Investigating the complexities of microRNA-target interactions in cardiac biology

Carly Hynes

January 2014

A thesis submitted for the degree of

Doctor of Philosophy

of the Australian National University

ORIGINALITY STATEMENT

I hereby declare that this submission is my own work and to the best of my knowledge it contains no materials written by another person, or substantial proportions of material which have been accepted for the award of any other degree or diploma at ANU or any other educational institution, except where due acknowledgement is made in thesis. Any contribution made to the research by others, with whom I have worked at ANU or elsewhere, is explicitly acknowledged in this thesis. I also declare that the intellectual content of this thesis is the product of my own work, except to the extent that assistance from others in the project's design and conception or in style, presentation and linguistic expression is acknowledged.

Signed .....

Date 12/11/2014.....

ABSTRACT

One in three Australians die of cardiovascular disease and despite advances in treatment it remains the leading cause of death in Australia. MiRNAs play critical roles in the heart and specific miRNAs can promote or prevent cardiac hypertrophy, sparking interest to use miRNAs as therapeutic targets. However, cells can produce both miRNAs and their mRNA targets in multiple processing variants increasing the complexity of miRNA-mediated control. It is therefore critical to elucidate the role of these variants in cardiac disease. The aim of this thesis was to profile miRNA and mRNA 3' UTRs in cardiac hypertrophy and to interrogate this data for evidence of processing variation.

We established a murine model of left ventricular hypertrophy using transverse aortic constriction as confirmed using haemodynamic, physiological and biochemical measurements. Cardiomyocytes were isolated from pre-hypertrophic, hypertrophic and control hearts and RNA extracted for next-generation sequencing of miRNAs and mRNA 3' ends. We identified 17 miRNAs and 776 mRNAs that changed in level during the hypertrophy, with the majority being upregulated. Overall, more mRNAs changed expression in the pre-hypertrophic cardiomyocytes, while the miRNAs were altered mostly in the hypertrophic samples. This is the first study to document the expression of miRNAs and mRNAs simultaneously in purified cardiomyocytes and their change in hypertrophy.
Analysing the datasets for processing diversity revealed that nearly all miRNAs in cardiomyocytes showed some degree of processing diversity including expression from both hairpin arms, 5' and 3' isomiRs, or the presence of non-templated additions. There was little directional change in miRNA processing during hypertrophy but two examples were identified where the bias for particular variants changed significantly. This is the first documentation of variable miRNA processing in purified cardiomyocytes. Such variants provide a mechanism to change the mRNA targeting properties of miRNAs from a given locus without changing the genomic sequence. In support of this, *in vitro* luciferase assays demonstrated that the different 5' isomiRs of the cardiac miR-133a-3p have distinct targeting preferences.

Analysis of the 3' end sequencing data demonstrated that mRNAs expressed in cardiomyocytes had on average 2.35 distinct 3' UTRs with an average length of 3.7 kb. The proportion of short versus long 3' UTRs was altered for 424 mRNAs, with 272 3' UTRs getting shorter and 152 getting longer, in the pre-hypertrophic cardiomyocytes. This is the

first study to provide an in-depth mapping of mRNA 3' UTR diversity in purified cardiomyocytes and document changes to the proportion of these variants in cardiac hypertrophy. Alterations in the length of 3' UTRs can lead to inclusion of or exclusion of miRNA target sites on mRNAs and thus change their sensitivity to regulation by miRNAs. Overall, the datasets reported here provide global information on expression changes to both miRNA sequence and mRNA 3' UTR lengths forming the basis for an improved systems level understanding of miRNA-regulation during cardiac hypertrophy. The realisation that cardiac miRNAs and their targets exist as currently under-appreciated variants with potentially complex effects on target specificities has important implications for the role of miRNAs in cardiac disease.

ACKNOWLEDGEMENTS

My thesis is now complete. Wow. I feel a mix of emotions (dazed, excited, exhausted to name a few). It seemed at times as though this PhD would never end, but here I am at the finish line of my PhD journey. I could not have finished this journey alone so I would like to say thank you to all those that have helped me along the way.

To Prof. Thomas Preiss, thank you for the guidance, encouragement and opportunities you have offered. Special thanks for reading my thesis drafts over the Christmas holidays.

To my co-supervisor Cath Suter, thanks for your understanding and your support. Thank you for your encouraging emails and phone calls, I am incredibly grateful.

To Nicola Smith, words cannot express how thankful I am to have you as both a mentor and a friend. Thank you for listening to me whinge and cry, for giving me invaluable advice, for reading everything I've ever written and for your brutal honesty (I know of no one who can edit a document in quite the same way).

To Jennifer Clancy, Hardip Patel and Brian Parker. Thanks for your contribution to this work, it was a tough project to crack and without your help I wouldn't have gotten to the end.

To Terry Neeman, thanks for your invaluable advice about statistical models, without which I would not have been able to complete this thesis.

To Tennille Sibbritt, Eloisa Pagler and Yalin Liao. Thank you for all the coffee breaks, lunch breaks, picnics and dinners that I needed to keep me sane. Thank you for listening to me whinge, and for sharing my laughs and tears. Special thanks to Tennille, who has managed to study, work, gym, holiday and live with me for 2 years and we're still the best of friends! I couldn't have done it without you.

This Journey began at the Victor Chang Institute in Sydney, a place that will always bring back fond memories. To all of the friends I made there; Ali McCorkindale, Stan Artap, Ting Wai Yiu, Alistair Stewart, Sara Ballouz and Chu Kong. Thanks for the laughs, the smiles and understanding the pain that comes with writing a PhD.

Thankyou to Prof Robert Graham and his laboratory for allowing me to tag along with their animal experiments. Thanks to Jianxin Wu for performing all of the animal surgeries, and to Ming Li and Sara Holman for helping me with the cardiomyocyte isolations.

To the other members of the Preiss lab who I have not yet mentioned. Those who I met when I started this journey at the Victor Chang Institute and those who were at the finish line at ANU, I offer my thanks. David Humphreys (who taught me everything I know about cloning and sequencing preparation), Grace Wei, Stuart Archer, Rina Soetanto (I wish you all the luck in continuing this work), Andrew Shafik, Natalie Beveridge and Hao Yan, thank you.

Finally, thanks to all my family and friends who were there when science got me down. I'm sorry for neglecting you and I'm so thankful for your understanding when I said "I can't I'm writing.". I would like to thank my Mum and Dad for supporting me in everything I do, without you I would not have made it this far and I hope I've made to proud. The biggest thanks go to Chris, my Fiancé, for putting up with me during my PhD (I know it hasn't been easy). Thank you for the pep talks when I just wanted to give up and for my PhD motto "When you're going through hell, keep going". I kept going because of your support and now I've crossed the finish line :).

Conferences, Publications and Awards

Conferences

Poster presentations:

Hynes C.J., Patel H.R., Humphreys D.T., Smith N.J., Beilharz T., Graham R.M., Clancy J.L., Preiss T. (2013) At the heart of the tail: microRNA variants and alternative polyadenylation in cardiac hypertrophy. *ComBio 2013, Perth, Australia.*

Hynes C.J., Humphreys D.T., Smith N.J., Patel H.R., Graham R.M., Clancy J.L., Preiss T. (2013) Shifting targets: microRNA variants and alternative polyadenylation in cardiac hypertrophy. *18th Annual Meeting of the RNA Society, Davos, Switzerland.*

Hynes C.J., Humphreys D.T., Smith N.J., Patel H.R., Graham R.M., Clancy J.L., Preiss T. (2013) Moving targets: microRNA and mRNA variants in cardiac hypertrophy. *Keystone Symposia "Noncoding RNAs in Development and Cancer", Vancouver, Canada.*

Hynes C.J., Humphreys D.T., Clancy J.L., Patel H.R., Suter C., Preiss T. (2012) The role of microRNA variants and alternative polyadenylation in cardiac hypertrophy. *Lorne Genome Conference 2012, Lorne, Australia.*

Hynes C.J., Humphreys D.T., Clancy J.L., Patel H.R., Suter C., Preiss T. (2011) Exploring the role of miRNA and target processing variability in cardiac hypertrophy. *EMBO Conference on Protein synthesis and translational control, Heidelberg, Germany.*

Hynes C.J., Clancy J.L., Patel H.R., Suter C., Humphreys D.T., Preiss T. (2011) Exploring the role of miRNA and target processing variability in cardiac hypertrophy. *ASMR Medical Research Week, ACT New Investigators Forum, Canberra, Australia.*

Hynes C.J., Clancy J.L., Patel H.R., Suter C., Humphreys D.T., Preiss T. (2011) Exploring the role of miRNA and target processing variability in cardiac hypertrophy. *Lorne Genome Conference 2011, Lorne, Australia.*

Hynes C.J., Humphreys D.T., Clancy J.L., Patel H.R., Wei G., Preiss T. (2010) *St Vincents and Mater Health Sydney Research Symposium, Sydney, Australia.*

Oral presentations:

Hynes C.J., Humphreys D.T., Smith N.J., Patel H.R., Graham R.M., Clancy J.L., Preiss T. (2012) The tell-tail heart: miRNA variants and 3' UTRs in cardiac hypertrophy. *ComBio 2012, Adelaide, Australia.*

Hynes C.J., Humphreys D.T., Smith N.J., Patel H.R., Graham R.M., Clancy J.L., Preiss T. (2012) A tail of the heart: miRNAs and 3' UTRs in cardiac hypertrophy. *ASMR Medical Research Week, ACT New Investigators Forum, Canberra, Australia.*

Hynes C.J., Humphreys D.T., Smith N.J., Patel H.R., Graham R.M., Clancy J.L., Preiss T. (2011) At the heart of the tail: 3' UTRs and isomiRs in cardiac hypertrophy. *Sydney next-generation sequencing special interest group meeting, Sydney, Australia.*

Hynes C.J., Clancy J.L., Patel H.R., Suter C., Humphreys D.T., Preiss T. (2011) Exploring the role of miRNA and target processing variability in cardiac hypertrophy. *NSW miRNA users workshop, Sydney, Australia.*

PUBLICATIONS

Hynes, C.J., Clancy, J.L. and Preiss, T. (2012) microRNAs in cardiac disease: sitting duck or moving target? IUBMB Life, 64(11).

Humphreys, D.T., **Hynes, C.J.**, Patel, H.R., Wei, G.H., Cannon, L., Fatkin, D., Suter, C.M., Clancy, J.L. and Preiss, T. (2012) Complexity of murine cardiomyocyte miRNA biogenesis, sequence variant expression and function. PLoS ONE, 7(2).

AWARDS

I was privileged to receive the following throughout my PhD studies:

2013:

- ASBMB Fellowship to attend the 18th Annual meeting of the RNA Society in Davos, Switzerland (ASBMB).
- ASBMB student poster prize, ComBio 2013 (ASBMB).

2012:

- Heart Foundation Travel Grant award (The Heart Foundation) to attend the Keystone Symposium "Noncoding RNAs in Development and Cancer" in Vancouver, Canada.
- ASBMB Student member ComBio 2012 Bursary (ASBMB).
- 2012 Winter School Student Travel Bursary (ARC Centre of Excellence in Bioinformatics, University of Queensland, Bioplatforms Australia and EMBL Australia).

2011:

- EMBL Corporate partnership registration fee fellowship (EMBO Conference on Protein Synthesis and Translational Control, Heidelberg, Germany).

COMMONLY USED ABBREVIATIONS

A-to-I	Adenosine to Inosine
Actc1	Actin, alpha, cardiac muscle 1
ADAR	Adenosine Deaminase acting on RNA
AGO	Argonaute
α-MHC	Alpha myosin heavy chain
Ang II	Angiotensin II
ANP	Atrial natriuretic peptide
APA	Alternative polyadenylation
ARF	ADP ribosylation factor
AS	Aortic stenosis
ATF-2	Activating transcription factor 2
α-SkA	Alpha skeletal actin
auxDSE	Auxiliary downstream element
β-AR	Beta adrenergic receptor
Bcl2	B-cell lymphoma 2
BMP	Bone morphogenetic protein
BNP	Brain natriuretic peptide
β-MHC	Beta myosin heavy chain
BP	Blood pressure
BPM	Beats per minute
CaMK	Calmodulin-dependent protein kinase
CDC42	Cell division cycle homolog 42
cDNA	Complementary DNA
CFIm	Cleavage factor I
CFIIm	Cleavage factor II
cGMP	Cyclic guanosine monophosphate
CnA	Calcineurin
Col1a1	Collagen, type I, alpha 1
Col1a2	Collagen, type I, alpha 2
Col3a1	Collagen, type III, alpha 1
CPM	Counts per million
CPSF	Cleavage and polyadenylation specificity factor
CstF	Cleavage stimulation factor
CTGF	Connective tissue growth factor
CVD	Cardiovascular disease
DAG	Diacyl glycerol
DCM	Dilated cardiomyopathy
Ddr2	Discoidin domain receptor tyrosine kinase 2
DGCR8	DiGeorge syndrome critical region 8
DSE	Downstream element
ECM	Extracellular matrix
edgeR	Empirical analysis of digital gene expression in R
EDP	End diastolic pressure
eEF2	Eukaryotic elongation factor 2
eIF2B	Eukaryotic initiation factor 2B
e-PCR	Emulsion PCR
EST	Expressed sequence tag
ERK	Extracellular-signal-regulated kinase
ET-1	Endothelin 1
FC	Fold change
FDR	False discovery rate
FGF	Fibroblast growth factor
GDP	Guanosine diphosphate
GO	Gene ontology

GPCR	G protein-coupled receptor
G protein regulatory protein	Guanine-nucleotide
GSK3β	Glycogen synthase kinase 3β
GTP	Guanosine triphosphate
HDAC	Histone deacteylases
HITS-CLIP	High-throughput sequencing of RNAs isolated by cross-linked immunoprecipitation (IP)
Hsp70	Heat shock protein 70
ICM	Ischemic cardiomyopathy
IGF	Insulin growth factor
IL-6	Interleukin-6
iPS	Induced pluripotent stem cells
IP_3	Inositol 1,4,5-trisphosphate
JNK	c-Jun N-terminal kinases
LV	Left ventricle
MAPK	Mitogen activated protein kinase
MAPKKKs	MAP kinase kinase kinases
MAPKKs	MAP kinase kinase
MBNL1	Muscleblind-like splicing regulator 1
MDS	Multi-dimensional scaling
MEF2	Myocyte enhancing factor 2
MEK	Mitogen-activated protein kinase kinase
Min	Minute
miRNA	MicroRNA
Mlc2v	Myosin light chain 2v
mRNA	Messenger RNA
mTOR	Mammalian target of rapamycin
NFATc	Nuclear factor of T cells
NGS	Next generation sequencing
nt	Nucleotide
NTA	Non-templated additions
PACdb	Poly(A) cleavage site and 3' UTR database
PAP	Poly(A) polymerase
PAS	Polyadenylation signal
PAR-CLIP	Photoactivatable ribonucleoside-enhanced cross-linking and IP
PCR	Polymerase chain reaction
PDK-1	Phosphoinositide-dependent kinase-1
PE	Phenylephrine
PGAM1	Phosphoglycerate mutase 1
PIP_2	Phosphatidylinositol 4,5-bisphosphate
PI3K	Phosphoinositide 3-kinase
$PI(3,4,5)P_3$	Phosphatidylinositol-3,4,5-triphosphate
$PI(4,5)P_2$	Phosphatidylinositol-4,5-bisphosphate
PKC	Protein kinase C
PKD	Protein kinase D
PKG-1	Protein kinase G-1
PLC	Phospholipase C
Poly(A)	Polyadenylation
PPARγ	Peroxisome proliferator-activated receptor Y
Pre-miRNA	Precursor miRNA
Pri-miRNA	Primary miRNA
QV	Quality value
Rab	Ras-related gtp-binding protein
Rac1	Ras-related C3 botulinum toxin substrate 1
Raf	Rapidly accelerated fibrosarcoma
Ran	Ras-related nuclear protein
RAS	Renin angiotensin system

RhoA	Ras homolog gene family Member A
RISC	RNA induced silencing complex
RNA	Ribonucleic acid
RNAi	RNA interference
rRNA	Ribosomal RNA
RT	Room temperature
RT-qPCR	Real time quantitative PCR
RV	Right Ventricle
SEM	Standard error of the mean
SR	Sarcoplasmic reticulum
SRF	Serum response factor
TAC	Transverse aortic constriction
TGF- β	Transforming growth factor beta
THRAP1	T3 receptor coregulator 1
TMM	Trimmed mean of M-values
Tnnt2	Cardiac muscle troponin T
TRBP	Transactivation response element RNA-binding protein 2
tRNA	Transfer RNA
USE	Upstream element
UTR	Untranslated region
UQ	Upper quartile
WW	Wet weight
ZCCHC11	Zinc finger CCHC domain containing 11
4E-BP1	eIF4E-binding proteins

List of Figures

List of Tables

Table of Contents

Chapter One

Introduction

1.1 Cardiac Hypertrophy

1.1.1 Heart disease and Cardiac Hypertrophy

One in three Australians die of cardiovascular disease (CVD) and despite advances in treatment it remains the leading cause of death in Australia (*1*). CVD accounts for 18% of the overall burden of disease and injury within Australia, which is a measure of the years of healthy life an individual, or population, loses as a result of disease or injury (*1*). It has the greatest health expenditure compared to other diseases in Australia, costing the country $5.9 billion in 2004 and 2005. CVD covers all pathologies that involve the heart or blood vessels. In Australia, the main types of CVD are coronary heart disease, stroke and heart failure. Heart failure is a loss in the ability of the heart to adequately pump blood in response to the body's needs. It can be induced by a number of pathological conditions, including aortic valve stenosis, dilated cardiomyopathies, ischemia associated with coronary artery disease or myocardial infarction, and chronic hypertension (*2*).

Prior to heart failure, the pathological conditions listed above first induce cardiac hypertrophy, defined as an increase in the size of the cardiomyocyte without cell division (*3*). This is initially a compensatory mechanism to maintain cardiac output, but prolonged hypertrophy often leads to heart failure, arrhythmia and/or sudden death. Hypertrophic signalling results in the rapid expression of the early response genes, c-fos, c-jun and c-myc, as well as heat shock protein genes (hsp70) (*4*). This is followed by re-expression of cardiac foetal genes, such as atrial natriuretic peptide (ANP), brain natriuretic peptide (BNP), β-myosin heavy chain (β-MHC) and alpha skeletal actin (α-SkA), which are now used as molecular markers of the hypertrophic phenotype (*5*).

1.1.2 Cardiac hypertrophy signalling pathways

Cardiac hypertrophy develops from the interaction of several signal transduction pathways, and as such dozens of ligands, receptors, transducers and transcriptional regulators modulating cardiac hypertrophy have been discovered (reviewed in (*2*)). The signalling pathways identified thus far include: calcineurin/nuclear factor of T cells (NFATc), cyclic guanosine monophosphate (cGMP)/protein kinase G-1 (PKG-1), phosphoinositide 3-kinase (PI3K)/Akt, G protein-coupled receptor (GPCR), mitogen activated protein kinases

(MAPK), interleukin-6 (IL-6) cytokine, JAK-STAT and the renin angiotensin system (RAS) (Reviewed in (*2, 6*)). The main signalling pathways involved in hypertrophy are summarized in Figure 1.1. There is significant cross talk and interaction between these signalling pathways, forming an integrated network that regulates heart homeostasis. These signalling pathways end in the nucleus, where they activate numerous transcription factors that then drive the expression of stress response and foetal cardiac genes.

1.1.2. A CALCINEURIN-NFAT SIGNALLING

Calcineurin (CnA) is a calmodulin-dependent phosphatase that interacts with members of NFATc family, dephosphorylating their serine residues, which allows them to translocate to the nucleus and activate immune response genes (*7*). NFATc factors themselves bind weakly to DNA and so interact with other transcription factors such as GATA-4, c-myc and myocyte enhancing factor 2 (MEF2) to modulate transcription (*8*).

CnA-NFAT signalling is activated in response to intracellular calcium concentrations. As calcium concentrations increase, the calcium saturates calmodulin, which leads to a stronger interaction between calmodulin and CnA. This interaction generates a conformational shift and subsequent activation of CnA (*9*). Overexpression of CnA in the heart is sufficient to increase heart size and induce expression of the hypertrophic markers β-MHC, α-SkA and BNP (*10*). This suggests a role for CnA-NFAT signalling in the hypertrophic response.

Calmodulin activates calmodulin-dependent protein kinase (CaMK), resulting in the phosphorylation of class II histone deacetylases (HDACs) (*11, 12*). Phosphorylated HDACs are then exported from the nucleus to the cytoplasm, removing the inhibitory effects of HDACs from MEF2, which orchestrates the expression of stress-responsive genes (*12*). These genes include connective tissue growth factor (CTGF) and procollagens (Col1a2 and Col3a1), which promote cardiac fibrosis.

1.1.2.B G PROTEINS AND G PROTEIN-COUPLED RECEPTORS

Cardiomyocytes sense changes in hormones and growth factors through a number of receptors including GPCRs. GPCRs are seven transmembrane domain receptors, which play an important role in heart homeostasis and disease. The majority of prescriptions for CVD are drug therapies that target a subclass of GPCRs either directly or indirectly (*13, 14*). Binding of a ligand induces a conformational change in the GPCR, allowing for coupling with the small or heterotrimeric guanine-nucleotide regulatory proteins (G proteins) (Reviewed in (*15*)). G proteins exist in two states, the inactive guanosine diphosphate (GDP) bound form and the active guanosine triphosphate (GTP) bound form (*16*). The heterotrimeric G proteins consist of α, β and γ subunits and, after coupling with the GPCR, the GTP is converted to GDP on the α subunit, which dissociates the α from the βγ subunits to activate intracellular signalling pathways (*17, 18*).

Endothelin-1 (ET-1), angiotensin II (Ang II) and catecholamines (such as phenylephrine (PE)), play important roles in cardiac growth and hypertrophy and bind to GPCRs that are coupled to the $G\alpha_{q/11}$ subclass of G proteins. These G proteins activate phospholipase C (PLC) resulting in the hydrolysis of phosphatidylinositol 4,5-bisphosphate (PIP_2) into inositol 1,4,5-trisphosphate (IP_3) and diacyl glycerol (DAG) (*19*). DAG activates protein kinase C (PKC), which promotes protein synthesis and activates protein kinase D (PKD). PKD stimulates export of HDAC5 from the nucleus via phosphorylation, allowing the activation of hypertrophic genes (*20, 21*). Simultaneously, IP_3 accumulates and binds to the IP_3 receptors on the sarcoplasmic reticulum (SR), releasing calcium into the cytoplasm. This sustained calcium release increases cytosolic calcium concentration, feeding into the calcineurin-NFAT signalling pathway (see above).

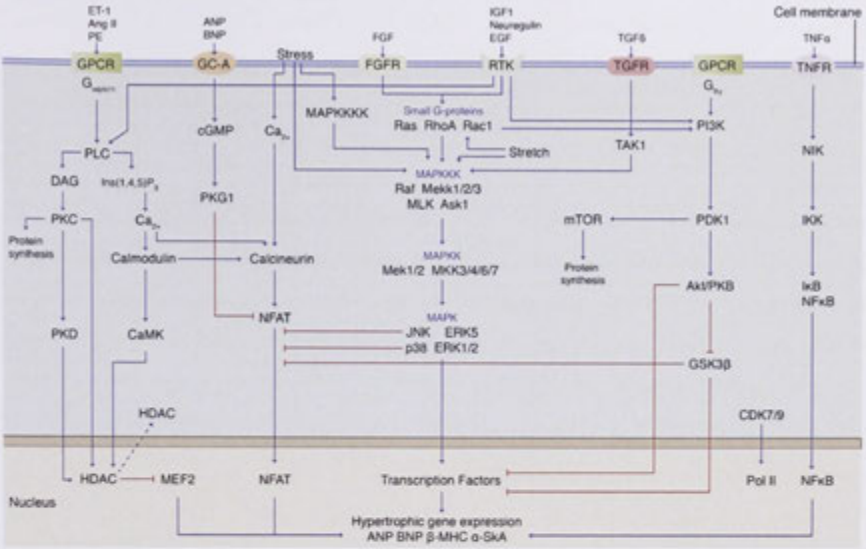


Figure 1.1. Signalling pathways controlling the hypertrophic response.

The extracellular stimulus initiating signalling and their respective receptors are shown at the top, with the cell membrane represented by the blue bar. The signalling downstream of the membrane receptors have been simplified for ease of interpretation. The Nucleus (brown bar representing the Nuclear membrane, Nucleus shown by brown background) is shown at the bottom. ANP, atrial natriuretic peptide; Ang II, angiotensin II; Ask1, apoptosis signal-regulating kinase 1; α-SkA, alpha skeletal actin; BNP, brain natriuretic peptide; β-MHC, beta myosin heavy chain; Ca_{2+}, calcium; CaMK, calmodulin-dependent kinase; CDK, cyclin-dependent kinase; cGMP, cyclic guanosine monophosphate; DAG, diacylglycerol; EGF, epidermal growth factor; ERK, extracellular signal-regulated kinase; ET-1, endothelin-1; FGF, fibroblast growth factor; FGFR, FGF receptor; GC-A, guanyl cyclase-A; GPCR, G protein-coupled receptor; GSK3β; glycogen synthase kinase-3β; HDAC, histone deacetylase; IκB, inhibitor of NFκB; IGF-1, insulin-like growth factor-1; IKK, inhibitor of NFκB kinase; Ins(1,4,5)P_3, inositol-1,4,5-triphosphate; JNK, c-jun N-terminal kinase; MAPK, mitogen-activated protein kinase; MAPKK, MAPK kinase; MAPKKK, MAPKK kinase; MAPKKKK, MAPKKK kinase; MEF, myocyte enhancing factor; MEK, mitogen-activated protein kinase kinase; MEKK, MEK kinase; MLK, mixed-lineage kinase; MKK, MAPK kinase; mTOR, mammalian target of rapamycin; NFAT, nuclear factor of activated T cells; NFκB, nuclear factor κB; NIK, NFκB-inducing kinase; PDK, phosphoinositide-dependent kinase; PI3K, phosphatidylinositol 3-kinase; PKB, protein kinase B; PKC, protein kinase C; PKD, protein kinase D; PKG, protein kinase G; PLC, phospholipase C; Pol II, RNA polymerase II; Rac1, ras-related C3 botulinum toxin substrate 1; Raf, rapidly accelerated fibrosarcoma; RhoA, Ras homolog gene family

member A; RTK, receptor tyrosine kinase; TAK, TGFβ – activated kinase; TGFβ, transforming growth factor β; TGFR, TGF receptor; TNFα, tumour necrosis factor α; TNFR, TNFα receptor. Figure modified from (*2*).

In addition to the heterotrimeric G proteins that interact with GPCRs, small G proteins also play important roles in cardiac signalling. The small G proteins are single polypeptides of low molecular weight, and are divided into 5 subfamilies: Ras (Ras, Rap and Ral), Rho (RhoA, Rac1 and CDC42), Rab, Arf and Ran (*22*). Members of both the Ras and Rho subfamilies have been described to participate in hypertrophic signalling (Reviewed in (*23*)). Ras, Rho and Rac can be activated by the $G\alpha_{q/11}$ agonists ET-1, PE and Ang II (Reviewed in (*24*)). Administration of ET-1 or PE to neonatal cardiomyocytes was sufficient to activate Ras, RhoA and Rac1 within minutes (*24*). These small G proteins can also be activated by the gradual increase in Ca^{2+} induced by IP_3 (*4*), and by stretching of the cardiomyocyte (*25*). Furthermore, binding of growth factors, such as insulin growth factor 1 (IGF-1) and fibroblast growth factor (FGF), with their respective tyrosine kinase receptors leads to Ras activation (*26, 27*). Once activated, downstream pathways regulate the cardiomyocyte cytoskeleton and amplify hypertrophic gene expression (*4*). The small G proteins stimulate the MAPK pathways (see below); Ras activates extracellular-signal-regulated kinases 1 and 2 (ERK1 and 2), c-Jun N-terminal kinases (JNKs) and p38 indirectly through Rac/CDC42. Conversely, Rac1 only stimulates the ERK cascade by promoting c-Raf activity (*28*).

Both GPCRs and small G proteins lead to modulation of transcription factor activity, which in turn regulates hypertrophic gene expression.

1.1.2.c MAPK PATHWAY

MAPK cascades are highly conserved signalling pathways that lead to the activation of the MAPKs JNK, ERK and p38 (*29*). These pathways consist of three signalling levels; MAP kinase kinase kinases (MAPKKKs) activate MAP kinase kinases (MAPKK) by phosphorylating serine and threonine residues, which in turn allows MAPKKs to activate MAPKs by phosphorylation of tyrosine and threonine residues (*4*). For example, the ERK MAPKs are activated by two MAPKKs, MEK1 and MEK2, which are activated by the MAPKKKs Raf-1, A-Raf, B-Raf, and MEKK1-3 (*29*).

The ERK pathway is activated by binding of hormones (such as ET-1 and Ang II) to GPCRs, binding of growth factors to receptor tyrosine kinases (RTKs, for example IGF-1)(*2*), and by mechanical stress (reviewed in (*30*)). The ERK family of MAPKs has been associated with cell growth and survival, while the p38 and JNK pathways are known to induce hypertrophic gene expression and increase protein synthesis (*31, 32*). Both JNK and p38 are also activated by myocyte stretch and activation of GPCRs (*33-36*). Once activated, the MAPKs phosphorylate and activate numerous transcription factors, including c-jun, Elk-1, activating transcription factor-2 (ATF-2), MEF2, NFAT, GATA4 and NF-κB (*35*). Activation of these transcription factors collectively induces the reprogramming of cardiomyocyte gene expression during hypertrophy. MAPKs also phosphorylate and activate other protein kinases, amplifying the hypertrophic signal.

1.1.2.D Protein Kinase C

The PKC family consists of at least 12 serine/threonine kinases (*37*). Conventional PKCs are Ca^{2+} dependent and are activated by both phosphatidylserine and DAG (produced by PLC, described in GPCRs) (*37*). PKC activates PKD by phosphorylating two serine residues in the activation loop of PKD (*38*). PKD phosphorylates class II HDACs, which normally represses pathological cardiac gene expression by interacting with the transcription factor MEF2 and its partner transcription factors such as NFAT and GATA4 (*20*). Once phosphorylated, HDAC5 binds to 14-3-3 proteins, which result in export of HDAC5 from the nucleus, allowing the expression of hypertrophic genes. It is possible that activated PKCs can also directly phosphorylate HDAC5, bypassing PKD and leading to the activation of hypertrophic genes (*20*).

1.1.2.E The PI3K-AKT Pathway

The PI3K-Akt pathway occurs downstream of GPCRs and receptor tyrosine kinases (Figure 1.1) (*4*). Activated PI3K converts the plasma membrane lipid phosphatidylinositol-4,5-bisphosphate [$PI(4,5)P_2$] to phosphatidylinositol-3,4,5-triphosphate [$PI(3,4,5)P_3$] (*39*). Proteins with pleckstrin-homology (PH) domains can now accumulate at the cell membrane by binding to $PI(3,4,5)P_3$. Association at the membrane brings phosphoinositide-dependent kinase-1 (PDK-1) into close proximity to Akt/PKB, allowing PDK-1 to phosphorylate and activate Akt/PKB (*40*). Now catalytically active, Akt phosphorylates a number of proteins generally inhibiting their action. For example, Akt

phosphorylates Bad (an apoptosis-inducer), which prevents Bad from binding to B-cel lymphoma 2 (Bcl-2) family members and in turn, releases them to orchestrate a cell survival response (*41*). Akt also phosphorylates glycogen synthase kinase 3- β (GSK3β). GSK3β is normally active and inhibits transcription factors such as c-Myc, GATA4 and NFAT (*39*). GSK3β also negatively regulates the eukaryotic initiation factor 2B (eIF2B) (*39, 42*), and the structural protein β-catenin that interacts with transcription factors to induce gene expression (*43*). Phosphorylation of GSK3β by Akt represses its catalytic activity, allowing the activation of the downstream proteins normally repressed by GSK3β.

The Akt/PKB pathway can also lead to increased protein synthesis via activation of mammalian target of rapamycin (mTOR) (*39, 42*). mTOR regulates multiple proteins involved in the regulation of mRNA translation, including the translational repressors eIF4E-binding proteins (4E-BP1/2) and eukaryotic elongation factor 2 (eEF2) kinase that inactivates eEF2 (*44*). The net result of mTOR activation is stimulation of protein synthesis and an increase in the components required for protein synthesis (*44*).

1.1.2.F SIGNALLING PATHWAY SUMMARY

Hypertrophy pathways can be induced by various stimuli and involve numerous cell-surface receptors, signalling pathways and transcription factors. However, the interactions that feed into and control these signalling pathways are still not completely understood. It is now evident that miRNAs are regulated during hypertrophy, and can interact with mRNAs that are interlinked with hypertrophic signalling pathways (*45*).

1.2 MIRNAS

1.2.1 CANONICAL BIOGENESIS OF MIRNAS

The first step to produce a mature miRNA is transcription of the primary miRNA (pri-miRNA) from the Genome (Figure 1.2). Akin to protein-coding genes, pri-miRNAs are transcribed primarily by RNA polymerase II and are capped and polyadenylated (*46, 47*). In the human genome ~50% of miRNAs are intergenic with their own promoter elements, while the other 50% are co-transcribed from the introns of protein-coding genes (and thus under the same transcriptional regulation as the host gene). Pri-miRNAs are larger transcripts that can contain one or several miRNA sequences already folded into hairpin

structures, for example the cardiac specific miRNAs, miR-133a-1 and miR-1-2, are co-transcribed to produce one bicistronic primary transcript (*48*). Maturation of miRNAs is mediated by the two RNase III endonucleases Drosha and Dicer (*49*). While still in the nucleus, the pri-miRNA is cropped by the "microprocessor" complex, consisting of Drosha and its cofactor DGCR8/Pasha (DiGeorge syndrome critical region gene 8), which assists Drosha to bind to its substrates (*46, 50*). Cleavage by Drosha ~11 nucleotides (nt) from the base of the hairpin yields a 60-70 nt long stem-loop with a 2 nt 3' overhang termed the precursor miRNA (pre-miRNA) (*51*). The pre-miRNA is then transported to the cytoplasm via an exportin-5-dependent mechanism (*52*). Once in the cytoplasm, the 2 nt 3' overhang is recognized and the pre-miRNA is then cleaved by Dicer ~22 nt from the Drosha cleavage site, with or without cofactors such as transactivation response element RNA-binding protein 2 (TRBP2) and Argonaute proteins (AGOs) (*53, 54*). The recognition of the hairpin and cleavage site choice appear to rely on certain structural features of the precursor (*55*). Similar observations were made in *C.elegans*, suggesting the structural preferences of Drosha and Dicer are conserved (*56, 57*). The cellular levels of pri-, pre- and mature miRNAs are often not well correlated and the processing steps are regulated through mechanisms that remain largely unknown.

The sequential cleavage of Drosha and Dicer yields a miRNA duplex, from which usually one strand is selectively loaded into an AGO protein in the RNA Induced Silencing Complex (RISC). The mammalian AGOs have a strong preference for the first base of the loaded miRNA to begin with a uridine (*56, 58*) and a 5' uridine bias is also observed in the cardiomyocyte miRNA population (*55*). Furthermore, there is a bias for the first base-pairing to be thermodynamically unstable (*59, 60*), which may help in unwinding of the duplex by the wedge-shaped domain of AGO (*61*). However, as there are numerous examples that run counter to these preferences (see below), there are likely to be additional regulatory factors at play. Furthermore, unlike in Drosophila (and plants), very few mammalian miRNAs are preferentially associated with a particular AGO (*62*), leaving the specific roles of the four mammalian AGOs in miRNA-mediated repression a continuing area of debate (the function of AGOs is reviewed in (*63*)).

1.2.2 Principles of miRNA:mRNA interactions

Once processed, one strand of the miRNA duplex (normally the strand with lower stability of base-pairing at the 5' end) is transferred onto an AGO protein to form an active RISC (*53, 59*). The AGO family proteins are the critical and best-characterised components of RISCs. Mammals contain four AGO proteins, all of which function in miRNA repression, however only AGO2 shows slicer activity and catalyses the miRNA-induced endonucleolytic cleavage of the target (*64*). This RNA interference (RNAi)-like cleavage is common in plants as miRNAs generally base pair to mRNAs with nearly perfect complementarity in this system. However, in mammals this pathway is rarely used and in most cases miRNAs bind imperfectly to their targets resulting in translational repression or exonucleolytic decay of the mRNA transcript (Figure 1.2)(reviewed in (*65, 66*)).

Early experiments identified Watson-Crick base pairing of the seed region, which comprises nucleotides 2-7 of the miRNA, with the mRNA 3' UTR as an important determinant of miRNA-mediated repression of target mRNA (*67*). *In vitro* reporter assays indicate perfect complementarity to the miRNA seed alone is in some cases sufficient to induce repression (*68*). On the other hand, more recent work shows that seed pairing is neither necessary nor sufficient for all miRNA:mRNA interactions and relying on seed-based pairing alone results in large false positive and negative rates (*69*). Additionally, it has also been suggested that the importance of seed matching for the binding of miRNAs to their targets may have been overestimated in miRNA overexpression studies (*70*). Our understanding of the miRNA:mRNA interaction has altered in response to a move from reporter-based experiments, to measurements of genome-wide endogenous miRNA:mRNA interactions using cross-linking techniques such as HITS-CLIP (high-throughput sequencing of RNAs isolated by cross-linked immunoprecipitation (IP)) and PAR-CLIP (photoactivatable ribonucleoside-enhanced cross-linking and IP) on RISC complexes containing miRNA and mRNA (*71, 72*). Contemporary bioinformatic predictions of targets therefore take into account the non-seed based binding of miRNAs to their targets as well as evidence from AGO-mRNA association studies in relevant cell types (*73-75*).

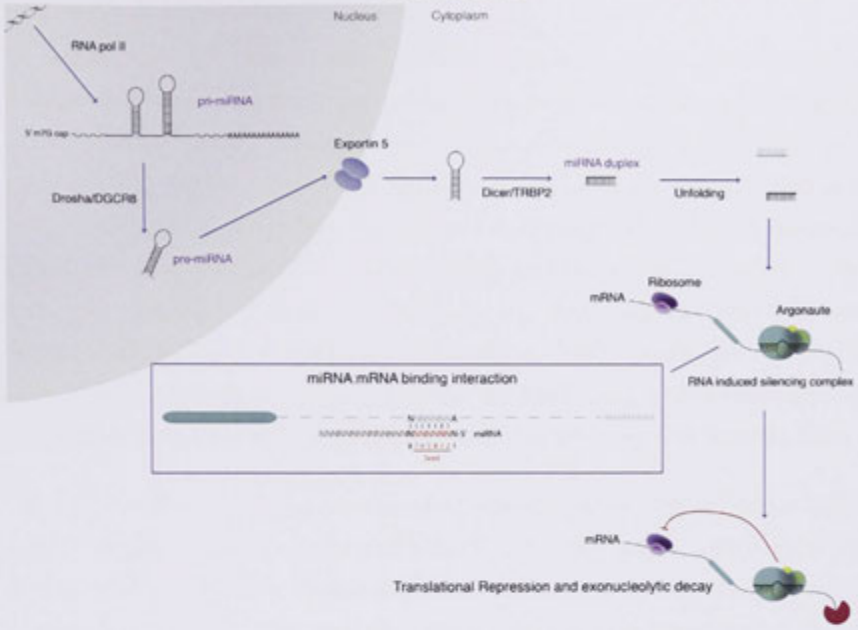


Figure 1.2. miRNA biogenesis and targeting in mammalian cells.

Primary miRNAs (pri-miRNAs) are transcribed from DNA in the nucleus by RNA pol II. They are then processed by the RNase III enzyme Drosha to form precursor miRNA (pre-miRNAs) that are exported to the cytoplasm with the assistance of Exportin-5. A second RNase III enzyme, Dicer, then cleaves the pre-miRNA to produce a miRNA duplex. The miRNA duplex is unwound and it was previously thought that only one miRNA strand, the mature miRNA, is incorporated into the RNA-induced silencing complex (RISC) where it associates with a member of the Argonaute (AGO) family and gene silencing takes place. In animals, the miRNA binds with imperfect complementarity, predominantly by the seed regions (shown within the miRNA:mRNA binding interaction box) and this leads to translational repression and/or exonucleolytic decay.

Of importance for future cardiac miRNA research, a study performed AGO2 immunoprecipitations and RNA-sequencing on murine ventricular tissue to identify the subset of cardiac mRNAs enriched in the RISC complex (*76*). This approach defined the mRNAs being targeted by all the miRNAs in ventricular tissue and was used as the background signal for subsequent experiments. To discover miRNA-specific targets, the investigators then overexpressed specific miRNAs, for example miR-133a or miR-499, to "reprogram" the RISC complex to favor the overexpressed miRNAs. Overexpression of specific miRNAs and AGO2 immunoprecipitation was highly specific and several of the mRNA targets immunoprecipitated were not predicted using a seed-based algorithm (Targetscan), supporting the role of alternative miRNA:mRNA interactions in the heart.

The miRNA:mRNA interaction is nevertheless sequence-specific, thus any alteration to the miRNA sequence or the sequence within the 3' UTR could reduce or abrogate miRNA repression of a target. It is evident from these studies that our understanding of how miRNAs interact with their targets is evolving, as is our understanding of the complexity of the miRNA population, leading to more sophisticated and (hopefully) accurate target prediction methods.

1.2.3 miRNA processing variants

It was originally thought that one pre-miRNA hairpin gives rise to a single dominant mature miRNA, however, it has recently become apparent that processing of one hairpin can result in several miRNA variants, potentially with altered mRNA target specificity. This can arise from use of both strands of the miRNA:miRNA star (miRNA*) duplex with similar frequency, the occurrence of 5' and 3' isomiRs as well as RNA editing (Figure 1.3A). It is postulated that these miRNA modifications could alter miRNA biogenesis and half-life, subcellular localization and/or targeting specificity.

Initial miRNA studies suggested that one strand of the pre-miRNA hairpin (known as the mature strand) was functional and incorporated into RISC while the other strand (known as the star or passenger strand) was non-functional and degraded. Strand selection is in part determined by the strength of the base pairing at the ends of the duplex (*51*). It is now evident that both the mature and star strands of the miRNA hairpin can be functional, with many miRNA* variants detected in RISC complexes suggesting that they can repress mRNAs *in vivo (62)*. Furthermore, their function *in vivo* has been highlighted by their role in human

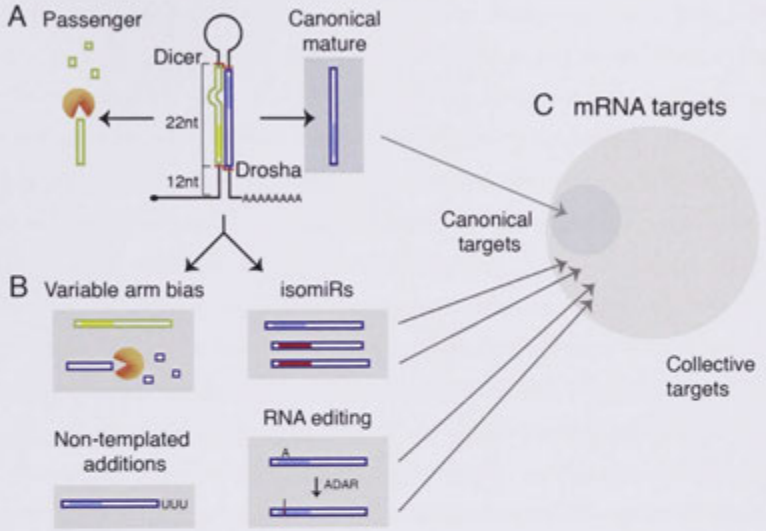


Figure 1.3. miRNA processing and variability and consequences for mRNA targeting.

A) The canonical view on miRNA hairpin precursor processing, leading to accumulation of one mature miRNA. (B) Next generation sequencing reveals that variable processing, editing, and non-templated additions can lead to the expression of miRNA sequence variants that (C) expand the mRNA targeting repertoire of a given miRNA locus. mRNA targets of canonical mature miRNA: small circle; collective mRNA target range of all expressed miRNA variants: large circle; miRNA seed regions: filled boxes. Image from (*77*).

disease (*52, 78*). For instance, miR-199a* (now called miR-199a-3p) is decreased in human hepatocellular carcinoma, where it normally functions to suppress the MEK/ERK pathway and cell growth (*78*).

Interestingly, the preferred strand can switch between tissue-types and cellular contexts, giving rise to the term 'arm switching' (*79, 80*). This suggests that additional regulatory mechanisms may determine which strand is functionally active in a certain cell type or state (*81*). In agreement with this notion, version 18 of miRBase (database of published miRNA sequences and annotation, www.mirbase.org) removed the mature miRNA/miRNA* nomenclature and reannotated all miRNAs based on the pre-miRNA arm they were derived from, with addition of the -5p or -3p suffix (*82*). The concept that miRNA* are functional greatly impacts the design and interpretation of cardiac miRNA studies, especially when the expression of the entire miRNA precursor is altered, potentially confounding the interpretation of phenotype and mRNA targets. Furthermore, prior to the advent of next generation sequencing (NGS), the expression and importance of many miRNA* were under appreciated in cardiac-related studies, as they were not included on microarray platforms. Thus our understanding of the role of miRNAs in cardiac biology is lacking in regards to miRNA*.

The availability of NGS has added further complexity to our definition of a functional miRNA. NGS studies have identified length and sequence variants of known miRNAs giving rise to the term isomiRs (*79, 83, 84*). IsomiRs have been detected in Argonaute complexes and in translating ribosomes, implying that they are functional miRNAs (*62, 85, 86*). IsomiRs with alternative 5' start sites (5' isomiRs) are generated by differential Drosha and Dicer cleavage, which results in shifted seed sequences and, as the seed sequence is commonly a major determinant of miRNA targeting, it can be postulated that isomiRs target different sets of mRNAs. There are several recent studies that support this notion. Firstly, microarray transcriptome analysis after deletion of miR-223 (which has two major 5' isomiRs in neutrophils) revealed substantial re-expression of mRNAs matching the seed region of both the canonical miRNA and its co-expressed 5' isomiR (*79*). This was not due to an overlap in binding sites, suggesting that 5' isomiRs can direct the repression of distinct sets of targets. A study by Cloonan et al. used biotin labeled miRNAs or isomiRs to pull-down endogenous targets and found that 5' isomiRs bind both common and unique targets (*85*). These experiments demonstrate the potential for isomiRs to act individually or

in concert to regulate gene expression, however there is much conjecture as to what is the biological role of isomiR expression. One hypothesis states that regulated changes to the 5' start site essentially generates a new miRNA with a different target spectrum, increasing the complexity of the miRNA pool (*85*). Another hypothesis is that co-expression of isomiRs leads to reinforced repression of core mRNA targets, while diluting off-targeting events between several isomiR species (*85*). Another possibility is that isomiRs are differentially located and regulated, and may associate with diverse effector proteins to create distinct outcomes (e.g. translational repression versus mRNA decay). The consequence of 5' isomiRs is more obvious in plants, where the first 5' nt of a miRNA is the major determinant for AGO protein association (*87*). For instance, addition of a U to the 5' end of Arabidopsis miR-156 leads to association with AGO5 instead of AGO1 (which is the major AGO for miRNA post-transcriptional gene silencing in plants) (*86*). The analysis of isomiR expression and function will be an important aspect of future cardiac miRNA analysis.

Even more enigmatic is the function of the extremely common 3' isomiRs. Sequencing data suggests that miRNA 3' ends are more heterogeneous compared to the 5' end (*77, 88, 89*). The mechanisms of 3' end modification of miRNAs are not fully understood and appear to be varied, including 3' trimming, non-templated addition and alteration to Drosha and Dicer cleavage sites (*62, 90, 91*). However, it is likely that these modifications are biologically important, as there is evidence that some modifications alter miRNA turnover (*92*), can alter preferential loading into one of 4 mammalian Argonautes (*62*) and are developmentally regulated (*83*). Although 3' isomiRs share a common seed sequence, 3' end changes can impact target selection or efficiency of repression in some cases. For example, terminal uridylation of miR-26 (which has high expression in the heart) in immune cells abrogates repression of interleukin (*93*). MiRNA 3' isomiRs need to be taken into account when designing primers for qPCR detection, as well as mimics and inhibitors for experimental and therapeutic manipulations.

MiRNA variation can additionally arise from RNA editing of internal nucleotides. Most RNA editing in human cells occurs by Adenosine Deaminases Acting on RNA (ADARs) that deaminate adenosine to create inosine (A-to-I), which then base pairs with cytosine (*94*). This editing can alter miRNA precursors and regulate miRNA biogenesis, as well as change the sequence of the mature miRNA (*95*). The biological implications of RNA

editing are highlighted by miR-376-a1-5p, 41% of which undergoes editing in the human medulla oblongata (*96*). The A-to-I edited site occurs within the seed sequence, changing the binding specificity of the miRNA. This alteration allows miR-376-a1-5p to target phosphoribosyl pyrophosphate synthetase 1 (PRPS1), an essential housekeeping enzyme in the uric acid synthesis cycle, to tightly regulate uric acid levels in select tissues such as the brain cortex (*96*). This form of RNA editing is most common in neuronal cells and there is little evidence for this modification in cardiomyocytes. However, as a splice variant of ADAR2 is highly expressed in the heart (*97*), there may be cardiac environments where this modification is important to miRNA action and regulation.

Thus, the initial concept that miRNA processing leads to a single mature miRNA that targets a canonical set of mRNAs has to be modified. It is necessary to accommodate the potential for a broader collective targeting range of all the miRNA processing variants derived from a given miRNA locus (Figure 1.3).

1.2.4 Regulation of miRNA biogenesis

Transcriptional regulation of miRNAs generally occurs via standard promoter element-determined mechanisms, whether via their own promoter or the promoter of the gene they are embedded in (*98*). Additionally, the epigenetic mechanisms that transcriptionally regulate mRNAs can also contribute to silencing of miRNA genes. For example, all three loci of the tumour suppressor miR-9 are hypermethylated in gastric cancer, which leads to repression of miR-9, and subsequently proliferation of gastric cancer cells (*99*).

It is often observed that the levels of pri-miRNA and mature miRNAs are poorly correlated (*100*), demonstrating the central importance of posttranscriptional regulation to the genesis of a specific miRNA population. And it appears that the regulation of Drosha cleavage is a central step. While DGR8/Drosha are the minimal components of the "microprocessor' required for pri-miRNA cleavage, *in vivo* Drosha resides in multiple protein complexes (which may commonly contain the p68 and p72 helicases) and associates with multiple regulatory proteins. For example, ERα and p53 interact with the microprocessor (probably through p68 and p72) to inhibit and stimulate miRNA biogenesis, respectively (*101, 102*). Interestingly, the biogenesis of miR-21 in smooth muscle is promoted by transforming growth factor beta (TGF-β) and bone morphogenetic protein (BMP) signalling, via interaction of SMAD1 with p68 and direct interaction of R-

SMADs with the pri-miR-21 in a ligand and sequence-dependent manner (*103*). It remains to be seen if this regulation is also important for the central role of miR-21 in the heart (*104*).

There is a plethora of other RNA binding proteins that directly bind to miRNA hairpins and regulate their processing. Of importance to cardiac function is the regulation of miR-1 expression by LIN28 and musclebind-like splicing regulator 1 (MBNL1), which both compete for binding to the loop of pre-miR-1. Binding of LIN28 to pre-miR-1 stimulates Zinc finger CCHC domain containing 11 (ZCCHC11)-mediated uridylation of the precursor and blocks Dicer processing, whereas binding of MBNL1 is permissive to Dicer processing (*105*). This regulation is perturbed in myotonic dystrophy, where overproduction of a CCUG repeats sequesters MBNL1 and leads to downregulation of miR-1 expression and multiple cardiac phenotypes. Similar regulation of miRNA biogenesis is observed during the developmental fluctuations in let-7 expression, which result from competition between LIN28A (which again stimulates uridylation of the pre-miR) and KH-type splicing regulator protein (KSPR) for binding to the loop of pre-let-7 (*106*). Interestingly LIN28B also represses let-7 processing, but this time in the nucleus through inhibition of microprocessor processing, highlighting the multiple levels of regulation microRNAs can be subjected to (*107*).

Another example of hairpin binding provides insight into the differential expression of clustered miRNAs, which are co-transcribed but commonly differentially expressed. The highly abundant hnRNP A1 protein is essential for pri-miR-18a processing but not required for the expression of other members of the miR-17-92 cluster. In this case, hnRNP A1 binds to the loop of the hairpin to relax the stem region and facilitate Drosha cleavage (*108*). Decreased expression of this cluster, including miR-18a, in cardiomyocytes has been linked to age-related heart failure (*109*).

Importantly, the hairpin binding proteins described above have all been demonstrated to regulate and/or bind other miRNA precursors and there are likely to be further proteins discovered with similar functions. It is likely there are many more examples of signalling-dependent miRNA processing.

1.2.5 Core cardiac miRNAs: their function and their variation

1.2.5.a Abundant and specific cardiac miRNAs

In most, if not all, cell/tissue types studied to date a small number of miRNAs are highly expressed, potentially enforcing tissue identity (*110*). In the adult mouse heart the top 18 highly expressed miRNAs account for >90% of the miRNA population (*111*) and include members of the let-7, miR-22, miR-208, miR-133 and miR-30 families. This abundance can result from individual regulation or regulation of the "host" gene. For example, the two miR-1/miR-133a clusters are highly expressed in both skeletal and cardiac muscle, whereas the miR-206/miR-133b cluster is only highly expressed in skeletal muscle (*112, 113*). This can be attributed to their independent regulation by key myogenic transcription factors, such as serum response factor (SRF) and MEF2 (*48, 112*). Alternatively, another heart-specific miRNA, miR-208a, is encoded within an intron of α-myosin heavy chain (α-MHC) gene Myh6 (*114*). The β-MHC gene (Myh7) encodes another miR-208 family member, miR-208b (*115*). Unlike the miR-1/miR-133 clusters that contain their own promoters, the tissue specificity of miR-208a and miR-208b is controlled by the expression of α-MHC and β-MHC, respectively (*115*), with α-MHC and miR-208a highly expressed in murine adult cardiomyocytes. Interestingly, miR-133a exists as two isomiRs, which are produced from both of the miR-133a clusters (*79*). The individual roles of these miR-133a variants in cardiac function and disease have yet to be addressed.

1.2.5.b Expression patterns of miRNAs in cardiac disease

The expression patterns of miRNAs in various cardiac hypertrophy-associated diseases in both human patients and murine models have been extensively examined (reviewed in (*116-118*)), for the most part using microarray analysis of canonical, miRbase-defined miRNA species and show that the majority of highly expressed cardiac miRNAs are differentially regulated in cardiac disease (*119*).

The first profile of cardiac miRNAs was described by Van Rooij et al. in 2006, who measured miRNA expression in two models of cardiac hypertrophy; transverse aortic constriction (TAC) and transgenic mice overexpressing CnA (*120*). Collectively, 60

miRNAs were upregulated and 20 were downregulated, with 21 and 7 up or down in both models, respectively. Both, anti-apoptotic miRNAs, such as miR-21, and miRNAs critical for cell proliferation, such as miR-23, were upregulated during hypertrophy and it was hypothesised that they play an important role in hypertrophic cell enlargement. Furthermore, overexpression of miRNAs downregulated in hypertrophy, including miR-133a, caused a reduction in cell size suggesting these miRNAs may function to suppress cell growth (*120*). The expression changes of miRNAs identified by subsequent murine studies are highlighted in Table 1.1. Importantly, different hypertrophic stimuli, such a phenylephrine (PE) and FGF, can activate distinct molecular pathways. In this case, PE triggers an increase in intracellular calcium, which stimulates hypertrophic gene transcription via NFATc, while FGF acts through MAPK signalling (*4*). Given that multiple pathways can initiate and transduce hypertrophy, it is not surprising that miRNA profiles are not always consistent between different models of hypertrophy. For example, miR-142-3p is upregulated in hearts after TAC but unaltered in cardiomyocytes treated with PE (*121*).

The expression of many miRNAs alters significantly in other human heart diseases, including ischemic cardiomyopathy (ICM) and dilated cardiomyopathy (DCM), and it is thought these miRNA changes contribute to the progression of heart disease and subsequent heart failure (*120, 122-124*). Cardiomyopathies are heart muscle diseases that result in a deterioration of heart function, which usually leads to heart failure. DCM is characterized by dilatation of the cardiac chambers and impaired contractile function of one or both ventricles (*125*). It is a leading cause of hospitalization and death in Western countries and can result from several factors, such as pressure or volume overload, inherited mutations and ischemia (*125*). Alternatively, ICM, which accounts for >50% of heart failure cases, develops when coronary heart disease or heart attack leads to myocardial ischemia and/or infarction (*126*). It is hypothesised that, despite the underlying

Table 1.1. miRNAs differentially expressed in murine models of hypertrophy. (A) TAC and constitutively activated calcineurin (CnA) mice *(120)*. (B) TAC at 24 hours, 7 days and 14 days *(127)*. (C) Phenylephrine (PE) treated cardiomyocytes and 14 days post TAC (*121*). (D) TAC at 7 days, 14 days and 21 days (*128*). (E) Early and end-stage hypertrophic cardiomyopathy (HCM) using a transgenic mouse *(129)*. (F) TAC at 7 days *(130)*. (G) TAC at 7 days *(131)*. # confirmed by Northern blot, ^ confirmed with RT-qPCR, X could not be confirmed.

miRNA	TAC & CnA [A]	TAC [B]			PE CM [C]	TAC [C]	TAC [D]			Transgenic HCM mice [E]				TAC [F]	TAC [G]
		Day 1-14	Day 7-14	Day 14		Day 14	Day 7-21	Day 14-21	Day 21	5 do	10 do	14 do	16 do	Day 7	Day 7
let-7b-5p				↑											
let-7c-5p				↑											
let-7d-3p				↓											
miR-1-3p		↓#								↓#	↓^	↓#			
miR-10a-5p				↓											
miR-10b-5p	↑			↓							↑^				
miR-15b-5p				↑										↑	
miR-17-5p						↑								↑	
miR-18b-5p					↑	↑									
miR-19a-3p	↑														
miR-19b-3p				↑											
miR-20b-5p					↑	↑									
miR-21-5p	↑#		↑#		↑#	↑#	↑#^				↑^	↑^	↑^	↑	↑
miR-21-3p															↑
miR-23a-3p	↑#			↑	↑#										
miR-23b-3p	↑#		↑											↑	
miR-24-3p	↑#			↑											
miR-25-3p	↑														
miR-26a-5p				↓											
miR-26b-5p				↓											
miR-27a-3p	↑		↑				↑#								
miR-27b-3p	↑		↑				↑#								
miR-27b-5p															↑
miR-29a-3p				↓			↓#								
miR-29b-3p							↓^								
miR-29c-3p	↓#			↓			↓#								
miR-30a-5p			↓												
miR-30a-3p				↓											
miR-30b-5p			↓			↓						↓^			
miR-30c-5p			↓			↓						↓^			
miR-30d-5p			↓												
miR-30e-5p	↓			↓			↓#					↓^			
miR-30e-3p				↓											
miR-31-5p			↑								↑^	↑^	↑^		
miR-34a-5p														↑	
miR-34b-5p														↑	
miR-34b-3p												↑^	↑^	↑	
miR-34c-5p														↑	↑
miR-93-5p	↓#														
miR-103-3p				↑											
miR-106-5p					↑	↑									
miR-107-3p				↑											
miR-125b-5p	↑#		↑#		↑#	↑#									
miR-126-5p	↑						↓^		↑^						
miR-127-3p			↑											↑	
miR-130b-3p														↑	
miR-132-3p											↑^	↑^	↑^		
miR-133a-3p	↓				↑		↓^			↓		↓^			
miR-133b-3p	↓						↓^					↓^			
miR-134-5p														↑	
miR-139-5p				↓											
miR-140-3p				↑		↑								↑	
miR-140-5p				↑											
miR-142-3p						↑						↓^	↑^	↑	
miR-146-5p							↓	↑^						↑	↑
miR-149-5p			↓				↓^								
miR-150-5p	↓#		↓			↓	↓#^					↓^			
miR-151-3p				↓											
miR-153-3p						↑									
miR-154-5p	↑													↑	
miR-155-5p				↓											
miR-181b-5p	↓#														
miR-184b-3p						↑								↑	↑
miR-185-5p			↓				↓#^								

Table 1.1. cont. miRNAs differentially expressed in murine models of hypertrophy. (A) TAC and constitutively activated calcineurin (CnA) mice (*120*). (B) TAC at 24 hours, 7 days and 14 days (*127*). (C) Phenylephrine (PE) treated cardiomyocytes and 14 days post TAC (*121*). (D) TAC at 7 days, 14 days and 21 days (*128*). (E) Early and end-stage hypertrophic cardiomyopathy (HCM) using a transgenic mouse (*129*). (F) TAC at 7 days (*130*). (G) TAC at 7 days (*131*). [#] confirmed by Northern blot, [*] confirmed with RT-qPCR, [X] could not be confirmed.

miRNA	TAC [A] & CnA	TAC [B] Day 1-14	TAC [B] Day 7-14	TAC [B] Day 14	PE [C] CM	TAC [D] Day 14	TAC [D] Day 7-21	TAC [D] Day 14-21	TAC [D] Day 21	Transgenic HCM mice [F] 5 do	Transgenic HCM mice [F] 10 do	Transgenic HCM mice [F] 14 do	Transgenic HCM mice [F] 16 do	TAC [G] Day 7	TAC [H] Day 7
miR-187-3p					↓										
miR-191-5p														↑	
miR-193b-3p														↑	
miR-194-5p				↓											
miR-195-5p	↑#			↑											
miR-199a-5p	↑#		↑											↑	↑
miR-199a-3p	↑#		↑												
miR-199b-5p															↑
miR-199b-3p			↑												
miR-200a-3p						↑									
miR-200b-3p														↑	
miR-208-3p						↑								↑	
miR-210-3p						↑					↑*				
miR-211-5p						↑									
miR-212-3p														↑	↑
miR-212-5p															↑
miR-214-3p	↑#		↑				↑#				↑X	↑*	↑X	↑	↑
miR-214-5p															↑
miR-217-5p	↑#														
miR-218-5p	↑			↓											
miR-221-3p			↑			↑									↑
miR-221-5p															↑
miR-222-3p			↑			↑						↑*	↑X	↑	
miR-223-3p														↑	↑
miR-292-5p					↓										
miR-298-5p														↑	
miR-299-3p														↑	
miR-300-3p														↑	↑
miR-322-5p							↓	↑#							↑
miR-324-3p														↑	
miR-326-3p														↑	
miR-330-5p	↑														↑
miR-335-3p														↑	
miR-337-5p														↑	
miR-341-3p							↑#								
miR-351-5p	↑		↑											↑	↑
miR-378-3p				↓											
miR-379-5p														↑	↑
miR-381-3p															↑
miR-382-5p														↑	
miR-409-3p														↑	
miR-410-3p														↑	↑
miR-411-5p														↑	↑
miR-423-3p															↑
miR-431-5p														↑	
miR-433-5p														↑	
miR-434-5p														↑	↑
miR-434-3p															↑
miR-451							↓#								
miR-455-3p														↑	
miR-466-3p					↓										
miR-484															↑
miR-486-5p							↓#					↓*			
miR-497-5p														↑	
miR-532-3p															↑
miR-541-5p														↑	↑
miR-542-3p											↓*				
miR-574-3p														↑	↑
miR-671-5p														↑	
miR-671-3p															↑
miR-673-5p														↑	
miR-674-5p														↑	
miR-1193-3p														↑	
miR-3096-5p														↑	

etiology that initiates heart failure, there is a final common pathway in advanced or end-stage heart failure (*132, 133*). These similarities are reflected in the transcriptome and proteome, with little detectable differences between ICM and DCM (*132*). It is of no surprise then, that numerous miRNAs are also consistently up or downregulated in the progression to heart failure, for example miR-195 increases and miR-222 decreases in end-stage ICM and DCM (expression changes in human heart disease are shown in Table 1.2). MiR-195 is also upregulated early in murine hypertrophic hearts and cardiac specific overexpression of miR-195 induced adverse cardiac remodeling which led to heart failure (*120*). While miR-195 regulates numerous cell cycle genes during development (*134*), it has been postulated that the role of miR-195 in heart failure is due to the downregulation of multiple pro-survival proteins (*135*).

Nevertheless, while there are commonalities, there are distinct subsets of miRNAs expressed for each heart condition. For example, miR-19 is downregulated in DCM but unaltered in ICM patients, suggesting some miRNA changes do not simply reflect end-stage heart disease but, instead, disease-specific gene expression changes (*124*). Distinguishing between ICM and DCM patients is currently difficult, with incorrect classification of ICM and DCM heart failure estimated to occur in 13% of heart failure patients (*136*). Encouragingly, the miRNA profiles are sufficiently dissimilar between disease classes that clinical diagnosis of DCM, ICM and aortic stenosis (AS) could be predicted in 69% of cases (*124*). Although promising, miRNA expression profiles are of limited clinical utility as biopsy of the heart is required. However, measurement of cardiac miRNAs in the plasma may prove informative. For instance, miR-1 and miR-208a are both elevated in the blood of patients following myocardial infarction (*137, 138*) and circulating levels of miR-423-5p increase specifically in heart failure patients (*139*). One definite outcome of these profiling studies is the identification of cardiac disease-related miRNAs that warrant functional interrogation in the heart.

Table 1.2. miRNAs differentially expressed in human heart disease. (A) Dilated cardiomyopathy (DCM) and ischemic cardiomyopathy (ICM) patients *(123)*. (B) DCM, ICM and patients with aortic stenosis (AS) *(124)*. (C) End stage heart failure (HF) patients *(122)*. # confirmed by Northern blot, ^ confirmed with RT-qPCR, X could not be confirmed.

miRNA	DCM[A]	ICM[A]	DCM[B]	ICM[B]	AS[B]	End stage HF[C]
let-7a-5p		↓				
let-7b-5p			↑	↑		↑
let-7c-5p		↓	↑	↑	↑	↑ ^
let-7d-5p						↑
let-7d-3p		↓				
let-7e-5p					↑	↑
let-7f-5p		↓				
miR-1-3p		↓	↓		↓	↑ ^
miR-10a-5p					↓	
miR-10b-5p	↓					
miR-15b-5p			↑		↑	
miR-17-5p			↓			
miR-19a-3p			↓		↓	
miR-19b-3p			↓		↓	
miR-20a-5p	↓		↓		↓	
miR-20b-5p			↓		↓	
miR-21-5p						↑ ^
miR-22-3p	↓					
miR-23a-3p	↑		↑	↑	↑	
miR-23b-3p					↑	
miR-24-3p				↑	↑	
miR-26b-5p		↑			↓	
miR-27a-3p				↑	↑	
miR-27b-3p					↑	
miR-28-5p		↑	↓			
miR-29a-3p						↑ ^
miR-30a-5p						↓
miR-30c-5p	↓					
miR-30e-5p			↓		↓	
miR-32-5p						↑
miR-34b-3p						↑
miR-92-3p	↓ ^	↓ ^				
miR-93-5p					↑	
miR-99b-5p			↑		↑	
miR-100-5p	↑ ^	↑ #	↑	↑	↑	
miR-101-3p			↓		↓	
miR-103-3p			↑	↑	↑	
miR-106-5p			↓			
miR-125a-5p						↑
miR-125b-5p	↑		↑		↑	
miR-126-5p					↓	
miR-126-3p			↓	↓	↓	
miR-130a-3p						↑
miR-132-3p						↑
miR-133a-3p	↓ X	↓ X				
miR-133b-3p	↓ ^	↓ ^				
miR-139-5p	↓					
miR-140-5p			↑	↑	↑	
miR-145-5p					↑	
miR-150-5p	↓ X	↓ X				
miR-181a-5p					↑	
miR-181b-5p	↑					
miR-182-5p						↓
miR-191-5p				↑	↑	
miR-195-5p	↑ ^	↑ X	↑	↑		
miR-199a-5p		↑				
miR-199a-3p			↑	↑		
miR-199b-3p		↑				
miR-212-3p						↑ ^
miR-214-3p			↑	↑	↑	
miR-221-3p	↓	↓				
miR-222-3p		↓	↓	↓	↓	
miR-224-5p		↓				
miR-302a-3p						↑
miR-320-3p				↑	↑	↑ ^
miR-322-5p						↑
miR-324-3p		↑	↑		↑	
miR-365-3p						↑
miR-374-5p					↑	
miR-378a-3p	↓	↓	↓	↓		
miR-382-5p	↑					
miR-423-3p						↑ ^
miR-429-3p						↑
miR-451					↑	
miR-484		↓				
miR-486-5p	↓	↓				
miR-494-3p						↓
miR-499-5p					↑	
miR-515-5p						↓

1.2.5.c MiRNAs ARE INVOLVED IN HYPERTROPHIC SIGNALLING PATHWAYS

Profiling studies have been extended to experiments designed to understand the role of single miRNAs in the induction and progression of pathological cardiac hypertrophy to heart failure. Gain- and loss- of-function experiments have uncovered both pro- and anti-hypertrophic miRNAs. A limited number of mRNA targets have also been identified, allowing us to begin to integrate miRNAs into the complexity of hypertrophic signalling.

miR-1 and miR-133a are anti-hypertrophic miRNAs and are downregulated during cardiac hypertrophy, which leads to stimulation of hypertrophic pathways via the de-repression of targets such as IGF-1, NFATc, calmodulin and calcineurin (*118*). Downregulation of miR-1 results in increased levels of IGF-1, which leads to activation of the small GTP protein Ras and subsequent activation of the MAPK signalling pathway. As mentioned previously, the MAPK pathway results in the dual phosphorylation and activation of p38, JNKs and ERKs, which each phosphorylate multiple targets and transcription factors that induce hypertrophic gene expression. IGF-1 also results in activation of PI3K, which turns on the serine/threonine kinase Akt, and leads to activation of mTOR and inhibition of GSK3β. Activated GSK3β normally negatively regulates hypertrophy transcription factors such as GATA4 and NFAT, while mTOR increases protein synthesis. Additionally, both miR-1 and miR-133a have been implicated in the calcineurin-NFAT circuit. Activation of calcineurin dephosphorylates NFAT, which allows NFAT to translocate to the nucleus and activate the heart failure gene program. MiR-1 negatively regulates calmodulin, whilst miR-133a represses both calcineurin and NFATc4. Thus, miR-1 and miR-133a are targeting ligands (IGF-1), effectors (calcineurin) and transcriptional regulators (NFATc4) of hypertrophy that ensures appropriate regulation of signalling pathways at all levels. In line with their anti-hypertrophic role, inhibition of miR-133a *in vivo* by an antagomiR (a small synthetic RNA that is perfectly complementary to the target miRNA, silencing its expression) is sufficient to induce the hypertrophic response (*140*).

