaesthesia therefore echocardiography data taken from mice with heart rates <400 bpm were not included in the analysis. HR: heart rate. FAC%: fractional area change. LV Endo-EDA: Left ventricle endocardial end diastolic area. MyoCSA: Myocardial cross sectional area. Error bars depict mean ± SD.

5.7.3 Tamoxifen affects systolic heart function independent of genotype

Having shown no obvious effects of the miR31 and Cre transgenes on systolic heart function, I then asked whether the use of tamoxifen in my experimental design would cause unintended effects on these same parameters. There are currently no reports on the effects of tamoxifen and Cre induction in the *Col1a2-CreER(T)* mouse strain on heart function, however there is evidence that the activity of Cre following tamoxifen administration in other experimental systems has adverse cardiac effects independent of the presence of a floxed transgene. For example, α*MHC-MerCreMer* mice showed significantly reduced ejection fraction by echocardiography four weeks following tamoxifen compared to vehicle control α*MHC-MerCreMer* mice (Lexow et al. 2013). The same study showed that this was not reproducible in WT mice treated with tamoxifen suggesting that the effect was specific to an active Cre in α*MHC-MerCreMer* mice and not due to tamoxifen alone. The post-tamoxifen washout period in this study was four weeks and therefore they could not comment on the acute effects of tamoxifen on cardiac function. However, a similar study in α*MHC-MerCreMer* mice did show a depressed ejection fraction one week following completion of tamoxifen treatment but this study used a vehicle control and therefore couldn't comment on the general effect of tamoxifen on WT mice (Bersell et al. 2013).

To query the effects of tamoxifen in my own experimental system, I compared pre- and post-tamoxifen echocardiography data. To control for the potentially confounding effects of tamoxifen induced Cre-recombinase, and/or the presence of the floxed miR31 transgene, I chose to look at the effects of tamoxifen *within* each genotype group (i.e. control for genotype effects). In addition, my pre- and post-tamoxifen murine echocardiography images were taken at 8 and 11 weeks respectively which is a period during which mice show significant increase in mass and therefore heart dimensions are usually normalised to total body weight. I therefore first queried whether my tamoxifen regimen was retarding the predicted weight increase during this age range to confirm that it would indeed be appropriate to normalise my data to mouse weight. Results indicated that mice of all genotypes significantly increased in body weight despite tamoxifen treatment in all genotype groups and therefore echocardiography data were normalised to body weight (*figure 5.19*).

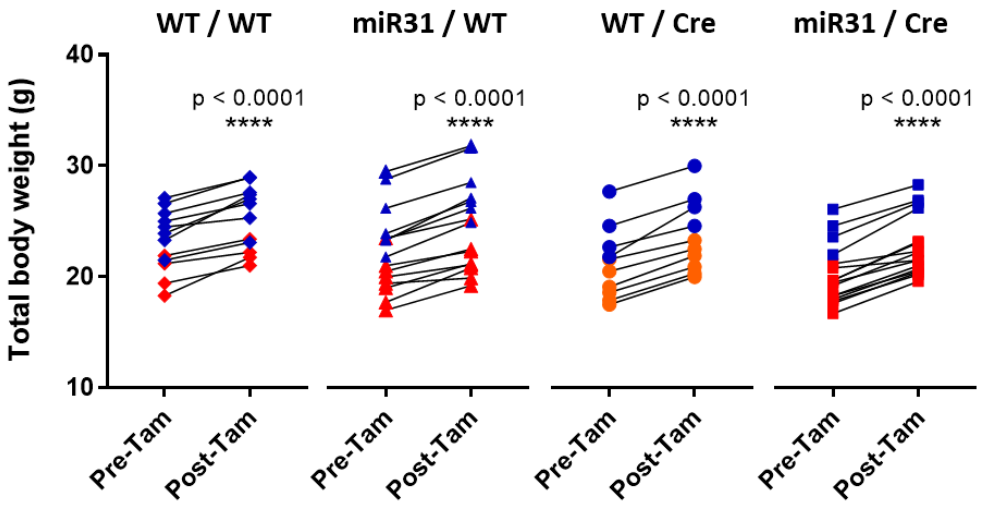


Figure 5.19 ***Before-after plots showing changes in mouse weight during intraperitoneal tamoxifen treatment protocol***

Mice of all four genotype variations were treated at 8 weeks of age with tamoxifen for seven days (2 mg / 30 g mouse / day). Genotype groups from left to right n = 4, 8, 6 and 12 females and 8, 7, 4 and 4 males. Means were compared with paired Student's *t*-test as data was shown to be normally distributed with K-S test.

Having shown no overt and unintended effects of tamoxifen on weight during my experimental procedure, I proceeded to normalise pre- and post-tamoxifen heart dimensions to body weight. Results indicated that tamoxifen had no effect on heart rate in all groups, and that tamoxifen did not cause left-ventricular dilatation or hypertrophy in any of the genotype groups. However results did show that tamoxifen administration significantly impaired ejection fraction ($p = 0.027$) compared to baseline values in wild-type mice and that a similar trend was observed in other genotype groups (*figure 5.20*). Given this small but significant effect, future studies seeking to investigate the effects of transgene expression on ejection fraction and heart muscle contractility should allow for a more prolonged tamoxifen washout period. Indeed a previous study showed no change in ejection fraction in WT C57BL/6 mice after a four week post-tamoxifen washout period when compared to vehicle controls (Lexow et al. 2013). In addition, future experiments using tamoxifen-inducible transgenic mice should consider using genotype controls, as opposed to vehicle controls, given the observation here that tamoxifen in itself depresses cardiac function in mice independent of the presence of Cre or a floxed transgene.

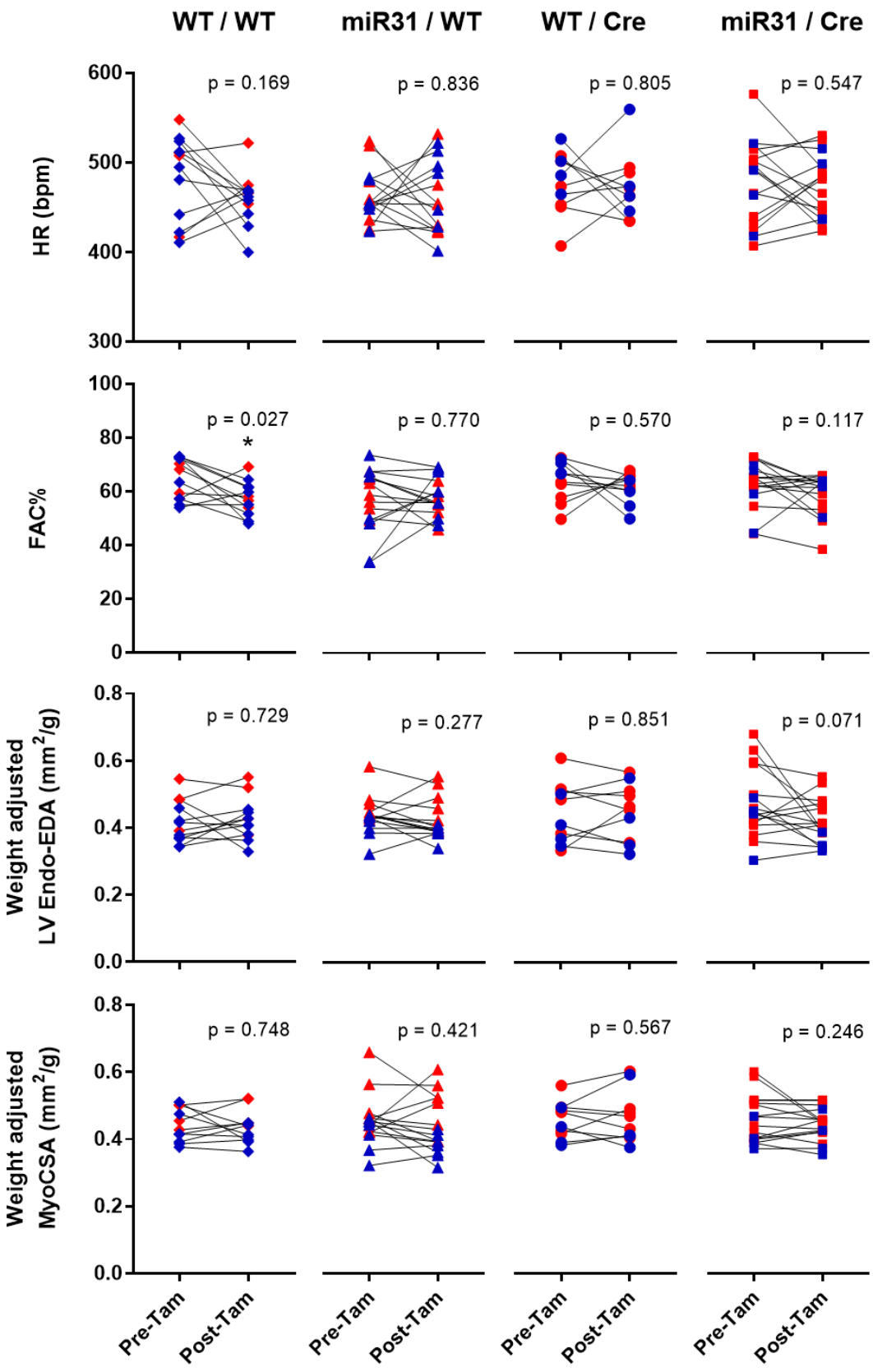

WT / WT
miR31 / WT
WT / Cre
miR31 / Cre
HR (bpm)
600
500
400
300
p = 0.169
p = 0.836
p = 0.805
p = 0.547
FAC%
100
80
60
40
20
0
p = 0.027
*
p = 0.770
p = 0.570
p = 0.117
Weight adjusted LV Endo-EDA (mm²/g)
0.8
0.6
0.4
0.2
0.0
p = 0.729
p = 0.277
p = 0.851
p = 0.071
Weight adjusted MyoCSA (mm²/g)
0.8
0.6
0.4
0.2
0.0
p = 0.748
p = 0.421
p = 0.567
p = 0.246
Pre-Tam
Post-Tam
Pre-Tam
Post-Tam
Pre-Tam
Post-Tam
Pre-Tam
Post-Tam

Figure 5.20 ***Before-after plots showing the effects of tamoxifen on echocardiographic readouts within genotype groups.***

Echocardiography readouts were taken before tamoxifen administration (8 weeks of age) and two weeks following cessation of tamoxifen (11 weeks of age). K-S test determined that heart rate, LV-Endo-EDA and MyoCSA measurements were normally distributed except within the miR31/WT group. For FAC% measurements, data was normally distributed except for the miR31/Cre group. Paired Student's *t*-Test was used for parametric data and paired Wilcoxon Test for non-parametric data. Blue and red symbols depict males and females respectively. Mice with HR <400 bpm omitted. LV-Endo-EDA and MyoCSA data adjusted to total body weight to correct for weight increase between pre- and post-tamoxifen time points. Genotype groups from left to right n = 4, 8, 6 and 12 females and 8, 7, 4 and 4 males.

5.7.4 Increased miR31 expression does not affect systolic heart function

I demonstrated previously that following tamoxifen administration, mice carrying both the miR31 and Cre transgenes show significantly increased atrial and left-ventricular miR31 expression compared to wild types. Conversely, mice carrying the Cre or miR31 transgenes alone did not show increased miR31 (*figure 5.12 above*). In light of this I queried whether miR31 overexpression would affect systolic heart function. Results indicated no statistically significant change between genotype groups in any of the parameters measured (*figure 5.21*). The results of this comparison demonstrate that two weeks following the induction of increased miR31 expression with tamoxifen, miR31 does not affect systolic heart function. However, given that the effects of structural remodelling and fibrosis on heart function are usually observed over months, follow up experiments are needed to conclusively understand the role of miR31 in this experimental system. Further, as tamoxifen depresses systolic heart function within the time-course of this experiment, extending the duration of the experiment is warranted anyway to remove the confounding effects of tamoxifen.

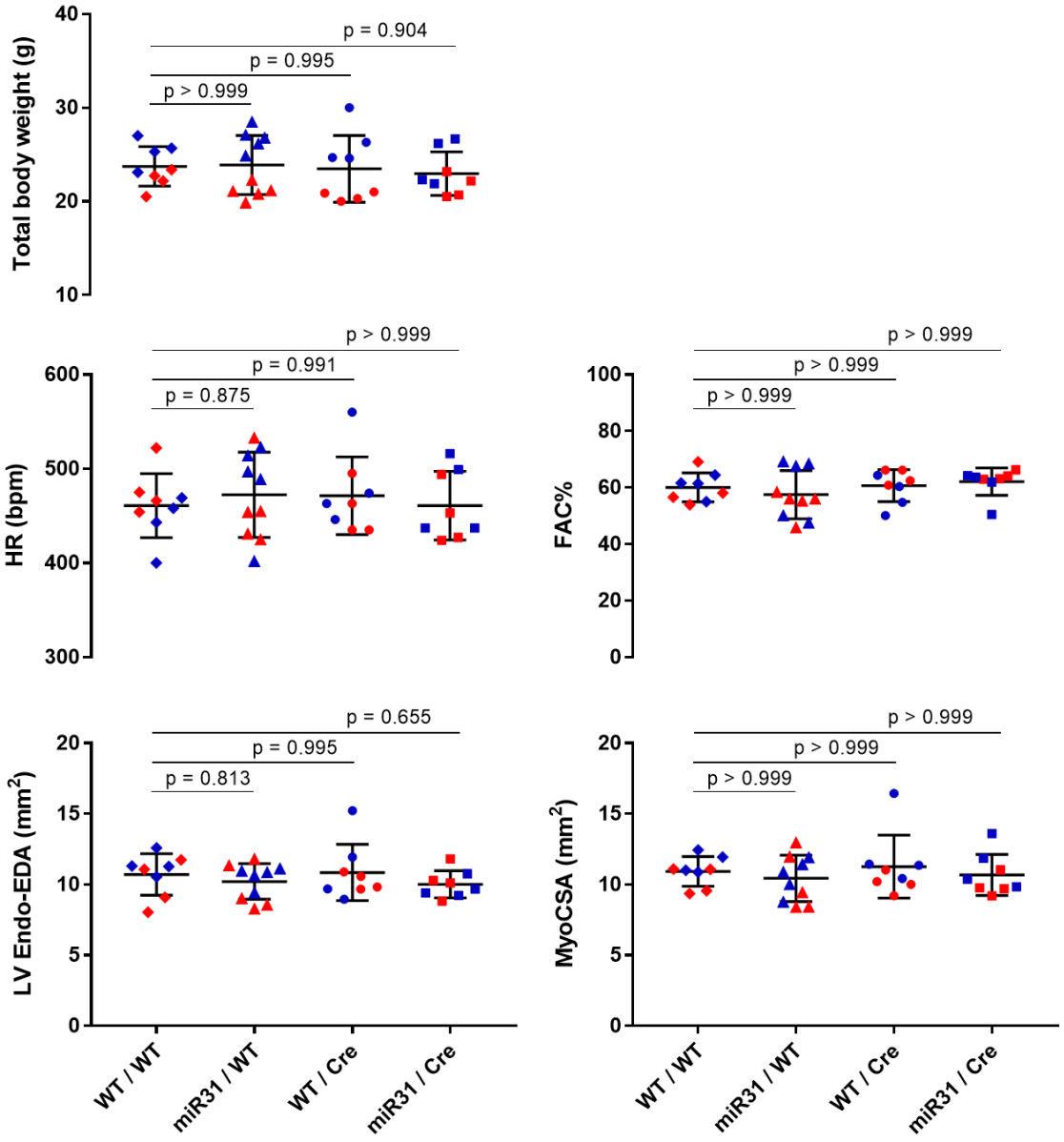


Figure 5.21 ***Echocardiographic parameters post-tamoxifen***

Figure shows the effects of genotype on echocardiographic readouts following one week of tamoxifen administration and a two week post-tamoxifen washout period in mice 11 weeks of age. From left to right genotype group n = 8, 10, 8 and 8. Each genotype group contains equal number of males (blue) and females (red) to balance for gender effects. K-S statistical test determined that all data was normally distributed except for FAC% in the miR31/Cre group and MyoCSA in the *WT/Cre* group. Means of normally distributed data was compared to *WT/WT* control using one-way ANOVA with Dunnett's correction for multiple comparison of means. Means of non-parametric data were compared to *WT/WT* genotype group using Kruskal-Wallis test with Dunn's correction for multiple comparison of means. Mice with HR <400 bpm omitted. Error bars depict mean ± SD.

Overall these echocardiography experiments have provided three pieces of information, firstly that at eight weeks of age, mice carrying the Cre and miR31 transgenes do not show obvious changes in systolic heart function. Secondly, a week of tamoxifen administration followed by a two week washout period can have a depressive effect on heart contractility, and hence ejection fraction. Thirdly, increased fibroblast-specific miR31 abundance does not have an immediate effect on systolic heart function. These findings will allow myself and future scientists to study the consequences of miR31 in this mouse model without the potentially confounding effects of tamoxifen and/or the Cre and miR31 transgenes on cardiac function. For example, future experiments should consider a longer post-tamoxifen washout period, the use of genotype controls as opposed to vehicle controls, and measure changes in systolic and diastolic heart function over a longer time-course. A summary of all echocardiography data generated in these experiments is provided in *table 5.4.*

Table 5.4 Echocardiographic readouts of all mice used in this study expressed as mean ± SD

	WT / WT		**miR31 / WT**		**WT / Cre**		**miR31 / Cre**	
	Pre-Tam	Post-Tam	Pre-Tam	Post-Tam	Pre-Tam	Post-Tam	Pre-Tam	Post-Tam
Females	(n=4)		(n=8)		(n=6)		(n=12)	
Total body weight (g)	20.2 ± 1.65	22.1 ± 1.01	19.8 ± 2.02	21.5 ± 1.84	19.2 ± 1.57	21.5 ± 1.30	18.9 ± 1.40	21.4 ± 1.21
HR (bpm)	496 ± 55.7	479 ± 29.8	469 ± 36.3	452 ± 37.7	465 ± 37.4	464 ± 25.8	481 ± 49.0	472 ± 34.9
FAC%	63.8 ± 6.46	59.5 ± 6.67	60.0 ± 6.43	56.1 ± 5.39	60.5 ± 7.91	64.8 ± 2.66	63.6 ± 8.10	58.8 ± 8.28
LV Endo-EDA (mm^2)	9.33 ± 1.25	9.99 ± 1.72	9.13 ± 1.02	9.71 ± 1.38	9.27 ± 1.94	10.1 ± 1.49	9.63 ± 2.12	9.46 ± 1.32
MyoCSA (mm^2)	9.24 ± 0.72	10.3 ± 0.94	9.61 ± 1.62	10.6 ± 1.65	9.37 ± 1.03	10.23 ± 1.46	9.43 ± 1.40	9.74 ± 0.85
Males	(n=8)		(n=7)		(n=4)		(n=4)	
Total body weight (g)	24.7 ± 1.83	26.9 ± 1.94	25.3 ± 2.96	28.1 ± 2.67	24.2 ± 2.61	27.0 ± 2.25	24.1 ± 1.72	27.0 ± 0.90
HR (bpm)	476 ± 46.1	449 ± 24.3	457 ± 20.5	471 ± 45.8	495 ± 26.1	485 ± 50.8	474 ± 44.2	472 ± 41.3
FAC%	65.0 ± 8.62	56.3 ± 6.33	53.4 ± 16.1	59.9 ± 8.98	69.3 ± 2.72	57.4 ± 6.25	60.5 ± 11.4	60.1 ± 6.5
LV Endo-EDA (mm^2)	9.70 ± 1.01	11.1 ± 1.18	10.2 ± 1.04	10.7 ± 0.63	10.2 ± 1.72	11.5 ± 2.82	10.6 ± 2.03	9.77 ± 0.69
MyoCSA (mm^2)	10.5 ± 1.18	11.4 ± 0.78	10.4 ± 1.33	10.5 ± 1.06	10.7 ± 1.28	12.4 ± 2.72	10.2 ± 1.06	11.4 ± 1.68

5.7.5 Increased miR31 alters mRNA and protein abundance of selected fibrogenesis related genes in mouse atria

Given the observed increase in miR31 expression in *pre-miR31-Inv / Cre* mice (*figure 5.12 above*) I queried whether miR31 was facilitating a change in mRNA transcripts and protein abundance of different pro-fibrotic genes of interest in the atria of these mice. These experiments would serve two purposes, firstly it would confirm that miR31 from the exogenous *pre-miR31-Inv* transgenic construct was acting predictably *in vivo* which would validate our experimental model further. Secondly, this experiment would provide the preliminary evidence that miR31 has a role in fibrogenesis in the intact mouse. Previous studies have demonstrated that carrying a functional Cre recombinase alone can alter gene expression profiles (Semprini et al. 2007, Thyagarajan et al. 2000) and therefore, to ensure that any observable changes in mRNA and protein abundance were specific to increased miR31 expression and not Cre alone, I used *WT / Cre* mice as controls. I first measured the mRNA abundance of five different housekeeping genes that have been reported as reproducibly stable in previous mouse heart experiments (Ruiz-Villalba et al. 2017). Results indicated that actinin alpha 2 (*Actn2)*, ribosomal protein 18S (*Mprs18*), peptidylprolyl isomerase A (*Ppia*) and *Tbp* differed significantly between control and treatment groups while *Hprt1* remained relatively stable and was therefore used as a housekeeping gene for normalisation in forthcoming RT-qPCR experiments (*figure 5.22*).

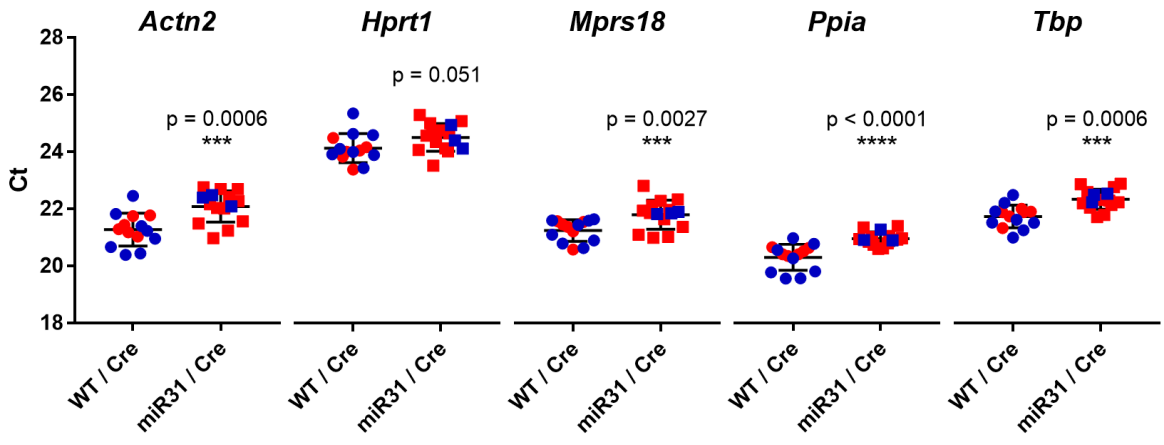


Figure 5.22 ***Raw RT-qPCR values of housekeeping gene transcripts in mouse atria***

Both genotype groups were subject to our seven day intraperitoneal tamoxifen and following a two-week washout period, atrial RNA was extracted and assessed by RT-qPCR for changes in housekeeping gene mRNA abundance. K-S test confirmed data was parametric and Student's *t*-test used for statistical analysis (n = 12 in both groups). Blue symbols indicate males, red females. Error bars depict mean ± SD.