MiR-9 is another anti-hypertrophic miRNA whose expression decreases after hypertrophy stimulated by aldosterone and isoproterenol. Administration of a miR-9 mimic into the adult mouse heart was capable of inhibiting the hypertrophic response after stimulation with isoproterenol (a β- adrenergic agonist that stimulates of β- adrenergic receptor, a type of GPCRs) (*141*). MiR-9 overexpression prevented the increase in heart weight,

cardiomyocyte size and upregulation of hypertrophic markers as seen in the control animals (*141*). Furthermore, administration of miR-9 improved fractional shortening, the ratio between the diameter of the LV when it is relaxed and contracted and is a measure of the contractile function of the left ventricle (LV), consolidating its anti-hypertrophic role (*141*). MiR-9 inhibits hypertrophy by directly targeting myocardin, a transcriptional coactivator that mediates the signalling of NFATc3 by inducing expression of hypertrophic genes such as ANF, BNP, α-SkA and β-MHC (*141, 142*). Isoproterenol leads to an increase in the dephosphorylated and active form of NFATc3, which transcriptionally upregulates myocardin to convey the hypertrophic signal. The decrease of miR-9 in isoproterenol and aldosterone stimulated hypertrophy removes the negative regulation of myocardin, and thus the calcineurin/NFAT-signalling pathway is amplified. To date, changes in the expression levels of miR-9 in hypertrophy induced by pressure overload have not been documented. However, as myocardin levels also increase in the hearts of mice subjected to TAC, it is plausible that miR-9 regulation of myocardin is downstream of several hypertrophic stimuli (*142*). Interestingly, the length of both miR-9b and miR-9c are dynamically regulated at the 3' end during Drosophila development (*83*). It is currently not known if miR-9 has variable 3' ends in mammalian cardiomyocytes and if they play a role in cardiac development and hypertrophy.

Profiling studies have also revealed miRNAs whose expression increases during the hypertrophic response. Several of these miRNAs, such as miR-21, miR-208a and miR-195, are pro-hypertrophic and stimulate cardiac remodeling and hypertrophy. Cardiac specific overexpression of miR-208a was sufficient to induce hypertrophy, and mice lacking miR-208a are resistant to fibrosis and hypertrophy after cardiac stress (*114, 115*). The pro-hypertrophic properties of miR-208a are partially mediated by the repression of T3 receptor coregulator 1 (THRAP1) and myostatin, both negative regulators of muscle growth and hypertrophy (*115*). Myostatin normally inhibits phosphorylation of p38, a downstream effector of the MAPK pathway, and Akt (see above)(*143*). Repression of myostatin by miR-208a would thus increase hypertrophic signalling via the MAPK and PI3K/Akt pathways. It was recently shown that miR-27b is also pro-hypertrophic and can positively regulate hypertrophy *in vivo*. Wang et al. generated transgenic mice expressing miR-27b under the control of the α-MHC promoter (*144*). The miR-27b overexpressing mice were born normal but developed severe cardiac hypertrophy at 3 months of age, characterized by increased heart weight, LV enlargement and expression of hypertrophy

markers like ANF. Fractional shortening was also decreased in miR-27b mice indicating impaired cardiac function. The levels of miR-27b normally increase in hearts subjected to pressure overload by TAC, while inhibition of miR-27b levels using antagomiRs in mice already subjected to TAC reversed the signs of hypertrophy and returned LV contractile function to normal. MiR-27b directly targets peroxisome proliferator-activated receptor γ (PPARγ), a transcription factor that inhibits the development of cardiac hypertrophy likely through inhibition of NF-κB (*144*).

It is evident from studies thus far that miRNAs regulate signalling pathways critical to cardiac hypertrophy. An increased understanding of miRNA function and their targets within cardiomyocytes will allow us to incorporate miRNAs into the signalling cascades that occur after cardiac insult.

1.3 Alternative polyadenlyation

miRNA-mediated regulation depends equally on the presence of the corresponding binding site(s) within the mRNA. This principle not only underlies differential targeting of mRNAs originating from different genes but also applies to mRNA variants derived from the same gene. This is because cells, in addition to promoter choice and variable splicing, can also use multiple sites for alternative mRNA 3' end cleavage and polyadenylation (APA) to generate further transcript diversity.

1.3.1 Polyadenylation

The 3' UTRs of mRNAs contain various *cis*-acting elements involved in post-transcriptional gene regulation, including miRNA and RNA binding protein binding sites, as well as one or more polyadenylation signals (PAS). Polyadenylation involves recognition of the PAS and subsequent cleavage at the downstream polyadenylation (poly(A)) cleavage site, followed by polymerization of the adenosine tail by poly(A) polymerase (PAP) to a species-specific length (~150-250 nucleotides in mammals) (*145, 146*).

The human 3' processing complex consists of approximately 85 proteins, including cleavage and polyadenylation specificity factor (CPSF), cleavage stimulation factor (CstF), cleavage factor I and II (CFIm and CFIIm) and PAP, some of which interact with sequence elements surrounding the poly(A) cleavage site (Figure 1.4) (*147*). A conserved

consensus hexamer AAUAAA, the PAS, is found 10-35 nt upstream of the poly(A) cleavage site and is recognized by the multi-subunit CPSF, which consists of at least five subunits (CPSF 160, CPSF 100, CPSF73, CPSF 30 and hFip1) (*148*). CPSF160 interacts with the PAS (*149*), CPSF73 is responsible for the cleavage of the mRNA (*150*), and hFip1 (*151*) and CPSF30 (*152*) bind to the U-rich sequences that surround the cleavage site. The other predominant sequence element is a G/U or U-rich region located 30 nt downstream of the cleavage site (downstream sequence element, DSE). The DSE is recognised by CstF64, a component of the heterotrimeric CstF, which increases the efficiency of 3' end processing (*153*). In the absence of the canonical PAS, cleavage of the 3' end depends on an upstream UGUA sequence, which is recognised by CFIm (*154*). Additional sequences located upstream (upstream sequence elements, USE) or downstream (auxiliary downstream element, auxDSE) of the cleavage site act as accessory sequences for 3' processing. USEs are U-rich and provide additional interaction sites for the canonical 3' end processing machinery or recruit auxiliary processing factors (*155*). The auxDSEs are normally G-rich and regulate cleavage by providing a binding site for regulatory factors enhancing 3' end processing efficiency (*156, 157*). The sequence elements and associated proteins are critical for the regulation of poly(A) site usage.

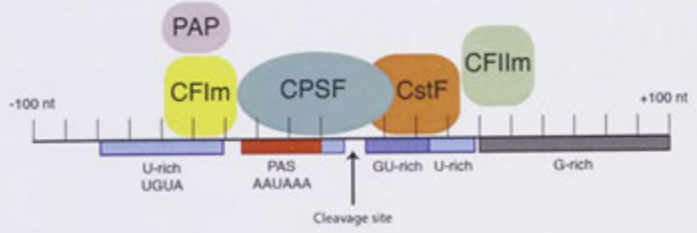


Figure 1.4. 3' end processing machinery and sequence elements.

Components of the 3' end processing machinery and the sequence elements they interact with around the poly(A) cleavage site. CFIm; Cleavage factor I, CFIIm; Cleavage factor II, CPSF; Cleavage and polyadenylation specificity factor, CstF; Cleavage stimulation factor. Image modified from (*158, 159*).

1.3.2 Alternative polyadenylation

More than half of the genes in the human genome can be subject to alternative cleavage and polyadenylation, meaning that the terminal exon contains multiple poly(A) sites that can result in mRNA isoforms with different 3' UTR lengths (*160*) (Figure 1.5). Longer 3' UTRs generated from the distal poly(A) site are usually 2-fold longer than shorter 3' UTRs and thus place the mRNA under the regulation of more *cis*-acting factors (*161*). For example, 52% of human miRNA target sites exist downstream of the first poly(A) site (*162*).

Alternative polyadenylation (APA) appears to be tightly regulated; it is controlled by both differentiation and proliferation states as 3' UTRs lengthen during embryonic development and shorten in the generation of induced pluripotent stem (iPS) cells (*161, 163*). Cancer cells frequently exhibit short 3' UTRs to escape regulation by miRNAs and other *cis*-acting elements, resulting in the over-expression of oncogenes (*164*).

Studies using expressed sequence tag (EST) data have revealed that several tissue types are biased in their poly(A) site usage. For example, mRNAs in the brain have long 3' UTRs while testes have short 3' UTRs (*165, 166*). Despite the technological advances allowing measurement of 3' UTRs, the phenomenon of APA has not been extensively studied in cardiac biology.

1.3.3 APA in cardiac biology

There is some evidence for alternative 3' UTR regulation of mRNA transcripts encoding important cardiac functions in existing EST databases, such as the poly(A) cleavage site and 3' UTR database (PACdb)(*167*). For example, cell division cycle 42 (CDC42), a small GTP-binding protein that amplifies hypertrophic gene expression (*4*), contains multiple poly(A) sites along the length of its 3' UTR (Figure 1.6). ESTs suggesting use of both the distal and proximal poly(A) cleavage sites for CDC42 have been derived from the heart. An important implication for its post-transcriptional regulation is the presence of binding sites for several cardiac miRNAs (including miR-133a and miR-1) after the proximal poly(A) cleavage site, which are not present in the shorter 3' UTR isoform (Figure 1.6). Therefore, use of different poly(A) sites along the CDC42 3' UTR substantially changes the degree of expected miRNA-mediated repression.

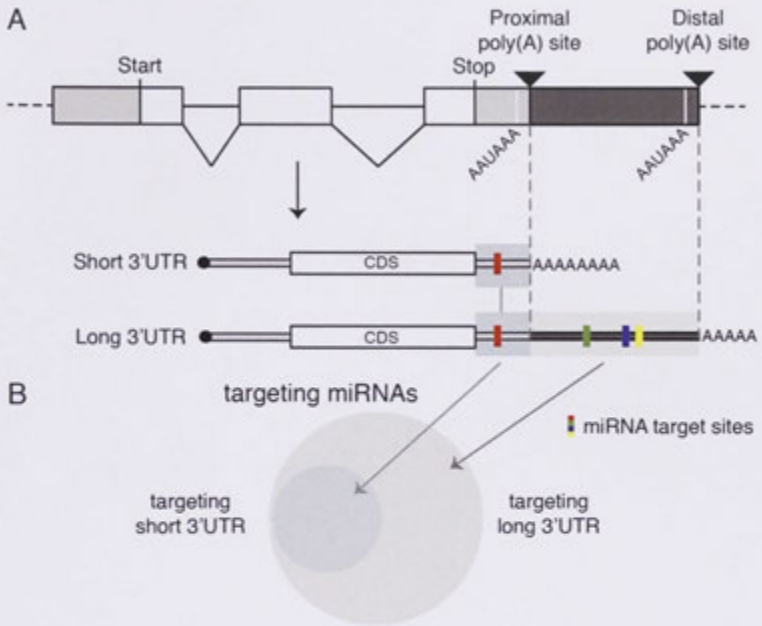


Figure 1.5. Alternative mRNA 3' end formation and consequences for targeting by miRNAs.

(A) Schematic of exon/intron organization of a gene. 5' and 3' UTR: grey areas; coding sequence: white area. mRNA 3' end cleavage and polyadenylation at alternative sites (poly(A) signal: AAUAAA) can generate mRNAs of different 3' UTR lengths, thus (B) expanding the range of miRNAs that may target a gene. Core miRNA targeting sites included in all mRNA variants: small circle; targeting sites in extended mRNA variants: large circle; miRNA target sites: upright bars. Image taken from (*77*) with permission.

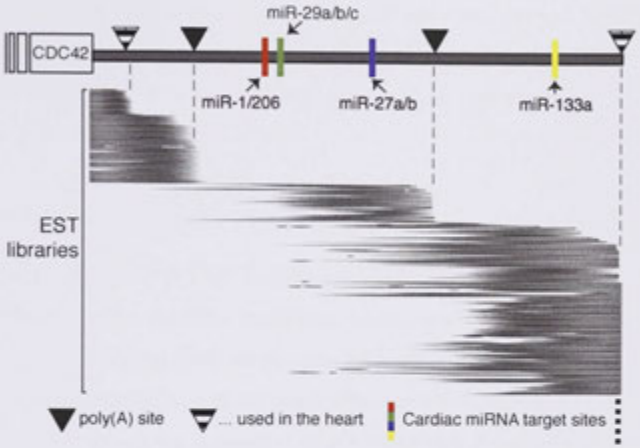


Figure 1.6. CDC42 mRNA exists as multiple 3' UTR variants.

Terminal exon of CDC42 indicating multiple functional poly(A) sites along its length. Lines represent EST data from PACdb (*167*) supporting these sites. The EST data supporting the distal site (on the right) was extensive, for graphical reasons the number of ESTs below this site has been abbreviated (indicated by the dots). Poly(A) cleavage sites: triangles (striped if based on cardiac-derived ESTs); target sites (confirmed/predicted) for miRNAs with cardiac function: coloured bars. Image taken from (*77*) with permission.

While initial global-based studies had insufficient depth (*166*), one recent study examined gene regulation in cardiac hypertrophy using genome-wide exon microarrays and analysed their data for altered 3' UTR usage (*168*). It was observed that 315 mRNAs changed 3' UTR length, with 211 becoming shorter, including Cdk13 and Tia1, suggesting there may be a tendency towards mRNAs losing miRNA-mediated regulation in hypertrophy. However, as this microarray analysis relied on poorly annotated murine 3' UTRs in public databases and lacks the sensitivity to unambiguously define expressed UTRs, further work is required to determine UTR usage in the heart and to understand UTR dynamics in hypertrophy.

1.3.4 Regulation of APA

While our appreciation that 3' UTR lengths are regulated is maturing, our knowledge of how this regulation occurs is still in its infancy. Genome-wide studies have indicated that there is a correlation between levels of proteins involved in polyadenylation and the global 3' UTR length. For instance, two of the core polyadenylation factors, CPSF and CstF, are strongly upregulated during the generation of iPS cells (*163*). A higher concentration of 3' processing factors would allow the proximal poly(A) site, which often has a weaker PAS, to be chosen and, thus, shorter 3' UTRs generated. In agreement, mRNAs have shorter 3' UTRs in iPS cells (*163*). Interestingly, the mRNAs corresponding to CstF proteins are upregulated as early as three days post-transverse aortic constriction, indicating that 3' UTR isoform regulation may be a component of very early stage hypertrophy (*169*).

The functional outcomes of 3' UTR length variation have been best studied in cancer, where it is recognized that oncogenes are often expressed with short/minimal 3' UTRs and become refractory to miRNA-mediated repression, greatly increasing their expression (*164*, *170*). Tissue-specific 3' UTRs have also been observed in Zebrafish, which presumably alter the ability of specific miRNAs to repress the expression of the respective mRNAs in different cellular environments (*171*). Commonly, multiple 3' UTR isoforms are co-expressed and therefore different portions of the mRNA population fall under the regulation of miRNAs (*170*). Furthermore, short 3' UTRs often still encode miRNA-binding sites, suggesting a shift to a different form of miRNA-mediated regulation in these cases. It is likely that the gene expression outcome of 3' UTR length variation will be specific for individual mRNAs and the miRNA environment they find themselves in. The observation that some mRNAs change their 3' UTR length in hypertrophy according to

microarray analysis suggests that complex changes to miRNA:mRNA networks are occurring in this response that warrant further analysis (*168*). The possibility of directing 3' UTR usage as a therapeutic intervention (*172*) further highlights this as an area of research interest for the heart and its disease states.

There have been inconsistent findings in the literature examining the expression levels of polyadenylation factors in the mouse embryo. One study showed high levels of some polyadenylation factors but low levels of others in the embryonic heart (*163*), while another study showed low levels of all polyadenylation associated factors (*166*). Furthermore, the global studies making use of ESTs have focused mainly on embryonic samples so little is known about 3' UTRs in the adult and very little data is available on APA regulation in disease (with the exception of cancer). However, the fact that there is a correlation between the levels of the proteins involved in polyadenylation and global 3' UTR length suggests that 3' UTR length is dynamically regulated.

1.4 Aims

MiRNAs play critical roles in the heart; they regulate the molecular pathways that are deregulated during the progression of heart disease. It is now known that cells can produce both miRNAs and their mRNA targets in multiple processing variants, which increases the complexity of miRNA-mediated control. Thus, the hypothesis underpinning these studies is that during the cardiac hypertrophic response there are changes to miRNA processing, as well as 3' UTR usage, which will alter miRNA:mRNA interactions.

The overall aim of this thesis is to examine variants of both miRNAs and mRNAs that can modify the miRNA:mRNA interactions in both normal and hypertrophic tissue. A mouse model of left ventricular hypertrophy will be established and cardiomyocytes isolated to examine changes occurring specifically in this cell type. Next generation sequencing will then be used to generate miRNA and mRNA 3' end datasets. The data will be interrogated for expression level changes during the development of hypertrophy, and for evidence of processing variants.

1.4.1 Detailed Aims

To achieve the overall aim of identifying miRNA and mRNA 3' UTR variant changes during cardiac hypertrophy, the project was divided into the following specific aims:

Aim 1: To establish a murine model of left ventricular hypertrophy (Chapter 3).

Aim 2: To identify changes in miRNA expression and processing in normal, pre-hypertrophic and hypertrophic cardiomyocytes (Chapter 4).

Aim 3: To determine mRNA 3' UTR usage and differential gene expression in normal, pre-hypertrophic and hypertrophic cardiomyocytes (Chapter 5).

Aim 4: To investigate the functional and mechanistic implications of these processing variants on miRNA action (Chapter 6).

Chapter two

Materials and Methods

2.1 Materials

2.1.1 Chemicals and Reagents

Amresco (Solon OH, USA)	RNA EZ-Vision® dye
BD Biosciences (Franklin Lakes, NJ, USA)	Bacto™ tryptone peptone digest, Bacto™ yeast extract, Bacto™ agar
Corning (Tewksbury, MA, USA)	T-75 flasks, 6-well plates, *0.22* µm *Vacuum filter*
GE Healthcare (Little Chalfort, BKM, UK)	GE Hybond-N^+ membrane
Life Technologies (Carlsbad CA, USA)	Novex® TBE-Urea Sample Buffer, SYBR® Gold Nucleic Acid Gel Stain, DMSO, BigDye® Terminator v1.1 Ready Reaction Mix, Ampicillin, AmpliTaq Gold® 360 Master Mix, Glycogen, ATP, Lipofectamine™ 2000, Opti-MEM® Reduced Serum Media, TRIzol® Reagent, Phenol, UltraPure™ DNase/RNase-Free Distilled Water, NaAc, DMEM, fetal bovine serum (FBS), Fast SYBR® Green Master Mix
Merck (Darmstadt, Hesse, Germany)	Ethanol, isopropanol, Na_2HPO_4, Pellet Paint® Co-precipitant
National Diagnostics (Atlanta, GA, USA)	SequaGel - UreaGel System
New England Biolabs (Ipswich, MA, USA)	Bovine serum albumin (BSA), Random primer mix, 10 x RNL2 buffer, NEB Next® Small RNA Library Prep Set for SOLiD™, dNTPs, Streptavidin magnetic beads
PerkinElmer (Waltham, MA, USA)	[γ-^{32}P]-ATP
Promega (Madison, WI, USA)	Passive Lysis Buffer, ammonium persulfate , X-gal
Roche Applied Science (Indianapolis, IN, USA)	LightCycler® 489 SYBR Green I Master
Sigma	Agarose – Low gelling temperature, Claycomb media,

(St. Louis, MO, USA)	Penicillin-streptomycin (Pen-Strep), noradrenaline, L-glutamine, Trypsin-EDTA, fibronectin, gelatin from bovine skin, TEMED, acrylamide/bis-acrylamide 40% solution, Perfect Hyb Buffer, 24% Peg 8000, EtBr solution, Tris-HCl. KCl Bioultra, chloroform, trisodium citrate, HCl, KH_2PO_4, $MgSO_4$-$7H_2O$, $NaHCO_3$, Na-Hepes, taurine, BDM, ascorbic acid, phosphate buffered saline (PBS), NaOH, NaOAc, EDTA, bromophenol blue, xylene cyanol, formamide
VWR International (Radnor, PA, USA)	NaCl, glucose, $CaCl_2$, urea

2.1.2 Kits

Agilent	RNA 6000 Nano LabChip kit, High Sensitivity DNA Analysis kit
Life Technologies	Poly(A) Tailing kit
Promega	PureYield™ plasmid miniprep system, Dual-Luciferase® Reporter Assay System
Qiagen (Hilden, North Rhine-Westphalia, Germany)	QIAquick gel extraction kit

2.1.3 Enzymes

Most enzymes were sourced from New England Biolabs. Other enzymes were supplied by:

Affymetrix (Santa Clara, CA, USA)	Optikinase
Life Technologies	Type II Collagenase, Turbo DNase, RNaseOUT™ Recombinant Ribonuclease Inhibitor, TURBO™ DNase, Superscript® III Reverse Transcriptase
Promega	Shrimp Alkaline Phosphatase (SAP)
Roche Applied Science	RNase T1, Fast Start Taq DNA polymerase
Worthington (Lakewood, NJ, USA)	Type II Collagenase

2.1.4 Buffers and solutions

2.1.4.A Buffers and solutions for general molecular biology

1 x TAE:

40 mM Tris-HCl (pH 8.2), 20 mM NaAc and 10 mM EDTA (pH 8.2).

2 x RNA loading dye:

1 mM EDTA, 0.1% bromophenol blue, 0.1% xylene cyanol, 95% formamide.

1 x TE:

10 mM Tris-HCl (pH7.5) and 1 mM EDTA.

1 x PAGE elution buffer:

2.5 M NaAc in TE buffer.

1 x Low TE:

10 mM Tris-HCl (pH8) and 0.1 mM EDTA.

0.5 x TBE:

45 mM Tris-borate (pH8), 1 mM EDTA.

10 x digest buffer:

3 M NaCl, 10% Tris (pH7) in double distilled water (ddH_20).

2 x ligation buffer:

40 µl 24% Peg 8000 and 10 µl 10 x RNL2 buffer

2.1.4.B Buffers and solutions for Northern blotting

20 x SSC:

3 M NaCl, 0.3 M trisodium citrate with the pH adjusted to 7 using HCl.

2.1.4.C Buffers and solutions for cardiomyocyte isolation

CM isolation perfusion buffer:

120.4 mM NaCl, 14.7 mM KCl, 0.6 mM KH_2PO_4, 0.6 mM Na_2HPO_4, 1.2 mM $MgSO_4$-$7H_2O$, 4.6 mM $NaHCO_3$, 10 mM Na-Hepes, 30 mM taurine, 10 mM 2,3-butanedione

monoxime (BDM), 5.5 mM Glucose. pH 7.3 and filtered through 0.22 µM filter. Prepared fresh each day.

Collagenase digestion buffer:

CM isolation perfusion buffer containing 2.4 mg/ml Type II Collagenase.

Collagenase stop buffer:

CM isolation perfusion buffer containing 1 mg/ml BSA and 125 µM $CaCl_2$.

2.1.4.D Buffers and solutions for tissue culture

Noradrenaline:

10 mM Noradrenaline made up in 30 mM ascorbic acid, filter-sterilized using 0.2 µm Acrodisc syringe filter. Noradrenaline was made fresh monthly.

Gelatin/Fibronectin:

0.02% gelatin solution made up in distilled water then autoclaved. Fibronectin was then added to the gelatin solution to a final concentration of 5 µg/ml.

Supplemented Claycomb media:

Claycomb media containing 10% FBS, 100 µg/ml Pen-Strep, 0.1 mM Norepinephrine and 2 mM L-Glutamine. Supplemented Claycomb media was stored in aluminum foil for a maximum of two weeks.

HL-1 freezing media:

95% FBS, 5% DMSO

HL-1 wash media:

Claycomb media containing 5% FBS and 100 µg/ml Pen-Strep.

2.1.5 Plasmids

IDT	pIDTSMART
Promega	psiCHECK™-2, pGEM®-T

2.1.6 Oligonucleotides

Synthetic DNA primers for q-PCR and RNA oligonucleotides for small RNA Northern blots were synthesized by Integrated DNA Technologies (Coralville, IA, USA). Mission® miRNA mimics (Sigma) were used for miRNA functional studies.

2.1.7 Bacterial strains

Escherichia coli (E.coli) DH5α

2.1.8 Bacterial growth media

Luria Broth Agar:

10 g/L Bacto® tryptone peptone digest, 5 g/L Bacto® yeast extract, 10 g/L Sodium chloride, 15 g/L Agar: agar was made up to a volume of 1 L with reverse osmosis (RO) H_2O before autoclaving. One the media had cooled to approximately 55°C, ampicillin (100 µg/ml) was added, the plates poured and then stored at 4°C.

Luria Broth:

10 g/L Bacto® tryptone peptone digest, 5 g/L Bacto® yeast extract, 10 g/L Sodium chloride made up to a volume of 1L with RO H_2O before autoclaving.

Transformation Buffer 1 (TFB1):

0.294 g Potassium Acetate, 0.989 g MnCl, 0.147 g $CaCl_2$. 15 ml 100% Glycerol were added to 75 ml H_2O and the pH brought to 5.8 with HAc. 1.2g of RbCl was then added and the final volume brought to 100 ml with H_2O. The buffer was then filter-sterilised using a 0.22 *µm* Vacuum filter and stored at 4 ℃.

Transformation Buffer 2 (TFB2):

0.294 g NaMOPs pH 7.0 was added to 1.1 g $CaCl_2$. 0.12g RbCl and 15 ml 100% Glycerol. The final volume was brought to 100 ml with H_2O. The buffer was then filter-sterilised using a 0.22 *µm* Vacuum filter and stored at 4 ℃.

2.1.9 DNA markers

Affymetrix Low molecular weight marker

BioLine (London, UK)	Hyperladder™-V
Invitrogen	10bp DNA ladder, 1kb plus DNA ladder

2.1.10 Columns

Bio-Rad	Micro Bio-Spin™ Chromatography Columns
Eppendorf	Phase Lock Gel™ Tubes
GE Healthcare	G-25 Microspin Column
Sigma	Corning® Costar® Spin-X® Plastic Centrifuge Tube Filters
Qiagen	MaXtract High Density Tubes

2.2 Methods

2.2.1 Cloning

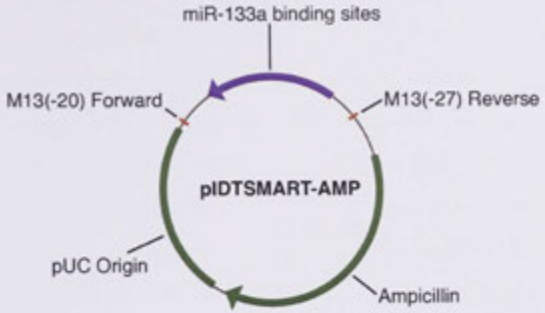
2.2.1.a Construction of three site isomiR vectors

Four miRNA binding cassettes (CTGF, CTGF MUT, PGAM1 and PGAM1 MUT binding sites) were ordered as one miniGENE within the pIDTSMART plasmid (Figure 2.1). Each cassette contained a specific restriction site at the 3' end (BsiWI, SpeI, SacI and NdeI respectively), and a NotI restriction site separated each cassette. The plasmids CTGF, CTGF MUT, PGAM1 and PGAM1 MUT were created by digesting pIDTSMART with *Not*I and the fragments inserted into the *Not*I site (see Digestion of psiCHECK™-2 with *Not*I below) behind the *Renilla* luciferase gene in the psiCHECK™-2 Vector (Figure 2.1). The specific constructs were selected using the restriction sites specific to each cassette and the final plasmid constructs were verified by sequencing at the Australian Cancer Research Foundation (ACRF) at the Garvan Institute, Sydney. The sequencing of all psiCHECK-2 vectors are listed in Appendix 8.14

2.2.1.b Digestion of psiCHECK-2 with Not1

5 µg of psiCHECK™-2 was digested with 10 units (U) of *Not*I for 1 hour at 37°C; 300ng was loaded onto a 2% agarose gel to check the digestion was successful. Dephosphorylation of the plasmid was necessary to ensure the plasmid did not recircularise

A



B

UAGCGGCCGCAUUAAAUUUGUUGAGAGUGUGACCAAAUGUCGACUAUUUGU
UGAGAGUGUGACCAAAUAUAAUAUUUGUUGAGAGUGUGACCAAACGUACGU
AGCGGCCGCAUUAAAUUUGUUGAGAGUGUGACCGAAAUGUCGACUAUUUGU
UGAGAGUGUGACCGAAAUAUAAUAUUUGUUGAGAGUGUGACCGAAAGACUA
GUUAGCGGCCGCAUUAAGCUGCUGGCCCUAGAAGGUUGGGAUCAAUCUGUC
GACUGCUGCUGGCCCUAGAAGGUUGGGAUCAAUCUAUAAUGCUGCUGGCCC
UAGAAGGUUGGGAUCAAUCGAGCUCUAGCGGCCGCAUUAAGCUGCUGGCCC
UAGAAGGUUGGGAUCGAAUCUGUCGACUGCUGCUGGCCCUAGAAGGUUGGG
AUCGAAUCUAUAAUGCUGCUGGCCCUAGAAGGUUGGGAUCGAAUCCAUAUGU
AGCGGCCGCUUAUU

— Not1 — BsiW1 — Sac1 — CTGF binding site
— Sal1 — Spe1 — Nde1 — PGAM1 binding site

C

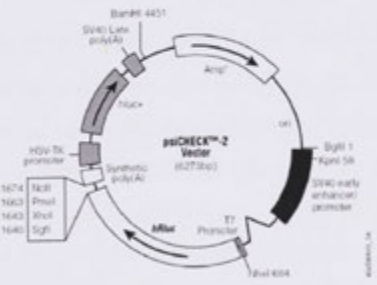


Figure 2.1. Plasmid maps of pIDTSMART and psiCHECK™-2.

(A) pIDTSMART-AMP containing the miR-133a-3p binding sites for all four luciferase vectors. (B) The miR-133a-3p binding site insert in pIDTSMART-AMP. The insert contained all three luciferase vectors with binding sites for miR-133a-3p (CTF binding site) and the isomiR (PGAM1 site) as well as mutated binding sites, separated by restriction enzyme sites. (C) psiCHECK™-AMP luciferase vector. The miR-133a-3p binding sites were cloned behind the renilla luciferase gene using the *Not*I site.

without the insert sequence. Therefore, the cut psiCHECK™-2 was treated with 1 U of Shrimp alkaline phosphatase (SAP) in 10% SAP buffer at 37°C for 30 minutes (min).
The dephosphorylated psiCHECK™-2 was then purified using a phenol-chloroform extraction. An equal volume of phenol-chloroform was added to the sample, mixed briefly then centrifuged at 13 000g at room temperature (RT) for 5 min. The top layer was transferred to a fresh tube and topped up to a final volume of 100 µl with 1 x TE solution. An equal volume of chloroform was added and the sample centrifuged at 13 000g at RT for 5 min, again the top layer was transferred to a fresh tube and topped up to 100 µl with with 1 x TE solution. The construct was then precipitated with 2.5x 100% ethanol (EtOH) containing 10% sodium acetate (NaOAc) and 1 µl pellet paint and spun at 22 000g at 4°C for 5 min. The supernatant was discarded and the pellet resuspended in 400 µl 70% EtOH, quickly spun and the supernatant discarded. The pellet was then dried under UV light, resuspended in 20 µl H_2O and stored at 4°C.

2.2.1.c Reducing isomiR sites in psiCHECK-2

2 sites:

To keep 2 binding sites, 1µg of each construct was digested with 10 U of *Sal*I and 10 U of *Xho*I for 1 hour at 37°C, followed by heat inactivation at 65°C for 20 min. Klenow enzyme was then used to create blunt ends; 1 µL Klenow with 0.5 µL dNTP were incubate at 37°C for 30 min, followed by heat inactivation at 75°C for 20 min. The fragments were separated by low melt agarose gel electrophoresis at 80 volts (V) for 20 min, the linearized plasmids excised and extracted with the QIAquick gel extraction kit. 250 ng of plasmid was then blunt ligated with 200 U of T4 DNA Ligase at RT for 2 hours. The vectors were then propogated in *E.coli* (see Transformation into prokaryotic cells and propagation of vectors). Plasmids extracted from individual colonies were checked with *Not*I for insert size and redigested with *Sal*I and *Xho*I (the sites should be removed from the Klenow and ligation reactions and therefore should not be digested). The final constructs were verified by conventional sequencing. The sequencing of all psiCHECK-2 vectors are listed in Appendix 8.14

1 site:

To keep 1 binding site, 1 µg of each construct was digested with *Sal*I and the restriction site specific to each construct (Table 2.1). The DNA Polymerase I, large (Klenow)

Fragment was then used to create blunt ends; 1 U Klenow enzyme with 120 µM dNTPs was incubated at 37°C for 30 min, followed by heat inactivation at 75°C for 20 min. The fragments were separated by low melt agarose gel electrophoresis (80V, 1 hour 20 min), the linearized plasmids excised and extracted with the QIAquick gel extraction kit. 250 ng of each plasmid was then blunt ligated with 0.5 µl T4 DNA Ligase at RT for 2 hours. The vectors were then propogated in *E.coli* (See Transformation into prokaryotic cells and propagation of vectors). The constructs were checked by digestion with *Not*I for insert size, the final constructs were verified by conventional sequencing. The sequencing of all psiCHECK-2 vectors are listed in Appendix 8.14

Table 2.1. Restriction site digestion conditions for miR-133a constructs.

Construct	Restriction Enzymes	Buffer Conditions	Reaction Conditions	Deactivation
CTGF	*Sal*I + *BsiW*I	NE buff. 3	37°C for 30min 55°C for 30min	80°C for 20min
CTGF Mut	*Sal*I + *Spe*I	NE buff 4. 0.2 µl BSA	37°C for 1 hour	80°C for 20min
PGAM1	*Sal*I + *Sac*I	NE buff 4. 0.2 µl BSA	37°C for 1 hour	65°C for20 min
PGAM1 Mut	*Sal*I + *Nde*I	NE buff 1.	37°C for 1 hour	65°C for 20min

2.2.1.D PREPARATION OF DHF ALPHA E.COLI

A scraping of frozen DHFα cells were put into 1 ml LB media and grown for 3 hours at 37°C with shaking at 220 rpm. The 1 ml was then transferred into 100 ml LB media and grown overnight at 37°C with shaking at 220 rpm. The optical density at 600 nm (OD600) was monitored to ensure the *E.coli* was in the exponential growth rate. Once the cells had an OD600 between 0.4-0.5, the cells were collected and alliquoted into 2 x 50 ml Falcon tubes, then chilled on ice for 10 min. The cells were centrifuged at 2 000 g for 5 min at 4°C. The supernatant was removed and the cells resuspended gently in 30 ml ice cold TFB1 and the two aliquots combined. The cells were chilled on ice overnight and centrifuged again at 2 000 g for 5 min at 4°C. The supernatant was removed and the cells

resuspended gently in 4 ml of ice cold TFB2, aliquoted into 0.1 ml per eppendorf tubes and immersed immediately under dry ice. The cells were then stored at -80°C.

2.2.1.E Transformation into prokaryotic cells and propagation of vectors.

Competent *E.coli* bacteria were used to propagate standard expression vectors. For transformation, approximately 50 ng DNA was incubated for 5 min on ice with 75 µl thawed *E.coli* cell suspension. The bacteria were exposed to heat-shock at 37°C for 1 min and then incubated for 5 min on ice. The bacteria were then plated on ampicillin agar plates and incubated overnight at 37°C for growth of colonies. Single colonies were picked from the plate and inoculated into LB media containing 100 µg/ml ampicillin, then incubated overnight at 37°C with shaking at 220 rpm. Overnight bacterial cultures were incubated in vessels with a volume of at least four times that of the culture media to allow for sufficient aeration. Plasmid purification was performed using the PureYield™ Plasmid Miniprep System kit according to the manufacturer's instructions. Purified plasmid DNA was quantified using the Nanodrop spectrometer.

2.2.1.F Conventional sequencing of plasmids

40ng of plasmid DNA and 3.2 pmoles of primer were submitted to the Australian Cancer Research Foundation (ACRF) at the Garvan Institute, Sydney.

2.2.1.G Agarose gel electrophoresis

1-2% agarose gels were used in the analysis of PCR products and plasmid constructs. The agarose powder was added to 1 x TAE, heated until the agarose powder was dissolved and ethidium bromide (EtBr) added to a final concentration of 0.5 mg/ml. Gels were then poured into horizontal gel boxes to set. One-tenth the volume of Ez-Vision® Three DNA Dye was added to the samples and size markers before loading.
DNA gels, once immersed in 1 x TAE buffer, were electrophoresed at 80-100V for 40-60 min. The EtBr was visualized by medium wavelength UV light and photographed using the Typhoon™ FLA 9000 (GE Healthcare).

2.2.2 Tissue Culture

2.2.2.a Maintenance of HL-1 cells

Flasks or wells were pre-coated with gelatin/fibronectin and incubated at 37°C overnight, the excess gelatin/fibronectin solution was removed by aspiration just prior to adding cells to the flasks. HL-1 cells were maintained in supplemented Claycomb media, with fresh media added daily. The cells were passaged only after full confluence. Each flask was rinsed with PBS warmed to 37°C, and 3 ml of 0.05% trypsin/EDTA per T-75 flask (flask with 75 cm^2 surface area) added and incubated at 37°C for 1 min. The trypsin/EDTA was removed, and fresh 0.05% trypsin/EDTA added then incubated for an additional 2 min at 37°C. If any cells remained adherent then the flask was tapped on the benchtop to dislodge remaining cells. 7 ml of HL-1 wash media was added and the cells collected in a 15 ml centrifuge tube. After centrifuging for 5 min at 500g, the supernatant was removed and the pellet gently resuspended in 3 ml of supplemented Claycomb media. For HL-1 culture maintenance the cells were split 1:3.

2.2.2.b Cryopreservation of HL-1 cells

Mammalian cell lines were preserved by cryopreservation in liquid nitrogen. Cells were frozen to -80°C at a rate of ~1°C/min using a Nalgene freezing container. One confluent T-75 flask (~$1x10^7$ *cells) was frozen into one cryovial. Cells were harvested as described above, and then resuspended in 1.5 ml of freezing media. HL-1 cells were thawed quickly (2 min) in a 37*°C water bath and transferred into 10 ml of pre-warmed HL-1 wash media. The cells were then centrifuged for 5 min at 500g, the wash media removed and the cells resuspended in 5 ml supplemented Claycomb media, and added to 10 ml of pre-warmed supplemented Claycomb media already in a T-75 flask. The media was replaced with fresh supplemented Claycomb media 4 hours after plating.

2.2.2.c Maintenance of HeLa cells

HeLa cells were maintained in DMEM containing 5% FCS and 100 µg/ml Pen-Strep. 0.05% trypsin/EDTA was used to detach the cells from the flask or dish.

2.2.2.D Luciferase assays

HeLa cells were maintained in DMEM with 5% FCS. For 24-well transfection, cells were plated at 5 x 10^4 cells per well into DMEM without antibiotics 24 hours prior to transfection. psiCHECK™-2 plasmids containing miRNA binding sites were cotransfected with MISSION® microRNA mimics using Lipofectamine™ 2000 according to manufacturer's instructions for transfection of plasmids. All transfections were performed in triplicate. 1µl Lipofectamine™ 2000/well was incubated in 50 µl Opti-MEM transfection medium/well for 5 min. This was then combined with 50 µl Opti-MEM containing 30 ng of plasmid and microRNA mimics at a final concentration of 10 nM and incubated at RT for 20 min to allow Mimic/Plasmid:Lipofectamine™ 2000 complexes to form. Immediately prior to transfection, 100 µl of media was removed from each well and replaced with 100 µl of transfection solution. The media was replaced after 6 hours and 24 hour post-transfection the cells were harvested for luciferase assays. The growth medium was removed and the cells washed once in 1 x PBS solution. 100 µl 1 x passive lysis buffer was added to each well and incubated at RT for 20 min with gentle shaking. 10 µl of cell lysate was used for luciferase assays; 50 µl of Luciferase Assay Reagent (LAR) and 50 µl of Stop and Glo® reagent were used to determine the expression of Firefly luciferase and Renilla Luciferase respectively. Addition of the substrates and measurement of the luminescence signals were performed using the FLUOstar Optima plate reader (BMG Labtech), with a gain of 3500. R-luc measurements were normalised against the corresponding F-luc value to correct for transfection efficiency.

2.2.3 RNA Isolation

All RNA work was done using UltraPure™ DNase/RNase-Free Distilled Water.

2.2.3.A RNA extraction from Heart tissue

RNA was extracted from left ventricular heart tissue using TRIzol®. Samples were completely homogenized in TRIzol®; to ensure high quality RNA the samples were kept on dry ice before adding the TRIzol® and periodically chilled during homogenization. 200 µl chloroform/1ml TRIzol® was added and the sample mixed vigorously for 15 seconds (sec) prior to incubation at RT for 3 min. The TRIzol® mix was then added to a pre-spun Maxtract high-density tube and spun at 12 000g for 4 min at 4°C. The upper aqueous phase, which contains RNA, was transferred to a new 1.5 ml Eppendorf tube containing 5

µl glycogen (5 mg/ml) and the RNA precipitated using 500 µl isopropanol. The precipitated RNA was mixed, incubated at RT for 10 min, centrifuged at 12 000g for 10 min at 4°C and the resultant RNA pellet was washed with 75% EtOH. The RNA was then vortexed and centrifuged at 7 500g at 4°C for 5 min. Following removal of the supernatant and briefly drying the pellet under a heat lamp, the RNA was redissolved using RNase-free water. 5 µL 3 M NaOAc was added and mixed, followed by 137.5 µl 100% EtOH and the RNA frozen at -20°C overnight. The next day, the sample was centrifuged at 20 000g for 15 min at 4°C and the supernatant removed. The RNA was then washed again with 75% EtOH and centrifuged at 7 000g for 5 min at 4°C. Following removal of the ethanol, the pellet was dried under a heat lamp then redissolved in 30 µl water. Spectrophotometric absorbance was measured by Nanodrop. The concentration of RNA was determined by the absorbance at 260nm, with one A_{260} OD unit equivalent to 40 µg/ml of RNA. After RNA isolation, the RNA quality was assessed using an Agilent 2100 Bioanalyzer (Agilent) together with the reagents in the RNA 6000 Nano LabChip kit. Samples with an RNA Integrity Number (RIN) >8 were considered suitable for downstream analysis. All RNA was aliquoted and stored at -80°C.

2.2.3.B RNA EXTRACTION FROM PRIMARY CARDIOMYOCYTES

5 ml of TRIzol® was added per heart to the isolated cardiomyocytes, 1 ml aliquots were then transferred into Eppendorf tubes. The RNA extraction was then performed as described in Section 2.2.3.A. It was necessary to further purify the cardiomyocyte RNA samples due to low 260/230 ratios (indicative of protein, salt or solvent contamination). The RNA volume was brought up to 50 µl and passed through a Micro Bio-Spin™ Chromatography Columns according to manufacturers instructions. The RNA was then reprecipitated by adding 137.5 µl 100% EtOH, 1 µl glycogen (5 mg/ml) and 5 µl 3 M NaOAc and frozen at -80°C overnight. The RNA was centrifuged at 20 000g for 15 min at 4°C and the supernatant removed. The RNA was then washed twice with 75% EtOH and centrifuged at 7 000g for 5 min at 4°C. The EtOH was removed, the pellet dried under a heat lamp and redissolved in 30 µl water.

2.2.3.C DNase treatment of RNA

RNA that was used for real-time quantitative PCR (RT-qPCR) analysis of 3' UTR variants and mRNAs was first DNase treated to avoid genomic contamination (as the primers cannot be designed across an exon-exon junction). 2 U TURBO DNase was added to 1 µg of total RNA, in 1 x TURBO DNase buffer at a final volume of 20 µl. The reaction was incubated at 37°C for 30 min. To inactivate the DNase, 3µl DNase inactivator was added and the sample vortexed and incubated at RT for 30 sec, this step was repeated 3 times. The RNA was then centrifuged in a mini-centrifuge for 1.5 min and the supernatant (approximately 20 µl) transferred to a new 0.2 ml Eppendorf tube.

2.2.4 Real time quantitative PCR (RT-qPCR)

2.2.4.A Stem-loop reverse transcription (microRNA)

RNA was extracted using TRIzol®, as described above. 300 ng of total RNA, 1 µl 10 mM dNTPs and 1 µl of a microRNA specific stem-loop primer (Table 2.2) were incubated at 65°C for 5 min followed by 3 min on ice. 1 µl 0.1 M dithiothreitol (DTT), 40 U RNaseOUT™, 200 U Superscript® III, 1 x First-strand buffer and water to a total volume of 20 µl were added. To perform reverse transcription, the reaction mixture was incubated at 16°C for 30 min, 42°C for 30 min, then 75°C for 15 min. cDNA was stored at -20°C and diluted 1:5 for individual expression analysis by RT-qPCR.

2.2.4.B Reverse transcription (mRNA)

RNA was extracted using TRIzol®, as described above. 10 µl of DNase treated RNA (~500 ng), 2 µl random primer mix and 6 µl water were incubated at 65°C for 5 min then ice for 3 min. The random primer mix contains a mixture of hexamers and anchored-dT primers, providing greater coverage of the 3' ends of RNA templates in comparison to traditional random priming by hexamer (*173*). 1 µl 0.1 M DTT, 40 U RNaseOUT™, 200 U Superscript® III, 1 x First-strand buffer and nuclease free water to a total volume of 20 µl were then added. To perform reverse transcription the reaction mixture was incubated at 50°C for 2 hours, then 75°C for 15 min. cDNA was stored at -20°C and diluted 1:10 for individual expression by RT-qPCR.

2.2.4.C RT-QPCR FOR MICRORNA USING STEM-LOOP RT PRIMERS

Forward specific and universal reverse primers (See Table 2.3) were used with the LightCycler® 480 SYBR Green I Master Mix or the Fast SYBR® Green Master Mix for rt-RT-qPCR. Each reaction was performed in triplicate using 2 µl of the diluted cDNA mixed with 8 µl Master Mix (5 µl SYBR, 0.5 µl of 5µM each primer and 2 µl water). The Master Mix was dispensed using the epMotion robot (Eppendorf). PCR amplification was performed in 384-well plate format on the LightCycler® 480 system (Roche) or the QuantStudio™ 12K Flex system (Life Technologies), the annealing temperature for each primer set was optimized and the final reaction conditions were: heating at 95°C for 10 min, 50 cycles of 95°C for 10 sec, annealing temp for 10 sec, 72°C for 4 sec, followed by melting curve analysis.

2.2.4.D RT-QPCR FOR MRNA

Forward and reverse specific primers (See Table 2.4) were used with LightCycler® 480 SYBR Green I Master Mix or the Fast SYBR® Green Master Mix for RT-qPCR. Each reaction was performed in triplicate using 2 µl diluted cDNA mixed with 8 µl Master Mix (5 µl SYBR, 0.5 µl of 5 µM each primer and 2 µl water). The Master Mix was dispensed using the epMotion robot (Eppendorf). PCR amplification was performed in 384-well plate format on the LightCycler® 480 system (Roche) or the QuantStudio™ 12K Flex system (Life Technologies), the annealing temperature for each primer set was optimized and the final reaction conditions were: heating at 95°C for 10 min, 50 cycles of 95°C for 10 sec, annealing temp for 10 sec, 72°C for 6 sec, followed by melting curve analysis.

Table 2.2. Stem-loop reverse transcription primers for miRNAs.

Name	Sequence	Gene
mSNO135_RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGAC CTTCAG	snoRNA-135
mSNO234_RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGAC TCTCAG	snoRNA-234
mSNO55_RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGAC TTGCTC	snoRNA-55
mSNO412_RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGAC TCTCAG	snoRNA-412
mSNO202_RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGAC CATCAG	snoRNA-202
miR-21_RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGAC TCAACA	miR-21-5p
mmiR-451_RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGAC AACTCA	miR-451
mmiR-214-5p_RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGAC GCACAG	miR-214-5p
mmiR-324-3p_RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGAC AGCAGC	miR-324-3p
mmiR-9-5p_RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGAC TCATAC	miR-9-5p
mmiR-199a-5p_RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGAC GAACAG	miR-199a-5p
mmiR-199a-3p_RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGAC TAACCA	miR-199a-3p
mmiR-31-5p_RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGAC CAGCTA	miR-31-5p
miR-214-3p_RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGAC ACTGCC	miR-214-3p
miR-199b-3p_RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGAC TAACCA	miR-199b-3p
miR-299-5p_RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGAC ATGTAT	miR-299-5p
miR-379-5p_RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGAC CCTACG	miR-379-5p
miR-143-5p_RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGAC CCAGAG	miR-143-5p
Hyp_miR-101a-3p_RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGAC TCAGTT	miR-101a-3p
Hyp_miR-10a-5p_RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGAC ACAAAT	miR-10a-3p
Hyp_miR-150-5p_RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGAC CACTGG	miR-150-5p

Table 2.3. Stem-loop RT-qPCR primers and experimental conditions for miRNAs.

Name	Sequence	Gene	Annealing Temp (°C)	Primer type
miR_unirev	GTGCAGGGTCCGAGGT	Universal Reverse	60	Reverse
mSNO135_For	TAGTGGTGAGCCTCTGGTTTT	snoRNA-135	60	Forward
mSNO234_For	TTAACAAAAATTCGTCACTACCA	snoRNA-234	60	Forward
mSNO55_For	CACTGACGACTCCATGTGTC	snoRNA-55	60	Forward
mSNO412_For	CTAACCTGATGCAAGTTACAAATTAC	snoRNA-412	60	Forward
mSNO202_For	AGTACTTTTGAACCCTTTTCCA	snoRNA-202	60	Forward
miR-21_For	CCGCTAGCTTATCAGACTGA	miR-21-5p	60	Forward
mmiR-451_For	GCGAAACCGTTACCATTAC	miR-451	60	Forward
mmiR-214-5p_For	GCGTGCCTGTCTACACTTG	miR-214-5p	60	Forward
mmiR-324-3p_For	GTCCACTGCCCCAGGT	miR-324-3p	60	Forward
mmiR-9-5p_For	CGGCTCTTTGGTTATCTAGCT	miR-9-5p	60	Forward
mmiR-199a-5p_For	GGCCCAGTGTTCAGACTAC	miR-199a-5p	60	Forward
mmiR-199a-3p_For	GCGACAGTAGTCTGCACAT	miR-199a-3p	60	Forward
mmiR-31-5p_For	TTAGGCAAGATGCTGGCA	miR-31-5p	60	Forward
miR-214-3p_For	CATACAGCAGGCACAGACA	miR-214-3p	60	Forward
miR-199b-3p_For	GCGACAGTAGTCTGCACAT	miR-199b-3p	60	Forward
miR-299-5p_For	CTATTGGTTTACCGTCCCAC	miR-299-5p	60	Forward
miR-379-5p_For	CGCCTGGTAGACTATGGAA	miR-379-5p	60	Forward
miR-143-5p_For	TAAGGTGCAGTGCTGCAT	miR-143-5p	60	Forward
Hyp_miR-101a-3p_For	ACGCCGTACAGTACTGTGATA	miR-101a-3p	60	Forward
Hyp_miR-10a-5p_For	CGTACCCTGTAGATCCGA	miR-10a-3p	60	Forward
Hyp_miR-150-5p_For	GATCTCTCCCAACCCTTGTA	miR-150-5p	62	Forward

2.2.5 3' RACE VALIDATION OF 3' ENDS

2.2.5.A TVN REVERSE TRANSCRIPTION

A superscript mix was made up with the following components: 6 µl 5 x First-strand buffer, 1.5 µl 0.1 M DTT, 60 U RNaseOUT™, 300 U Superscript® III and 1.5 µl TVN (12x T, V=C,G,A, N=A,C,T,G) anchor. 0.5 µg of RNA and 1 µl 10 mM dNTPs were then added to H_2O to give a final volume of 12 µl. The reaction was incubated at 65°C for 5 min, then 50°C for 5 min. 8 µl of the superscript mix was added to the RNA mix at 50°C, then the reaction was mixed at 50°C for 1 hour, followed by 75°C for 15 min. The cDNA was diluted 1:10 for PCR amplification. The sequences of primers and oligonucleotides used for 3' race validation are shown in Table 2.5.

2.2.5.B PCR AMPLIFICATION AND PURIFICATION OF PCR PRODUCT

2 µl of diluted cDNA was added to 0.4 µl 10 mM dNTPs, 1 U Fast start Taq DNA polymerase, 2 µl each of forward specific primer and reverse anchor primer (5 µM), 1 x buffer in a total volume of 20 µl. The reaction was then amplified using the following conditions: heating at 94°C for 15 min, 36-38 cycles of 94°C for 15 sec, 60°C for 30 sec, 70°C for 30 sec, followed by 70°C for 1 min. The PCR products were run with RNA EZ-Vision® dye on a 1.5% agarose gel at 100V for 45 min. The appropriate bands were excised and extracted using the QIAquick gel extraction kit according to manufacturer's instructions.

2.2.5.C CLONING AND SEQUENCING OF PCR PRODUCTS

4 µl of the PCR product was ligated into 50 mg of the pGEM®-T Vector with 400 U of T4 DNA Ligase in 1 x T4 DNA Ligase buffer at RT overnight. The ligation mixture was then transformed into competent *E.coli* cells by adding 50 µl *E.coli* suspension to each ligation mixture and incubating on ice for 30 min. The cells were then heat shocked at 42°C for 45 sec and put back onto ice for 2 min. 200 µl LB media was added and the cells allowed to recover by incubating them at 37°C for 1 hour with shaking at 220 rpm. 150 µl was then spread onto ampicillin plates coated with 25 µl ampicillin (100 µg/ml), 25 µl X-gal (20 mg/ml) and 40 µl 0.1 M IPTG and incubated overnight at 37°C. Single white colonies were picked from the plate and inoculated into 10 ml LB media containing 100 µg/ml

ampicillin, then incubated overnight at 37°C with shaking at 220 rpm. Overnight bacterial cultures were incubated in vessels with a volume of at least four times that of the culture media to allow for sufficient aeration. Plasmid purification was performed using the PureYield™ Plasmid Miniprep System kit according to the manufacturer's instructions as outlined. The insert size was confirmed by digesting pGEM®-T with 10 U EcoR1 at 37°C for 2 hours and running the digestion products on an agarose gel. 4 µl of pGEM®-T containing the PCR insert was combined with 1 µl BigDye® Terminator v1.1 Ready Reaction Mix, 3.2 µl M13 reverse primer (1 pmole/µl) in 1 x buffer and H_2O to a total volume of 20 µl. The sequencing PCR was run as follows: 94°C for 15 min, 30 cycles of 96°C for 10 sec, 50°C for 5 sec, 60°C for 4 min. The PCR product was then precipitated by adding 1 µl 3 M NaOAc, 5 µl 125 mM EDTA and 50 µl 100% EtOH, vortexing and centrifuging for 20 min. The supernatant was removed and the pellet washed twice with 75% EtOH. Once dried, the samples were taken to the Biomolecular Resource Facility, Canberra, Australia for BigDye® Terminator Sequencing.

2.2.6 Northern blots of small RNAs

2.2.6.a Polyacrylamide gel electrophoresis

A 12% urea polyacrylamide gel (12% acrylamide:bisacrylamide 19:1, 7 M urea in 0.5 x TBE) was prepared in 16 x 20 cm gel plates (the large gel is necessary to separate processing variants of miRNAs). To prevent the gel from heating, all electrophoresis was performed in a cold room at 4°C. The gel was prerun at 180V for 30 min, and the wells flushed carefully with a syringe. 5-15 µg of total RNA was mixed with an equal volume of 2 x RNA loading dye, denatured at 95°C for 5 min and cooled briefly on ice. Single-stranded RNA mimics of the miR-133a-3p were also loaded onto the gel for size comparison (miR-133a-3p of 23 nt in length, 5'-UUUGGUCCCCUUCAACCAGCUGU-3'; IsomiR of 22 nt in length, 5'-UUGGUCCCCUUCAACCAGCUGU-3'). The RNA was loaded onto the gel, and the gel run at 20 milliamps (mA) until the samples entered the gel. The gel was then run at 30-40mA (maximum voltage at 600 V) until the dye was 2 cm from the bottom of the gel. Prior to transfer, the gel was stained with EtBr (0.5 µg/ml) for 5 min and visualized by medium wavelength UV light and photographed using the Typhoon™ FLA 9000 (GE Healthcare).

Table 2.4. RT-qPCR primers and experimental conditions

Name	Sequence	Gene	Annealing Temp (°C)	Primer type
mGAPDH_qpcr_for	CTTGGGCTACACTGAGGAC	GAPDH		Forward
mGAPDH_qpcr_rev	CTGTTGCTGTAGCCGTATTC	GAPDH	60	Reverse
mCANX_qpcr_for	GCTTTGCCAGTGTTCCTT	CANX1		Forward
mCANX_qpcr_rev	TTCATCCCTCTTGTTCTTCTC	CANX1	60	Reverse
mHPRT_for	CCTAAGATGAGCGCAAGTTGAA	HPRT		Forward
mHPRT_rev	CCACAGGACTAGAACACCTGCTAA	HPRT	60	Reverse
mARBP_for	GGACCCGAGAAGACCTCCTT	ARBP		Forward
mARBP_rev	GCACATCACTCAGAATTTCAATGG	ARBP	60	Reverse
mDDR2_for	GCATTCTAAGGCTTGCTGCT	DDR2		Forward
mDDR2_rev	GAAGACGGAGTTGGATCTGG	DDR2	58	Reverse
mCol1a1_for	GTGTGATGGGATTCCCTGGACCTA	COL1A1		Forward
mCol1a1_rev	CCTGAGCTCCAGCTTCTCCATCTT	COL1A1	58	Reverse
mCol3a1_for	ACCCCTGGTCCACAAGGATTA	COL3A1		Forward
mCol3a1_rev	ACGTTCTCCAGGTGCACCAGAAT	COL3A1	58	Reverse
ma-vim_for	GATTTCTCTGCCTCTGCCAACCTT	α-VIM		Forward
ma-vim_rev	CATTGATCACCTGTCCATCTCTGG	α-VIM	58	Reverse
mctnt_for	GATCTCTGCAGATGCCATGATGCA	TNNT2		Forward
mctnt_rev	CAGTGCATCGATATTCTTGCGCCA	TNNT2	58	Reverse
m-mlc2v_for	TCTCCGTGGGTAATGATGTG	MLC2V		Forward
m-mlc3v_rev	ACTATGTCCGGGAGATGCTG	MLC2V	58	Reverse
mnppa_for	ATTGACAGGATTGGAGCCCAGAGT	NPPA		Forward
mnppa_rev	TGACACACCACAAGGGCTTAGGAT	NPPA	60	Reverse
maSka_for	TGAGACCACCTACAACAGCA	α-SKA		Forward
maSka_rev	CCAGAGCTGTGATCTCCTTC	α-SKA	60	Reverse
mNppb_for	GCCAGTCTCCAGAGCAATTCA	NPPB		Forward
mNppb_rev	TGTTCTTTTGTGAGGCCTTGG	NPPB	60	Reverse
mbMHC_for	CTACAGGCCTGGGCTTACCT	β-MHC		Forward
mbMHC_rev	TCTCCTTCTCAGACTTCCGC	β-MHC	60	Reverse
mmu_Hadh-FOR	TGC TGG CCT CCA CTT TTT CA	HADH		Forward
mmu_Hadh-REV	CAC CAA GAG TCG GTT CAC GA	HADH	60	Reverse
mmu_Cnn1-FOR	CGG CGT CAC CTC TAT GAT CC	CNN1		Forward
mmu_Cnn1-REV	TCA AAG ATC TGC CGC TTG GT	CNN1	60	Reverse
mmu_CLU-FOR	AGC CGT GCG GAA TGA GAT AG	CLU		Forward
mmu_CLU-REV	TTC TTC CCG AGA GCA GCA AG	CLU	60	Reverse
mmu_Klf15-FOR	TGT ACA CCA AGA GCA GCC AC	KLF15		Forward
mmu_Klf15-REV	ACA CTG GTA CGG CTT CAC AC	KLF15	60	Reverse
mmu_Pdk4-FOR	CTGCCTGACCGCTTAGTGAA	PDK4		Forward
mmu_Pdk4-REV	TGCCTTGAGCCATTGTAGGG	PDK4	60	Reverse

Table 2.4 cont. RT-qPCR primers and experimental conditions

Name	Sequence	Gene	Annealing Temp (°C)	Primer type
SPARC-CDS-FOR	ATTGGCGAGTTTGAGAAGGTA	SPARC		Forward
SPARC-CDS-REV	CAGGGGGCGATGTATTTG	SPARC	60	Reverse
mmu_Ndufa1-FOR	TTATGGGGGTGTGCTTGGTC	NDUFA1		Forward
mmu_Ndufa1-REV	ACGTCTATCGCGTTCCATCA	NDUFA1	60	Reverse
Acsl1-CDS-FOR	GCAGGCCTTTCTCATAGCAG	ACSL1		Forward
Acsl1-CDS-REV	AAGGCCGTTGTCAATAGAAAAT	ACSL1	60	Reverse
mmu_Calm1-FOR	GGCACCATTGACTTCCCAGA	CALM1		Forward
mmu_Calm1-REV	TCTGCCGCACTGATGTAACC	CALM1	60	Reverse
MTUS1_3UTR_Start_FOR	CAGGGCCAGTTTATCTTGGA	MTUS1 short and long 3'UTR		Forward
MTUS1_3UTR_Start_REV	GCTGGAAGAATGCTCAGAGG	MTUS1 short and long 3'UTR	60	Reverse
MTUS1_3UTR_Mid_FOR	TTGTGGCTGAAATTGGCTCA	MTUS1 long 3'UTR		Forward
MTUS1_3UTR_Mid_REV	AAACACTTAAGACCACAGCTGTTCC	MTUS1 long 3'UTR	60	Reverse
CCND2_3UTR_Start2_FOR2	GACTTGCCCTCACATTCCAT	CCND2 short and long 3'UTR		Forward
CCND2_3UTR_Start2_REV2	CCCTTTCGCAAAAACACAAT	CCND2 short and long 3'UTR	60	Reverse
CCND2_3UTR_Mid_FOR	CTGAGTCTGGTTGGTGCTGA	CCND2 long 3'UTR		Forward
CCND2_3UTR_Mid REV	ACACCCGAGACCACGAAAC	CCND2 long 3'UTR	60	Reverse
CAMKK2_3UTR_Start2_FOR2	CCCAGCTTCGAAGAACAAAG	CAMKK2 short and long 3'UTR		Forward
CAMKK2_3UTR_Start2_REV2	TTGCACGCACTCTACCAAAG	CAMKK2 short and long 3'UTR	60	Reverse
CAMKK2_3UTR_MID_FOR	TGGTTGAAGTTTGGGTGTGA	CAMKK2 long 3'UTR		Forward
CAMKK2_3UTR_MID_REV	GGATAGCGGCTAGACCACAG	CAMKK2 long 3'UTR	60	Reverse
PPM1K_3UTR_Start_FOR	AGTAGTGGTGCCCTTTGGTG	PPM1K short and long 3'UTR		Forward
PPM1K_3UTR_Start_REV	CAATGACACACGGGTACAGG	PPM1K short and long 3'UTR	60	Reverse
PPM1K_3UTR_Mid_FOR	TCCCTTAGGAGGGGTCAGTT	PPM1K long 3'UTR		Forward
PPM1K_3UTR_Mid_REV	TCCCACCCTGAAACATTCTC	PPM1K long 3'UTR	60	Reverse
GSK3b start_2_FOR	TACTGCCACACGCACATACA	GSK3B short and long 3'UTR		Forward
GSK3b start_2_REV	AGTACACCACTCACCGCACA	GSK3B short and long 3'UTR	60	Reverse
GSK3b middle_2_FOR	ATAATTGCCGTGCGAAGAC	GSK3B long 3'UTR		Forward
GSK3b middle_2_REV	GACTCTAGCACGCACACCAA	GSK3B long 3'UTR	60	Reverse
Etf1_Start_FOR	AAACTGCCTCCTGTCAGCTC	Etf1 short and long 3'UTR		Forward
Etf1_Start_REV	TTGCTGCATTTCGCAGTGAC	Etf1 short and long 3'UTR	60	Reverse
Etf1_Mid_For	AAACTGCCTCCTGTCAGCTC	Etf1 long 3'UTR		Forward
Etf1_Mid_Rev	TTGCTGCATTTCGCAGTGAC	Etf1 long 3'UTR	60	Reverse
Utp6_Start_FOR	GGAAAGCCGTCCATTTTCGG	Utp6 short and long 3'UTR		Forward
Utp6_Start_REV	TGTCCGCAGTTCTCAGGTTT	Utp6 short and long 3'UTR	60	Reverse
Utp6_Mid_For	AATCGTTTCCTGTGCCACCT	Utp6 long 3'UTR		Forward
Utp6_Mid_Rev	TCACCGAAGTGAGGCTGAAC	Utp6 long 3'UTR	60	Reverse
Htatip3_Start_For	GGCCGATAAGTCCAGCAGTT	Htatip2 short and long 3'UTR		Forward
Htatip3_Start_Rev	CAAGAGTCTGGCAGAGAGCC	Htatip2 short and long 3'UTR	60	Reverse
Htatip2_Mid_For	TGACGGTGGTTAGAGCGATG	Htatip long 3'UTR		Forward
Htatip2_Mid_Rev	CCTCCAGCATGGTCACAGTT	Htatip long 3'UTR	60	Reverse

Table 2.5. PCR primers and experimental conditions for validation of 3' ends.

Name	Sequence	Gene	Annealing Temp (°C)	Reverse primer	Number of Cycles	Primer type
CDC42_1400_qR60_for	CGAACCAATGCTTTCTTGT	CDC42 3'UTR	60	IC_anchor rev1	36	Forward
CDC42_50_qR60_for	CTCCAGAACCGAAGAAGAG	CDC42 3'UTR	60	IC_anchor rev1	38	Forward
Usf2_1400_qR60_for	GACATCGACCGTGTACTCAT	Usf2 3' UTR	60	IC_anchor rev1	38	Forward
Usf2_630_qR67_for	GTATGCCGCTGTGTGCGT	Usf2 3' UTR	60	IC_anchor rev2	38	Forward
SPARC_1100_qR60_for	CTCTCACAGCAATGTGAAGAG	SPARC 3'UTR	60	IC_anchor rev1	38	Forward
SPARC_100_QR60_for	GCTGCAGTCCTGAACTCTC	SPARC 3'UTR	60	IC_anchor rev1	38	Forward
SRF_2250_qR60_for	CATGGTCACATCTTCTGTACAG	SRF 3'UTR	60	IC_anchor rev1	38	Forward
SRF_900_qR60_for	CCCAGCTGCTTTACTTAAGAG	SRF 3'UTR	60	IC_anchor rev1	38	Forward
GSK3b short 3'UTR_4	TCTGCATCAGCTTCCAACTC	GSK3B 3'UTR	60	IC_anchor rev1	38	Forward
GSK3b long 3'UTR_4	TGTGGAAAGTATGCTCAATGTTT	GSK3B 3'UTR	60	IC_anchor rev1	38	Forward
M13_rev	CAGGAAACAGCTATGAC	pGEM-T reverse primer				Sequencing primer
IC_TVNanchor	CCA GTG CAG GGT CCG AGG TAT TTT TTT TTT TTV N	TVN anchor				TVN Anchor Reverse Transcription
IC_anchor rev1	GTG CAG GGT CCG AGG TA	Anchor Reverse 1				Reverse
IC_anchor rev2	CCA GTG CAG GGT CCG AGG TA	Anchor Reverse 2				Reverse

2.2.6.B Transfer and cross-linking of membrane RNA

After staining with EtBr, the gel was soaked in 0.5 x TBE for 15 min. Six sheets of 3MM Whatman filter paper and a Hybond-N$^+$ membrane were presoaked in 0.5 x TBE. Three filter papers were put onto a transfer cassette, followed by the gel, membrane and remaining filter paper. The RNA was transferred using a wet electrotransfer system (Biorad Trans-Blot transfer cell) for 1 hour at 20V. The transfer was performed in 0.5 x TBE and in the cold room. The transfer was checked by restaining the gel with EtBr, and by visualizing the membrane using medium wavelength UV light. The RNA was immobilized on the membrane by UV irradiation and baking: the membrane was UV cross-linked (120mJ/cm^2) using the Stratalinker® 1800 UV Crosslinker (Stratagene) then baked at 80°C for 30 min.

2.2.6.C Probe labelling and hybridisation

A DNA probe complementary to miR-133a-3p (5'-ACAGCTGGTTGAAGGGGA CCAA-3') was labeled with γ-^{32}P. 15 pmoles of DNA oligonucleotide was added to 50 µCi of γ-32P-ATP and 10 U OptiKinase in 1 x OptiKinase buffer (total volume of 25 µl), then incubated at 37°C for 1 hour. The reaction was terminated by heating at 65°C for 10 min. The probe was then purified using a Sephadex G-25 column. The resin in the column was resuspended by vortexing, the bottom closure opened and the column centrifuged for 1 min at 735g to remove the liquid. The column was put into a fresh 1.5 ml Eppendorf tube, the kinase reaction slowly added to the center of the resin and eluted by spinning at 735g for 2 min. The membrane was pre-hybridised by rotating for 1 hour at 37°C in Perfect Hyb Buffer. The labeled probe was denatured at 95°C for 5 min, added to 5 ml Perfect Hyb Buffer and hybridized to the membrane by rotating overnight at 37°C. The membrane was then washed three times with 20 ml of 2 x SSC containing 0.1% SDS for 20 min at 37°C, and the membrane exposed onto a Fujifilm phosphoimager cassette overnight. The Northern blot was visualized using the Typhoon™ FLA 9000 (GE Healthcare). If the membrane needed to be reprobed, it was incubated at 80°C with 0.2 x SSC and 0.1% SDS for 30 min.

2.2.7 Murine model of Transverse Aortic Constriction

This research was performed following the guidelines, and with the approval, of the Garvan Institute of Medical Research/St. Vincent's Animal Experimentation Ethics Committee (Project number 11/25). Male 11 week C57Bl/6 mice were obtained from the Animal Resources Centre, Adelaide, Australia. The mice were housed in the Victor Chang Cardiac Research Institute, Sydney.

2.2.7.a Time course of cardiac hypertrophy

Transverse Aortic Constriction (TAC) or a sham procedure (identical surgery except no constriction of the aorta) was performed on 12 week old male mice (*174*). Animals were anaesthetized with an itraperitoneal injection (IP) of xylazine (20 mg/kg), ketamine (75 mg/kg) and atropine (1.2 mg/kg). Once unconscious, as determined by loss of toe pinch reflex, the mice were intubated and ventilated with oxygen and 1.5-2% isoflurane for maintenance. The aorta was constricted between the right and left carotid arteries using a 27-gauge (27-G) needle and 7/0 silk suture. After surgery and the following morning, analgesia was provided with Buprenorphine (0.07 mg/kg) subcutaneously. After 48 hours or 7 days the ventricular systolic and diastolic function was assessed by micromanometry.

2.2.7.b Micromanometry

Micromanometry was used to assess cardiac function, including arterial blood pressure and heart rates as well as LV function (*175*). Anaesthesia was induced by placing the mice in a sealed chamber, with 5% isoflurane flowing and from which gas was scavenged. Once unconscious, the mice were intubated and ventilated with oxygen and 1.5-2% isoflurane for maintenance. A 1.0F micro-tip catheter (Millar Instruments) was inserted into the right carotid artery for arterial pressure, and then progressed into the LV for LV pressure. The micromanometry data was analysed using Acqknowledge software (Biopac systems Inc.).

2.2.7.c Isolation of cardiomyocytes

Cardiomyocytes were isolated from TAC and Sham mice according to a standard protocol (*176*). 500 ml of CM isolation perfusion buffer was prepared per heart; 50 ml of which was used to prepare the collagenase digestion buffer. The mouse was first injected with 100 U/ml heparin to prevent coagulation of blood in the coronary arteries. The mouse was

anaesthetized, the chest opened and the heart rapidly removed and mounted on a perfusion apparatus via a canula placed in the aorta just above the coronary. The heart was immediately perfused with the calcium-free perfusion buffer (4 ml/min at 37°C); blood rapidly clearing from the coronary arteries indicated correct cannulation and successful perfusion. After 4 min, the heart was perfused with collagenase digestion buffer (4 ml/min at 37°C), to digest the extracellular matrix, for 3 min. The collagenase digestion buffer was collected and discarded. After 2-3 min, the heart was digested for a further 8 min at 4ml/min (digestion times varied slightly from heart to heart) with collagenase digestion buffer containing 40 µM $CaCl_2$. A well-perfused heart became swollen, flaccid and slightly pale.

Once digestion was completed, the heart was cut just below the atria from the cannula (leaving the atria attached to the perfusion apparatus) and placed in a dish with 2.5 ml of collagenase digestion buffer containing 40 µM $CaCl_2$. The ventricles were teased into 10-12 small pieces and 5 ml collagenase stopping buffer containing 2 mM freshly added ATP at RT was added to the dish. The pieces were then pipetted gently with a plastic transfer pipette for 60-90 sec to further disperse cells. The cell suspension was transferred to a 15 ml tube; the plate was rinsed with 2.5 ml collagenase stopping buffer and the buffer also added to the tube (final volume of 10 ml containing 2mM freshly added ATP). The heart tissue was dissociated further until all the large pieces were dispersed in cell suspension (3-5 min). Unlike the protocol for short and long-term culture the calcium concentration was not increased at each wash step. At this point, 80 µl of cell suspension was aliquoted to a microcentrifuge tube, and duplicate 10 µl aliquots were used to count rod-shaped and round cardiomyocytes in a hemocytometer. The total number of cardiomyocytes and the number of rod-shaped cardiomyocytes were calculated and recorded as the initial cell numbers. If the total cardiomyocyte yield was low (<1 million) or the percent of rod-shaped myocytes was low (<50%), which indicates a less than optimal isolation, the cells were excluded from deep sequencing analysis.

Whilst counting, the remaining cardiomyocytes sedimented passively by gravity at RT. The cells were then centrifuged at 20g for 3 min, the supernatant collected as Wash 1, and the cells were resuspended in 10 ml collagenase stopping buffer containing 2 mM freshly added ATP (final calcium concentration 12.5 µM). Wash 1 supernatant was spun at 168g for 3 min and 1 ml TRIzol® was added to the pellet. The cardiomyocytes were washed a further 2 times (total of 3 washes) with 10 ml myocyte stopping buffer (final calcium

concentration 12.5 µM and 2 mM ATP). The rod-shaped and rounded cardiomyocytes were counted again using a hemacytometer (as above). The total number of cardiomyocytes, number of rod-shaped cardiomyocytes and percentage of rod-shaped cardiomyocytes were calculated and recorded as the final cell numbers. If the total cardiomyocyte number was low (total cells < 8×10^5) or cardiomyocyte purity poor (<60% rod-shaped myocytes), then the cells were excluded from deep sequencing analysis. The final cells were washed once in PBS, to remove residual BSA contained in the wash buffer, and spun at 42g for 1 min. 5 ml of TRIzol® per heart was then added to the cell pellet, the pellet was dissolved in TRIzol® using a vortex and split into 1 ml aliquots for RNA extraction.

2.2.7.D Dissection of LV and lungs from murine hearts

After micromanometry analysis, a cohort of mice were euthanized by cervical dislocation. The chest area was dissected to expose the heart. The mouse was perfused by inserting a needle into the apex of the LV, and flushing the blood out with 10 ml of PBS until clear fluid ran out. The major vessels were cut at the point they enter or exit the heart and the heart was excised from the chest. The heart was gently squeezed onto a surgical gauze to remove all remaining liquid and the heart weight recorded. The atria were then removed, and the RV free wall cut from its attachment to the LV leaving only the LV and the septum. The LV weight was then recorded, the LV sectioned into 3 separate samples and transferred into a 1.5 ml Eppendorf tube and submerged in liquid nitrogen for tissue preservation. The samples were then stored at -80°C until used for RNA extraction. The lungs were also removed by gently tugging on the trachea while snipping away the connective tissue and weighed to establish the lung wet weight.

2.2.8 Next-generation sequencing

2.2.8.A Small RNA sequencing

Libraries were created using from RNA using the NEBNext® Small RNA Library Prep Set for SOLiD™ as per manufacturer's instructions. An overview of the library preparation method is shown in Figure 2.2. In brief, 2 µg of total cardiomyocyte RNA was added to the initial ligation reaction together with the NEBNext® 5' and 3' Small RNA Adaptors. The PCR amplification was modified from the NEBNext® protocol to include the SOLiD® 5' PCR Primer and SOLiD® 3' Barcodes for each sample (Barcodes 2-9), and 13 cycles of

amplification were used. The PCR products were added to 5 µl Novex® TBE-Urea Sample Buffer and separated on a polyacrylamide gel (5% stacking gel, 8% resolving gel, SequaGel –UreaGel System) at 300 V for 45 min. Hyperladder-V was run as a size marker. The gel was stained with 1 X SYBR® Gold staining solution for 5 min, then visualized by medium wavelength UV light and photographed using the Typhoon™ FLA 9000 (GE Healthcare). PCR products approximately 110-120 nt in size were excised from the gel. The gel was dissolved by putting the gel slice containing the PCR product into a 0.2 ml Eppendorf tube that had been pierced with a 19-G needle, then spun at 14 000g into a 1.5 ml Eppendorf tube. 400 µL of PAGE elution buffer was added to the spun gel and rotated overnight at RT. The tube was briefly centrifuged and the supernatant transferred to a tube containing a Corning® Costar® Spin-X® Plastic Centrifuge Tube Filter, and centrifuged at 10 000g for 30 sec. The product was precipitated by adding 2 µl glycogen (5 mg/ml) and 1 ml 100% EtOH and storing at -80°C overnight. The samples were then spun at 20 000g at 4°C for 30 min and the supernatant discarded. The products were washed with 500 µl 70% EtOH and centrifuged at 7 000g for 5 min at 4°C. The supernatant was discarded, the pellet dried and then resuspended in 10 µl Low TE Buffer. The size and concentration of the PCR products was confirmed by running 1 µl of the PCR product on an Agilent High Sensitivity DNA Chip. The products were diluted to similar concentrations (300 pM) and prepared for sequencing using Emulsion PCR (e-PCR) on the SOLiD® EZ Bead™ Emulsifier according to manufacturer's instructions. The libraries were then sequenced using the SOLiD™ 5500 platform.

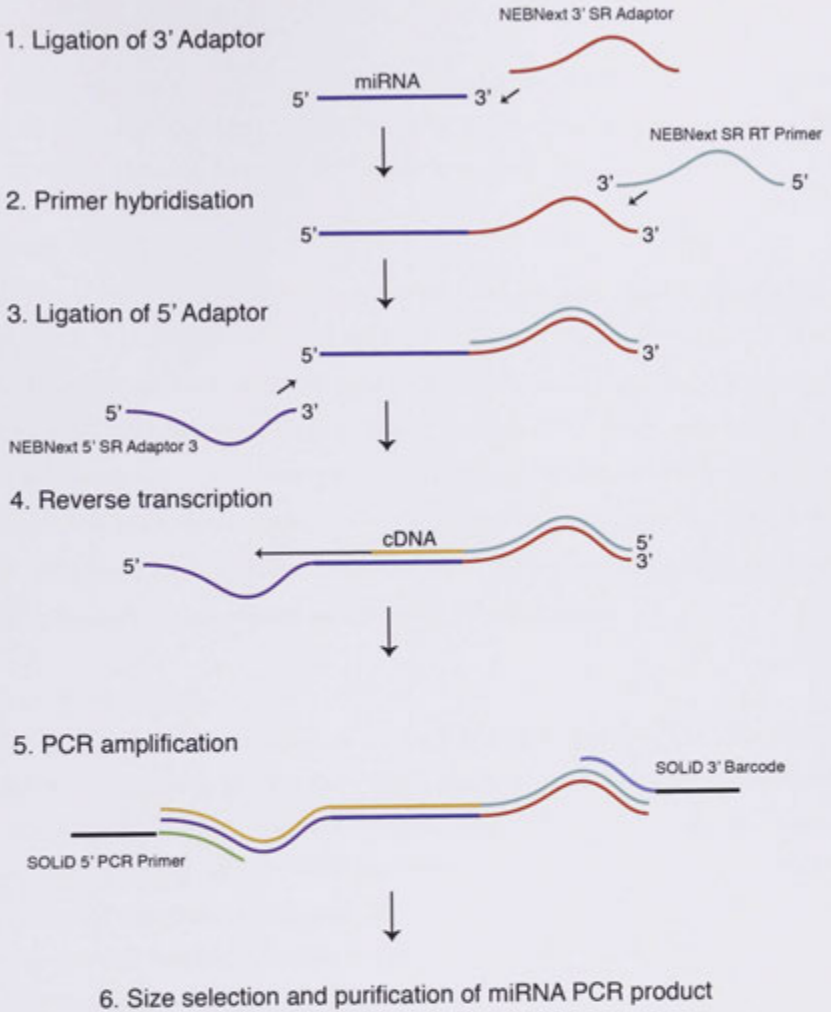


Figure 2.2. Small RNA library preparation method.

The NEBNext® Small RNA Library Prep Set for SOLiD™ was used for small RNA sequencing. Step 1: The NEBNext® 3' adaptor is ligated to the 3' end of the miRNA. Step 2: The NEBNext® ST primer is annealed to the complementary sequence on the 3' adaptor. Step 3: The NEBNext® 5' SR Adaptor 3 is ligated onto the 5' end of the miRNA. Step 4: The SR RT primer is used to reverse transcribe the miRNA and 5' adaptor. Step 5: The sequences within the 5' and 3' adaptors are used to prime the SOLiD® 3' Barcode and 5' PCR Primer for PCR amplification. Picture modified from NEBNext® Multiplex Small RNA Library Prep Set for SOLiD™ manual.

2.2.8.B 3' END SEQUENCING

The cardiomyocyte RNA was sent to Dr Traude Beilharz and the libraries were created using the PAT-Seq method developed by Dr Beilharz at Monash University (*177*). An outline of the library preparation method is shown in Figure 2.3.

End Extension:

2 µg of RNA in 11 µL ddH_20 was added to 1 µl of 100 µM Biotin labeled oligo (Bio-EE oligo; Biotin-5'-CTGCTGTACGGCCAAGGCGTTTTTTTTTTTTT-3'), heated at 80°C for 5 min and briefly centrifuged. The heated RNA was then added to 1 µl 0.1 M DTT, 40 U RNaseOUT, 1 µl 10mM dNTPs, 5 U Klenow polymerase Exo in 1X SuperScript™ III First-strand buffer at 37°C. The reaction was mixed, briefly centrifuged and transferred to a preheated PCR machine to incubate at 37°C for 1 hour, followed by 80°C for 5 min. While still at 80°C, 80 µl digest solution (10 µl 10 x digest buffer, 1 µg carrier RNA and 70 µl ddH_20) was added and then incubated at 80°C for a further 10 min.

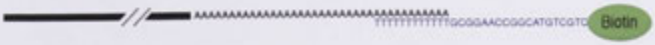
Limited RNase T_1 digestion:

The end-extended RNA was cooled on ice for 5 min, then added to a pre-cooled Phase Lock Gel tube containing 100 µl cold 10 x digest buffer and 10U RNase T1 and incubated on ice for 1 min. Then, the tube was vortexed briefly, returned to ice and the reaction terminated by adding 200 µl 50:50 phenol:chloroform. The reactions were shaken vigorously and centrifuged at top speed for 5 min, then extracted once with chloroform. The top layer was transferred to a new 1.5 ml tube containing 100 µl 5mg/ml streptavidin magnetic beads prewashed in 1 x binding buffer. The DNA:RNA duplexes were allowed to bind the beads for 15 min at 37°C with shaking. The beads were concentrated using a magnetic rack (MagnaRack™, Invitrogen) and washed once with 300 µl 1 x RT buffer at 55°C. The beads were concentrated again and then washed once with 100 µl 1 x T4 DNA ligase buffer (containing ATP) at 37°C with shaking. The beads were concentrated a third time and resuspended in 20 µl 1 x T4 DNA ligase buffer.

Addition of 5' Phosphate to fragments:

10 U Polynucleotide Kinase (PNK) was added to the 20 µl RNase T1 fragments, and incubated at 37°C for 30 min, followed by heat inactivated at 70°C for 10 min. Once

1. Anneal oligo d(T) anchor

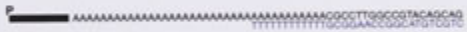


2. End extension with Klenow Polymerase

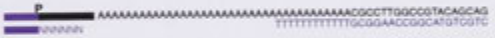
3. Limited RNase T1 digestion

4. Pull down 3' end Fragments

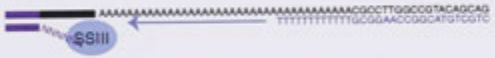
5. 5' phosphorylate fragments with PNK



6. Ligate 5' SOLID linker



7. Reverse Transcription



8. Size selection of cDNA

9. PCR amplification

Figure 2.3. 3' sequencing library preparation method.

Steps 1-2: Polyadenylated RNA is extended by generating a 3' tag on polyadenylated RNA, ensuring that reverse transcription is only possible from true 3' ends. Step 3: The RNA is fragmented with RNase T1, which cuts after G residues. A limited digestion ensures all poly(A)-tails are left intact. Step 4: The 3' end fragments containing a poly(A) tail are pulled down using Streptavidin beads. Step 5: The RNase T1 reaction generates a 5' hydroxyl group, which is not compatible with ligating the 5' SOLiD linker. A 5' phosphate is added. Step 6: The 5' SOLiD linker is ligated onto the 3' end fragment, which is necessary for reverse transcription. Step 7: The 3' tag is used with SSIII to reverse transcribe the full poly(A) sequence and remaining 3' end fragment. Step 8: The cDNA is run on a 6% urea polyacrylamide gel and bands of 70-300nt purified. Step 9: The sequences within the 5' SOLiD linker and the 3' tag are used to amplify the cDNA product.

cooled to room temperature, 200 µl of 1 x RT buffer was added to the beads and then all liquid removed.

Ligation of 5' SOLiD linker:

The ligation of the 5' SOLiD adaptor is based on the SOLiD™ Total RNA-Seq kit. The beads were then added to 2 µl 5' Adaptor mix, 3 µl Hyb solution and 2 µl ddH_2O on ice, then incubated at 65°C for 10 min and 16°C for 5 min. 10 µl 2 x ligase buffer and 2 µl RNA ligase 2 were added, and the mix incubated at 16°C for 16 hours.

Reverse transcription:

To reverse transcribe the fragment containing the 5' linker, the ligation product was added to 2 µl 0.1M DTT, 40 U RNaseOUT, 2 µl 10 mM dNTPs, in 1X SuperScript™ III First-strand buffer (total volume 39 µl), then incubated at 70°C for 5 min, briefly centrifuged, incubated at 55°C for 1 min before adding 200 U Superscript III. The reaction was mixed by rapid inversion, briefly centrifuged and incubated at 55°C for a further 60 min.

Size selection of cDNA:

The cDNA products were size selected using PAGE. A 6% urea-polyacrylamide gel was pre-run at 180 V for 30 min, and the wells flushed carefully with a syringe. The products were added to 2 x RNA loading dye, heated at 80°C for 5 min and then loaded onto the gel, along with a 100 nt RNA ladder for size comparison. The gel was run at 180 V until the bromophenol blue was 2 cm from the bottom of the gel. The gel was stained with 1 x SYBR® Gold staining solution for 5 min, then visualised by medium wavelength UV light. The bands from 70-300 nt (50 nt sequence, up to 18 nt 3' end extension and poly(A)) were excised. The gel was dissolved by putting the gel containing PCR product into a 0.5 ml Eppendorf tube that had been pierced with a 21-G needle, then spinning at 13 000g into a 2 ml Eppendorf tube. 300 µl PAGE elution buffer was added to the shredded material, and incubated overnight with gentle agitation at 4°C. The buffer was transferred to a new tube, and another 300 µl PAGE elution buffer added and incubated at 37°C for 1-2 hours with gentle agitation. The eluted material and remaining gel were put into an Eppendorf tube containing a Corning® Costar® Spin-X® Plastic Centrifuge Tube Filter, and centrifuged at 10 000g for 30 sec. The product was precipitated by adding 2 µl glycogen (5 mg/ml) and

700 µl isopropanol and centrifuging at 13 000g for 20 min. The pellet was then washed once with 75% EtOH, the supernatant removed and the pellet dried.

PCR amplification:

The size-selected cDNA was then PCR amplified by adding 8 µl of the cDNA to 1 µl 100 µM SOLiD seq primer, 1 µl barcoded reverse primer and 50 µl 2 x Amplitaq 360 master mix in a total volume of 100 µl, and 15 cycles of amplification. The size and concentration of the PCR products was confirmed by running 10 µl of the PCR product on a Urea-polyacrylamide gel. Emulsion PCR (e-PCR) was then performed on the SOLiD® EZ Bead™ Emulsifier according to manufacturer's instructions. The libraries were then sequenced using the SOLiD™ 5500 platform.

2.2.8.C Small RNA sequence analysis pipeline

An individual deep sequence read is referred to as a tag and the number of times it occurs as a count.

Data filtering:

Prior to mapping a custom Perl script was used to trim adaptors that are ligated at the end of the miRNA from the sequence tags and remove all tags with an average quality value of less than 18 across the tag (Appendix 8.1 - adaptortrim.pl). Each colour-call in SOLiD™ chemistry is associated with a quality value ranging from -1 to 33; -1 being the lowest and 33 being the highest quality. These quality values (q) reflect the probability of error (p) according to following equation.

$$q = -10\log_{10}(p)$$

All tags with a mean quality value <18 (~98% accuracy) and tags containing an ambiguous color call, represented by ".", were removed to retain high quality tags for further analysis. Tags were retained only if the length was at least 20 nt after trimming, corresponding to the length of a mature miRNA

Mapping to the reference genome:

The remaining high quality tags were then mapped to the male mouse genome reference (assembly version mm9, 18S rRNA gi|374088232 and 28S rRNA gi|120444900 rRNA) by using Bowtie software (*178*), version 0.12.8, that supports colour-space data mapping (Appendix 8.1 - miRNA.bowtie.SGE.pl). Mapping parameters were chosen such that up to

2 colour-space mismatches were allowed in the first 20 nt seed region, and the alignment was extended to the full-length of the tag if the sum of the quality values at mismatch positions did not exceed 70.

Mapping protocol for total genome analysis:

Mapping to the genome. Tags were mapped such that one random mapping position was selected for tags mapping to more than 1 locations by invoking the –M 1 option in Bowtie command. This set of alignments is referred to as *genomic alignments* from here on, and was used for counting tag distribution in various genomic features.

Bowtie Command, Genomic alignments:

```
bowtie -f -C –Q input.QV.qual --integer-quals -l 20 --nomaqround --
maxbts 800 --tryhard --chunkmbs 2048 -M 1 -a --best --strata --snpfrac
0.01 --col-cqual --col-keepends --sam --mapq 20 --offrate 2 --threads
12 --shmem reference_bowtie_colorspace_index input.csfasta >output.sam
```

Tag counting for distributions of various genomic features:

The genomic alignments were used to determine the proportion of tags mapping to various genomic features. Coordinates for genomic features such as protein coding loci, non-coding RNA loci and genomic repeats (e.g. LINE, SINE, simple repeats) were obtained from Ensembl v67. In addition, miRBase v18 annotations were used for miRNA loci, piRNA coordinates were obtained from the piRNABank (*179*) and protract (*180*), and 18/28S rRNA were obtained from the GenBank. A custom BAM file parser was designed (Appendix 8.1 - countGff) by using the BamTools API (*181*) to identify the number of tags that overlap each genomic feature. Tags were preferentially assigned to a miRNA loci, even if multiple, overlapping genomic features were present for that locus. If the mapped position of the tag overlapped a non-miRNA locus with multiple, overlapping genomic features, the tag was randomly assigned to one genomic feature.

Mapping protocol for miRNA-focused analysis:

Mature miRNAs derived from multiple genomic loci can be identical in sequence. Therefore randomly assigning tags to all of the identical loci can effectively reduce the number of tags assigned to each individual locus. This random assignment can affect the downstream expression analysis. Therefore, we chose to assign tags to all of the identical loci to retain the true levels of tag counts observed for any mature miRNA in our data. To

assign the number of tags at each miRNA locus, tags were allowed to map to up to 20 locations. If a tag mapped equally to more than 20 locations, one location was randomly chosen. These set of alignments are referred to as *miRNA alignments* from here on.

Bowtie Command, *miRNA alignments*:

```
bowtie -f -C –Q input.QV.qual --integer-quals -l 20 --nomaqround --
maxbts 800 --tryhard --chunkmbs 2048 -M 20 -a --best --strata --
snpfrac 0.01 --col-cqual --col-keepends --sam --mapq 20 --offrate 2 --
threads 12 --shmem reference_bowtie_colorspace_index input.csfasta
>output.sam
```

Tag counting for miRNA loci:

Tags were assigned to a mature miRNA if their mapped position was within +/-3nt of the miRBase annotated 5' start site and were between 20-26nt in length inclusive (*82*). These mature miRNA tags are referred to as *miRNA tags* and the count of these tags is referred to as the *total tag count for a miRNA* from here on.

Tag count information tables were generated by using miRNAcountsBAM.pl (Appendix 8.1). These tables were then loaded into MySQL database. Using miRNAcounts.mysql.pl (Appendix 8.1), which used the mmu.features (Appendix 8.1) and MySQL database as input, final counts tables were generated.

Many mature miRNAs are derived from multiple genomic locations but the mature miRNA sequence is identical. These miRNAs have the same miRNA name followed by a suffix, for example miR-199a-1 and miR-199a-2. Owing to the same nucleotide and seed sequence, it is assumed that they act similarly in mRNA regulation pathway in a given biological context. Therefore, we randomly chose one mature miRNA in such instances for analysis of miRNA.

Differential expression analysis:

The total tag count for a miRNA in each sample was obtained for differential expression analysis as described above. Effective library size for each sample was set to the sum of tag counts mapping to miRNAs in each sample. The edgeR package was used for the following differential expression analysis (*182*). Tag counts were normalised for RNA output by the trimmed mean of M-value (TMM) scaling method described in (*183*) to obtain normalised

CPM (nCPM). All miRNA with mean nCPM across all samples below 10 were discarded from further analysis. Tests for the differential expression of miRNA between samples were carried out as described in the edgeR user guide. The script used for all miRNA differential expression and associated analysis is listed in Appendix 8.2. The data was normalised separately using upper quartile normalization, the same script was used as listed above but the function calcNormFactors (x, method="upperquartile") was used.

miRNA processing analysis:

miRNAs can exist as 5' and 3' isomiRs, exhibit arm bias and undergo post-transcriptional modifications. In light of these properties, we analysed the sequencing data for such post-transcriptional events to better understand these processes in our model. We focused on *a)* hairpin arm selection *b)* 5' isomiR *c)* 3' isomiR *d)* internal editing and *e)* non-templated additions. Distinct genomic regions can give rise to canonical mature miRNAs that are identical in sequence. Nevertheless, difference in precursor sequence at 5' upstream or 3' downstream positions may lead to their differential processing for isomiR generation, arm bias, internal editing, and non-templated additions. Therefore, to systematically assay these processing events, each genomic position of a mature miRNA was analyzed individually.

Hairpin arm selection:

The ratio between the total tag count for a miRNA (as described earlier) on the 5' arm of a given miRNA hairpin *and* the sum of the total tag counts for miRNAs on both 5' and 3' arms of the hairpin was calculated for this analysis.

5' and 3' isomiR ratio:

miRBase v18 annotated start and end positions for a mature miRNA were considered as the canonical start and canonical end positions, respectively. miRNA tags differing in their 5' start positions compared to the canonical start position were considered as 5' isomiRs and those that differed at the 3' end positions compared to the canonical end position were considered as 3' isomiRs. The percentage of canonical miRNA tags for each mature miRNA in each sample was calculated with respect to total tags for that miRNA,

Internal editing:

miRNA tags with a nucleotide mismatch before the last two nucleotides of the tag and before the last two nucleotides of the canonical end position were considered as tags containing possible editing events. For each mature miRNA the number of tags with possible editing event and the identity of the base substitution was determined. The mean percentage and tange of internally edited miRNAs was then calculated.

Non-templated addition (NTA):

If miRNA tags (as described earlier) differed from the reference sequence in the last two nucleotides of the tag, then they were selected to contain non-templated nucleotide addition. For each mature miRNA the number of tags with NTA was recorded as well as the identity of the non-templated base. The percentage of NTA tags was then determined relative to the total miRNA tags for a given mature miRNA. The mean percentage of NTA was then calculated along with the range for all miRNAs.

Statistical model of miRNA processing variants:

Differential processing of miRNA variants was determined using the proportion a type of processing variant contributes to each miRNA. The data was transformed to an approximate Guassian distribution by an arc sin of square root transformation. Count noise was then incorporated into the model by quadratic fit of variance versus total read counts for each miRNA, using the Voom method of Smyth et al. (*184*). A weighted linear model was then used for differential processing using Limma (*184*). The FDR was estimated using the Benjami-Hockberg method.

GO-term enrichment for miR-133a-3p predicted targets:

Target prediction was done using the can/23 nt and iso/22 nt seed sequences of miR-133a and Targetscan (*67*). Enrichment of gene function terms was determined with IPA software (Ingenuity® Systems), using the Benjamini-Hochberg multiple testing correction method and a threshold p-value <0.01.

2.2.8.D 3' END SEQUENCE ANALYSIS PIPELINE

Data filtering:

Prior to mapping to a reference, the tags were filtered to retain only high quality reads; any tags with a mean quality value <18 were removed and tags with ambiguous colour calls were also removed using the perl script csqvfilter.pl (Appendix 8.3).

Mapping to the reference genome:

All clean sequenced tags were then mapped to the mouse genome reference (assembly version mm9, 18S rRNA gi|374088232 and 28S rRNA gi|120444900 rRNA) with Bowtie Software (*178*), version 1.0.0 that supports colour-space mapping. Mapping parameters were chosen to allow 3 mismatches in the 21 nt seed region, and the alignment was

extended to the full-length if the sum of the quality values at mismatch position was less than 150 (parameter --maqerr 150). Tags were allowed to map to up to 20 locations.

Bowtie command, *Genome alignments:*

```
bowtie -f -C --integer-quals --seedmms 3 --nomaqround --maxbts 800 --
tryhard --chunkmbs 512 -M 20 --best --strata --maqerr 150 --col-cqual
--col-keepends --sam --threads 12 --mm -Q filtered.QV.qual --un
GNeg.csfasta mouse.e65chr.ncbirRNA.fa filtered.csfasta
```

To capture exon junction-spanning reads, tags that failed to align to the genome were further mapped to the mouse cDNA and ncRNA sequences (Ensembl v67) using the same Bowtie parameters with the exception of allowing tags to match to up to 200 locations.

Bowtie command, *Transcriptome alignments:*

```
bowtie -f -C --integer-quals --seedmms 3 --nomaqround --maxbts 800 --
tryhard --chunkmbs 512 -M 200 --best --strata --maqerr 150 --col-cqual
--col-keepends --sam --threads 12 --mm -Q GNeg.csfasta.qual --un
TNeg.csfasta Mus_musculus.NCBIM37.67.cdna.ncrna.fa GNeg.csfasta
```

Read alignments to RNA sequences were projected to genomic locations and merged with genomic alignments using transcript2genome.pl (Appendix 8.3). All tags that did not map to either genome or expressed RNA sequences were mapped to the genome by trimming the last 54 nt (retaining the first 21 nt seed) using the Bowtie software, allowing up to 3 mismatches in the first 21 nt. Up to 200 locations for each tag with mismatches ranging between 0 and 3 were identified and processed further to identify non-templated poly(A) tail additions on reads.

Bowtie command, *Seed alignments:*

```
bowtie -f -C --integer-quals -v 3 --trim3 54 -M 200 --col-cqual --col-
keepends --sam --threads 12 --mm -Q TNeg.csfasta.qual --un
SNeg.csfasta mouse.e65chr.ncbirRNA.fa TNeg.csfasta
```

A custom Perl script (Appendix 8.3 - cspolyaextender.pl) was used to perform extension of alignments such that the mismatch proportion remained less than 10% in either the templated region of the read or non-templated region of the read. All alignments were merged into a single bam file for each sample and processed further to identify 3' sequencing peaks and poly(A) tail start positions.

Defining 3' sequencing peaks:

Contiguous stretches of nt along the length of the genome with at least 1 read coverage were binned into peaks representative of 3' end sequencing of RNA using the script coveragePeaks.pl. Tags linking two or more exons were identified using the following command; `intersectBed -bed -wo -s -split -abam T2G.merged.bam -b ThreeUTRPeaks.6.gff >ThreeUTRPeaks.6.intersect.` The script linkCoveragePeaks.pl (Appendix 8.3) was run to turn links between 3' seq peaks into transcript units.

Information for each peak such as the length of the transcript, the exon IDs that make up the transcript, the number of reads that contribute to the peak, presence of poly(A) and poly(A) signal were defined using the script getTranscripts.pl (Appendix 8.3). For a poly(A) cleavage site to be called, the start position for the cleavage site had to be supported by at least two reads with different start positions. The peaks were filtered by retaining only peaks longer than 75 nt and with tags present in at least two samples. Any downstream filtering was done in R.

Counting peaks for distributions of various genomic features:

Genomic features were identified using the canonical transcript as defined in Ensembl v67 using processTable.pl. The peaks were labeled as Annotated, Orphan, Exonic or Intronic based on the following criteria:

Annotated peaks:

- If the peak overlaps with a known gene and it is within 20 kb downstream of the end/stop codon of a known gene then it is assigned as Annotated.
- If the peak does not overlap with a known gene but it is within 20 kb downstream of an upstream end/stop codon of a known gene then it is assigned as Annotated.
- If the peak overlaps with the intron of a known gene but in the antisense direction, and it is within 20 kb of an upstream end/stop codon of a known gene in the sense direction then it is assigned as Annotated to the gene in the sense direction.

Orphan peaks:

- If the peak does not overlap with a known gene and it is not within 20 kb of an upstream end/stop codon then it is assigned as an Orphan peak.

Exon peaks:

- If the peak overlaps with the exon of a known gene then it is assigned as an Exon peak.

Intronic peaks:

- If the peak overlaps with the sense strand of an intron of a known gene then it is assigned as an Intronic_sense peak.
- If the peak overlaps with the antisense strand of an intron of a known gene and is not within 20kb of an upstream end/stop codon then it is assigned as an Intronic_antisense peak.

Calculating distances to stop codon:

These were calculated from the stop codon of the canonical transcript of a protein coding gene or from the 3' end of the canonical transcript of a non-coding gene (as defined in Ensembl v67).

Calculating novel or known 3' ends:

All transcript isoforms defined in Ensembl v67 were used to determine if a peak was novel or known. If a peak overlapped a known 3' UTR and was within 100 nt from the end of the 3' UTR then it was considered as known.

Features of the 3' sequencing peaks:

- Searching for the canonical and non-canonical polyadenylation singals (PAS).
    - These were searched 50 nt upstream of the poly(A) start site with at least 2 reads supporting the start site. Where the poly(A) tail was not present, the nearest PAS was identified to the 3' end of the peak (entire peak upstream and 50 nt downstream). Both the canonical and non-canonical PAS were recorded if present, then a preference was given to the canonical PAS. The distance from the start of the poly(A) tail or the end of the 3' seq peak was calculated.

2.2.8.e Merging 3' seq peaks for gene expression

The tag counts for 3' seq peaks assigned to the same gene were added together for differential expression analysis. The effective library size for each sample was set to the sum of tag counts assigned to genes. The edgeR package was used for the following

differential expression analysis (*182*). Tag counts were normalised for RNA output by the trimmed mean of M-value (TMM) scaling method described in (*183*) to obtain normalised CPM (nCPM). The 98th percentile was retained for differential expression testing. Tests for the differential expression of miRNA between samples were carried out as described in the edgeR user guide. The script used for all mRNA differential expression and associated analysis is listed in Appendix 8.4.

GO-term enrichment of differentially expressed genes:

Gene ontology enrichment was performed using the database for annotation, visualization, and integrated discovery (DAVID)(*185*). The analysis was performed in R, using the script GOterm.pl. The enrichment was performed using differentially expressed mRNA transcripts (Appendix 8.5) with a FDR <10% and the top 98th percentile as the background. Enriched GO_FAT terms with FDR < 30% for biological process (GO_BP_FAT), cellular component (GO_CC_FAT) and molecular function (GO_MF_FAT) were recorded.

2.2.8.F Detection of 3'UTR changes:

Using protein coding 3' UTR associated peaks with coverage >10 and tags in at least two samples.

Noise reduction filter:

A peak needed to contribute greater than 10% of total tags to a gene across all samples to be considered for analysis. An additional filter requiring ≥ 4 reads across TAC and Sham at each time point.

Differential poly(A) model:

Used a model that eliminates changes that are confounded by overall gene expression level changes. Alternative polyadenylation was detected as a proportion statistic between proximal and distal poly(A) sites. All possible pairs of 3' seq peaks for a given transcript were used. The data was first transformed to an approximate Gaussian distribution by an arc sin of square root transformation. Count noise was then incorporated into the model by quadratic fit of variance versus total read counts over each peak, using the Voom method of (*184*). A weighted moderate linear model was used for differential 3' end usage

estimation using Limma(*184*). The FDR was then estimated using the Benjamini-Hockberg method.

GO-term enrichment of 3'UTRs with differential poly(A):

GO-term enrichment was performed as described in 2.2.7.E but using the mRNA transcripts with alterations in the proportion of 3' UTRs (FDR <70%).

2.2.8.G MDS PLOTS AND HEATMAPS

Multidimensional scaling plots were drawn using the plotMDS function in the edgeR package. The plotMDS function takes the top 500 genes with the largest biological variation between libraries (largest tagwise dispersion). Then the distance between each pair of libraries, termed the biological coefficient of variation is calculated. The spearman correlation was calculated using the function cor and the method "spearman", which refers to the Spearman's *rho* or rank statistic.

Heatmaps were drawn using the heatmap.2 function in the gplots package. Heatmap.2 performs complete linkage clustering, using a Euclidean distance measure. The dendrogram is then reordered using the row or column means.

Chapter three

Hypertrophy model

3.1 Introduction

To examine the relevance of miRNA and 3' UTR variants in cardiac disease, a model of cardiac hypertrophy was used in collaboration with Professor Robert Graham's laboratory at the Victor Chang Cardiac Research Institute.

Transverse Aortic Constriction (TAC) is a commonly used and reproducible mouse model of pressure overload-induced left ventricular hypertrophy and heart failure (*186*). TAC closely mimics the clinical scenario in patients with aortic coarctation (a congenital condition resulting in narrowing of the aorta) and somewhat resembles aortic stenosis or hypertension (*187*). TAC is a microsurgical technique that involves tying off the aorta with a suture against a needle to yield a consistent constriction of the transverse aorta (Figure 3.1). It has been previously shown that in 8-week-old C57BL/6 mice the diameter of the aorta at the site of constriction is 1.2 mm, and aortic constriction with a 7-0 nylon suture against a 27-G needle results in a 0.4 mm reduction in diameter once the needle is removed (*186*). This reduction in the diameter of the aorta increases the pressure that the heart has to push against. The increased pressure can be offset by an increase in LV wall thickness, an adaptive response that results from an increase in the size of the cardiomyocytes.

TAC is initially a model of compensated hypertrophy, where the heart maintains cardiac output, but the response to chronic pressure overload can eventually become maladaptive and result in heart failure. The progression to heart failure can be accelerated depending on the tightness of the aortic constriction, with severe TAC leading to impaired LV function as early as 7 days (*188*). For this reason, we subjected 12-week-old C57BL/6 mice to mild TAC (using a 27-G needle) and analysed the hypertrophic response at 7 days where there is a reported change in cardiac structure but the cardiac function remains unaffected (*188*). To analyse pre-hypertrophic changes, animals were also analysed 48 hours post-TAC surgery. Increased load on the heart results in the parallel addition of sarcomeres in the cardiomyocyte, resulting in an increase in cardiomyocyte width and subsequent increase in LV wall thickness (*3*). It has been previously shown that there is no significant difference in LV wall thickness, and thus no hypertrophy, at 24 hours post-TAC but the LV wall thickness progressively increases until 10 days post-TAC surgery (*189*). Therefore, it was hypothesised that changes at 48 hours may be causative rather than a downstream result of hypertrophic signalling.

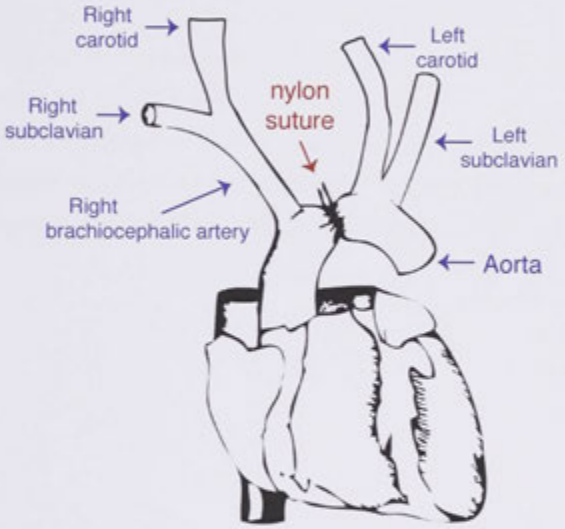


Figure 3.1. Transverse Aortic Constriction.

Schematic of TAC surgical procedure; the aorta is tied off with a nylon suture around a needle that is placed parallel to the aorta, between the left carotid artery and the right brachiocephalic artery. Picture modified from (*186*).

Previous studies looking at miRNA and mRNA expression in the hypertrophic/failing heart have used whole LV samples (*130, 190*). The heart consists of multiple cell types including cardiomyocytes, cardiac fibroblasts and vascular cells such as the endothelium. Cardiac fibroblasts account for 60-70% of the total number of cardiac cells, and are essential in generating and maintaining the extracellular matrix of the myocardium (*191*). Although cardiomyocytes are fewer in number, they take up the bulk of the volume of the heart and are the contractile cells that provide mechanical force to the heart (*192*). Both cardiomyocytes and fibroblasts are intrinsically involved in the progression to maladaptive hypertrophy and heart failure. Studies looking at whole LV samples have therefore been looking at the gene expression changes in a number of cell types. We hypothesised that to correctly establish miRNA interactions involved in hypertrophy it would be important to isolate cardiomyocytes. For this reason, once the animals had undergone haemodynamic studies using micromanometery thereby confirming successful pressure overload, the cardiomyocytes were enriched for subsequent RNA analysis. This would allow detection of miRNA changes and mRNA 3' UTR usage in cardiomyocytes independent of the other cell types.

3.2 Results

3.2.1 Establishing a model of LV hypertrophy

Tight TAC (constriction of the aorta to the size of a 27-G needle) was performed on 12-week C57BL/6 mice and the hypertrophic response analysed at 48 hours and 7 days post surgery. 64 mice were used (32 Sham and 32 TAC), with 44 mice used for isolation of the whole LV and other tissues and 20 mice used for cardiomyocyte enrichment. An overview of the experimental procedure to obtain samples for SOLiD™ next-generation sequencing is shown in Figure 3.2. The animal surgeries and micromanometry were performed by Dr Jianxin Wu. Dr Nicola Smith and myself assisted in the pre- and post-operative care of the animals, performed the dissections after micromanometry and data analysis.

Micromanometry using a micromanometer catheter placed in the aorta and then the LV cavity was performed on all animals to confirm differential pressure across the aortic band and to assess the changes in cardiac function. As cardiac function is closely correlated with heart rate (*193*), the micromanometry measurements were taken when the heart rates for both Sham and TAC groups were comparable (Figure 3.3.A). Both aortic and LV systolic blood pressures (BP) were significantly increased in TAC banded animals at 48 hours and 7 days (Figure 3.3.B). Aortic systolic pressure is the maximal pressure in the aorta following ejection of blood from the left ventricle. The aortic systolic BP was increased from 116.8±2.9 mmHg in Sham animals to 155.9±3.3 mmHg ($p<0.0001$) in TAC animals at 48 hours, with a further increase to 176.1±6.3 mmHg ($p<0.0001$) in TAC animals at 7 days (Figure 3.3.B). This indicates that the aortic constriction was indeed increasing the pressure that the LV must push against. Increased pressure at a point in a system causes a pressure backup to the preceding elements within the system. As the aortic pressure rises, the ventricular pressure will soon rise due to pressure backing up within the system. The increased pressure was reflected by an increase in LV systolic BP at 48 hours from 113.9±1.8 mmHg in the Sham animals to 149.2±4.3 mmHg ($p<0.001$) in the TAC banded animal. Similarly, the LV systolic BP also increased from 110.9±2.3 to 166.5±9.0 mmHg ($p<0.0001$) 7 days post-surgery. LV systolic BP is an indication of the total load on the heart, as it is a function of the pressure gradient from the TAC in addition to the systemic BP (*194*). The increased pressure occurs immediately after constricting the aorta, with increases in systolic pressure previously detected as early as one hour post-surgery (*195*).

The increased pressure indicates the TAC banding successfully decreased the diameter of the aorta resulting in an increase in pressure that the heart must pump against.

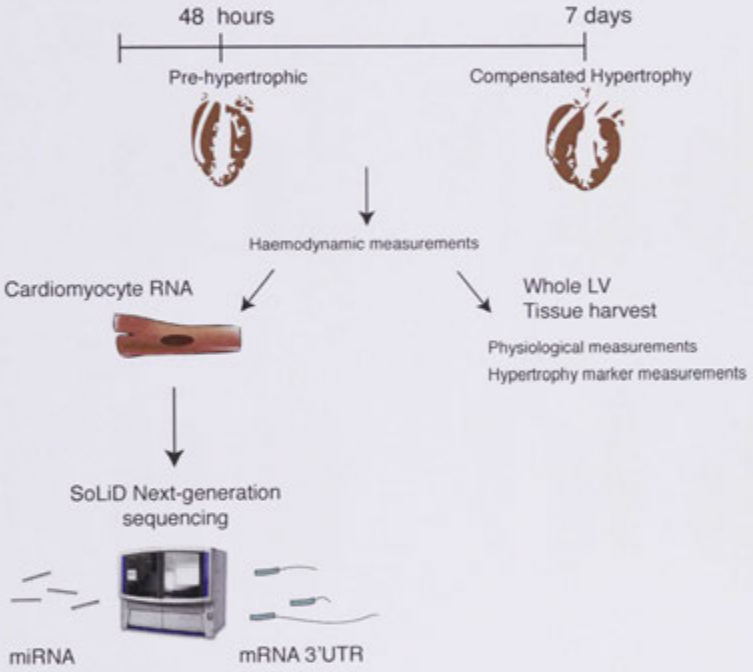


Figure 3.2. TAC: a model of left ventricular (LV) hypertrophy.

Overview of experimental approach: TAC was performed on 12 week C57BL/6 mice with a 27-G needle and hearts isolated 48 hours or 7 days post-surgery. Haemodynamics measurements were taken using micromanometery on all animals (n=64). Physiological and biochemical measurements were taken from tissues and whole LVs harvested from a subset of mice (n=44). The remaining 20 mice were used for cardiomyocyte enrichment and subsequent SOLiD™ next-generation sequencing of small RNAs and 3' end whole transcriptome.

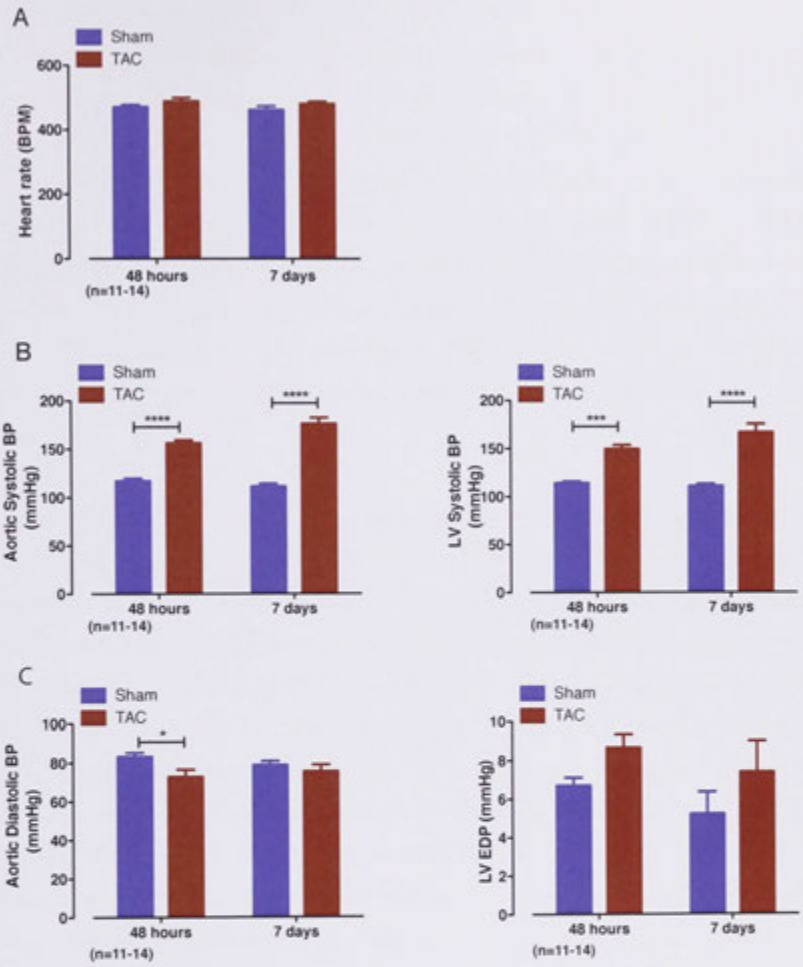


Figure 3.3. TAC leads to pressure overload.

Haemodynamic measurements were taken by micromanometry. (A) Heart rate (beats per minute, BPM) is comparable between Sham and TAC animals. (B) TAC increases both aortic and LV systolic blood pressure (BP). (C) TAC leads to a slight decrease in Aortic Diastolic BP at 48 hours, which normalised by 7 days. There was no significant difference in LV end diastolic pressure between Sham and TAC. (**** $p < 0.0001$. *** $p < 0.001$, * $p < 0.05$, Two way-ANOVA with Bonferonni's post-hoc analysis). Data is shown as mean with SEM (n=11-14).

Diastolic pressure, which is the lowest pressure in the aorta just before the ventricle ejects blood, was also measured (Figure 3.3.C). There was a slight decrease in aortic diastolic BP from 83.3±2.1 to 72.9±3.6 mmHg ($p<0.05$) at 48 hours, which normalised at 7 days (78.8±2.1 mmHg in Sham animals versus 75.5±3.5 mmHg in TAC animals). There was no significant change in LV end diastolic pressure (EDP), which is the pressure at the end of diastole just prior to the heart's next contraction, at either time point; 6.7±0.4 and 8.7±0.7 mmHg for 48 hour Sham and TAC respectively, and 5.2±1.2 and 7.4±1.6 mmHg for 7 day Sham and TAC respectively. LVEDP is a measure of ventricular performance, with increasing LVEDP associated with heart failure (*196*). Thus, as there is not a significant increase in LVEDP in the TAC banded animals there is no indication of cardiac dysfunction.

To further determine if the TAC banding negatively impacted on heart function, the dP/dt_{max} and dP/dt_{min} were calculated (Figure 3.4). The dP/dt_{max} and dP/dt_{min} represent the maximum and minimum rate of pressure change in the LV, with dP/dt_{max} an indicator of systolic function and dP/dt_{min} a measure of diastolic function. These values are commonly used as an index of LV function and contractility (*197*). There was a significant decrease in dP/dt_{max} from 11099.8±760.9 mmHg/sec in Sham animals to 8790.6±372.6 mmHg/sec ($P<0.05$) in TAC animals at 48 hours (Figure 3.4.A). There was no significant difference in dP/dt_{max} at 7 days; 9458.9±320.0 and 9757.6±710.4 mmHg/sec for 7 day Sham and TAC respectively. This suggests the LV systolic function deteriorated soon after TAC but normalised by day 7. This is consistent with previous studies that show LV systolic function decreases immediately after TAC, but recovers with the development of compensated hypertrophy (*189*). Reflecting the slight decrease in Aortic diastolic BP, there was a trend for higher dP/dt_{min} at 48 hours (-9765.6±364.1 and -6871.2±515.7 mmHg/sec for 48 hour Sham and TAC respectively) but this was not significant (Figure 3.4.B). There was no difference in dP/dt_{min} at 7 days; -8527.6±1611.6 and -8762.5±728.9 mmHg/sec for 7 day Sham and TAC respectively.

The micromanometry data shows that the aortic banding was successful, resulting in increased blood pressure and subsequent pressure overload of the heart.

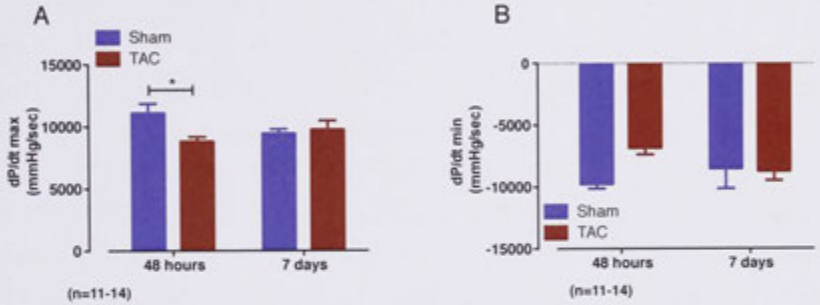


Figure 3.4. Contractility in pre-hypertrophic and hypertrophic hearts.

(A) TAC lead to a decrease in dP/dt max at 48 hours, with no detectable difference at 7 days. (B) TAC lead to a slight increase in dP/dt min at 48 hours (not significant) but was unchanged at 7 days. (* $p < 0.05$, Two way-ANOVA with Bonferonni's post-hoc analysis). Data is shown as mean with SEM (n=11-14).

The increase in LV pressure resulted in LV hypertrophy at 7 days with an increase in the whole heart and LV weight, relative to tibia length or body weight (Figure 3.5). There was a significant increase in heart weight relative to body weight (5.1±0.1 to 5.9±0.1 mg/g, $p<0.001$) at 48 hours, but this increase in heart size did not reach significance when normalised to tibia length. Weight loss can occur as the health of the animal deteriorates. Loss of more than 20% body weight was considered an ethical endpoint in this study, however, none of the mice reached this endpoint indicating they were still in a healthy weight range. Normalising the changes in heart weight to tibia length is more accurate as it removes any bias due to loss in body weight. Heart weight was significantly higher in TAC-banded animals at 7 days, relative to both body weight and tibia length (32.6% and 34.8% increase respectively). As expected, this increase in heart weight was derived from the LV. The TAC animals had significantly heavier LV at 7 days ($p<0.0001$), compared to body weight and tibia length (44.8% and 46.8% increase respectively). Whilst the LV weight tended to be heavier at 48 hours, this slight increase was not significant. Importantly, there was no increase in right ventricular (RV) weight at 48 hours and 7 days compared to both body weight and tibia length (Table 3.1), reinforcing that the hypertrophy was derived only from the left ventricle. Furthermore, there was no increase in lung-wet weight (WW), indicating the model did not result in heart failure (which can be characterised by oedema of the lung and increased lung WW, Figure 3.6.A). The body weights of all Sham and TAC animals were comparable both pre-surgery and at the Endpoint of the study (Figure 3.6.B).

At the molecular level, cardiac hypertrophy can be characterised by the re-expression of cardiac foetal genes; atrial natriuretic peptide (ANP), brain natriuretic peptide (BNP), beta-myosin heavy chain (β-MHC) and alpha-skeletal actin (α-Ska) (*198*). These foetal genes are commonly used as markers of hypertrophy. TAC lead to re-expression of these foetal genes in the whole LV samples; BNP was re-activated as early as 48 hours post-TAC ($p<0.0001$), with upregulation of the remaining markers reaching significance in the 7 day post-TAC animals (Figure 3.7).

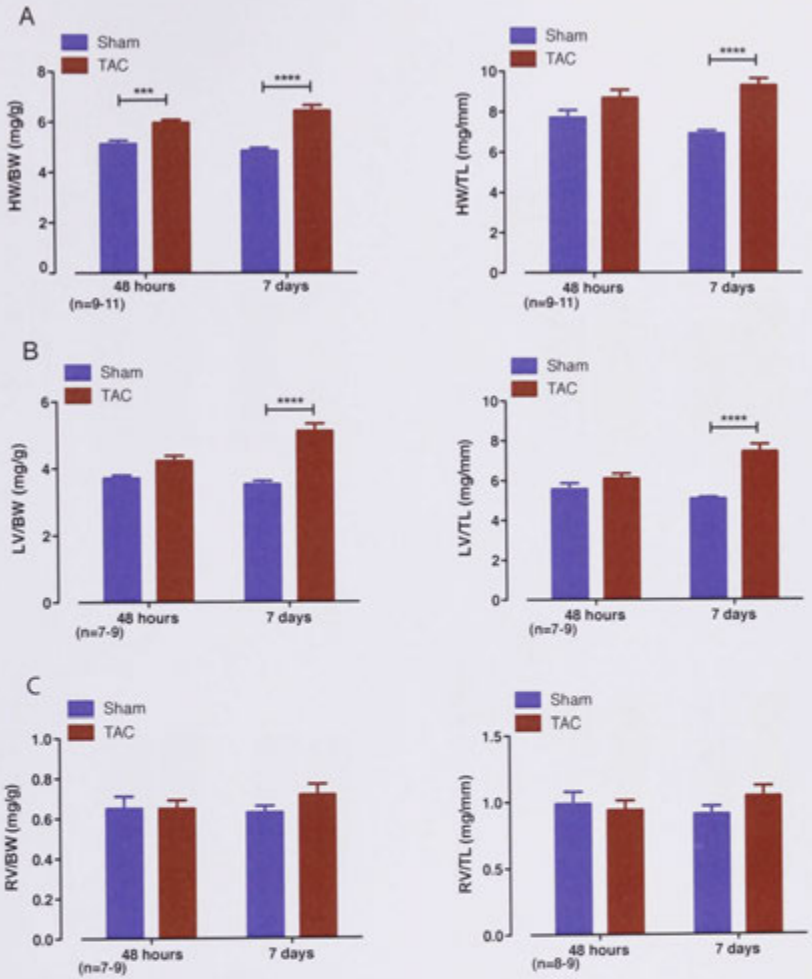


Figure 3.5. TAC leads to LV hypertrophy.

(A) TAC leads to a significant increase in heart weight normalised to body weight as early as 48 hours (left). Significant increase is detectable only at 7 days when normalising heart weight to tibia length (right). (B) The increase in heart weight is derived from the LV; significant increase in LV weight compared to both body weight and tibia length at 7 days. (C) The increased heart weight is not derived from the RV; no difference between the RV weight of Sham and TAC animals. (*** $p < 0.001$, **** $p < 0.0001$, Two way-ANOVA with Bonferonni's post-hoc analysis). Data is shown as mean with SEM (n=7-11).

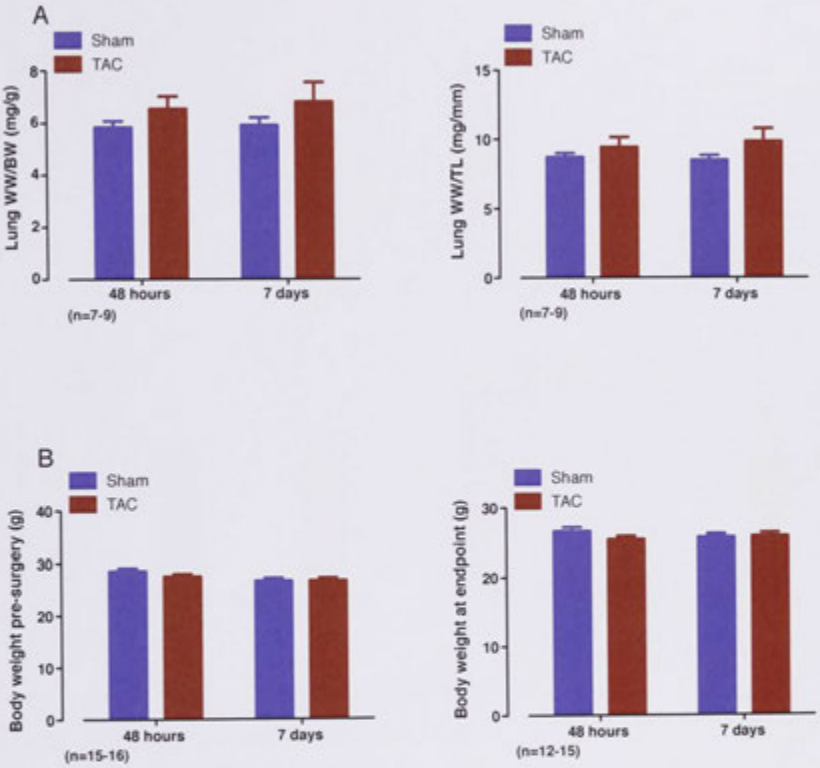


Figure 3.6. TAC does not lead to Heart Failure.

The lung wet weight (WW) did not increase in TAC animals compared to body weight (left) or tibia length (right). (B) The body weights of Sham and TAC animals were comparable at both pre-surgery and the endpoint. Data is shown as mean with SEM (n=7-16).

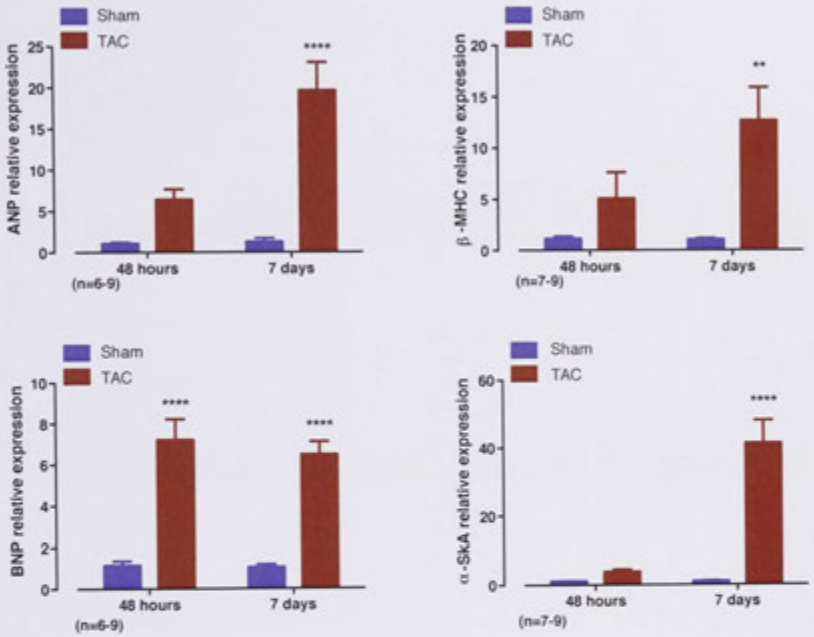


Figure 3.7. TAC leads to re-activation of cardiac foetal genes.

The hypertrophy markers atrial natriuretic peptide (ANP), beta-myosin heavy chain (β-MHC), brain natriuretic peptide (BNP) and alpha-skeletal actin (α-Ska) were upregulated 7 days post-TAC in the whole LV samples. BNP was upregulated pre-hypertrophy at 48 hours. (** $p<0.01$, **** $p<0.0001$, Two way-ANOVA with Bonferonni's post-hoc analysis). Data is shown as mean with SEM (n=6-9).

Table 3.1 Assessment of Hypertrophy.

	48 hours		7 days	
	Sham	TAC	Sham	TAC
Physiological data				
HW/BW (mg/g)	5.1 ± 0.1	5.9 ± 0.1	4.8 ± 0.1	6.4 ± 0.2
HW/TL (mg/mm)	7.7 ± 0.4	8.7 ± 0.4	6.9 ± 0.2	9.3 ± 0.3
LV/BW (mg/g)	3.7 ± 0.1	4.2 ± 0.2	3.5 ± 0.1	5.1 ± 0.2
LV/TL (mg/mm)	5.6 ± 0.3	6.1 ± 0.3	5.1 ± 0.1	7.5 ± 0.4
RV/BW (mg/g)	0.7 ± 0.1	0.7 ± 0.0	0.6 ± 0.0	0.7 ± 0.1
RV/TL (mg/mm)	1.0 ± 0.1	0.9 ± 0.1	0.9 ± 0.1	1.0 ± 0.08
Lung WW/BW (mg/g)	5.8 ± 0.2	6.5 ± 0.5	5.9 ± 0.3	6.8 ± 0.8
Lung WW/TL (mg/tl)	8.7 ± 0.3	9.4 ± 0.7	8.5 ± 0.3	9.8 ± 0.9
Initial BW (g)	28.5 ± 0.6	27.4 ± 0.5	26.6 ± 0.6	26.6 ± 0.5
Endpoint BW (g)	26.5 ± 0.6	25.4 ± 0.5	25.7 ± 0.5	25.8 ± 0.5
Haemodynamics				
Aortic SP (mmHg)	116.8 ± 2.9	155.9 ± 3.3	9458.9 ± 320	176.1 ± 6.3
LV SP (mmHg)	113.9 ± 1.8	149.2 ± 4.3	111.3 ± 2.8	166.5 ± 9.0
Aortic DP (mmHg)	83.3 ± 2.1	72.9 ± 3.6	110.9 ± 2.3	75.5 ± 3.5
LV EDP (mmHg)	6.7 ± 0.4	8.7 ± 0.7	78.8 ± 2.1	7.4 ± 1.6
HR (BPM)	457.0 ± 6.7	487.0 ± 10.8	5.2 ± 1.2	477.2 ± 7.8
dP/dT max (mmHg/sec)	11099.8 ± 760.9	8790.9 ± 372.6	459.3 ± 11.3	9757.6 ± 710.4
dP/dT min (mmHg/sec)	-9765.6 ± 364.1	-6871.2 ± 515.7	-8572.6 ± 1611.6	-8762.5 ± 728.9

The abbreviations used in this table are: BW, body weight; DP, diastolic pressure; EDP, end diastolic pressure; HR, heart rate; HW, heart weight; LV, left ventricle; RV, right ventricle; SP, systolic pressure; TL, tibia length; WW, wet weight. (n=7-16).

3.2.2 Cardiomyocyte enrichment

To select for transcripts expressed only by cardiomyocytes, these cells were enriched from the hearts of 5 Sham and 5 TAC animals at each time point (20 animals in total) according to a standard protocol (*176*). In brief, the hearts were perfused with collagenase to digest the extracellular matrix and release the cardiomyocytes, and the ventricles isolated for tissue dissociation to disperse the cardiomyocytes into a single-cell suspension. Differential centrifugation was then used to separate cardiomyocytes from other cell types, such as fibroblasts, based upon the density difference between cardiomyocytes and other cells (Figure 3.8). The wash fractions, which were presumed to contain fibroblasts and other cells, were retained to test for cell purity. Dr Ming Li performed the first stages of cardiomyocyte isolations up until the ventricles were isolated, subsequent steps were performed by Sara Holman with my assistance. I performed the RNA isolation and RT-qPCR analysis, with assistance from Nadine Mourad.

Table 3.2 shows a summary of the cardiomyocyte enrichment. The total number of cells and the proportion of rod-shaped cardiomyocytes were measured before and after differential centrifugation. On average, 2.2 million cells were isolated from each heart, of which 63% were rod-shaped cardiomyocytes. After washing and centrifugation, 1.3 million cells remained, with 66% derived from rod-shaped cardiomyocytes. Samples with a low number of total cells or less than 60% rod-shaped cardiomyocytes were not considered for next generation sequencing. TRIzol® was added directly to the isolated cardiomyocytes for RNA extraction and the RNA quality assessed by an Agilent Bioanalyzer Chip. The extracted RNA was of high quality, indicated by a RIN higher than 7.2 (Table 3.2).

To test the enrichment of cardiomyocytes in the samples, expression of several marker genes of fibroblasts and cardiomyocytes were analysed by RT-qPCR. The fibroblast markers α-Vimentin, collagen type I alpha 1 (Col1a1), collagen type III alpha 1 (Col3a1) and discoidin domain receptor tyrosine kinase 2 (Ddr2), were enriched in the wash fractions compared to cardiomyocytes and whole LV tissue (Figure 3.9). Similarly, myosin light chain 2v (Mlc2v) and cardiac muscle troponin T (Tnnt2), markers of cardiomyocytes, were highly expressed only in the cardiomyocyte fraction (Figure 3.9). Importantly, these enriched cardiomyocytes still displayed reactivation of the hypertrophic markers α-SkA and ANP 7 days post-surgery (Figure 3.10). This confirmed there was an enrichment of

cardiomyocytes from both pre-hypertrophic and hypertrophic hearts, indicating that the samples were appropriately pure for deep sequencing.

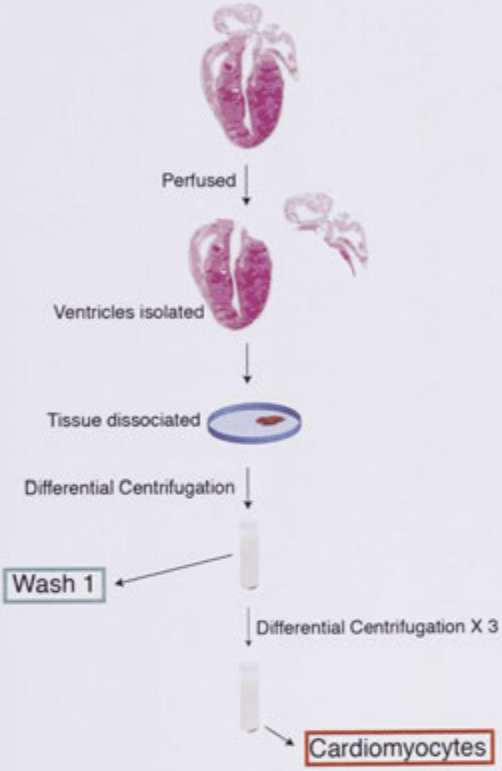


Figure 3.8. Method of Cardiomyocyte enrichment.

Overview of cardiomyocyte isolation from Sham and TAC hearts; hearts were excised rapidly and perfused with collagenase to digest the extracellular matrix. The hearts were cut just below the atria to release the ventricles, and the ventricles dissociated into a single-cell suspension. Repeated differential centrifugation was used to isolate the cardiomyocytes. The supernatant from the first wash, containing fibroblasts and cell debris, was retained to assess enrichment.

Table 3.2. Assessment of cardiomyocyte enrichment.

Samples highlighted in red were used for NGS.

	Cell counts					Fibroblast markers (fold depletion)				CM markers (fold enrichment)		Hypertrophy markers	
Mouse ID	Cell no. pre wash	Rods (%)	Cell no. post wash	Rods (%)	RIN	α-vim	Ddr2	Col3a1	Col1a1	Mlc2v	cTnT	α-SkA	Nppa
48 hour Sham													
48 Sham 3	NC	NC	1.30E+06	55	7.8	33.0	12.5	40.1	13.3	5.9	7.0	0.43	0.35
48 Sham 1	2.90E+06	66	1.30E+06	69	7.2	39.8	10.3	58.0	10.3	5.9	7.3	0.46	0.28
48 Sham 2	2.60E+06	75	1.70E+06	79	8.2	32.4	5.5	13.2	127.7	3.7	5.8	0.27	0.55
48 Sham 4	2.45E+06	81	1.70E+06	73	7.4	19.1	9.4	11.9	5.5	10.7	11.0	0.46	0.45
48 Sham 5	1.30E+06	41	3.00E+05	22	NA	11.3	3.3	12.1	3.8	3.3	3.0	0.15	0.17
48 hour TAC													
48 TAC 3	2.40E+06	83	1.90E+06	76	7.6	48.3	8.4	101.4	59.0	2.3	3.0	0.50	0.37
48 TAC 1	2.30E+06	81	1.60E+06	73	8.2	27.4	11.0	10.2	17.3	7.0	11.7	0.76	1.22
48 TAC 2	2.50E+06	68	1.40E+06	73	8.3	24.5	11.1	9.1	17.0	7.3	6.9	0.89	1.45
48 TAC 4	3.20E+06	57	1.10E+06	58	7.8	37.5	4.9	48.9	18.0	0.02	2.9	0.43	0.18
48 TAC 5	1.40E+06	61	1.00E+06	63	8.5	18.7	5.6	3.7	18.0	9.2	16.9	0.66	0.95
7 day Sham													
7 Sham 3	2.50E+06	57	1.50E+06	67	7.8	26.1	8.8	12.0	13.0	7.5	6.3	0.21	0.15
7 Sham 1	2.00E+06	70	1.70E+06	68	7.9	30.1	3.0	57.5	28.0	2.9	3.7	0.16	0.20
7 Sham 4	1.80E+06	52	9.00E+05	58	7.8	23.3	5.5	40.0	14.5	3.5	3.5	0.31	0.16
7 Sham 5	1.50E+06	46	3.00E+05	64	7.5	9.9	4.0	0.8	4.2	1.7	3.2	0.10	1.05
7 Sham 2	2.80E+06	63	1.00E+06	72	7.6	21.1	3.7	7.5	6.7	3.5	4.6	0.14	0.21
7 day TAC													
7 TAC 3	2.76E+06	61	1.80E+06	69	8.3	18.2	12.7	6.0	8.5	51.5	34.0	2.46	0.58
7 TAC 4	1.20E+06	47	1.20E+06	44	7.7	3.6	3.5	8.0	3.8	17.4	14.3	1.66	1.60
7 TAC 1	2.40E+06	83	1.80E+06	75	8.1	20.1	9.7	5.1	10.0	3.8	6.8	2.65	3.59
7 TAC 2	2.02E+06	63	9.00E+05	65	8.2	19.4	17.9	58.6	32.0	13.7	11.6	1.53	1.10
7 TAC 5	2.10E+06	58	1.40E+06	59	7.9	NA	NA	NA	NA	NA	NA	2.97	3.22

The abbreviations used in this table are: Ddr2, discoidin domain receptor tyrosine kinase 2; Col3a1, collagen type III alpha 1; Col1a1, collagen type I alpha 1; cTnT, cardiac muscle troponin T; Mlc2v, myosin light chain 2v; α-SkA, alpha skeletal actin; Nppa, atrial natriuretic peptide; α-vim, alpha vimentin; RIN, RNA Integrity number; CM, cardiomyocyte; NA, not available – no fibroblast sample available; NC, not counted.