I showed previously using RNA-seq, 4824 and 1144 genes significantly altered in hAFs (adj. $p < 0.05$) following overexpression and depletion of miR31 respectively (*see results section 3.4*). Of these significantly altered genes, 161 genes were thought to be direct targets of miR1 given their simultaneous reduction, and increase, in miR31 mimic and inhibitor datasets respectively (*see results section 3.4.1, figure 3.9*). A further 22 of these were also predicted as direct targets using Targetscan (Release 7.2) of which three genes including BAP1, TNS1 and PKN2 were of interest given their potential roles in fibrosis (see results section *3.4.1, table 3.7*). I therefore measured changes in these gene transcripts following increased atrial miR31 expression in my mouse model. Despite having no documented role in fibrosis, I also measured STK40 expression as a positive control as STK40 is a previously validated miR31 target and showed a robust response in the hAF RNA-seq data. Results indicated that *Stk40*, *Bap1*, *Tns1* and *Pkn2* transcripts were all significantly reduced in the *pre-miR31-Inv / Cre* group compared to *WT / Cre* control, these data therefore corroborated my *in vitro* results in primary hAFs (*figure 5.23A*). Following these qPCR results, I then queried protein abundance of STK40, BAP1 and TNS1 in mouse atria, PKN2 was not pursued in the interest of time. Immunoblotting results showed a congruent and significant reduction in STK40, but no change in BAP1 and a significant *increase* in TNS1. The increase in TNS1 atrial protein was incongruent with the *reduction* in atrial TNS1 mRNA abundance which reiterates the importance of using different methodological approaches when studying biological systems (*figure 5.23B*).

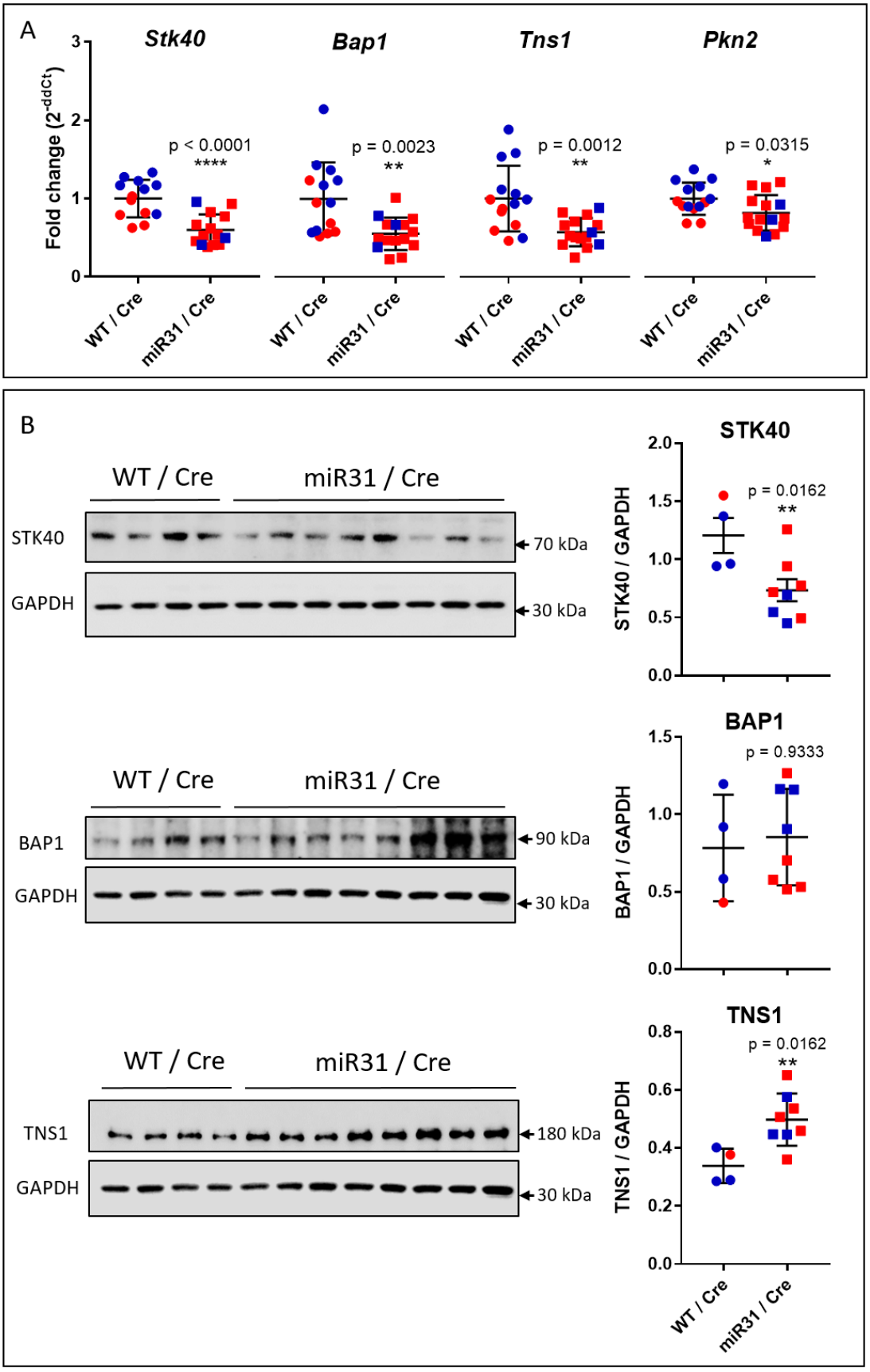

A
Stk40
Bap1
Tns1
Pkn2
Fold change (2-ddCt)
p < 0.0001
****
p = 0.0023
**
p = 0.0012
**
p = 0.0315
*
WT / Cre
miR31 / Cre
B
WT / Cre
miR31 / Cre
STK40
GAPDH
70 kDa
30 kDa
STK40 / GAPDH
p = 0.0162
**
BAP1
90 kDa
BAP1 / GAPDH
p = 0.9333
TNS1
180 kDa
TNS1 / GAPDH
p = 0.0162
**

Figure 5.23 ***Mouse atrial mRNA and protein changes of genes predicted as direct miR31 targets from hAF RNA-seq and Targetscan7.2***

A: mouse atrial RT-qPCR results normalised to HPRT1 and expressed as fold change of *WT/Cre* control. K-S test showed that data was parametric thus unpaired Student's *t*-test was used to compare means (n = 12 in both groups) **B**: mouse atrial immunoblots and graphs of band intensities normalised to GAPDH. Band intensities were determined using ChemiDoc software (version 5.1). K-S test determined that data was non-parametric and therefore groups were compared using Mann–Whitney U test. *WT/Cre* group n = 4, *miR31/Cre* group n = 8. Blue indicate males, red females. Error bars depict mean ± SD.

To investigate the effects of increased mouse atrial miR31 further, I again used the hAF RNA-seq data to identify pro-fibrotic genes that significantly changed with increased miR31 expression, but didn't necessarily change with miR31 inhibition and were not necessarily Targetscan predicted direct targets of miR31. Such genes were still of interest as they were showing a response with increased miR31 in hAFs and could therefore conceivably show a similar response in mouse atria following miR31 overexpression. Fibroblast Growth Factor 1 (FGF1), Vascular Endothelial Growth Factor B (VEGFB), and Integrin beta 3 (ITGB3) were pro-fibrotic genes that *decreased* following increased miR31 in hAFs and these were investigated in mouse atria. Conversely, COL3A1 and TGFβ2, were genes of interest that *increased* with miR31 overexpression in hAFs and these were also queried in mouse atria. FGFs are fibroblast secreted proteins with paracrine and autocrine signalling roles in both physiological and disease states of the heart (Itoh et al. 2013). VEGFB is a ligand for VEGF receptor 1 (VEGFR1), and among other pro-fibrotic roles, VEGF signalling has been shown to enhance TGFβ production in fibroblasts thereby facilitating fibrosis (Zhao et al. 2014) (Zhang et al. 2019). ITGB3 is a membrane receptor for fibronectin, laminin, MMP2 and thrombospondin, and mediates signal transduction and cell-adhesion to the ECM. The role of ITGB3 in cardiac fibrosis is unclear however one study showed significantly reduced ITGB3 expression following inhibition of cardiac fibroblast specific TGFβ-Smad2/3 signalling (Khalil et al. 2017).

Results in mouse atria showed that *Fgf1* and *Vegfb* mRNA was significantly reduced and that *Itgb3* followed a similar trend (*figure 5.24A*). Following this observation, I queried FGF1 atrial protein content with immunoblotting which showed a similar but non-significant trend, in the interest of time VEGFB and ITGB3 immunoblots were not performed (*figure 5.24B*). Further, results indicated a significant increase in *Col3a1* but no change in *Tgf-β2* transcripts (*figure 5.25A*), consequently immunoblotting of COL3A1 was performed which also showed a significant increase (*figure 5.25B*).

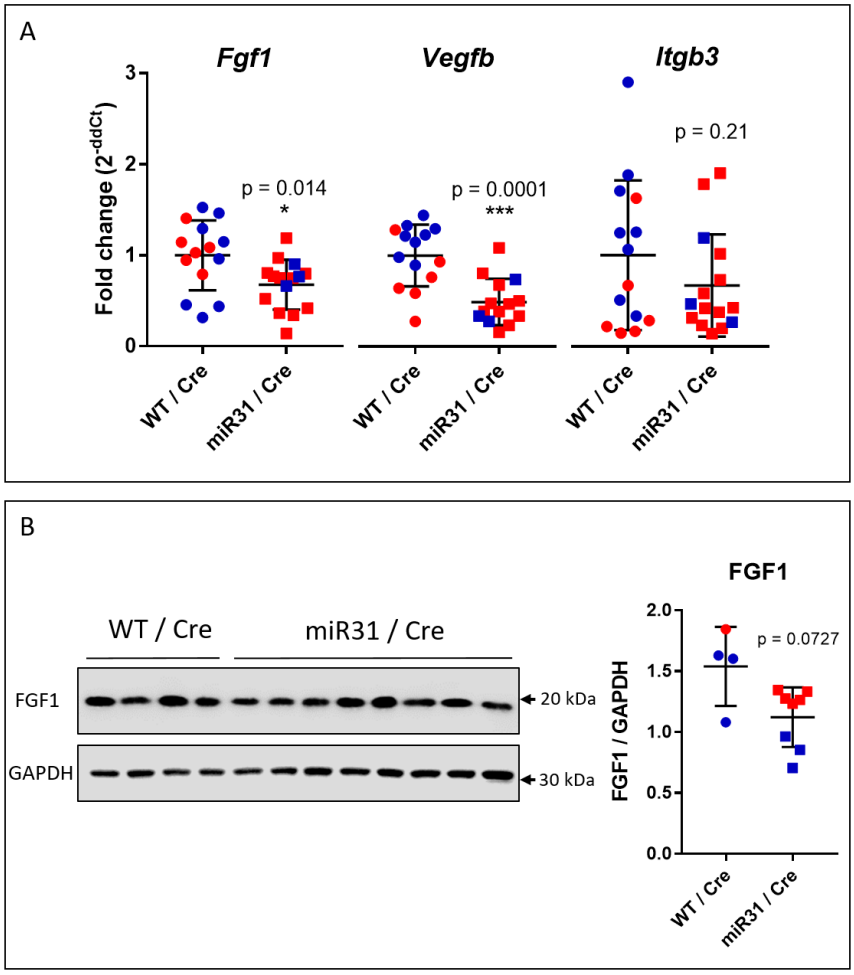


Figure 5.24 ***Mouse atrial mRNA and protein changes of genes that decreased in hAF RNA-seq following increased miR31 but didn't necessarily change after miR31 depletion***

A: mouse atrial RT-qPCR results normalised to HPRT1 and expressed as fold change of *WT/Cre* control. K-S test showed that data was parametric thus unpaired Student's *t*-test was used to compare means (n = 12 in both groups). **B**: mouse atrial immunoblots and graphs of band intensities normalised to GAPDH. Band intensities were determined using ChemiDoc software (version 5.1). K-S test determined that data was non-parametric and therefore groups were compared using Mann–Whitney U test. *WT/Cre* group n = 4, *miR31/Cre* group n = 8. Blue indicate males, red females. Error bars depict mean ± SD.

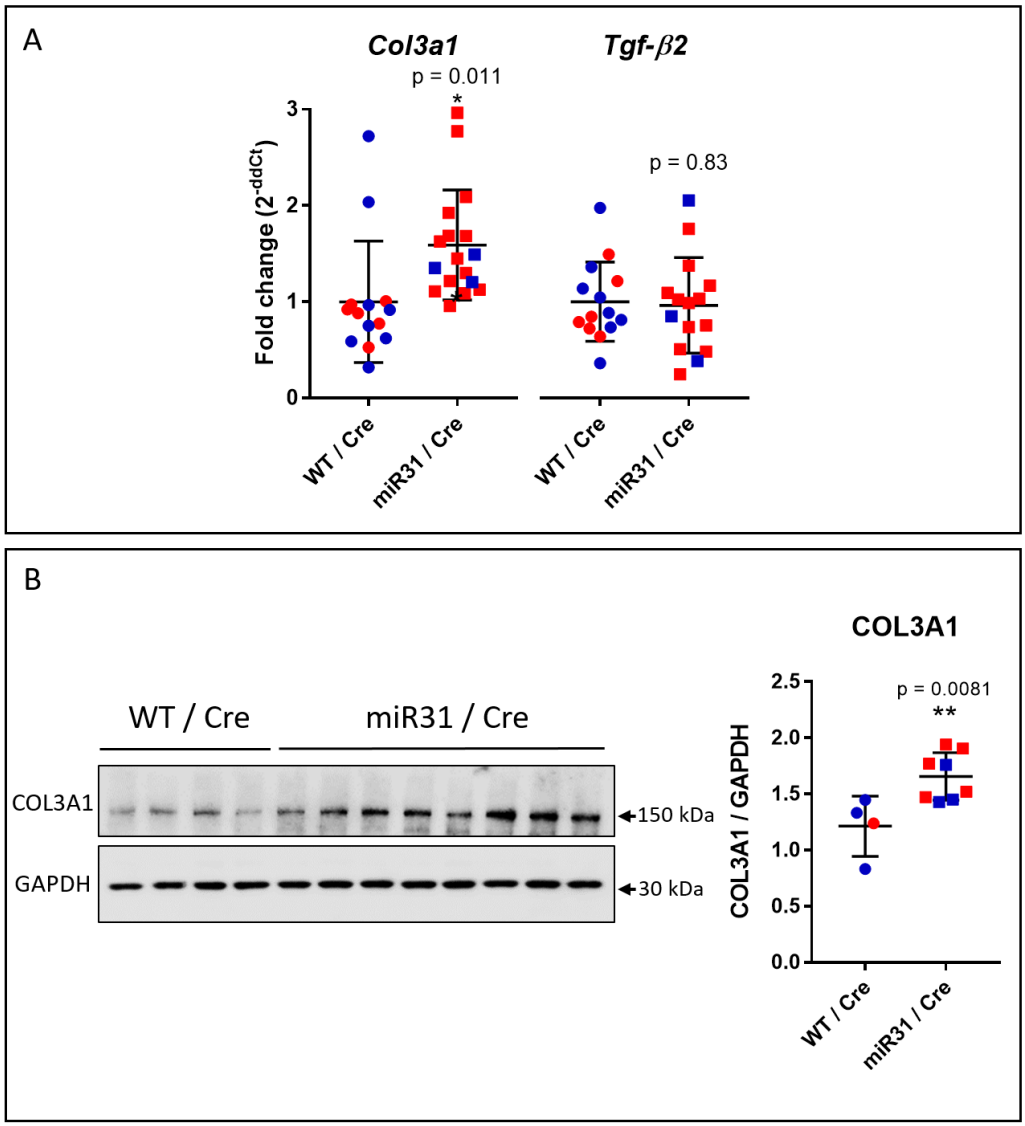


Figure 5.25 ***Mouse atrial mRNA and protein changes of genes that <u>increased</u> in hAF RNA-seq following increased miR31 but didn't necessarily change after miR31 depletion***

A: mouse atrial RT-qPCR results normalised to HPRT1 and expressed as fold change of *WT/Cre* control. K-S test showed that data was parametric thus unpaired Student's *t*-test was used to compare means (n = 12 in both groups). **B**: mouse atrial immunoblots and graphs of band intensities normalised to GAPDH. Band intensities were determined using ChemiDoc software (version 5.1). K-S test determined that data was non-parametric and therefore groups were compared using Mann–Whitney U test. *WT/Cre* group n = 4, *miR31/Cre* group n = 8. Blue indicate males, red females. Error bars depict mean ± SD.

As a further investigation of the effects of increased miR31 on transcript and protein abundance in mouse atria, I looked for changes in candidate pro-fibrotic genes of interest but did not change in hAF RNA-seq. *Calcr, Fn1 and Pdgfb* transcripts were significantly reduced in mouse atria following increased miR31, however *Col1a1*, *Postn, TGFβ1, Ctgf, Mmp2, Mmp9* and *Timp4* showed no change (*figure 5.26*). Our lab has recently shown that disturbed calcitonin (CT) signalling through *Calcr* facilitates atrial fibrosis and atrial fibrillation (Moreira et al. 2020). Thus showing here that increased miR31 abundance reduces *Calcr* mRNA was of particular interest and consequently CALCR protein abundance was checked which also showed significant reduction (*figure 5.27 upper panel*). Despite no evident change in *Col1a1* mRNA, immunoblotting of COL1A1/2 did show a significant increase (*figure 5.27 upper middle panel*). Further no significant change was observed in FN1 immunoblots (*figure 5.27 lower middle panel*). αSMA protein is a known myofibroblast marker and therefore immunoblotting was also performed for αSMA which showed a significant increase (*figure 5.27 lower panel*).

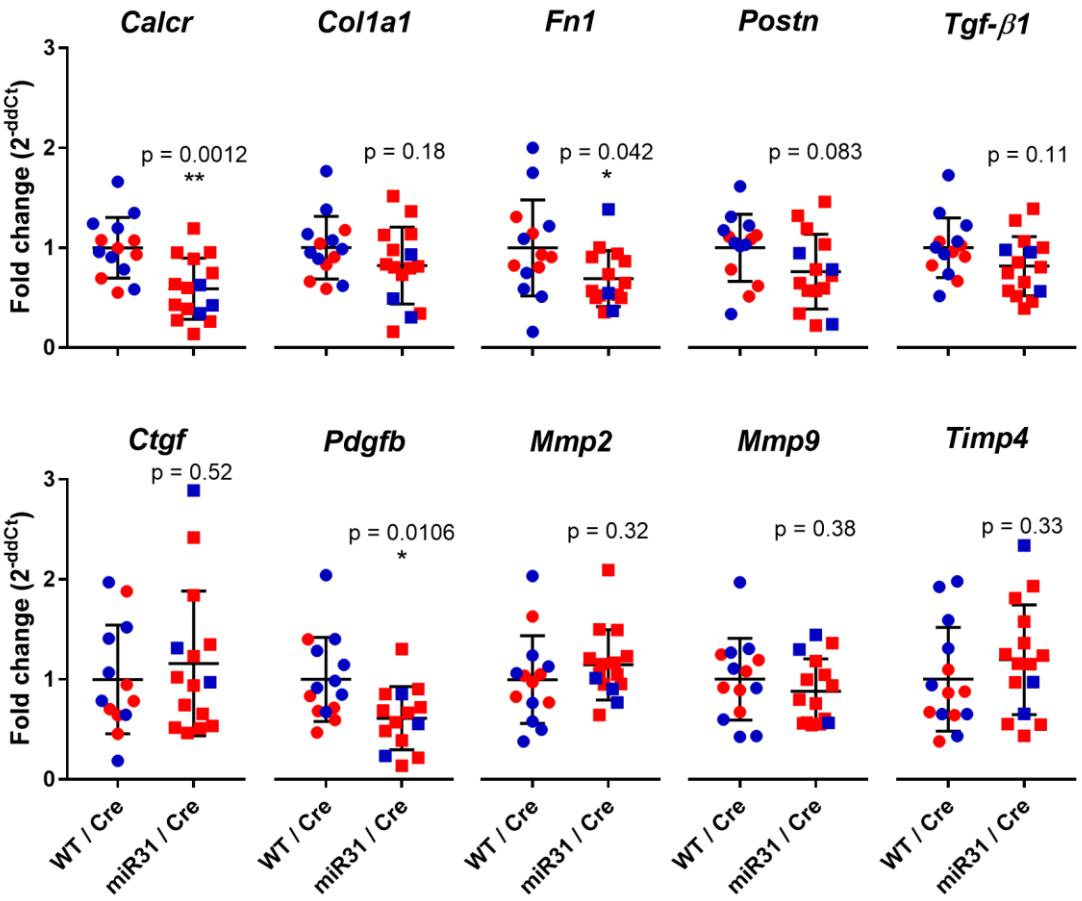


Figure 5.26 mRNA fold changes in mouse atria of candidate pro-fibrotic genes

Mouse atrial RNA was assessed by RT-qPCR for changes in mRNA abundance of various pro-fibrotic genes of interest that did not significantly change in hAF RNA-Seq. Gene Ct values were normalised to HPRT1 and expressed as fold change of *WT/Cre* control. K-S test showed that data was parametric thus unpaired Student's *t*-test was used to compare means (n = 12 in both groups). Error bars depict mean ± SD. Blue symbols indicate males, red indicate females.