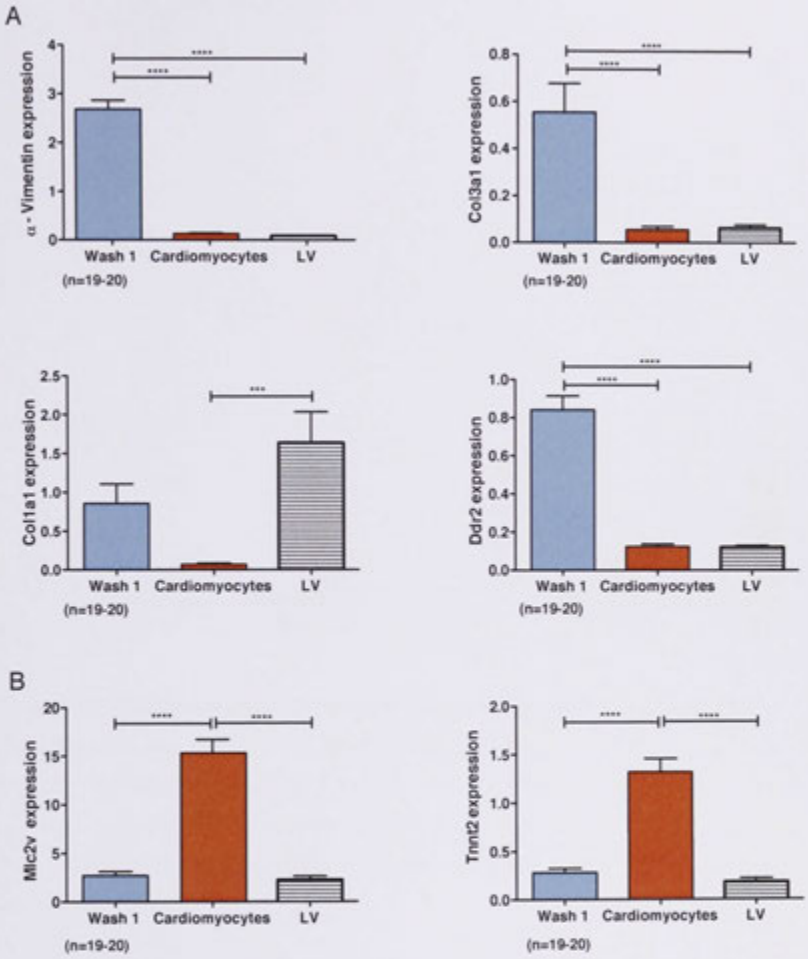


Figure 3.9. Enrichment of Cardiomyocytes for RNA extraction.

Cardiomyocytes were purified from whole hearts; RT-qPCR was performed on RNA extracted from Wash 1, whole LV and the cardiomyocyte fractions (CM) and normalised to HPRT expression. (A) There was significantly higher expression of fibroblast markers α-vimentin, Col3a1, Col1a1 and Ddr2 in wash 1 or whole LV compared to the CM fraction. (B) There was also significant enrichment of cardiomyocyte markers Mlc2v and Tnnt2 in the CM fraction compared to LV and wash 1. (** p <0.01, *** p<0.001, **** p<0.0001, One way-ANOVA with Bonferonni's post-hoc analysis). Data is shown as mean with SEM (n=19-20).

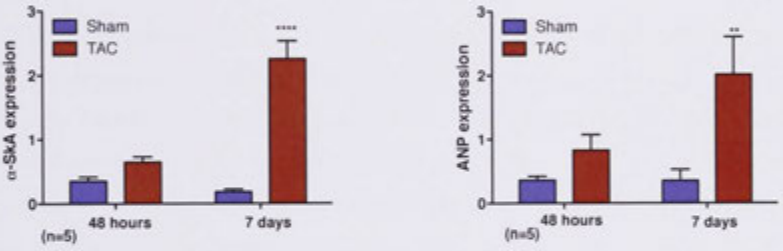


Figure 3.10. Hypertrophy markers are upregulated in the enriched cardiomyocytes.

Cardiomyocytes were purified from whole hearts; RT-qPCR was performed on RNA extracted from the cardiomyocyte fractions (CM) and normalised to HPRT expression. There was a significant upregulation of the cardiac foetal genes α-sKa and ANP at 7 days post-TAC. (** $p<0.01$, **** $p<0.0001$, Two way-ANOVA with Bonferonni's post-hoc analysis). Data is shown as mean with SEM (n=5).

3.3 Discussion

In order to document the mRNA and miRNA changes that occur during cardiac hypertrophy, it was essential to set up a valid and consistent model of the disease. For this reason, transverse aortic constriction was used. The surgeries were performed on 12- week old male C57BL/6 mice of comparable weight (Figure 3.7.B). The LV hypertrophy was confirmed in the TAC banded animals by haemodynamics, physiological and molecular measurements. The hearts were then enriched for cardiomyocytes using a standard method, for subsequent next-generation sequencing.

3.3.1 Pre-hypertrophic and compensated hypertrophy

The process of hypertrophy is a major undertaking for a cardiomyocyte and involves molecular, structural and physiological changes. The overall aim of this thesis was to examine the changes in miRNA processing and 3' UTR lengths during the development of left ventricular hypertrophy, specifically within a cardiomyocyte. Thus, the aim of this chapter was to establish a robust and validated model of LV cardiac hypertrophy, and to isolate pre-hypertrophic hearts and compensated hypertrophic hearts prior to their progression to heart failure. But this begs the question, what is the definition of pre-hypertrophic and compensated hypertrophy? Does hypertrophy begin when the appropriate molecular changes occur or when there is an increase in heart weight? The development of hypertrophy has previously been described in three stages: the initial myocardial damage and impairment of contractile function before there is an increase in heart mass, the second stage where normal function is restored due to an increase in cardiomyocyte mass, and the third stage that results in a gradual deterioration of cardiac function and output leading to heart failure (*199*). We have used the first stage of hypertrophy development to define pre-hypertrophic hearts; there is a slight impairment of ventricular performance but no increase in the weight of the LV. However, it is evident from our molecular data that the hypertrophic signalling has already been initiated at this stage, as shown by the increase in BNP levels at 48 hours. This needs to be taken into account when considering the gene expression changes found in the next-generation sequencing data.

The 7-day time point was used to capture compensated hypertrophic hearts. During this stage there is a temporary enhancement of cardiac contractility due to the thickening of the LV wall. Hypertrophy results in fibrosis, which is the deposition of extracellular matrix (ECM) proteins such as collagen that stiffens the ventricle leading to decompensation and a decrease in the strength of the cardiac contraction. Importantly, the transition from compensated to decompensated hypertrophy is a gradual process, making it difficult to determine an exact point for compensated changes. It is known that LVED increases with worsening fibrosis and cardiomyocyte damage, which is then followed by a decrease in cardiac output (*200*). As there was not a significant increase in LVED in our experiments we can conclude that the hearts were compensated and still maintaining cardiac output. Furthermore, the absence of right ventricular hypertrophy or pulmonary oedema indicates that the mouse hearts were free of overt heart failure.

3.3.2 Haemodynamic measurements using micromanometry

Micromanometry was used to measure the changes in pressure and cardiac function of the TAC banded animals. Micromanometry is an invasive technique that involves dissection of the right carotid artery, and was therefore used only at the end-point of the study. Echocardiography is another method of evaluating cardiac function; it is a non-invasive ultrasound technique that can also be used to visualise cardiovascular structures. Echocardiography would have allowed the recording of dynamic changes to systolic and diastolic function during the hypertrophy time course. Furthermore, echocardiography allows the direct measurement of LV wall thickening (*193*) and thus a more accurate time-point where the LV weight increases could have been determined. Despite the invasiveness of micromanometry, it still revealed the pressure overload in the TAC banded animals and the subsequent changes to cardiac function.

3.3.3 TAC generated LV hypertrophy

The physiological data indicated the TAC banded animals were undergoing LV hypertrophy but had not yet entered heart failure (Figure 3.5-3.6). The LV weight was significantly increased only at 7 days, with no change in the RV or the lung wet weight. As expected, the hypertrophy markers ANP and BNP were re-expressed in the TAC banded animals at 7 days post surgery. ANP and BNP have natriuretic, diuretic and vasorelexant activities, which alter blood pressure and volume. Both activate guanylyl cyclase-linked

receptors leading to an increase in intracellular cGMP concentrations (*201*). The natriuretic peptides are antihypertrophic, with the increase in cGMP inhibiting hypertrophic signalling pathways such as calcineurin-NFAT and MAPK/ERK (*202, 203*). Increased levels of ANP and BNP also decrease the amount of sodium in the blood and thus blood volume, which in turn lowers blood pressure and increases cardiac output. Interestingly, BNP was the only marker to be upregulated with a significant increase already at 48 hours. Consistent with this, it has been previously shown that the concentration of BNP in plasma dramatically increases in patients with acute myocardial infarction, while the ANP concentrations increase only mildly (*204*). The transcription of BNP is more rapidly induced than ANP during cardiomyocyte hypertrophy *in vitro* (*205*), and it has been suggested that BNP is an emergency hormone in response to heart stress (*206, 207*).

Pressure overload also resulted in the induction of the foetal isoforms of sarcomeric proteins α-SkA and β-MHC. These isoforms have different functional properties that are beneficial to the increase in hemodynamic load (*208*). The contractility and ATP consumption of the heart depends on the number and the composition of sarcomeres (*209*). β-MHC is bioenergetically more efficient than the adult α-MHC isoform and it has been hypothesised that the shift from α- to β-MHC is an adaptive response to conserve energy (*210, 211*). However, β-MHC has lower adenosine triphosphate activity and lower filament sliding velocity than α-MHC, ultimately decreasing contractile function (*212, 213*). Therefore, the increase in β-MHC in the compensated TAC hearts could later become maladaptive and contribute to heart failure. As mentioned above, the levels of α-SkA also increase in the hypertrophic heart. There is a linear relationship between the induction of α-SkA and the progression of LV hypertrophy (*214*). Like β-MHC, the increase in α-SkA is an adaptive response to maintain normal cardiac contractility and function, as the levels of α-SkA are positively correlated with myocardial contractility measurements (*214*). Consistent with these previous observations, the alterations in dP/dt_{max} and dP/dt_{min} observed at 48 hours are corrected by 7 days when the levels of α-SkA are upregulated (Figure 3.4, 3.7).

The haemodynamic measurements, physiological and molecular data indicate that TAC lead to pressure-overload induced LV hypertrophy. In agreement with previous studies using the TAC model, this hypertrophy was compensated and resulted in the re-expression

of natriuretic peptides and sarcomeric proteins, which impact on cardiac contractility and output.

3.3.4 Cardiomyocyte enrichment

Once the hypertrophy was confirmed, cardiomyocytes were purified from the hearts of the Sham and TAC animals. This was performed using a standard isolation protocol, which was modified for immediate isolation of the cells (*176*). In brief, the heart was perfused with collagenase and differential centrifugation was used to isolate the cardiomyocytes. Normally the cells are then cultured for 24 hours by plating onto laminin-coated dishes for attachment. Once attached, the non-attached cells, which are mostly rounded cardiomyocytes are washed off. The remaining cells in the flask are 90% cardiomyocytes, resulting in a highly pure primary cardiomyocyte sample. However, this protocol was modified to exclude the short-term culture and the cells were isolated immediately after washing. This resulted in 66% rod-shaped cardiomyocytes on average in our samples, compared to the 90% achieved with short-term culture. It was hypothesised that culturing the cardiomyocytes for 24 hours could alter both the miRNA and mRNA expression and turnover, potentially confounding the next-generation sequencing analysis. The half-life of miRNAs in embryonic fibroblasts is ~ 5 days, but the average half-life of miRNAs and mRNAs has not been measured in cardiomyocytes (*215*). The cardiomyocyte isolation procedure was completed within 3 hours to minimize the effects of RNA degradation. Furthermore, the rod-shaped cells mimic the cardiomyocytes in the intact heart while a majority of the round cells are hypercontracted cardiomyocytes (*176*). This only becomes important when using the culture for signalling and mechanical assays. It is a lesser concern in the protocol used for this thesis, where the cells were immediately extracted for RNA. The percentage of cardiomyocytes was therefore also potentially underestimated, as the hypercontracted cardiomyocytes were not counted. A more accurate measurement of cardiomyocyte enrichment may have been to sort the cells using flow cytometry based on cardiomycoyte and fibroblast cell surface markers. This was not performed as incubation of the cells with antibodies and subsequent flow cytometry analysis would have been lengthy and potentially damaging to the cardiomyocytes, again potentially impacting on the miRNA and mRNA expression profiles.

The cardiomyocyte enrichment was confirmed using RT-qPCR against multiple cardiomyocyte and fibroblast markers. The majority of markers (α-vimentin, col1a1, col3a1

and tnnt2) were identified from a study that used laser microdissection to isolate pure cardiomyocytes and fibroblasts at a single-cell level (*216*). Alike to the results presented in Figure 3.9, α-vimentin, col1a1 and col3a1 are highly abundant in fibroblasts and tnnt2 is expressed in isolated cardiomyocytes. Mlc2v (one of the major myosin light chain isoforms in the heart), is specific to ventricular cardiomyocytes and was highly enriched in the cardiomyocyte samples (*217*). Furthermore, an additional fibroblast marker, Ddr2, was significantly depleted in the purified cardiomyocytes compared to the fibroblasts. Ddr2 is a collagen specific receptor tyrosine kinase that is expressed on cardiac fibroblasts, but not on cardiomyocytes or endothelial cells (*191*). All of the cardiomyocyte preparations were enriched for the expression of cardiomyocyte markers and depleted in fibroblast markers. There was however variability in the expression of markers between samples, with some samples having high fold enrichment of one marker but not another (Table 3.2). This could reflect differences in enrichment or variation in the expression of a marker within a sample. The eight samples used for next-generation sequencing had 65-79% rod shaped cardiomyocytes and are highlighted in Table 3.2. The identical samples were used for next-generation sequencing of miRNAs (Chapter Four) and mRNA 3' ends (Chapter Five).

Chapter Four

miRNA Sequencing

4.1 Introduction

miRNAs have in the past been measured using microarrays, PCR or northern blotting. All of these approaches require *a priori* knowledge of miRNAs and their sequences. The original studies on global miRNA expression in the normal and hypertrophic murine heart were performed using microarray platforms that rely on hybridisation to an oligonucleotide probe (*120, 121, 127, 128*) or RT-qPCR arrays (*129*). The use of such platforms does not allow for discrimination of processing variants and often does not include probes against what was referred to as the miRNA*. More recently, two studies published in 2012 used NGS to profile the miRNA expression in whole LV samples 7 days after TAC banding (*130, 131*). In both of these studies, all of the sequences that matched a known strand of a miRNA locus were clustered and counted towards that miRNA strand. They did not investigate the sequencing variants of the expressed miRNAs. Furthermore, there are no published studies of miRNAs in purified cardiomyocytes during hypertrophy. Therefore, information regarding miRNA processing and variants in cardiac hypertrophy is currently lacking, which could impact the way we think about miRNA interactions and their targets. NGS allows the discovery of miRNA variants in a population as it has single-nucleotide resolution.

The aim of this chapter was to sequence the cardiomyocyte miRNA population of normal (Sham), pre-hypertrophic (48 hour TAC) and hypertrophic (7 day TAC) samples. Firstly, to identify those miRNAs differentially regulated in pre- and hypertrophic cardiomyocytes. Secondly, to generate a reference for all the miRNA processing variants in murine cardiomyocytes, and then to establish if any of these processing variants change in frequency during the induction of hypertrophy. To do this we used SOLiD™ sequencing and a bioinformatics analysis pipeline modified from a previous study in which we documented the miRNA population in the HL-1 cardiomyocyte cell line (*55*).

4.2 Results

This chapter describes the small RNA sequencing of normal, pre- and hypertrophic cardiomyocytes, focusing on differential expression of miRNAs and miRNA processing variants.

4.2.1 Conversion of miRNAs into libraries for sequencing

The RNA extracted from purified cardiomyocytes (Chapter Three) was used for NGS of small RNAs. The library preparation was performed under the guidance of Dr David Humphreys at the Victor Chang Cardiac Research Institute, while Dr David Humphreys completed the e-PCR and subsequent steps. Eight samples in total, two Sham and two TAC from each time point, were subjected to sequencing. The libraries were prepared using the NEBNext® Small RNA Library Prep Set for SOLiD™, an outline of the method is shown in Chapter Two, Figure 2.2. In brief, adaptor sequences are ligated to the 3' and 5' end of the miRNA, which are then used to reverse transcribe the miRNA into cDNA. The samples were multiplexed for sequencing so that all samples could be sequenced simultaneously. This was achieved by including unique barcode sequences into the PCR primers for each sample. After PCR amplification, the miRNA products were size selected using polyacrylamide gel electrophoresis. The expected size of a 21 nt miRNA after addition of the adaptor sequences is ~ 114 nt. Therefore, PCR products corresponding to ~ 110-125 nt were excised from the gel to allow for variations in the length of a miRNA (Figure 4.1). The larger bands represent other RNA species such as tRNAs and rRNAs that would have been cloned using the library preparation method. The miRNA-derived products were purified and the sizes confirmed with a second polyacrylamide gel (Figure 4.2). Note that less of the 7 Sham 2 sample was loaded onto the gel, however this was accounted for when diluting for subsequent steps.

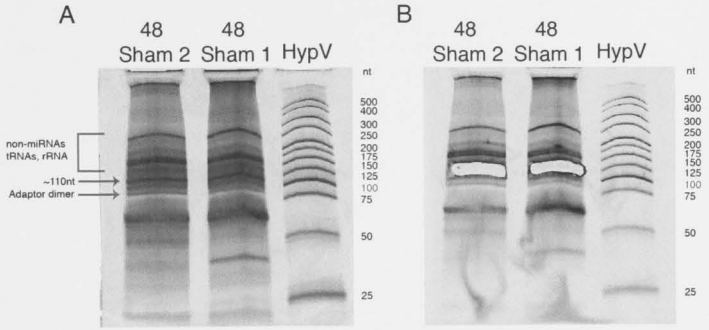


Figure 4.1. Size selection of miRNA PCR products for next-generation sequencing.

Representative polyacrylamide gels for two sham samples (48 Sham 1, 48 Sham 2) before (A) and after (B) excision of bands corresponding to 110-125 nt in length. The expected locations of the miRNA are indicated by the red arrow, and expected size of adaptor dimers and non-miRNA species are indicated in blue. HypV; hyperladder V.

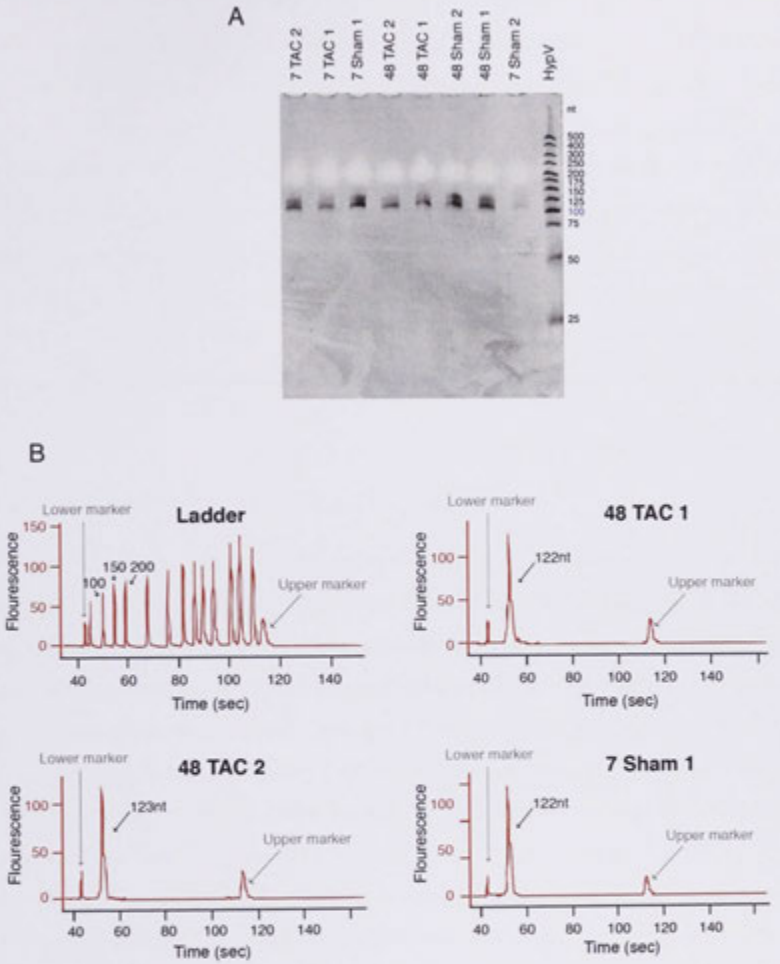


Figure 4.2. Confirmation of miRNA library size.

(A) Polyacrylamide gel for eight of the samples subjected to sequencing after size selection. The isolated PCR products are between 100-125 nt. (B) The sizes and concentrations of the sequencing libraries were then checked on an Agilent Bioanalyzer Chip. The results are shown for the size ladder and three of the samples. The grey arrows refer to the lower and upper markers, and the black arrow in each sample refers to the average size of the main peak. The sizes of the ladder corresponding to the same elution time as the samples are also indicated. HypV; hyperladder V.

After purification, the PCR products were brought to the same concentration and analysed using an Agilent Bioanalyzer Chip. Figure 4.2.B shows the Agilent Chip results for three of the sequencing libraries, the apparent size of the libraries was 122-123 nt. This is consistent with the instructions for the NEBNext® Small RNA Library Prep Set for SOLiD™, where the expected peak size on the Agilent Bioanalyzer is described as between 114 and 123 nt. It was determined that all the samples were within a comparable concentration range and were used for emulsion PCR (e-PCR). In the e-PCR reaction, the PCR products are mixed into a water-in-oil emulsion, such that each water droplet contains a single PCR product as the template attached to a single bead coated in primers to form a clonal colony. The beads were then deposited onto three chambers or lanes of a glass flow cell and sequenced on the SOLiD 5500™.

4.2.2 General features of the libraries

The following section documents the features of the small RNA libraries as well as details for mapping the data. Dr Hardip Patel from ANU performed the mapping of the miRNA libraries, and I completed the secondary analysis, including graphical representations and differential expression, using the R programming language. The libraries were sequenced on average to a depth of 3.97×10^7 tags (up to 35 nt in length), yielding a total of 3.18×10^8 tags. The shortest known products of Dicer cleavage are 19 nt and anything shorter than this is unlikely to be a miRNA (*218*). Therefore, the tags were restricted to a length greater than or equal to 19 nt. Approximately 72% of the tags met this criterion as shown in Figure 4.3.A. As the samples were diluted to equivalent concentrations prior to sequencing, the yield of sequenced tags was relatively even across libraries, although the 48 Sham 2 had approximately twice the number of tags seen with the other libraries. This suggests that more of the 48 Sham 2 library was added, resulting in more tags. This difference in tags is accounted for when normalising the libraries.

During SOLiD™ sequencing, each nucleotide or ligation cycle is associated with a colour call and associated quality value. The quality value is related to the strength of the signal and is a measure of confidence that the colour called for that ligation cycle is correct. The data was filtered to retain only tags with high average quality values (representing more than 98% accuracy). Importantly, all of the libraries were of high quality, as 99.2% of tags were retained after applying this quality value filter (Figure 4.3.A).

To identify known miRNAs and their processing variants, tags were mapped to the mouse genome (mm9 assembly including rRNA). miRBase version 18 annotations were used to determine miRNA coordinates and tags that started within ± 3 nt of a miRBase annotated 5' start site were assigned to that miRNA.

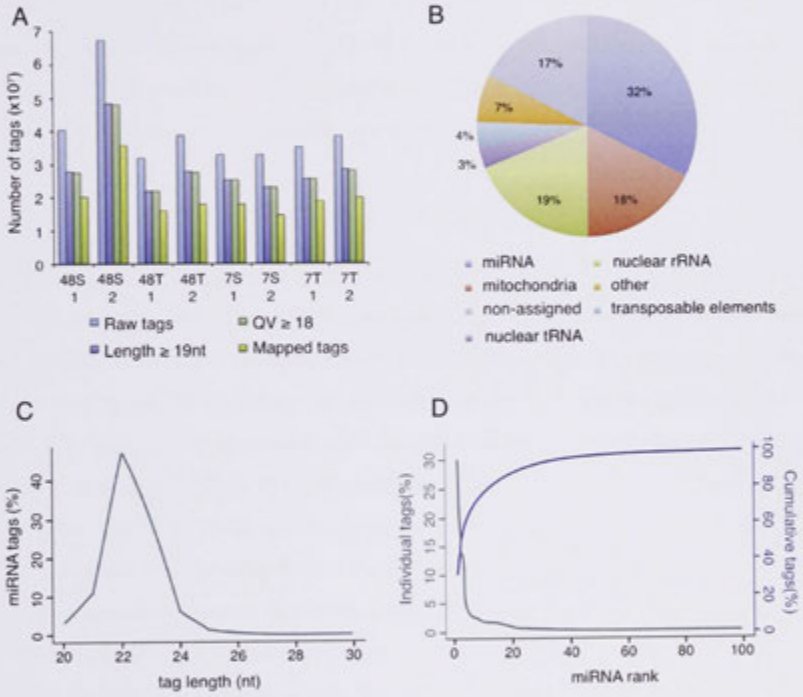


Figure 4.3. General features of the small RNA library.

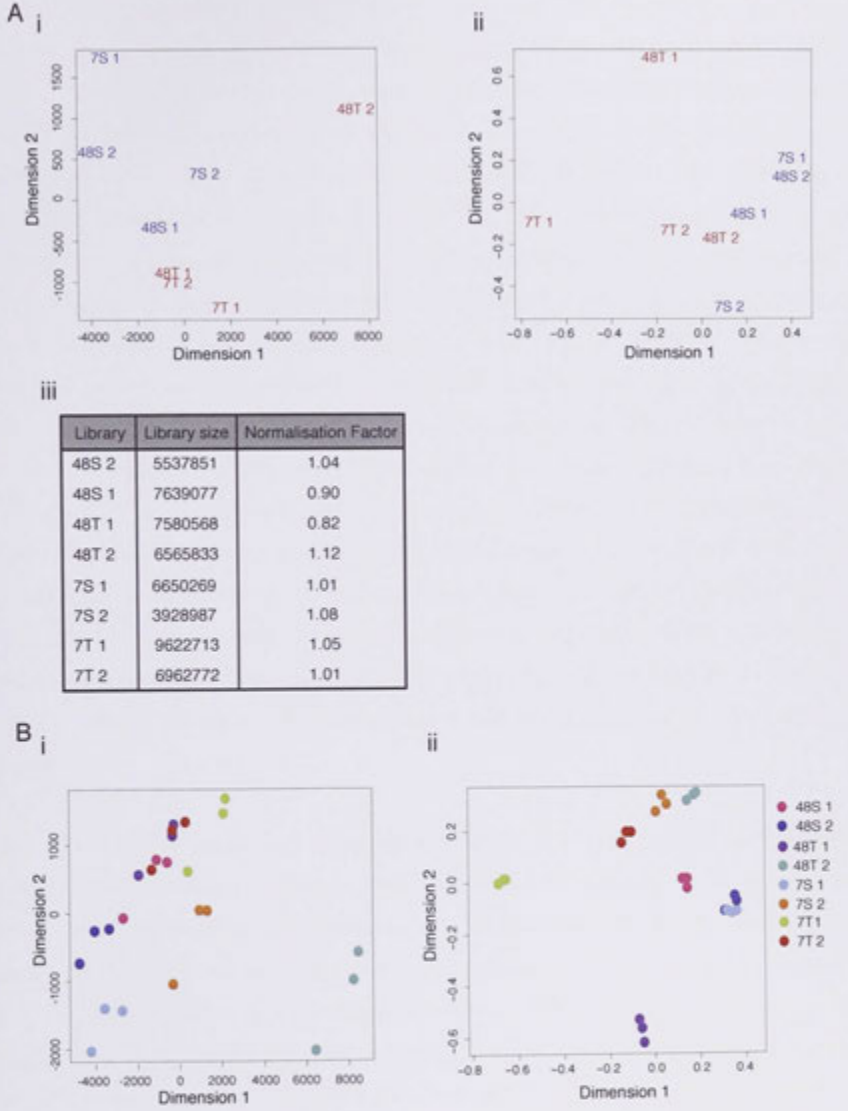
(A) Number of tags in sequential order of filtering steps; number of raw tags, number of tags with length ≥ 19 nt, number of tags with quality value (QV) ≥ 18 and miRBase v18-mapped tags. S, Sham; T, TAC (B-D) Combined features of all eight small RNA libraries. (B) Distribution of 2.36×10^8 mappable tags across the mm9 mouse genome assembly after length and QV filter. (C) Distribution of tag lengths for all miRNA-mapped tags. (D) Distribution of tag counts across the one hundred most abundant miRNAs. Contribution of individual miRNAs to total tag count (green), or cumulative tag contribution (blue) is plotted against ranked miRNA abundance.

On average, 32% of raw tags mapped to a miRNA locus after restricting the length to ≥ 19 nt and the QV ≥ 18, detecting 806 miRNAs (Figure 4.3.B). Figure 4.3.B-D illustrates the mapping statistics and general features of all small RNA libraries combined. A large proportion of tags mapped to other RNA species, including rRNA and tRNA. Again, this could also be due to the size range taken during the library preparation. The modal tag length of miRNA-mapped tags was 22 nt (Figure 4.3.C), which is consistent with miRBase where the majority of miRNA sequences are 21-22 nt (*82*). To remove tags representing very lowly expressed miRNAs and spurious tags, the miRNAs were filtered for those that have average expression at levels of at least 10 counts per million (CPM) across the eight libraries, resulting in 260 confidently detectable miRNAs. The 20 most abundant miRNAs contributed 87.3% of all miRNA-mapped tags in the combined libraries (Figure 4.3.D), with the most abundant miRNA, cardiac specific miR-133a, contributing 30.2% of the tags.

4.2.3 Normalisation of miRNA tag counts

In order to use NGS data for differential expression analysis, tag counts must be normalised for each library to remove technical bias while maintaining true biological signal. As each library had a different total miRNA tag count, the libraries were initially normalised to the size of the library (based on miRNA mapped tags) by converting the tags to CPM. However, highly expressed miRNAs can sometimes take up a lot of "sequencing real estate" in a sample (*218*). The composition of the RNA population in each library is not the same, which can influence the detection of true differentially expressed transcripts (*219*). For this reason, we utilised the normalisation method 'trimmed mean of M-values' (TMM) that takes into account the composition of each library. This method assumes that most transcripts are not differentially expressed and their true levels remain constant, except for the data points that lie within the extreme M-value (log expression ratio) and A-value ranges (log-intensity averages). It then uses a simple linear scaling factor after trimming the data points that are located in these extreme ranges (*219*). All of the miRNA tags were used for normalisation and then the data was filtered for an average expression of at least 10 CPM.

To measure the effect of normalising the miRNA libraries, multi-dimensional scaling plots (MDS) for the eight miRNA NGS libraries using CPM and the TMM normalised samples were generated (Figure 4.4.A). The distances in the MDS correspond to leading log2-fold changes between each pair of RNA samples



iii

Library	Library size	Normalisation Factor
48S 2	5537851	1.04
48S 1	7639077	0.90
48T 1	7580568	0.82
48T 2	6565833	1.12
7S 1	6650269	1.01
7S 2	3928987	1.08
7T 1	9622713	1.05
7T 2	6962772	1.01

Figure 4.4. Normalisation of the miRNA sequencing libraries.

(A) MDS plots of miRNA CPM to miRNA mapped tags (i) and TMM normalised miRNA libraries (ii). The miRNA libraries were normalised by TMM normalisation factors (iii). (B) MDS plots of individual lanes both before, represented as CPM (i), and after (ii) TMM normalisation. Each sample is represented by a different colour and was sequenced on three lanes.

and can be interpreted as a type of unsupervised clustering. The leading log2-fold change is defined as the average of the largest absolute log2-fold changes between each pair of samples for the genes that distinguish those samples (*182*). In essence, MDS plots show the relative similarities between samples. Dimension 1 separates the Sham from the TAC samples after the data has been normalised using TMM (Figure 4.4.A.ii). There is not a complete separation of the Sham from the TAC after normalisation. However, the normalisation factors calculated using the TMM method are close to one, indicating that all of the libraries are fairly similar in composition (Figure 4.5.A.iii). As mentioned previously, the libraries were multiplexed and sequenced on three lanes of a SoLiD™ flow cell, with each lane being a technical replicate of each library. This data was exploited to determine the efficiency of the normalisation method. The samples on each lane were treated as individual samples and TMM normalisation applied. It is expected that after normalisation the individual lanes for each sample will cluster together as the biological signal is equivalent. The MDS plots based on individual lane data are shown in Figure 4.4.B, where each sample is a different colour and represented three times (lanes 1-3). It is clear that after normalisation the technical replicates are brought together and the three lanes for each sample overlap (Figure 4.4.B.ii). To conclude, the miRNA libraries are similar in composition but the use of TMM normalisation still brings the biological replicates closer together.

The data was also normalised independently using upper quartile (UQ) normalisation. Here, the normalisation factors are calculated from the 75% quantile of the tags for each library, after removing transcripts that are zero in all libraries (*220*). All of the miRNAs with a FDR <10% are equivalent between normalisation methods, and 13 of the 17 miRNAs with a FDR <30% are consistent between TMM and UQ normalisation. This suggests that the normalisation methods trialled give fairly consistent normalised tag counts for differential expression. The miRNAs identified as differentially expressed using UQ normalisation are listed in Appendix 8.7.

The miRNA populations in all the cardiomyocyte samples are very similar, regardless of their treatment. This is indicated by the Spearman rank correlation, which is a measure of the strength of the relationship between two samples (Figure 4.5). The lowest Spearman correlation is 0.92, which occurs when comparing the 48 TAC 2 and the 7 TAC 1 samples. Hierarchical clustering based on the Spearman correlation clusters the Sham and the TAC separately, suggesting the TAC samples are more closely related to each other than to the

Sham samples and vice versa. This provides further support for the use of TMM normalisation, which assumes that the samples are highly similar and majority of miRNAs do not change (*219*).

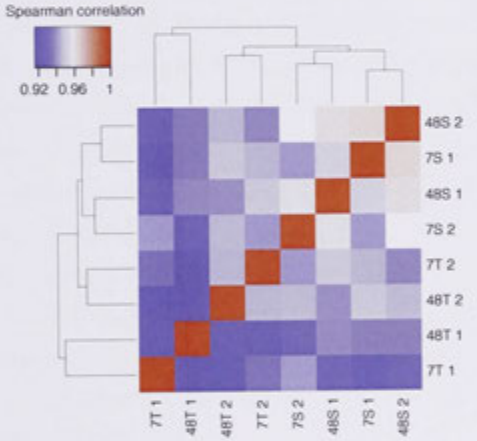


Figure 4.5. The cardiomyocyte samples have similar miRNA populations.

The heatmap shows the clustering of samples according to their Spearman rank correlation. The lowest Spearman correlation is 0.92.

4.2.4 miRNA differential expression

One of the main interests of this project was to look at miRNA processing variants. However, the data presented an opportunity to look at differential expression of miRNAs in purified cardiomyocytes during cardiac hypertrophy, which has not previously been done. To detect miRNAs that are differentially expressed between the TAC banded animals and the controls, pairwise comparisons were made between biological replicates using the empirical analysis of digital gene expression in R (edgeR) package (*182*). The edgeR package requires biological replicates and detects differential expression using an overdispersed Poisson model that accounts for increased variability and an empirical Bayes procedure to moderate the degree of overdispersion (*182*). A parameter then estimates the spread between replicates and separates the true biological variation from the technical variation. This package has previously been used in other studies to analyse the expression of miRNAs (*221-223*). Several pairwise comparisons were made (Figure 4.6, listed in Table 4.1) to maximise the chances of finding differentially expressed miRNAs. Using this method, 44 miRNAs with a p value <0.05 were detected. However, when making multiple comparisons a p value of 0.05 can result in a large number of false positives. For this reason, the p values need to be adjusted to take into account multiple testing. The edgeR package uses the Benjamini-Hochberg method to adjust p values and gives a false discovery rate (FDR) (*224*). The FDR is a measure of false positive, where a FDR of 10% means that 10% of significant tests will result from false positives. A FDR less than 30% was chosen to detect differentially expressed miRNAs as we could test the expression changes in the extended cardiomyocyte sample set. The fold change and expression levels of these miRNAs are illustrated in Figure 4.6. After adjusting the p values, 17 miRNAs had a false discovery rate (FDR) less than 30% (Figure 4.6-4.7).

The miRNAs with an FDR less than 30% are shown in Table 4.1, with many of the miRNAs being significant in multiple pairwise comparisons such as miR-21-5p. The 48 hour TAC vs remaining samples identified miR-324-3p as a candidate, however the levels of this miRNA increase 3-fold in both the 7 day Sham and TAC samples compared to the 48 hour sample. Thus, this miRNA was excluded from subsequent analysis. Importantly, when a comparison was made between the 48 hour and 7 day Sham samples, there were only two miRNAs with a p value <0.05. Furthermore, these miRNAs had a FDR close to 100% indicating no differences could be found between the 48 hour and 7 day Sham

samples and therefore they were suitable controls (details for all pairwise comparisons are shown in Appendix 8.6).

Table 4.1. Differentially expressed miRNAs with a FDR <30%.

Pairwise comparisons were made using the edgeR package. Each pairwise comparison is indicated by the bold text.

TAC vs Sham				
miRNA	logFC	logCPM	p value	FDR (%)
miR-21-5p	2.48	13.83	1.21E-09	3.14E-05
miR-223-3p	1.73	5.70	1.74E-05	0.23
miR-34c-5p	2.02	3.57	0.001	8.78
miR-451	2.07	7.54	0.002	11.77
miR-143-5p	-1.63	5.48	0.004	23.02
miR-10a-5p	-1.01	8.41	0.006	24.60
48hr TAC vs Sham				
miRNA	logFC	logCPM	p value	FDR (%)
miR-451	2.74	7.36	2.91E-04	7.24
miR-21-5p	1.94	13.61	0.001	11.24
7 day TAC vs Sham				
miRNA	logFC	logCPM	p value	FDR (%)
miR-21-5p	3.24	14.12	1.9402E-08	5.06E-04
miR-34c-5p	3.11	4.27	1.04E-04	1.36
miR-150-5p	-1.61	10.31	0.003	26.3
7 day TAC vs remaining samples				
miRNA	logFC	logCPM	p value	FDR (%)
miR-34c-5p	2.52	3.57	1.25623E-07	3.27E-03
miR-214-3p	2.03	6.77	2.03236E-06	0.03
miR-199a-5p	1.53	9.25	2.23636E-05	0.19
miR-199b-3p	1.46	10.61	6.79032E-05	0.37
miR-199a-3p	1.46	10.60	7.03649E-05	0.37
miR-299-5p	1.30	3.94	8.36E-04	3.62
miR-379-5p	1.18	3.74	0.002	8.25
miR-101a-3p	-1.45	10.29	0.003	8.25
miR-150-5p	-1.26	10.27	0.005	13.46
miR-31-5p	1.38	8.98	0.007	17.59
miR-410-3p	1.09	3.56	0.007	17.59
miR-224-5p	1.00	3.87	0.011	23.41
miR-21-5p	1.55	13.83	0.012	23.41
48 hr TAC vs remaining samples				
miRNA	logFC	logCPM	p value	FDR (%)
miR-324-3p	-4.06	3.89	9.29671E-05	2.42

Log is log2; FC, fold change; CPM, normalised counts per million; FDR; false discovery rate.

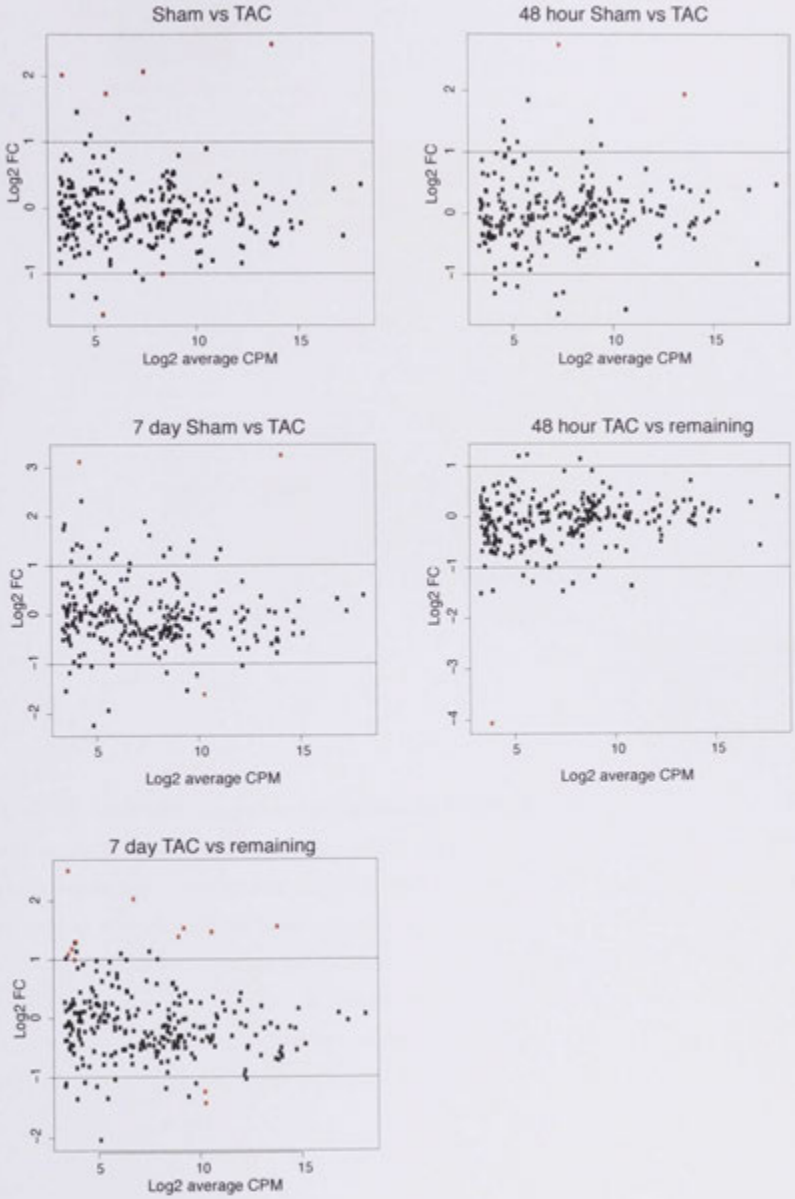


Figure 4.6. Identification of differentially expressed miRNAs.

Log2-fold change (FC) is plotted against the Log2- average counts per million (CPM) for the pairwise comparisons used to identify differentially expressed miRNAs. The lines indicate a log2-FC of 1 and the miRNAs identified as differentially expressed with a FDR <30% are shown in red.

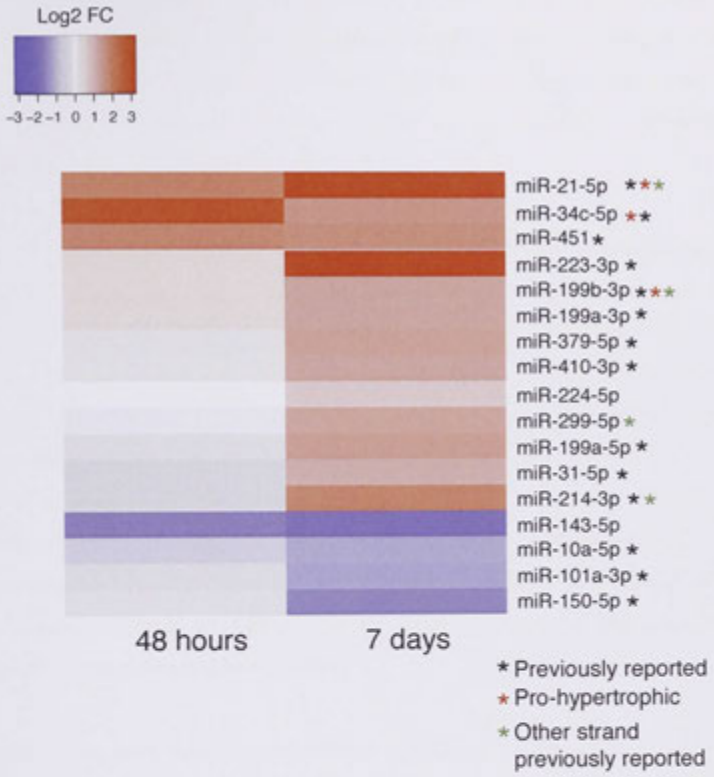


Figure 4.7. Differentially expressed miRNAs with a FDR <30%.

Pairwise comparisons were made with edgeR package and miRNAs with a FDR <30% were considered differentially expressed. miRNAs with a black star have been previously reported to be regulated in the same direction in a model of hypertrophy. miRNAs with a red star are pro-hypertrophic (functional studies have shown they promote the hypertrophic response). miRNAs with a green star indicate that the other strand of the miRNA hairpin has been previously reported to be regulated in the same direction in a model of hypertrophy. Data is shown as TAC/Sham log2 fold change (average of duplicate samples) at each time point.

Intriguingly, 13 of the miRNAs were upregulated and only 4 were downregulated. As shown in Table 4.1, more miRNAs were differentially regulated at 7 days compared to 48 hours. In fact, only miR-21-5p and miR-451[a] had significant changes at 48 hours. This suggests that the induction and processing of most miRNAs may take longer than 48 hours to reach a significant level of change.

Reassuringly, 13 of the 17 differentially expressed miRNAs have been previously reported to change levels in the same direction. These changes were reported in similar TAC experiments (*120, 121, 127, 128, 130, 131*), PE-treated neonatal cardiomyocytes (*121*), transgenic mice with early and end-stage hypertrophic cardiomyopathy (*129*), as well as human patients with DCM or ICM (*123, 124*) and patients with end-stage heart failure (*122*). For instance, miR-214-3p, which is highly upregulated 7 days post-TAC, is well documented to be upregulated at this point in several other studies (*127, 128, 130, 131*). Its expression also increases in transgenic mice with hypertrophic cardiomyopathy (*129*), as well as in patients with DCM, ICM and AS (*124*). Detection of miRNAs previously found to alter expression levels in hypertrophy provides further evidence that the approach to find differentially expressed miRNAs was valid.

For four miRNAs, the other strand on the miRNA hairpin has been previously reported to be regulated in hypertrophy. For instance, miR-299-3p is upregulated 7 days after TAC (*130*), but in our data 99% of miR-299 is derived from the 5p arm and miR-299-3p does not make our threshold for expression. Similarly, both strands of miR-199a are significantly upregulated 7 days post-TAC, suggesting the hairpin is upregulated at the transcriptional level. The detection of a different strand than previously reported, such as miR-299-3p, could indicate a difference in the processing of the miRNA hairpin.

Interestingly, three pro-hypertrophic miRNAs (miR-21-5p, miR-34c-5p and miR-199b-3p) are strongly upregulated at 7 days, with both miR-34c-5p and miR-21-5p being upregulated as early as 48 hours (*225-227*). This is consistent with the physiological data that suggests that at 7 days the hearts are hypertrophic. The upregulation of miR-34c-5p and miR-21-5p at 48 hours may form part of the initial response to pressure overload that results in cardiac hypertrophy.

4.2.4.A miRNA validation using RT-qPCR

The expression of the miRNAs in Figure 4.7 was tested using RT-qPCR. As mentioned in Chapter Three, a total of 20 cardiomyocyte samples were originally isolated for RNA extraction (2 Sham and 2 TAC for each time-point). Only eight samples were subjected to next-generation sequencing but all samples were used for RT-qPCR validation. miRNA expression was detected using stem-loop reverse transcription (see section 2.2.4 for a description). The stem-loop primers are more efficient and provide better specificity for miRNAs than conventional reverse transcription (*228*). Furthermore, as the reverse transcription is a different method than the preparation of the next-generation sequencing library it should not validate changes based on sequencing bias alone. The miRNA expression was normalised to the expression of five small nucleolar RNAs (snoRNAs); snoRNA-135, snoRNA-202, snoRNA-55, snoRNA-412 and snoRNA-234. snoRNAs are better for miRNA normalisation than control genes such as GAPDH, as they are closer in length (60-300 nt) to miRNAs than mRNA transcripts and they are constitutively expressed across multiple tissues. Furthermore, the levels of snoRNAs are detected using the same stem-loop reverse transcription and RT-qPCR reactions as miRNAs. In addition, it has been previously shown that snoRNA-202 and snoRNA-234 show the least variability across tissues and cell lines (*229*). The expression of the snoRNAs was consistent across all 20 cardiomyocyte samples (Figure 8.1) and as a result, the average of the five snoRNAs was used for normalisation. Rina Soetanto from the ANU performed the RT-qPCRs for four of the miRNAs.

Of the eight miRNAs considered for testing the expression changes of six followed the same trend in the samples subjected to NGS (Figure 4.8.A-B). The levels of four of these could be validated by RT-qPCR in the extended cardiomyocyte dataset (Figure 4.8.A). miRNA expression changes were considered as validated if a two-way ANOVA with Bonferoni's post-hoc analysis on the RT-qPCR data resulted in an adjusted p value of less than 0.05. The levels of miR-31-5p, miR-199a-3p, miR-21-5p and miR-223-3p were significantly upregulated in the extended cardiomyocyte sample set for 7 days. miR-223-3p was also significantly upregulated at 48 hours. On the contrary, the change in levels of miR-451 could be seen in the samples used for sequencing, but this increase in expression level was not seen when looking at the larger pool of cardiomyocyte samples. Similarly, miR-143-5p was decreased by more than 2-fold at both time points according to the sequencing

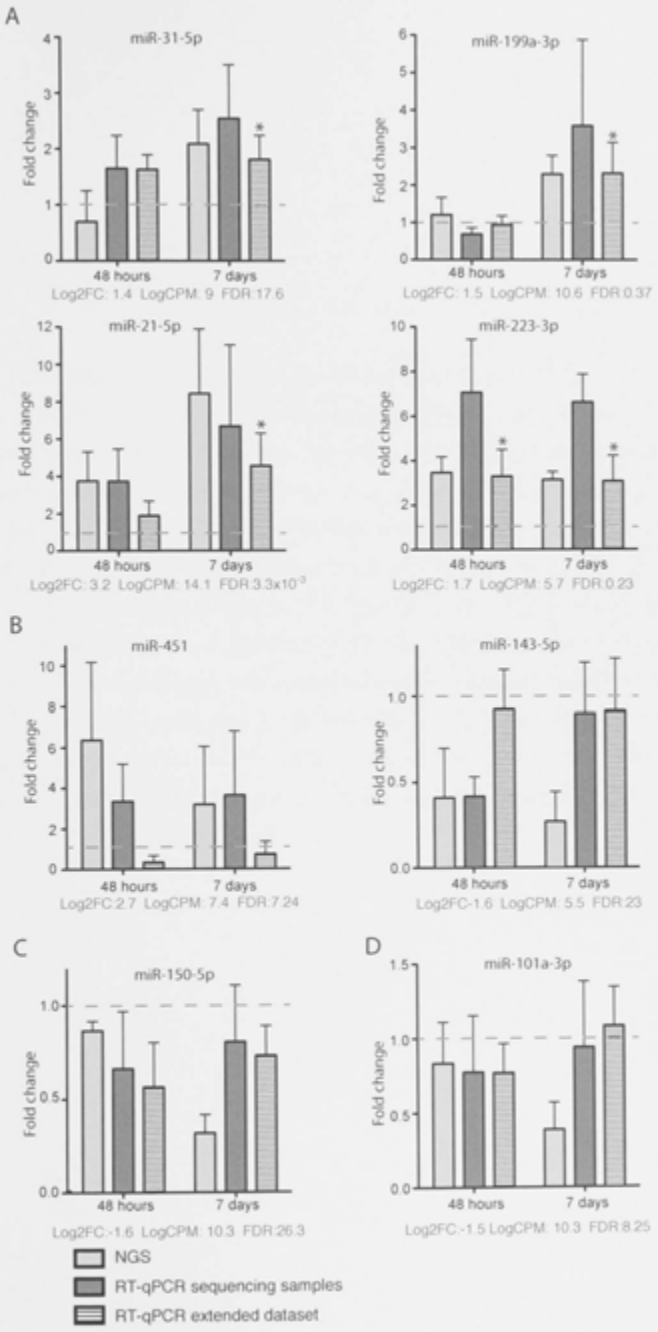

A
miR-31-5p
Fold change
48 hours
7 days
Log2FC: 1.4 LogCPM: 9 FDR:17.6
miR-199a-3p
Log2FC: 1.5 LogCPM: 10.6 FDR:0.37
miR-21-5p
Log2FC: 3.2 LogCPM: 14.1 FDR:3.3x10-3
miR-223-3p
Log2FC: 1.7 LogCPM: 5.7 FDR:0.23
B
miR-451
Log2FC:2.7 LogCPM: 7.4 FDR:7.24
miR-143-5p
Log2FC-1.6 LogCPM: 5.5 FDR:23
C
miR-150-5p
Log2FC:-1.6 LogCPM: 10.3 FDR:26.3
D
miR-101a-3p
Log2FC:-1.5 LogCPM: 10.3 FDR:8.25
NGS
RT-qPCR sequencing samples
RT-qPCR extended dataset

Figure 4.8. Validation of differentially expressed miRNAs by RT-qPCR.

Fold change (TAC/Sham) values for the next-generation sequencing data (NGS), RT-qPCR on the eight samples subjected to next generation sequencing (RT-qPCR sequencing samples) and RT-qPCR on the extended cardiomyocyte sample set (RT-qPCR extended dataset). (A) miRNA expression changes could be validated by RT-qPCR in both the samples used for sequencing and the extended dataset. (B) miRNA expression changes could be validated in the samples used for sequencing but not the extended dataset. (C) The decrease in miR-150-5p could be observed in the extended dataset but did not reach significance. (D) The decrease in miR-101a-3p could not be observed in the sequencing samples or the extended dataset. The dashed grey line indicates a fold change of 1. Data shown as mean fold change with SEM. (n=4-5 for RT-qPCR sequencing sample and RT-qPCR extended dataset, n=2 for NGS * $p < 0.05$, Two way-ANOVA with Bonferonni's post-hoc analysis). The log2FC, log counts per million (CPM) and false discovery rate (FDR) for the NGS data are shown below each graph.

data. RT-qPCR on the samples used for sequencing also showed the dramatic decrease at 48 hours but only a 20% decrease at 7 days. Furthermore, the extended dataset indicated a minor 10% drop in the levels of miR-143-5p at both 48 hours and 7 days that did not reach significance. A slight decrease in the levels of miR-150-5p was seen in the extended dataset at 7 days but this too was not significant (Figure 4.8.C). Lastly, the levels of miR-101a-3p showed no difference at 7 days in the RT-qPCR data in comparison to the decrease seen in the NGS data (Figure 4.8.D). RT-qPCR primers were also designed against the major isoforms of miR-34c-5p, miR-299-5p, miR-379-5p, miR-214-3p and miR-10a-5p for validation (primers provided in Methods). Unfortunately, the RT-qPCR assays for the above miRNAs could not be optimised to detect a specific product and these miRNAs did not go forward for validation. In summary, the same directional trend could be observed for seven of the eight miRNAs tested, with four reaching statistical significance in the extended cardiomyocyte sample set.

4.2.4.B Previously documented miRNAs

Numerous miRNAs have already been documented to be deregulated in the hypertrophic response, with some miRNAs consistently reported among studies. Therefore, we examined the expression levels of a 53 miRNAs that have been reported to change in more than two studies (listed below) across the time course regardless of their FDR (Figure 4.9). These miRNAs were derived from similar studies that looked at the expression of miRNAs across various experimental models and patients with heart disease. These studies include LV isolated after TAC experiments (*120, 121, 127, 128, 130, 131*), PE-treated neonatal cardiomyocytes (*121*), LV from transgenic mice with early and end-stage hypertrophic cardiomyopathy (*129*), and human patients with DCM or ICM (*123, 124*) and patients with end-stage heart failure (*122*). For majority of the miRNAs, the trend remains the same but the miRNAs have not made the cut-off for statistical significance. Importantly, pro-hypertrophic miRNAs are mostly upregulated at 7 days, while anti-hypertrophic miRNAs are downregulated. There are 12 miRNAs that are documented to be upregulated whose levels increase by more than 1.5 fold at either 48 hours or 7 days, five of which have a FDR less than 30% and are present in Figure 4.8. The remaining seven miRNAs increase in expression but have a high FDR. For example, both strands of miR-208b, miR-208b-5p and miR-208b-3p, appear to have increased levels at 7 days. On closer inspection, miR-208b-3p has very low levels in one of the 7 day Sham samples that is driving the

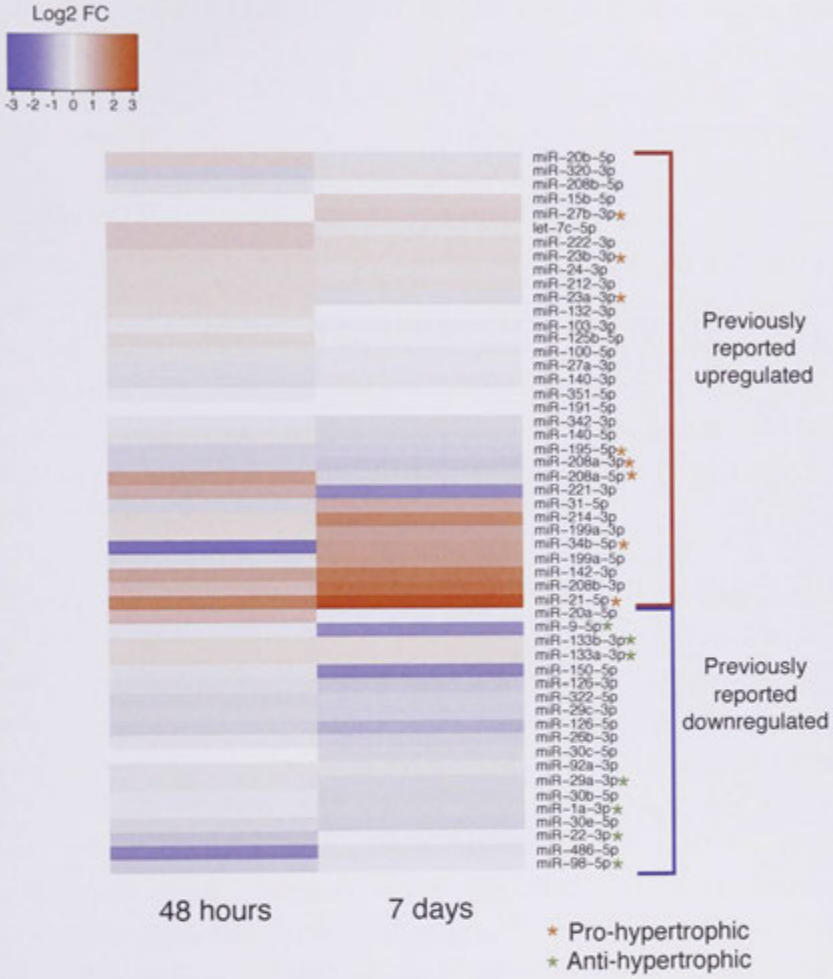


Figure 4.9. miRNAs previously reported to be regulated in hypertrophy show similar expression patterns.

Data is shown as TAC/Sham log-2 fold change at each time point, for miRNAs previously reported to be up- or down-regulated in hypertrophy. These studies include TAC experiments (*120, 121, 127, 128, 130, 131*), PE-treated neonatal cardiomyocytes (*121*), transgenic mice with early and end-stage hypertrophic cardiomyopathy (*129*), and human patients with DCM or ICM (*123, 124*) and patients with end-stage heart failure (*122*). The red star indicates miRNAs that promote or induce hypertrophy, while the green star indicates miRNAs that can prevent or suppress the hypertrophic response.

upregulation in the 7 day TAC samples. Furthermore, miR-208b-5p is barely detectable in the dataset and does not make our threshold of 10 CPM. There are no tags representing miR-208b-5p in 5 of the samples, and only 5 tags in one of the 7 day TAC samples making it appear that there is a strong upregulation at 7 days. There are also a large number of miRNAs that have been reported to be upregulated, but for which we see no change (miR-27b-3p to miR-191-5p in Figure 4.9). For instance, miR-23a-3p is reported to be upregulated in TAC banded animals (*120, 127*), PE-treated neonatal CM (*121*) as well as in patients with DCM, ICM and AS (*123, 124*). Yet miR-23a-3p has very small fold change values in our experiment. This could be due to differences in experimental design, such as the time point investigated after hypertrophy, as well as the platform on which the measurements were made. In line with this, a recent paper that used next-generation sequencing to examine miRNAs 7 days post-TAC actually used ten miRNAs that remained unchanged in their data as normalisers (*131*). This included miR-23a-3p, miR-24-3p and miR-23b-3p which also remain unchanged in our dataset.

All of the miRNAs previously reported to be downregulated, with the exception of miR-133a-3p and miR-133b-3p, have lower expression levels either at 48 hours or at 7 days. miR-22-3p, miR-486-5p and miR-98-5p look to be strongly downregulated at 48 hours. However, variation between biological replicates confounds the fold change values and results in a high FDR. For instance, the levels of miR-486-5p appears to decrease 3-fold at 48 hours but the FDR is 64% because the 48 hour TAC 1 sample has considerably lower levels. miR-133a-3p and miR-133b-3p show little to no regulation at 7 days, but these miRNAs were also used as normalisers in the study mentioned above as their expression levels did not change between the control and the TAC banded mice 7 days after surgery (*131*). In agreement with the previously upregulated miRNAs, the downregulated miRNAs did not reach significance due to variation between biological replicates or the levels remained the same between the Sham and TAC samples. To summarise this comparison, 23 miRNAs previously reported to be up- or down-regulated in other studies showed the same trend in our data with an absolute fold-change greater than 1.5, four miRNAs did not make our 10 CPM threshold for expression level, three were regulated in the opposite direction and 21 did not change expression in response to pressure overload.

In conclusion, we have detected a small number of differentially expressed miRNAs that change levels between the Sham and TAC samples. Only 50% of those tested for expression could be validated in the extended sample set of cardiomyocytes, mostly due to

variation in the levels between biological replicates. Despite this, several of the miRNAs we detected have been previously reported to change in the same direction, suggesting we are detecting changes consistent with other studies.

4.2.5 miRNA processing variants

One of the main aims of this thesis was to look for miRNA processing variants. The processing variants of miRNAs in purified cardiomyocytes, let alone cardiomyocytes from pre and hypertrophic hearts, have not been documented thus far. As mentioned earlier, one miRNA hairpin can result in several miRNA variants, potentially with altered mRNA target specificity. This can arise from use of both strands of the miRNA:miRNA star duplex with similar frequency, the occurrence of 5' and 3' isomiRs as well as RNA editing. It is postulated that these miRNA modifications could alter miRNA biogenesis and half-life, subcellular localisation and/or targeting specificity. For this reason, we interrogated all of the mature miRNAs that had expression levels greater than 10 CPM for the presence of processing variants. Once a static picture had been generated, Dr Brian Parker from the ANU applied a weighted linear model to each type of processing variant looking at the proportion of variants for each miRNA.

4.2.5.a Arm bias

We first looked at the arm bias of miRNAs, which is the preference of the miRNA to be derived from one or both strands of the miRNA hairpin (Figure 4.10.A). Appendix 8.8 documents the proportion of 5' arm bias for all miRNAs with greater than 10 CPM. There were 37 miRNAs that had greater than 20% of their tags derived from each strand of the miRNA hairpin, meaning both strands could function to target mRNAs (Table 4.2). For example, let-7d has 52.2% 5' arm hairpin bias, meaning that 52% of let-7d is derived from the 5p arm and 48% is derived from the 3p arm. Critically, this list includes 15 miRNAs with known expression and/or function in the heart. This data becomes important when interpreting the functional studies associated with a miRNA locus.

Changes in arm bias in hypertrophy

It is known that miRNAs can switch strands in a tissue specific manner (*79, 80*) so we hypothesised that the strands may also change under conditions of cellular stress like hypertrophy. Once we had established that numerous miRNAs express both strands of the hairpin, the data was interrogated to look for examples where this proportion changed

across the hypertrophy time course. In doing so, 35 miRNAs were identified where the proportion of arm bias changed in range by more than 20% across the eight cardiomyocyte samples (blue datapoints in Figure 4.10.B, Table 4.3). This included 19 miRNAs that had previously been reported to change expression in hypertrophy and/or have a known role in the heart (Figure 4.11). 14 of which also have more than 20% derived from both strands at any one time (Table 4.1). However, on closer inspection of some individual examples, it appears that there is no directionality to the changes in arm bias proportions and the large range is due to variation in the biological replicates (Figure 4.11). To confirm that there was no change, a weighted linear model looking at the proportion statistic for each miRNA hairpin was applied to the data (the results for all processing variants are shown in Appendix 8.13). This did not result in detection of any miRNA hairpin bias change with a FDR <80%. It can thus be concluded that we did not detect any change in arm bias between the Sham and TAC samples.

The majority of miRNAs give rise to one predominant strand, with only 12.5% of the miRNA hairpins detected giving rise to both strands simultaneously. In addition, it is evident from this data that while there is fluctuation in the arm bias for 35 miRNA hairpins it does not correlate with the hypertrophic response.

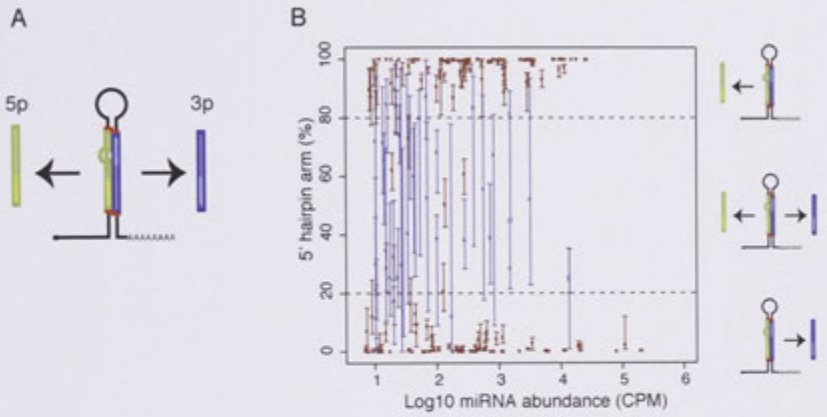


Figure 4.10. MiRNAs exhibit arm bias in cardiomyocytes.

(A) MiRNA arm bias; the miRNA can be derived from both strands of the miRNA hairpin in varying proportions. (B) The mean and range of 5p strand bias of each miRNA. MiRNAs with average tag count greater than 10 CPM were used for analysis. MiRNAs with a range less than 20% are shown in red, those with a range greater than 20% are indicated by blue. Diagrams to the side of the graph indicate the strand bias for that portion of the graph.

Table 4.2. miRNA hairpin precursors with >20% of tags derived from both the 5p and 3p arm.

miRNA	Arm bias (5p tag counts/hairpin tag counts)					
	Mean	Range	48 hour Sham	48 hour TAC	7 day Sham	7 day TAC
let-7d	52.2	70.75	45.02	39.50	73.20	51.09
let-7i	38.93	34.13	32.24	36.11	42.11	45.24
miR-126#	24.66	15.35	29.23	23.50	24.10	21.82
miR-140#	38.14	23.78	36.01	42.40	41.48	32.66
miR-151#	67.22	72.22	77.10	79.75	67.13	44.91
miR-154#	71.99	71.42	54.00	76.54	91.91	65.51
miR-1839	62.15	12.62	66.19	56.90	63.50	62.02
miR-1843	75.14	35.27	77.16	65.67	79.68	78.06
miR-199a#	28.45	23.12	34.49	24.67	27.05	27.57
miR-208a#*	55.56	69.84	35.66	72.93	55.50	58.14
miR-322#	60.75	12.05	62.01	59.85	61.32	59.79
miR-330#	30.19	64.31	32.53	44.25	36.94	7.05
miR-335#	68.13	21.48	74.54	60.64	75.64	61.69
miR-337#	29.97	26.46	28.79	34.83	24.08	32.18
miR-338	32.18	11.71	28.72	31.06	34.61	34.33
miR-362	65.47	31.05	72.85	64.80	56.77	67.46
miR-378#	25.34	34.32	29.62	14.32	28.50	28.92
miR-423#	38.73	37.86	45.12	43.57	43.16	23.08
miR-434#	28.76	16.57	29.15	32.91	24.46	28.53
miR-490	34.67	59.10	45.35	62.43	17.14	13.75
miR-501	71.49	36.40	83.96	65.23	69.16	67.63
miR-503	25.54	27.08	24.76	16.05	27.91	33.45
miR-5117	25.35	38.52	22.16	35.36	27.14	16.72
miR-532#	59.97	42.60	66.08	59.07	66.84	47.89
miR-542	21.98	20.54	19.24	30.11	21.47	17.09
miR-582	50.43	16.35	55.42	46.97	52.19	47.12
miR-652	29.63	96.77	9.91	9.71	47.77	51.12
miR-674#	52.61	64.81	42.59	35.18	68.54	64.16
miR-676	20.85	54.77	16.71	21.64	34.15	10.87
miR-700	73.15	18.02	69.03	71.58	71.84	80.17
miR-744	67.52	52.22	72.06	51.80	69.83	76.37
miR-7a-1	20.7	16.37	23.90	26.15	14.95	17.79
miR-872	74.07	20.10	71.71	73.79	77.31	73.45
miR-874	40.54	78.25	24.05	30.70	55.55	51.85
miR-9*	32.51	48.85	27.11	29.13	43.20	30.60
miR-98*	79.71	56.53	87.92	61.10	85.82	84.01

miRNAs with known expression changes in hypertrophy. * miRNAs with known function in the heart. miR-199a includes miR-199a-1 and miR-199a-2. miR-9 includes miR-9-1, miR-9-2 and miR-9-3.