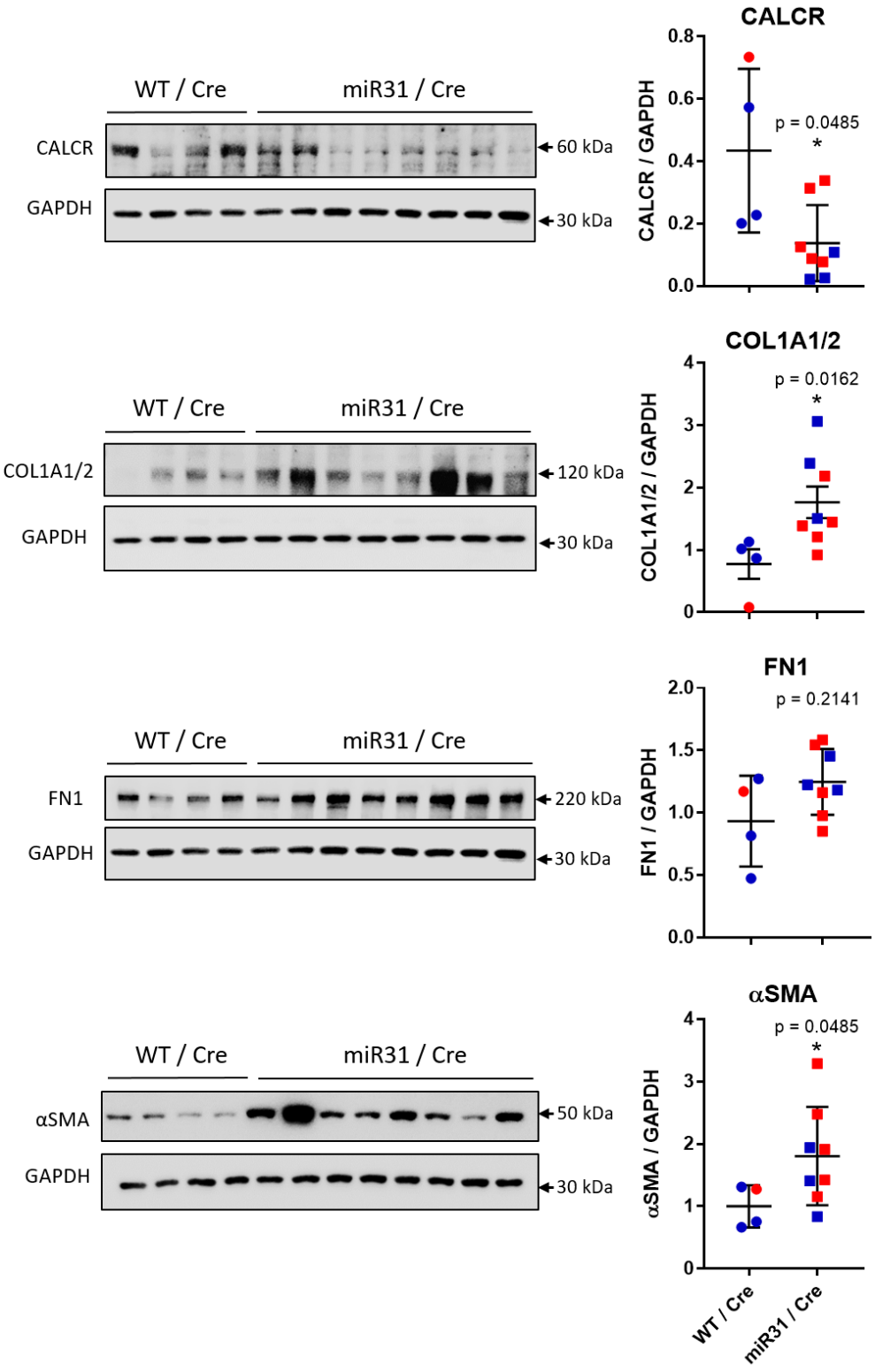


Figure 5.27 ***Protein changes in mouse atria of candidate pro-fibrotic genes***

Atrial protein was assessed by immunoblotting for changes in various pro-fibrotic genes of interest that did not significantly change in hAF RNA-seq. Band intensities were normalised to GAPDH and determined using ChemiDoc software (version 5.1). K-S test showed that the data was non parametric and therefore groups were compared using Mann–Whitney U test. WT/Cre group n = 4, miR31/Cre group n = 8. Blue indicate males, red females. Error bars depict mean ± SD.

In this part of the project, RT-qPCR and immunoblotting was used to query changes in pro-fibrotic genes following increased miR31 abundance in mouse atria. My initial investigation was focused on a select few genes by considering a gene list of bioinformatically predicted miR31 targets generated by Targetscan7.2 as well as considering changes in the hAF cell RNA-seq. I then looked at genes that had not been identified in my RNA-seq but were still of interest given their role in fibrosis. A summary of all these results is given in *Table 5.5* below.

Table 5.5 Summary of all genes explored in mouse atria and relevant hAF RNA-seq data

Gene	miR31 mimic hAF RNA-seq	miR31 inhib hAF RNA-seq	miR31 target (Targetscan)	Mouse Atria RNA	Mouse Atria Protein
STK40	↓	↑	Yes	↓	↓
BAP1	↓	↑	Yes	↓	—
TNS1	↓	↑	Yes	↓	↑
PKN2	↓	↑	Yes	↓	NA
VEGFB	↓	↑	No	↓	NA
ITGB3	↓	↑	No	—	NA
FGF1	↓	—	No	↓	—
COL3A1	↑	↑	No	↑	↑
TGFβ2	↑	—	No	—	NA
CALCR	—	—	Yes	↓	↓
FN1	—	—	No	↓	—
PDGFB	—	—	No	↓	NA
αSMA	—	—	No	NA	↑
COL1A1	—	—	No	—	↑
POSTN	—	—	No	—	NA
TGFβ1	—	—	No	—	NA
CTGF	—	—	No	—	NA
MMP2	—	—	No	—	NA
MMP9	—	—	No	—	NA
TIMP4	—	—	No	—	NA

Significant increase (↑), significant decrease (↓), no change (—), not measured (NA)

Overall, these experiments provide further evidence that our newly generated mouse model behaves as intended and that increased miR31 expression alters mRNA and protein abundance of different genes involved in fibrosis *in vivo.* Further, these results showed many congruent mRNA transcript changes following increased miR31 between the hAF *in vitro* model and the *in vivo* mouse model (e.g. *Stk40, Bap1, Tns1, Pkn2, Vegfb, Fgf1 and Col3a1*) thereby validating these separate experimental systems. There were however incongruent findings between the two experimental systems, for instance *Itgb3 and Tgf-β2* significantly changed in hAFs but didn't change in mouse atria following overexpression of miR31. Conversely, *Calcr, Fn1 and Pdgfb* significantly changed in mouse atria but showed no change in hAFs. These incongruences can be explained by three factors. First, the difference in models (*in vitro* human vs. *in vivo* mouse). Second, the large difference between miR31 abundance in these models - in hAFs miR31 expression was ~400-fold greater than controls, while in mouse atria miR31 was ~3-fold greater than controls. Third, my data from mouse atria are derived from RNA extracted from all atrial cell types including cardiomyocytes, fibroblasts and endothelial cells as I did not sort fibroblasts from mouse atria before extracting RNA. Therefore comparing RNA derived from a *pure* population of hAFs would predictably show larger differences.

5.7.6 Increased miR31 abundance results in elevated total atrial collagen composition

The former immunoblotting experiments showed a significant increase in both Col1a1/2 and Col3a1 atrial protein following increased miR31 expression in our animal model. Collagen is the main structural protein of the extracellular matrix in mouse atria and ~13.5% of collagen's amino acid content is hydroxyproline. Due to the highly restricted distribution of hydroxyproline in collagen, hydroxyproline content in tissues can be measured as a surrogate for collagen. Thus I sought to measure total hydroxyproline content from heart tissues extracted from our mouse model using a colorimetric based assay (Chondrex: 6017). I chose to first query changes in LV collagen production as LV tissue mass is significantly higher than atria and therefore more dispensable. Following optimisation in LV, atrial collagen content was then measured. Results showed no significant change in LV collagen content, however there was an increase in atrial collagen in our *miR31 / Cre* mice compared to *WT / Cre* controls (*figure 5.28*). This change was not statistically significant with the number of biological replicates used, however, given that these data indicate a trend, and that this is an important finding, this experiment will be repeated with increased sample size.

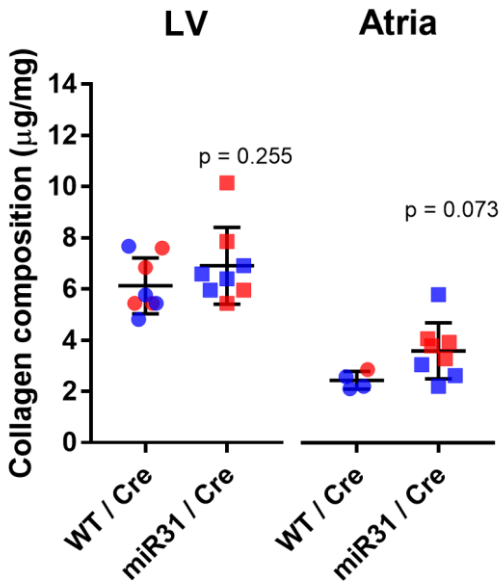


Figure 5.28 ***LV and atria collagen content***

LV and atrial hydroxyproline was purified and measured to estimate collagen tissue composition (µg collagen per mg tissue) in *miR31 / Cre* and compared to *WT / Cre* control following IP tamoxifen and two week washout period. Unpaired Student's *t*-test used for statistics. Blue indicate males, red females. Error bars depict mean ± SD.

5.8 Discussion and plans for future experiments

Results in this chapter first showed that in mice carrying both the *Cre* and *miR31* transgenes, tamoxifen successfully induced Cre-dependent inversion of the loxP flanked *pre-miR31-Inv* allele, leading to increased miR31 expression in mouse LV and atria. This Cre-dependent inversion switch has been employed before with larger gene inserts, however, our mouse is the first in which this has been achieved with microRNAs *in vivo*. Echocardiography results then showed that there were no overt changes in LV structure and function, specifically no hypertrophy or dilatation, as well as no change in ejection fraction, in mice before miR31 induction (8 weeks of age), or after miR31 induction (11 weeks of age). These parameters were looked at as they are widely used to screen for obvious changes in cardiac phenotype (Lindsey et al. 2018). Following these experiments I queried whether increased fibroblast specific miR31 affected the molecular makeup of mouse atria, again in 11-week old mice, by looking at mRNA and protein changes of key genes given their pre-defined roles in atrial fibrosis. This approach revealed a significant increase in mRNA and protein content of Col3a1 in mouse atria following increased miR31 expression, similarly there was a significant increase in Col1a1 protein and consequently when total atrial collagen was measured in these mice using a hydroxyproline assay, this showed a borderline increase.

Experiments are ongoing to increase the sample size of the hydroxyproline collagen assay to definitively measure total collagen content in mouse atria at this age, furthermore we have samples of mouse hearts embedded in paraffin and will assess crude cardiac fibrosis and ECM deposition with picrosirius red and Masson's trichrome staining. In addition to studying ECM and collagen content, when atrial αSMA protein was measured with immunoblotting, which is well known to increase during fibroblast differentiation into pro-fibrotic myofibroblasts, this also showed a significant increase in miR31 transgenic mice. In light of these molecular based findings, future experiments will isolate both atrial fibroblasts and myocytes from these mice. MiR31, Col3a1, Col1a1 and αSMA mRNA and protein expression could then be measured in these cells separately, this information would be used to demonstrate that the molecular changes observed in total mouse atria are likely being driven by miR31 abundance in atrial fibroblasts, and not other resident cell types.

Given that the effects of molecular changes and fibrosis on heart function, in response to increased miR31, would likely develop slowly over an extended duration of time, future experiments could investigate the wider phenotypic outcomes of these observed molecular changes in aged mice. For example, it would be useful to obtain echocardiography measurements in mice at 18 months of age which is considered equivalent to 62 years for humans and when age-related changes are more easily detectable (Dutta et al. 2016). These could be combined with LV catheterisation studies to directly measure ventricular filling pressures and parameters of ventricular relaxation to produce a more comprehensive *in vivo* phenotype. For example, increases in LV end-diastolic pressure or tau, the time constant for isovolumic relaxation, would be indicative of diastolic dysfunction. In addition, simply checking whether there are any changes in systemic blood pressure would be useful, since this could drive structural changes in the heart via secondary means. One might also consider whether spontaneous arrhythmias develop over time and this could be measured using implantable ECG radiotelemetry.

As an alternative to looking at the long term effects of increased systemic miR31 on heart structure and function, future experiments could look at the susceptibility to cardiac disease in younger mice using interventions known to induce cardiac dysfunction. This would help reduce the husbandry costs associated with longer animal experiments. For example, it could be determined whether miR31 transgenic mice are more susceptible to fibrosis, measured with histological and molecular based techniques including qPCR and immunoblotting, when subject to pressure overload by aortic banding or treatment with isoproterenol or angiotensin-II osmotic minipumps. Echocardiography could then be performed to assess ventricular geometry and function at 4 and 12 weeks after induction of pressure overload. The susceptibility to AF induction, as well as AF duration, could also be investigated using trans-oesophageal burst pacing in these mice. In addition, the delivery of miR31 inhibitors in WT mice subject to the same aforementioned experimental procedures could be used to corroborate our finding in the transgenic miR31 model. These results would help us understand the potential pharmacological efficacy of inhibiting miR31 in patients with AF.

Chapter 6

Discussion

Cardiac fibrosis is characterised by the excessive accumulation of interstitial ECM and is seen in a number of cardiovascular diseases including atrial fibrillation. The main aims of this work were to: (i) investigate the impact of manipulating miR31 abundance in human atrial fibroblasts *in vitro* on pro-fibrotic gene expression and phenotype; (ii) determine whether the transcription factor ETS1 has a role in regulating miR31 abundance by binding at the LOC554202 promoter which encodes miR31; (iii) explore whether tamoxifen inducible increased fibroblast specific miR31 expression in a novel transgenic mouse strain is sufficient to alter pro-fibrotic milieu of the atria and ventricles.

The first aim was approached by using human atrial fibroblasts obtained from healthy donors that were verified by positive immunofluorescence staining of intracellular procollagen-1a1 and vimentin as well as surface expression of CD90 which was determined using FACS. In addition, FACS confirmed that these cells did not express surface PDGFRα which has been shown to increase in fibroblasts as they differentiate into disease state myofibroblasts (Chong et al. 2013, Ivey et al. 2016), suggesting that these cells were healthy and viable for experimentation. The role of miR31 in these cells was queried by either inhibiting or enriching miR31, with miR31 mimic or inhibitor oligonucleotides. Changes in miR31 abundance were first confirmed using RT-qPCR and then the poly-A tail mRNA transcriptome was sequenced which showed significant changes in transcript abundance of genes involved in a vast array of cellular processes. The observation that simply manipulating miR31 had such profound effects on gene mRNA abundance was in itself an important finding given that fibroblast differentiation involves drastic changes in cell behaviour (motility, proliferation and ECM turnover), and thus any regulators of this process would need to show far reaching effects on gene expression.

MicroRNAs elicit their influence through sequencing complementarity between their internal seed sequence and binding sites on mRNA transcripts. Given that the microRNA seed sequence length is relatively short (7 nucleotides), there is high probability for site matching through the genome, this is compounded by the fact that targeting can be achieved without exact sequence homology. The wide reach of individual microRNAs has been supported with experiments that demonstrate that microRNA cell signatures inversely correlate with their target mRNA abundance and imply that the expression of any given microRNA is fine-tuned to maintain or drive cell specific behaviour. In my experiments, transfection of hAF cells with miR31 mimic resulted in supra-physiological abundance of miR31 and future experiments may wish to more carefully titrate miR31 transfection to bring it closer to endogenous levels. Saturating hAF cells with miR31 would have certainly resulted in off target consequences, for instance by increasing chance binding to non-homologous mRNA transcripts, as well as excessively occupying the RISC complex thereby out-competing other cellular microRNAs. Indeed the miR31 mimic RNA-seq experiment showed changes in transcript abundance of close to five thousand genes which far out-weighs any reasonable prediction of miR31 behaviour and does not accurately reflect the role of increased miR31 as seen in a natural disease state like AF.

In light of this caveat, all experiments I performed which involved the overexpression of miR31 in hAFs were done so in tandem with experiments in which endogenous miR31 was inhibited with its reverse complementary oligonucleotide. Furthermore, when each RNA-seq dataset was analysed for pathway enrichment with IPA, I applied a very stringent cut off (adj. $p < 0.00000001$) to the mimic dataset in an attempt to focus on biologically meaningful targets and this whittled the list down to 861 genes. Conversely, the usual cut-off was applied to the inhibitor dataset (adj. $p < 0.05$) which gave a list of 1144 genes. When these lists were analysed with IPA, the miR31 mimic dataset showed a significant *increase* in pro-fibrotic pathways including TGFβ, ERK/MAPK and FGF signalling, cardiac hypertrophy, hepatic fibrosis, renin angiotensin signalling and integrin signalling. Conversely, in the inhibitor dataset, there was a *reduction* of enrichment in TGFβ, BMP, PDGF, hepatic fibrosis and integrin signalling. These data supported our hypothesis that increased miR31 abundance in human atrial fibroblasts drives pro-fibrotic gene expression changes in these cells.

I also manually screened these lists for different genes of interest given their pre-described role in fibrotic signalling and/or use as markers for fibroblast differentiation (this was done with a more relaxed cut off, adj $p < 0.05$ for both datasets). Of note, increased miR31 showed an increase in TβR1 and TGFβ2 which would facilitate canonical TGFβ activity, as well as an increase in COL3A1, COL4A5 and COL14A1. Conversely, inhibition of miR31 showed a reduction in AT1R which is the primary receptor for angiotensin II. Neither dataset showed any changes in markers frequently used to monitor fibroblast differentiation including CD90, PDGFRα, Periostin, VAP1, αSMA and Vimentin, however miR31 mimic did show changes in SCA1 and TBX18 which are implicated in fibroblast development. The lack of significant change in these usual markers may be explained by the experimental design which only looked at changes at a single time point - 48 hours following transfection with miR31 mimic and inhibitor - thereby focusing on the immediate role of miR31 and not its long term effects. Furthermore mRNA abundance does not always reflect protein abundance, and therefore follow up immunoblotting experiments would be justified to check the protein presence of these usual markers.

To validate the RNA-seq data, I narrowed the investigation to genes that showed significant but *opposing* changes in both datasets. This was done as by definition they were genes showing the most reproducible behaviour with miR31 manipulation and could therefore be easily used for validation experiments. I only looked at genes that *decreased* with miR31 and *increased* with miR31 inhibition, as these were more likely direct targets of miR31. Following this approach, I used qPCR to show congruent changes in STK40, BAP1, TNS1 and PKN2 thereby validating the sequencing data. BAP1 and PKN2 were of interest given their roles in promoting cell motility by rearranging cytoskeletal protein (Lachmann et al. 2011, Artegiani et al. 2019), while TNS1 was of interest given that one study showed TNS1 was essential for myofibroblast differentiation in pulmonary fibroblasts (Bernau et al. 2017). Despite being able to validate these genes with qPCR, their protein expression was not determined and future experiments may wish to complete this to fully explore the role of these genes in the fibrotic response.

Having shown significant mRNA enrichment in genes involved in major pro-fibrotic signalling axes, I then went on to study the effects of manipulating miR31 on fibroblast phenotype. Results of initial experiments using cells from donor 1 (which was the donor also used for the RNA-seq) showed a significant increase in hAF proliferation following treatment with miR31 mimic, conversely hAF proliferation was impeded following miR31 inhibition. However, when this experiment was repeated in hAFs from donor 2, no discernible change in proliferation was observed. Similarly, when collagen production was measured following manipulation of miR31 in donor 2 hAFs no change was observed. The inter-donor variability may be explained by the fact that the donors were not age and sex matched. The first donor was a 50 year old female, the second was as 61 year old male. Furthermore, the variability of sensitivity to miR31 manipulation between the two donors may be understood when considering that hAFs from donor 2 showed no response to TGFβ1 and had presumably reached a saturation point in their pro-fibrotic potential. These results reiterated the importance of using positive controls in experimental design and all future *in vitro* assays with different donor cells should utilise the response to TGFβ1 as a standardized point of reference.

To understand the role of ETS1 in hAFs, which was my second experimental aim, I conducted experiments whereby ETS1 was knocked down in hAFs with chemically modified anti-sense oligonucleotides called Gapmers. Results showed that inhibiting ETS1 elevated cellular miR31 abundance, however this result was not statistically significant. Upon inspection of the data, the raw Ct value of miR31 indicated that miR31 had indeed significantly increased, but this observation was being washed out by variability in RNU58A abundance which was the housekeeping microRNA used to normalise miR31 expression. Consequently this experiment could have been improved by identifying more stable microRNAs to which miR31 expression could be compared. Regardless, I also conducted experiments whereby ETS1 protein was enriched in hAF cells using a plasmid expression vector. Remarkably this experiment resulted in an *increase* in miR31, and not the predicted decrease in miR31 abundance. This result may reflect the fact that ETS1 itself is not known to inhibit or repress transcription, but is thought rather to simply bind DNA to facilitate the role of other factors like EMSY and KDM5B (Vire et al. 2014). Consequently, increased abundance of ETS1 would likely have a range of transcriptional effects - this would have been further compounded by the fact that ETS1 plasmid resulted in supra-physiological ETS1 abundance.

Furthermore, it was noted that transfection of hAFs with ETS1 expression vector resulted in the detection of two bands, one of identical molecular weight to endogenous ETS1 and one heavier band which was believed to be either phosphorylated or acetylated ETS1. ETS1 is inhibited by phosphorylation at residues Ser282 and 285 (Fleischman et al. 1995, Vetter et al. 2005), and acetylation of Lys8 and Lys18 is known to inhibit the DNA binding activity of ETS1 (Chen et al. 2017). Consequently future experiments could consider investigating ETS1 phosphorylation and acetylation status in hAFs and this may help explain the apparent increase in miR31 with ETS1 overexpression.

To further explore the role of ETS1 on miR31 abundance, I commenced experiments to understand the functional relevance of eleven putative ETS1 binding sites at the LOC554202 promoter using a dual glow firefly/ renilla reporter assay. Firefly luciferase constructs were designed, each carrying the WT LOC554202 promoter, but with sequential mutation of the predicted ETS1 binding sites. These constructs were to be introduced into HEK293 cells, together with ETS1 expression vector, to visualise the effects of retarding ETS1 binding sites on LOC554202 transcription. Results did show successful expression of ETS1 from the vector construct in HEK293 cells however the effects on LOC554202 transcription were not determined due to time constraints. Future experiments will finalise this assay and will validate ETS1 recruitment to the LOC554202 promoter in hAF cells using chromatin immunoprecipitation followed by qPCR with primers flanking the identified ETS1 binding sites.

For the third experimental aim, which sought to understand the role of fibroblast specific miR31 overexpression *in vivo*, we utilised the Col1a2-Cre driver line which expresses tamoxifen inducible Cre in all mouse fibroblasts. The Cre line was used in conjunction with a transgenic mouse developed in collaboration with the Wellcome Trust's transgenic core facility in Oxford. They generated a transgenic construct which contained the mouse *pre-miR31* sequence, cloned upstream of *eGFP* cDNA. This sequence was flanked by loxP variants which would later be used for targeting by Cre. The entire construct was inserted, in the reverse orientation, at the *Gt(ROSA26)Sor* locus which had been previously genetically modified to accept construct delivery by PhiC31 integrase (Dolatshad et al. 2015). This approach meant that only a single copy of the transgenic construct was integrated into the genome thereby facilitating more

control over gene dosage. The inserted construct is flanked by two sets of loxP variants and this enabled the use of Cre to flip the construct linking it to the CAG promoter to drive miR31 expression. This Cre-dependent inversion switch has been employed before with larger gene inserts however our mouse is the first in which this has been achieved with microRNAs *in vivo*. Given the novelty of this mouse, much of my attention was focused on carefully validating the molecular processes of construct inversion, recombination and trapping. Results of these experiments successfully showed the predicted behaviour of the transgenic construct *in vivo* and therefore this finding in itself will be of interest to current fields studying microRNAs in mouse development and/or disease models.

With the transgenic mouse model optimised, I went on to demonstrate that simply inheriting and carrying either the Cre or miR31 transgene, or both together, had no impact on mouse body weight and did not cause overt visible anomalies. Importantly, at eight weeks of age when baseline measurements were made, mouse hearts that were visualised echocardiography, appeared to have no changes in heart structure (specifically no hypertrophy and no dilatation), and when ejection fraction was estimated from these measurements, again there was no observable change. These parameters were looked at as they are widely used to assess systolic heart function and therefore served as metrics to screen for obvious changes in LV structure and function. Pulse wave and tissue Doppler echocardiography were also performed to assess ventricular filling and hence diastolic function, however these images were not analysed since it was not possible to obtain a full data-set in all animals and at that stage of the project I was simply screening for overt unintended consequences of genotype on cardiac function. This was further reasoned by the fact that diastolic images require artificially low heart rates and are prone to angulation error, so it was unlikely that the young and otherwise healthy mice used would exhibit a quantifiable phenotype anyway.

Having shown no obvious structural and functional change in these transgenic mice, I then queried whether the use of tamoxifen in my experimental design would affect the structural and systolic functional echocardiography readouts. Using a time course experiment I had previously determined that seven days of tamoxifen administration induced Cre activity and subsequent miR31 expression, however it was unknown as to

whether tamoxifen itself would affect heart function and potentially confound the interpretation of echocardiographic data. Using a before-after paired analysis of pre- and post-tamoxifen echocardiography data, I did find that, independent of mouse genotype, tamoxifen had a small but significant depressive effect on ejection fraction, and consequently future studies may wish to prolong the tamoxifen washout period to avoid this. However, given that I was using genotype controls (i.e. all genotypes were receiving tamoxifen), I could still assess the data while accepting a blanket effect of tamoxifen across all my experimental groups. When I subsequently compared the genotypes post tamoxifen, no change in heart structure or function was observed. This indicated that there was no acute effect of systemic miR31 overexpression on systolic heart structure and function.

Having completed baseline phenotyping in these transgenic mice, I then went on to query whether miR31 overexpression was causing pro-fibrotic molecular changes in mouse atria. When doing this I initially utilised the hAF cell sequencing results, in combination with online tools that predict microRNA targets, to focus on gene transcripts likely to be directly inhibited by miR31. This approach showed a congruent reduction in atrial mRNA of STK40, BAP1, TNS1 and PKN2 which validated these separate experimental approaches and systems. When protein was measured, STK40 did show a reduction, no change in BAP1 was observed, and TNS1 increased. These data again reiterated the importance of using different experimental approaches to validate molecular changes. I also looked at mRNA and protein changes of key genes of interest given their pre-defined roles in atrial fibrosis, but which hadn't necessarily changed in the sequencing data. This approach revealed that both mRNA and protein of Col3a1 increased in mouse atria following increased miR31 expression. Similarly there was an observed increase in Col1a1 protein and consequently when total atrial collagen was measured in these mice, this too showed an increase. Similarly when atrial αSMA protein content was measured, which is well known to increase during fibroblast differentiation into pro-fibrotic myofibroblasts, this also showed an increase.

Our lab has just recently shown that disrupting calcitonin receptor (CALCR) signalling in mice causes atrial fibrosis and that atrial-specific knockdown of calcitonin promotes atrial fibrosis (Moreira et al. 2020). CALCR mRNA carries predicted binding sites for miR31 and consequently I checked atrial CALCR mRNA and protein abundance following increased miR31. This investigation did indeed show reduced CALCR mRNA and protein, which pointed to miR31 as an upstream regulator of this novel pro-fibrotic signalling mechanism. Alongside the changes observed in atrial collagen, this finding may uncover miR31 as a potential target for therapeutic manipulation in AF. This is encouraged by the fact that increased miR31 in cardiomyocytes facilitates pro-arrhythmic electoral remodelling in these cells (Reilly et al. 2016). Given these observed molecular changes, future experiments using these mice have been planned and are discussed in Chapter 5 (*section 5.8 discussion and plans for future experiments*).

6.1 Conclusion

Through experiments conducted in this project, I make the following conclusions. Firstly, increased miR31 abundance in atrial fibroblasts seen in patients with sustained atrial fibrillation, likely has a direct role in promoting pro-fibrotic gene expression changes. This was demonstrated by the fact that increasing miR31 in hAFs *in vitro* promoted pro-fibrotic gene expression changes, while inhibiting miR31 had an opposing effect. Secondly, ETS1 seems to be involved in the upstream regulation of miR31 transcription and consequently the observed reduction of atrial ETS1 seen in AF patients may explain the upregulation of miR31 in atrial fibroblasts. Thirdly, animal models can be created to induce fibroblast specific expression of microRNAs, and in this project the fibroblast specific induction of miR31 significantly altered the mRNA and protein abundance of previously defined, as well as novel, markers of atrial fibrosis. These are all publishable findings and we hope that they may be utilised someday to improve therapeutic outcomes in patients suffering with AF. Currently however, these findings cannot comment on the extended and long term consequences of miR31 in hAFs *in vitro*, or in mice *in vivo*, but given the significant molecular changes observed in these separate experimental systems, future investigations, that are indeed currently underway, are warranted.

References

Acharya A, Baek T, Banfi S, Eskiocak B, Tallquist M (2011) Efficient inducible Cre-mediated recombination in Tcf21 cell lineages in the heart and kidney. Genesis 49:870-877

Acharya A, Baek S, Huang G, Eskiocak B, Goetsch S, Sung C, Banfi S, Sauer M, Olsen G, Duffield J (2012) The bHLH transcription factor Tcf21 is required for lineage-specific EMT of cardiac fibroblast progenitors. Development 139:2139-2149

Adam O, Lohfelm B, Thum T, Gupta S, Puhl S, Schafers H, Bohm M, Laufs U (2012) Role of miR-21 in the pathogenesis of atrial fibrosis. Basic Research in Cardiology 107:278-290

Agarwal V, Bell G, Nam J, Bartel D (2015) Predicting effective microRNA target sites in mammalian mRNAs. Elife 4:e05005

Ali R, Ranjbarvaziri S, Talkhabi M, Zhao P, Subat A, Hojjat A, Kamran P, Müller S, Volz S, Tang Z, Red-Horse K, Ardehali R (2014) Developmental heterogeneity of cardiac fibroblasts does not predict pathological proliferation and activation. Circulation Research 115:625-635

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

Andersson T, Magnuson A, Bryngelsson I, Frøbert O, Henriksson K, Edvardsson N, Poçi D (2013) All-cause mortality in 272,186 patients hospitalized with incident atrial fibrillation 1995-2008: a Swedish nationwide long-term case-control study. European Heart Journal 34:1061-1067

Andrade J, Khairy P, Dobrev D, Nattel S (2014) The clinical profile and pathophysiology of atrial fibrillation: relationships among clinical features, epidemiology, and mechanisms. Circulation Research 114:1453-1468

Annes J, Munger J, Rifkin D (2003) Making sense of latent TGFβ activation. Journal of Cell Science 116:217-224

Armendariz-Borunda J, Katayama K, Seyer J (1992) Transcriptional mechanisms of type I collagen gene expression are differentially regulated by interleukin-1 beta, tumor necrosis factor alpha, and transforming growth factor beta in Ito cells. The Journal of Biological Chemistry 267:14316-14321

Artegiani B, van Voorthuijsen L, Lindeboom R, Seinstra D, Heo I, Tapia P, Lopez-Iglesias C, Postrach D, Dayton T, Oka R, Hu H, van Boxtel R, van Es J, Offerhaus J, Peters P, van

Rheenen J, Vermeulen M, Clevers H (2019) Probing the tumour suppressor function of BAP1 in CRISPR-engineered human liver organoids. Cell Stem Cell 24:927-943

Asli N, Xaymardan M, Harvey R (2014) Epicardial origin of resident mesenchymal stem cells in the adult mammalian heart. Journal of Developmental Biology 2:117-137

Augoff K, McCue B, Plow E, Sossey-Alaoui K (2012) miR-31 and its host gene lncRNA LOC554202 are regulated by promoter hypermethylation in triple-negative breast cancer. Molecular Cancer 11:1476-4598

Avula U, Abrams J, Katchman A, Zakharov S, Mironov S, Bayne J, Roybal D, Gorti A, Yang L, Iyer V, Waase M, Saluja D, Ciaccio E, Garan H, Marks A, Marx S, Wan E (2019) Heterogeneity of the action potential duration is required for sustained atrial fibrillation. Journal of Clinical Investigation 4:e128765

Bagchi R, Lin J, Wang R, Czubryt M (2016) Regulation of fibronectin gene expression in cardiac fibroblasts by scleraxis. Cell and Tissue Research 366:381-391

Bageghni S, Hemmings K, Zava N, Denton C, Porter K, Ainscough J, Drinkhill M, Turner N (2018) Cardiac fibroblast-specific p38alpha MAP kinase promotes cardiac hypertrophy via a putative paracrine interleukin-6 signaling mechanism. Federation of American Societies for Experimental Biology 32:4941-4954

Balik M, Matousek V, Maly M, Brozek T (2017) Management of arrhythmia in sepsis and septic shock. Anaesthesiology Intensive Therapy 49:419-429

Banerjee I, Fuseler J, Price R, Borg T, Baudino T (2007) Determination of cell types and numbers during cardiac development in the neonatal and adult rat and mouse. American Journal of Physiology - Heart and Circulatory 293:1883-1891

Barski A, Cuddapah S, Cui K, Roh T-Y, Schones E, Wang Z, Wei G, Chepelev I, Zhao K (2007) High-resolution profiling of histone methylations in the human genome. Cell 129:823-837

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

Bartel D (2009) MicroRNAs: target recognition and regulatory functions. Cell 136:215-233

Bartos D, Grandi E, Ripplinger C (2015) Ion channels in the heart. Comprehensive Physiology 5:1423-1464

Bashey R, Donnelly M, Insinga F, Jimenez S (1992) Growth properties and biochemical characterization of collagens synthesized by adult rat heart fibroblasts in culture. Journal of Molecular and Cellular Cardiology 24:691-700

Benjamin E, Wolf P, D'Agostino R, Silbershatz H, Kannel W, Levy D (1998) Impact of atrial fibrillation on the risk of death: the Framingham heart study. Circulation 98:946-952

Bernau K, Torr E, Evans M, Aoki J, Ngam C, Sandbo N (2017) Tensin 1 is essential for myofibroblast differentiation and extracellular matrix formation. American Journal of Respiratory Cell and Molecular Biology 56:465-476

Bersell K, Choudhury S, Mollova M, Polizzotti B, Ganapathy B, Walsh S, Wadugu B, Arab S, Kühn B (2013) Moderate and high amounts of tamoxifen in αMHC-MerCreMer mice induce a DNA damage response, leading to heart failure and death. Disease Model Mechanisms 6:1459-1469

Bochaton-Piallat M, Gabbiani G, Hinz B (2016) The myofibroblast in wound healing and fibrosis: answered and unanswered questions. Faculty of 1000 Research 5:e752

Bohnen M, Iyer V, Sampson K, Kass S (2015) Novel mechanism of transient outward potassium channel current regulation in the heart: implications for cardiac electrophysiology in health and disease. Circulation Research 116:1633-1635

Borschel C, Schnabel R (2019) The imminent epidemic of atrial fibrillation and its concomitant diseases - myocardial infarction and heart failure - a cause for concern. International Journal of Cardiology 287:162-173

Braitsch C, Kanisicak O, van Berlo J, Molkentin J, Yutzey K (2013) Differential expression of embryonic epicardial progenitor markers and localization of cardiac fibrosis in adult ischemic injury and hypertensive heart disease. Molecular Cell Cardiology 65:108-119

Burdett P, Lip G (2020) Atrial fibrillation in the United Kingdom: predicting costs of an emerging epidemic recognising and forecasting the cost drivers of atrial fibrillation-related costs. European Heart Journal - Quality of Care and Clinical Outcomes 93:1-8

Burstein B, Comtois P, Michael G, Nishida K, Villeneuve L, Yeh H, Nattel S (2009) Changes in connexin expression and the atrial fibrillation substrate in congestive heart failure. Circulation Research 105:1213-1222

Burstein B, Nattel S (2008) Atrial fibrosis: mechanisms and clinical relevance in atrial fibrillation. Journal of the American College of Cardiology 51:802-809

Butt R, Laurent G, Bishop J (1995) Collagen production and replication by cardiac fibroblasts is enhanced in response to diverse classes of growth factors. European Journal Cell Biology 68:330-335

Calvo D, Filgueiras-Rama D, Jalife J (2018) Mechanisms and drug development in atrial fibrillation. Pharmacological Reviews 70:505-525

Cardin S, Guasch E, Luo X, Naud P, Le K, Shi Y, Tardif J, Comtois P, Nattel S (2012) Role for MicroRNA-21 in atrial profibrillatory fibrotic remodeling associated with experimental postinfarction heart failure. Circulation Arrhythmia and Electrophysiology 5:1027-1035

Carrel A, Ebeling A (1926) The fundamental properties of the fibroblast and the macrophage: the fibroblast. The Journal of Experimental Medicine 44:261-287

Chen F, Sha M, Wang Y, Wu T, Shan W, Liu J, Zhou W, Zhu Y, Sun Y, Shi Y, Bleich D, Han X (2016) Transcription factor Ets-1 links glucotoxicity to pancreatic beta cell dysfunction through inhibiting PDX-1 expression in rodent models. Diabetologia 59:316-324

Chen J, Fu Y, Day S, Sun Y, Wang S, Liang X, Gu F, Zhang F, Stevens M, Zhou P, Li K, Zhang Y, Lin R, Smith H, Zhang J, Sun K, Melero-Martin M, Han Z, Park J, Zhang B, Pu T (2017) VEGF amplifies transcription through ETS1 acetylation to enable angiogenesis. Nature Communications 8:383

Chen P, Antzelevitch C (2017) Mechanisms of cardiac arrhythmias and conduction disturbances. Hurst's The Heart, edition 14, chapter 79, pages 1-55

Chen P, Chen L, Fishbein M, Lin S, Nattel S (2014) Role of the autonomic nervous system in atrial fibrillation. Circulation Research 114:1500-1515

Chen Y, Surinkaew S, Naud P, Qi X, Gillis M, Shi Y, Tardif J, Dobrev D, Nattel S (2017) JAK-STAT signaling and the atrial fibrillation promoting fibrotic substrate. Cardiovascular Research 13:310–320

Chong J, Reinecke H, Iwata M, Torok-Storb B, Stempien-Otero A, Murry C (2013) Progenitor cells identified by PDGFR-alpha expression in the developing and diseased human heart. Stem Cells Development 22:1932-1943

Christophersen I, Ellinor P (2016) Genetics of atrial fibrillation: from families to genomes. Journal of Human Genetics 61:61-70

Chugh S, Havmoeller R, Narayanan K, Singh D, Rienstra M, Benjamin E, Gillum R, Kim Y, McAnulty J, Zheng Z, Forouzanfar M, Naghavi M, Mensah G, Ezzati M, Murray C (2014) Worldwide epidemiology of atrial fibrillation: a global burden of disease 2010 Study. Circulation 129:837-847

Collins K, Korcarz C, Lang R (2003) Use of echocardiography for the phenotypic assessment of genetically altered mice. Physiological genomics 13:227-239

Conway S, Molkentin J (2008) Periostin as a heterofunctional regulator of cardiac development and disease. Current Genomics 9:548-555

Cooley N, Cowley M, Lin R, Marasco S, Wong C, Kaye D, Dart A, Woodcock E (2012) Influence of atrial fibrillation on microRNA expression profiles in left and right atria from patients with valvular heart disease. Physiological Genomics 44:211-219

Corcoran L, Pandit V, Gordon B, Bhattacharjee A, Kaminski N, Benos V (2009) Features of mammalian microRNA promoters emerge from polymerase II chromatin immunoprecipitation data. Public Library of Sciences 4:e5279

Czuwara-Ladykowska J, Sementchenko I, Watson K, Trojanowska M (2002) Ets1 is an effector of the transforming growth gactor β (TGF-β) signaling pathway and an antagonist of the profibrotic effects of TGF-β. Journal of Biological Chemistry 277:20399-20408

Dallas M, Yang Z, Boyle J, Boycott H, Scragg J, Milligan C, Elies J, Duke A, Thireau J, Reboul C, Richard S, Bernus O, Steele D, Peers C (2012) Carbon monoxide induces cardiac arrhythmia via induction of the late Na+ current. American Journal of Respiratory Critical Care Medicine 186:648-656

Davis R, Hobbs F, Kenkre J, Roalfe A, Iles R, Lip G, Davies M (2012) Prevalence of atrial fibrillation in the general population and in high-risk groups: the ECHOES study. The European Journal of Pacing, Arrhythmias and Cardiac Electrophysiology 14:1553-1559

Dawson K, Wakili R, Ordog B, Clauss S, Chen Y, Iwasaki Y, Voigt N, Qi X, Sinner M, Dobrev D, Kaab S, Nattel S (2013) MicroRNA29: a mechanistic contributor and potential biomarker in atrial fibrillation. Circulation 127:1466-1475

Denham N, Pearman C, Caldwell J, Madders G, Eisner D, Trafford A, Dibb K (2018) Calcium in the pathophysiology of atrial fibrillation and heart failure. Frontiers in Physiology 9:e1380

Dhamoon A, Jalife J (2005) The inward rectifier current (IK1) controls cardiac excitability and is involved in arrhythmogenesis. Heart Rhythm 2:316-324

Ding W, Gupta D, Lip G (2020) Atrial fibrillation and the prothrombotic state: revisiting Virchow's triad in 2020. Heart 106:1463-1468

Dolatshad H, Biggs D, Diaz R, Hortin N, Preece C, Davies B (2015) A versatile transgenic allele for mouse overexpression studies. Mammalian Genome 26:598-608

Donati S, Ciuffi S, Brandi M (2019) Human circulating miRNAs real-time qRT-PCR-based analysis: an overview of endogenous reference genes used for data normalization. International Journal of Molecular Sciences 20:4353-4372

Dorn L, Petrosino J, Wright P, Accornero F (2018) CTGF/CCN2 is an autocrine regulator of cardiac fibrosis. Molecular Cell Cardiology 121:205-211

Duisters F, Tijsen J, Schroen B, Leenders J, Lentink V, Made I, Herias V, Leeuwen R, Schellings W, Barenbrug P, Maessen J, Heymans S, Pinto M, Creemers E (2009) miR-133 and miR-30 regulate connective tissue growth factor. Circulation Research 104:170-178

Dutta S, Sengupta P (2016) Men and mice: relating their ages. Life Sciences 152:244-248

Dzeshka M, Lip G, Snezhitskiy V, Shantsila E (2015) Cardiac fibrosis in patients with atrial fibrillation: mechanisms and clinical implications. Journal of the American College of Cardiology 66:943-959

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

Everett T, Wilson E, Verheule S, Guerra J, Foreman S, Olgin J (2006) Structural atrial remodeling alters the substrate and spatiotemporal organization of atrial fibrillation: a comparison in canine models of structural and electrical atrial remodeling. American Journal of Physiology Heart and Circulatory Physiology 291:2911-2923

Farh K, Grimson A, Jan C, Lewis B, Johnston W, Lim L, Burge C, Bartel D (2005) The widespread impact of mammalian microRNAs on mRNA repression and evolution. Science 310:1817-1821

Ferreira C, Providencia R, Ferreira M, Goncalves L (2015) Atrial fibrillation and non-cardiovascular diseases: a systematic review. Arquivos Brasileiros Cardiology 105:519-526

Fleischman L, Holtzclaw L, Russell J, Mavrothalassitis G, Fisher R (1995) ETS1 in astrocytes: expression and transmitter-evoked phosphorylation. Molecular and Cellular Biology 15:925-931

Forrester S, Booz G, Sigmund C, Coffman T, Kawai T, Rizzo V, Scalia R, Eguchi S (2018) Angiotensin II signal transduction: an update on mechanisms of physiology and pathophysiology. Physiological Reviews 98:1627-1738

Franklyn J, Boelaert K (2012) Thyrotoxicosis. The Lancet 379:1155-1166

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

Frustaci A, Caldarulo M, Buffon A, Bellocci F, Fenici R, Melina D (1991) Cardiac biopsy in patients with "primary" atrial fibrillation. Histologic evidence of occult myocardial diseases. Chest 100:303-306

Frustaci A, Chimenti C, Bellocci F, Morgante E, Russo M, Maseri A (1997) Histological substrate of atrial biopsies in patients with lone atrial fibrillation. Circulation 96:1180-1184

Fu X, Khalil H, Kanisicak O, Boyer J, Vagnozzi R, Maliken D, Sargent M, Prasad V, Valiente-Alandi I, Blaxall C, Molkentin J (2018) Specialized fibroblast differentiated states underlie scar formation in the infarcted mouse heart. Journal of Clinical Investigations 128:2127-2143

Fu X, Liu Q, Li C, Li Y, Wang L (2020) Cardiac fibrosis and cardiac fibroblast lineage-tracing: recent advances. Frontiers Physiology 11:416-416

Gaborit N, Steenman M, Lamirault G, Le Meur N, Le Bouter S, Lande G, Léger J, Charpentier F, Christ T, Dobrev D, Escande D, Nattel S, Demolombe S (2005) Human atrial ion channel and transporter subunit gene-expression remodelling associated with valvular heart disease and atrial fibrillation. Circulation 112:471-481

Go A, Hylek E, Phillips K, Chang Y, Henault L, Selby J, Singer D (2001) Prevalence of diagnosed atrial fibrillation in adults: national implications for rhythm management and stroke prevention: the anticoagulation and risk factors in atrial fibrillation (ATRIA) Study. Journal of the American Medical Association 285:2370-2375

Goren Y, Meiri E, Hogan C, Mitchell H, Lebanony D, Salman N, Schliamser J, Amir O (2014) Relation of reduced expression of miR-150 in platelets to atrial fibrillation in patients with chronic systolic heart failure. The American Journal of Cardiology 113:976-981

Grant O (2009) Cardiac ion channels. Circulation Arrhythmia and Electrophysiology 2:185-194

Greenberg J, Lancaster T, Schuessler R, Melby S (2017) Postoperative atrial fibrillation following cardiac surgery: a persistent complication. European Journal of Cardio-Thoracic Surgery 52:665-672

Griffiths-Jones S, Grocock R, van Dongen S, Bateman A, Enright A (2006) miRBase: microRNA sequences, targets and gene nomenclature. Nucleic Acids Research 34:140-144

Guenther G, Levine S, Boyer A, Jaenisch R, Young A (2007) A chromatin landmark and transcription initiation at most promoters in human cells. Cell 130:77-88

Guo Y, Pan W, Liu S, Shen Z, Xu Y, Hu L (2020) ERK/MAPK signalling pathway and tumorigenesis. Experimental and Therapeutic Medicine 19:1997-2007

Hara H, Takeda N, Komuro I (2017) Pathophysiology and therapeutic potential of cardiac fibrosis. Inflammation and Regeneration 37:1-10

He X, Zhang K, Gao X, Li L, Tan H, Chen J, Zhou Y (2016) Rapid atrial pacing induces myocardial fibrosis by down-regulating Smad7 via microRNA-21 in rabbit. Heart and Vessels 31:1696-1708

Healey J, Connolly S, Gold M, Israel C, Van Gelder I, Capucci A, Lau C, Fain E, Yang S, Bailleul C, Morillo C, Carlson M, Themeles E, Kaufman E, Hohnloser S (2012) Subclinical atrial fibrillation and the risk of stroke. The New England Journal of Medicine 366:120-129

Heijman J, Voigt N, Nattel S, Dobrev D (2014) Cellular and molecular electrophysiology of atrial fibrillation initiation, maintenance, and progression. Circulation Research 114:1483-1499

Heijman J, Voigt N, Wehrens X, Dobrev D (2014) Calcium dysregulation in atrial fibrillation: the role of CaMKII. Frontiers in Pharmacology 5:e30

Herum K, Lunde I, McCulloch A, Christensen G (2017) The soft- and hard-heartedness of cardiac fibroblasts: mechanotransduction signalling pathways in fibrosis of the heart. Journal of Clinical Medicine 6:e53

Hong K, Bjerregaard P, Gussak I, Brugada R (2005) Short QT Syndrome and Atrial Fibrillation Caused by Mutation in KCNH2. Journal of Cardiovascular Electrophysiology 16:394-396

Hsia L, Ashley N, Ouaret D, Wang L, Wilding J, Bodmer W (2016) Myofibroblasts are distinguished from activated skin fibroblasts by the expression of AOC3 and other associated markers. Proceedings of the National Academy of Sciences 113:2162-2171