Table 4.3. miRNA hairpins with a range of arm bias proportion greater than 20% across the cardiomyocyte samples.

miRNA	Arm bias (5p tag counts/hairpin tag counts)					
	Mean	Range	48 hour Sham	48 hour TAC	7 day Sham	7 day TAC
let-7d	52.2	70.8	45.0	39.5	73.2	51.1
let-7i	38.9	34.1	32.2	36.1	42.1	45.2
miR-140#	38.1	23.8	36.0	42.4	41.5	32.7
miR-151#	67.2	72.2	77.1	79.7	67.1	44.9
miR-154#	72.0	71.4	54.0	76.5	91.9	65.5
miR-1843	75.1	35.3	77.2	65.7	79.7	78.1
miR-1843b	84.6	25.3	86.6	78.7	85.5	87.4
miR-18a	90.0	23.0	95.3	84.9	88.3	91.5
miR-193#	11.5	21.9	9.5	10.4	10.3	15.5
miR-199a#	28.5	23.1	34.5	24.7	27.1	27.6
miR-208a#*	55.6	69.8	35.7	72.9	55.5	58.1
miR-210#	12.0	77.4	1.3	39.1	4.8	2.7
miR-31#	83.4	58.1	87.8	63.5	89.2	93.0
miR-324#	87.1	27.0	92.5	98.6	76.0	81.4
miR-330#	30.2	64.3	32.5	44.2	36.9	7.0
miR-335#	68.1	21.5	74.5	60.6	75.6	61.7
miR-337#	30.0	26.5	28.8	34.8	24.1	32.2
miR-362	65.5	31.1	72.9	64.8	56.8	67.5
miR-378#	25.3	34.3	29.6	14.3	28.5	28.9
miR-423#	38.7	37.9	45.1	43.6	43.2	23.1
miR-486#	81.2	43.6	87.0	67.2	84.8	86.0
miR-490	34.7	59.1	45.4	62.4	17.1	13.7
miR-501	71.5	36.4	84.0	65.2	69.2	67.6
miR-503	25.5	27.1	24.8	16.1	27.9	33.4
miR-5117	25.4	38.5	22.2	35.4	27.1	16.7
miR-532#	60.0	42.6	66.1	59.1	66.8	47.9
miR-542	22.0	20.5	19.2	30.1	21.5	17.1
miR-652	29.6	96.8	9.9	9.7	47.8	51.1
miR-674#	52.6	64.8	42.6	35.2	68.5	64.2
miR-676	20.9	54.8	16.7	21.6	34.2	10.9
miR-744	67.5	52.2	72.1	51.8	69.8	76.4
miR-872	74.1	20.1	71.7	73.8	77.3	73.5
miR-874	40.5	78.3	24.1	30.7	55.5	51.8
miR-9*	32.5	48.8	27.1	29.1	43.2	30.6
miR-98*	79.7	56.5	87.9	61.1	85.8	84.0

miRNAs with known expression changes in hypertrophy. * miRNAs with known function in the heart. miR-199a includes miR-199a-1 and miR-199a-2. miR-9 includes miR-9-1, miR-9-2 and miR-9-3.

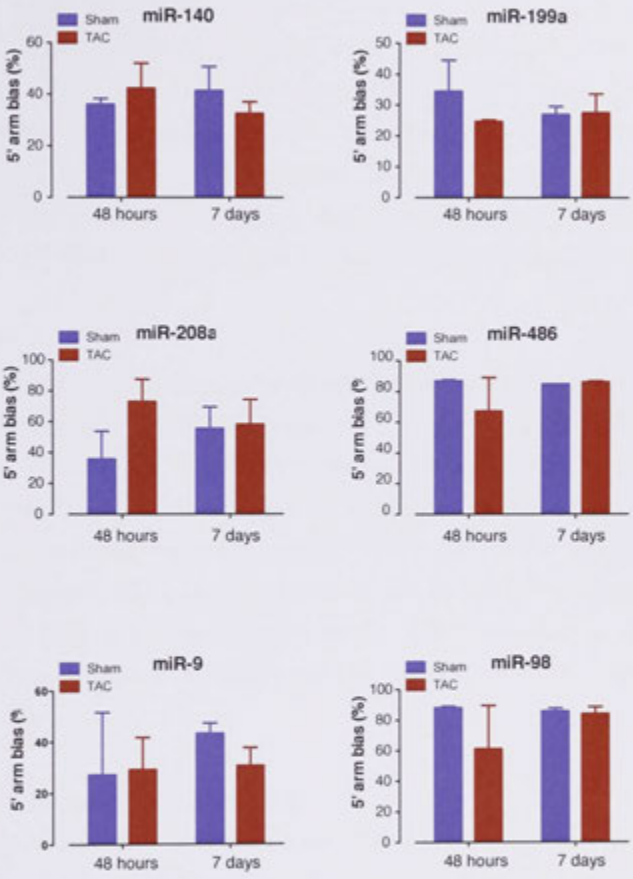


Figure 4.11. Cardiac miRNAs with arm bias across the hypertrophic samples.

The 5' arm bias is shown for several cardiac miRNA with more than 20% range across the eight samples. The arm bias represents the amount the 5' or 5p arm contributes to the miRNA hairpin, with 0% indicating all of the miRNA is derived from the 3p arm and 100% indicating all of the miRNA is derived from the 5p arm. Data is presented as the mean and SEM of 5' arm bias (%) derived from the small RNA sequencing data (n=2).

4.2.5.B 5' AND 3' ISOMIRS

Next-generation sequencing also allows the examination of a miRNA strand for variation at the start or end position, termed 5' and 3' isomiRs respectively. Unlike arm bias, the presence of isomiRs represents a subtler variation. IsomiRs will bind to the same microarray probe and will often be detected with the same RT-qPCR primers as the canonical miRNA because the majority of the sequence remains the same. Next-generation sequencing can detect these subtle differences, as the sequence of the adaptors originally ligated onto the miRNA is known. Thus, the exact start and end position of the miRNA can be established. We have mined the data for miRNAs with evidence of 5' and 3' isomiRs. We have defined isomiRs as a miRNA tag that starts/ends within 3nt upstream or downstream of the miRBase canonical defined start/end site (Figure 4.12.A).

5' isomiRs

As shown in Figure 4.12.B, the majority of miRNAs start where expected, with close to 0% being derived from a 5' isomiR (the proportion of 5' isomiRs for all miRNAs with CPM greater than 10 are listed in Appendix 8.8). However, 55 miRNAs have more than 20% of their tags derived from a 5' isomiR (Table 4.4). Importantly, these miRNAs have a large range of tag counts indicating that the presence of 5' isomiRs is not confined to only lowly expressed and potentially spurious miRNAs. This list includes several cardiac miRNAs, such as miR-133a, miR-9 and miR-98, which act as anti-hypertrophic miRNAs by supressing the expression of transcripts that play into the hypertrophic signalling pathway (*140, 141, 230*).

Variation in 5' isomiRs during hypertrophy

After determining that 5' isomiRs were detectable in 55 miRNAs, we wanted to know if the proportion of 5' isomiRs for each miRNA changed between the Sham and TAC cardiomyocytes. Only eight miRNAs have a range greater than 20%, suggesting that while several miRNAs have a large proportion of isomiRs, the processing remains static between the control and the hypertrophic hearts (Table 4.5). The contribution of 5' isomiRs at each time point for these eight miRNAs is shown in Figure 4.13. In support of this observation, analysis using the weighted linear model did not yield any miRNAs with a FDR < 60% (Appendix 8.13). As seen with the variation in arm bias, we could not detect any directional change in 5' isomiRs with hypertrophy induction across the time course.

3' isomiRs

In contrast to 5' isomiRs, 248 out of the 260 miRNAs with coverage greater than 10 CPM have more than 20% 3' isomiRs (Figure 4.12.C). Further information about the 3' isomiR percentage for each of these miRNAs can be found in Appendix 8.8. As majority of the miRNAs have 3' isomiRs we did not interrogate the miRNAs for their individual involvement with cardiac hypertrophy.

Changes in 3' isomiRs during hypertrophy

The 3' isomiR proportions were then analysed for any change between the Sham and TAC samples. The percentage of 3' isomiRs changes more than 20% in 50 miRNAs across the time course (Table 4.6). This includes 16 miRNAs whose expression has been documented to change in hypertrophy or with known functional roles in the heart. After applying the weighted linear model to look for changes in the proportion of 3' isomiRs, one miRNA with a FDR of 31% was identified with a difference between the Sham and TAC samples at 48 hours (Appendix 8.13). miR-125a-5p exists as three predominant 3' isomiRs, the canonical miR-125a-5p (Can miR-125a-5p), a 3' isomiR that finished one nucleotide upstream (Iso1 miR-125a-5p) and another 3' isomiR that finished two nucleotides upstream (Iso2 miR-125a-5p) as illustrated in Figure 4.14.A. The proportion of Can miR-125a-5p increases in the 48 hour TAC samples relative to the other 3' isomiRs (Figure 4.14.B.ii). This change in proportion was also visible in the 7 day samples. The expression levels of miR-125a-5p decrease slightly at both time points (Figure 4.14.B.i), however the change in expression is small with the largest absolute log2-FC of 0.3. To sum up, 19% of miRNAs have a difference in the proportion of 3' isomiRs across samples, with one miRNA having a directional change with hypertrophy.

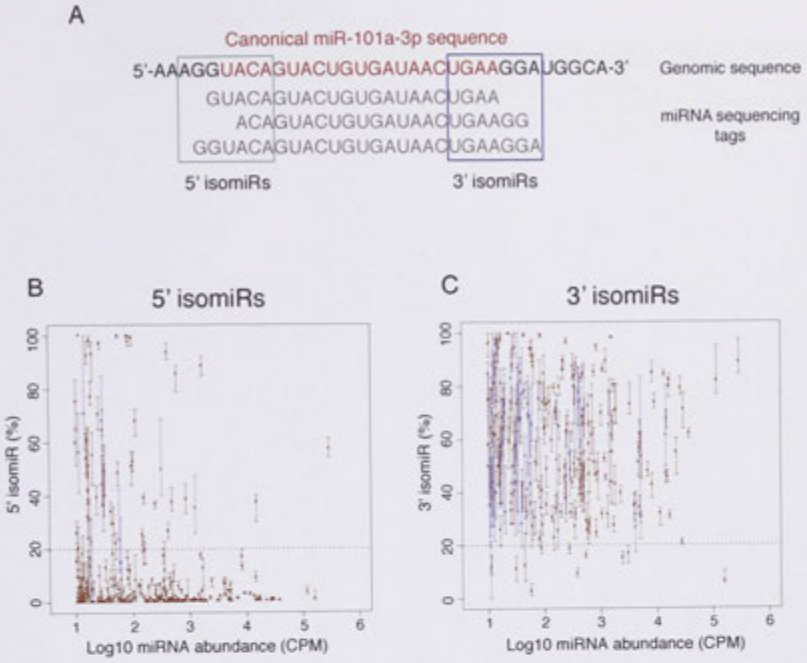


Figure 4.12. 5' and 3' isomiRs are detected in cardiac miRNAs.

(A) IsomiRs were defined as miRNAs that started or ended within ± 3nt from the annotated start position of a miRNA. (B) 5' isomiRs in the hypertrophic heart; the proportion of tags representing 5' isomiRs for each miRNA. (E) 3' isomiRs in the hypertrophic heart; the proportion of tags representing 3' isomiRs for each miRNA. MiRNAs with a range less than 20% are shown in red, those with a range greater than 20% are indicated by blue.

Table 4.4. miRNAs with >20% 5' IsomiRs.

miRNA	Mean CPM	5' IsomiRs (%)					
		Mean	Range	48 hour Sham	48 hour TAC	7 day Sham	7 day TAC
miR-133a-3p#*	292300	57.1	7.2	57.15	59.95	54.64	56.77
miR-126-3p#	15030	37.2	10.1	33.78	36.38	39.17	39.59
miR-345-5p	1721	88.4	7.3	90.65	86.93	88.25	87.61
miR-101a-3p#	1256	35.0	21.6	27.76	30.01	37.00	45.06
miR-29b-3p#	861.1	38.3	11.2	35.82	40.75	34.52	42.26
miR-378b	610	85.5	9.8	83.00	88.35	84.32	86.23
miR-145-3p	490.3	37.3	8.8	37.08	37.21	39.41	35.48
miR-133a-5p*	418.8	93.4	6.2	94.15	91.17	95.37	92.87
miR-30a-3p#	418.6	26.4	5.8	26.76	26.94	24.74	27.27
miR-101b-3p#	322.5	49.7	28.9	43.96	45.31	48.81	60.70
miR-140-3p#	250.1	36.5	3.0	36.50	37.05	36.78	35.67
miR-322-3p	157.1	38.9	3.8	37.40	38.52	39.97	39.73
miR-7a-1-3p	146.6	25.3	6.7	25.81	26.95	23.43	24.81
miR-5105-3p-novel	115.4	67.8	10.1	71.39	70.68	63.69	65.46
miR-582-5p	102.3	52.5	7.3	51.27	55.90	52.83	50.00
let-7g-3p	100.0	98.7	3.5	99.29	98.62	98.72	98.26
miR-582-3p	96.9	50.7	12.6	47.58	50.88	52.76	51.64
miR-5109	86.9	97.6	3.6	96.77	99.03	96.67	97.86
miR-28c	84.6	99.8	0.7	99.75	99.78	99.66	99.84
miR-5097	57.7	23.3	30.8	18.57	27.08	16.17	31.39
miR-101c	57.2	99.8	0.8	99.62	99.64	100.00	99.89
miR-24-1-5p	55.6	48.4	9.0	48.73	45.27	47.33	52.29
miR-101a-5p	50.0	22.1	9.9	19.04	22.12	22.97	24.08
miR-455-3p#	45.6	29.3	10.2	27.50	32.26	30.54	27.05
miR-1843-5p	41.6	33.8	8.5	31.44	31.71	34.39	37.64
miR-504-5p	32.6	49.0	19.3	54.49	43.96	42.81	54.72
miR-3107-3p	31.4	39.6	19.7	35.92	42.18	38.24	41.84
miR-30c-1-3p	30.5	36.4	32.0	38.59	29.59	32.81	44.70
miR-330-3p	30.2	64.4	15.2	66.44	63.28	65.31	62.68

miRNAs with known expression changes in hypertrophy. * miRNAs with known function in the heart. miR-9 includes miR-9-1, miR-9-2 and miR-9-3.

Table 4.4 cont. miRNAs with >20% 5' IsomiRs.

miRNA	Average CPM	5' IsomiRs (%)					
		Average	Range	48 hour Sham	48 hour TAC	7 day Sham	7 day TAC
miR-345-3p	29.9	67.9	22.5	74.41	65.08	65.60	66.34
miR-361-3p	29.6	44.0	19.0	44.04	39.31	45.71	46.97
miR-467b-5p	27.6	96.7	2.8	97.41	96.93	96.78	95.76
miR-503-3p	25.5	70.4	12.3	65.73	72.25	69.55	73.95
miR-3963	23.5	39.2	10.7	43.41	38.78	33.55	41.06
miR-5117-3p	20.2	92.9	14.5	93.28	87.23	96.44	94.82
miR-1843b-5p	19.9	55.0	19.8	57.16	61.67	49.67	51.58
miR-9-3p#*	18.9	77.0	8.5	77.11	74.72	80.71	75.40
miR-208b-3p#*	18.9	22.6	31.7	22.76	29.88	9.62	28.08
miR-5099	18.3	98.0	3.8	97.72	98.71	97.78	97.72
miR-425-3p	17.3	30.7	17.3	35.56	26.35	34.98	25.72
miR-466b-3p	16.8	42.1	14.7	39.93	46.89	42.06	39.38
miR-466c-3p	16.7	41.9	14.2	39.10	46.50	42.43	39.47
miR-466e-3p	16.6	61.1	17.1	62.90	55.84	62.07	63.43
miR-466p-3p	16.6	41.5	14.1	38.88	45.85	41.99	39.22
miR-1949	16.4	92.4	8.5	93.70	92.01	92.50	91.19
miR-29a-5p	16.3	31.9	15.1	35.57	26.35	34.82	30.65
miR-466a-3p	16.2	60.1	18.5	61.85	53.61	61.93	62.80
miR-203-3p	15.5	53.8	15.4	53.58	57.98	49.80	53.88
miR-324-3p#	14.6	70.7	42.9	68.29	83.33	67.06	64.04
miR-1306-5p	11.9	100.0	0.0	100.00	100.00	100.00	100.00
miR-98-3p*	11.5	56.1	29.9	56.35	63.02	47.80	57.35
miR-30d-3p	11.0	20.1	13.3	20.20	19.30	21.94	19.12
miR-337-3p	10.4	75.1	18.8	76.12	78.12	76.46	69.77
miR-107-5p	10.3	25.8	8.5	26.10	27.78	23.36	26.05
miR-505-3p	10.2	59.7	14.1	63.60	58.64	58.69	57.68

miRNAs with known expression changes in hypertrophy. * miRNAs with known function in the heart. miR-9 includes miR-9-1, miR-9-2 and miR-9-3.

Table 4.5. miRNAs with a range >20% 5' isomiRs across the timecourse.

miRNA	Average CPM	5' IsomiR (%)					
		Average	Range	48 hour Sham	48 hour TAC	7 day Sham	7 day TAC
miR-101a-3p#	1256.0	35.0	21.6	27.76	30.01	37.00	45.06
miR-101b-3p	322.5	49.7	28.9	43.96	45.31	48.81	60.70
miR-5097	57.7	23.3	30.8	18.57	27.08	16.17	31.39
miR-30c-1-3p	30.5	36.4	32.0	38.59	29.59	32.81	44.70
miR-345-3p	29.9	67.9	22.5	74.41	65.08	65.60	66.34
miR-208b-3p#*	18.9	22.6	31.7	22.76	29.88	9.62	28.08
miR-324-3p#	14.6	70.7	42.9	68.29	83.33	67.06	64.04
miR-98-3p*	11.5	56.1	29.9	56.35	63.02	47.80	57.35

miRNAs with known expression changes in hypertrophy. * miRNAs with known function in the heart

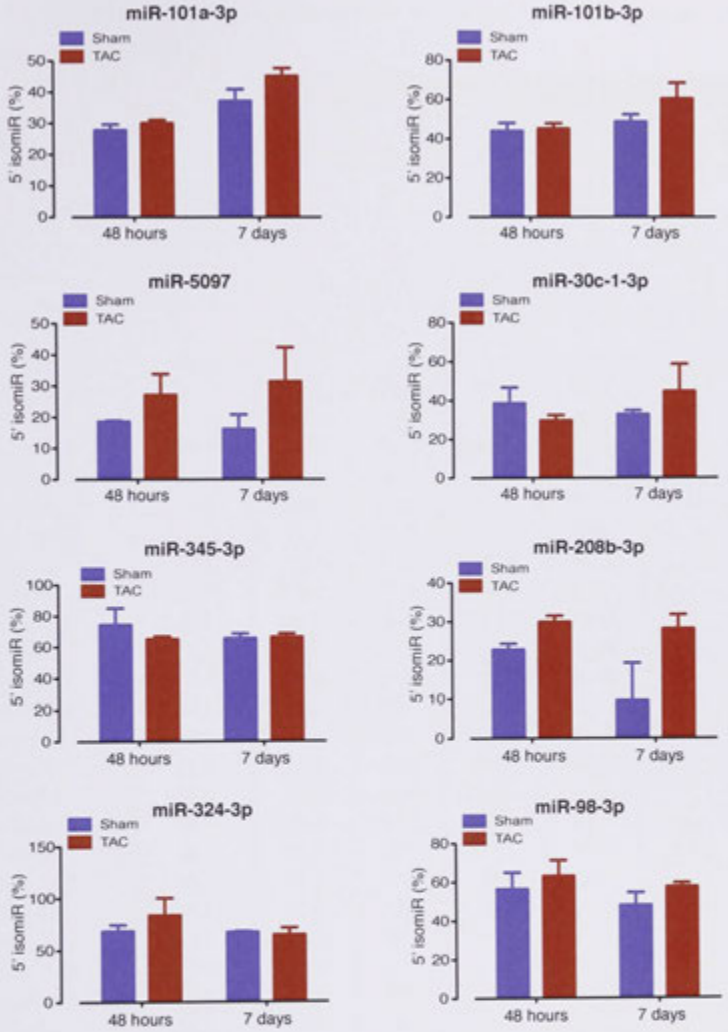


Figure 4.13. miRNAs with variation in the proportion of 5' isomiRs across the cardiomyocyte samples.

The contribution of 5' isomiRs for each miRNA with more than 20% difference in 5' isomiRs across the eight samples. Data is presented as mean and SEM of 5' isomiR (%) derived from the small RNA sequencing data (n=2).

Table 4.6. miRNAs with >20% difference in the proportion of 3' isomiRs across the cardiomyocyte samples.

miRNA	Average CPM	3' IsomiR (%)					
		Average	Range	48 hour Sham	48 hour TAC	7 day Sham	7 day TAC
miR-21-5p#*	14550.0	61.6	32.3	58.52	78.33	50.08	59.48
miR-378-5p	5161.0	30.8	61.4	22.57	53.52	23.22	24.01
miR-27b-3p#*	4179.0	26.7	22.7	26.29	20.85	36.02	23.58
miR-125a-5p	1429.0	65.4	36.3	77.47	59.98	76.02	48.14
miR-31-5p#	504.6	70.6	28.4	67.42	82.77	67.99	64.28
let-7i-5p	436.2	74.6	22.9	79.42	73.45	77.65	67.75
miR-28-5p#	418.4	73.9	40.5	70.46	81.07	81.05	62.88
miR-106b-5p#	407.1	77.6	22.0	77.24	82.57	71.32	79.35
miR-320-3p#	357.5	75.0	36.7	69.42	80.55	68.71	81.23
miR-146a-5p#	281.5	54.6	22.6	58.00	58.73	57.55	43.97
miR-210-3p#	244.9	46.3	30.7	42.94	50.92	51.98	39.50
miR-450a-5p	173.0	25.6	21.3	26.57	33.04	21.37	21.48
miR-93-5p#	167.9	52.4	43.9	52.37	53.61	56.88	46.67
miR-324-5p	84.9	34.2	37.0	26.85	35.59	46.18	28.18
miR-190-5p	80.4	75.2	21.2	74.19	80.35	66.46	79.92
miR-127-3p#	61.1	65.7	20.8	67.40	61.67	63.09	70.55
miR-101c	57.2	43.3	47.8	61.58	53.28	34.59	23.70
miR-24-1-5p	55.6	68.8	25.8	72.77	64.60	75.85	61.84
miR-497-5p#	51.7	69.5	37.2	74.66	70.22	74.69	58.47
miR-101a-5p	50.0	41.4	20.6	36.55	46.34	41.12	41.69
miR-744-5p	45.7	53.9	31.8	59.70	42.46	58.16	55.20
miR-1843-5p	41.6	77.3	22.8	73.42	83.66	77.79	74.43
miR-532-5p	39.1	28.2	24.9	27.41	19.57	34.83	30.95
miR-15b-3p	36.5	67.7	23.0	67.99	70.72	69.42	62.79
miR-874-3p	35.3	74.9	21.8	70.66	68.70	72.20	87.87
miR-652-3p	34.1	87.0	75.0	61.13	92.75	97.79	96.28
miR-504-5p	32.6	44.1	30.4	52.17	41.26	40.93	42.13
miR-30c-1-3p	30.5	43.2	22.6	39.28	50.61	40.61	42.19
miR-330-3p	30.2	58.3	27.1	54.27	70.86	53.09	55.08
miR-18a-5p	30.1	40.6	26.4	46.07	32.63	42.84	40.82
miR-671-5p#	26.2	83.1	21.3	76.23	85.59	86.31	84.31
miR-500-3p	20.2	73.3	26.6	75.75	72.16	80.28	64.92
miR-5117-3p	20.2	32.7	25.6	33.38	46.30	24.46	26.50

miRNAs with known expression changes in hypertrophy. * miRNAs with known function in the heart

Table 4.6. cont. miRNAs with >20% difference in the proportion of 3' isomiRs across the cardiomyocyte samples

miRNA	Average CPM	3' IsomiR (%)					
		Average	Range	48 hour Sham	48 hour TAC	7 day Sham	7 day TAC
miR-208b-3p#*	18.9	59.2	50.1	65.46	64.88	60.53	45.75
miR-1839-5p	17.5	58.1	20.9	64.83	59.17	60.88	47.41
miR-466a-3p	16.2	42.4	21.8	34.80	43.86	47.59	43.18
miR-17-3p	16.2	53.9	34.2	66.24	53.41	42.50	53.40
miR-490-3p	15.2	89.1	23.5	79.91	92.09	92.48	91.91
miR-874-5p	13.6	84.3	25.8	82.36	90.96	79.03	84.88
miR-379-5p#	12.9	40.7	32.8	51.30	34.36	44.85	32.37
miR-154-5p#	12.6	38.4	32.4	36.76	37.99	32.89	46.11
miR-676-3p	12.5	76.2	26.0	77.57	76.02	83.56	67.66
miR-139-3p	12.4	88.0	20.5	84.04	92.81	88.53	86.64
miR-1843-3p	12.2	43.8	21.3	42.96	39.35	50.05	43.00
miR-1306-5p	11.9	70.2	20.2	63.02	73.21	74.23	70.14
miR-301a-3p	11.6	42.2	26.3	40.31	43.25	31.82	53.38
miR-30d-3p	11.0	35.7	33.7	37.68	29.39	43.44	32.19
miR-9-5p*	10.6	35.1	30.1	32.13	47.25	35.69	25.18
miR-1839-3p	10.5	70.7	21.8	75.95	71.94	67.46	67.36
miR-107-5p	10.3	50.1	20.2	53.08	48.40	48.21	50.67

miRNAs with known expression changes in hypertrophy. * miRNAs with known function in the heart. miR-9 includes miR-9-1, miR-9-2 and miR-9-3.

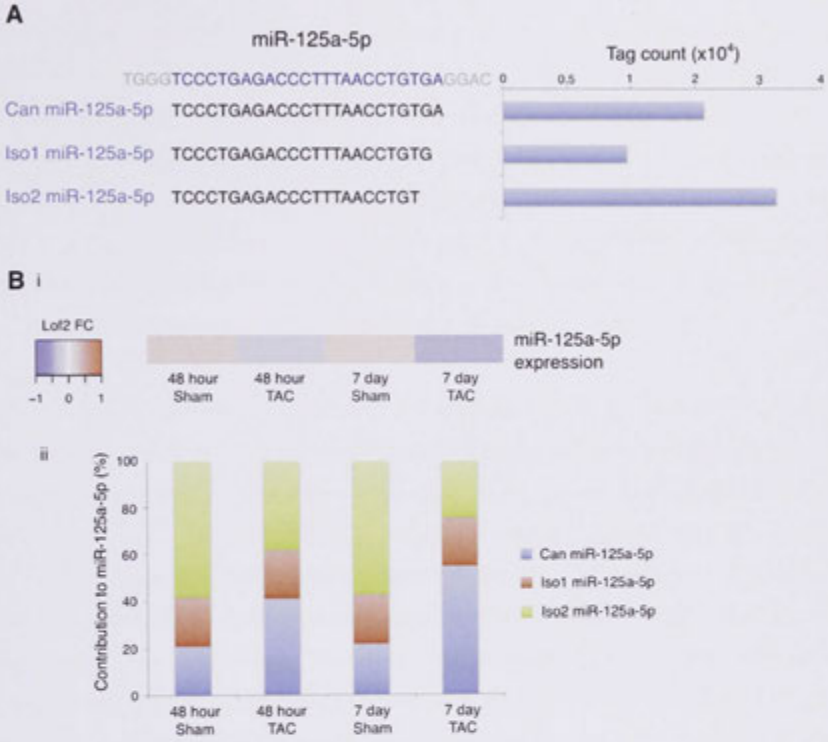


Figure 4.14. The 3' isomiRs of miR-125a-5p alter in cardiac hypertrophy.

(A) The canonical miR-125a-5p is defined as the sequence corresponding to miR-125a-5p in miRBase (shown in blue). The major forms of miR-125a-5p found in the cardiomyocyte NGS are listed below with the corresponding tag counts on the right. (B) (i) miR-125a-5p decreases slightly in the both the 48 hour and 7 day TAC samples. Data is shown as average Log2 FC (n=2). (ii) Contribution of each miR-125a-5p 3' isomiR to the miR-125a-5p locus.

miR-133a-3p exists as 5' and 3' isomiRs:

A predominant example of 5' and 3' isomiRs is the cardiac specific miR-133a, which is the most abundant miRNA in the sequencing libraries. miR-133a exists as multiple 5' and 3' isomiRs derived from both the 5p and 3p arms, which could greatly increase its target mRNA spectrum. The predominant species of miR-133a-3p is actually a 5' isomiR that starts one nucleotide downstream of the annotated start site (Figure 4.15.A). The 5' isomiRs are also associated with multiple 3' ends. The same miR-133a isomiRs were identified by our group in an earlier study that documented the miRNA population of the HL-1 cardiomyocyte cell line (*55*). In further support of miR-133a-3p isomiRs, the top sequencing tags for miR-133a-3p listed in miRBase also include the 5' isomiR that begins one nucleotide downstream of the canonical start site (Figure 4.15.B).

MiR-133a-3p is anti-hypertrophic and its downregulation is sufficient to induce hypertrophy in mice (*120*). As the isomiRs differ by only one or two nucleotides and the remainder of the miRNA sequence remains the same, it is not possible to validate the existence of 5' or 3' isomiRs by conventional PCR methods. For this reason, a northern blot protocol was developed to probe against the miR-133a-3p sequence. A large amount of RNA is required for northern blotting, which was not available from the cardiomyocyte samples used for sequencing. Therefore, the northern blot was run using HL-1 cardiomyocyte RNA which contains the same proportion of miR-133a-3p isomiRs(*55*), and RNA from a murine whole LV sample (Figure 4.15.C). The size of the endogenous miR-133a was compared to artificial single stranded RNA molecules that mimic the sequences of the canonical and the isomiR (labelled can/23 nt and iso/22 nt respectively). The 23 nt and 22 nt mimics are clearly visible on the blot, and when mixed at 50:50 proportions the bands can be distinguished. While faint, the two bands corresponding to 22 nt and 23 nt are visible for the HL-1 cardiomyocyte RNA. The signal for miR-133a-3p is stronger in the whole LV sample, with the predominant species being 22 and 23 nt in length. This confirmed that miR-133a-3p is present as multiple lengths in both a cardiomyocyte cell line and primary LV tissue.

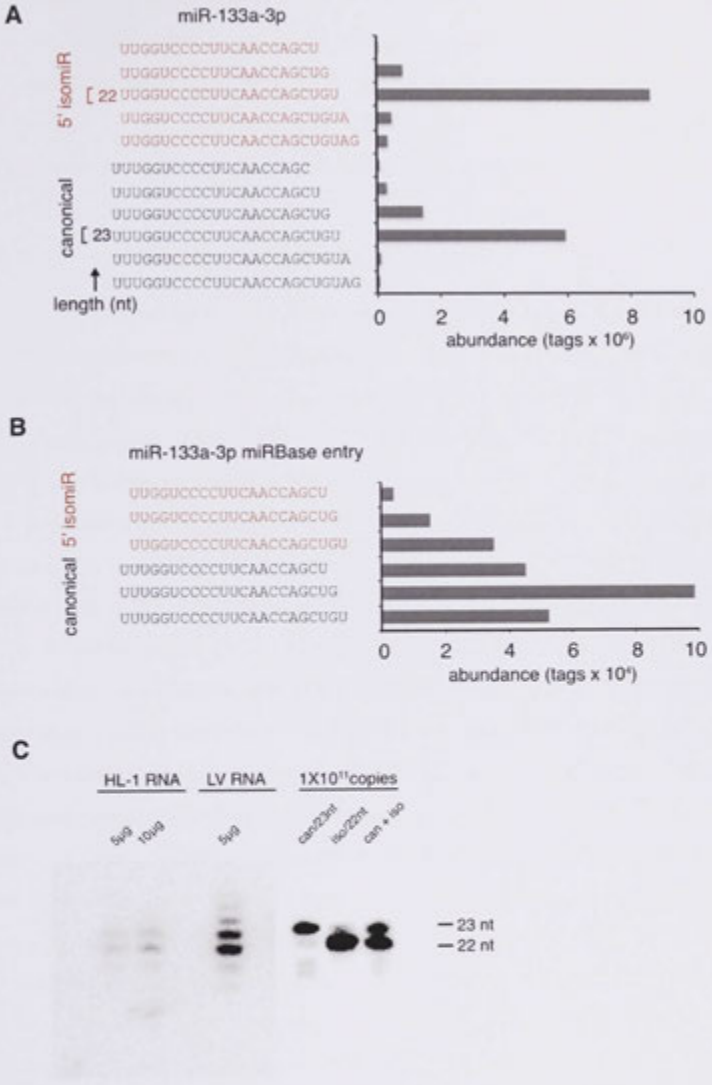


Figure 4.15. miR-133a-3p exists as both 5' and 3' isomiRs.

(A) Major mature miR-133a-3p species and their abundance in cardiomyocytes. Sequence tags are grouped into those with canonical (black, 'can') and +1 (red, 'iso') start sites. The mimics used for northern blots are indicates by the brackets (B) Northern blot of miR-133a-3p in the HL-1 cardiomyocyte cell line compared to the iso/22nt miRNA mimic. (C) Northern blot of miR-133a-3p in the HL-1 cardiomyocyte cell line and RNA from whole murine LV, compared to can/23nt and iso/22nt miRNA mimics.

After examination of the isomiRs present in cardiomyocytes purified from murine hearts, it is evident that several miRNAs exist as 5' isomiRs and majority of miRNAs have variation at the 3' end. The proportion of both 5' and 3' isomiRs remains relatively unchanged between normal, pre-hypertrophic and hypertrophic cardiomyocytes suggesting that regulation of isomiRs is not a primary mechanism to alter miRNA regulation in the heart.

4.2.5.c Non-templated additions

miRNA sequences can also differ from the canonical sequence by addition of non-templated nucleotides (NTA) to the ends of the miRNA (*231*). Numerous studies using next-generation sequencing have now observed NTA that differ from the genomic sequence at the 3' end of a miRNA (*79, 231-233*). The miRNA data was examined for evidence of NTA, defined as nucleotides that did not match the genomic sequence within the last 2 nt of the sequencing tag. Using this criterion, we identified 131 miRNAs that had evidence of NTA (more than 20% of the tags had NTA, Appendix 8.8). This is consistent with a prior study where 50% of the miRNAs expressed in normal human prostate tissue had NTA to some extent (*233*). Previous publications have observed that the most frequently added nucleotides to mammalian miRNAs are U and A (*79, 232*). Indeed, we also found that U and A were the predominant nucleotide added to individual tags (Figure 4.16.B). When comparing a single nucleotide position across multiple tags, U was the most frequently added nucleotide (Figure 4.16.C). It has been postulated that NTA has a strong preference for some miRNAs over others, with certain miRNAs consistently being modified in multiple tissue and cell types (*233*). For instance, miR-143-3p is frequently extended with 59-77% NTA reported in the murine brain (*233*) and 63% NTA reported in another study looking at murine brain, ovary and testes (*79*). Notably, miR-143-3p has 64% NTA in our data which is in agreement with these previous studies. In fact, of the ten miRNAs with the highest proportion of NTA according to Wyman et al., nine had NTA in our data while one miRNA was not expressed at all (*233*). Furthermore, the average NTA for eight of these miRNAs was within 10% of the proportion documented in Wyman et al. (*233*). In short, 50% of the miRNAs in purified cardiomyocytes have evidence of NTA with the U the most frequently added nucleotide.

NTA events are altered in hypertrophy

The addition of nucleotides, in particular A and U, can change both the stability of the miRNA and its function (*234, 235*). Therefore, we investigated if the proportion of NTA

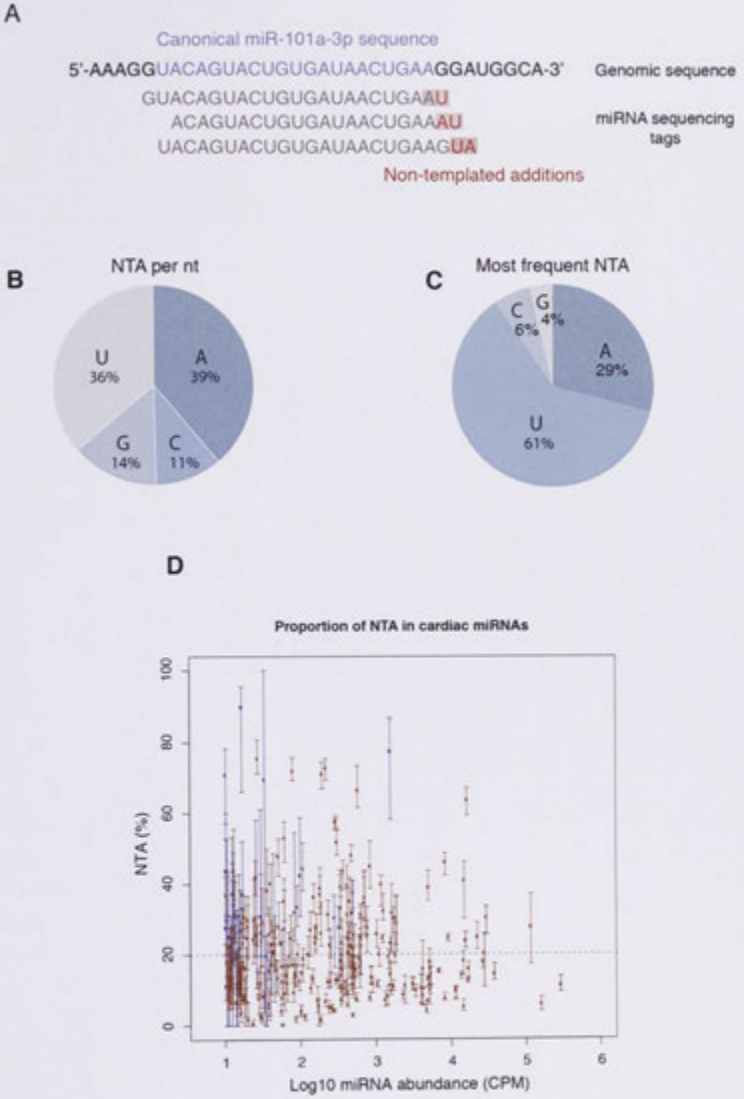


Figure 4.16. Cardiac miRNAs undergo non-templated additions.

(A) NTA were defined as mismatches to the genomic sequence in the last two nucleotides of the sequencing tag. (B) The contribution of each nucleotide to NTA position in individual tags. (C) The most common nucleotide for each NTA position in individual tags. (D) The proportion of tags representing non-templated additions per miRNA. MiRNAs with average tag count greater than 10 CPM were used for analysis. MiRNAs with a range less than 20% are shown in red, those with a range greater than 20% are indicated by blue.

Table 4.7. miRNAs where the proportion of NTA changes by more than 20% across the cardiomyocytes.

miRNA	Average CPM	NTA (%)					
		Average	Range	48 hour Sham	48 hour TAC	7 day Sham	7 day TAC
miR-92a-2-3p#	1595	77.4	28.5	81.9	74.5	84.0	69.1
miR-145-3p	490.3	31.8	27.2	24.0	33.1	30.4	39.9
miR-101b-3p#	322.5	27.7	20.8	31.3	22.4	30.5	26.4
miR-146a-5p#	281.5	30.4	20.5	33.4	32.9	33.5	21.8
let-7g-3p	100.0	42.3	28.0	43.4	41.5	34.4	50.0
miR-28c	84.6	32.0	30.6	31.3	28.1	29.5	39.2
miR-674-5p#	61.4	27.3	26.3	29.0	21.2	34.9	24.0
miR-143-5p	44.5	20.6	41.8	7.8	30.2	13.9	30.4
miR-532-5p	39.1	15.3	20.3	16.3	8.1	19.6	17.0
miR-874-3p	35.3	19.6	25.5	20.2	24.7	11.3	22.3
miR-652-3p	34.1	69.4	100.0	42.6	74.6	93.6	66.7
miR-330-3p	30.2	26.1	23.4	20.9	33.9	27.6	22.1
miR-345-3p	29.9	30.8	56.1	15.6	33.0	35.1	39.6
miR-671-5p	26.2	42.0	47.6	28.6	46.6	46.1	46.7
miR-425-3p	17.2	37.1	24.8	36.1	33.8	45.3	33.1
miR-5100	17.2	89.7	29.4	91.8	80.7	92.7	93.5
miR-17-3p	16.2	26.8	22.7	28.4	21.0	29.2	28.6
miR-203-3p	15.5	31.0	20.9	28.5	35.9	31.9	27.8
miR-490-3p	15.2	19.3	21.4	15.9	18.9	17.3	25.0
miR-324-3p#	14.6	20.5	32.8	27.8	10.0	20.9	23.1
miR-224-5p#	14.2	10.5	20.8	8.3	7.8	16.9	9.0
miR-125b-2-3p#	13.1	34.1	25.6	35.2	30.7	42.9	27.5
miR-379-5p#	12.9	27.4	24.0	25.4	23.8	36.0	24.2
miR-154-5p#	12.6	20.5	36.5	24.6	14.1	13.3	30.2
miR-676-3p	12.5	37.1	25.3	38.9	33.5	44.1	32.1
miR-139-3p	12.4	18.7	24.8	21.1	9.1	24.4	20.4
miR-1306-5p	11.9	20.0	20.7	18.5	20.7	26.8	14.2
let-7e-3p	11.4	25.0	25.6	22.5	32.0	24.0	21.7
miR-34c-5p#*	11.1	11.0	21.9	17.4	6.4	10.6	9.5
miR-9-5p*	10.6	29.7	52.7	17.0	29.4	35.0	37.5
miR-337-3p	10.4	43.6	29.1	41.6	35.0	54.4	43.4
miR-547-3p	10.4	70.8	21.1	72.8	73.0	72.0	65.5
miR-505-3p	10.2	27.5	22.2	30.3	24.4	31.4	23.9

miRNAs with known expression changes in hypertrophy. * miRNAs with known function in the heart. miR-9-5p consists of miR-9-1-5p, miR-9-2-5p and miR-9-3-5p.

changed for any miRNA between Sham and TAC. We found 35 miRNAs where the proportion of NTA changes more than 20% across the eight cardiomyocyte samples (Table 4.7). The weighted linear model comparing the proportion of NTA for each miRNA was applied, and this identified one miRNA with a FDR of 21% that had a different proportion of NTA between the Sham and TAC samples at 48 hours (Appendix 8.13). miR-143-5p has an increase in the proportion of NTA that correlates with a decrease in the overall expression of the miRNA (Figure 4.17). On average, 20.6% of miR-143-5p tags have NTA in the last two nucleotides of the tag, the most prominent addition being a non-templated A (Figure 4.17.A). miR-143-5p is significantly downregulated at both 48 hours and 7 days in the TAC samples (Figure 4.17.B). The downregulation of miR-143-5p correlates with an increase in the NTA of Adenosine to the 3' end of miR-143-5p.

A large portion of miRNAs have evidence of NTA. However, only one example could be found where this NTA is correlated to changes in expression levels. This suggests that in general, the NTA may be a part of normal miRNA regulation and turnover and is not influenced by the stress induced on the cell during hypertrophy.

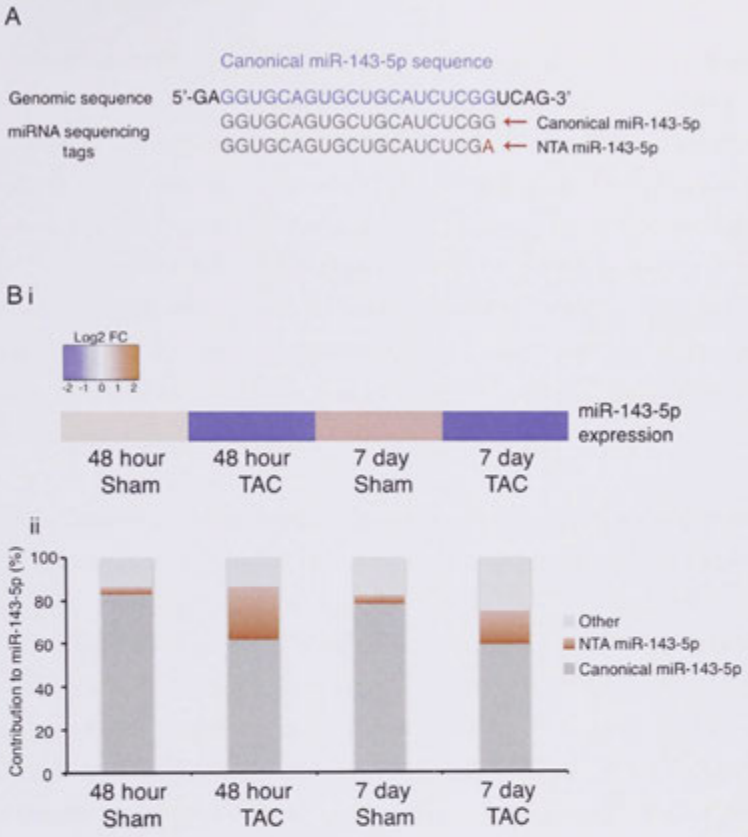


Figure 4.17. miR-143-5p is modified at the 3' end with NTA.

(A) The major forms of miR-143-5p found in the cardiomyocyte next-generation sequencing. The canonical miR-143-5p is defined as the sequence corresponding to miR-143-5p in miRBase. The NTA miR-143-5p is defined as a variant of miR-143-5p that contains an non-templated A at the 3' end. (B) The NTA miR-143-5p negatively correlates with the expression level of miR-143-5p. (i) miR-143-5p expression decreases in both the 48 hour and 7 day TAC samples. Data is shown as Log2 FC (n=2). (ii) Contribution of the canonical miR-143-5p, NTA miR-143-5p and all other miR-143-5p variants to the miR-143-5p locus.

4.2.5.D INTERNAL EDITING

The other form of post-transcriptional modification reported for miRNAs is internal editing, most commonly the deamination of adenosine (*95, 232*). Deamination of adenosine to inosine (A-to-I editing) is detected as an A-to-G change in miRNA sequencing products (*95*). Internal editing was defined as a nucleotide mismatch before the last two nucleotides of the canonical end position and before the last two nucleotides of the tag. There is notably very little internal editing compared to NTAs (proportion of tags with internal editing for all miRNAs with CPM greater than 10 are listed in Appendix 8.8). Only five miRNAs were detected with internal editing in more than 20% of their tags (Figure 4.18.A). The miRNAs with more than 20% internal editing or greater than 20% difference across the timecourse are listed in Table 4.8.

Each miRNA in Table 4.8 was checked manually for the location of the nucleotide change, a potential match to another genomic location and if the same editing was found in miRBase. In 50% of the tags for miR-101c, there is a U at the 5' end instead of the genomically encoded C. While the same editing is found in miRBase, the edited sequence actually matches 100% to miR-101a-3p, which is identical to miR-101c except for the additional U at the 5' end. Thus, it cannot be concluded that miR-101c is edited as the tags may be derived from the miR-101a-3p locus. Both miR-652-3p and miR-210-3p have more than 20% difference in internal editing events across the samples. However, both miRNAs have extremely low tag counts in at least two samples that is confounding the proportion of editing events. For instance, miR-652-3p has only two tag counts in the 370S sample, one of which is not an identical match to the genomic sequence resulting in an apparent editing of 50%.

miR-3963, miR-490-3p and miR-5097 are more convincing candidates for internal editing events. The main variant of miR-3963 contains a non-templated C instead of a U at the 5' end of the miRNA. This change is also detected at very low levels in the miRBase sequencing data, and importantly the edited sequence does not match any other miRNA in miRBase. miR-5097 has a high level of internal editing, with 85% of tags containing an A to G mismatch in the 3' end of the miRNA. While this was an ideal candidate for A-to-I editing, miR-5097 has now been removed from the last release of miRBase as it is derived from a fragment of a tRNA and is not loaded into Ago2 (*236*). miR-490-3p is edited in 32% of its tags, accounted for by an A to U or G transition towards the 3' end of the

miRNA (Figure 4.18.B). The edited sequence does not match another miRNA deposited in miRBase. While there is large difference in the proportion of miR-490-3p edited between the different samples, there is no correlation between expression level of miR-490-3p and the editing events.

Internal editing does not change during cardiac hypertrophy

We next investigated if the proportion of internal editing is altered for a miRNA during cardiac hypertrophy. The proportion of internal editing changed more than 20% in 5 miRNA (listed in Table 4.8). However, only one miRNA, miR-221-3p, was identified with a FDR of 6.2% when applying the weighted linear model comparing the proportion of internal editing (Details in Appendix 8.9). On closer inspection, miR-221-3p is edited on average in 3.1% of miR-221-3p tags, and the change in the proportion of miR-221-3p was only 2.8%. Due to the low levels of editing, miR-221-3p was not investigated further.

Table 4.8. miRNAs with >20% internal editing and/or with a range >20% internal editing

miRNA	Average CPM	Internal editing (%)					
		Average	Range	48 hour Sham	48 hour TAC	7 day Sham	7 day TAC
miR-101c	57.2	48.82	26.11	60.11	50.12	45.93	39.12
miR-3963	23.54	58.93	9.51	54.88	58.80	63.49	58.56
miR-490-3p	15.21	32.04	44.07	20.26	26.47	53.63	27.82
miR-5097	57.63	86.16	22.36	90.30	86.42	90.16	77.77
miR-652-3p	34.07	21.43	100	3.02	4.96	51.18	26.55
miR-210-3p#	245.03	6.988	23.356	2.83	13.62	6.57	4.93
miR-324-3p#	14.56	14.42	40	12.06	20.00	11.96	13.67

miRNAs with known expression changes in hypertrophy

To summarise, there is very little editing detected in the miRNAs expressed in murine cardiomyocytes. miR-490-3p is the most convincing candidate for internal editing, which contains an A-to-G transition that reflects A-to-I editing. This alteration to the miRNA sequence occurs in the 3' end of the miRNA making it difficult to postulate what the biological function of such a change could be.

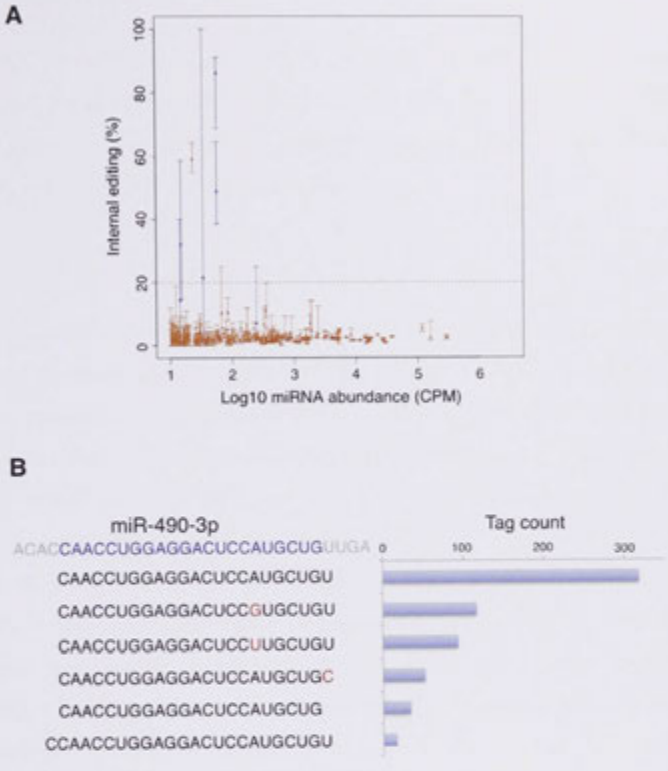


Figure 4.18. A small number of miRNAs have evidence of internal editing.

(A) Internal editing in the hypertrophic heart; the proportion of tags representing internal editing events per miRNA. MiRNAs with average tag count greater than 10CPM were used for analysis. MiRNAs with a range less than 20% are shown in red, those with a range greater than 20% are indicated by blue. (B) miR-490-3p contains evidence of internal editing; the miRBase defined sequence is shown in blue, with the major isoforms in the cardiomyocyte data shown below. The red nucleotides indicate mismatches to the genomic sequence.

4.3 Discussion

In this chapter, I performed genome-wide profiling of miRNAs in control, pre-hypertrophic and hypertrophic cardiomyocytes. The data was analysed for differential expression of miRNAs, the presence of miRNA processing variants and indications of variant changes between Sham and TAC samples.

4.3.1 Features of the small RNA library

The library construction is based on size selection of PCR products corresponding to 110-125 nt to avoid contamination by other small RNAs such as tRNA and primer dimers. However, there was still a large proportion of reads that mapped to tRNA and rRNA. Due to the nature of the library preparation, tRNA-derived RNA fragments, which can give rise to functionally active miRNAs (*237*), are also incorporated into small RNA libraries. Another study using 454 sequencing, made small RNA libraries from fragments corresponding to 17-26 nt in length and found miRNAs contribute to only 40% of the total tags (*238*). A large proportion of the library was derived from tRNA, rRNA, protein coding loci and transposable elements (*238*). The ligation of adaptors requires a 5' phosphate and 3' OH group, so while non-phosphorylated degradation products would not be incorporated, any other RNA species of the correct size could be sequenced (*239*). These other RNA populations have sizes ranging from 19-26 nt and therefore would be present in our sequencing libraries along with miRNAs (*239*).

In addition to other RNA species, 17% of our reads did not overlap with any features in the mouse genome. Another study using SOLiD sequencing for small RNAs found that 37-43% of the mapped reads that did not overlap with annotated features were non-annotated small RNAs (*240*). It is therefore possible that the tags that do not overlap with known features are novel small RNAs, which remain to be explored.

The contribution of each miRNA to the library was highly heterogeneous, with the top 20 miRNAs contributing 87.3% of the tags and the highly abundant miR-133a alone accounting for 30.2% of tags. This distribution of tags has been found in other studies. For example, the top miRNA in the Zaragosi et al. study also accounted for 30% of tags (*240*). On the other hand, a previous study in the lab indicated that miR-145 was the most

abundantly expressed miRNA contributing only 13% of tags in the HL-1 cardiomyocyte cell line (*55*). It could be postulated that these differences are due to different miRNA profiles in the cell line compared to cardiomyocytes derived from the mouse heart. However, the apparent abundance of miRNAs in next-generation sequencing depends heavily on the library preparation method and the sequencing platform used (*241*). While both of these studies were conducted using the SOLiD™ platform, our earlier work used the SOLiD™ Small RNA Expression Kit and the libraries discussed here were prepared with the NEBNext® Small RNA Library Prep Set for SOLiD™. The biases introduced by the sequencing platform and preparation method are highly reproducible, meaning that next-generation sequencing is suitable to make comparisons of relative expression changes between samples, whereas the absolute expression or rank of a miRNA between different studies is not as accurate (*242, 243*).

4.3.2. Sample heterogeneity

There is no defined or best practice for normalisation of miRNA sequencing data, and only a handful of studies have made comparisons of the available normalisation methods (*244, 245*). Some of the earlier studies have normalised only to the total number of miRNA mapped reads (*242, 246*). But as mentioned previously, this doesn't take into account differences in the RNA composition that may occur between samples. Other normalisation strategies used for miRNA sequencing include normalisation to miRNAs that do not change or to the levels of synthetic miRNAs. Normalisation to other miRNAs involves choosing a subset of miRNAs that remain stable between samples according to RT-qPCR data (*131*). Alternatively, a synthetic oligonucleotide can be added to the first step of library preparation and can then be used to normalise the levels of small RNAs between samples (*247*). Both of these methods would account for the input of RNA in each sample but would not account for RNA composition differences in the initial samples. For this reason, we used TMM normalisation, which takes into account the RNA population of each sample (*219*).

Our samples are closely related as reflected by the high Spearman correlations. The top expressing miRNAs in our libraries, such as miR-133a-3p, miR-133b-3p and miR-22-3p, remain unchanged between the Sham and the TAC, which would contribute to the large Spearman correlations. In line with this, another study looking at miRNA expression 7 days post-TAC found that more moderately expressed miRNAs change compared to the most

abundant miRNAs (*131*). Despite the large similarity between samples, the Spearman correlation does cluster the Sham and the TAC samples separately, indicating that there are distinct expression differences between the two groups.

4.3.3 Differential expression of miRNAs

This was the first study to document the differential expression of miRNAs in purified cardiomyocytes during the hypertrophic response. A small number of miRNAs were identified with a significant expression difference between the Sham and TAC samples. More miRNAs were upregulated than those downregulated. This is consistent with another study that used Illumina sequencing to document the miRNA changes in the whole LV 7 days post-TAC (*131*). This study also found more miRNAs upregulated and hypothesised that there may be a lag period for miRNA elimination resulting in less detectable downregulated miRNAs (*131*).

There are multiple mechanisms by which a miRNA can be regulated, including processing and turnover. Several of the upregulated miRNAs are under the control of transcription factors known to be involved in cardiac hypertrophy. For example, miR-199b is upregulated in the 7 day TAC samples (Figure 4.7) and is a target of the calcineurin/NFAT pathway that is upregulated in response to pressure overload (*10, 227*). Other miRNAs are under the control of their host gene. For instance, miR-208b is encoded within an intron of the β-MHC gene and miR-208a is located within α-MHC (*248*). The ratio of β-MHC to α-MHC increases during the hypertrophic response (refer to Chapter 3.3.3) to conserve energy and change cardiac output (*210, 211*). We observed a slight increase in β-MHC at 48 hours and a significant overexpression at 7 days (Figure 3.7). We also observed an increase in the levels of miR-208b-5p at both 48 hours and 7 days albeit not significant. Similarly, the levels of both strands of miR-208a decreased at 7 days, possibly reflecting the decrease in α-MHC.

It was originally anticipated that miRNAs would be a component of the initial response to cardiac stress and thus we would see expression changes on or before the 48 hour time point. In fact, only three miRNAs had sufficient dysregulation at 48 hours to make statistical significance (miR-451, miR-223-3p and miR-21-5p). miR-451 is in a cluster with miR-144 and is directly regulated by the transcription factor GATA-4 (*249*). GATA-4 lies downstream of multiple hypertrophic signalling pathways, such as the calcineurin/NFAT

and the PI3K-Akt pathways (reviewed in (*4*)). It is known that in response to TAC or pressure overload, the DNA binding activity of GATA-4 increases, which leads to the activation of genes such as α-MHC, ANP and BNP (*250, 251*). Since BNP is strongly induced at 48 hours (Figure 3.7), it is possible that the upregulation of miR-451 is also due to stimulation by GATA-4.

In addition to transcriptional control, regulation of miRNA biogenesis and turnover can influence the levels of a miRNA. In line with this, certain miRNAs are expressed in a tissue or developmental stage specific manner without variation in the levels of the pre-miRNA, suggesting regulatory mechanisms that act on the mature miRNA (*252*). A study looking at miRNA decay disrupted transcription and found that 95% of human miRNAs remained stable for at least 8 hours (*253*). This could account for the small number of downregulated miRNAs in our study. If the locus was transcriptionally silenced, it could take a considerable amount of time before the existing miRNAs are cleared. While miRNAs are generally considered to be very stable, it is also known that the turnover rate can be miRNA specific. For instance, miR-208a has a half-life greater than 12 days in rat hearts, so even after repression of the miR-208a locus it would take several days before any downregulation would be visible (*114*). While other miRNAs are unstable and can be rapidly degraded by enzymes such as Xrn1 and Eri1 (*253, 254*). It is not yet known if there are specific enzymes, such as exonucleases, that regulate miRNA turnover in the mammalian heart. But it is possible that the miRNAs downregulated in this study have high turnover rates in cardiomyocytes.

It is also interesting that we only see strong upregulation of pro-hypertrophic miRNAs, while miRNAs found to be anti-hypertrophic are notably absent from Figure 4.7. It could be hypothesised that anti-hypertrophic miRNAs are upregulated early in the timecourse to try and combat the hypertrophic signalling, and the levels of these miRNAs then decrease once the hypertrophic response is established. It is possible that we did not sample the cardiomyocytes early enough; if the anti-hypertrophic miRNAs are upregulated before the 48 hour time point. On the other hand, while these miRNAs are known to be antihypertrophic, the heart is trying to compensate for the increase in pressure and thus the hypertrophy is an adaptive mechanism. Thus, the levels of anti-hypertrophic miRNAs may be high in normal cardiomyocytes to prevent unnecessary hypertrophy from occurring. It has been shown that anti-hypertrophic miRNAs are more important during regression of hypertrophy once the stimulus has been removed and thus we would not observe their

upregulation in our study (*255*). This data provides more insight into the regulation of hypertrophy prior to the onset of symptoms. The miRNAs upregulated early could provide a clue as to which miRNAs to target for therapeutic intervention against heart disease.

4.3.3.A Validation of miRNA expression

The expression of the differentially expressed miRNAs were tested using RT-qPCR. Importantly, the PCR primers were designed against the major variant of the miRNA present in the sequencing data. This is particularly important for the 3' end of the miRNA, as the ligation of the stem-loop reverse transcription primer relies only on the last six nucleotides. The RT-qPCR was performed on the identical eight samples subjected to small RNA sequencing, plus the additional cardiomyocyte samples as documented in Chapter 3. The expression levels of six miRNAs followed the same trend in the RT-qPCR on the same samples used for sequencing, with four of these having a statistically significant difference in the extended dataset. There was often a large variation in the levels between biological replicates. The most prominent example is miR-451, which has a large increase at 48 hours in the sequencing data. The RT-qPCR for the samples subjected to sequencing follows exactly the same trend as the next-generation sequencing itself. However, large variations in the expression levels across the extended dataset actually result in a decrease in expression between Sham and TAC.

Discrepancies between the expression levels could be attributed to variation in the cardiomyocyte enrichment, as there was a difference in the number of cardiomyocytes and expression of cardiomyocyte markers between samples. The question was then asked if any of the miRNAs tested for validation were cell type specific. However, the only known miRNA that is purely cardiomyocyte specific is miR-208a (*114*), the remaining miRNAs just have enhanced expression is certain cells. Three of the miRNAs that could be validated in the extended dataset are abundant in cardiomyocytes; miR-199 is predominantly expressed in cardiomyocytes (*256*), miR-223-3p is localized to cardiomyocytes in cryosections in the heart (*257*), and while miR-21 is highly expressed in fibroblasts it is still abundant in cardiomyocytes (*258*). On the other hand, miR-31-5p was also validated but is mainly expressed in vascular smooth muscle cells (*259*). Of those that could not be validated, miR-143-5p and miR-101a-3p have the highest expression in the walls of the aorta and cardiac fibroblasts respectively (*260, 261*). The specificity of miR-451 and miR-150-5p are less defined, both are known to have a functional role in cardiomyocytes and

miR-451 is regulated by GATA-4 so it is possible they are cardiomyocyte specific (*249, 262, 263*). It is plausible then that if the miRNAs tested for validation have cell specificity then the levels could be variable between samples. The expression of these miRNAs could also be tested in the Wash samples obtained during cardiomyocyte enrichment. This would help ascertain if these miRNAs are altering expression due to enrichment or due to the hypertrophic response. The downstream analysis, such as correlation with mRNA levels, should concentrate on the miRNAs that could be validated in the extended cardiomyocyte sample set.

4.3.3.B Novel miRNA changes in hypertrophy

Of the miRNAs we found to change expression levels, only three are novel to cardiac hypertrophy (miR-143-3p, miR-224-5p and miR-299-5p). This begs the question as to why they have not been documented previously? miR-143-3p was previously called miR-143*, as it has low expression levels in comparison to miR-143-5p. In our dataset, miR-143-3p contributes only 0.3% of the tags that map to the miR-143 hairpin. miR-143-5p also has lower levels in the 7 day TAC samples, but variable expression in the 7 day Sham samples resulted in a high FDR. It is known that miR-143-5p has the highest expression levels in the heart, and is most abundant in the walls of the aorta and coronary vessels (*260*). miR-143 is encoded within a cluster along with miR-145, and is regulated by the transcription factors SRF and Nkx2-5 (*264*). Therefore, miR-143-3p would also be under the control of the same transcription factors, and if regulated transcriptionally, would have the same localisation as miR-143-5p. Studies thus far have concentrated on the role of miR-143 in vascular smooth muscle cells (VSMCs), where a downregulation of miR-143-5p and miR-145 results in the de-differentiation, migration and proliferation of VSMCs (*260*). In a previous study, the expression of miR-143-5p did not change in the myocardium after TAC but did decrease in the aorta immediately upstream and downstream of the TAC site (*260*). The levels of miR-143-3p were confirmed by RT-qPCR in the sequencing samples at 48 hours but could not be validated in the extended dataset. As mentioned above, this could be attributed to the expression being limited to VSCMs. The differences in miR-143-3p expression could therefore be due to subtle differences in the purification of the cardiomyocyte samples.

miR-224-5p increases at 7 days, but it is moderately to lowly expressed. It has been previously shown to decrease expression in patients with ICM but its role in cardiac disease

remains unknown (*123*). It is known that miR-224-5p is upregulated in several cancer types, and it is postulated to play a role in the regulation of apoptosis and proliferation (*265-267*). The levels of miR-224-5p increase after stimulation with TGF-β1 in follicular cells (*266*). TGF-β1 is upregulated in the heart in response to pressure overload (*268*) and feeds into the MAPK pathway (see section 1.1.2.C). The increase in miR-224-5p 7 days post-TAC could therefore be attributed to activation of the MAPK pathway.

miR-299-5p increases at 7 days, but it is also moderately to lowly expressed. It was previously annotated as miR-299*, and the other strand miR-299-3p has previously been reported to increase 7 days post-TAC (*130*). However, miR-299-3p has next to no expression in our dataset. It is highly likely that because miR-299-5p and miR-143-3p were previously annotated as the star form of a miRNA hairpin they were not present on the earlier microarrays and thus they have not been reported to date. The only hint on the function of miR-299-5p is that it controls haematopoietic progenitor fate and is upregulated in megakaryoblasts (the precursor to megakaryocytes that produce platelets) (*269*).

The three novel miRNAs to be regulated in our study have been missed previously due to cell specificity and the low expression that may have been confounded in whole LV samples. Nevertheless, two of them are linked to signalling pathways that are activated during the hypertrophic response suggesting they are upregulated in response to pressure overload.

4.3.3.C Comparison with other studies

We also examined the fold changes for several miRNAs that have been shown to change in one or more previous studies of hypertrophy. There were 21 miRNAs that appeared unchanged in our dataset or only had slight changes in their expression. There are a number of reasons why this could be the case. Firstly, the miRNAs may be enriched or specific to cell types other than cardiomyocytes. For example, members of the miR-23~27~24 cluster are reported to be upregulated 7 days post-TAC in multiple studies, yet the fold changes found in our data are minimal (*127, 128, 130*). The miR-23~27~24 cluster is highly expressed in endothelial cells and thus the levels of these miRNAs would have been reduced during cardiomyocyte enrichment (*270*). The same applies for miR-126-5p, the miR-30 family and the miR-17~92 cluster (that includes miR-20a and miR-92a), which

are enriched in endothelial cells or cardiac fibroblasts (*271, 272*). Therefore, although yet undefined, it is possible that some of the other miRNAs that did not make our cut off for statistical significance are derived from other cell types in the heart.

Secondly, there is a difference in the detection of miRNAs depending on the platform and the library preparation used (*242*). Each library-preparation method preferentially captures a distinct set of miRNAs, therefore some of the miRNAs previously detected may not make our thresholds for expression. Finally, the criteria used to define "differentially expressed" miRNAs are not consistent between studies. Some of the earlier studies did not apply a statistical threshold but instead used only the fold change as an indication of differential regulation (*120*). Furthermore, another next-generation sequencing study had five biological replicates per group and could therefore confidently detect changes as small as 25% (*131*). It is known that when the number of transcripts, in this case miRNAs, is small or the fold changes are small then you need a larger number of replicates to control for the FDR (*273*). It is possible that if we changed the criteria used to detect differential expression, or included more biological replicates for sequencing then we would increase the overlap of our data with miRNAs others have previously found to be regulated in hypertrophy.

4.3.4 miRNA processing variants in the hypertrophic heart

This was the first study to document the miRNA processing variants in the hypertrophic heart. We have documented that 97% of miRNAs in cardiomyocytes have some form of variation in their processing. Individual examples were identified where this processing changes during cardiac hypertrophy but broadly speaking there was no wholesale change in processing variants.

4.3.4.a ARM BIAS

Our analysis of miRNA arm bias identified that 82.3% of miRNA hairpins produce mature miRNAs predominantly from one strand. Our previous study looking into the miRNA population in HL-1 cardiomyocytes found similar results, with 87.5% of miRNA hairpins producing one predominant strand (*55*). We reported 36 miRNA hairpins with more than 20% of the tags derived from both strands. Of these, 35 were expressed in HL-1 cardiomyocytes, of which 26 had more than 20% derived from each strand (*55*). While majority of miRNA hairpins give rise to one strand, the predominant strand in our data is

not necessarily the miRNA strand that has been studied to date. This is because the previous view on miRNA processing was that only one miRNA strand from the miRNA hairpin was functional and the other strand was degraded. It is now acknowledged that both strands of the hairpin are functional and the predominant miRNA strand can change (*79, 80*). This change in view is reflected by miRBase, which dispensed with the mature/miRNA* nomenclature and renamed all miRNAs based on the strand they were derived from (*82*). It is important to determine the predominant strand expressed in the biological system of interest, particularly when looking at the functional or therapeutic potential of a miRNA. For instance, miR-9-5p was previously defined as the mature mir-9 and functional studies have been performed to elucidate its role and targets in systems including the heart. Yet in our data, 67.5% of miR-9 is actually derived from the miR-9-3p strand (Figure 4.19). It is known that miR-9-5p directly targets myocardin in the heart to suppress the hypertrophic signalling that is stimulated by NFATc3 (*141*). Additionally, several other cardiac related mRNAs were identified as miR-9 targets by a study that determined miRNA:mRNA interactions in AGO proteins in the mouse brain (*71*). Albeit in the brain, mRNAs that regulate fibrosis and cardiac contractility were found cross-linked with miR-9-5p (*71*). The targets of miR-9-3p on the other hand, remain unknown. Strand expression is important when knocking out or overexpressing a whole pre-miRNA hairpin as you cannot assume that only one strand will be expressed.