Hu J, Chen C, Liu Q, Liu B, Song C, Zhu S, Wu C, Liu S, Yu H, Yao D, Kang J, Zhu L (2015) The role of the miR-31/FIH1 pathway in TGF-beta-induced liver fibrosis. Clinical Science 129:305-317

Hudon-David F, Bouzeghrane F, Couture P, Thibault G (2007) Thy-1 expression by cardiac fibroblasts: lack of association with myofibroblast contractile markers. Journal of Molecular Cell Cardiology 42:991-1000

Hughes-Davies L, Huntsman D, Ruas M, Fuks F, Bye J, Chin F, Milner J, Brown A, Hsu F, Gilks B, Nielsen T, Schulzer M, Chia S, Ragaz J, Cahn A, Linger L, Ozdag H, Cattaneo

E, Jordanova S, Schuuring E, Yu S, Venkitaraman A, Ponder B, Doherty A, Aparicio S, Bentley D, Theillet C, Ponting C, Caldas C, Kouzarides T (2003) EMSY links the BRCA2 pathway to sporadic breast and ovarian cancer. Cell 115:523-535

Indroneal B, Fuseler J, Price R, Borg T, Baudino T (2007) Determination of cell types and numbers during cardiac development in the neonatal and adult rat and mouse. American Journal of Physiology-Heart and Circulatory Physiology 293:1883-1891

Itoh N, Ohta H (2013) Pathophysiological roles of FGF signaling in the heart. Frontiers in Physiology 4:247-247

Ivey M, Tallquist M (2016) Defining the cardiac fibroblast. Japanese Circulation Journal: 80:2269-2276

January C, Wann L, Alpert J, Calkins H, Cigarroa J, Cleveland J, Conti J, Ellinor P, Ezekowitz M, Field M, Murray K, Sacco R, Stevenson W, Tchou P, Tracy C, Yancy C (2014) AHA/ACC/HRS Guideline for the management of patients with atrial fibrillation: executive summary. Journal of the American College of Cardiology 64:2246-2280

Kalajzic I, Kalajzic Z, Kaliterna M, Gronowicz G, Clark SH, Lichtler A, Rowe D (2002) Use of type I collagen green fluorescent protein transgenes to identify subpopulations of cells at different stages of the osteoblast lineage. Journal of the American Society for Bone and Mineral Research 17:15-25

Kanisicak O, Khalil H, Ivey J, Karch J, Maliken D, Correll N, Brody J, Aronow J, Tallquist D, Molkentin D (2016) Genetic lineage tracing defines myofibroblast origin and function in the injured heart. Nature Communications 7:12260

Kaur H, Takefuji M, Ngai Y, Carvalho J, Bayer J, Wietelmann A, Poetsch A, Hoelper S, Conway J, Möllmann H, Looso M, Troidl C, Offermanns S, Wettschureck N (2016) Targeted ablation of periostin-expressing activated fibroblasts prevents adverse cardiac remodeling in mice. Circulation Research 118:1906-1917

Kawai S, Mukai Y, Inoue S, Yakabe D, Nagaoka K, Sakamoto K, Takase S, Chishaki A, Tsutsui H (2019) Non-pulmonary vein triggers of atrial fibrillation are likely to arise from low-voltage areas in the left atrium. Scientific Reports 9:e12271

Khalil H, Kanisicak O, Prasad V, Correll R, Fu X, Schips T, Vagnozzi R, Liu R, Huynh T, Lee S, Karch J, Molkentin J (2017) Fibroblast-specific TGF-beta-Smad2/3 signaling underlies cardiac fibrosis. Journal of Clinical Investigation 127:3770-3783

Khalil H, Kanisicak O, Vagnozzi R, Johansen A, Maliken B, Prasad V, Boyer J, Brody M, Schips T, Kilian K, Correll R, Kawasaki K, Nagata K, Molkentin J (2019) Cell-specific

ablation of Hsp47 defines the collagen-producing cells in the injured heart. Journal of Clinical Investigation 4:e128722

Kim J, Nakashima K, de Crombrugghe B (2004) Transgenic mice expressing a ligand-inducible Cre recombinase in osteoblasts and odontoblasts: a new tool to examine physiology and disease of postnatal bone and tooth. The American Journal of Pathology 165:1875-1882

Kirchhof P, Benussi S, Kotecha D, Ahlsson A, Atar D, Casadei B, Castella M, Diener H, Heidbuchel H, Hendriks J, Hindricks G, Manolis S, Oldgren J, Popescu B, Schotten U, Van Putte B, Vardas P (2016) 2016 ESC Guidelines for the management of atrial fibrillation developed in collaboration with EACTS. European Heart Journal 37:2893-2962

Kisanuki Y, Hammer E, Miyazaki J, Williams C, Richardson A, Yanagisawa M (2001) Tie2-Cre transgenic mice: a new model for endothelial cell-lineage analysis in vivo. Developmental Biology 230:230-242

Kompa A, See F, Lewis D, Adrahtas A, Cantwell D, Wang B, Krum H (2008) Long-term but not short-term p38 mitogen-activated protein kinase inhibition improves cardiac function and reduces cardiac remodeling post-myocardial infarction. Journal of Pharmacology and Experimental Therapeutics 325:741-750

Kounis N, Koniari I, Tzanis G, Soufras G, Velissaris D, Hahalis G (2020) Anaphylaxis-induced atrial fibrillation and anaesthesia: pathophysiologic and therapeutic considerations. Annals of Cardiac Anaesthesia 23:1-6

Lachmann S, Jevons A, De Rycker M, Casamassima A, Radtke S, Collazos A, Parker J (2011) Regulatory domain selectivity in the cell-type specific PKN-dependence of cell migration. The Public Library of Science 6:e21732

Laffer C, Elijovich F, Sahinoz M, Pitzer A, Kirabo A (2020) New insights into the renin-angiotensin system in chronic kidney disease. Circulation Research 127:607-609

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

Landgraf P, Rusu M, Sheridan R, Sewer A, Iovino N, Aravin A, Pfeffer S, Rice A, Kamphorst A, Landthaler M, Lin C, Socci N, Hermida L, Fulci V, Chiaretti S, Foà R, Schliwka J, Fuchs U, Novosel A, Müller R, Schermer B, Bissels U, Inman J, Phan Q, Chien M, Weir D, Choksi R, De Vita G, Frezzetti D, Trompeter H, Hornung V, Teng G, Palkovits M, Di Lauro R, Wernet P, Macino G, Rogler C, Nagle J, Ju J, Papavasiliou F, Benzing T, Lichter P, Tam W, Brownstein M, Bosio A, Borkhardt A, Russo J, Sander C, Zavolan

M, Tuschl T (2007) A mammalian microRNA expression atlas based on small RNA library sequencing. Cell 129:1401-1414

Landry N, Cohen S, Dixon I (2017) Periostin in cardiovascular disease and development: a tale of two distinct roles. Basic Research in Cardiology 113:e1

Landry, N, Rattan G, Dixon I (2019) An improved method of maintaining primary murine cardiac fibroblasts in two-dimensional cell culture. Scientific Reports 9:e12889

Laurence J, Elhadad S, Robison T, Terry H, Varshney R, Woolington S, Ghafoory S, Choi M, Ahamed J (2017) HIV protease inhibitor-induced cardiac dysfunction and fibrosis is mediated by platelet-derived TGF-β1 and can be suppressed by exogenous carbon monoxide. The Public Library of Science 12:e0187185

Laurila E, Kallioniemi A (2013) The diverse role of miR-31 in regulating cancer associated phenotypes. Genes, Chromosomes and Cancer 52:1103-1113

LeBleu V, Neilson E (2020) Origin and functional heterogeneity of fibroblasts. The Federation of American Societies for Experimental Biology 34:3519-3536

Lei A, Chen L, Zhang M, Yang X, Xu L, Cao N, Zhang Z, Cao Y (2019) EZH2 regulates protein stability via recruiting USP7 to mediate neuronal gene expression in cancer cells. Frontiers in Genetics 10:422-440

Levy S (1997) Factors predisposing to the development of atrial fibrillation. Pacing and Clinical Electrophysiology 20:2670-2674

Lexow J, Poggioli T, Sarathchandra P, Santini M, Rosenthal N (2013) Cardiac fibrosis in mice expressing an inducible myocardial-specific Cre driver. Disease Models and Mechanisms 6:e1470

Li Y, Li Z, Zhang C, Li P, Wu Y, Wang C, Bond Lau W, Ma XL, Du J (2017) Cardiac fibroblast-specific activating transcription factor 3 protects against heart failure by suppressing MAP2K3-p38 signalling. Circulation 135:2041-2057

Li Y, Song D, Mao L, Abraham D, Bursac N (2020) Lack of Thy1 defines a pathogenic fraction of cardiac fibroblasts in heart failure. Biomaterials 236:e119824

Lim W, Corden B, Ye L, Viswanathan S, Widjaja A, Xie C, Su L, Tee N, Schafer S, Cook S Antibody-mediated neutralization of IL11 signalling reduces ERK activation and cardiac fibrosis in a mouse model of severe pressure overload. Clinical and Experimental Pharmacology and Physiology 48:605-613

Lin C, Lai L, Lin J, Sun Y, Hsu C, Chen C, Mao S, Huang S (2007) Increased expression of extracellular matrix proteins in rapid atrial pacing-induced atrial fibrillation. Heart Rhythm 4:938-949

Lindsey M, Kassiri Z, Virag J, de Castro Bras L, Scherrer-Crosbie M (2018) Guidelines for measuring cardiac physiology in mice. American Journal Heart Circulation and Physiology 314:733-752

Ling L, Kistler P, Kalman J, Schilling R, Hunter R (2016) Comorbidity of atrial fibrillation and heart failure. Nature Reviews Cardiology 13:131-147

Lip G, Brechin C, Lane D (2012) The global burden of atrial fibrillation and stroke: a systematic review of the epidemiology of atrial fibrillation in regions outside North America and Europe. Chest 142:1489-1498

Lip G, Fauchier L, Freedman S, Van Gelder I, Natale A, Gianni C, Nattel S, Potpara T, Rienstra M, Tse H, Lane D (2016) Atrial fibrillation. Nature Review Disease Primers 2:e16016

Liu T, Xiong F, Qi X, Xiao J, Villeneuve L, Abu-Taha I, Dobrev D, Huang C, Nattel S (2020) Altered calcium handling produces reentry-promoting action potential alternans in atrial fibrillation–remodeled hearts. Journal of Clinical Investigation 5:e133754

Liu Z, Zhou C, Liu Y, Wang S, Ye P, Miao X, Xia J (2012) The expression levels of plasma micoRNAs in atrial fibrillation patients. The Public Library of Science 7:e44906

Lundby A, Ravn L, Svendsen J, Haunsø S, Olesen S, Schmitt N (2008) KCNE3 mutation V17M identified in a patient with lone atrial fibrillation. Cellular Physiology and Biochemistry 21:47-54

Luo Y, Hitz B, Gabdank I, Hilton J, Kagda M, Lam B, Myers Z, Sud P, Jou J, Lin K, Baymuradov U, Graham K, Litton C, Miyasato S, Strattan J, Jolanki O, Lee W, Tanaka F, Adenekan P, O'Neill E, Cherry J (2020) New developments on the Encyclopedia of DNA Elements (ENCODE) data portal. Nucleic Acids Research 48:882-889

Ma X, Qi S, Duan Z, Liao H, Yang B, Wang W, Tan J, Li Q, Xia X (2017) Long non-coding RNA LOC554202 modulates chordoma cell proliferation and invasion by recruiting EZH2 and regulating miR-31 expression. Cell Proliferation 50:e12388

Maharaj S, Perez-Downes J, Seegobin K, Chang S (2018) Consideration of atrial arrhythmias associated with cardiac tamponade and pericarditis. The American Journal of Emergency Medicine 36:e338

Mahnkopf C, Badger T, Burgon N, Daccarett M, Haslam T, Badger C, McGann C, Akoum N, Kholmovski E, Macleod R, Marrouche N (2010) Evaluation of the left atrial substrate in

patients with lone atrial fibrillation using delayed-enhanced MRI: implications for disease progression and response to catheter ablation. Heart Rhythm 7:1475-1481

Maik-Rachline G, Hacohen-Lev-Ran A, Seger R (2019) Nuclear ERK: mechanism of translocation, substrates, and role in cancer. International Journal of Molecular Sciences 20:e1194

Marson A, Levine S, Cole M, Frampton G, Brambrink T, Johnstone S, Guenther M, Johnston W, Wernig M, Newman J, Calabrese J, Dennis L, Volkert T, Gupta S, Love J, Hannett N, Sharp P, Bartel D, Jaenisch R, Young R (2008) Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells. Cell 134:521-533

Martinez E, Lilyanna S, Wang P, Vardy L, Jiang X, Armugam A, Jeyaseelan K, Richards A (2017) MicroRNA-31 promotes adverse cardiac remodeling and dysfunction in ischemic heart disease. Journal Molecular Cell Cardiology 112:27-39

Martins R, Kaur K, Hwang E, Ramirez R, Willis B, Filgueiras-Rama D, Ennis S, Takemoto Y, Ponce-Balbuena D, Zarzoso M, O'Connell R, Musa H, Guerrero-Serna G, Avula U, Swartz M, Bhushal S, Deo M, Pandit S, Berenfeld O, Jalife J (2014) Dominant frequency increase rate predicts transition from paroxysmal to long-term persistent atrial fibrillation. Circulation 129:1472-1482

Mase M, Grasso M, Avogaro L, D'Amato E, Tessarolo F, Graffigna A, Denti M, Ravelli F (2017) Selection of reference genes is critical for miRNA expression analysis in human cardiac tissue. A focus on atrial fibrillation. Scientific Reports 7:41127-41127

Mayourian J, Ceholski D, Gonzalez D, Cashman T, Sahoo S, Hajjar R, Costa K (2018) Physiologic, pathologic, and therapeutic paracrine modulation of cardiac excitation-contraction coupling. Circulation Research 122:167-183

McManus D, Lin H, Tanriverdi K, Quercio M, Yin X, Larson M, Ellinor P, Levy D, Freedman J, Benjamin E (2014) Relations between circulating microRNAs and atrial fibrillation: data from the Framingham offspring study. Heart Rhythm 11:663-669

Mendez M, Kojima S, Goldman R (2010) Vimentin induces changes in cell shape, motility, and adhesion during the epithelial to mesenchymal transition. Federation of American Societies for Experimental Biology. 24:1838-1851

Meyer A, Wang W, Qu J, Croft L, Degen J, Coller B, Ahamed J (2012) Platelet TGF-β1 contributions to plasma TGF-β1, cardiac fibrosis, and systolic dysfunction in a mouse model of pressure overload. Blood 119:1064-1074

Michaud G, Stevenson W (2018) Atrial Fibrillation. Harrison's Principles of Internal Medicine, edition 20, chapter 246, pages 1-8

Miyasaka Y, Barnes M, Gersh B, Cha S, Bailey K, Abhayaratna W, Seward J, Tsang T (2006) Secular trends in incidence of atrial fibrillation in Olmsted County, Minnesota, 1980 to 2000, and implications on the projections for future prevalence. Circulation 114:119-125

Mohammadzadeh N, Lunde IG, Andenæs K, Strand M, Aronsen J, Skrbic B, Marstein H, Bandlien C, Nygård S, Gorham J, Sjaastad I, Chakravarti S, Christensen G, Engebretsen K, Tønnessen T (2019) The extracellular matrix proteoglycan lumican improves survival and counteracts cardiac dilatation and failure in mice subjected to pressure overload. Scientific Reports 9:e9206

Molkentin J, Bugg D, Ghearing N, Dorn L, Kim P, Sargent M, Gunaje J, Otsu K, Davis J (2017) Fibroblast-specific genetic manipulation of p38 mitogen-activated protein kinase in vivo reveals its central regulatory role in fibrosis. Circulation 136:549-561

Monfredi O, Dobrzynski H, Mondal T, Boyett M, Morris G (2010) The anatomy and physiology of the sinoatrial node-a contemporary review. Pacing and Clinical Electrophysiology 33:1392-1406

Moore-Morris T, Guimarães-Camboa N, Banerjee I, Zambon A, Kisseleva T, Velayoudon A, Stallcup W, Gu Y, Dalton N, Cedenilla M, Gomez-Amaro R, Zhou B, Brenner D, Peterson K, Chen J, Evans S (2014) Resident fibroblast lineages mediate pressure overload-induced cardiac fibrosis. Journal of Clinical Investigation 124:2921-2934

Morales O, Price L, Goldsmith C (2005) Expression of discoidin domain receptor 2 (DDR2) in the developing heart. Microscopy and Microanalysis 11:260-267

Moreira L, Takawale A, Hulsurkar M, Menassa D, Antanaviciute A, Lahiri S, Mehta N, Evans N, Psarros C, Robinson P, Sparrow A, Gillis M, Ashley N, Naud P, Barallobre-Barreiro J, Theofilatos K, Lee A, Norris M, Clarke M, Russell P, Casadei B, Bhattacharya S, Zajac JD, Davey A, Sirois M, Mead A, Simmons A, Mayr M, Sayeed R, Krasopoulos G, Redwood C, Channon K, Tardif C, Wehrens X, Nattel S, Reilly S (2020) Paracrine signalling by cardiac calcitonin controls atrial fibrogenesis and arrhythmia. Nature 587:460-465

Murtha L, Schuliga M, Mabotuwana N, Hardy S, Waters D, Burgess J, Knight D, Boyle A (2017) The processes and mechanisms of cardiac and pulmonary fibrosis. Frontiers in Physiology 8:e777

Nattel S (2016) How does fibrosis promote atrial fibrillation persistence: in silico findings, clinical observations, and experimental data. Cardiovascular Research 110:295-297

Nattel S (2017) Molecular and cellular mechanisms of atrial fibrosis in atrial fibrillation. The Journal of the American College of Cardiology: Clinical Electrophysiology 3:425-435

Nattel S, Heijman J, Zhou L, Dobrev D (2020) Molecular basis of atrial fibrillation pathophysiology and therapy: a translational perspective. Circulation Research 127:51-72

Ni H, Zhang H, Grandi E, Narayan S, Giles W (2019) Transient outward K(+) current can strongly modulate action potential duration and initiate alternans in the human atrium. American Journal of Physiology: Heart Circulation Physiology 316:527-542

Nielsen B (2012) MicroRNA in situ hybridization. Next-generation microRNA expression profiling technology: methods and protocols. Humana Press 822:67-84

Nielsen J, Bentzen B, Olesen M, David J, Olesen S, Haunsø S, Svendsen J, Schmitt N (2014) Gain-of-function mutations in potassium channel subunit KCNE2 associated with early-onset lone atrial fibrillation. Biomarkers in Medicine 8:557-570

Odutayo A, Wong C, Hsiao A, Hopewell S, Altman D, Emdin C (2016) Atrial fibrillation and risks of cardiovascular disease, renal disease, and death: systematic review and meta-analysis. British Medical Journal 354:e4482

Okamura K, Hagen J, Duan H, Tyler D, Lai E (2007) The Mirtron Pathway Generates microRNA-Class Regulatory RNAs in Drosophila. Cell 130:89-100

Olesen M, Bentzen B, Nielsen J, Steffensen A, David J, Jabbari J, Jensen H, Haunsø S, Svendsen J, Schmitt N (2012) Mutations in the potassium channel subunit KCNE1 are associated with early-onset familial atrial fibrillation. British Medical Journal: Medical Genetics 13:e24

Omelyanenko N, Slutsky L, Mironov S (2014) Connective tissue: histophysiology, biochemistry, molecular biology, edition 1, chapter 3, pages 121-139

Oparil S, Acelajado M, Bakris G, Berlowitz D, Cífková R, Dominiczak A, Grassi G, Jordan J, Poulter N, Rodgers A, Whelton P (2018) Hypertension. Nature reviews Disease primers 4:18014-18014

Ozsolak F, Poling L, Wang Z, Liu H, Liu X, Roeder R, Zhang X, Song J, Fisher D (2008) Chromatin structure analyses identify miRNA promoters. Genes and Development 22:3172-3183

Palmiter R, Brinster R (1986) Germ-line transformation of mice. Annual Review of Genetics 20:465-499

Pan C, Lin J, Lai L, Chen C, H S, Lin C (2007) Downregulation of angiotensin converting enzyme II is associated with pacing-induced sustained atrial fibrillation. Federation of European Biochemical Societies 581:526-534

Pandit S, Berenfeld O, Anumonwo J, Zaritski R, Kneller J, Nattel S, Jalife J (2005) Ionic determinants of functional reentry in a 2-D model of human atrial cells during simulated chronic atrial fibrillation. Biophysical Journal 88:3806-3821

Papa S, Choy P, Bubici C (2019) The ERK and JNK pathways in the regulation of metabolic reprogramming. Oncogene 38:2223-2240

Park J, Lee C, Leshem E, Blau I, Kim S, Lee JM, Hwang J, Choi B, Lee M, Hwang H (2019) Early differentiation of long-standing persistent atrial fibrillation using the characteristics of fibrillatory waves in surface ECG multi-leads. Scientific Reports 9:e2746

Park S, Ranjbarvazirj S, Lay FD, Zhao P, Miller J, Dhaliwal S, Huertas-Vazquez A, Wu X, Qiao R, Soffer M, Mikkola H, Lusis A, Ardehali R (2018) Genetic regulation of fibroblast activation and proliferation in cardiac fibrosis. Circulation 138:1224–1235

Patrick D, Montgomery R, Qi X, Obad S, Kauppinen S, Hill J, van Rooij E, Olson E (2010) Stress-dependent cardiac remodeling occurs in the absence of microRNA-21 in mice. Journal of Clinical Investigation 120:3912-3916

Peretto G, Sala S, Rizzo S, De Luca G, Campochiaro C, Sartorelli S, Benedetti G, Palmisano A, Esposito A, Tresoldi M, Thiene G, Basso C, Della P (2019) Arrhythmias in myocarditis: state of the art. Heart Rhythm 16:793-801

Perna F, Leo M (2012) Epidemiology, classification and description of cardiac arrhythmias. Sports Cardiology: from Diagnosis to Clinical Management 2020:155-177

Piccini J, Yen Ho S, Becker A (2017) Electrophysiologic anatomy. Hurst's The Heart, edition 14, chapter 78, pages 1-32

Pillai I, Li S, Romay M, Lam L, Lu Y, Huang J, Dillard N, Zemanova M, Rubbi L, Wang Y, Lee J, Xia M, Liang O, Xie Y, Pellegrini M, Lusis A, Deb A (2017) Cardiac fibroblasts adopt osteogenic fates and can be targeted to attenuate pathological heart calcification. Cell Stem Cell 20:218-232

Pinto A, Ilinykh A, Ivey M, Kuwabara J, D'Antoni M, Debuque R, Chandran A, Wang L, Arora K, Rosenthal N, Tallquist M (2016) Revisiting cardiac cellular composition. Circulation Research 118:400-409

Platonov P, Mitrofanova L, Orshanskaya V, Ho S (2011) Structural abnormalities in atrial walls are associated with presence and persistency of atrial fibrillation but not with age. Journal of the American College of Cardiology 58:2225-2232

Prystowsky E, Halperin J, Kowey P (2017) Atrial fibrillation, atrial flutter, and atrial tachycardia. Hurst's The Heart, edition 14, chapter 83, pages 1-32