While several miRNA hairpins had miRNAs derived from both strands, we could not find evidence that strand preference changed between Sham and TAC. It has been postulated that certain sequence and thermodynamic properties of the pre-miRNA and miRNA duplex can influence strand selection into RISC (*59, 274*). For instance, the strand whose 5'-end is less tightly paired is incorporated into RISC and the highly expressed strand tends to have a U-bias at the 5' end (*274, 275*). These rules however, do not account for the strand preference that occurs in certain tissues, implying that there are additional regulatory mechanisms determining which miRNA strands are predominant in the heart. This lead us to speculate if there would be an advantage to switching strands in response to stress or

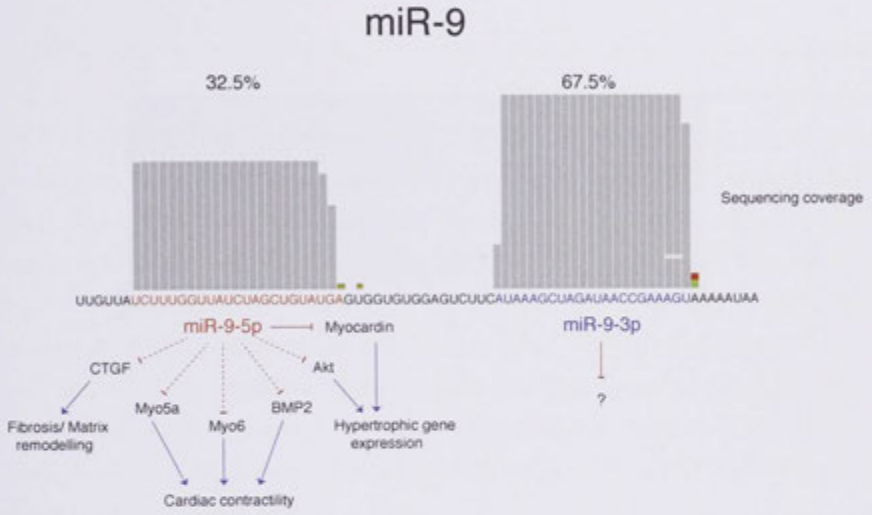


Figure 4.19. The role of miR-9 in the heart.

The sequencing coverage is shown for miR-9-5p and miR-9-3p, with miR-9-3p being the predominant strand. The target interactions of both strands are shown below the sequencing coverage. Several targets of miR-9-5p have been documented, these targets and the pathways they regulate are listed. The solid red line indicates a direct target of miR-9-5p in the heart (*141*), while the dashed lines indicate targets identified from an AGO IP in the murine brain (*71*).

disease. While the targets between the strands is different, it may be more efficient to upregulate a miRNA from a different locus via transcription. Although the strand regulation is consistent between Sham and TAC, knowledge about the strands expressed is critical for understanding the functional role of a miRNA hairpin.

4.3.4.B 5' AND 3' ISOMIRS

Approximately 20% of the miRNAs we detected had greater than 20% of their tags derived from 5' isomiRs, including 13 miRNAs that have previously been studied in hypertrophy. Of the 55 miRNAs with greater than 20% 5' isomiRs, 41 were also expressed in HL-1 cardiomyocytes, of which 31 were also found to have more than 20% 5' isomiRs (*55*). It has been postulated that Drosha and Dicer cleavage with relaxed specificity is the reason behind miRNA start and end heterogeneity (*88, 276, 277*). As the miRNA relies on the 5' end for targeting specificity, there is an evolutionary pressure to maintain constant 5' ends of the miRNA (*57*). Consistent with other studies, 95% of miRNAs had variation at the 3' end (*88, 232, 276*), suggesting that 3' isomiRs may also derive from post-processing addition and/or trimming events. In agreement, *Drosophila* miR-34 has multiple 3' ends that result from the 3' to 5' exoribonuclease Nibbler trimming the 3' end while miR-34c is within the RISC complex (*278*).

Multiple cardiac miRNAs existed as 5' isomiRs with varying 3' ends. This included miR-133a-3p, which has been extensively studied in the heart (*140, 279, 280*). The impact of these isomiRs on the targeting specificity of miR-133a-3p is documented in Chapter Six. The length heterogeneity of miR-133a-3p found in the sequencing data was confirmed using high resolution Northern blotting. Length variation such as this has been detected using Northern blotting for several miRNAs across different cell lines (*57*). It is important to note that the results of the Northern blot confirm two major lengths of miR-133a-3p, which correspond to the major 5' isomiRs, but the exact sequence of these two lengths could not be established with this method.

Notably, there was no obvious difference in processing of 5' isomiRs between the Sham and the TAC samples. One miRNA, miR-125a-5p, had variation in 3' isomiRs between samples, with more of the canonical miR-125a-5p 3' end in the TAC samples. As mentioned in the Introduction, alteration of the 3' end of a miRNA can alter its loading into specific AGO complexes or its turnover (*62, 92*). There is a slight decrease in the

levels of miR-125a-5p in the TAC samples so it is possible the 3' end is altering the stability of the miRNA but this remains to be confirmed. In summary, there is evidence for 5' and 3' isomiRs in several miRNAs but the isomiRs produced remain consistent during the cellular stress induced by cardiac hypertrophy.

4.3.4.c Non-templated additions

Approximately half of the miRNAs in our study have low levels of non-templated nucleotides in the last two positions of the sequencing tag. This was not confined to just lowly expressed miRNAs, which is consistent with a previous study that showed the amount of NTAs is not correlated with expression levels (*281*). As mentioned previously, a subset of miRNAs are frequently modified by A and U additions, irrespective of their expression (*281*). In agreement, several of the miRNAs we detected with NTA such as miR-143-5p, have similar proportions in other studies (*79, 233*).

It is important to note that the level of NTA we have detected may be an underestimate. There is a 25% chance that the NTA will coincidentally match the genomic sequence masking any true NTA events. This was explored in a study by Westholm et al., where they used mirtrons (miRNAs spliced out of introns independently of Drosha) to examine NTA as the exact cleavage position is defined (*282*). They found that miRNAs with NTA are quite abundant once the nucleotides that fortuitously match the genomic sequence are accounted for.

The nucleotide most frequently added in the NGS data was a U. Urdiylation is catalysed by ZCCHC11 and commonly occurs on the pre-miRNA to block the uptake into Dicer, while adenylation is catalysed by PAD4 and occurs after cleavage by Dicer (*281*). You would therefore expect a larger proportion of miRNAs to be derived from the 3p arm to have a larger proportion of U addition, as this is the 3' end available for uridylation in the pre-miRNA hairpin. We found a slight preference for uridylation on the 3p arm (9% more than the 5p arm), which is consistent with this notion. It has been shown before that an increase in U addition on the 3' end of a miRNA results in a decrease in miRNA abundance. However, we could only find one example where the proportion of NTA consistently changed in the same direction between Sham and TAC (*283*). The levels of miR-143-5p decrease as the amount of A addition to the 3' end increases, implying that an increase in A destabilises the miRNA or marks it for degradation. The addition of an A to the mature

miRNA can reduce its ability to interact with AGO2 and AGO3 (*281*). In agreement, there is a bias towards mono uridylation and away from mono adenylation among miRNAs found in AGO complexes (*231*). The binding of a miRNA to the AGO protein stabilizes the miRNA, with increasing levels of AGO proteins correlating with increased levels of mature miRNAs (*284*). Therefore, the addition of an A to miR-143-5p may reduce its levels in the TAC samples by preventing its incorporation into the RISC complex.

To conclude, at least 50% of the miRNAs in cardiomyocytes have NTA but in general the levels of NTA do not change in the pre- and hypertrophic cardiomyocytes.

4.3.4.D Internal Editing

The other form of posttranscriptional modification we searched for was internal editing of the miRNA. RNA editing is common in the murine brain, but our data would suggest that A-to-I editing in mature miRNAs is incredibly rare in the heart. The editing of pri-miRNAs is more common then mature miRNAs, and can alter their processing by supressing either Drosha or Dicer cleavage (*96, 285*). It is possible then that A-to-I editing does occur more frequently in the heart on pri-miRNA segments that do not get cleaved into the mature miRNA strand, and thus are not detected in out data. There was one plausible example of editing in miR-490-3p. It is conceivable that miR-490-3p is modified at the pri-miRNA level, which could hinder the expression of the miRNA. In support of this hypothesis, miR-490-3p is lowly expressed in the sequencing library. However, the impact of such a modification on the mature miR-490-3p is not clear. There is no correlation between the extent of internal editing and the expression levels of miR-490-3p, implying the modification is not altering the degradation or stability of the miRNA. It is possible that this base change could alter the targeting potential of miR-490-3p. The editing does not occur within the seed region, which is conventionally thought of as the main determinant of miRNA binding. However, binding within the 3' region of the miRNA can enhance target interactions and supplement mismatches within the seed region (*286*). Thus, there may be subtle differences in the mRNA target spectrum of the internally edited miR-490-3p.

4.3.5 Conclusions

This aim of this chapter was to analyse NGS data for differentially expressed miRNAs, to detect miRNA processing variants and their regulation in cardiomyocytes derived from control, pre-hypertrophic and hypertrophic hearts. This was the first study to examine miRNA differential expression during hypertrophy in purified cardiomyocytes, and analysis revealed a small number of miRNAs that are deregulated during the hypertrophic response. Importantly, the majority of these miRNAs have been previously reported to change in the same direction, suggesting we have captured the correct miRNA changes with our approach to differential expression analysis.

The most novel aspect of this work is the analysis of miRNA processing variants. We found ample evidence of arm bias, 5' and 3' isomiRs and NTA for cardiac miRNAs. All of these variants provide a mechanism to change the sequence and potentially the function of a given miRNA locus without changing the genomic sequence. There is a high correlation between the total miRNA population and the miRNAs found within the RISC complex, suggesting that the miRNAs we have detected are almost entirely a bioactive population (*131*). I detected individual examples that have variable processing in hypertrophy but overall there was no wholesale change in processing. Despite this, the existence of miRNA variants is still vital when determining the mRNA network that one miRNA can regulate. This network is further complicated by the processing variants of mRNAs that the miRNA interacts with. Chapter Five investigates the differential expression of mRNA transcripts, as well as 3'UTR variants that will determine the mRNA pool available for miRNA regulation.

Chapter Five

3' UTRs in the Murine Heart

5.1 Introduction

The number of investigations into 3' UTR usage on a global scale has increased in the past five years. The initial studies used microarrays that had a select number of probes against alternative 3' UTRs (*161, 287*). This included one study on cardiac hypertrophy, which documented that for the mRNAs observed, there was a tendency towards mRNAs becoming shorter in hypertrophy (*168*). However, the microarray relies on *a priori* knowledge of the 3' UTRs present in the sample and thus the full spectrum of 3' UTRs expressed in the heart remain unknown.

The first study to use next-generation sequencing (NGS) technology to investigate 3' UTRs was published in 2010 and used a modified RNA-seq protocol (*288*). Since then, a number of methods to interrogate the 3' ends of mRNA transcripts with NGS have been developed (*289-292*). All of these methods rely on capturing the poly(A) tail at the 3' end of the mRNA. However, these methods can also result in internal priming of poly(A) stretches encoded within an mRNA transcript which would be detected as a false 3' end. To overcome this caveat, another method was published that relies on the ligation of a double stranded primer to the 3' end of the poly(A) tail and thus cannot capture internal poly(A) stretches (*293*). These methods have been used to document the 3' UTRs in several cell lines, as well as in mouse and human tissues (*289-293*). However, 3' UTRs have not been examined using NGS in the heart or during the progression of cardiac hypertrophy.

The aim of this chapter was to sequence the 3' ends of mRNA transcripts from normal, pre-hypertrophic and hypertrophic cardiomyocytes. Firstly, to document the number of 3' UTRs for each mRNA transcript, and secondly to establish if the 3' UTRs of these mRNAs change during the induction of hypertrophy. To do this we used a 3' sequencing protocol with the SOLiD™ sequencing platform and developed a bioinformatics analysis pipeline to annotate 3' ends.

5.2 Results

The RNA extracted from purified cardiomyocytes (Chapter Three) was used for NGS of mRNA 3' ends. The cardiomyocyte RNA was sent to Dr Traude Beilharz and the libraries were created using the PAT-seq method developed by Dr Beilharz at Monash University (*177*). Similar to the method published by Jan et al. (*293*), this method relies on the presence of a single-stranded poly(A) tail and thus will only capture the 3' end of the poly(A) tail and not internal poly(A) stretches (see Figure 2.3 for details). For full details of the method, refer to section 2.2.7.

5.2.1 General features of the library

The libraries were sequenced on the SOLiD 5500™ to an average depth of 4.54 x 10^7 tags (up to 75 nt in length), yielding a total of 3.63 x 10^8 tags (Figure 5.1.A). Dr Hardip Patel from ANU performed the mapping of the 3' seq libraries, and I completed the secondary analysis, including filtering, graphical representations and differential expression, using the R programming language. The data was then filtered for tags with an average quality value greater than or equal to 18 (98% accuracy), which retained 82% of the 3' sequencing tags. The filtered tags were mapped directly to the mouse genome assembly mm9 including rRNA. Those tags that did not map to the genome were then mapped to mouse transcriptome sequences to account for tags that may span exon-exon junctions. All tags that did not directly map to either genome or transcriptome may contain the beginning of the poly(A) tail in the tag. For this reason, the remaining tags were trimmed and the first 21 nt of the tag mapped to the genome and then extended until the start of a poly(A) stretch. The majority of the tags (68%) mapped to the mouse genome, with only a small proportion that mapped to the transcriptome or contained poly(A) sequences, 2.7% and 5% respectively (Figure 5.1.A).

The mapped tags were then merged to form functional units, defined as 3' sequencing peaks (3' seq peaks) (Figure 5.1.B). Overlapping or contiguous stretches of tags along the genome with at least one read coverage were merged into 3' seq peaks, resulting in 412 332 peaks. Each sequencing tag was 75 nt, therefore to remove clonal peaks, the peaks were filtered for those that had a length greater than 75 nt which resulted in 169 700 3' seq peaks. A minimum threshold was then applied, considering only peaks with an average

coverage greater than 10 and the presence in two or more samples, which resulted in 106 869 3' seq peaks. This minimum threshold was used for all subsequent analyses.

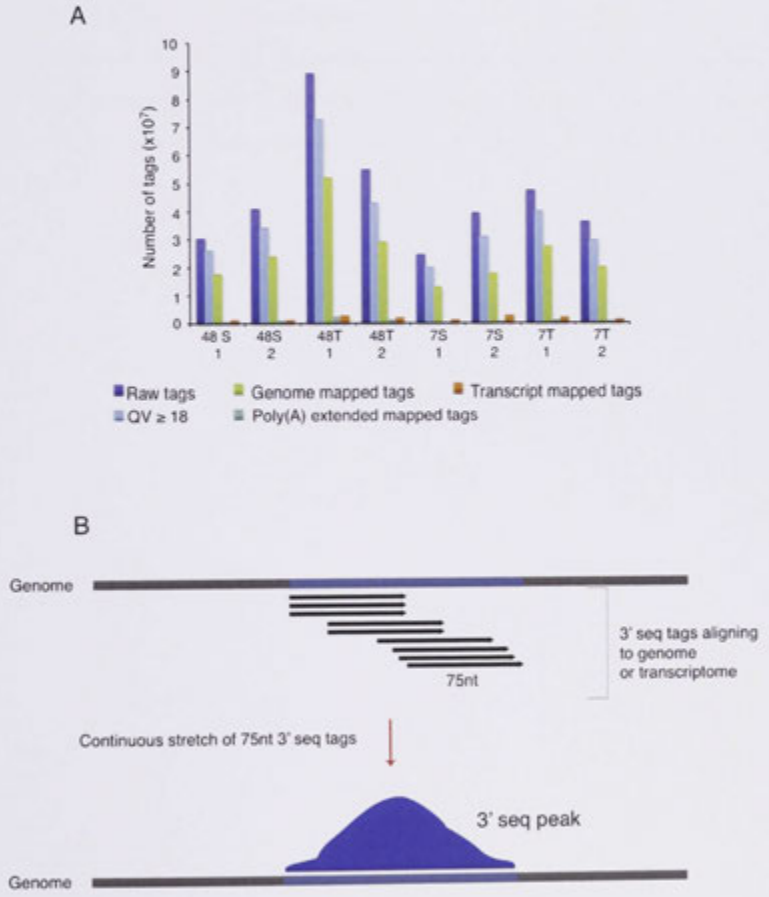


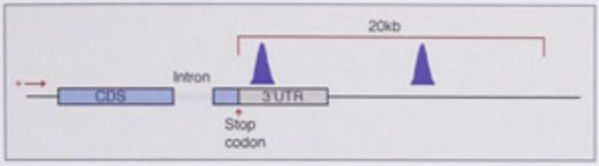
Figure 5.1. Alignment of 3' seq tags.

(A) Number of raw tags, number of tags with quality value ≥ 18, genome mapped tags, transcript mapped tags and poly(A) extended mapped tags. S, Sham; T, TAC; QV, quality value (B) Defining 3' seq tags; a continuous stretch of 75 nt 3' seq tags with at least one tag coverage were merged into 3' seq peaks representative of 3' UTR ends.

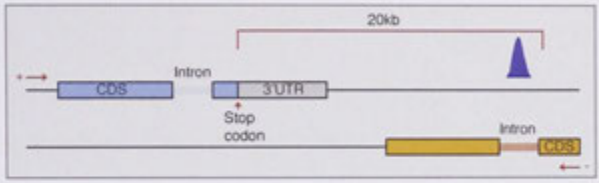
The 3' seq peaks were then assigned to genomic features using the canonical transcripts defined in Ensembl v67 (Figure 5.2). Peaks that did not overlap a known gene and were further than 20 kb downstream from the closest gene were annotated as Orphan peaks (Figure 5.2.B). The peaks that overlapped a known gene were annotated based on their location relative to the annotated stop codon or canonical transcript end. Peaks located upstream of the stop codon were termed Exon or Intron peaks respectively (Figure 5.2.C-D). Peaks located within 20 kb downstream of a stop codon or transcript end were defined as downstream peaks (Figure 5.2.A). The breakdown of peak annotations is shown in Figure 5.3. 35% of 3' seq peaks are annotated as downstream peaks (Figure 5.3.A). Non-coding RNAs known to undergo polyadenylation, such as lincRNAs and miRNAs, represent a small portion of peaks (*294, 295*). 90.6% of downstream peaks are associated with the 3' UTR of protein coding mRNAs, referred to from now on as 3' UTR peaks (Figure 5.3.B).

A Downstream Peaks

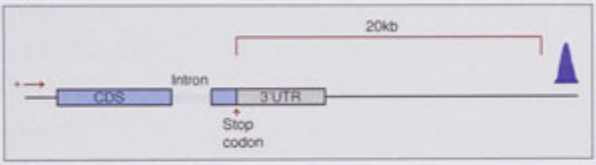
i



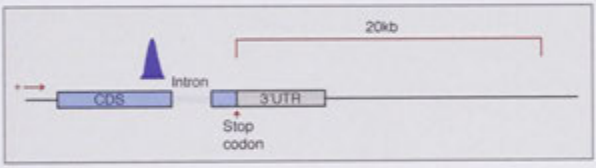
ii



B Orphan Peaks

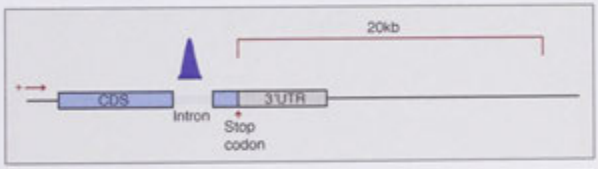


C Exon Peaks



D Intron Peaks

i



ii

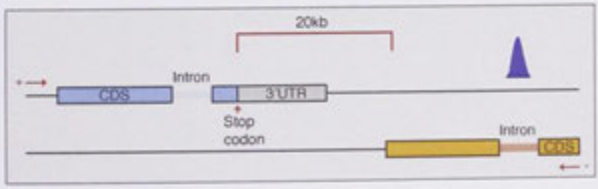


Figure 5.2. Strategy to annotate 3' sequencing peaks.

Peaks were annotated as a Downstream peak if (i) the peak overlapped a known gene and was downstream of the stop codon or the peak was within 20 kb downstream of a stop codon for protein coding transcripts or the transcript end for non-coding RNAs. (ii) the peak overlapped an intron on the opposite DNA strand and was within 20 kb downstream of a stop codon or transcript end. (B) Orphan peaks were defined as those that did not overlap with a known genomic feature and were more than 20 kb downstream of a stop codon or transcript end. (C) Exon peaks were defined as those peaks that overlapped with an exon of a known gene but upstream of the stop codon or transcript end. (D) Intron peaks were defined as those that are (i) within an intron on the same strand of the DNA or (ii) within an intron on the opposite strand of the DNA and greater than 20 kb downstream of a stop codon or transcript end on the same strand. The red arrows indicate the direction of the DNA strand. CDS, Coding region; 3' UTR, 3' untranslated region.

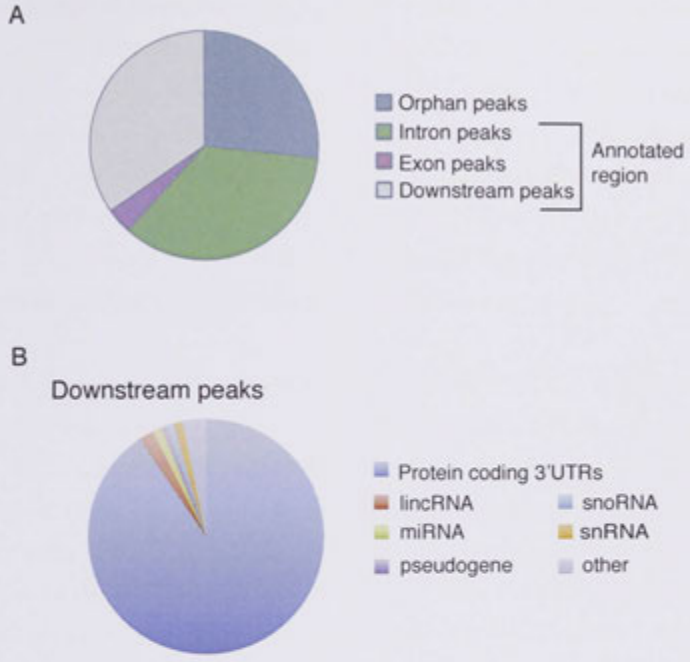


Figure 5.3. Annotation of 3' sequencing peaks.

(A) Proportion of 3' sequencing tags that were assigned as Orphan peaks, Intron peaks, Exon peaks and Downstream peaks. (B) Distribution of 36 994 Downstream peaks across genomic features. Data is shown as the sum of all eight sequencing libraries.

5.2.2 Features of Polyadenylation

Once the 3' peaks had been categorised, we looked for features of polyadenylation. Due to the nature of polyadenylation and the library preparation, it was expected that some of the peaks would contain the beginning of the poly(A) tail (Figure 5.4.A). For instance, as RNAse T1 was used to digest the RNA, the cleavage site from the start of the poly(A) tail is variable as shown in Figure 5.4.A. Figure 5.4.B-C shows the proportion of peaks that contain a stretch of A's at the 3' end of the peak, for all peaks (5.4.B) or those peaks associated with protein coding 3' UTRs (5.4.C). A higher proportion of 3' UTR peaks contain evidence of a poly(A) tail compared to all of the 3' seq peaks.

We also looked for the presence of the polyadenylation signal (PAS) in the 3' seq peaks. The conserved canonical PAS (AAUAAA) or single-nucleotide variants, defined as a non-canonical PAS, are the main signals used to determine the site of cleavage and polyadenylation (*160*). The canonical PAS was searched for within 50 nt upstream of the poly(A) for peaks containing the poly(A), or within the 3' seq peak and 50 nt downstream of the 3' end of the peak for peaks without poly(A). Peaks without the canonical PAS in this region were interrogated for the presence of a non-canonical PAS. 69% of 3' seq peaks contain either the canonical PAS or one of the twelve non-canonical PAS variants. This proportion increases to 76% when considering only the peaks associated with protein coding 3' UTRs.

The peaks were then split further into those containing poly(A) stretches and those without poly(A) stretches. The breakdown for poly(A) containing peaks and polyadenylation signals is shown in Table 5.1. The peaks containing evidence of a poly(A) tail had a higher proportion of tags with either the canonical or a non-canonical PAS. The most likely reason for this is that the region to search for the PAS in those peaks containing a poly(A) stretch is well defined, whereas the PAS for those peaks without a poly(A) stretch may be further downstream of the peak and thus the PAS was not detected. Despite this, between 65% and 80% of peaks contained either a poly(A) stretch and/or a PAS.

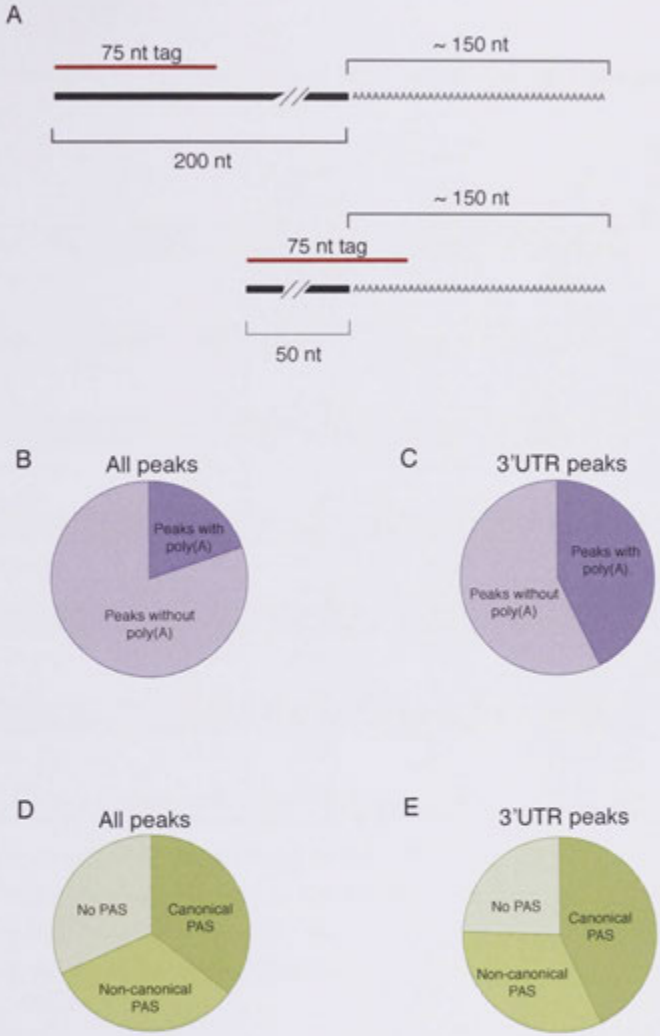


Figure 5.4 Polyadenylation features are enriched in peaks associated with protein 3' UTRs. A) Variation in the RNAse T1 digestion site can change the distance of the tag start site to the poly(A) tail. (B-C) The proportion of peaks that contain evidence of a poly(A) tail at the 3' end of the peak in total peaks (B) or protein 3' UTR peaks (C). (E-F) The proportion of peaks that contain the canonical PAS or non-canonical PAS in total peaks (E) or protein 3' UTR peaks (F).

Table 5.1. Proportion of peaks containing evidence of a poly(A) tail, canonical and non-canonical PAS.

		Canonical PAS (%)	Non-canonical PAS (%)	No PAS (%)	Peaks with a poly(A) feature (%)
Total peaks	106869	36.05	33.10	32.00	
Poly(A) peaks (%)	19.75	57.29	35.56	12.91	70.55
No poly(A) (%)	80.25	30.82	32.49	36.69	
Biotype associated peaks	36994	42.71	32.19	25.10	
Poly(A) peaks (%)	40.27	59.57	30.06	10.37	79.07
No poly(A) (%)	59.73	31.34	33.62	35.04	
Orphan peaks	28487	35.30	29.21	35.49	
Poly(A) peaks (%)	9.85	53.26	29.32	17.42	66.22
No poly(A) (%)	90.15	33.34	29.19	37.47	
Intron peaks	37470	31.49	32.67	35.84	
Poly(A) peaks (%)	6.91	52.14	28.33	19.53	65.51
No poly(A) (%)	93.09	29.96	32.99	37.05	
Exon peaks	3918	22.05	43.06	34.89	
Poly(A) peaks (%)	20.83	45.83	31.37	22.79	69.86
No poly(A) (%)	79.17	15.80	46.13	38.07	
Protein 3' UTR peaks	33507	43.07	32.59	24.35	
Poly(A) peaks (%)	42.87	59.79	30.13	10.08	79.97
No poly(A) (%)	57.13	30.52	34.43	35.05	

The total number of tags for each annotation group is shown in the grey box. The proportion of peaks with or without evidence of poly(A) tail are listed below. The proportion of tags that contain a canonical PAS, non-canonical PAS or no PAS are listed in the rows, with breakdowns for the total peaks, poly(A) peaks and no poly(A) containing peaks. The right hand row documents the proportion of tags that contain a poly(A) signature and/or a PAS.

5.2.3 Peaks associated with 3' UTRs

One of the main aims of this thesis was to look at the 3' UTRs of mRNA transcripts and alternative polyadenylation. Therefore, the remainder of the analysis on 3' seq peaks was restricted to those peaks associated with the 3' UTRs of protein coding genes (termed 3' UTR peaks). As mentioned above, 43% of 3' UTR peaks contain evidence of the canonical PAS. The location of the PAS relative to the 3' UTR peak was then determined. As discussed in section 1.1.3, a number of sequence elements determine cleavage and polyadenylation (Figure 5.5.A). This includes the PAS which is located 10-35 nt upstream of the poly(A) cleavage site (*148*), U-rich elements upstream and around the cleavage site and a GU-rich DSE (reviewed in (*160*)). While all of these elements contribute to and enhance cleavage and polyadenylation, the PAS is the dominant signal (*160*). Therefore, we determined the location of the PAS relative to the start of the poly(A) stretch or the end of the peak for those 3' peaks without evidence of a poly(A) stretch. The canonical PAS was found predominantly 17.5 nt upstream of the end of the 3' seq peak, which is within the expected region of the PAS(Figure 5.5.B.i). The signal is also found at low levels beyond the peak, which can be attributed to those peaks without a poly(A) stretch where the poly(A) cleavage site may be further downstream. In agreement, a higher proportion of peaks are located 16 nt upstream of the start of the poly(A) stretch (Figure 5.5.B.ii).

We next looked at the peaks that did not contain the canonical PAS, but one of the 12 non-canonical PAS (Figure 5.6.A) (*160*). For all protein 3' UTR peaks, the second most abundant PAS was AUUAAA, which is consistent with a previous study that documented mouse PAS (*160*). When we considered the peaks with and without a poly(A) stretch separately, AUUAAA was the more predominant PAS in peaks with a poly(A) (Figure 5.6.C), while AAGAAA was the most frequent PAS in peaks without a poly(A) stretch (Figure 5.6.B). As for the canonical PAS, the distance of the non-canonical PAS from the peak end or poly(A) stretch was then determined. When considering all protein coding peaks, AUUAAA has a predominant peak between 10-25 nt upstream of the peak end, while AAGAAA has a peak that centres around 0. Again, the location of the PAS is more obvious in the peaks that contain a poly(A) stretch, where most PAS variants have a peak in the expected location, with the exception of UUUAAA that centres around the start of the poly(A) tail.

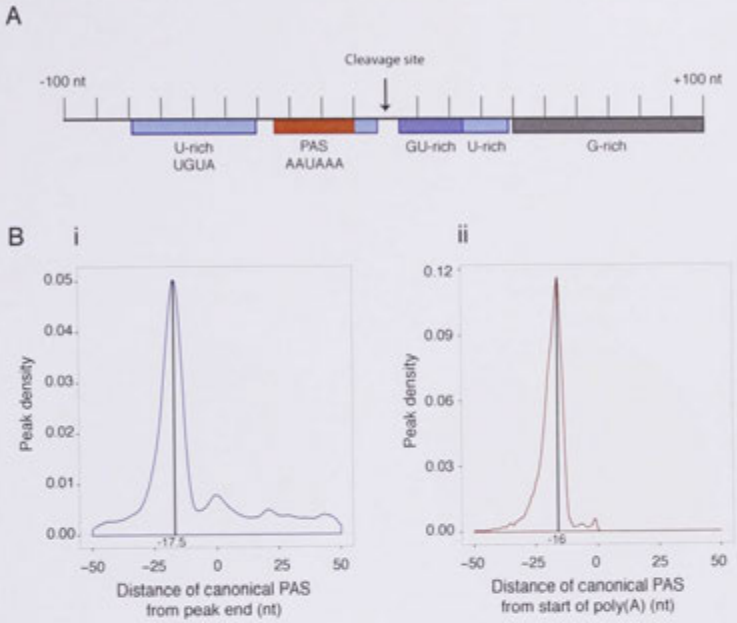


Figure 5.5. Canonical PAS is found at the expected position from the cleavage site.

(A) Schematic of sequence elements that surround the poly(A) cleavage site. Modified from (*159*). (B) The location of the canonical PAS (AAUAAA) in peaks that contain the canonical PAS. (i) The density of the canonical PAS from the end of the 3' seq peak. (ii) The density of the canonical PAS from the start of the poly(A) in peaks that contain evidence of poly(A) tail. The density is a proportion, with the area under the curve equal to 1.

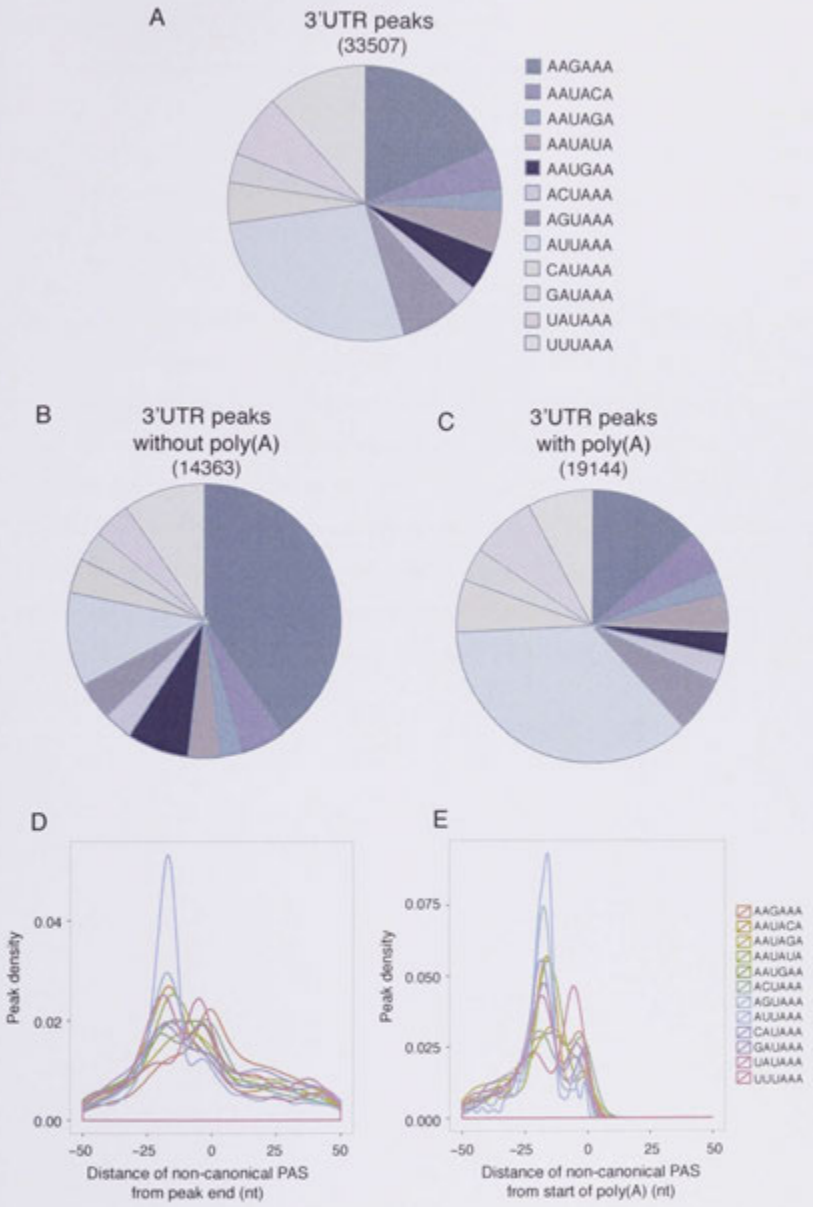


Figure 5.6. Non-canonical PAS in protein coding associated peaks.

(A) The proportion of non-canonical PAS in protein coding associated peaks that do not contain the canonical PAS. (B-C) Breakdown of non-canonical PAS in protein coding peaks with (C) and without (B) evidence of a poly(A) tail. (D) The location of non-canonical PAS in all protein coding associated peaks. (E) The location of non-canonical PAS in protein coding associated peaks that contain a poly(A) tail. The density is a proportion, with the area under the curve equal to 1.

We next analysed the 3' UTR peaks for their association with mRNAs. On average, one mRNA transcript was associated with 2.35 3' seq peaks representing individual 3' UTRs (Figure 5.7.A). Despite allowing a distance of 20 kb downstream of a stop codon, more 3' seq peaks occurred closer to the stop codon, with a mean distance of 3.7 kb, median distance of 1.6 kb and the maximum peak density at approximately 350 nt downstream of the stop codon (Figure 5.7.B).

To determine how many novel 3' ends were in our dataset, the 3' seq peaks were compared to the transcript coordinates in Ensembl v67. A 3' seq peak was considered a known 3' end if the peak overlapped the last exon of a transcript and the 3' end of the peak was within 100 nt of the end of the transcript (Figure 5.7.C). 34% of 3' seq peaks aligned with a known 3' end and were annotated as known 3' UTRs. 37% did not overlap with a known 3' UTR, and the remaining 29% overlapped the last exon of a transcript but were further than 100 nt away from the known 3' end, representing shorter 3' UTRs. These 3' UTRs represent novel 3' ends. The 34% of tags that aligned with a known 3' end accounted for 69% of the tags, suggesting a tendency of highly expressed 3' UTRs to already be annotated.

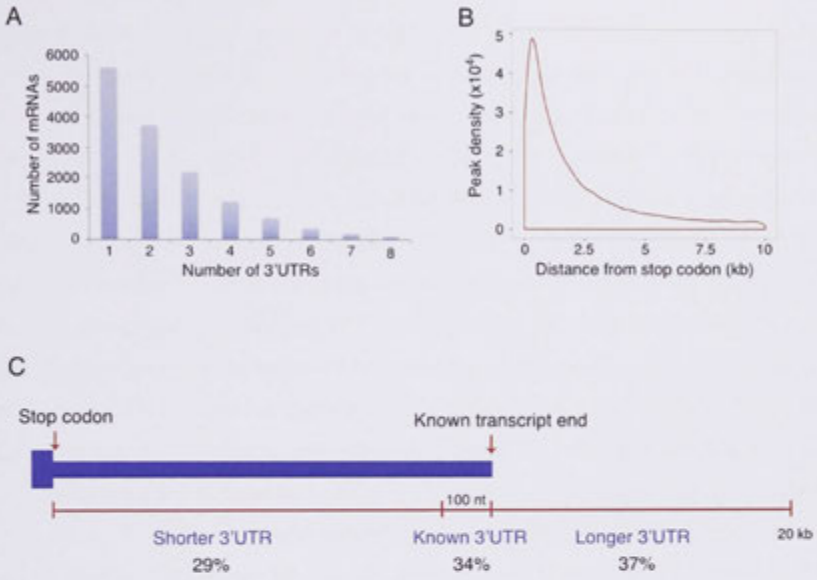


Figure 5.7. Majority of protein coding transcripts have multiple 3' ends.

(A) The number of transcripts with 1-8 3' UTRs is plotted. (B) The location of peaks from the stop codon in protein coding transcripts. The Peak density is a proportion, with the area under the curve equal to 1. (C) Assigning peaks to known 3' ends; peaks overlapping a known 3' UTR and within 100 nt from a known transcript end were annotated as Known 3' UTRs. Peaks that overlap the 3' UTR but were further than 100 nt from the transcript end are Shorter 3' UTRs, and peaks that do not overlap the 3' UTR were annotated as Longer 3' UTRs.

5.2.4 Cardiac mRNAs with multiple 3' ends

Firstly we wanted to establish the static picture of mRNA 3' UTRs. In particular, the 3' UTR status of mRNAs implicated in cardiac hypertrophy. 803 genes related to cardiac hypertrophy were collected from the GeneCards database (*296*). Of these, 652 could be detected in our sequencing data. The number of 3' seq peaks associated with each gene ID was then calculated. Similar to the overall dataset, on average each gene had 2.65 3' ends (Figure 5.8). However, this list was filtered further by looking only at 3' seq peaks that contributed at least 10% of the tags towards the mRNA transcript. In doing so, the average number of 3' ends per gene was reduced to 1.44, indicating that a large number of 3' seq peaks contributed only a small proportion to the whole mRNA transcript (Figure 5.8).

Table 5.2 summarises the number of 3' UTRs for the detectable genes associated with cardiac hypertrophy. Cardiac hypertrophy genes that contain one or more novel 3' ends are highlighted in blue in Table 5.2. Notably, for some of the genes with only one 3' end this 3' UTR is novel. It is important to note that if the peak did not overlap with a known 3' UTR it was considered novel regardless of its distance to the closest upstream transcript end (Figure 5.7.C). Examples of cardiac hypertrophy related mRNAs with two, three or four 3' ends are illustrated in Figure 5.9. For the mRNAs visually inspected, the novel 3' UTRs tend to be shorter than the annotated 3' UTR.

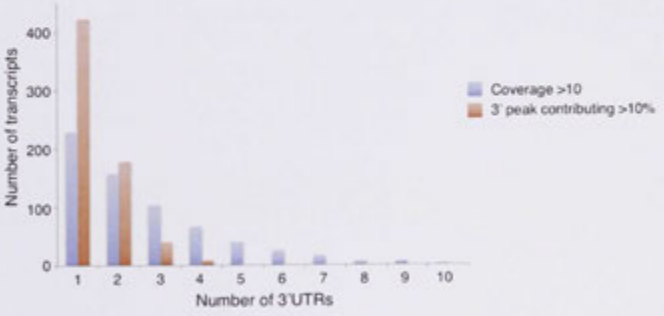


Figure 5.8. Number of 3' UTRs for cardiac hypertrophy associated genes.

The number of cardiac hypertrophy genes with 1-10 3' seq peaks, which represent individual 3' UTRs. Data is shown for all 3' seq peaks associated with cardiac hypertrophy genes (blue) or only 3' seq peaks that contribute more than 10% to the mRNA.

Table 5.2. The number of 3' UTRs for mRNAs associated with cardiac hypertrophy.

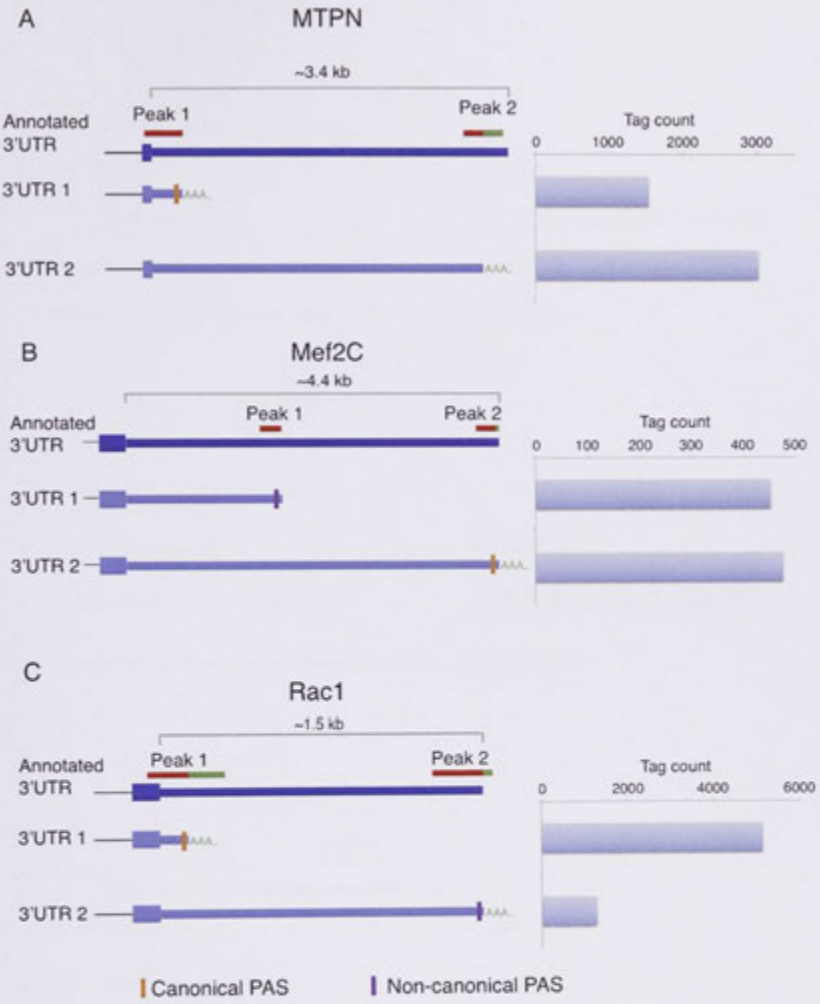
The number of 3' UTRs is based on peaks with a coverage >10, presence in more than 2 samples and contribution of greater than 10% to the mRNA transcript. Genes highlighted in blue indicate that the gene has one or more novel 3' ends.

One 3' UTR
Abcg2, Abra, Acadm, Acadvl, Acta1, Acta2, Actb, Actc1, Actg2, Actn2, Ada, Adcy3, Adcy6, Adipoq, Adk, Adm, Adra1b, Adra2a, Adra2c, Adrb3, Agk, Agt, Agtr1a, Aifm1, Akap1, Akap13, Akt1, Aldh2, Ankrd1, Ankrd2, Anxa5, Apoa1, Atf2, Atf3, Atf4, Atp2a1, Atp2a2, Atp2a3, Bak1, Bcar1, Bdkrb2, Bglap, Bmp1, Bmp10, Bmp4, C3, Cabin1, Cacna1a, Cacna1b, Cacna1g, Cacna1h, Cacna1s, Cacna2d1, Cacna2d2, Cacnb3, Cacng1, Cacng2, Cacng4, Cacng5, Cacng7, Calcrl, Calm2, Calr3, Camk1, Camk2g, Camk4, Camta2, Casp9, Cat, Cav1, Cav3, Cck, Cdk7, Cdkn1b, Cflar, Chkb, Chrm2, Ckm, Ckmt1b, Cma1, Cntf, Col1a1, Col3a1, Corin, Cox6c, Cpe, Creb3, Creb3l4, Cryab, Cs, Csk, Csrp3, Cst3, Ctgf, Ctnnb1, Ctsl1, Cxcr2, Cyba, Cybb, Cycs, Cyp19a1, Cyp1b1, Cyp2e1, Ddah2, Des, Dusp1, Ece1, Edf1, Edn1, Ednra, Egf, Egfr, Egr1, Eif2b1, Eif2b2, Eif2b4, Eif2b5, Eno2, Ep300, Epor, Eprs, Erbb2, Esrrg, F2rl1, Fabp3, Fgf9, Fgfr2, Fgfr3, Fhl1, Fkbp4, Flnb, Fn1, Fos, G6PD, Gaa, Gadd45b, Galr2, Gas6, Gast, Gata4, Gata6, Gca, Gdf1, Gh1, Ghr, Gja1, Glrx3, Gna12, Gna13, Gnai1, Gnai3, Gnb2, Gnb3, Gng3, Gng4, Gpbar1, Gpr119, Gprc6a, Gpsm2, Gsk3a, Gstp1, H2afz, Hand1, Hand2, Hbegf, Hcn2, Hcn4, Hdac1, Hdac10, Hdac11, Hdac2, Hdac5, Hdac6, Hdac7, Hdac8, Hdac9, Hey1, Hey2, Hgf, Hmga1, Hmox1, Hnf1a, Hopx, Hp, Hsf1, Hspa4, Hspa5, Hspb1, Hspb8, Htr2a, Htr2b, Igf1, Igf2, Igfbp3, Ikbkb, Il11, Il15, Il4R, Il6st, Inpp5f, Ins, Insr, Irs1, Itga1, Itga2, Itga2b, Itga3, Itga4, Itga5, Itga6, Itga8, Itgav, Itgax, Itgb1, Itgb1bp2, Itgb2, Itgb5, Itgb6, Itgb7, Itgb8, Itpr2, Jak3, Jph2, Kat2b, Kcnh2, Kdm4a, Klf15, Krit1, Lama2, Lect1, Lepr, Lmcd1, Lmna, Lpl, Map2k1, Map2k5, Map3k11, Map3k12, Map3k13, Map3k14, Map3k4, Map3k5, Map3k6, Mapk11, Mapk12, Mapk13, Mapk14, Mapk3, Mapk4, Mapk6, Mapkapk3, Met, Mgp, Mkks, Mkl1, Mkl2, Mme, Mmp2, Mmp3, Mterf, Mtor, Mybpc1, Mybpc2, Mybpc3, Myef2, Myh6, Myh7, Myl3, Myl4, Mylk2, Myocd, Myog, Myot, Myoz2, Nedd4, Nes, Nexn, Nfatc1, Ngf, Nisch, Nppa, Nppb, Nppc, Npr1, Nr1h2, Nr3c2, Nras, Nrg1, Nrip1, Ntrk3, Nupr1, P2rx4, Pcna, Pde1a, Pde1c, Pde5a, Pdgfrb, Pdha1, Pik3r2, Pik3r3, Pik3r4, Pitx2, Plat, Plcd4, Plce1, Plcg1, Plcg2, Plcl1, Pld2, Pnkd, Polr1c, Pomc, Ppara, Ppard, Pparg, Ppbp, Ppig, Ppp2r4, Ppp3cb, Ppp3cc, Ppp3r1, Prkab1, Prkaca, Prkar1b, Prkar2b, Prkcb, Prkcd, Prkch, Prkcz, Prkg1, Psmc4, Ptgs1, Pthlh, Ptk2b, Ptpn11, Rac2, Rac3, Ramp2, Ramp3, Rapgef3, Rcan1, Retn, Rgs2, Rgs4, Rhod, Rnls, Rock1, Rps27a, Rrad, Rras, Ryr2, S100a1, Scarb1, Sdhb, Selp, Serpine1, Setd2, Sfn, Sfrp2, Shbg, Shc1, Sirt1, Slc29a2, Slc2a1, Slc2a4, Slc9a1, Sln, Smad3, Smn1, Socs3, Sod1, Sos1, Sos2, Sox4, Spon2, Spp1, St3gal4, Stat3, Stk11, Tac1, Tbpl2, Tcap, Tfb1m, Tgfa, Tgfb3, Tgfbr1, Tgfbr3, Thbs1, Timp1, Tkt, Tnc, Tnf, Tnfaip3, Tnfrsf11b, Tnfrsf1a, Tnnc1, Tnni3, Tnni3k, Tnnt2, Tpm1, Tpm4, Trap1, Trim63, Trpc3, Trpc6, Ttn, Ttr, Tubb, Txn, Txnip, Ubtf, Vcam1, Vcl, Vegfa, Vegfb, Vim, Vip, Wnt5a, Xdh, Xirp1, Ywhag, Zfpm2
Two 3' UTRs
Abcc9, Ablim3, Acadl, Ace, Actg1, Actn1, Adam17, Adra1a, Ahr, Ak1, Akap5, Akt2, Apln, Aplnr, Asph, Atf1, Atp2b4, Bmp6, Braf, Cacna1c, Cacnb1, Cacnb2, Cacnb4, Calm1, Calm3, Camk2b, Camk2d, Camp, Casq2, Cast, Ccl2, CD36, CDC42, Cdh2, Cdk9, Cdkn3, Chka, Cntfr, Creb1, Crebbp, Creg1, Ctf1, Dmd, Dusp19, Ednrb, Eif2b3, Eif4e, Eif4ebp1, Elk1, Eln, Enpp1, Erbb3, Fblim1, Fbxo32, Fgf1, Fgf23, Fgfr1, Fgfr4, Fkrp, Foxo1, Gal, Gdf15, Gna11, Gnai2, Gnaq, Gnb1, Gnb5, Gng2, Grb2, Grk5, Hdac3, Hdac4, Hdc, Hfe, Hmgcr, Hras, Idua, Il1b, Il6, Itga10, Itgb4, Jak2, Junb, Kcnj5, Lgals3, Lifr, Lpar1, Lrp6, Map2k2, Map2k3, Map2k4, Map2k6, Map3k1, Map3k10, Map3k3, Map3k7, Map3k8, Mapk1, Mapk10, Mapk8, Mapk9, Mapkapk2, Mef2a, Mef2c, Mef2d, Mras, Mstn, Mthfr, Mtpn, Myc, Mylk, Mylk3, Nfatc4, Nfkb1, Nkx2-5, Nos2, Npr2, Npr3, Odc1, Osbpl3, Osm, Parp1, Pdgfa, Pdgfra, Pdk1, Pdlim5, Pdp1, Pdpk1, Pik3ca, Pik3cb, Pik3r1, Plcb2, Plcb4, Pld1, Pln, Pol2rl, Ppargc1a, Prkaa1, Prkacb, Prkag2, Prkar1a, Prkar2a, Prkce, Prkci, Prkcq, Pten, Ptgs2, Ptp4a3, Pxn, Rac1, Raf1, Rgs5, Rhoa, Rps6ka1, Rps6ka5, Rras2, Slc8a1, Smad2, Smad4, Sp3, Srf, Stat6, Tagln, Tead1, Tfam, Tfb2m,

Table 5.2 cont. The number of 3' UTRs for mRNAs associated with cardiac hypertrophy.

The number of 3' UTRs is based on peaks with a coverage >10, presence in more than 2 samples and contribution of greater than 10% to the mRNA transcript. Genes highlighted in blue indicate that the gene has one or more novel 3' ends.

Two 3' UTRs
Tgfb1, Tgfbr2, Thbs2, Thra, Timp2, Trdn, Twf1, Ubc, Uts2r, Vdr, Wt1, Yes1, Ywhaq, Yy1
Three 3' UTRs
Ace2, Acp1, Adra1d, Apc, Atf6, Camk2a, Ccnd2, Cpt1a, Csnk1a1, Erbb4, F2, F2r, Fgf2, Foxp1, Fzd4, Gla, Gnal, Gnb4, Hexim1, Id2, Ift81, Itga9, Itpr1, Jak1, Jun, Map3k2, Myl2, Nab1, Pde4a, Plcb3, Plxna2, Prkca, Prkd1, Rock2, Rps6kb1, Scn4a, Soat1, Sp1, Tead4, Zbtb16
Four 3' UTRs
Gsk3b, Igf1r, Itga11, Ppp3ca, Prkaa2, Ptk2, Smc3, Tgfb2, Vcan


A
MTPN
~3.4 kb
Peak 1
Peak 2
Annotated 3'UTR
3'UTR 1
3'UTR 2
Tag count
0
1000
2000
3000
B
Mef2C
~4.4 kb
Peak 1
Peak 2
Annotated 3'UTR
3'UTR 1
3'UTR 2
Tag count
0
100
200
300
400
500
C
Rac1
~1.5 kb
Peak 1
Peak 2
Annotated 3'UTR
3'UTR 1
3'UTR 2
Tag count
0
2000
4000
6000
Canonical PAS
Non-canonical PAS

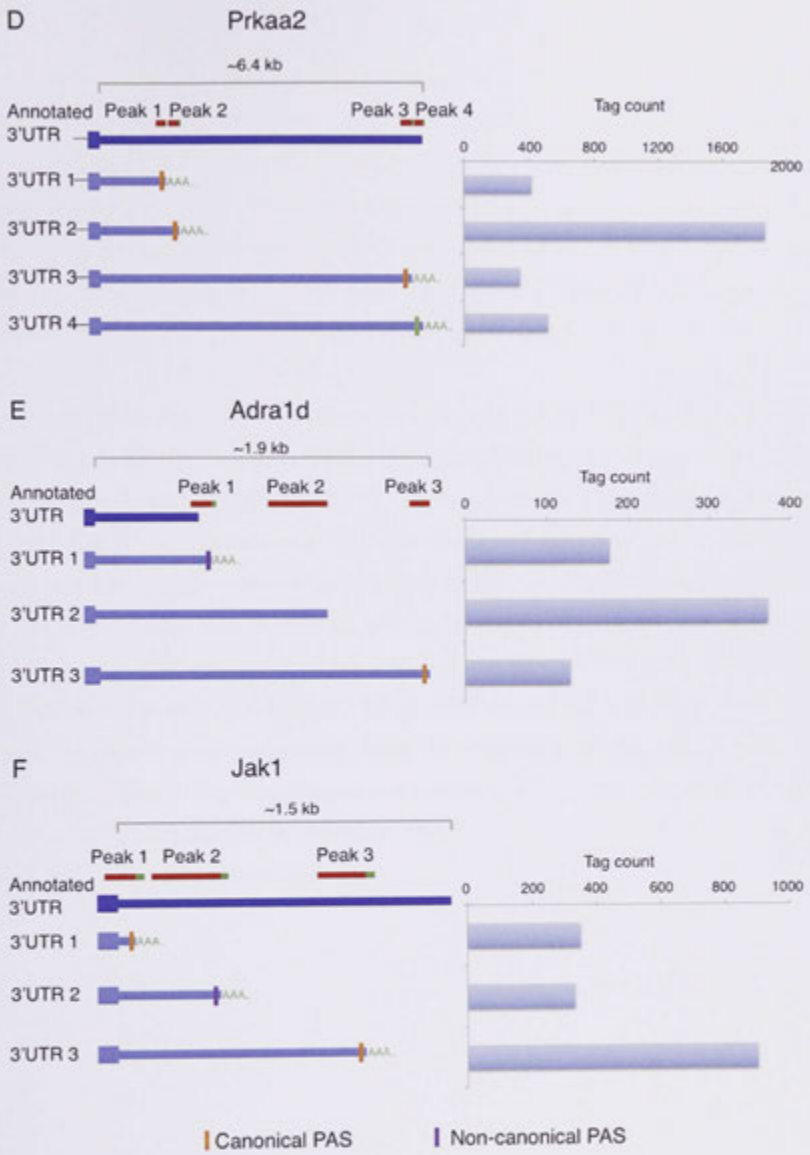


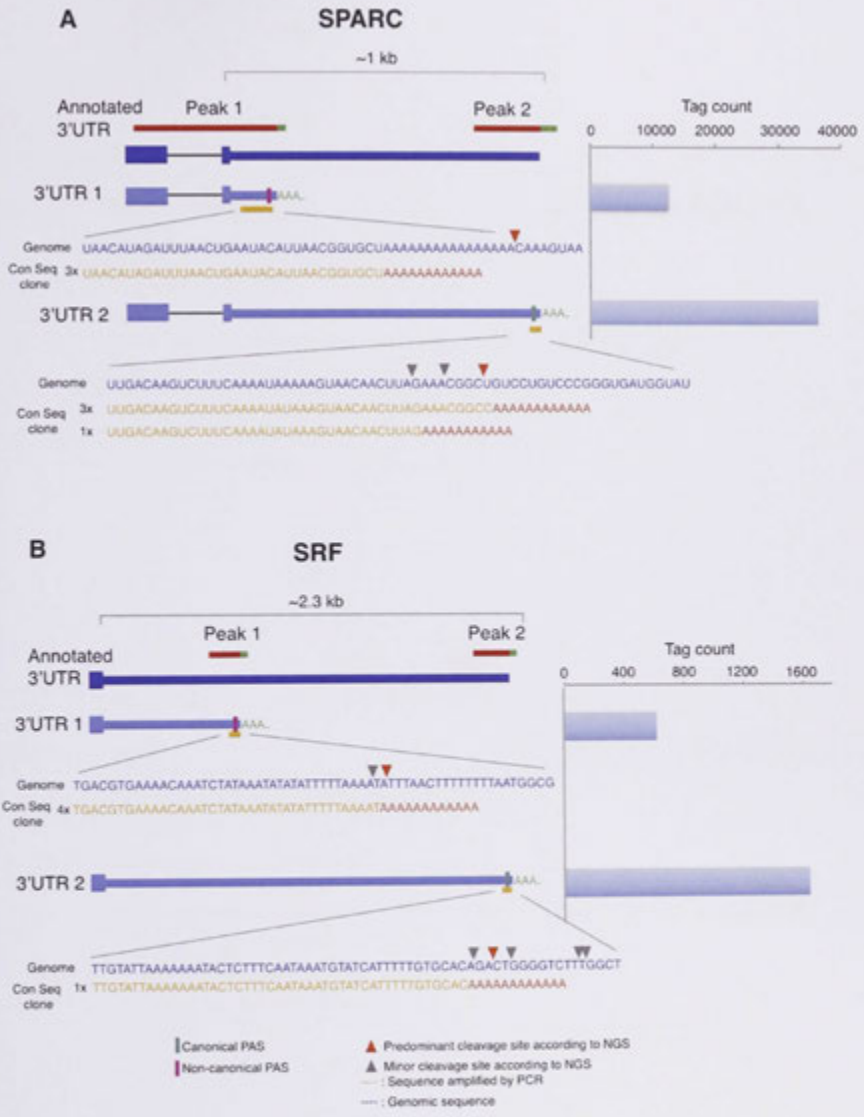
Figure 5.9 Cardiac mRNAs with multiple 3' ends.

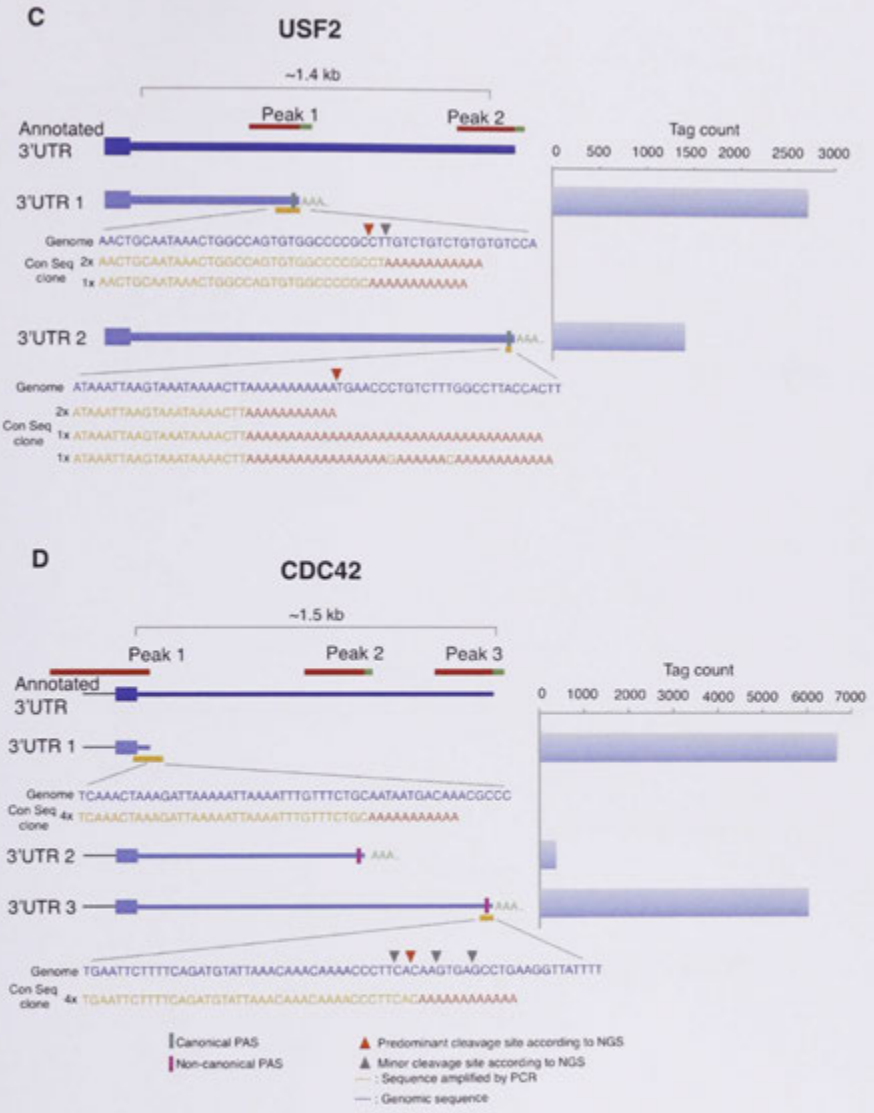
Schematics of the annotated 3' UTR and 3' seq peaks are shown for some cardiac mRNAs with two, three or four 3' seq peaks. The dark blue boxes represent the annotated 3' UTRs, with the location of the 3' seq peaks indicated above by the red lines. Peaks that contained a poly(A) tail are indicated by the green line. Each 3' UTR represented by the 3' seq peak is drawn below, with their corresponding tag count on the right. The coloured bars represent the canonical or non-canonical PAS.

5.2.5 Validation of mRNAs with multiple 3' ends

The 3' ends of several cardiac mRNAs were mapped by an independent 3' RACE method. cDNA was made using a TvN anchor, which captures the end of the mRNA at the point where the poly(A) tail begins. This is achieved by twelve T's, followed by a G,A or C and then any nucleotide. After reverse transcription, PCR is performed using a reverse primer that detects the anchor sequence and a forward primer specific to the 3' UTR of interest. The PCR products were size selected and then subjected to conventional sequencing. Rina Soetanto from the ANU performed the validation of 3' ends.

Six genes with multiple 3' UTRs were tested. Of these, all of the 3' ends detected with 3' sequencing could be validated using conventional sequencing (Figure 5.10). The full sequences are listed in Appendix 8.10. The transition point into a poly(A) tail found in the conventional sequencing data were within the locality of the poly(A) tail as determined by NGS. In some cases, if multiple poly(A) start positions were present in the 3' seq peak then not all cleavage sites were confirmed with conventional sequencing. This is most likely because only a small number of clones were used for analysis. One peak, CDC42 Peak 1, did not contain evidence of a poly(A) tail in the 3' sequencing data (Figure 5.10.D). As a result, the conventional sequencing maps the beginning of the poly(A) tail slightly downstream of the 3' seq. To summarise, the location of the 3' seq peak could be validated for all mRNAs investigated. Furthermore, the cleavage sites present in the 3' sequencing data were often visible using conventional sequencing.


A
SPARC
~1 kb
Annotated 3'UTR
Peak 1
Peak 2
Tag count
0
10000
20000
30000
40000
3'UTR 1
3'UTR 2
Genome
Con Seq clone
3x
1x
B
SRF
~2.3 kb
Peak 1
Peak 2
Annotated 3'UTR
Tag count
0
400
800
1200
1600
3'UTR 1
3'UTR 2
Genome
Con Seq clone
4x
Canonical PAS
Non-canonical PAS
Predominant cleavage site according to NGS
Minor cleavage site according to NGS
: Sequence amplified by PCR
: Genomic sequence


C
USF2
~1.4 kb
Peak 1
Peak 2
Annotated 3'UTR
Tag count
0 500 1000 1500 2000 2500 3000
3'UTR 1
AAA
Genome AACTGCAATAAACTGGCCAGTGTGGCCCCGCCTTGTCTGTCTGTGTGTCCA
Con Seq 2x AACTGCAATAAACTGGCCAGTGTGGCCCCGCCTAAAAAAAAAAAA
clone 1x AACTGCAATAAACTGGCCAGTGTGGCCCCGCAAAAAAAAAAAA
3'UTR 2
AAA
Genome ATAAATTAAGTAAATAAAACTTAAAAAAAAAAATGAACCCTGTCTTTGGCCTTACCACTT
Con Seq 2x ATAAATTAAGTAAATAAAACTTAAAAAAAAAAA
clone 1x ATAAATTAAGTAAATAAAACTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
1x ATAAATTAAGTAAATAAAACTTAAAAAAAAAAAAAAAAGAAAAAACAAAAAAAAAAAA
D
CDC42
~1.5 kb
Peak 1
Peak 2
Peak 3
Annotated 3'UTR
Tag count
0 1000 2000 3000 4000 5000 6000 7000
3'UTR 1
Genome TCAAACTAAAGATTAAAAATTAAAATTTGTTTCTGCAATAATGACAAACGCCC
Con Seq 4x TCAAACTAAAGATTAAAAATTAAAATTTGTTTCTGCAAAAAAAAAAA
clone
3'UTR 2
AAA
3'UTR 3
AAA
Genome TGAATTCTTTTCAGATGTATTAAACAAACAAAACCCTTCACAAGTGAGCCTGAAGGTTATTTT
Con Seq 4x TGAATTCTTTTCAGATGTATTAAACAAACAAAACCCTTCACAAAAAAAAAAAA
clone
Canonical PAS
Non-canonical PAS
Predominant cleavage site according to NGS
Minor cleavage site according to NGS
: Sequence amplified by PCR
: Genomic sequence

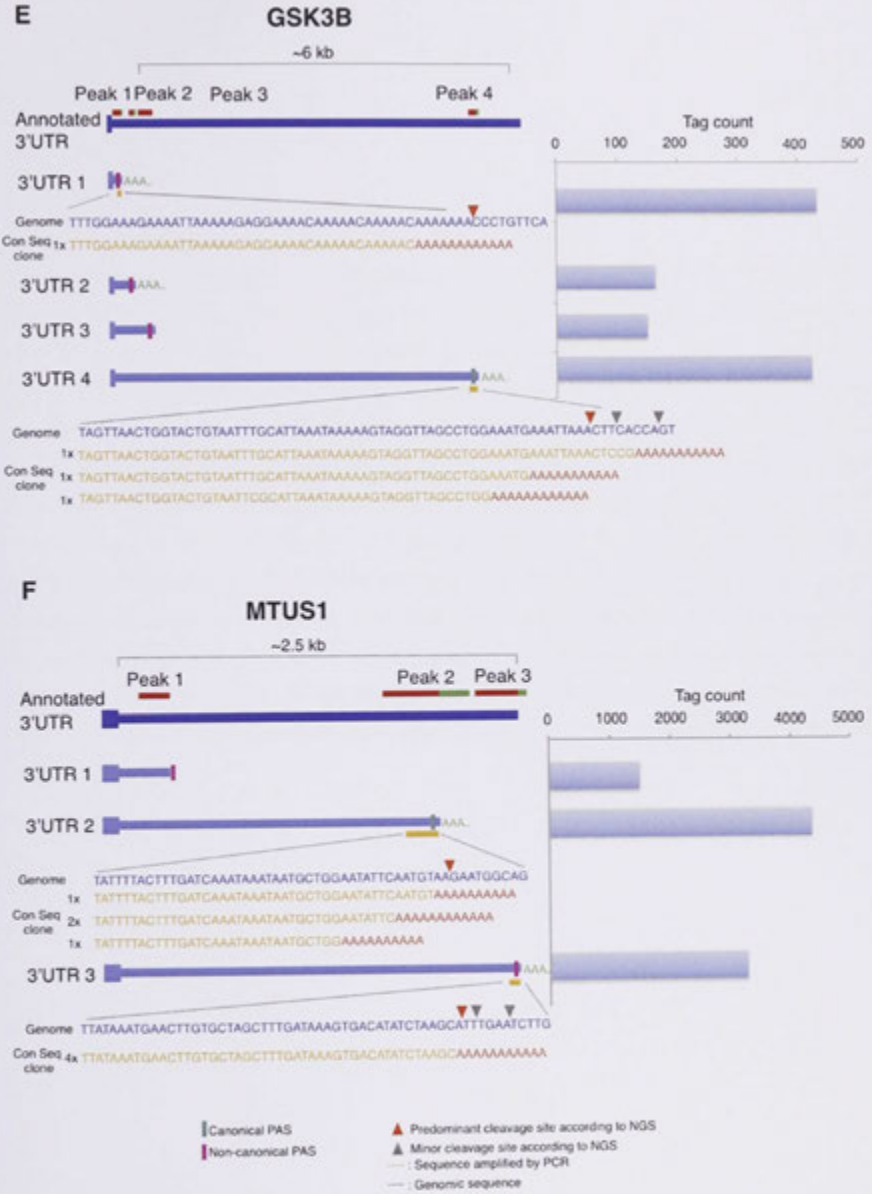

E
GSK3B
~6 kb
Peak 1
Peak 2
Peak 3
Peak 4
Annotated 3'UTR
Tag count
0
100
200
300
400
500
3'UTR 1
3'UTR 2
3'UTR 3
3'UTR 4
Genome
Con Seq clone
F
MTUS1
~2.5 kb
Peak 1
Peak 2
Peak 3
Tag count
0
1000
2000
3000
4000
5000
Canonical PAS
Non-canonical PAS
Predominant cleavage site according to NGS
Minor cleavage site according to NGS
: Sequence amplified by PCR
: Genomic sequence

Figure 5.10. Validation of 3' ends detected with 3' sequencing.