Purohit A, Rokita A, Guan X, Chen B, Koval O, Voigt N, Neef S, Sowa T, Gao Z, Luczak E, Stefansdottir H, Behunin A, Li N, El-Accaoui R, Yang B, Swaminathan P, Weiss R, Wehrens X, Song L, Dobrev D, Maier L, Anderson M (2013) Oxidized Ca(2+)/calmodulin-dependent protein kinase II triggers atrial fibrillation. Circulation 128:1748-1757

Ravens U, Odening K (2017) Atrial fibrillation: therapeutic potential of atrial K+ channel blockers. Pharmacology and Therapeutics 176:13-21

Ravens U, Wettwer E (2010) Ultra-rapid delayed rectifier channels: molecular basis and therapeutic implications. Cardiovascular Research 89:776-785

Reilly S, Liu X, Carnicer R, Recalde A, Muszkiewicz A, Jayaram R, Carena M, Wijesurendra R, Stefanini M, Surdo N, Lomas O, Ratnatunga C, Sayeed R, Krasopoulos G, Rajakumar T, Bueno-Orovio A, Verheule S, Fulga T, Rodriguez B, Schotten U, Casadei B (2016) Up-regulation of miR-31 in human atrial fibrillation begets the arrhythmia by depleting dystrophin and neuronal nitric oxide synthase. Science Translational Medicine 8:340-374

Ribatti D, Tamma R (2019) Giulio Gabbiani and the discovery of myofibroblasts. Inflammation Research 68:241-245

Rienks M, Papageorgiou A, Frangogiannis N, Heymans S (2014) Myocardial extracellular matrix: an ever-changing and diverse entity. Circulation Research 114:872-888

Ripamonti A, Provasi E, Lorenzo M, De Simone M, Ranzani V, Vangelisti S, Curti S, Bonnal R, Pignataro L, Torretta S, Geginat J, Rossetti G, Pagani M, Abrignani S (2017) Repression of miR-31 by BCL6 stabilizes the helper function of human follicular helper T cells. Proceedings of the National Academy of Sciences 114:12797-12802

Rosenkranz S (2004) TGF-β1 and angiotensin networking in cardiac remodeling. Cardiovascular Research 63:423-432

Ruby G, Jan C, Bartel D (2007) Intronic microRNA precursors that bypass Drosha processing. Nature 448:e83

Ruiz-Villalba A, Mattiotti A, Gunst Q, Cano-Ballesteros S, van den Hoff M, Ruijter J (2017) Reference genes for gene expression studies in the mouse heart. Scientific Reports 7:e24

Sakaguchi T, Takefuji M, Wettschureck N, Hamaguchi T, Amano M, Kato K, Tsuda T, Eguchi S, Ishihama S, Mori Y, Yura Y, Yoshida T, Unno K, Okumura T, Ishii H, Shimizu Y, Bando YK, Ohashi K, Ouchi N, Enomoto A, Offermanns S, Kaibuchi K, Murohara T (2019) Protein kinase N promotes stress-induced cardiac dysfunction through phosphorylation of myocardin-related transcription factor A and disruption of its interaction with actin. Circulation 140:1737-1752

Santiago J, Dangerfield A, Rattan, S, Bathe K, Cunnington H, Raizman E, Bedosky K, Freed H, Kardami E, and Dixon, I (2010) Cardiac fibroblast to myofibroblast differentiation in vivo and in vitro: Expression of focal adhesion components in neonatal and adult rat ventricular myofibroblasts. Developmental Dynamics 239: 1573-1584

Schafer S, Viswanathan S, Widjaja A, Lim W, Moreno-Moral A, DeLaughter M, Ng B, Patone G, Chow K, Khin E, Tan J, Chothani P, Ye L, Rackham L, Ko J, Sahib E, Pua J, Zhen G, Xie C, Wang M, Maatz H, Lim S, Saar K, Blachut S, Petretto E, Schmidt S, Putoczki T, Guimarães-Camboa N, Wakimoto H, van Heesch S, Sigmundsson K, Lim L, Soon L, Chao T, Chua L, TanE, Evans M, Loh J, Jamal H, Ong K, Chua C, Ong H, Chakaramakkil J, Seidman G, Seidman E, Hubner N, Sin K, Cook SA (2017) IL-11 is a crucial determinant of cardiovascular fibrosis. Nature 552:110-115

Schmidt A, Durgan J, Magalhaes A, Hall A (2007) Rho GTPases regulate PRK2/PKN2 to control entry into mitosis and exit from cytokinesis. European Molecular Biology Organization 26:1624-1636

Schmittgen T, Livak K (2008) Analyzing real-time PCR data by the comparative CT method. Nature Protocols 3:1101-1108

Schroer A, Merryman W (2015) Mechanobiology of myofibroblast adhesion in fibrotic cardiac disease. Journal of Cell Science 128:1865-1875

See F, Thomas W, Way K, Tzanidis A, Kompa A, Lewis D, Itescu S, Krum H (2004) p38 mitogen-activated protein kinase inhibition improves cardiac function and attenuates left ventricular remodeling following myocardial infarction in the rat. Journal of the American College of Cardiology 44:1679-1689

Semba T, Sammons R, Wang X, Xie X, Dalby K, Ueno N (2020) JNK signalling in stem cell self-renewal and differentiation. International Journal of Molecular Sciences 21: e2613

Semprini S, Troup T, Kotelevtseva N, King K, Davis J, Mullins L, Chapman K, Dunbar D, Mullins J (2007) Cryptic loxP sites in mammalian genomes: genome-wide distribution

and relevance for the efficiency of BAC/PAC recombineering techniques. Nucleic Acids Research 35:1402-1410

Shan H, Zhang Y, Lu Y, Zhang Y, Pan Z, Cai B, Wang N, Li X, Feng T, Hong Y, Yang B (2009) Downregulation of miR-133 and miR-590 contributes to nicotine-induced atrial remodelling in canines. Cardiovascular Research 83:465-472

Shi Y, Tao M, Wang Y, Zang X, Ma X, Qiu A, Zhuang S, Liu N (2020) Genetic or pharmacologic blockade of enhancer of zeste homolog 2 inhibits the progression of peritoneal fibrosis. Journal of Pathology 250:79-94

Shifera A, Hardin J (2009) PMA induces expression from the herpes simplex virus thymidine kinase promoter via the activation of JNK and ERK in the presence of adenoviral E1A proteins. Archives of Biochemistry and Biophysics 490:145-157

Shinde A, Humeres C, Frangogiannis N (2017) The role of α-smooth muscle actin in fibroblast-mediated matrix contraction and remodeling. Molecular Basis of Disease 1863:298-309

Sigel A, Centrella M, Eghbali-Webb M (1996) Regulation of proliferative response of cardiac fibroblasts by transforming growth factor-β1. Journal of Molecular and Cellular Cardiology 28:1921-1929

Smith C, Baek S, Sung C, Tallquist M (2011) Epicardial-derived cell epithelial-to-mesenchymal transition and fate specification require PDGF receptor signaling. Circulation Research 108:15-26

Sobral M, Hicks M, Parrish K, McCann S, Hsieh J, Goodspeed A, Costello C, Black C, Jedlicka P (2020) KDM3A/Ets1 epigenetic axis contributes to PAX3/FOXO1-driven and independent disease-promoting gene expression in fusion-positive Rhabdomyosarcoma. Molecular Oncology. 14:2471-2486.

Spinale F (2007) Myocardial matrix remodeling and the matrix metalloproteinases: influence on cardiac form and function. Physiological Reviews 87:1285-1342

Stellato M, Czepiel M, Distler O, Błyszczuk P, Kania G (2019) Identification and isolation of cardiac fibroblasts from the adult mouse heart csing two-color Flow Cytometry. Frontiers in Cardiovascular Medicine 6:e105

Stephenson M, Lenihan S, Covarrubias R, Huttinger R, Gumina R, Sawyer D, Galindo C (2016) Scanning electron microscopy of macerated tissue to visualize the extracellular matrix. Journal of Visualized Experiments 14:e54005

Stewart S, Hart C, Hole D, McMurray J (2002) A population-based study of the long-term risks associated with atrial fibrillation: 20-year follow-up of the Renfrew/Paisley study. The American Journal of Medicine 113:359-364

Svensson V, Vento-Tormo R, Teichmann S (2018) Exponential scaling of single-cell RNA-seq in the past decade. Nature Protocols 13:599-604

Tadevosyan A, Xiao J, Surinkaew S, Naud P, Merlen C, Harada M, Qi X, Chatenet D, Fournier A, Allen B, Nattel S (2017) Intracellular angiotensin-II interacts with nuclear angiotensin receptors in cardiac fibroblasts and regulates RNA synthesis, cell proliferation, and collagen secretion. Journal of the American Heart Association: Cardiovascular and Cerebrovascular Disease 6:e004965

Takeda N, Manabe I, Uchino Y, Eguchi K, Matsumoto S, Nishimura S, Shindo T, Sano M, Otsu K, Snider P, Conway S, Nagai R (2010) Cardiac fibroblasts are essential for the adaptive response of the murine heart to pressure overload. Journal of Clinical Investigation 120:254-265

Tallquist M (2020) Cardiac fibroblast diversity. Annual Review of Physiology 82:63-78

Tarbit E, Singh I, Peart J, Rose'Meyer R (2019) Biomarkers for the identification of cardiac fibroblast and myofibroblast cells. Heart Failure Reviews 24:1-15

Tello A, Hernandez D, Sanchez M, Valdes M, Marin F (2015) Lone atrial fibrillation - a diagnosis of exclusion. Current Pharmaceutical Design 21:544-550

Thum T, Gross C, Fiedler J, Fischer T, Kissler S, Bussen M, Galuppo P, Just S, Rottbauer W, Frantz S, Castoldi M, Soutschek J, Koteliansky V, Rosenwald A, Basson A, Licht D, Pena T, Rouhanifard S, Muckenthaler M, Tuschl T, Martin G, Bauersachs J, Engelhardt S (2008) MicroRNA-21 contributes to myocardial disease by stimulating MAP kinase signalling in fibroblasts. Nature 456:980-984

Thyagarajan B, Guimarães M, Groth A, Calos M (2000) Mammalian genomes contain active recombinase recognition sites. Gene 244:47-54

Trenor B, Cardona K, Saiz J, Noble D, Giles W (2017) Cardiac action potential repolarization revisited: early repolarization shows all-or-none behaviour. The Journal of Physiology 595:6599-6612

Tsai C, Lai L, Hwang J, Lin J, Chiang F (2008) Molecular genetics of atrial fibrillation. Journal of the American College of Cardiology 52:241-250

van den Berg W, Kawasaki M, Berger W, Neefs J, Meulendijks E, Tijsen A, de Groot J (2017) MicroRNAs in atrial fibrillation: from expression signatures to functional implications. Cardiovascular Drugs and Therapy 31:345-365

van der Velden H, Ausma J, Rook M, Hellemons A, van Veen T, Allessie M, Jongsma H (2000) Gap junctional remodeling in relation to stabilization of atrial fibrillation in the goat. Cardiovascular Research 46:476-486

Vetter M, Blumenthal S, Lindemann R, Manns J, Wesselborg S, Thomssen C, Dittmer J (2005) Ets1 is an effector of protein kinase Calpha in cancer cells. Oncogene 24:650-661

Vincent S, Settleman J (1997) The PRK2 kinase is a potential effector target of both Rho and Rac GTPases and regulates actin cytoskeletal organization. Molecular and Cellular Biology 17:2247-2256

Vire E, Curtis C, Davalos V, Git A, Robson S, Villanueva A, Vidal A, Barbieri I, Aparicio S, Esteller M, Caldas C, Kouzarides T (2014) The breast cancer oncogene EMSY represses transcription of antimetastatic microRNA miR-31. Molecular Cell 53:806-818

Voskoboinik A, McDonald C, Chieng D, O'Brien J, Gutman S, Ngu P, Sugumar H, Wong G, Kalman J, Taylor A, Kistler P (2021) Acute electrical, autonomic and structural effects of binge drinking: insights into the 'holiday heart syndrome. International Journal of Cardiology 331:100-105

Walker J, McLeod K, Kim S, Conway S, Hamilton D (2016) Periostin as a multifunctional modulator of the wound healing response. Cell Tissue Research 365:453-465

Wang J, Wang Y, Han J, Li Y, Xie C, Xie L, Shi J, Zhang J, Yang B, Chen D, Meng X (2015) Integrated analysis of microRNA and mRNA expression profiles in the left atrium of patients with nonvalvular paroxysmal atrial fibrillation: Role of miR-146b-5p in atrial fibrosis. Heart Rhythm 12:1018-1026

Wang T, Larson M, Levy D, Vasan R, Leip E, Wolf P, D'Agostino R, Murabito J, Kannel W, Benjamin E (2003) Temporal relations of atrial fibrillation and congestive heart failure and their joint influence on mortality: the Framingham Heart Study. Circulation 107:2920-2925

Wang X, Zhang X, Yuan J, Wu J, Deng X, Peng J, Wang S, Yang C, Ge J, Zou Y (2018) Evaluation of the performance of serum miRNAs as normalizers in microRNA studies focused on cardiovascular disease. Journal of Thoracic Disease 10:2599-2607

Wang X, Zhang Y, Jiang B, Zhang Q, Zhou R, Zhang L, Wang C (2017) Study on the role of Hsa-miR-31-5p in hypertrophic scar formation and the mechanism. Experimental Cell Research 361:201-209

Wang Y, Cai H, Li H, Gao Z, Song K (2018) Atrial overexpression of microRNA-27b attenuates angiotensin II-induced atrial fibrosis and fibrillation by targeting ALK5. Human Cell 31:251-260

Wolf P, Abbott R, Kannel W (1991) Atrial fibrillation as an independent risk factor for stroke: the Framingham study stroke 22:983-988

Woodall C, Woodall P, Gao E, Yuan A, Koch J (2016) Cardiac fibroblast GRK2 deletion enhances contractility and remodelling following ischemia/reperfusion injury. Circulation research 119:1116-1127

Xu G, Gan T, Tang B, Li Y, Guo X, Abudula A, Abutireheman M, Huo Q (2009) Differential expression of collagen and matrix metalloproteinases between left and right atria in patients with chronic atrial fibrillation. Sheng Li Xue Bao 61:211-216

Xu J, Cui G, Esmailian F, Plunkett M, Marelli D, Ardehali A, Odim J, Laks H, Sen L (2004) Atrial extracellular matrix remodeling and the maintenance of atrial fibrillation. Circulation 109:363-368

Xu J, Wu H, Chen S, Qi B, Zhou G, Cai L, Zhao L, Wei Y, Liu S (2018) MicroRNA-30c suppresses the pro-fibrogenic effects of cardiac fibroblasts induced by TGF-β1 and prevents atrial fibrosis by targeting TGFβRII. Journal of Cellular and Molecular Medicine 22:3045-3057

Yang Y, Kim H, Woo K, Cho D, Bang S (2017) Lipo-PGE1 suppresses collagen production in human dermal fibroblasts via the ERK/Ets-1 signaling pathway. The Public Library of Science12:e0179614

Yata Y, Scanga A, Gillan A, Yang L, Reif S, Breindl M, Brenner D, Rippe R (2003) DNase I–hypersensitive sites enhance α1(I) collagen gene expression in hepatic stellate cells. Hepatology 37:267-276

Yousefi F, Soltani B, Rabbani S (2021) MicroRNA-331 inhibits isoproterenol-induced expression of profibrotic genes in cardiac myofibroblasts via the TGFβ/smad3 signaling pathway. Scientific Reports 11:2548-2548

Yu T, Ma P, Wu D, Shu Y, Gao W (2018) Functions and mechanisms of microRNA-31 in human cancers. Biomed Pharmacotherapy 108:1162-1169

Yue L, Xie J, Nattel S (2011) Molecular determinants of cardiac fibroblast electrical function and therapeutic implications for atrial fibrillation. Cardiovascular Research 89:744-753

Zhang D, Gaussin V, Taffet, Belaguli N, Yamada M, Schwartz R, Michael L, Overbeek P, Schneider M (2000) TAK1 is activated in the myocardium after pressure overload and is sufficient to provoke heart failure in transgenic mice. Nature Medicine 6:556-563

Zhang J, Chu M (2019) Differential roles of VEGF: Relevance to tissue fibrosis. Journal of Cellular Biochemistry 120:10945-10951

Zhang J, Xu D, Li N, Li Y, He Y, Hu X, Lyu L, He L (2017) Downregulation of microRNA-31 inhibits proliferation and induces apoptosis by targeting HIF1AN in human keloid. Oncotarget 8:74623-74634

Zhang Q, Padi S, Tindall D, Guo B (2014) Polycomb protein EZH2 suppresses apoptosis by silencing the proapoptotic miR-31. Cell Death and Disease 5:e1486

Zhao T, Zhao W, Meng W, Liu C, Chen Y, Sun Y (2014) Vascular endothelial growth factor-C: its unrevealed role in fibrogenesis. American Journal of Physiology: Heart Circulation Physiology 306:789-796

Zhao Z, Zhang D, Wu F, Tu J, Song J, Xu M, Ji J (2021) Sophoridine suppresses lenvatinib-resistant hepatocellular carcinoma growth by inhibiting RAS/MEK/ERK axis via decreasing VEGFR2 expression. Journal of Cellular and Molecular Medicine. 25:549-560

Zheng B, Zhang Z, Black C, de Crombrugghe B, Denton C (2002) Ligand-dependent genetic recombination in fibroblasts: a potentially powerful technique for investigating gene function in fibrosis. The American Journal of Pathology 160:1609-1617

Appendices

1. hAF donor 1 certification

Lonza

Certificate of Analysis

eShop Dummy Customer GB
used in eShop only
UNITED KINGDOM

Print Date: 06-Jun-2019

Product Name: NHCF-A Human Cardiac Fibroblasts-Atrial
Material Number: CC-2903
Batch No: 0000634284
Manufacturing Date: 05-Jun-2017

Test	*RESULT*	*SPECIFICATION MIN*	*MAX*	*UNIT*
Tissue Acquisition Number	32677			
DONOR CHARACTERISTICS				
Age	50 Y			
Race	C			
Sex	FEMALE			
VIRUS TESTING				
HIV Test	Not Detected			
HBV Test	Not Detected			
HCV Test	Not Detected			
MICROBIAL TESTING				
Sterility Test	Negative	Negative	***	
Mycoplasma	Negative	***	***	
CELL PERFORMANCE TESTING				
Cell Passage Frozen	2			
Viability-Tryp.Blue Exclusion	84	>=70%	***	%
Cell Count (Cells/ml)	822000	>= 500,000	***	
Total Population Doublings	8	>=5	***	
Seeding Efficiency	67	>=40%	***	%
Doubling Time (hours)	30	15	48	hrs

Lonza Walkersville Inc.
8830 Biggs Ford Road
Walkersville, MD 21792 8415
Tel (301) 898 7025
Fax (301) 845 4024

Lonza

Certificate of Analysis

eShop Dummy Customer GB
used in eShop only
UNITED KINGDOM

Print Date: 06-Jun-2019

Product Name:	NHCF-A Human Cardiac Fibroblasts-Atrial
Material Number:	CC-2903
Batch No:	0000634284
Manufacturing Date:	05-Jun-2017

Test	***RESULT***	***SPECIFICATION MIN***	***MAX***	***UNIT***
Collagen Type I	Pass	Pass: >=90% positive		***
Factor VIII Related Antigen	Pass	Pass: <=10% positive		***

Additional Information:
These cells were isolated from donated human tissue after obtaining permission for their use in research applications by informed consent or legal authorization. This product is for research use only. Details concerning the use of our cell and media products can be downloaded from our website at www.lonza.com/cell-protocols.

Carolyn Carter
Electronically signed by Carolyn Carter
Date: 04-AUG-2017 21:09:58 CET
RELEASE (Inspection Lot: Usage Decision)

Lonza Walkersville Inc.
8830 Biggs Ford Road
Walkersville, MD 21792 8415
Tel (301) 898 7025
Fax (301) 845 4024

2. hAF donor 2 certification

Lonza

Certificate of Analysis

eShop Dummy Customer CH
used in eShop only
CH-

Print Date: 24-Dec-2020

Product Name: NHCF-A Human Cardiac Fibroblasts-Atrial
Material Number: CC-2903
Batch No: 18TL312486
Manufacturing Date: 24-Nov-2018

Test	*RESULT*	*SPECIFICATION MIN*	*MAX*	*UNIT*
Tissue Acquisition Number	36841			
DONOR CHARACTERISTICS				
Age	61 Y			
Race	C			
Sex	MALE			
VIRUS TESTING				
HIV Test	Not Detected			
HBV Test	Not Detected			
HCV Test	Not Detected			
MICROBIAL TESTING				
Sterility Test	Negative	Negative	***	
Mycoplasma	Negative	***	***	
CELL PERFORMANCE TESTING				
Cell Passage Frozen	2'			
Viability-Tryp.Blue Exclusion	96	>=70%	***	%
Cell Count (Cells/ml)	630800	>= 500,000	***	
Total Population Doublings	9	>=5	***	
Seeding Efficiency	64	>=40%	***	%
Doubling Time (hours)	28	15	48	hrs

This lot has been reviewed by Quality Assurance in compliance with requirements of Lonza's Quality System. This document was generated from a validated Part 11-compliant electronic system and thus handwritten signatures are not required.

Lonza
523 Davis Drive Suite 400B
Morrisville, NC, 27560

For Technical Assistance, call 1-800-521-0390

Lonza

Certificate of Analysis

eShop Dummy Customer CH
used in eShop only
CH-

Print Date: 24-Dec-2020

Product Name: NHCF-A Human Cardiac Fibroblasts-Atrial
Material Number: CC-2903
Batch No: 18TL312486
Manufacturing Date: 24-Nov-2018

Test	*RESULT*	*SPECIFICATION MIN*	*MAX*	*UNIT*
Collagen Type I	Pass	Pass: >=90% positive		***
Factor VIII Related Antigen	Pass	Pass: <=10% positive		***

Additional Information:
These cells were isolated from donated human tissue after obtaining permission for their use in research applications by informed consent or legal authorization. This product is for research use only. Details concerning the use of our cell and media products can be downloaded from our website at www.lonza.com/cell-protocols.

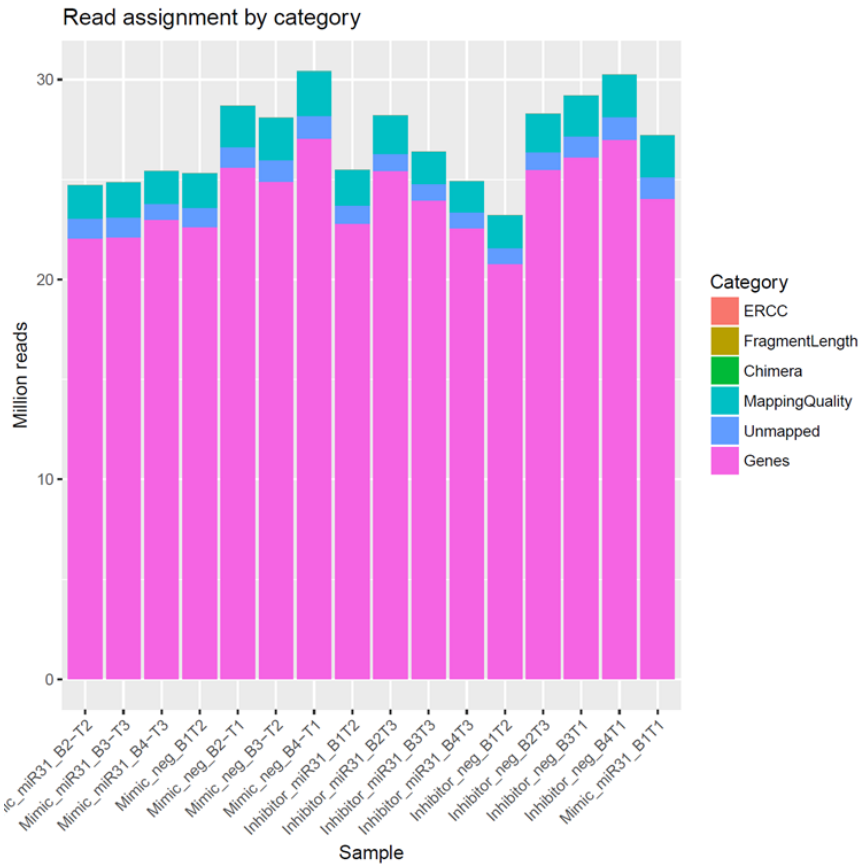
Cheryl Kitchen
Electronically signed by Cheryl Kitchen
Date: 31-JAN-2019 15:12:37 EST
RELEASE (Inspection Lot: Usage Decision)

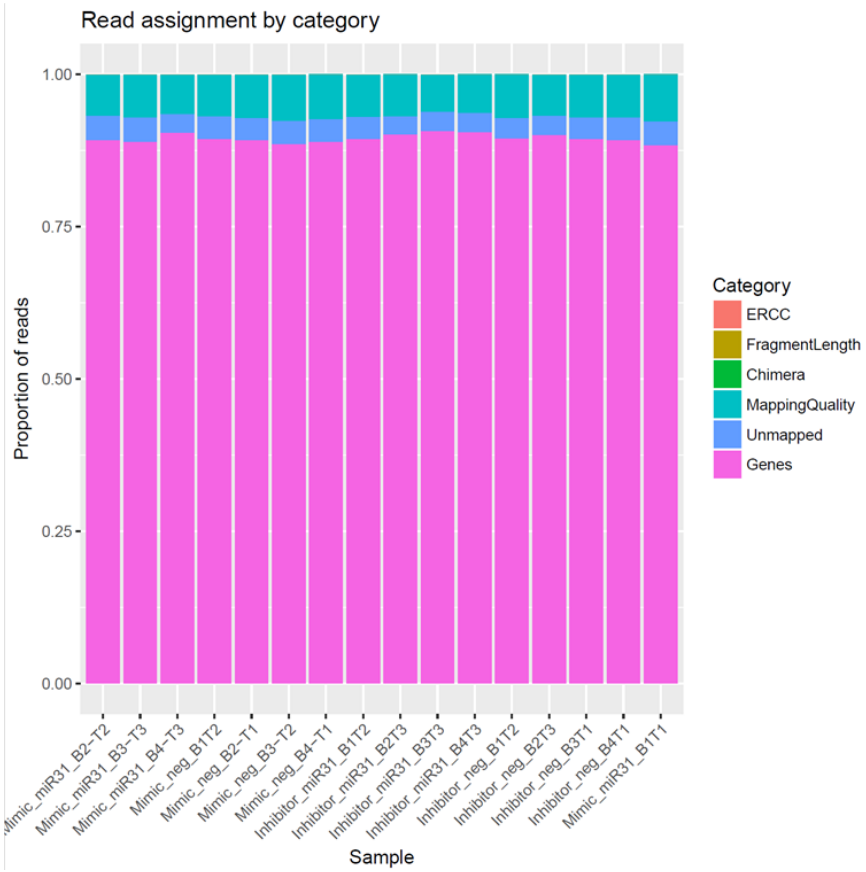
This lot has been reviewed by Quality Assurance in compliance with requirements of Lonza's Quality System. This document was generated from a validated Part 11-compliant electronic system and thus handwritten signatures are not required.

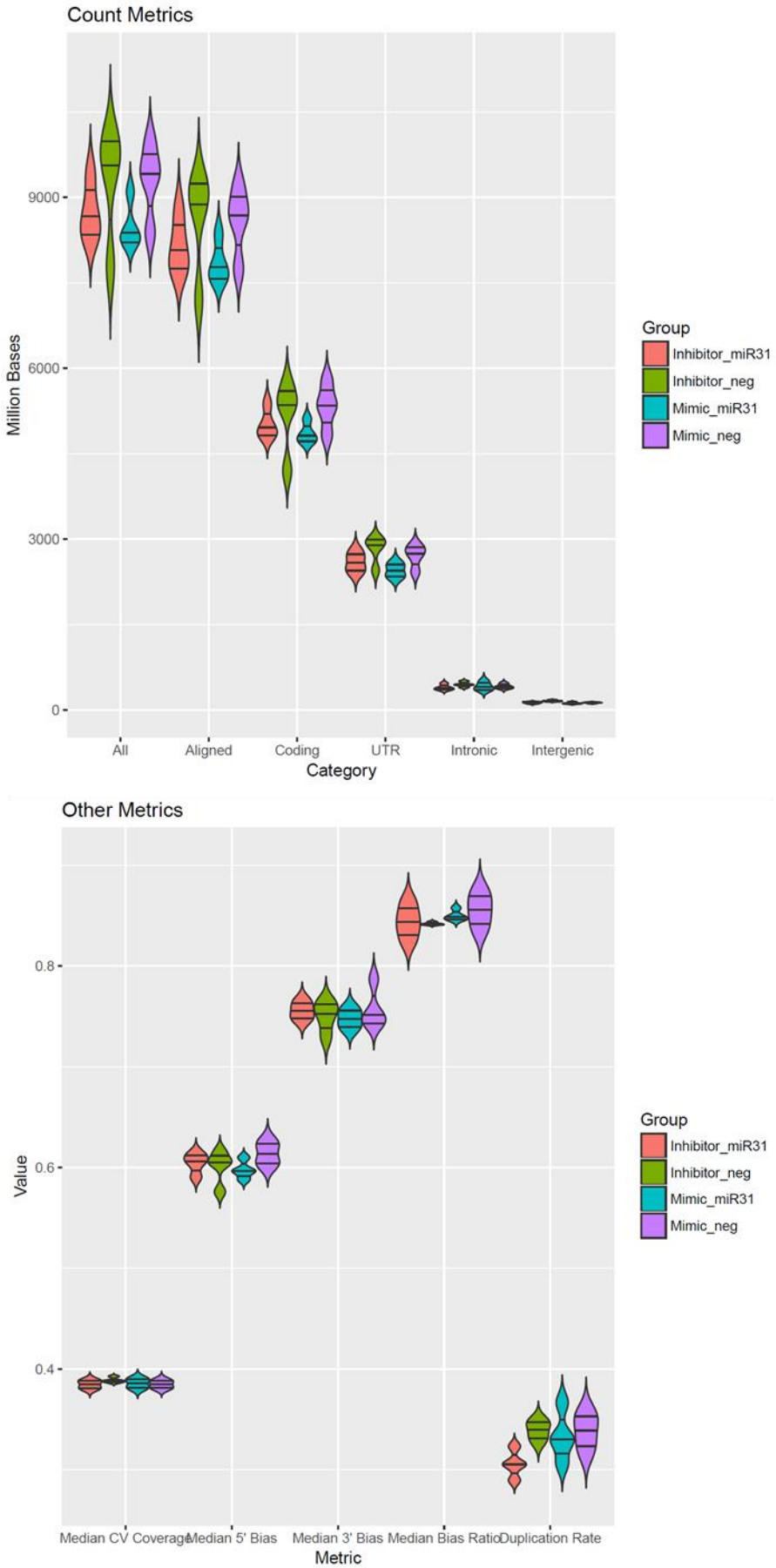
For Technical Assistance, call 1-800-521-0390

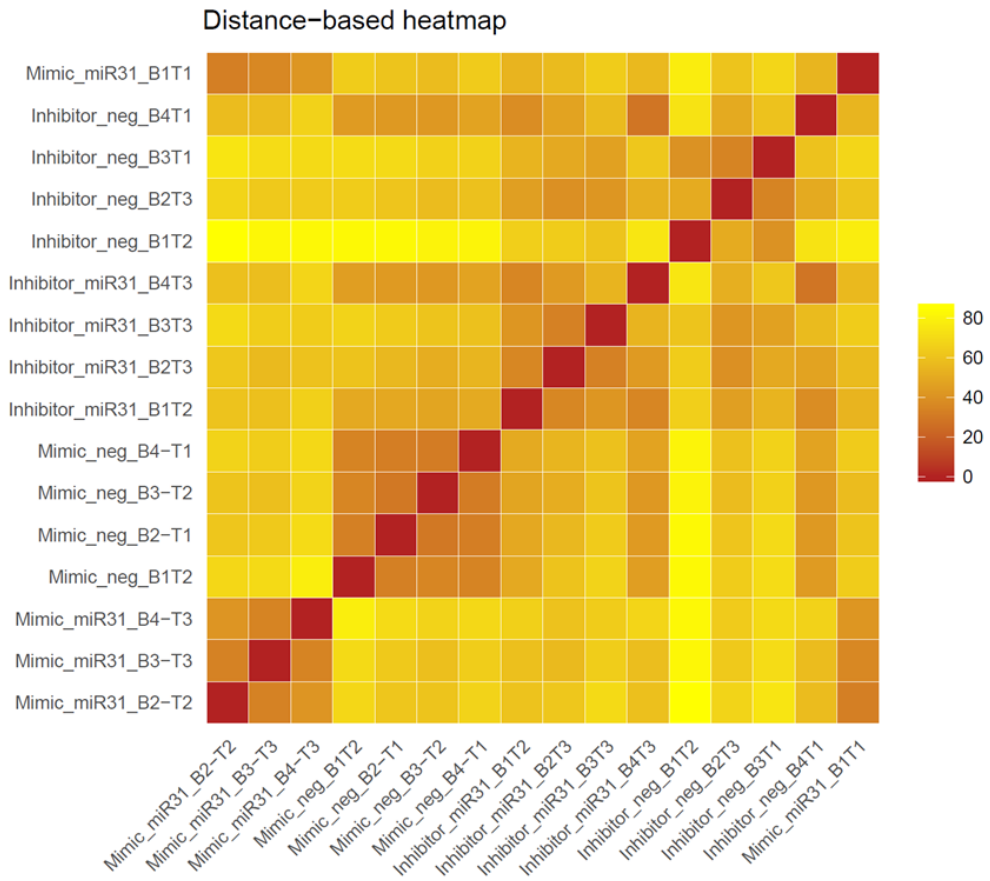
Lonza
523 Davis Drive Suite 400B
Morrisville, NC, 27560

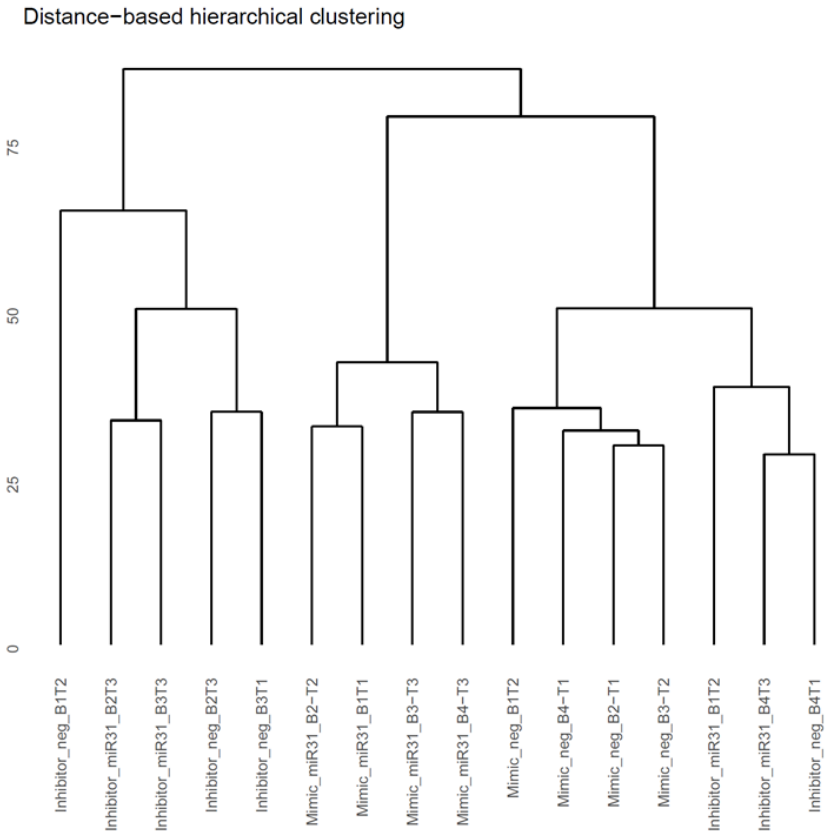
3. RNA-seq quality control

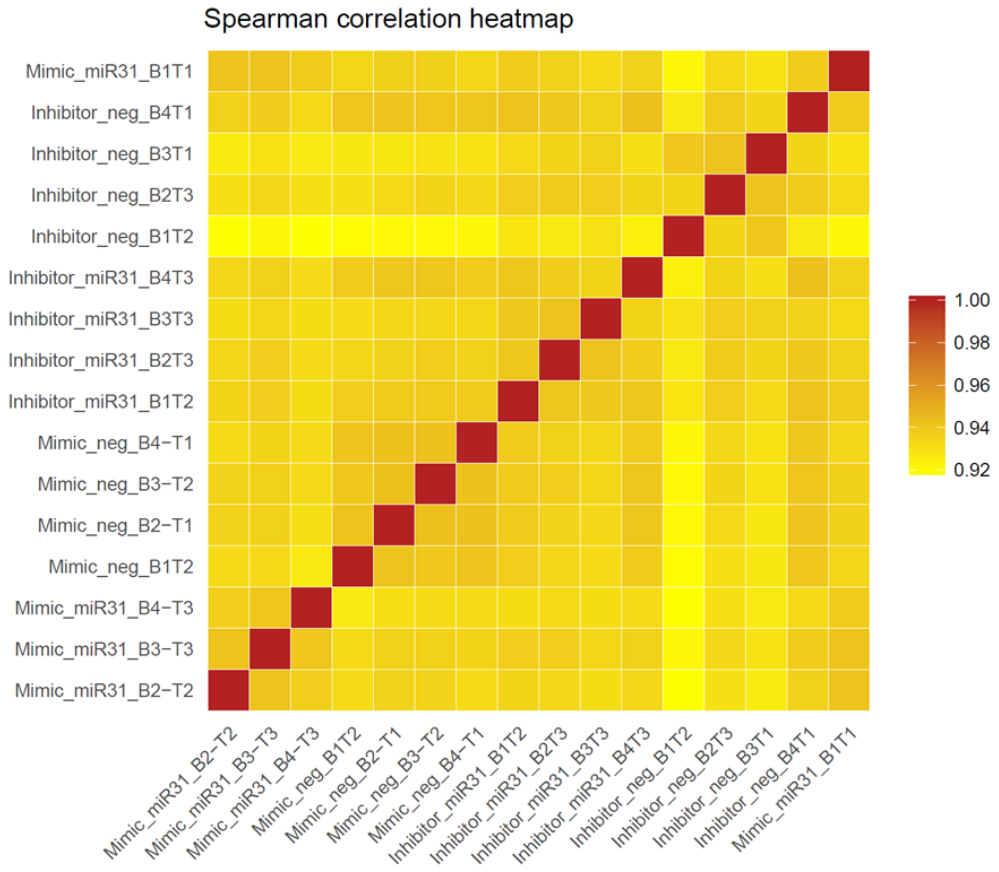


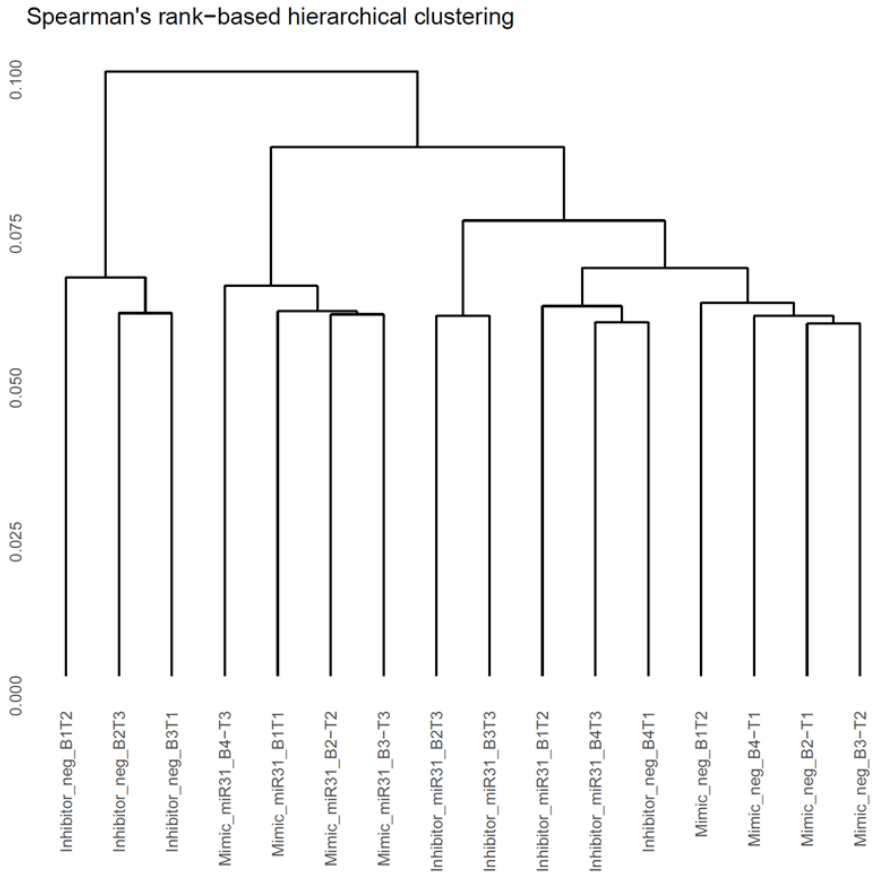


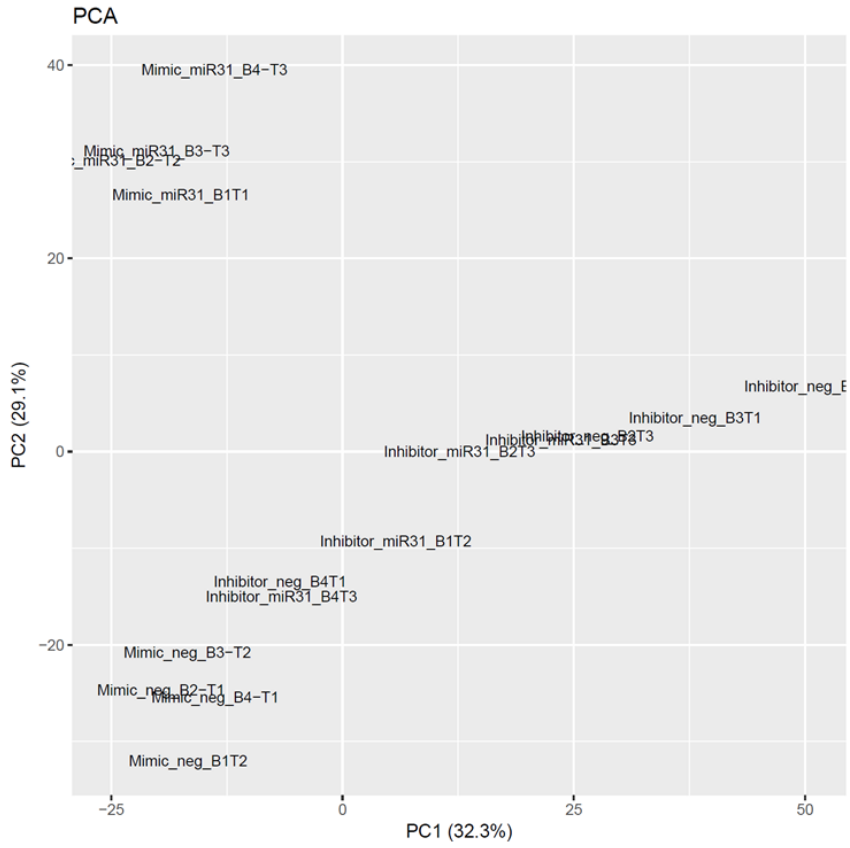

Count Metrics
Million Bases
9000
6000
3000
0
All
Aligned
Coding
UTR
Intronic
Intergenic
Category
Group
Inhibitor_miR31
Inhibitor_neg
Mimic_miR31
Mimic_neg
Other Metrics
Value
0.8
0.6
0.4
Median CV Coverage
Median 5' Bias
Median 3' Bias
Median Bias Ratio
Duplication Rate
Metric
Group
Inhibitor_miR31
Inhibitor_neg
Mimic_miR31
Mimic_neg

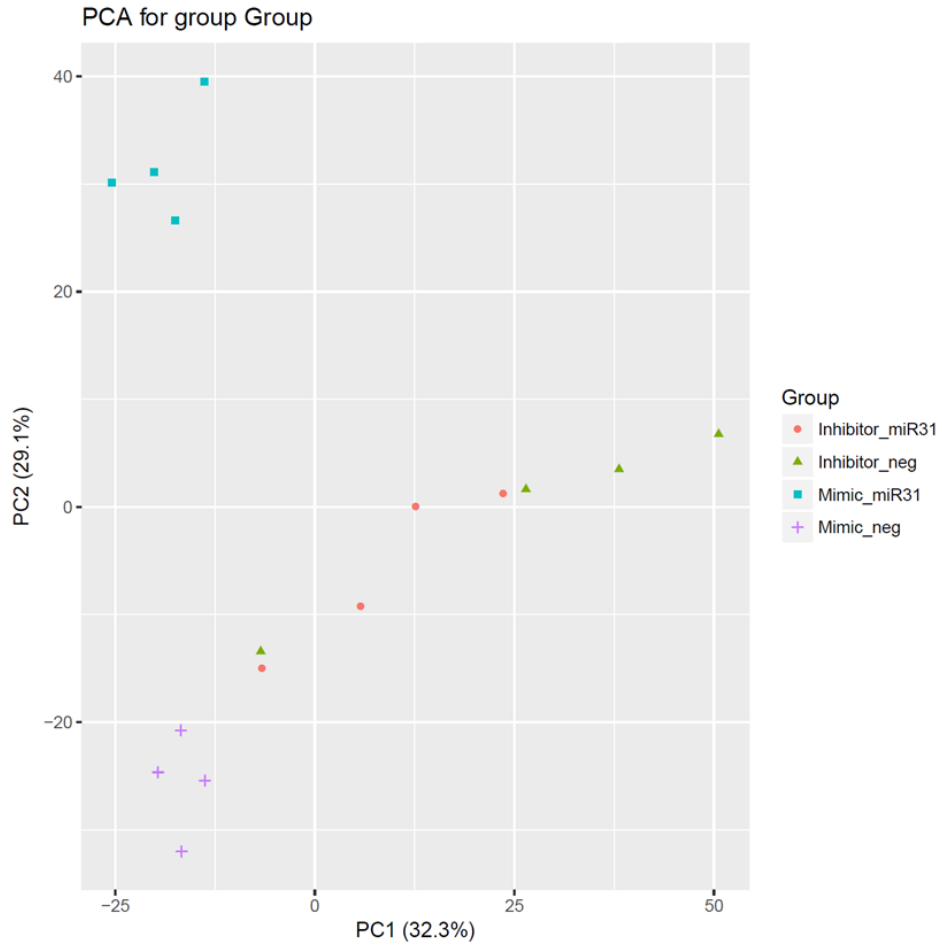

Distance−based heatmap
Mimic_miR31_B1T1
Inhibitor_neg_B4T1
Inhibitor_neg_B3T1
Inhibitor_neg_B2T3
Inhibitor_neg_B1T2
Inhibitor_miR31_B4T3
Inhibitor_miR31_B3T3
Inhibitor_miR31_B2T3
Inhibitor_miR31_B1T2
Mimic_neg_B4−T1
Mimic_neg_B3−T2
Mimic_neg_B2−T1
Mimic_neg_B1T2
Mimic_miR31_B4−T3
Mimic_miR31_B3−T3
Mimic_miR31_B2−T2
80
60
40
20
0
Mimic_miR31_B2−T2
Mimic_miR31_B3−T3
Mimic_miR31_B4−T3
Mimic_neg_B1T2
Mimic_neg_B2−T1
Mimic_neg_B3−T2
Mimic_neg_B4−T1
Inhibitor_miR31_B1T2
Inhibitor_miR31_B2T3
Inhibitor_miR31_B3T3
Inhibitor_miR31_B4T3
Inhibitor_neg_B1T2
Inhibitor_neg_B2T3
Inhibitor_neg_B3T1
Inhibitor_neg_B4T1
Mimic_miR31_B1T1