Schematics of the annotated 3' UTR and 3' seq peaks for validated mRNAs are shown (A) SPARC, secreted protein, acidic, cysteine-rich (B) SRF, serum response factor (C) USF2, upstream transcription factor 2 (D) CDC42, cell division cycle 42 (E) GSK3B, glycogen synthase kinase 3 beta (F) MTUS1, microtubule associated tumor suppressor 1. The dark blue boxes represent the annotated 3' UTRs, with the location of the 3'seq peaks indicated above by the red lines. Peaks containing a poly(A) tail are shown by the overlapping green line. Each 3' UTR represented by the 3' seq peak is drawn below, with their corresponding tag count on the right. The coloured vertical bars represent the canonical (green) or non-canonical (purple) PAS. The Yellow lines indicate the area that was validated by conventional sequencing and contained a poly(A) stretch. The region validated by conventional sequencing is zoomed, showing the genomic sequence (blue) and the sequence found by conventional sequencing (yellow). The number of clones supporting each transition is shown to the left. The red triangle indicates the predominant cleavage site found in the NGS, and the grey triangles represent minor cleavage sites. Con seq clone; conventional sequencing clone.

5.2.6 3' SEQ DATA AS PROXY FOR GENE EXPRESSION

Prior to looking at changes in 3' UTR usage, the 3' seq data was used as a measurement for overall differential gene expression. In order to get the overall expression for each mRNA transcript, the protein coding 3' UTR peaks associated with the same gene ID were added together (Figure 5.11.A). The 33 507 peaks associated with protein coding genes were merged to give 14 232 expressed genes. The contribution of individual genes is shown in Figure 5.11.B. Noticeably, there is a larger range in the expression values of the mRNAs compared to the miRNA data (see Figure 4.3.D). The top expressing mRNA, Cox2, contributed 6.9% of all the protein coding tags (Figure 5.10.B) and the top 20 mRNAs contribute 32.7% of tags. In fact, the top 1000 mRNAs contribute to 87% of the tags.

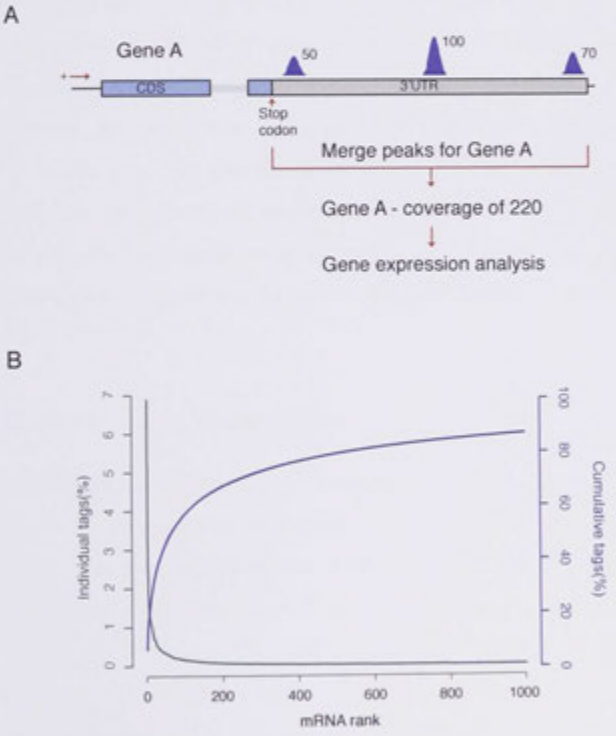


Figure 5.11 Merging 3' seq peaks for gene expression analysis.

(A) The tag counts for 3' seq peaks associated with the same gene ID were added together to generate total coverage for a gene. (B) Distribution of tag counts across the one thousand most abundant mRNA transcripts. Contribution of individual mRNAs to total tag count (green), or cumulative tag contribution (blue) is plotted against the ranked mRNA abundance.

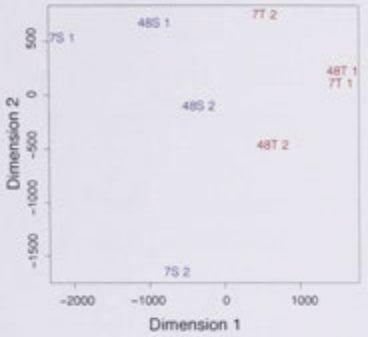
5.2.6.a Normalisation of mRNAs

The same 'trimmed mean of M-values' (TMM) normalisation method used for the miRNA sequencing was applied to the merged 3' seq data (See section 4.2.3 for further details). TMM normalisation and downstream use of the edgeR package has been applied to RNA sequencing data in a number of studies (*296-299*). Figure 5.12 shows multi-dimensional scaling plots (MDS) for the 8 cardiomyocyte samples showing CPM and the TMM normalised samples. As mentioned in Section 4.2.3, the distances in the MDS plots are a reflection of the similarities between samples. After normalisation, the Sham and TAC samples are separated more by Dimension 1 (Figure 5.11.A). The normalisation factors for the 3' seq peaks range between 0.9-1.1, indicating that the libraries are fairly similar in composition (Figure 4.5.A.iii). This is reflected by the Spearman correlation, with the lowest correlation value of 0.73 occurring between '7 Sham 2' and '48 TAC 1'. Hierarchical clustering groups three of the TAC samples together, while 7 TAC 1 is clustered with three of the Sham samples. Interestingly, the 7 Sham 2 sample is a separate cluster from the rest of the samples, indicating that the expression profile from this sample is dissimilar compared to the rest. The percentage of rod shaped cardiomyocytes, as well as the expression of the fibroblast and cardiomyocyte markers in the 7 Sham 2 sample were equivalent to the remaining samples used for sequencing, suggesting that the cardiomyocyte enrichment is not the reason for a difference in mRNA expression (Table 3.2).

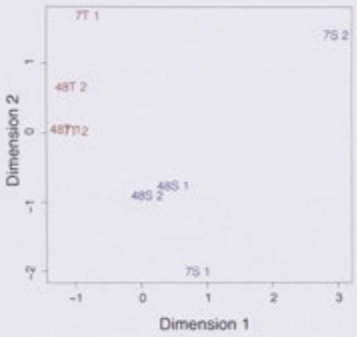
5.2.6.b Differential expression

To detect genes that are differentially expressed between the TAC banded animals and the controls, pairwise comparisons were made between biological replicates using the edgeR package (*182*). The data was filtered to retain the top 98th percentile of mRNAs, leaving approximately 6000 transcripts for analysis. Pairwise comparisons between 48 hour Sham vs TAC and 7 day Sham vs TAC were performed to determine differentially expressed mRNAs. In total, 776 mRNAs were differentially expressed with a FDR of less than 10% (Figure 5.13). The details for the normalisation and differential expression analysis are available in Appendix 8.11. The data was then divided into subsets based on direction of regulation and the time point, as shown in Figure 5.14. There were 469 mRNAs upregulated, with the majority (333 mRNAs) being regulated only at the 48 hour time point. 81 mRNAs were upregulated only at 7 days and a small proportion (56 mRNAs)

A i



ii



iii

Library	Library size	Normalisation Factor
48S 1	13976361	1.02
48S 2	18313928	0.97
48T 1	39968142	1.03
48T 2	22714198	1.00
7S 1	11135353	0.90
7S 2	15126969	1.07
7T 1	22030136	1.10
7T 2	16070001	0.92

B

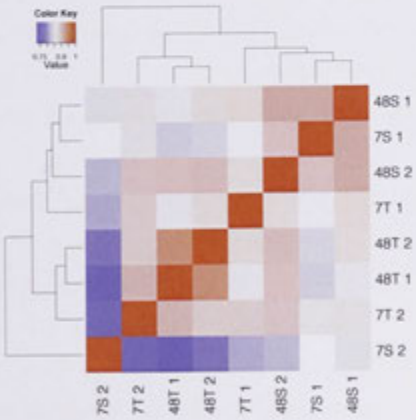


Figure 5.12 Normalisation of merged 3' seq peaks.

(A) MDS plot for merged 3' seq peaks expressed as CPM and (B) after normalisation using TMM. (C) The size of each library and TMM normalisation factor. (D) Heatmap of each sample based on the Spearman correlation.

were upregulated at both 48 hours and 7 days. In comparison, 303 mRNAs were downregulated, with 233 being regulated only at the 48 hour time point. A small number of mRNAs were downregulated at 7 days (37 mRNAs) or at both time points (37 mRNAs). To summarise, the majority of mRNA transcripts were regulated at 48 hours, with a larger proportion increasing in expression.

To start to discern the biological pathways that these mRNAs regulate, gene ontology (GO) enrichment was performed using the database for annotation, visualisation, and integrated discovery (DAVID)(*185*). GO-terms are grouped into three ontologies: biological process, cellular component and molecular function (*300*). A molecular function describes the activity of the gene at the molecular level, for example transporter activity. The biological process is a series of molecular functions. For instance, signal transduction is a broad biological process. The data was interrogated for enrichment of biological processes, molecular function, cellular component and KEGG pathways using the top 98% of mRNAs as the background. GO-terms with a FDR less than 30% were considered significant and are listed in Appendix 8.12. The enrichment within each GO-term category was related to the same pathway or function, so for ease of interpretation the Biological processes are presented. The mRNAs were first split into those upregulated or downregulated regardless of the timepoint (Table 5.3). The main GO-terms enriched in the upregulated transcripts are those involved in actin filament processes, translation and adhesion. Conversely, the majority of the GO-terms enriched in the downregulated transcripts are linked to fatty acid and lipid metabolism and oxidation reduction. Six genes related to the regulation of heart contraction are also downregulated.

The heatmap was then broken down into the subgroups illustrated in Figure 5.14 to establish the biological processes that may be specific to certain time points (Table 5.4). The mRNAs upregulated at 48 hours are involved in translation and the actin filament processes. The upregulated mRNAs at 48 hours are driving the enrichment of these processes when considering all of the upregulated transcripts. The 7 day upregulated transcripts are involved in cell adhesion but also introduce another term, proteolysis, which is the breakdown of proteins. Interestingly, when considering the mRNAs that are upregulated at both time points, numerous terms involved in angiogenesis and blood vessel development become significant.

The mRNAs downregulated at 48 hours are enriched for the same GO-terms when considering all of the downregulated transcripts, just with a slightly smaller number of genes contributing to each GO-term. For instance, fatty acid metabolic process and lipid catabolic process remain the GO-terms with the lowest FDR. In contrast, no biological processes had a FDR less than 30% for the 7 day downregulated transcripts. There were only 37 transcripts that were downregulated at both time points, and a small number of these were involved in the positive regulation of transcription. For both the up and downregulated transcripts, the 48 hour time point is driving the enriched GO-terms.

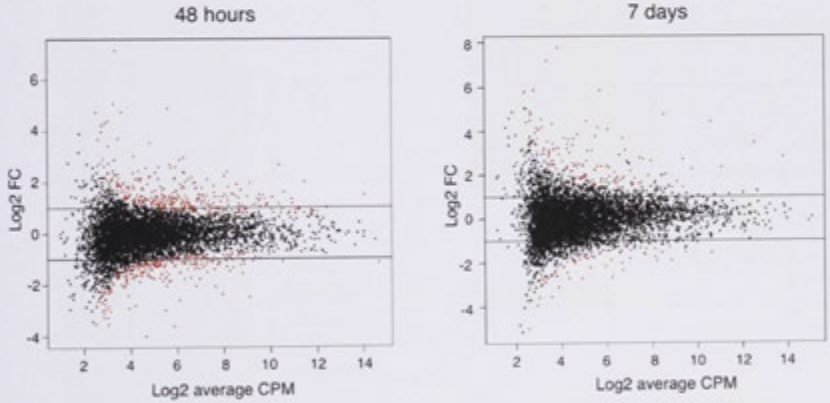


Figure 5.13. Identification of differentially expressed mRNA transcripts.

Log-2 fold change (FC) is plotted against the Log-2 average counts per million (CPM) at 48 hours and 7 days. The lines indicate a Log2 FC of 1 and the mRNAs identified as differentially expressed with a FDR <10% are shown in red.

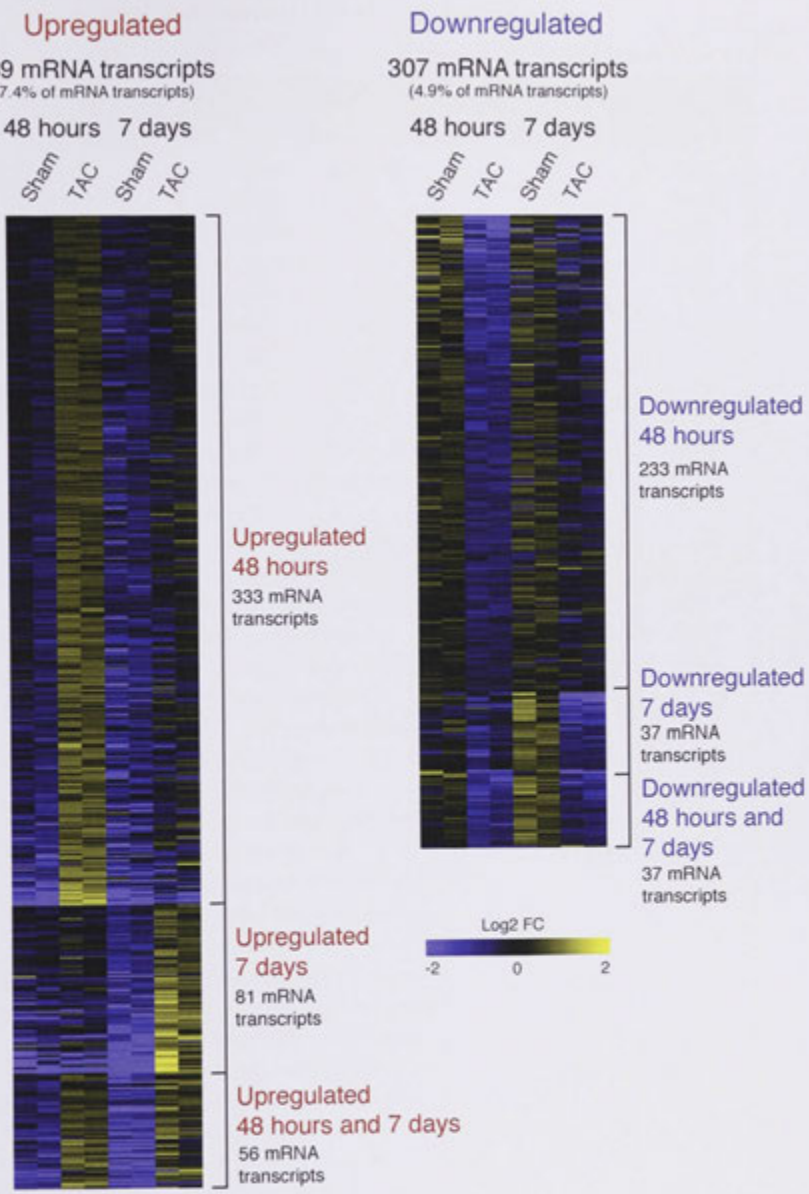


Figure 5.14. Differentially expressed mRNA transcripts with a FDR <10%.

Pairwise comparisons were made with the edgeR package, data is shown as log2-fold change for each cardiomyocyte sample.

Table 5.3. GO-TERM enrichment of up- and down-regulated genes.

Biological processed with FDR <30%, comparing all upregulated and downregulated transcripts.

TermID	TermDescription	Count	FDR
Upregulated			
GO:0007015	actin filament organisation	11	0.018
GO:0006412	translation	36	0.039
GO:0030036	actin cytoskeleton organisation	18	0.057
GO:0030029	actin filament-based process	19	0.060
GO:0007010	cytoskeleton organisation	21	2.881
GO:0030198	extracellular matrix organisation	8	11.426
GO:0042254	ribosome biogenesis	12	21.919
GO:0022610	biological adhesion	20	22.848
GO:0007155	cell adhesion	20	22.848
GO:0051017	actin filament bundle formation	4	24.164
GO:0010324	membrane invagination	10	27.719
GO:0006897	endocytosis	10	27.719
GO:0051258	protein polymerisation	5	28.788
Downregulated			
GO:0006631	fatty acid metabolic process	26	4.01E-10
GO:0016042	lipid catabolic process	13	0.001
GO:0044242	cellular lipid catabolic process	9	0.017
GO:0055114	oxidation reduction	34	0.044
GO:0009062	fatty acid catabolic process	6	0.407
GO:0034440	lipid oxidation	6	0.407
GO:0019395	fatty acid oxidation	6	0.407
GO:0006635	fatty acid beta-oxidation	5	0.782
GO:0042493	response to drug	8	0.863
GO:0006084	acetyl-CoA metabolic process	7	2.170
GO:0006071	glycerol metabolic process	5	3.203
GO:0030258	lipid modification	6	4.300
GO:0006637	acyl-CoA metabolic process	4	4.352
GO:0006732	coenzyme metabolic process	13	4.550
GO:0019400	alditol metabolic process	5	4.574
GO:0016054	organic acid catabolic process	8	4.866
GO:0046395	carboxylic acid catabolic process	8	4.866
GO:0015672	monovalent inorganic cation transport	11	9.788
GO:0006813	potassium ion transport	7	9.811
GO:0051186	cofactor metabolic process	14	12.478
GO:0008016	regulation of heart contraction	6	19.423
GO:0048878	chemical homeostasis	13	22.059
GO:0006085	acetyl-CoA biosynthetic process	3	23.268
GO:0009719	response to endogenous stimulus	9	28.310
GO:0006869	lipid transport	7	29.205

Table 5.4. GO-Term enrichment for up- and down-regulated transcripts at 48 hours or 7 days. FDR <30%.

Biological processes with a FDR <30%, comparing the subgroups illustrated in Figure 5.14.

TermID	TermDescription	Count	FDR
48 hour upregulated			
GO:0006412	translation	35	2.42E-05
GO:0007015	actin filament organisation	9	0.069
GO:0030029	actin filament-based process	14	0.646
GO:0030036	actin cytoskeleton organisation	13	0.914
GO:0006897	endocytosis	10	3.227
GO:0010324	membrane invagination	10	3.227
GO:0034470	ncRNA processing	12	7.442
GO:0042254	ribosome biogenesis	10	16.524
GO:0006364	rRNA processing	8	17.786
GO:0016072	rRNA metabolic process	8	19.599
GO:0022613	ribonucleoprotein complex biogenesis	11	21.079
GO:0007010	cytoskeleton organisation	14	29.386
7 day upregulated			
GO:0006508	proteolysis	12	9.125
GO:0022610	biological adhesion	7	14.514
GO:0007155	cell adhesion	7	14.514
48 hour and 7 day upregulated			
GO:0060284	regulation of cell development	5	0.506
GO:0001525	angiogenesis	5	3.620
GO:0050767	regulation of neurogenesis	4	4.496
GO:0051960	regulation of nervous system development	4	4.891
GO:0048514	blood vessel morphogenesis	5	10.951
GO:0045596	negative regulation of cell differentiation	4	17.604
GO:0001568	blood vessel development	5	19.790
GO:0001944	vasculature development	5	20.360
GO:0031589	cell-substrate adhesion	3	25.013
48 hour downregulated			
GO:0006631	fatty acid metabolic process	22	5.15E-09
GO:0016042	lipid catabolic process	12	5.80E-04
GO:0044242	cellular lipid catabolic process	8	0.025
GO:0006635	fatty acid beta-oxidation	5	0.253
GO:0055114	oxidation reduction	26	0.358
GO:0006084	acetyl-CoA metabolic process	7	0.462
GO:0019395	fatty acid oxidation	5	1.563
GO:0034440	lipid oxidation	5	1.563
GO:0009062	fatty acid catabolic process	5	1.563
GO:0006732	coenzyme metabolic process	11	4.756
GO:0046395	carboxylic acid catabolic process	7	5.055

Table 5.4 cont. GO-Term enrichment for up- and down-regulated transcripts at 48 hours or 7 days. FDR <30%.

TermID	TermDescription	Count	FDR
48 hour downregulated cont.			
GO:0016054	organic acid catabolic process	7	5.055
GO:0006869	lipid transport	7	8.418
GO:0051186	cofactor metabolic process	12	9.385
GO:0030258	lipid modification	5	9.440
GO:0006071	glycerol metabolic process	4	12.651
GO:0006085	acetyl-CoA biosynthetic process	3	13.831
GO:0010876	lipid localisation	7	15.855
GO:0009081	branched chain family amino acid metabolic process	4	16.078
GO:0019400	alditol metabolic process	4	16.078
GO:0009719	response to endogenous stimulus	8	19.842
GO:0043648	dicarboxylic acid metabolic process	4	19.872
GO:0006633	fatty acid biosynthetic process	5	26.607
GO:0019216	regulation of lipid metabolic process	5	29.707
GO:0006637	acyl-CoA metabolic process	3	29.889
GO:0006573	valine metabolic process	3	29.889
GO:0001676	long-chain fatty acid metabolic process	3	29.889
48 hour and 7 day downregulated			
GO:0045893	positive regulation of transcription, DNA-dependent	5	15.380
GO:0051254	positive regulation of RNA metabolic process	5	16.503
GO:0055114	oxidation reduction	7	17.039
GO:0042493	response to drug	3	19.144
GO:0006631	fatty acid metabolic process	4	22.188
GO:0045941	positive regulation of transcription	5	22.269
GO:0010628	positive regulation of gene expression	5	22.712
GO:0045935	positive regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	5	28.306
GO:0010557	positive regulation of macromolecule biosynthetic process	5	28.793

5.2.6.c Comparison to published data

To identify novel expression changes, the differentially expressed mRNAs were compared to other studies that have documented gene expression changes during hypertrophy. The first study performed TAC with a 28-G needle, resulting in a tighter constriction, and took whole LV samples at 1, 3 and 12 weeks post-surgery (*168*). Another performed TAC with a 27-G needle and isolated whole LV samples after 48 hours, 1, 2, 3 and 8 weeks (*301*). Similarly, a study by Zhao et al. performed TAC with a 27-G needle but isolated whole hearts after 48 hours, 10 days and 3 weeks (*187*). The final study induced hypertrophy with Isoprotenol or Angiotensin II via a mini-pump and isolated whole LV after 7 or 14 days respectively (*302*). All of these studies used microarrays for expression analysis. 215 of the differentially expressed mRNAs identified here were also reported in one or more of these studies (Figure 5.15.A). 189 have been shown to change expression in the same direction, while 28 were regulated in the opposite direction when compared to the published data (Figure 5.15.B). One of these studies also analysed microarray data for heart development and intriguingly, we found that 136 of the mRNAs regulated in our data were altered in the opposite direction during development (Figure 5.15.B), providing further evidence that the heart is returning to a foetal like gene program (*168*).

The mRNAs that had been previously identified to change expression were then removed from the heatmap and the GO-term enrichment was repeated to see if the depth of our sequencing uncovered any novel biological processes (Table 5.5). The full list of enriched GO-terms for the novel mRNAs is listed in Appendix 8.12. For both the upregulated and downregulated lists, the GO-terms remain essentially the same but the number of genes contributing to the biological process changes, which then alters the FDR. For example, 36 mRNAs are involved in translation when considering all upregulated mRNAs but this number is reduced to 31 when looking at only the novel mRNAs. The main processes are still enriched; with actin cytoskeletal processes enriched in the upregulated genes and fatty acid or lipid metabolism enriched in the downregulated transcripts. Similarly, the majority of the terms associated with either the 48 hour upregulated or downregulated mRNA transcripts remain the same but with a smaller number of genes contributing to each term (Table 5.6). After removing the previously identified transcripts, there was no enrichment of biological processes in the mRNAs upregulated at 7 days or those upregulated at 48 hours and 7 days. This suggests that the processes of proteolysis, cell adhesion and angiogenesis have been documented in the previous studies. On the other hand, the

number of genes associated with transcription remains unchanged when comparing all mRNAs downregulated at 48 hours and 7 days and those that have not been reported before, indicating that regulation of transcription is specific to the novel mRNAs.

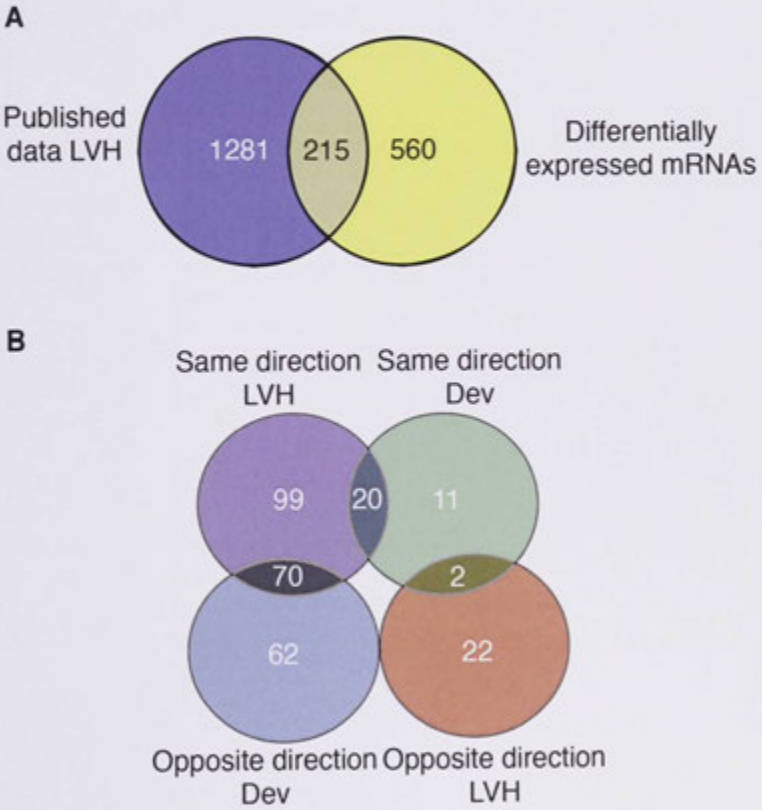


Figure 5.15. Comparison between differentially expressed mRNAs and published data.
(A) Overlap between differentially expressed mRNAs with a FDR <10% and mRNAs previously reported to be differentially expressed during LVH (*168, 187, 301, 302*). (B) Overlap between mRNAs reported to be differentially regulated in LVH and development (Dev) (*168*). Majority of mRNAs are regulated in the same direction during LVH or in the opposite direction during development.

Table 5.5. GO-TERM enrichment for novel mRNAs.

Biological processes with FDR <30%, comparing all upregulated and downregulated transcripts.

TermID	TermDescription	Count	FDR
Upregulated			
GO:0006412	translation	31	0.002
GO:0030036	actin cytoskeleton organisation	13	0.700
GO:0042254	ribosome biogenesis	12	1.282
GO:0030029	actin filament-based process	13	1.716
GO:0022613	ribonucleoprotein complex biogenesis	13	2.105
GO:0034470	ncRNA processing	12	5.978
GO:0007015	actin filament organisation	6	14.773
GO:0006364	rRNA processing	8	15.391
GO:0016072	rRNA metabolic process	8	17.003
Downregulated			
GO:0006631	fatty acid metabolic process	19	5.06E-06
GO:0016042	lipid catabolic process	12	7.79E-04
GO:0044242	cellular lipid catabolic process	9	0.002
GO:0009062	fatty acid catabolic process	6	0.116
GO:0019395	fatty acid oxidation	6	0.116
GO:0034440	lipid oxidation	6	0.116
GO:0006635	fatty acid beta-oxidation	5	0.282
GO:0016054	organic acid catabolic process	8	1.070
GO:0046395	carboxylic acid catabolic process	8	1.070
GO:0042493	response to drug	7	1.310
GO:0030258	lipid modification	6	1.338
GO:0055114	oxidation reduction	23	6.632
GO:0009791	post-embryonic development	6	14.057
GO:0043648	dicarboxylic acid metabolic process	4	21.245
GO:0006355	regulation of transcription, DNA-dependent	23	26.289

Table 5.6. GO-Term enrichment for up- and down-regulated novel transcripts at 48 hours or 7 days.

Biological processes with a FDR <30%, comparing the novel mRNAs within the subgroups illustrated in Figure 5.14.

TermID	TermDescription	Count	FDR
48 hour upregulated			
GO:0006412	translation	30	5.41E-05
GO:0034470	ncRNA processing	12	1.140
GO:0042254	ribosome biogenesis	10	3.705
GO:0022613	ribonucleoprotein complex biogenesis	11	4.414
GO:0006364	rRNA processing	8	5.212
GO:0016072	rRNA metabolic process	8	5.837
GO:0030036	actin cytoskeleton organisation	10	6.348
GO:0007015	actin filament organisation	6	6.367
GO:0030029	actin filament-based process	10	11.773
GO:0034660	ncRNA metabolic process	11	24.026
48 hour downregulated			
GO:0006631	fatty acid metabolic process	16	4.05E-05
GO:0016042	lipid catabolic process	11	5.82E-04
GO:0044242	cellular lipid catabolic process	8	0.005
GO:0006635	fatty acid beta-oxidation	5	0.099
GO:0009062	fatty acid catabolic process	5	0.630
GO:0019395	fatty acid oxidation	5	0.630
GO:0034440	lipid oxidation	5	0.630
GO:0046395	carboxylic acid catabolic process	7	1.509
GO:0016054	organic acid catabolic process	7	1.509
GO:0030258	lipid modification	5	4.079
GO:0043648	dicarboxylic acid metabolic process	4	10.743
GO:0055114	oxidation reduction	18	15.699
GO:0009791	post-embryonic development	5	24.651
48 hour and 7 day downregulated			
GO:0045893	positive regulation of transcription, DNA-dependent	5	3.914
GO:0051254	positive regulation of RNA metabolic process	5	4.240
GO:0045941	positive regulation of transcription	5	5.997
GO:0010628	positive regulation of gene expression	5	6.138
GO:0045935	positive regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	5	7.995
GO:0010557	positive regulation of macromolecule biosynthetic process	5	8.164
GO:0009891	positive regulation of biosynthetic process	5	9.041
GO:0051173	positive regulation of nitrogen compound metabolic process	5	9.041
GO:0031328	positive regulation of cellular biosynthetic process	5	9.041

Table 5.6 cont. GO-Term enrichment for up- and down-regulated novel transcripts at 48 hours or 7 days.

TermID	TermDescription	Count	FDR
48 hour and 7 day downregulated cont.			
GO:0006357	regulation of transcription from RNA polymerase II promoter	5	11.789
GO:0045944	positive regulation of transcription from RNA polymerase II promoter	4	15.948
GO:0010604	positive regulation of macromolecule metabolic process	5	17.228
GO:0006355	regulation of transcription, DNA-dependent	6	20.838
GO:0051252	regulation of RNA metabolic process	6	23.126

5.2.6.D Expression of hypertrophic markers

The hypertrophic markers ANP, BNP, α-Ska and β-MHC were all upregulated and detected in the mRNAs with a FDR less than 10% (Figure 5.14, Figure 5.16). The expression of these four markers was ascertained prior to sequencing using RT-qPCR on another lot of LV samples (Figure 3.7). Furthermore, the expression of α-Ska and ANP was confirmed to be upregulated in the purified cardiomyocyte samples, including those subjected to sequencing (Figure 3.10). This provided further validation that merging the peaks associated with the same gene could be used as a proxy of gene expression.

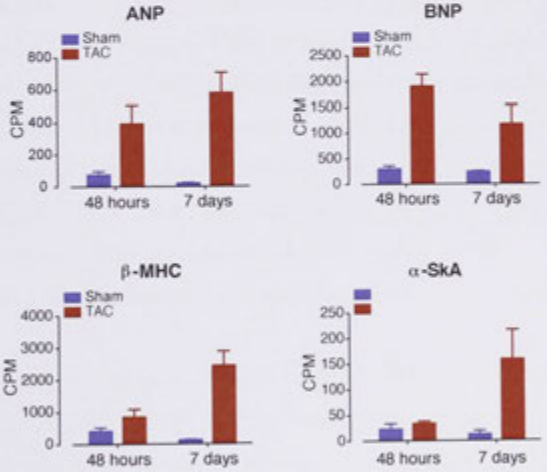


Figure 5.16. Cardiac hypertrophy markers are upregulated according to the 3' seq data.
The normalised CPM for the hypertrophy markers ANP, BNP, α-Ska and β-MHC. Data is presented as mean with SEM (n=2).

5.2.6.E Validation of merged 3' ends with RT-qPCR

The expression of nine mRNAs with varying CPM, fold change and FDR values from Figure 5.12 were tested using RT-qPCR. The expression was tested in the samples subjected to next-generation sequencing as well as in the additional cardiomyocyte samples. Of the nine mRNAs considered for testing, the expression changes of six could be validated in both the samples used for sequencing and the extended dataset (Figure 5.14). The mRNA expression changes were considered validated if a two-way ANOVA with Bonferonni's post-hoc analysis on the RT-qPCR data of the extended dataset resulted in an adjusted p value of less than 0.05.

The downregulation of Nduf1 and Acsl1 could be observed in the RT-qPCR data for the sequencing samples and the extended dataset, but this did not reach significance (Figure 5.17.B). In fact, the fold change for Acsl1 using the sequencing samples at 48 hours had a fold change greater than or equal to the NGS data. But as there are only two replicates for this analysis, the result was not significant. The extended dataset had a greater variability in the expression of Acsl1, leading to a higher standard error and adjusted p value greater than 0.05. Ndufa1 had a lower absolute fold change value in comparison to the other genes that could be validated. On the other hand, Calm1 did not show any upregulation in the sequencing samples or the extended dataset compared to the NGS data (Figure 5.17.C). Alike to Ndufa1, the log2-FC for Calm1 according to NGS was less than 1. The absolute log2-FC of the six mRNAs that could be validated was greater than 1, suggesting that this could be the factor that determines if the expression changes can be validated.

In summary, the expression levels of eight mRNAs followed the same trend as the NGS data when analysed using RT-qPCR. Furthermore, the expression of six mRNAs was statistically significant in the extended cardiomyocyte sample set.

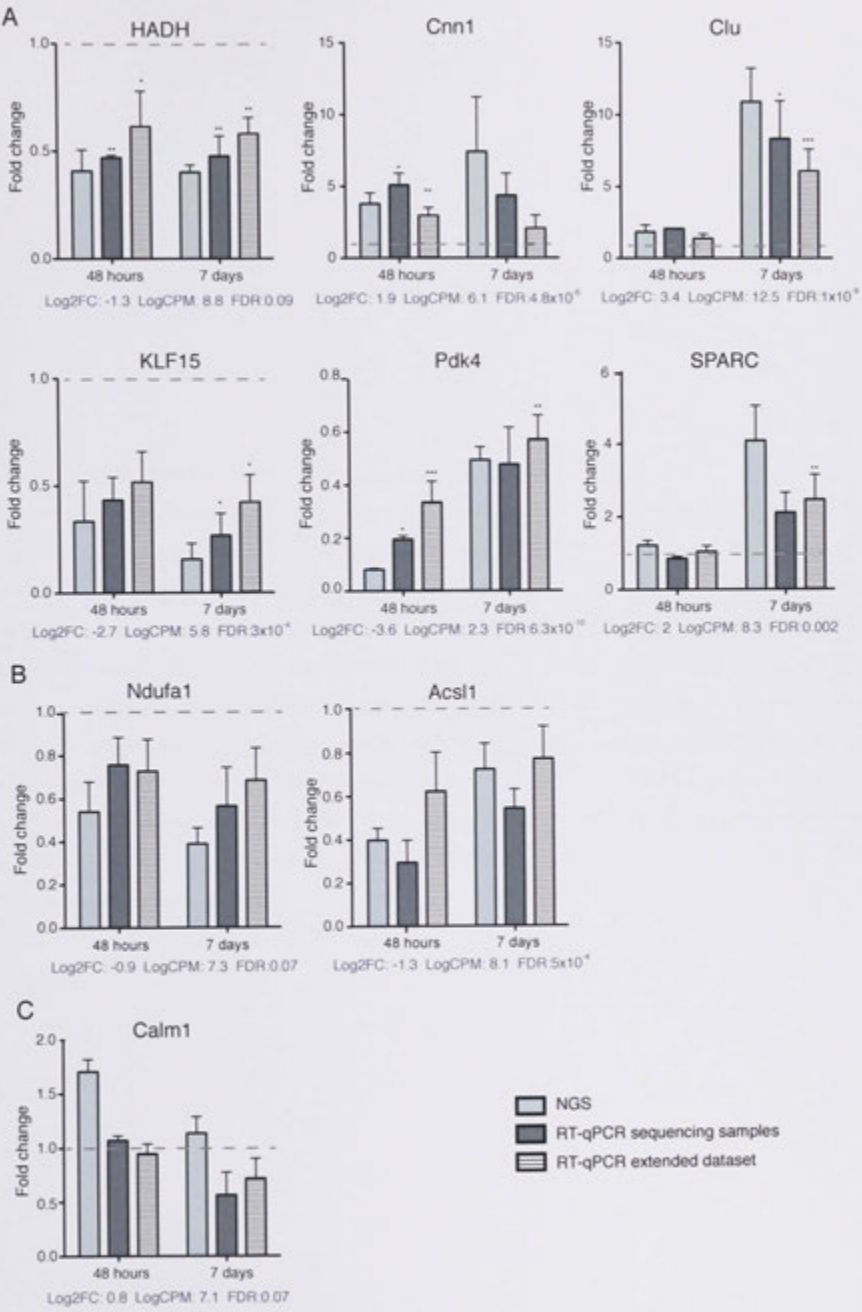

A
HADH
Fold change
48 hours
7 days
Log2FC: -1.3 LogCPM: 8.8 FDR:0.09
Cnn1
Log2FC: 1.9 LogCPM: 6.1 FDR:4.8x10⁻⁶
Clu
Log2FC: 3.4 LogCPM: 12.5 FDR:1x10⁻⁹
KLF15
Log2FC: -2.7 LogCPM: 5.8 FDR:3x10⁻⁴
Pdk4
Log2FC: -3.6 LogCPM: 2.3 FDR:6.3x10⁻¹⁰
SPARC
Log2FC: 2 LogCPM: 8.3 FDR:0.002
B
Ndufa1
Log2FC: -0.9 LogCPM: 7.3 FDR:0.07
Acsl1
Log2FC: -1.3 LogCPM: 8.1 FDR:5x10⁻⁴
C
Calm1
Log2FC: 0.8 LogCPM: 7.1 FDR:0.07
NGS
RT-qPCR sequencing samples
RT-qPCR extended dataset

Figure 5.17. Validation of differential gene expression by RT-qPCR.

Fold change (TAC/Sham) values for the next-generation sequencing data (NGS), RT-qPCR on the eight samples subjected to next-generation sequencing (RT-qPCR sequencing samples) and RT-qPCR on the extended cardiomyocyte sample set (RT-qPCR extended dataset). (A) mRNA expression changes that could be validated in both the samples used for sequencing and the extended dataset. (B) The same trend could be observed in the q-PCR for both the sequencing samples and the extended dataset but it was not significant. (C) The trend could not be observed in either the sequencing samples or the extended dataset. The dashed grey line indicates a fold change of 1. Data shown as mean log FC with SEM (n=4-5 for RT-qPCR sequencing sample and RT-qPCR extended dataset, n=2 for NGS * $p < 0.05$, Two way-ANOVA with Bonferonni's post-hoc analysis). The log2FC, log counts per million (CPM) and false discovery rate (FDR) for the NGS data are shown below each graph.

5.2.7 Detection of alternative 3' UTR usage

The next step was to determine if there was any evidence for changes in 3' UTR usage within the data. Dr Brian Parker from ANU conducted the analysis into the mRNA length changes and the weighted linear model for alternative polyadenylation. To detect changes in mRNA length overall, the mean 3' UTR length was calculated for the Sham and TAC samples. The distances were weighted according to the number of counts, with highly expressed mRNAs being more heavily weighted. The mean 3' UTR length was shorter in the TAC samples at both time points, with a greater difference at 48 hours. The weighted mean was reduced from 450 to 430 nt in the TAC samples at 48 hours, and from 482 to 448 nt at 7 days. However, while a trend could be observed, both of these analyses had a p value greater than 0.05 (0.13 at 48 hours and 0.22 at 7 days).

Alternative polyadenylation was then detected at the gene level by looking for a change in the proportion of 3' seq peaks for those genes with multiple 3' ends. The peaks associated with protein coding genes were used, and only peaks that contributed more than 10% to the tags associated with a gene were considered. Alternative polyadenylation was detected as a proportion statistic between proximal and distal sites using a weighted linear model (Refer to section 2.2.7.E for more details). Similar to the gene expression data, more candidates of alternative polyadenylation were detected at 48 hours compared to 7 days. However, the FDR for both time points was very high, therefore we ranked the mRNAs according to FDR and looked at mRNAs with an FDR less than 70%. There were no examples at 7 days with an FDR less than 70%, so we looked only at the changes occurring at 48 hours. We filtered this further by looking only at mRNAs that had an average tag count greater than 100, and had a change in proximal:distal ratio greater than 15% between the Sham and TAC samples (Figure 5.18). This resulted in 424 mRNAs with a change in 3' UTRs at 48 hours; 152 where the 3' UTR was longer in the TAC samples, and 272 mRNAs where the 3' UTR was shortened in the TAC samples. Therefore, more genes had a shortening of the 3' UTR, which is consistent with the analysis of genome wide averages.

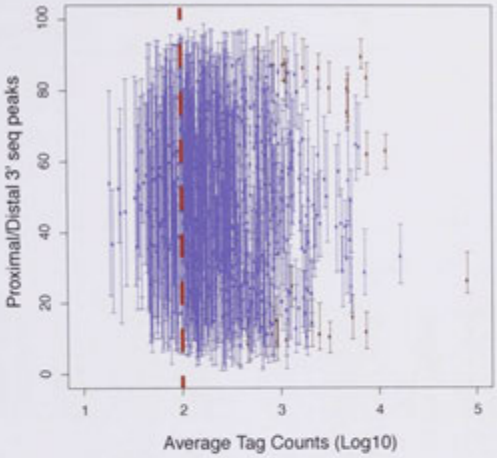
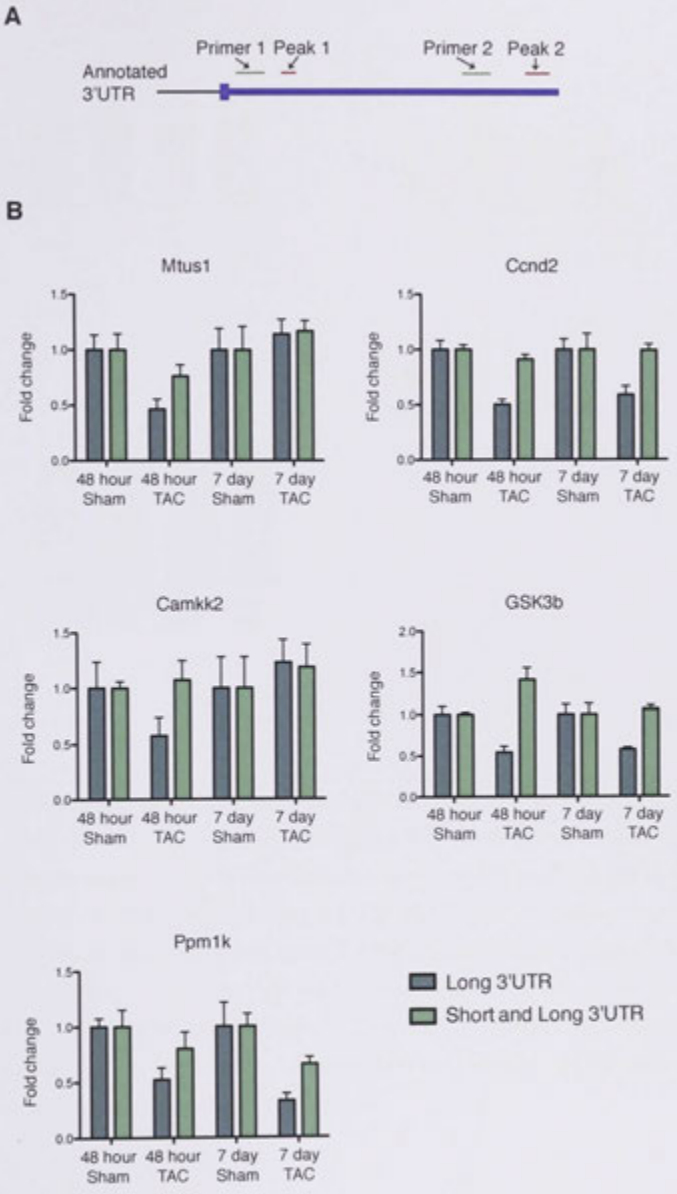


Figure 5.18. mRNAs with FDR <70% at 48 hour.

The mean and range of Proximal/distal ratio for mRNAs with a FDR less than 70% at 48 hours. mRNAs with a range less than 15% are shown in red, those with a range greater than 15% are indicated in blue. The mRNAs with a log10 average tag count greater than 2 (to the right of the red dashed line) and with a range greater than 15% were used for downstream analysis.

5.2.7.A Validation of 3' UTR changes

10 mRNAs were chosen for validation of 3' UTR changes, five linked to cardiac hypertrophy (Mtus1, Ccnd2, Camkk2, GSK3b and Ppm1k) and five chosen with varying FDRs and proportion changes (Atg3, Usf2, Etf1, Utp6 and Htatip2) (*303-307*). Rina Soetanto from the ANU performed the RT-qPCRs for 3' UTR changes. RT-qPCR was performed using primers downstream or upstream of the proximal 3' seq peak in the 3' UTR (Figure 5.19.A). The primers for Atg3 and Usf2 could not be optimised and were excluded from downstream analysis. If both 3' UTRs were regulated to the same extent, then the fold change values for both primer pairs would be equivalent. If the 3' UTR was getting shorter, then the fold change for the primers amplifying the Long 3' UTR would be lower. If a greater proportion of 3' UTRs was longer, then the fold change for the primers amplifying the Long 3' UTR would be higher. The decrease in the use of the long or distal 3' UTR could be confirmed with the extended samples at 48 hours for the five mRNAs tested for shortening 3' UTRs (Figure 5.19.B). In some cases (Ccnd2, Gsk3b and Ppm1k), this decrease in the long 3' UTR could also be seen at 7 days.


A
Primer 1
Peak 1
Primer 2
Peak 2
Annotated
3'UTR
B
Mtus1
Ccnd2
Camkk2
GSK3b
Ppm1k
Fold change
48 hour Sham
48 hour TAC
7 day Sham
7 day TAC
Long 3'UTR
Short and Long 3'UTR

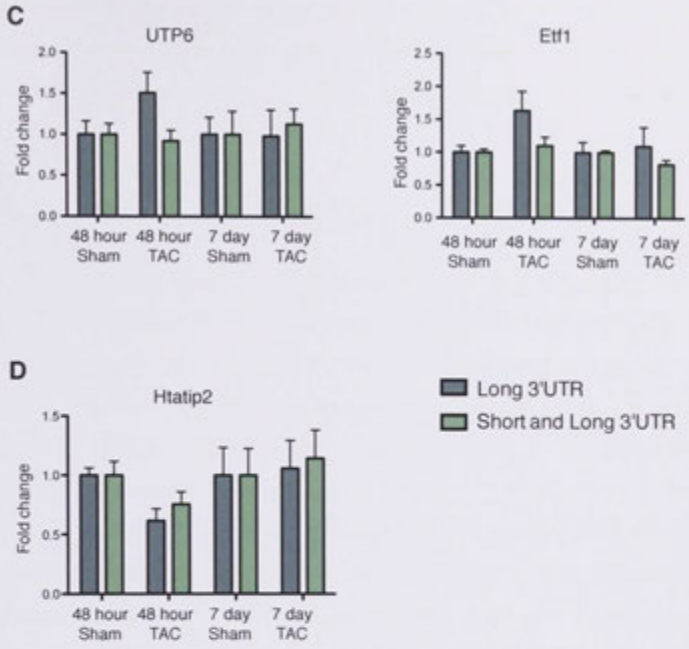


Figure 5.19. The change in 3' UTRs could be validated with conventional RT-qPCR.

(A) Schematic of a 3' UTR showing the position of two 3' sequencing peaks (red) and the primers used for validation (green). Primer 1 was designed before the proximal 3' seq peak and amplifies both variants of the 3' UTR (Short and Long 3' UTR) while Primer 2 is designed before the distal peak and amplifies only the long 3' UTR (Long 3' UTR). (B) mRNAs with shorter 3' UTRs in the 48 hour TAC samples. (C) mRNAs confirmed to have longer 3' UTRs in the 48 hour TAC samples. (D) Htatpip2 had more of the long 3' UTR in the TAC samples according to NGS, this could not be confirmed with RT-qPCR. RT-qPCR with the extended cardiomyocyte sample set, data is shown as mean fold change with SEM (n=2-5).

The increase in the use of the long or distal 3' UTR was validated in the extended samples at 48 hours for two of the three mRNAs tested (Figure 5.19.C). For the third mRNA, Htatip2, the increase in the long 3' UTR could not be confirmed using RT-qPCR on the extended dataset.

5.2.7.B Correlation with gene expression

To begin to untangle the possible outcome of alternative polyadenylation, the mRNAs detected to have alterations in the proportions of 3' UTRs were compared to the mRNAs differentially expressed in Figure 5.14. It could be hypothesised that a decrease in the length of the 3' UTR would increase the amount of the transcript overall, as there would be less regulatory elements present in the 3' UTR sequence. Only 29 mRNAs overlapped with the mRNAs differentially expressed with an FDR <10% (Figure 5.20). Of these, there was no obvious correlation between the direction of change in the 3' UTR and the expression overall. For instance, for mRNAs with a shortened 3' UTR, there are candidates where the overall gene expression increased or decreased. This analysis was extended to the 424 genes that change 3' UTRs, but again no correlation could be observed between the 3' UTR change and gene expression levels (data not shown).

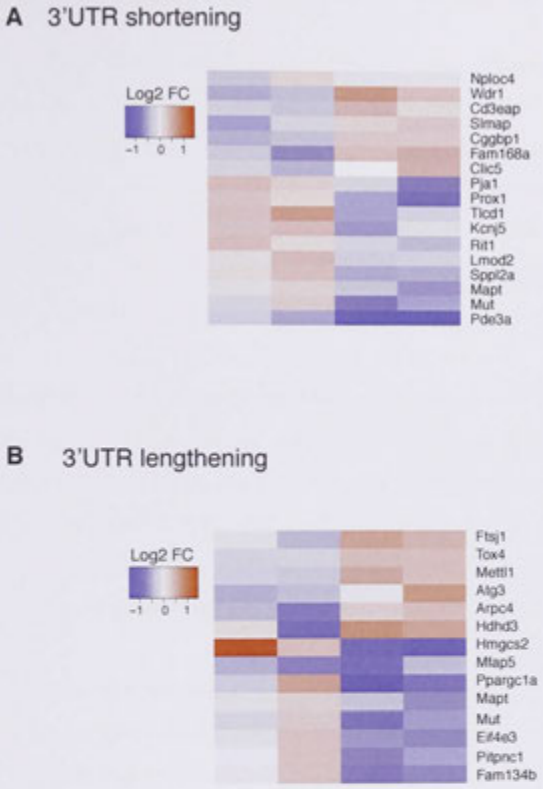


Figure 5.20. 3' UTR changes do not correlate with gene expression changes.

(A) Log2-fold change (FC) for mRNAs that shorten their 3' UTRs in the TAC sample. (B) Log2-fold change for mRNAs that lengthen their 3' UTRs in the TAC samples.

5.2.7.C Cardiac hypertrophy associated genes with APA

We next asked if any of the genes associated with cardiac hypertrophy altered their 3' UTRs in response to the TAC banding. 31 of the 652 genes from the "cardiac hypertrophy" genecard list had a change in their 3' UTR with a FDR less than 70% (Table 5.6). This list included genes involved in cell cycle regulation, transcription, energy and metabolism, as well as factors directly involved in the hypertrophic signalling pathways (*308-312*). In summary, the functions of the cardiac hypertrophy related genes with changing 3' UTRs are diverse.

These mRNAs were also investigated for their correlation with overall gene expression. Three genes, Ppargc1a, Ppp3ca and Cacna1c, had more of the distal 3' UTR and a decrease in their absolute expression levels (log2-FC less than -1). While only one gene, Rras2, had an increase in the proximal 3' UTR and a log2-FC greater than 1. This is consistent with the previous section, where there was no correlation overall between the 3' UTR changes and the log2 FC for the gene in entirety (Table 5.7).

Some of the mRNAs with three or four 3' UTRs were documented as having a change in ratio between the proximal and distal 3' seq peaks. In this case, one of the pairwise comparisons between two of the 3' seq peaks was significant. For example, Peak 1 and Peak 4 of GSK3b showed a change, with less of Peak 4 present in the TAC samples (Figure 5.10.E). For such examples, the change is still biologically meaningful as one of the 3' seq peaks is changing in comparison to the others.

The data provided in Table 5.7 is particularly important when considering the miRNA interactions with these hypertrophy related mRNAs. For instance, miR-495-3p is predicted to target GSK3b but this miRNA binding site is not present in the TAC samples once the proximal poly(A) site is used. Information such as this is critical when considering miRNAs or siRNAs against specific genes for therapeutic interventions.

Table 5.7. 3' UTR changes for cardiac hypertrophy associated genes.

Name	Number of 3' UTRs	3' UTR change	Overall Log2-FC
Cd36	2	Lengthen	-0.79
Polr2l	2	Lengthen	0.72
Ppargc1a	2	Lengthen	-1.85
Tfam	2	Lengthen	-0.05
Pten	2	Lengthen	-0.61
Gnb5	2	Lengthen	0.41
Jak1	3	Lengthen	0.35
Ppp3ca	4	Lengthen	-1.21
Cacna1c	2	Lengthen	-1.72
Prkaa2	2	Lengthen	-0.73
Pln	2	Shorten	-0.22
Acadl	2	Shorten	-0.99
Ak1	2	Shorten	0.05
Prkar1a	2	Shorten	0.07
Ubc	2	Shorten	-0.55
Ccnd2	3	Shorten	0.11
Akt2	2	Shorten	0.21
Csnk1a1	3	Shorten	-0.17
Jun	3	Shorten	-1.30
Slc8a1	2	Shorten	0.33
Rac1	2	Shorten	-0.16
Rras2	2	Shorten	1.24
Hfe	2	Shorten	-0.58
Kcnj5	2	Shorten	-1.02
Cpt1a	3	Shorten	-0.68
Atf6	3	Shorten	-0.46
Pdgfa	2	Shorten	-0.62
Mef2a	2	Shorten	-0.87
Smad4	2	Shorten	-0.44
Ablim3	2	Shorten	0.46
Gsk3b	4	Shorten	0.68

Overall Log2-FC refers to the fold change (TAC/Sham) for the mRNA transcript.

5.2.7.D GO-TERM ENRICHMENT OF MRNAS WITH 3' UTR CHANGES

To more systematically examine the genes undergoing 3' UTR changes, GO-term enrichment was performed to functionally annotate the genes with shortened or lengthened 3' UTRs in the 48 hour TAC samples. However, the mRNAs with a longer 3' UTR in the 48 hour TAC samples did not show any enrichment for a biological process with an FDR less than 30%. The shortened mRNAs were enriched for some biological processes, but the enriched GO-terms were very broad (data not shown). For example, cellular protein catabolic process was the most highly enriched GO-term. This result is consistent with a publication that examined UTR changes in TAC banded mice using microarray probes. No GO-terms were significantly associated with the mRNAs with shortened 3' UTRs (*168*). This suggests that the 3' UTR changes are a broad phenomenon in hypertrophic cardiomyocytes and are not confined to a functional subclass of mRNAs.

5.3 DISCUSSION

In this chapter, we performed genome-wide profiling of mRNA 3' UTRs in control, pre-hypertrophic and hypertrophic cardiomyocytes. The data was annotated for 3' UTRs and interrogated for features of polyadenylation, then used to generate a static picture of all 3' UTRs, with a focus on mRNAs involved in cardiac hypertrophy. The 3' sequencing data was successfully used as a measure of overall gene expression, and then analysed for the presence of alternative polyadenylation during hypertrophy.

5.3.1 FEATURES OF THE 3' SEQUENCING LIBRARY

5.3.1.A GENERAL FEATURES OF THE LIBRARY

As expected, the majority of the tags mapped directly to the mouse genome, with only a small number mapping to the transcriptome. The small proportion of tags that map to the transcriptome span exon:exon boundaries and likely represent short 3' UTRs where the 3' seq peak begins prior to the the stop codon. 20% of peaks contained evidence of a poly(A) tail towards the 3' end of the tag. This variation was expected, due to variable poly(A) tail length, the library preparation and size selection of the cDNA which ranged from 70-300 nt. As the sequencing tag was 75 nt in length, the sequencing will not always read far enough into the cloned product to detect the poly(A) tail. While this does not matter for defining the 3' UTRs of a given mRNA transcript, it does mean that the exact poly(A) cleavage site cannot always be determined. As next-generation sequencing improves and

the length of sequencing reads become longer, we will be able to read into the poly(A) tail and define the exact cleavage sites for more 3' ends.

5.3.1.B Assigning to genomic features

Peaks were assigned to a genomic feature if they overlapped the 3' UTR region for mRNAs or they were less than 20kb downstream of a stop codon or transcript end. Other studies have used similar criteria to define 3' ends but using 5 or 8 kb to assign 3' seq peaks to upstream genes (*171, 313*). While we allowed 20 kb, the mean distance from the stop codon of mRNA transcripts was 3.7 kb. In further support of the 20 kb range, a recent study discovered that putative lincRNAs were actually extended 3' UTRs of upstream genes and they validated 3' UTRs up to 19 kb in length using Northern blotting (*314*). In fact, the length of the 3' UTRs for six genes (Ppp1r7, Dnajc15, Tlmm17a, Hmbox1, Pdk3 and Ppargc1b) found in our dataset were also found in Miura et al.s data and were validated using Northern blotting (*314*). Therefore 20 kb is a suitable cut off to define 3' UTRs while retaining the long, potentially novel 3' UTRs.

Just over 25% of the peaks were annotated as Orphan peaks, meaning they did not overlap with any known genomic feature and were more than 20kb away from the closest upstream gene. This proportion is consistent with another study that found 22.5% of detected 3' end sites discovered using NGS in Zebrafish were unannotated (*171*). Despite accounting for one quarter of the peaks, the Orphan peaks contribute only a small proportion of tags (7.6% of total mapped tags), suggesting that they are expressed at low levels. The RNA sequencing by the Encyclopedia of DNA Elements (ENCODE) project revealed that 83.7% of the human genome undergoes pervasive transcription (*315*). It has been demonstrated that transcription of unannotated regions such as where the Orphan peaks map, results in the production of a large number of lincRNAs (*316*). The low but widespread coverage of our Orphan peaks is consistent with pervasive transcription and it is plausible that these Orphan peaks are derived from as yet unannotated lincRNAs.

Peaks that were derived from Exons upstream of the stop codon or Intron regions accounted for 22.4% and 10.9% of the tags respectively. For simplicity, we used the canonical transcript defined in Ensembl for annotation. Therefore, it is possible that these Exon and Intron peaks are still *bona fide* 3' ends and represent Intronic APA or alternative splicing. Alternative splicing can generate an mRNA transcript variant with a different last

exon, thus altering the location of the 3' UTR on the genome. Intronic APA results from cleavage at a PAS within the intron, extending the previous exon and making it the terminal one (*317*). While we did not investigate such peaks in this study, they can be interrogated in the future for changes in expression levels that may relate to APA of shorter mRNA transcript variants.

Of the peaks that were downstream of a stop codon or transcript end, the majority were derived from protein coding genes as expected. However, other RNA species such as lincRNAs, snoRNAs and miRNAs were also detected at low levels. lincRNAs are processed in a similar manner to mRNAs, including the addition of a poly(A) tail (*295*). Similarly, the primary miRNA precursor is capped and polyadenylated (*294*). In further support of our data, polyadenylation sites for miRNAs, lincRNAs and snoRNAs are documented in the expression and polyadenylation database (xPAD), which provides a map of polyadenylation sites in human cancer tissues and cell lines (*318*). In addition, another PolyA-seq study captured polyadenylated non-coding miRNAs in their data, including primary miRNA precursors and antisense RNAs (*289*). The polyadenylation of non-coding RNA species was not the focus of this study, but nonetheless our data will also provide a resource for the expression and polyadenylation of non-coding RNAs.

5.3.1.c Features of polyadenylation

After the peaks were assigned to genomic features, we searched for features of polyadenylation. As mentioned above, a proportion of tags and therefore a proportion of peaks contained evidence of a poly(A) tail at the 3' end. This proportion increased when considering only protein coding associated peaks. All transcripts containing a poly(A) tail should have been cloned into the sequencing library but the poly(A) tail may have been further downstream of the sequencing region. It's plausible that as the protein coding transcripts are the most abundant, contributing the most tags in the sequencing library, then there is a higher chance of detecting a poly(A) stretch within the 3' seq peak.

We also searched for the presence of the canonical PAS or one of its variants. The canonical PAS was found in 43% of protein coding associated peaks, with this number increasing to 60% when considering protein coding associated peaks that contain a poly(A) tail. This is consistent with previous studies that found between 53-58% of PAS hexamers in human are the canonical AAUAAA (*319*). The proportion of peaks where no PAS was

detected increases from 10% in protein coding peaks with a poly(A) stretch to 35% when considering protein coding peaks without any indication of a poly(A) tail. As mentioned previously, this could be attributed to the region in which the search for the PAS is conducted. For those peaks without the poly(A) tail, the cleavage site and thus the PAS may be much further downstream making it hard to determine the exact PAS. Approximately 20% of human poly(A) cleavage sites do not have an upstream PAS (*317*). These sites rely on strong binding of the CFIm 3' processing factor to the upstream UGUA element (*154*), or on Cstf to the DSE with other auxiliary elements (*320*). Therefore, the 3' seq peaks that do not contain a PAS could rely on other sequence elements for processing.

This analysis was then extended by looking at the location of the PAS. In agreement with the known literature, the canonical PAS was found 17.5 nt upstream of the cleavage site. The second most abundant PAS was AUUAAA, which has been reported previously in the mouse as the second most abundant PAS (*160*). Curiously, the most abundant non-canonical PAS is not consistent between peaks that do or do not contain a poly(A) stretch. AUUAAA remains the most abundant non-canonical PAS in poly(A) containing peaks and is present at 10 to 25 nt upstream of the cleavage site. However, AAGAAA is more abundant in the peaks without poly(A). This PAS variant has been reported to contribute to only a small proportion of PAS sites in the mouse (*160*), and it is centred around the cleavage site in our data. The algorithm to detect the PAS searches for the canonical PAS first, followed by the non-canonical PAS variants and reports the distance from the end of the 3' sequencing peak. As the actual cleavage site may be much further downstream, it is possible that the canonical PAS or a more common PAS variant exists further downstream. The third most abundant PAS was UUUAAA, which contains two peaks, one at the expected location between -10 and -25, and another centred around the 3' seq peak end. It is known that the region around the cleavage site and PAS is U-rich (*160*), therefore we may be detecting a U-rich region rather than an actual PAS. Conclusions about the use of PAS are more accurate for the peaks that contain the start of the poly(A) tail, and of these we observe the most common PAS.

5.3.2 APA IN PROTEIN CODING GENES

5.3.2.A NUMBER OF 3' UTRS FOR PROTEIN CODING GENES

We found that on average each protein coding gene had 2.35 3' UTRs. Similar studies in Humans and Zebrafish have identified that on average an mRNA has 2.5 and 2.8 3' UTRs respectively (*171, 315*). 34% of our 3' seq peaks associated with protein coding genes overlapped with a known 3' end and accounted for 69% of the tags, which is consistent with other studies that suggest that most 3' ends are novel but the majority of the tags are derived from known 3' ends (*171, 315*). Due to the depth of NGS, we are detecting the novel 3' ends that are expressed at lower levels. In some cases, the main 3' UTR detected was not at the exact cleavage site as defined in Ensembl. Changing the length of the 3' UTR can increase or decrease the area available for regulation by RNA binding proteins (RBPs) and miRNAs.

5.3.2.B CARDIAC MRNAS WITH MULTIPLE 3' ENDS

We detected several mRNAs implicated in cardiac hypertrophy that have multiple 3' UTRs. Having a static picture about the 3' UTRs available, regardless if they change between the Sham and TAC samples is important for understanding the regulation of the mRNA transcript. In some cases the mRNA had only one 3' UTR, but this 3' UTR was not the annotated 3' end in Ensembl. A detailed map of the 3' UTRs available in the heart can help more accurately predict miRNA binding sites.

5.3.2.C IMPLICATIONS OF MULTIPLE 3' UTRS

The 3' UTR can determine the stability, cellular localisation and translation efficiency of the transcript. The length of a 3' UTR has been linked to nonsense mediated decay (NMD) in a sequence-independent manner (*321*). The protein Upf1 regulator of nonsense transcripts homolog (Upf1) binds along the length of a 3' UTR, and as more Upf1 molecules can bind a longer 3' UTR the probability of Upf1 binding factors that regulate NMD increases (*321*). Regulatory sequence elements, such as miRNA and RBP binding sites, are also present within a longer 3' UTR. For instance, the Au-binding factor 1 (AUF1) protein binds to AU-rich regions in the 3' UTR and promotes the degradation of mRNA targets by recruiting the exosome (*322*). The interaction between miRNAs and the 3' UTR variants is discussed in Chapter Seven.

The localisation of mRNAs is mostly determined by *cis*-regulatory sequences within the 3' UTR. For example, the long 3' UTR of brain-derived neurotrophic factor (BDNF) targets the mRNA to the dendrites in neuronal cells, resulting in site-specific translation (*323*). Similarly, the calmodulin-dependent protein kinase II (CaMKIIα) is targeted to dendrites by a sequence just after the proximal poly(A) site (*324*). The 3' UTR variants of the cardiac mRNAs may therefore result in different subcellular localisations, which could have downstream implications for translation.

The 3' UTR choice can also influence the translation or protein output. Accurate conclusions about the translational efficiency of mRNA 3' UTRs can be made by analysing the levels of protein (*325*). It is known that shorter mRNAs can produce up to 40 times more protein than a longer variant, as they escape regulation by miRNAs and other RBPs (*164*). For example, the Rat ATPase, Na+/K+ transporting, beta 1 polypeptide (ATP1B1) mRNA uses three predominant PAS (*326*). Use of the proximal PAS results in a short 3' UTR that is translationally more efficient than the longer 3' UTR, as the longest 3' UTR contains a unique translational repressor sequence (*327*). There are also examples of long 3' UTRs producing more protein than the short 3' UTR variant or both producing comparable amounts (*323, 328, 329*). The long 3' UTR variant of the amyloid precursor protein (APP) has higher translation efficiency and produces more protein than the short 3' UTR (*323*). There are proteins that specifically bind to the short but not the long APP 3' UTR, reducing its translational efficiency (*323*). Similarly, the longer 3' UTR of the brain-derived neurotrophic factor (BDNF) produces more protein in neuronal cells (*329*). On the other hand, the first 24 nt of the lipoprotein lipase (LPL) 3' UTR contains a motif for translational inhibition and upstream or downstream elements do not affect translation (*330*). So while it is generally thought that a shorter 3' UTR produces more protein, it is clear that there are individual examples where the opposite is true. The 3' UTRs documented in this thesis can provide valuable information about sequence regulatory elements controlling a mRNA transcript and their potential impact on translational efficiency, localisation and stability.

5.3.3 3' SEQUENCING AS A PROXY FOR GENE EXPRESSION

5.3.3.A DIFFERENTIAL EXPRESSION

While the main aim of this chapter was to document the 3' UTRs of cardiac mRNAs, we took advantage of the available data to look at gene expression. The 3' seq peaks associated

with a protein coding gene (downstream of the stop codon) were merged together to generate tag counts for differential expression. Other studies have shown that NGS of 3' ends can accurately be used as a proxy of gene expression, with the method highly correlated with normal RNA sequencing, microarrays and RT-qPCR (*289, 331, 332*). In fact, 3' sequencing removes any length bias that standard RNA sequencing encounters as it does not span the entire transcript (*332*). It also has the advantage of identifying low to moderately expressed genes (*332*). This is reflected in the dynamic range of the data, with the top 1000 genes contributing 87% of the tags.