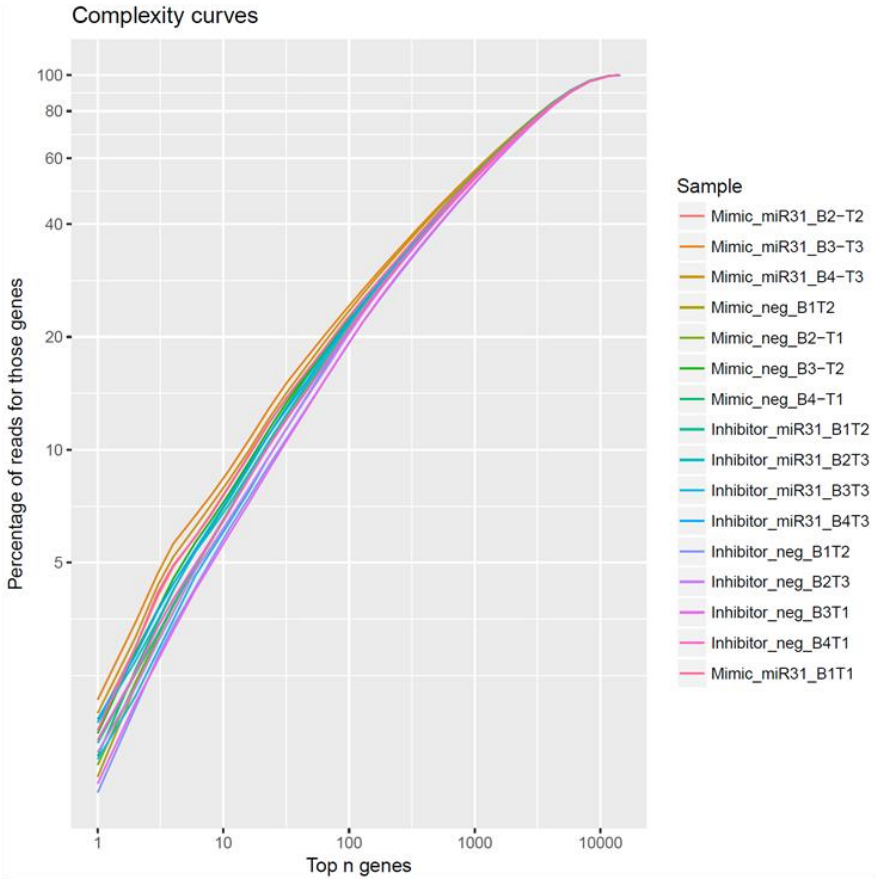

Distance−based hierarchical clustering
0
25
50
75
Inhibitor_neg_B1T2
Inhibitor_miR31_B2T3
Inhibitor_miR31_B3T3
Inhibitor_neg_B2T3
Inhibitor_neg_B3T1
Mimic_miR31_B2−T2
Mimic_miR31_B1T1
Mimic_miR31_B3−T3
Mimic_miR31_B4−T3
Mimic_neg_B1T2
Mimic_neg_B4−T1
Mimic_neg_B2−T1
Mimic_neg_B3−T2
Inhibitor_miR31_B1T2
Inhibitor_miR31_B4T3
Inhibitor_neg_B4T1

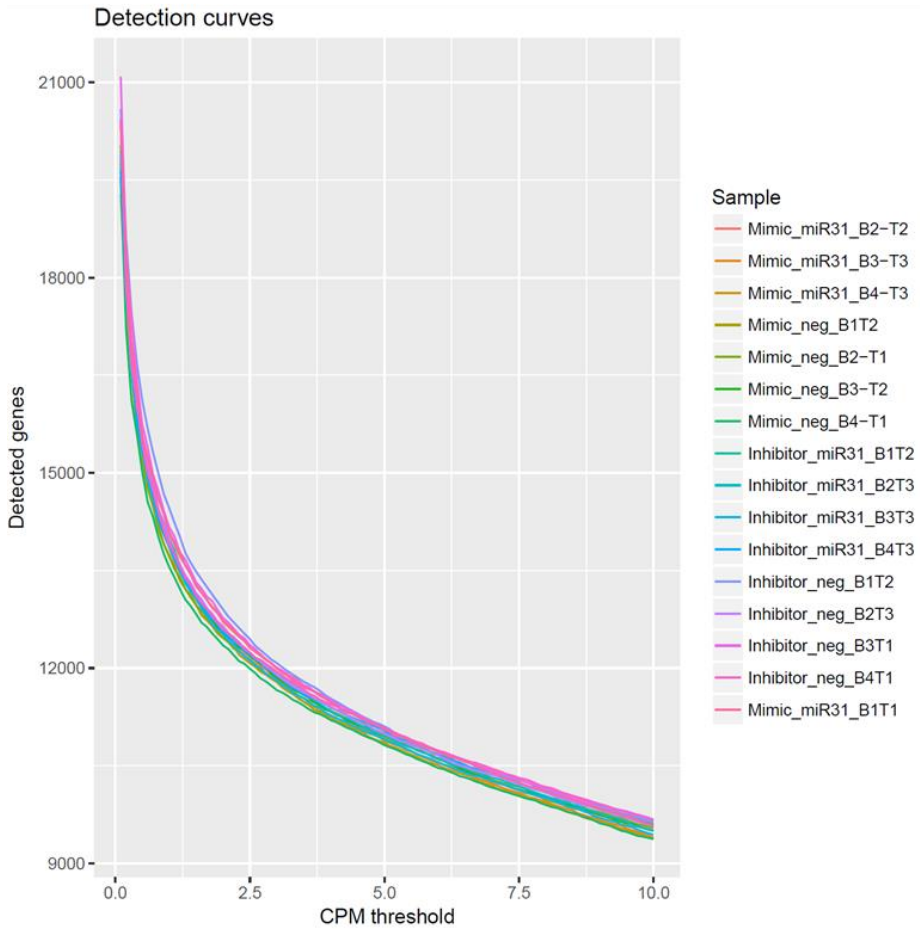

Spearman correlation heatmap
Mimic_miR31_B1T1
Inhibitor_neg_B4T1
Inhibitor_neg_B3T1
Inhibitor_neg_B2T3
Inhibitor_neg_B1T2
Inhibitor_miR31_B4T3
Inhibitor_miR31_B3T3
Inhibitor_miR31_B2T3
Inhibitor_miR31_B1T2
Mimic_neg_B4−T1
Mimic_neg_B3−T2
Mimic_neg_B2−T1
Mimic_neg_B1T2
Mimic_miR31_B4−T3
Mimic_miR31_B3−T3
Mimic_miR31_B2−T2
1.00
0.98
0.96
0.94
0.92
Mimic_miR31_B2−T2
Mimic_miR31_B3−T3
Mimic_miR31_B4−T3
Mimic_neg_B1T2
Mimic_neg_B2−T1
Mimic_neg_B3−T2
Mimic_neg_B4−T1
Inhibitor_miR31_B1T2
Inhibitor_miR31_B2T3
Inhibitor_miR31_B3T3
Inhibitor_miR31_B4T3
Inhibitor_neg_B1T2
Inhibitor_neg_B2T3
Inhibitor_neg_B3T1
Inhibitor_neg_B4T1
Mimic_miR31_B1T1

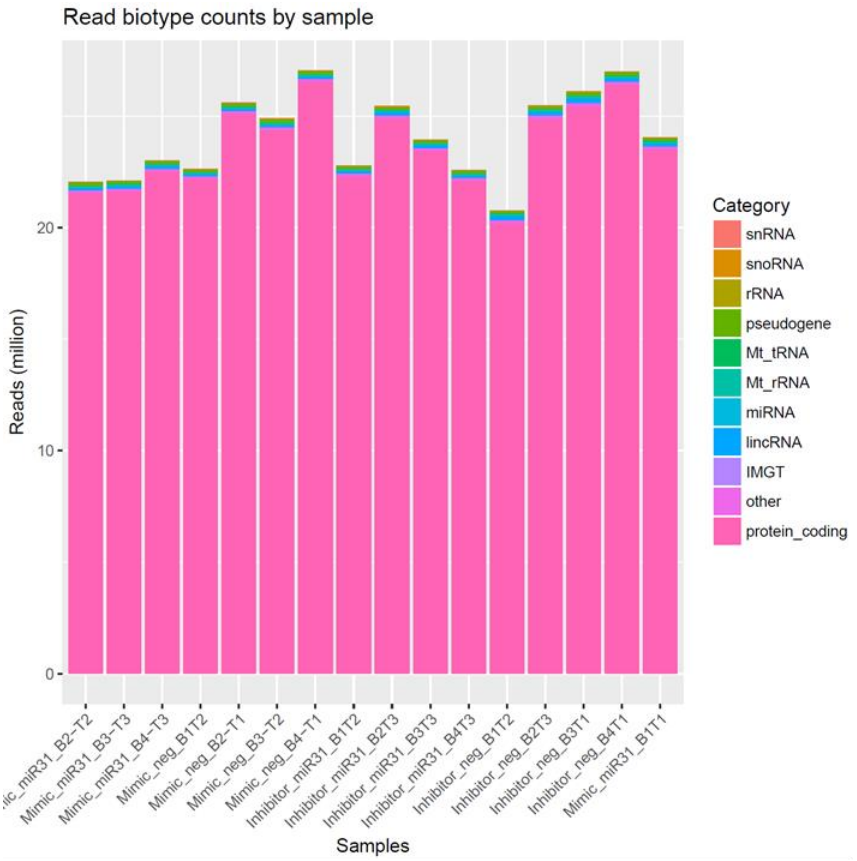

Spearman's rank−based hierarchical clustering
0.100
0.075
0.050
0.025
0.000
Inhibitor_neg_B1T2
Inhibitor_neg_B2T3
Inhibitor_neg_B3T1
Mimic_miR31_B4−T3
Mimic_miR31_B1T1
Mimic_miR31_B2−T2
Mimic_miR31_B3−T3
Inhibitor_miR31_B2T3
Inhibitor_miR31_B3T3
Inhibitor_miR31_B1T2
Inhibitor_miR31_B4T3
Inhibitor_neg_B4T1
Mimic_neg_B1T2
Mimic_neg_B4−T1
Mimic_neg_B2−T1
Mimic_neg_B3−T2

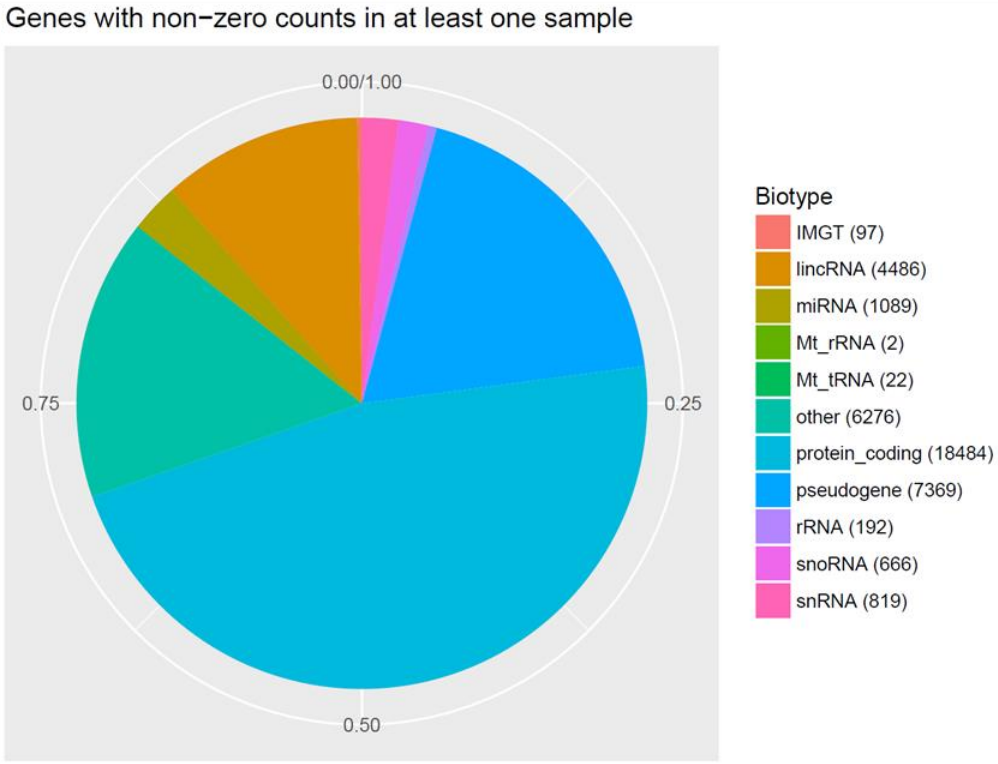

PCA
Mimic_miR31_B4−T3
Mimic_miR31_B3−T3
Mimic_miR31_B1T1
Inhibitor_neg_B3T1
Inhibitor_miR31_B1T2
Inhibitor_neg_B4T1
Inhibitor_miR31_B4T3
Mimic_neg_B3−T2
Mimic_neg_B4−T1
Mimic_neg_B1T2
PC2 (29.1%)
PC1 (32.3%)
40
20
0
−20
−25
0
25
50


PCA for group Group
PC2 (29.1%)
PC1 (32.3%)
Group
Inhibitor_miR31
Inhibitor_neg
Mimic_miR31
Mimic_neg
40
20
0
−20
−25
0
25
50


Complexity curves
Percentage of reads for those genes
100
80
60
40
20
10
5
1
10
100
1000
10000
Top n genes
Sample
Mimic_miR31_B2−T2
Mimic_miR31_B3−T3
Mimic_miR31_B4−T3
Mimic_neg_B1T2
Mimic_neg_B2−T1
Mimic_neg_B3−T2
Mimic_neg_B4−T1
Inhibitor_miR31_B1T2
Inhibitor_miR31_B2T3
Inhibitor_miR31_B3T3
Inhibitor_miR31_B4T3
Inhibitor_neg_B1T2
Inhibitor_neg_B2T3
Inhibitor_neg_B3T1
Inhibitor_neg_B4T1
Mimic_miR31_B1T1


Detection curves
Detected genes
21000
18000
15000
12000
9000
0.0
2.5
5.0
7.5
10.0
CPM threshold
Sample
Mimic_miR31_B2−T2
Mimic_miR31_B3−T3
Mimic_miR31_B4−T3
Mimic_neg_B1T2
Mimic_neg_B2−T1
Mimic_neg_B3−T2
Mimic_neg_B4−T1
Inhibitor_miR31_B1T2
Inhibitor_miR31_B2T3
Inhibitor_miR31_B3T3
Inhibitor_miR31_B4T3
Inhibitor_neg_B1T2
Inhibitor_neg_B2T3
Inhibitor_neg_B3T1
Inhibitor_neg_B4T1
Mimic_miR31_B1T1


Read biotype counts by sample
Reads (million)
20
10
0
Samples
Category
snRNA
snoRNA
rRNA
pseudogene
Mt_tRNA
Mt_rRNA
miRNA
lincRNA
IMGT
other
protein_coding


Genes with non−zero counts in at least one sample
0.00/1.00
0.25
0.50
0.75
Biotype
IMGT (97)
lincRNA (4486)
miRNA (1089)
Mt_rRNA (2)
Mt_tRNA (22)
other (6276)
protein_coding (18484)
pseudogene (7369)
rRNA (192)
snoRNA (666)
snRNA (819)

4. 477 predicted mRNA targets of miR31 by TargetScan7.2

RNF144B, RSBN1, SH2D1A, AK4, PAX9, LPP, PRKCE, NR5A2, TMEM145, ARHGEF2, WDR5, PSMB11, PEX5, FGF7, SYDE2, PDZD2, SSH1, TFRC, IL34, CIAPIN1, PC, TMPRSS11F, CRYBG3, EGLN3, TBXA2R, HIAT1, CACNG4, UCN2, UBE2K, TESK2, HIF1AN, SERTAD2, C17orf78, MGAT1, KHDRBS3, RGS4, FAM73A, AHCYL1, SPARC, KANK1, DCBLD2, FAM53B, FLOT1, AP000695.1, LBH, TMED10, SEMA6D, AMER1, FAM60A, SEPHS1, PAX5, RASA1, HOMER1, CD28, GXYLT1, NDRG3, MMP16, RNF144A, CREG1, ISL1, TIFAB, PRKAR2A, PIK3C2A, KCTD21, SGMS1, GCNT4, CAMK2D, VAPB, VEZT, KRT80, SLC1A2, JMJD8, VPS26B, NUMB, KCNN3, MIXL1, OAS2, KIAA1462, PCDH8, APBB2, IDE, TMEM43, CACYBP, PXK, MAP4K5, BMPER, LATS2, PTPLAD1, STK40, RNF183, PRELP, MBOAT2, FAM118A, PPP6R3, RAB31, KDELR2, EDNRB, GABBR1, PTGFRN, VTI1A, STX12, DDTL, ENY2, CCNC, RGAG1, RHOBTB1, BAHD1, FZD4, GCH1, MFAP3, DUSP7, ZFP30, OPRL1, WNK1, CHMP7, WBP11, SIKE1, SATB2, CCNT1, SLC6A6, GCLM, SLC2A4, KIAA1024, PPP2R2A, ZKSCAN8, ABLIM1, CLASP2, SLC15A4, UCK2, GJD2, SYT6, SLC31A1, MAP1B, GSTCD, ATL2, SIM1, SUPT16H, PYURF, SRC, PPP6C, HCN1, PHF21B, EMP1, COPS2, BACH2, RAB9B, ITPA, CDC42SE1, ELAVL1, SOX11, TMEM168, EHBP1, PPP3CA, ALKBH1, GLB1L, TACC1, YWHAE, MTFR1, TIA1, SLC30A6, TMX4, LITAF, MRPL48, ZBTB34, KLF13, ANKFY1, XKR6, SH3BGRL2, VAV3, NUP153, PDE4D, POU2F1, CEBPA, GSG1L, ST3GAL2, STARD13, QKI, SRSF10, AP2B1, PREPL, MYO19, RBM23, NCK2, FZD3, ADCY6, PSMB2, ARHGEF26, CDK1, EIF5, INSC, BOD1L2, MMS19, CYSLTR2, AGRN, SLC43A2, ZSWIM6, SALL2, APP, RBM8A, RBMS1, DPYSL5, SH3BP1, LRRC55, TENM4, GPRC5A, CLOCK, ALX4, GNA13, RAB27A, SCN2A, FXR1, TTC31, GLTSCR1, ARID1A, VPS39, PRSS8, KIAA1324L, AK2, SNX22, ZFP36L1, KLF7, DGCR8, FAM134B, AMOTL1, SAMD12, HHIP, SNN, SAMD10, EBF3, MYO5A, NUFIP2, VAMP4, RDX, TNS1, ZNF384, CNNM1, RPH3A, PPP1R9A, ETV3, PRKAA2, ZDHHC18, HSPD1, TACC2, AP1S1, HERPUD2, GRB10, PDZD4, DPY19L3, MBNL3, CTNND2, GPR137, OXSR1, C11orf87, RIMKLB, EIF1, ZDHHC20, AL590483.1, PLEKHA1, C17orf102, GPRIN3, FAM135B, NR2C2, ARID2, UBN1, SLC5A3, OSBP2, JAZF1, SPRED1, ARRDC3, EGLN1, IRGQ, FITM2, SLC17A8, CCDC42B, AXIN1, CCNJ, TM9SF3, CKAP4, ROCK1, PKN2, RYBP, CRISPLD2, FAM78A, HELZ, SLC24A3, SYDE1, AP000350.4, PAPOLA, CAMK2G, CTIF, RBM12, AGO3, CEP85L, CLSTN2, CLCN3, CEP85, KCTD20, ZFHX4, PVRL3, MAP3K1, GABRA3, CACUL1, TEX261, PANK3, SS18L1, AFF1, RREB1, ATF7IP, FBXL4, PARPBP, RAB5B, NFIC, C2ORF15, ATP8A1, CDIP1, GRIK3, DGKH, STAM2, ZC3H12C, FUBP1, SMIM13, RSF1, ANKRD52, ZSCAN25, KLF3, LRP10, PDPR, E2F2, FIGN, SLC39A13, VPS53, CACNB2, SNTB2, ARHGEF15, CHD7, PDE11A, UACA, EML5, MLEC, BAP1, C9orf114, PDPK1, CBL, TLN2, YPEL2, KIAA0226, WDR37, KIAA1211, PDE12, ELAVL2, RIMS3, EIF4E3, PTPRG, HEYL, RABL3, TTBK2, POLR3H, FRK, SLC39A9, CALCR, RAB19, BICD1, TTC5, PHF12, BMPR1A, SMARCD1, IFFO2, FCHSD2, KDM5A, SFXN1, DPP8, ABL2, NCOA2, HBEGF, DACH1, IKZF1, VWA3A, CBX5, MYCBP, UBASH3B, RAP2B, PLXNA4, LSAMP, TMEM41A, ATF6, MAFG, PLXNA2, IPO11, EDC3, FAM122A, GJA9, SLC46A1, NEK9, KIF1C, KIAA0141, NFAT5, ZNF704, PPL, CPSF2, NEDD4L, MOCS1, FMNL3, IRS4, GATAD2B, ALDH6A1, DGKI, DMD, COL5A1, CTH, SRGAP3, CSRNP3, PIGV, KCNC3, ATP11A, UBE2N, EPC1, IQSEC2, MLXIP, MTF1, CCDC113, NHS, EPB41L5, ARRB1, RAB6B, ZDHHC21, CHSY1, EIF5A2, GDAP1L1, LMAN2, STX1B, FRMPD4, TMEM260, EVC, WDR12, TAOK1, RHBDL3, STAU2, SF3B3, CASD1, TMEM220, ABCG8, GPR108, SRCAP, YARS2, VGLL3, DICER1, ATP5S, SLC25A16, USP49, RRP8, COQ7, FTO, BCL9L, MAGIX, POP4, AHSA2, RBM48, GTDC1, PHKG2, LPXN, TPD52, YLPM1, SLC45A4, CYB5B, WWC2, ZNF574, HNRNPA0, CENPP, CCS, ZBTB20, FNDC5, ZNF703, EMC7, ITIH5, CNTD1, CSMD1, ZNF652