Out of the 6000 mRNAs that were used for assessing differential expression, approximately 13% were differentially expressed at an FDR of less than 10%. The majority of the mRNAs were upregulated and more mRNAs were differentially regulated at 48 hours. There is a uniform increase in total RNA synthesis that accounts for the 30-50% increase in cell volume during hypertrophy (*333*). This is achieved by an increase in the activity of RNA pol I (*334*), pol II and pol III (*335*). In addition, there is a more dramatic increase in the expression of specific genes such as the cardiac hypertrophy markers (*336*). This could account in part for the larger number of upregulated mRNA transcripts compared to downregulated transcripts in the TAC samples.

5.3.3.B GO-TERM ENRICHMENT

GO-term enrichment of the differentially expressed mRNAs reveals processes that are well documented to change or regulate cardiac hypertrophy. The GO-terms for all of the upregulated transcripts are mainly involved in actin or cytoskeletal organisation, translation and cell adhesion. Compensated hypertrophy results in the accumulation of cytoskeletal proteins that counteract the pressure on the heart (*337*). The cytoskeleton in cardiomyocytes consists of the sarcomeric skeleton, the true cytoskeletal proteins, membrane-associated proteins and the proteins of the intercalated disc (*337*). It is the thin and thick filaments of the sarcomere that form the contractile units of a cardiomyocyte (*338*). The upregulated genes included alpha actin cardiac muscle 1 (Actc1), an alpha actin that forms part of the sarcomere (*339*). Other genes include alpha-Actinin and Filamin, which link the sarcomere and cytoskeleton to the extracellular matrix (ECM). Alpha-actinin is an intermediary molecule that transmits signals from Integrin receptors on the cell-surface to the actin filaments (*340*). Similarly, Filamin acts as a scaffold between the cytoskeleton and transmembrane receptors (*341*). Filamins help maintain the membrane

integrity, which is particularly important during increased mechanical stress that would occur during pressure overload. The enrichment of biological and cellular adhesion is also related to the ECM and cytoskeleton of the cardiomyocyte. Cellular adhesion is defined as the attachment of a cell to another cell or to a substrate such as the ECM via adhesion molecules. The focal adhesion complex is a cluster of Integrin transmembrane receptors that connect the ECM to the cell via actin binding proteins (*342*). These actin-binding proteins include Talin, alpha-Actinin and Filamin. In our data, the expression of both Talin1 and Integrin-alpha 5 increase after TAC. There is also an increase in proteins that are involved in the ECM, such as connective tissue growth factor (CTGF) that stimulates ECM synthesis (*343*).

The biological process translation is also enriched in the upregulated genes. It is known that there is an increase in protein synthesis in response to hypertrophy. Initially, this is due to an increase in the efficiency of translation within hours of the hypertrophic stimulus (*344*). However, this is not sufficient to account for the drastic increase in protein that accompanies cardiac hypertrophy and as a result, the amount of ribosomal protein also increases after 12-24 hours (*345*). In agreement, we saw a number of transcripts encoding ribosomal proteins as well as eukaryote initiation and elongation factors increase expression at the 48 hour time point. This increases the capacity for protein synthesis.

The downregulated transcripts were enriched for terms associated with fatty acid metabolism and oxidation-reduction. Energy is generated in the normal heart by the oxidation of fatty acids (*346*). In agreement with the re-expression of the foetal isoforms of MHC and skeletal actin, the metabolism of the heart reverts back to a foetal metabolic profile with an increase in carbohydrate utilisation during cardiac hypertrophy (*347*). Confirming our GO-term enrichment, it is well documented that Fatty acid metabolism is drastically decreased in the hypertrophic heart (*348-351*). Furthermore, it is known that this decrease is achieved in part by the downregulation of genes that encode fatty acid oxidation enzymes (*352, 353*). While we observed this decrease in genes involved with fatty acid metabolism, we did not see an upregulation of genes involved in glycolysis. However, an increase in the expression or capacity of glycolytic enzymes does not always correlate with increased glycolysis (*354*). Consistent with our data, other studies have found no significant change in the expression of glucose transporters or enzymes involved in the glycolytic pathway (*354-356*). Subsequently, it is thought that an increase in glucose uptake is the reason for an increase in glycolysis (*357*).

To better understand the molecular mechanisms of hypertrophy occurring at the different stages of disease, the GO-term enrichment was performed on subsets of the data. The most significant GO-terms enriched for the mRNAs upregulated at 48 hours remain translation and actin filament organisation, suggesting that this is an early response to the pressure overload. Another two terms, endocytosis and ncRNA processing are also enriched. The endocytosis machinery is activated in response to stimulation of β-adrenergic receptors (β-AR), a type of GPCR (*358*). The internalisation of β_1-AR is required for the activation of Akt, which then results in the expression of transcription factors that regulate the hypertrophic response (see Section 1.1.2.E for more details on the Akt pathway). It is possible then that the upregulation of genes involved in endocytosis mediate the internalisation of β-ARs for downstream signalling. As mentioned above, the other process enriched at 48 hours is ncRNA processing. The majority of the genes contributing to this enrichment are involved in rRNA or tRNA processing, which is consistent with the upregulation of translation. Interestingly, Dicer1 was also upregulated at 48 hours. The majority of miRNAs we observed to change expression levels occurred at 7 days. The miRNA processing pathway also involves several steps so the questions remains if upregulating one of the miRNA processing enzymes would drastically increase the levels of functional miRNAs if the remaining proteins do not change. In summary, changes in gene expression relating to translation and actin filament organisation occur before the heart has increased in size.

The hypertrophic cardiomyocytes at 7 days upregulated genes implicated in proteolysis and biological adhesion. As mentioned earlier, cardiac hypertrophy results in an increase in protein synthesis or translation, which was detected at the mRNA level at 48 hours. In the normal cell, up to 50% of proteins are degraded by the proteasome even prior to peptide elongation (*359*). Due to the stress on the cardiomyocyte, the increase in translation that occurs at 48 hours also results in an increase in denatured proteins that need to be degraded to avoid apoptosis (*360*). This suggests that between 48 hours and 7 days components of the protein degradation pathway are upregulated to compensate for the increase in translation. mRNAs involved in biological adhesion are also upregulated at 7 days. As discussed above, biological adhesion is related to the actin cytoskeletal network and the ECM. It is known that ECM deposition or fibrosis occurs during hypertrophy and this indicates that structure changes to the ECM occur on or before 7 days.

At both the pre- and hypertrophic stages of cardiac hypertrophy, angiogenesis and blood vessel formation factors are upregulated. During compensatory hypertrophy, the number of myocardial capillaries increases. However, after prolonged hypertrophy, the relative capillary density decreases as the cardiomyocytes increase in size, resulting in larger diffusion distances for oxygen and nutrients (*361*). This impairment in angiogenesis is one of the reasons the heart progresses from compensated to maladaptive hypertrophy (*362*). We observed enrichment of angiogenesis factors at 7 days. Two of these angiogenesis factors, CTGF and heparin-binding epidermal growth factor (HB-EGF) are known to be upregulated and secreted from cardiomyocytes in response to hypertrophy, suggesting that the enrichment of mRNAs associated with angiogenesis is not derived from other cell types (*363, 364*). The increase in these factors at 7 days is consistent with the notion that there is a larger density of capillaries when the heart is compensating against the pressure overload.

Genes that were downregulated at both 48 hours and 7 days had less significant enrichment, with the lowest FDR for genes that positively regulate transcription at 15.4%. These genes encode transcription factors or transcription co-activators including KLF15, Jun and Pparg1ca. KLF15 inhibits gene expression and the hypertrophic response (*365, 366*). It does this by repressing the transcription factors Mef2 and GATA4, and the transcriptional co-activator myocardin that stimulate hypertrophic growth (*365*). Downregulation of KLF15 therefore removes this repression and results in the re-expression of hypertrophic genes including ANF and α-SkA as observed in our data (*366*). Similarly, Jun (or c-jun) is a transcription factor that regulates the cytoskeleton and counteracts the increased fibrosis that occurs after pressure overload (*367*). The enrichment of transcription does not refer to an overall decrease in the levels of transcription, but rather refers to downregulation of transcription factors that mediate gene expression involved in cardiac hypertrophy.

5.3.3.c Comparison to other studies

The differentially regulated mRNAs were compared to other studies that have documented genome wide expression changes in hypertrophy. 24% of the mRNAs we detected as differentially expressed have been reported in other studies to change in the same direction during hypertrophy. The studies we made comparisons with have all used microarrays as a measure of gene expression. It is possible that there were not probes on the microarrays

for all of the genes we have detected. It is also plausible that these mRNAs changed within the data, but did not meet the statistical cut-offs used in other studies. Furthermore, these studies used different end points and/or different models of cardiac hypertrophy. For instance, van den Bosch et al. described gene expression changes at multiple time points after TAC banding, the latest of which was 8 weeks post-surgery (*301*). At this late stage in hypertrophy when it is progressing to decompensation, a different gene expression profile could be expected. Similarly, the study by Friddle et al. used Isoproterenol or Angiotensin-II to induce hypertrophy and thus activated different signalling pathways compared to pressure overload induced hypertrophy (*302*). In fact, there are no mRNAs that are differentially regulated in all of the four studies used as a comparison, highlighting the differences between these studies (data not shown). Despite the difference in the study design, 17.5% of the mRNAs were also regulated in the opposite direction in a heart development time course (*168*). This provides further confirmation that the heart is returning to a foetal like gene program.

After redoing the GO-term enrichment only for the mRNAs unique to our data, it was evident that the main biological processes remained enriched. A large number of genes still contribute to translation, actin cytoskeleton organisation and fatty acid metabolism. These processes are critical for the progression of compensated to decompensated hypertrophy. So while the same pathways have remained enriched, we may have identified novel mRNAs that could serve as therapeutic targets for heart disease.

While the expected biological pathways were identified, the genes that contribute to this enrichment represent a minor proportion of the mRNAs that were differentially expressed. This leads to the question, what are the remaining mRNAs doing? Are they contributing to the same biological pathways but they are not currently annotated sufficiently to be classified with a GO-term? The gene expression data presented in this chapter can be used as a resource for mRNAs that are altered during cardiac hypertrophy and derived primarily from cardiomyocytes.

5.3.3.D Limitations of GO-term enrichment analysis

There are also limitations to GO-term enrichment as an analysis of global gene expression changes. For instance, GO-term enrichment doesn't take into account the expression levels or changes associated with each gene and each gene is treated equally. The information

about the degree of change could be used to assign weighting to each gene, which could provide more information about the pathways that are changing (*368*). Furthermore, some biological processes have been studied in more depth than others, and subsequently have more annotated genes associated with that term. Therefore, these well-annotated pathways are more likely to be significantly enriched. Some annotations are also incomplete and not all genes are functionally annotated. This could explain why pathways already known to change with cardiac hypertrophy are significantly enriched but novel pathways are not.

5.3.4 Changing 3' UTRs

The second main aim of this chapter was to determine if mRNAs changed their 3' UTR usage in cardiac hypertrophy. We detected a slight reduction in the global 3' UTR length in the TAC samples and have documented 424 mRNAs with 3' UTR changes at 48 hours. This was based on pairwise comparisons with an FDR less than 70%. While this FDR was high, the changing 3' UTRs of seven out of eight mRNAs tested with FDRs ranging from 35-64% could be confirmed with RT-qPCR (Figure 5.19). A similar study conducted in Zebrafish did not apply a statistical model but filtered their data for genes that changed more than 30% in pairwise comparisons (*171*). Other publications have used more stringent criteria. For instance, a study looking at 3' UTR changes during mouse development calculated the use of the distal poly(A) site relative to the proximal site and then used a microarray analysis package to detect genes with an FDR less than 10% (*161*). As there is not consistency between the reporting or analysis of 3' sequencing data, we thought it was reasonable to use the mRNAs with an FDR less than 70%, then filter further for expression level and change in 3' UTR usage.

5.3.4.a Improving the 3' UTR statistics

There are a number of mRNAs with multiple 3' seq peaks with two or more peaks in close proximity. For example, GSK3b has three peaks at the proximal end of the 3' UTR and one peak at the distal end of the 3' UTR (Figure 5.10.E). The pairwise comparisons to detect changes in 3' UTR ratios were made between individual 3' seq peaks. However, merging peaks that are within close proximity to each other may identify changes in 3' UTR length that were not identified when considering the peaks individually. This would remain biologically meaningful if *cis*-regulatory elements were present between the short and long 3' UTRs.

5.3.4.B What controls APA?

The choice of poly(A) cleavage site and thus the 3' UTR length is dictated by the concentration of 3' processing factors, the interaction of trans-acting factors and transcription. The distal poly(A) site is usually stronger than the upstream site and thus is preferentially used by the 3' processing machinery (*160*). Upregulation of core processing factors can result in higher usage of a weaker proximal poly(A) site (*369, 370*). The was first illustrated for CstF-64 and the IgM mRNA, where high levels of CstF-64 during B cell maturation result in use of a proximal poly(A) cleavage site within an intron (*369*). The merged differential expression data was examined for the core components of the 3' processing machinery. However, there was no change in expression for any of the components examined (data not shown). It is still possible that the protein levels increase due to translational control. Nevertheless, this indicates that at least on the mRNA level there is no correlation between the global 3' UTR length and 3' processing factors during cardiac hypertrophy. This is not surprising as there is only a mild reduction in the mean 3' UTR length and thus it could be expected that there is no bulk change in 3' processing factors.

The choice of poly(A) cleavage site can also be influenced by trans-acting proteins. For instance, the poly(A)-binding protein nuclear 1 (PABN1) generates longer 3' UTRs by associating with the proximal poly(A) cleavage sites and preventing their cleavage (*292*). Conversely, the Cytoplasmic polyadenylation element binding protein 1 (CPEB1) increases the use of the weaker proximal poly(A) cleavage sites by binding and recruiting the 3' processing factors (*371*). The activity of such proteins and the control of alternative polyadenylation in the heart remains to be investigated.

Polyadenylation is closely coupled to transcription, with 3' processing factors associating with RNA Pol II and transcription factors (*372*). It is postulated that these interactions increase the efficiency of cleavage and polyadenylation and therefore increase the use of proximal poly(A) cleavage sites (*373*). Increases in transcriptional activity leads to increased cleavage at the proximal poly(A) cleavage sites (*374*). As transcription is known to increase during cardiac hypertrophy (*333*), this could account for the slightly shorter global 3' UTR length in the TAC samples. Another hypothesis for the link between transcription and 3' end processing is kinetic coupling. A decrease in the elongation rate of RNA Pol II results

in an increased use of the proximal poly(A) cleavage site, presumably because it takes a longer period of time to reach the distal poly(A) cleavage site and thus allowing more time for the proximal site to be cleavage (*375*). Both of these mechanisms would result in a net change in 3' UTR usage, with an increase in transcriptional activity resulting in a preference for the proximal 3' UTR.

5.3.4.C Correlation with gene expression

We could not observe a correlation between changes in 3' UTR length and overall transcript levels. Other studies have not observed a negative correlation between 3' UTR length and gene expression (*171, 287, 376*). This suggests that for most genes alternative polyadenylation does not greatly impact on transcript stability. As mentioned in section 5.3.2.C, the length of the 3' UTR can impact on translational efficiency and change the binding sites available for miRNAs and other RBPs. In eukaryotes, miRNAs can repress their targets by translational repression (*65*) . Therefore, while we would see a difference at the protein level for the mRNA 3' UTRs with different miRNA binding sites, we may not observe a difference at the RNA level.

5.3.4.D Cardiac hypertrophy related genes with changing 3' UTRs

A small proportion (4.7%) of cardiac hypertrophy related genes had different proportions of 3' UTRs between the Sham and TAC samples at 48 hours. The function of these genes is diverse, suggesting that 3' UTR changes are not specific to a cellular pathway. In addition, the signalling pathways that are activated in response to cardiac hypertrophy are regulated at the post-translational level (See section 1.1.2 for more details on the signalling pathways). For example, the MAPK pathway is a series of phosphorylation events which results in the phosphorylation and activation of transcription factors (*35*). It is no surprise then that components of this pathway are not regulated by alternative polyadenylation (data not shown).

5.3.5 Conclusions

In this chapter, we have described the 3' UTRs expressed in murine cardiomyocytes, and documented changes in 3' UTRs during the induction of cardiac hypertrophy. Each mRNA in the heart has on average 2.35 3' UTRs, which increases the area for regulatory sequences

and interactions. This data provides an important resource for the 3' UTRs used in the heart, particularly when most novel 3' UTRs are tissue specific (*289*). It can be used for miRNA target predictions, as well as to identify novel *cis-* or *trans-* regulatory elements. A number of mRNAs changed the proportion of their 3' UTRs in response to hypertrophy and this can be used to identify miRNA and RBP interactions with disease stage specific 3' UTRs.

CHAPTER SIX

ISOMIR FUNCTION

6.1 Introduction

The small RNA sequencing described in Chapter Four illustrated that the cardiac specific miR-133a exists as multiple start (5') and end (3') isomiRs derived from both the 5p and 3p arms (Figure 4.15.A). Furthermore, the major 5' isomiRs of the miR-133a-3p are associated with multiple 3' ends. In both human and mouse, miR-133a has two genomic locations and is transcribed as a bicistronic transcript with miR-1 (Figure 6.1.A). Both of the loci encoding the identical miR-133a-1 and miR-133a-2 in the mouse are activated by MEF2- and SRF-dependent enhancers (*48, 112*). The MEF2 transcription factor is activated by Calcineurin-NFAT signaling during hypertrophy and orchestrates the expression of stress-responsive genes, while SRF is a transcription factor that regulates pathological cardiac remodeling leading to heart failure (*377*) (See section 1.1.2.A for more details on MEF2). Several of the targets for miR-133a-3p have also been identified, including SRF that is repressed in a negative feedback loop (shown in Figure 6.1.B) (*378*). The targets identified thus far implicate miR-133a-3p in numerous pathways critical to the heart, including proliferation, cardiac conduction, fibrosis and cardiac hypertrophy signaling (*140, 378-381*). Nevertheless, these studies have not taken into account the multiple processing variants that can be generated from the pre-miR-133a hairpin.

The critical role of miR-133a in the heart has been highlighted by functional studies where the level of miR-133a has been altered. For instance, an antagomiR against miR-133a-3p that was continuously delivered into the murine heart resulted in an increase in hypertrophy markers such as wall thickness and the re-expression of the cardiac fetal genes (*140*). The same study also showed that overexpression of the entire miR-133a hairpin using an adenovirus reduced the size of cardiomyocytes and decreased the expression of hypertrophic markers in a transgenic mouse overexpressing Akt (*140*). While studies such as this suggest that the miR-133a locus is antihypertrophic, the methods used to knockdown miR-133a-3p or overexpress the miRNA hairpin actually manipulate the levels of both the strands of the miR-133a hairpin and their associated isomiRs. It is therefore critical to establish if these isomiRs have different targeting properties and function.

The aim of this chapter was to determine if the isomiRs of miR-133a-3p have different targeting specificities. This was addressed using luciferase vectors containing putative miR-133a-3p binding sites within the 3'UTR.

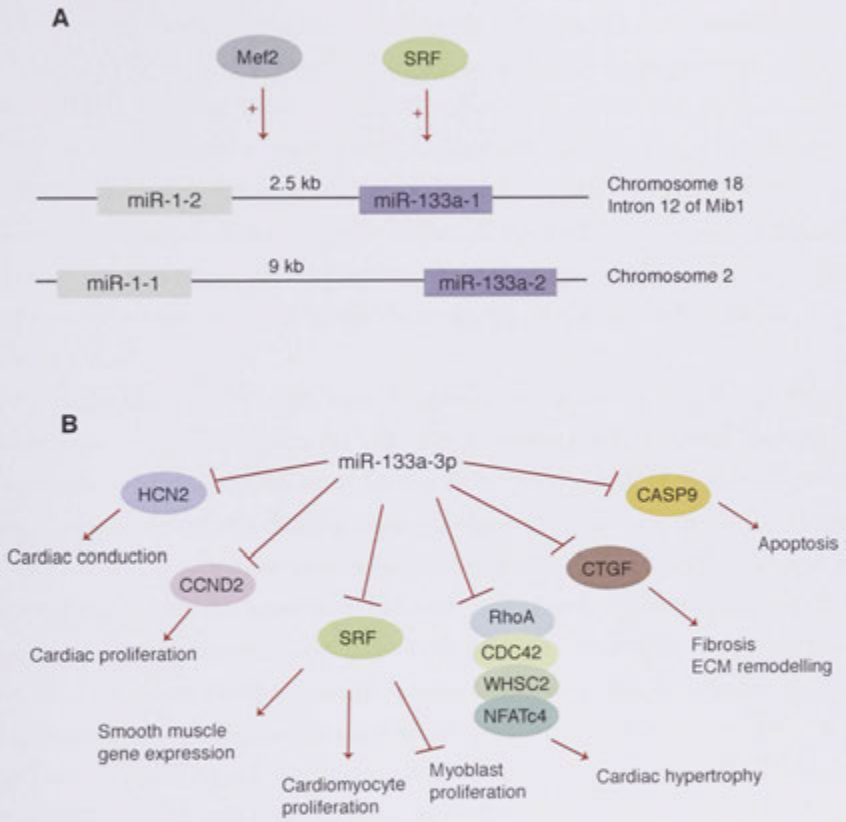


Figure 6.1. miR-133a-3p is a central regulator of the heart.

(A) miR-133a is encoded as two bicistronic elements with miR-1. In the mouse, miR-133a-1 is encoded with miR-1-2 in an Intron of Mib1, while miR-133a-2 is encoded in an intergenic region with miR-1-1 (*165*). The expression of miR-1 and miR-133a are regulated by the transcription factors MEF2 and SRF (*112, 165*). (B) Experimentally validated targets of the mouse miR-133a-3p (*140, 378-382*). miR-133a-3p regulates heart homeostasis by repressing key mRNAs in processes critical to the heart.

6.2 Results

6.2.1 miR-133a-3p targets

The first step to understanding the function of the miR-133a-3p isomiRs identified (Figure 6.2.A) was to examine the predicted targets for each. The respective seed sequences (nucleotides 2-8) of the major 5' isomiRs (designated can/23 nt and iso/22 nt, Figure 6.2.A) were used to predict mRNA targets using TargetScan (Figure 6.2.B). This analysis yielded both common and unique targets, with both 5' isomiRs being predicted to target many genes involved in cardiovascular disease as determined by GO-term enrichment, although few of the latter were in common. This implies that both 5' isomiRs may regulate similar cardiomyocyte functions though frequently through different mRNA targets.

To experimentally test whether these 5' isomiRs can indeed have different targeting specificities we next needed to select miR-133a-3p targets with the potential to specifically respond to either the can/23 nt or the iso/22 nt variant of miR-133a-3p. Due to the uncertainty associated with targeting prediction algorithms, it was not practical to base our selections on computational predictions alone. For this reason, a mRNA dataset was sourced from a miR-133a knockout mouse study (*279*). Importantly, the whole pre-miRNA sequence, which contains the isomiRs for both the 5p and 3p strands, was deleted in this model. Comparison of viable double knockout mice and wild-type hearts had been performed using microarray analysis at postnatal day 1. To find potential targets of isomiRs generated from the 3p arm, the data was filtered for mRNAs that were upregulated more than 2-fold in the miR-133a$^{-/-}$ hearts compared to the wild-type. This list was then inspected for predicted targets of miR-133a-3p. Each target prediction algorithm uses a different set of criteria to detect targets. For example, the degree of sequence complementarity, the thermodynamics of the binding interaction and/or the phylogenetic conservation of the binding site are often used to find and rank the putative targets (reviewed in (*383*)). For this reason, the list of upregulated genes was cross-referenced with the outputs of several target prediction algorithms (searches were performed in 2011); this included Targetscan for the canonical miR-133a-3p (*67*), a custom Targetscan search using the isomiR seed sequence, the miRanda algorithm at microRNA.org (*384*) and miRNA viewer (http://cbio.mskcc.org/cgi-bin/mirnaviewer/mirnaviewer.pl), which used an earlier version of the miRanda algorithm (*384*). Interestingly, the miRNA viewer used the iso/22

nt miR-133a-3p sequence for its target predictions, providing an excellent resource to identify targets that may be unique to one version of miR-133a-3p. This generated a list of predicted targets for the can/23 nt or the iso/22 nt miR-133a-3p that are upregulated after deletion of the miR-133a locus. This list was then visually inspected for target sites that may exhibit different binding preferences between the can/23 nt and iso/22 nt miR-133a. Two miRNA binding sites were chosen for downstream applications, derived from connective tissue growth factor (CTGF) mRNA for the canonical miR-133a-3p (can/23 nt) and phosphoglycerate mutase 1 (PGAM1) mRNA for the isomiR (iso/22 nt) (Figure 6.2.C). In fact, CTGF has been previously reported as an experimentally validated target of miR-133a-3p that is critical in the production of the ECM during cardiac hypertrophy (*385*).

Next, these target sites were cloned as three concatemers behind Renilla luciferase in a reporter construct (Figure 6.2.C). Expression of this reporter can be a measure of miRNA activity, where an interaction between the miRNA and the target sites results in a decrease in the level of Renilla luciferase protein. Mutated binding sites containing an additional nucleotide in the seed-binding region were also generated. The construction of the luciferase vectors and subsequent transfections are detailed in the next section.

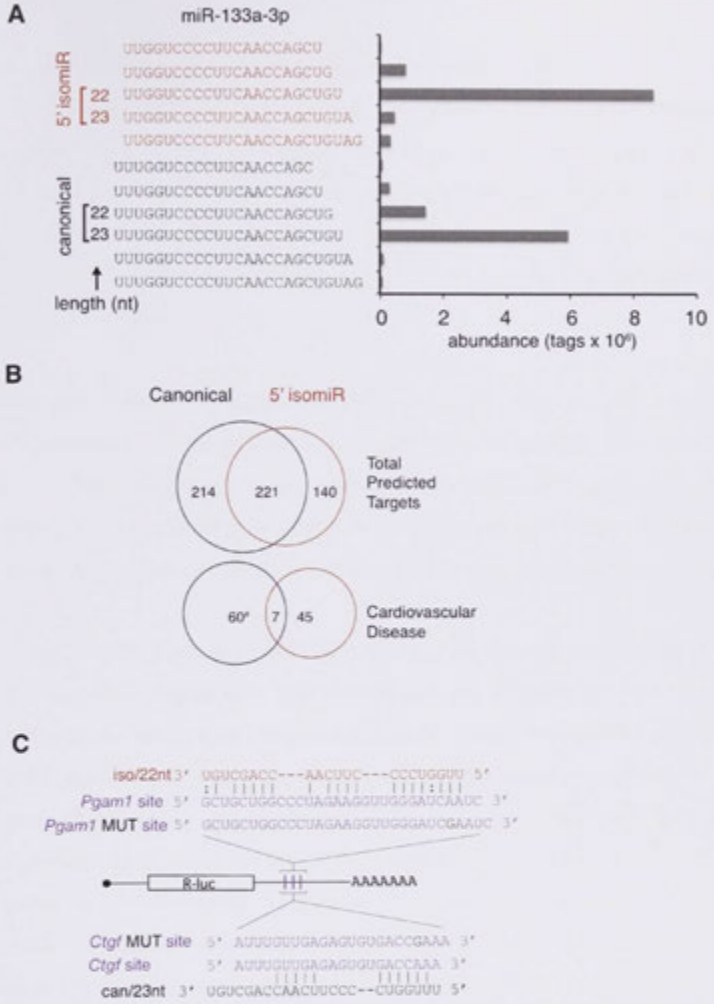


Figure 6.2. miR-133a-3p isomiRs have different predicted targets.

(A) Count frequency of miR-133a-3p isomiRs; isomiRs chosen for targeting analysis are indicated by the arrow. (B) Predicted targets (TargetScan) for the canonical and 5' isomiR of miR-133a-3p. Enrichment of gene function terms using IPA software identified targets involved in cardiovascular disease. (C) miR-133a-3p isomiR-specific binding sites (connective tissue growth factor (CTGF) and phosphoglycerate mutase 1 (Pgam1)) were cloned in three concatemers behind Renilla luciferase in a reporter construct. Mutated binding sites containing an additional nucleotide in the seed sequence were also cloned.

6.2.2 Constructing luciferase vectors

The miR-133a-3p binding sites were ordered as four miRNA binding cassettes within a miniGene in the pIDTSMART-AMP plasmid (see Figure 2.1 for the sequence). Each cassette contained three of each binding site (labeled as Pgam1 3 site, Pgam1 MUT 3 site, Ctgf 3 site, Ctgf MUT 3 site), and a *Not*I restriction site that separated the cassettes. Additionally, each cassette contained a unique restriction site at the 3' end, as well as a *Sal*I site between the first and second binding site. The structure of the pIDTSMART-AMP plasmid is shown in Figure 6.3.A.

The cassettes were isolated from pIDT-SMART using a *Not*I digestion (Figure 6.3.B). The bands corresponding to 100 nt (CTGF 3 sites) and 125 nt (PGAM1 3 sites) were isolated (Figure 6.3.C). These bands were then ligated into a psiCHECK™ vector using the *Not*I site behind the Renilla luciferase gene (psiCHECK™ digested with *Not*I is shown in Figure 6.3.B-C). After ligation, ten clones of each were used for subsequent analysis and cloning.

The ligated psiCHECK-2 products were then checked for the correct insertion size using another *Not*I digestion (Figure 6.4). The correct size was detected for the 125 nt insert (Figure 6.4.A) and the 100 nt insert (Figure 6.4.B) in all clones. The type of cassette and the orientation of the cassette were then determined by performing a restriction enzyme digestion with *Eco*RI, which occurs outside of the *Not*I insert, and the cassette specific restriction enzyme. An example for the CTGF cassettes (psiCHECK-2 with 100 nt insert) is shown in Figure 6.5. The correct cassette and orientation would result in a band ~ 715 nt in length, while an incorrect orientation or incorrect cassette would give bands 800 nt and 830 nt respectively. The gels for the psiCHECK-2 vectors containing the 100 nt inserts are shown in Figure 6.6. Three clones were identified that contained the CTGF 3 sites in the correct orientation, and two clones with the CTGF mutated 3 sites in the correct orientation. The same approach was applied for the 125 nt inserts, with three PGAM1 3 sites in the correct orientation identified and two PGAM1 mutated 3 sites (data not shown). The inserts were then confirmed using conventional sequencing (Appendix 8.14).

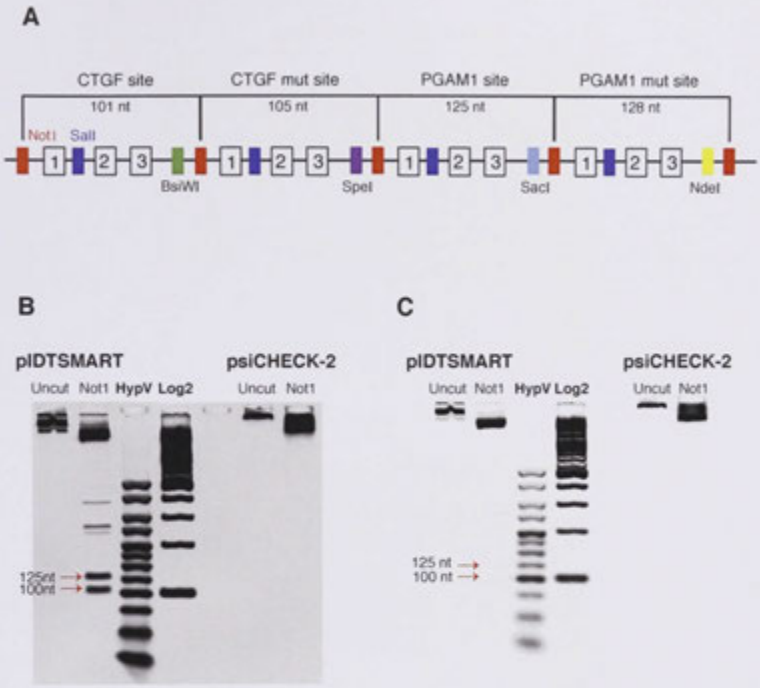


Figure 6.3. pIDTSMART-AMP minigene construct.

(A) Schematic showing the four miR-133a target cassettes (CTGF, CTGF mut, PGAM1 and PGAM1 mut) within the pIDTSMART-AMP minigene. A *Not*I site separated the cassettes. Each cassette contained three miR-133a or isomiR binding sites, with the first and second site separated by a SalI site, and a cassette specific restriction site after the third site. (B) Both pIDTSMART and psiCHECK-2 were digested with *Not*I. (C) The 125 nt (corresponding to PGAM1 cassettes) and 100 nt (corresponding to 125 nt cassettes) were excised and inserted into the *Not*I site of psiCHECK-2.

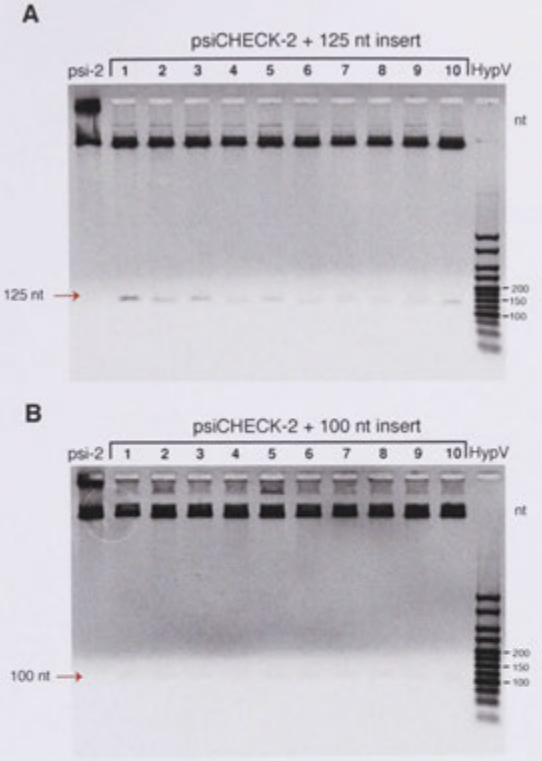


Figure 6.4. psiCHECK-2 contains inserts that correspond to the expected cassette size.

(A) The ten clones of psiCHECK-2 ligated with the 125 nt cassettes contain inserts approximately 125 nt in length. (B) The ten clones of psiCHECK-2 ligated with the 100 nt cassettes contain inserts approximately 100 nt in length. HypV; hyperladder V, psi-2; psiCHECK-2.

Figure 6.5. Strategy for determining correct construct and orientation.

The psiCHECK-2 plasmids thought to contain the CTGF cassettes were digested with *Eco*RI and the cassette specific enzyme (*Bsi*WI or *Spe*I) to determine if the cassette was correct. (A) If the cassette was correct and in the correct orientation, bands of 715 nt and 125 nt would be visible. (B) If the cassette was correct but in the incorrect orientation then a band ~ 800 nt would be seen. (C) If the cassette does not correspond to the chosen cassette specific enzyme, then a band ~ 830 nt would be visible.

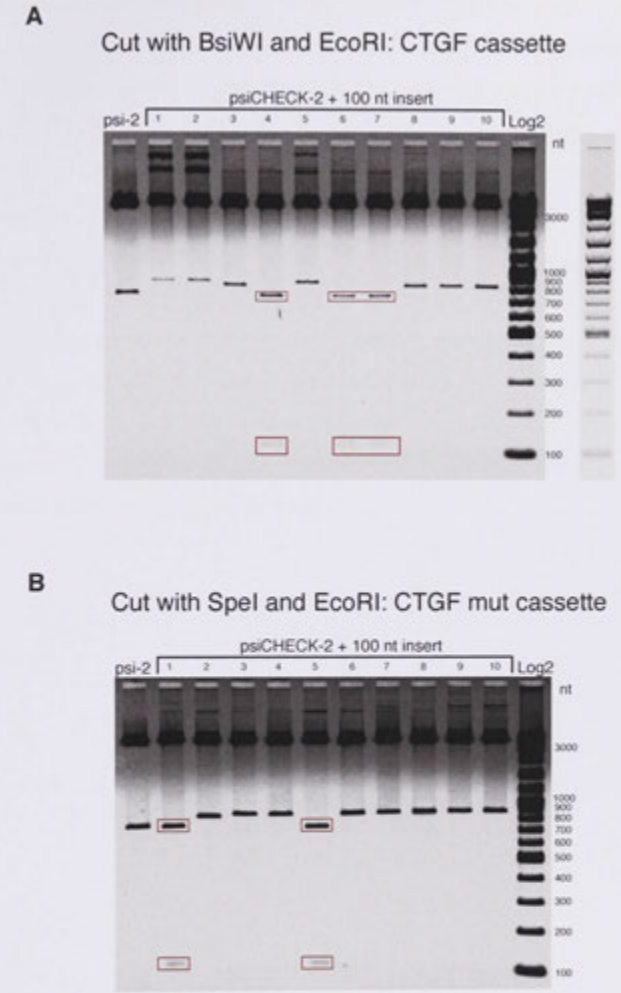


Figure 6.6. Detection of the correct cassettes within psiCHECK-2.

(A) psiCHECK-2 vectors with 100 nt insert were digested with *BsiW*I and *Eco*RI. Clones 4,6 and 7 contain the correct insert (CTGF sites) and are indicated by the red boxes, they have bands corresponding to 715 nt and 125 nt. (B) psiCHECK-2 vectors with 100 nt insert were digested with *Spe*I and *Eco*RI. Clones 1 and 5 contain the correct insert (CTGF mutated sites) and are indicated by the red boxes, they have bands corresponding to 715 nt and 125 nt. Psi-2, psiCHECK-2.

Once the constructs were determined to be correct, further restriction digestions were performed to reduce the number of binding sites in each cassette to two or one (Figure 6.7.A). The *Sal*I site was positioned between the first and second binding site in each cassette. This was used in conjuction with either the *Xho*I site (upstream of the insert site in psiCHECK-2), or the cassette specific restriction enzyme, to reduce the number of sites in the cassettes. Figure 6.7.A shows the excision and ligation to generate a cassette containing only one site. The removal of the site was confirmed by performing another digestion with *Not*I, and the number of sites determined by the size of the bands after gel electrophoresis. Cassettes containing three sites will appear as a larger band on an agarose gel. As shown in Figure 6.7.B, digesting with *Sal*I and the cassette specific enzyme reduced the size of the insert from ~150 nt to less than 100 nt. The same approach was used for generating cassettes with two sites (Figure 6.7.A, data not shown). The cassettes were also confirmed by conventional sequencing (Appendix 8.14).

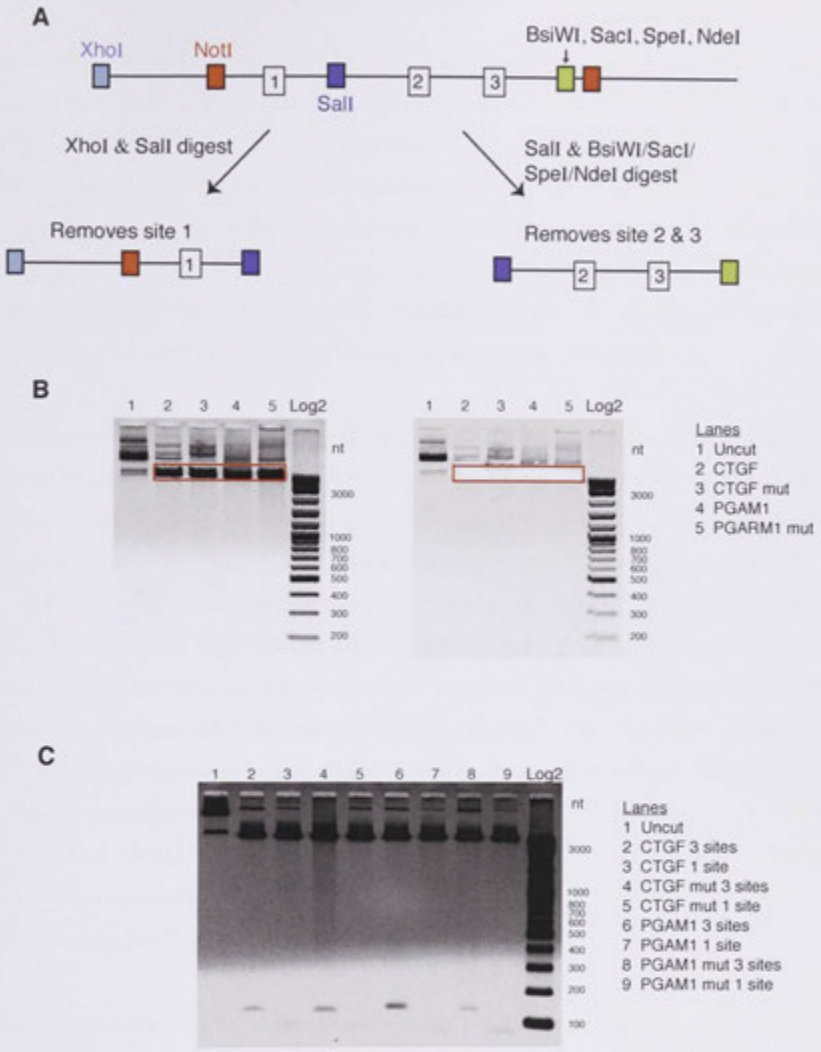


Figure 6.7. Reducing the number of binding sites in psiCHECK-2

(A) Strategy to reduce binding sites to two or one. Digestion with *Xho*I and *Sal*I removes Site 1, leaving two sites in psiCHECK-2. Digestion with *Sal*I and the cassette specific restriction enzyme removes Site 2 and 3, leaving one site in psiCHECK-2. (B-C) Reducing psiCHECK-2 to one site. (B) psiCHECK-2 were digested with *Sal*I and the cassette specific restriction enzyme, the bands were excised, blunt ended and ligated. (C) Digestion with *Not*I confirmed correct number of sites. The casettes with three sites gave rise to a product ~150 nt, while cassettes containing only one site were less than 100 nt.

6.2.3 Function of miR-133a-3p isomiRs

The function of the 5' isomiRs for miR-133a-3p (can/23 nt and iso/22 nt), were tested using the constructs generated in 6.2.2. Custom Synthetic MISSION miRNA mimics were designed to correspond to the can/23 nt and iso/22 nt variants (Shown in Figure 6.2.A). These mimics are double-stranded RNA molecules that do not undergo processing. This is critical when examining the function of isomiRs as the start and end site positions need to be exact. Furthermore, the opposing strand of the miRNA mimic is chemically modified so it cannot be incorporated into the RISC complex, minimizing any unintended effects by the passenger strand. A non-specific mimic was also generated as a control.

The mimics were transiently transfected into HeLa cells, which express low levels of miR-133a-3p, with one of the psiCHECK-2 vectors containing either CTGF 3 sites, CTGF 3 mutated sites, PGAM1 3 sites or PGAM1 3 mutated sites. To first determine the appropriate concentration for comparisons between the miRNA mimics, a serial dilution experiment was performed using the CTGF 3 sites with the can/23 nt mimic, and the PGAM1 3 sites with the iso/22 nt mimic (Figure 6.8.A-B). We used a concentration of 10nM in subsequent experiments as it resulted in sufficient repression for both combinations (Figure 6.8.A-B) and had minimal cell death (data not shown). The CTGF site was strongly repressed by the can/23 nt mimic, with a 20-fold repression observed at the higher concentrations of can/23 nt (Figure 6.8.A). In comparison, the maximal repression of the PGAM1 site with the iso/22 nt mimic was approximately 8-fold (Figure 6.8.B). This suggested that the CTGF site was more efficiently repressed by the can/23 nt miR-133a than the PGAM1 site was by the iso/22 nt.

Comparisons were then made between the can/23 nt and iso/22 nt miR-133a mimics and their effects on the 3 site luciferase constructs. This revealed marked preferential targeting by each variant (Figure 6.8.C). There was significantly greater repression of the CTGF 3 site construct by the canonical miR-133a-3p (can/23 nt) compared to the isomiR (iso/22 nt), and *vice versa* for the PGAM1 3 site construct. Minimal repression was observed for the mutated sites (CTGF MUT 3 sites and PGAM1 MUT 3 sites) and no repression was detected in the presence of the control mimic. This provides proof-of-principle evidence for preferential alternative mRNA targeting by 5' isomiRs of miR-133a-3p.

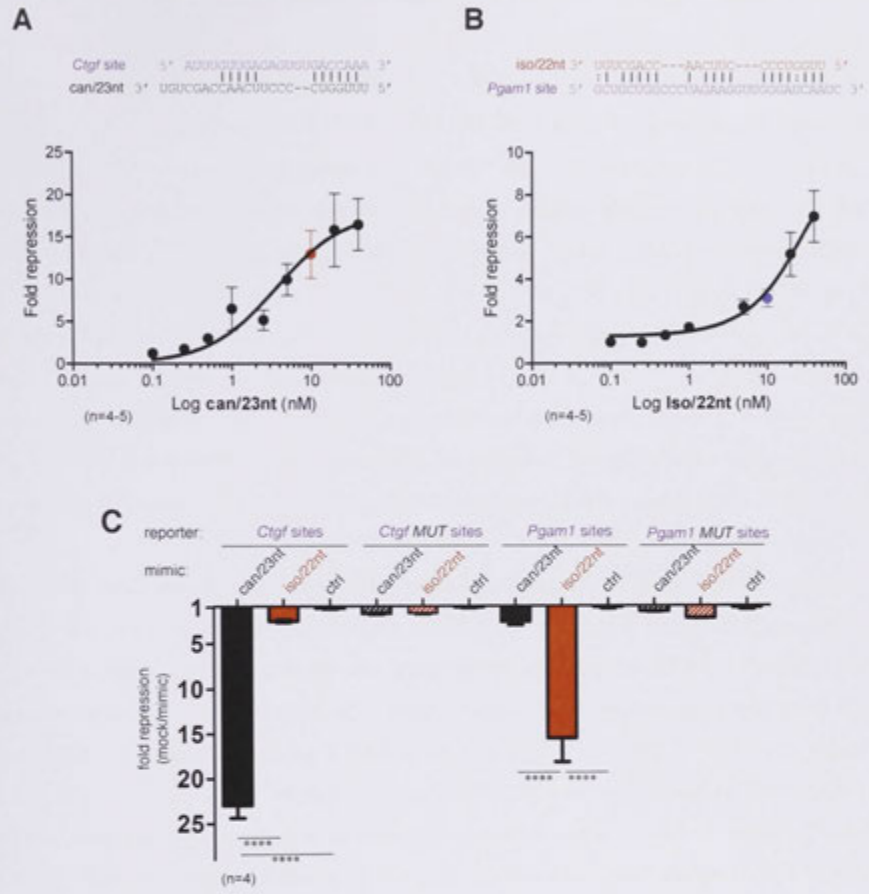


Figure 6.8. miR-133a isomiRs have different targeting properties.

(A) The fold repression of the CTGF sites psiCHECK-2 vector is shown for 0.1nM to 100nM of can/23 nt mimic. (B) The fold repression of the PGAM1 sites psiCHECK-2 vector is shown for 0.1nM to 100nM of can/23 nt mimic. The binding site and miRNA mimic is shown above each graph. 10nM of each miRNA mimic was used for downstream applications (C) miR-133a-3p 5' isomiRs exhibit binding site-dependent differential repression activity, not observed with mutated binding sites (**** $p < 0.0001$, one way-ANOVA with Bonferroni's post-hoc analysis). Data is presented as mean with SEM (n=4-5).

The endogenous mRNAs of CTGF and PGAM1 contain only one target site within their 3' UTR. Therefore to better match the endogenous mRNA, luciferase constructs with fewer sites were used for transfections. The cassette containing CTGF 2 sites was still robustly repressed by the can/23 nt miR-133a, with little to no repression by the iso/22 nt mimic or the control (Figure 6.9.A). However, the PGAM1 2 site cassette was repressed only slightly by the iso/22 nt mimic and this did not reach significance (Figure 6.9.B). When considering the cassettes that contain only 1 binding site, can/23 miR-133a still repressed the CTGF site to a significantly greater extent than the iso/22 nt mimic (Figure 6.10.A). However, when looking only at one site it is evident that the mutated CTGF site is also significantly repressed by both the can/23 nt and the iso/22 nt. Conversely, can/23 nt and iso/22 nt showed no differential repression of the PGAM1 1 site (Figure 6.10.B). While this repression was significant compared to the control miRNA, the mutated PGAM1 site is regulated to the same extent. In summary, the CTGF sites show selective repression by the can/23 nt miR-133a even when only one site is present in the 3'UTR.

The two mimics initially chosen varied in length (22 versus 23 nt). Therefore, we wanted to ensure that the 3' end or the length of the miRNA did not impact on targeting specificity. Two more mimics were generated that represented the other length for each miR-133a variant (can/22 nt and iso/23 nt). These mimics were then transfected with the psiCHECK-2 vectors containing 3 binding sites (CTGF 3 sites, PGAM1 3 sites, Figure 6.11). The targeting ability of the canonical miR-133a was not altered by the miRNA length, with the fold repression remaining the same for the can/22 nt mimic (Figure 6.11.A). The introduction of the iso/23 nt with the PGAM1 3 sites did result in a slightly lower repression compared to the iso/22 nt mimic, but this difference was not significant (Figure 6.11.A). In conclusion, for these two examples, the 3' end of the miRNA does not impact on the specificity of the miRNA.

The final question we asked was whether the miR-133a-3p isomiRs were acting cooperatively, or competing for the same binding sites within the 3' UTR. This was addressed by transfecting different proportions of two miR-133a-3p mimics as a mix while keeping the total level of transfected mimics constant, together with either the CTGF 3 site or PGAM1 3 site plasmids (Figure 6.11.B-C). Increasing the proportion of can/23 nt relative to the iso/22 nt correlated with the increased repression of the CTGF 3 site (Figure 6.11.B). By contrast, altering the proportions in a mix of can/23 nt and can/22 nt did not have an impact on the repression of the CTGF 3 site. Similarly, increasing the

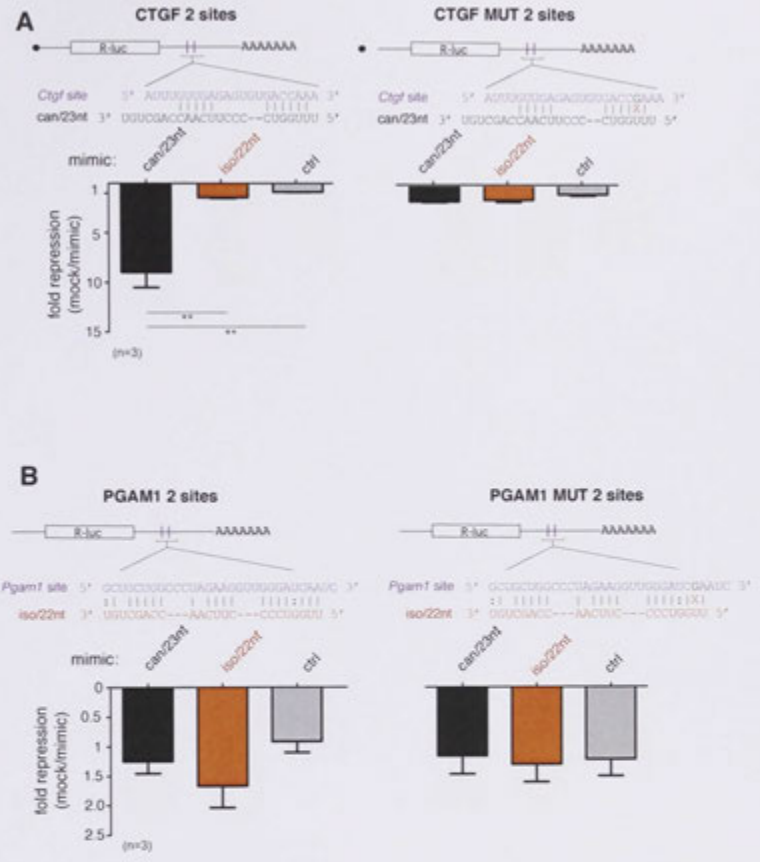


Figure 6.9. CTGF 2 sites is selectively repressed by the canonical miR-133a-3p.

(A) miR-133a-3p 5' isomiRs exhibit differential repression against two copies of the CTGF site; the can/23 nt miR-133a-3p represses this cassette while the iso/22 nt and the control mimic do not. No repression is seen for two copies of the CTGF mutated site. (B) The iso/22 nt mimic represses two copies of the PGAM1 site slightly more than the can/23 nt but this is not significant. (** $p < 0.01$, one way-ANOVA with Bonferroni's post-hoc analysis). Data is presented as mean with SEM (n=3).

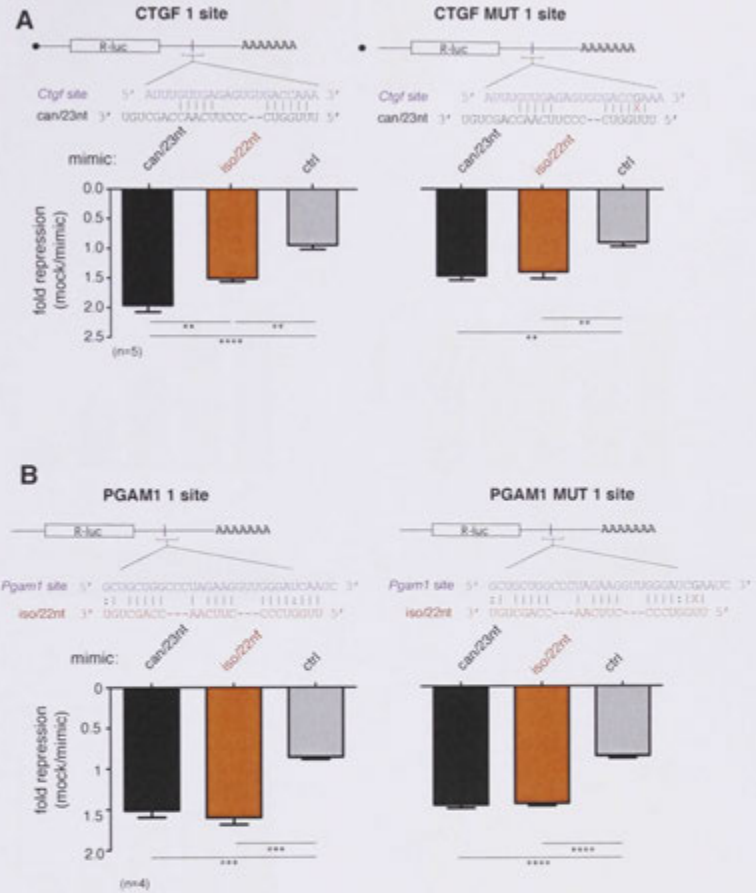


Figure 6.10. CTGF 1 site but not PGAM1 1 site is selectively repressed by the canonical miR-133a-3p.

(A) miR-133a-3p 5' isomiRs exhibit differential repression against one copy of the CTGF site; the can/23 nt miR-133a-3p represses this cassette 2-fold while the iso/22 nt mimic represses only 1.5-fold, and the control mimic does not. A small amount of repression is seen for one copy of the CTGF mutated site. (B) The iso/22 nt mimic and the can/23 nt mimics repress one copy of the PGAM1 site and one copy of the PGAM1 mut site to the same extent. (** $p < 0.01$, one way-ANOVA with Bonferroni's post-hoc analysis). Data is presented as mean with SEM (n=4-5).

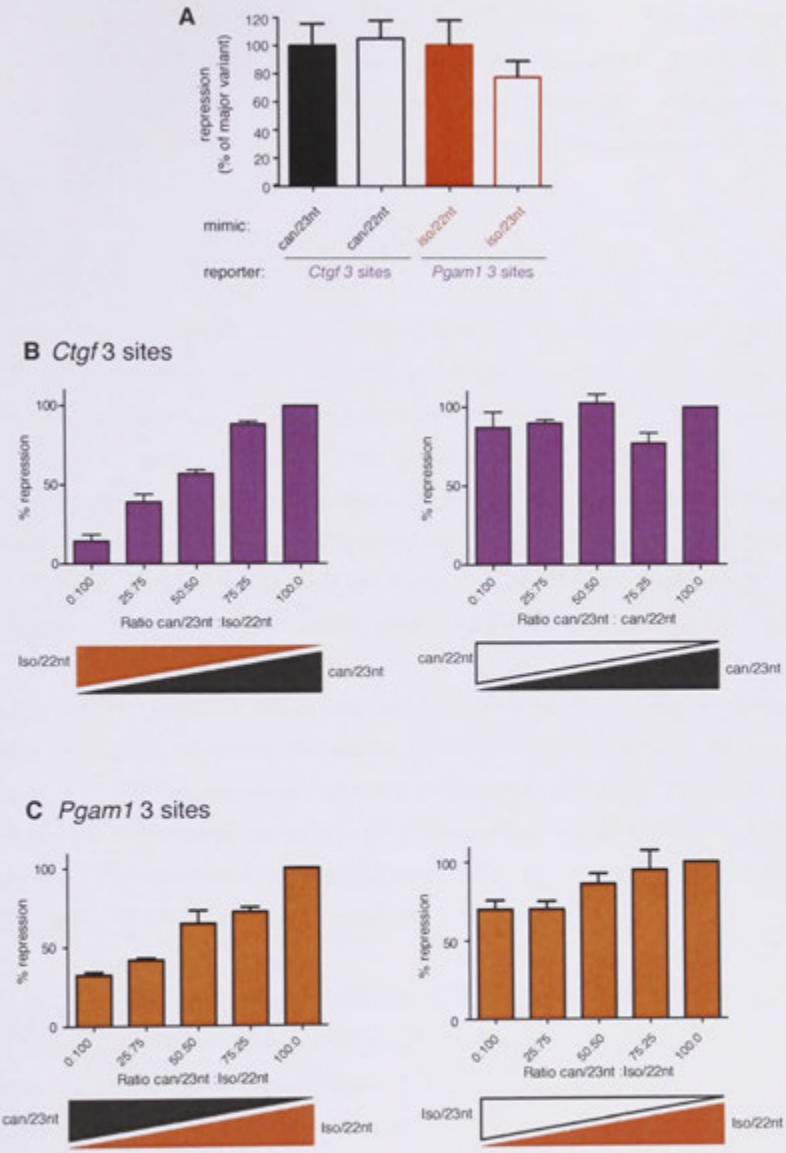


Figure 6.11. miR-133a-3p 3' isomiRs do not impact on targeting specificity.

(A) miR-133a-3p 3' isomiRs do not exhibit binding-site dependent repression at either the CTGF or PGAM1 3 sites. (B-C). miR-133a isomiRs do not competively bind to the target site. (B) Repression of the CTGF 3 sites with changing proportions of the iso/22 nt and can/23 nt or can/22 nt and can/23 nt. (C) Repression of the PGAM1 3 sites with the can/23 nt and iso/22 nt or iso/23 nt and iso/22 nt.

proportion of iso/22 nt relative to the can/23 nt correlated with the increased repression of the PGAM1 site (Figure 6.11.C). Again, changing the proportion of 3' isomiRs had minimal impact on the repression of the PGAM1 sites. Collectively, these results suggest that for the binding sites examined, the miR-133a-3p isomiRs do not compete for the same binding site.

6.3 Discussion

The miR-133a-3p 5' isomiRs are predicted to have overlapping and unique targets, with both being predicted to target mRNAs involved in cardiovascular disease. Using luciferase constructs containing a known and a predicted miR-133a binding site we have shown that miR-133a-3p 5' isomiRs can have different targeting specificity and efficiency. The largest difference in targeting was observed when using constructs that contained 3 binding sites. The experiment was designed to get maximal repression of the binding sites so that any differences in miRNA targeting could be observed. However, three consecutive miRNA binding sites for the same miRNA are not often observed within an endogenous 3' UTR.

It is important to note that due to technical limitations and the inability to discriminate between the isomiR sequences, the relative expression level and AGO incoportation of each isomiR was not measured following transfection. Differing concentrations and incorporation of the isomiR mimics would affect their activity and subsequent repression levels. However, one isomiR strongly repressed only one of the luciferase constructs, suggesting the concentrations were comparable.

The can/23 nt mimic had a stronger effect on the CTGF 3 sites, compared to the effect of iso/22 nt on the PGAM1 3 sites. The PGAM1 binding site contains a G:U pair (or G:U wobble) within the seed region (Figure 6.2.C). This G:U wobble is tolerated in a miRNA:mRNA interaction if there is sufficient binding at 3' end to compensate (*85*). It has been demonstrated that a G:U wobble within the seed region can still result in protein repression but is less effective (*386, 387*), which could explain the stronger repression exhibited by the CTGF 3 sites. This difference becomes more evident when the number of binding sites is reduced to 2 or 1. The CTGF site still remains selectively repressed by the can/23 nt mimic when only one site remains. However, when there is only one PGAM1 site both the can/23 nt and iso/22 nt variants of miR-133a repress the 3' UTR to the same weak extent. This might be attributable to the extended complementarity along the full

length of the miRNA, which is known to compensate for mismatches in the 5' end of the miRNA (*388*)(Figure 6.2.C). This would also explain why the PGAM1 mutated site (Figure 6.10.B) is repressed to the same weak extent as the PGAM1 site. Despite the PGAM1 sites showing no differential repression when only 1 site is present, these experiments still provide evidence that miRNA isomiRs can have different targeting specificity, in particular, that miRNA variants can have both common and unique mRNA targets.

Another recent study reported that 54% of miR-133a-3p in the rat left ventricle is equivalent to the iso/22 nt variant (*389*). They also demonstrated using the DIANA-microT targeting prediction algorithm that the canonical miR-133a and the iso/22 nt variant have overlapping and unique targets, which is consistent with the analysis presented in this Chapter (*390*). The 3' UTR for one of these targets, Gelsolin, was cloned behind a reporter gene and it was shown that the canonical miR-133a-3p was more efficient at suppressing this target compared to the iso/22 nt variant. This study provides further evidence that the 5' isomiRs of miR-133a-3p can have different targeting specificity.

6.3.1 Outlook:

To extend the conclusions obtained from our experiments and to appreciate the full function of miR-133a-3p isomiRs, similar studies should be conducted in a cell type relevant to its expression, such as the HL-1 cardiomyocyte cell line. This would allow for the identification of endogenous targets in the relevant cellular system. However, miR-133a-3p is already highly expressed in cardiomyocytes so its targets are likely already repressed. Furthermore, the high levels of endogenous miR-133a-3p would interact with the targets being interrogated making it difficult to untangle the effects of a synthetic miRNA mimic. Ideally, a system where the endogenous levels of miR-133a-3p were reduced and specific variants reintroduced would be used to identify their targets. One study has overexpressed miR-133a in the mouse heart and sequenced the RNA present in the RISC complex to identify mRNA targets (*76*). An experiment such as this using the can/23 nt and iso/22 nt would provide the information necessary to uncover the mRNA target network of miR-133a-3p 5' isomiRs.

There are several hypotheses for the biological function of 5' isomiRs. It may be that the isomiRs have completely different mRNA targets that increase the complexity of the miRNAs, or that the co-expression of isomiRs leads to reinforced repression of core

targets, and reduces off-target effects (*85*). Either way, the existence of such isomiRs needs to be taken into account when analysing the function of miRNA or considering a miRNA for therapeutic intervention against a disease.

Chapter Seven

General Discussion & Future directions

7.1 Discussion

In this thesis I have presented detailed miRNA and mRNA profiles of cardiomyocytes purified from normal, pre-hypertrophic and compensated hypertrophic murine hearts. A model of LV hypertrophy was successfully established and confirmed by haemodynamic, physiological and biochemical measurements. RNA extracted from purified cardiomyocytes was then used for NGS of both small RNAs and mRNA 3' ends. Differential expression of both miRNAs and mRNAs was detected, as well as processing variation in miRNA sequence and mRNA 3' UTRs. Firstly, a static picture of all the variants present in cardiomyocytes was generated, and then the data was analysed for differences in processing between control and hypertrophic samples. The functional role of 5' isomiRs for a cardiac specific miRNA was then determined using *in vitro* luciferase assays. In this chapter, I discuss the outcomes of these studies, integrating both the miRNA and 3' end datasets and the importance of this dataset for cardiac biology.

Cardiac hypertrophy is a dynamic process that involves several signalling pathways and multiple levels of gene regulation. Furthermore, we now appreciate that gene expression is highly regulated at the post-transcriptional level. The precise balance of miRNAs and mRNA 3' UTRs expressed within a cell control the amount of protein produced from a particular mRNA transcript. Our study makes a key advance in understanding this phenomena by considering alternative processing of miRNAs and mRNA 3' UTRs in these processes. Previous studies have often focussed on one aspect of gene regulation in cardiac hypertrophy. This is the first study to simultaneously document miRNA processing variants and mRNA 3' UTRs. Furthermore, this is the first study to use enriched cardiomyocytes for miRNA and mRNA differential expression. Other studies looking at miRNA expression in mouse models of cardiac hypertrophy have taken the whole LV and most have been based on microarray analysis (*120, 121, 127, 128*), resulting in incomplete datasets defined by the probesets represented on the arrays and confounded by the presence of non-cardiomyocytes in the sample. Two studies have used NGS to profile miRNAs after TAC, but again this was in the whole LV and they did not interrogate the data for miRNA processing variants (*130, 131*). Of the miRNA studies, only one other has looked at an earlier time point and they reported that 24 hours after TAC-banding only one miRNA, miRNA-1, changed in expression (*127*). In addition, these studies have not looked for the presence of miRNA processing variants or assessed for changes in cardiac

hypertrophy. The expression levels of mRNAs in response to TAC have been described at multiple time points, including 48 hours and 7 days (*168, 187, 301, 302*). Alike to the miRNA analysis, these studies have used the whole heart or whole LV and microarray analysis. One study has performed NGS for both miRNAs and mRNAs at 7 days post-TAC, as well as NGS of mRNA transcripts present within RISC to establish miRNA: mRNA interactions (*131*). However, this study did not take into account the presence of multiple 3' UTRs per mRNA transcript. Lastly, the 3' UTRs of some mRNA transcripts have been documented in whole LV samples using microarray analysis (*168*). This study also documented shortening of some 3' UTRs in TAC banded mice after 7 days, but the publication did not provide an extensive dataset of the 3' UTRs expressed as it relied on microarray probes. To summarise, this is the first study to document expression changes in miRNAs and mRNAs in cardiomyocytes purified from pre- and hypertrophic hearts. Furthermore, this is the first study extensively documenting the miRNA processing variants and 3' UTRs expressed in primary murine cardiomyocytes, and thus to identify miRNA variants and 3' UTRs that change in response to pressure overload-induced hypertrophy.

7.1.1 Transcriptional response to pressure overload

In addition to documenting processing variants, differential expression analysis was conducted for both miRNAs and mRNAs. There are some interesting observations when comparing the results of both analyses. Firstly, as indicated by the Spearman correlation, there was less variation between cardiomyocytes in the miRNA population compared to the mRNA transcriptome. The miRNA population was dominated by relatively few, highly expressed cardiomyocyte-enriched miRNAs that did not alter in expression levels, thus resulting in high Spearman correlations. This was also reflected in the differential expression analysis, with only a small number of miRNAs being statistically significant in comparison to the mRNA differential expression.

The second major difference in the expression profiles for miRNAs and mRNAs was the kinetics for the response to pressure overload. Two miRNAs had significant changes in expression at 48 hours, while 16 significantly changed expression at 7 days. In contrast, 469 mRNAs had significant differential expression at 48 hours and 307 were significantly altered at 7 days. Therefore regulation at the level of mRNA abundance occurred faster than alteration in miRNA levels. This could be due to a number of reasons. Firstly,

miRNAs have an average half-life of ~ 5 days and are up to 10 times more stable than mRNA transcripts (*215*). Thus a reduction in the level of a miRNA may take longer than that of an mRNA transcript. Secondly, many miRNAs are processed from the introns of protein-coding genes and thus the levels of these host genes could be expected to change first, followed by a subsequent change in the miRNA level. For example, miR-208b is encoded within β-MHC, which is increased slightly at 48 hours and significantly overexpressed at 7 days (*248*). The same trend is observed for miR-208b-5p, with a larger increase at 7 days but the fold change values are small and it does not reach significance. The processing time of a miRNA could also account for the delayed miRNA response. In support of this, the levels of the pri-miR and pre-miR are often upregulated well before the mature miRNA can be detected (*215, 391*). Therefore, the stability and processing of a miRNA could be why the mRNA levels change prior to the miRNAs. In support of this finding, another study looking at a time-course of miRNA and mRNA regulation in melanoma cells reported the same phenomenon (*392*). After stimulation with interferon gamma (IFN-γ), the mRNA expression levels change first and then the miRNA levels were altered by a second wave of transcription. It was hypothesised that the increase in the levels of transcription factors in the initial response was required for the activation of miRNAs. It is therefore possible that the mRNA levels increase soon after TAC banding to initiate a compensatory response, which is then subsequently followed by an increase in miRNAs to fine-tune or adjust the hypertrophic response (*392*).

7.1.2 miRNAs and mRNAs exist as processing variants

One of the major goals of this thesis was to determine the extent of sequence variation in miRNAs as well as mRNA 3' UTR usage in murine cardiomyocytes. I have presented evidence to suggest that both miRNAs and mRNA 3' UTRs have substantial variation in their processing. Nearly all of the miRNAs expressed in the heart showed some degree of sequence variation including arm bias, 5' and 3' isomiRs as well as non-templated additions. Similarly, the mRNAs do not always express the annotated 3' UTR and an mRNA has on average 2.35 3' UTRs. This complicates our view of the miRNA:mRNA interaction network, with several different interactions possible from the same miRNA and/or mRNA genomic loci depending on which variant is produced.

The static maps of mRNA 3' UTRs and miRNA processing variants have important implications for cardiac biology. It is critical when looking at the impact of a miRNA in a

cellular system to have the details about the miRNAs and mRNA 3' UTRs that are expressed. This is particularly important when predicting miRNA targets. The target prediction algorithms are mostly based on the 3' UTRs from cDNA sequence repositories such as Refseq (*393, 394*). Release 6 of Targetscan now includes all 3' UTRs from Refseq rather than using only the longest 3' UTR from each gene (*67*). However, this only includes annotated 3' UTRs and the target prediction algorithms don't take into account cell-specific 3' UTRs. Furthermore, the canonical miRNA defined in miRBase is used for target predictions. Thus, the accuracy of target predictions would be improved if the cell or tissue-specific 3' UTR and miRNA variants as described here were used as the input.

7.1.3 Processing changes with cardiac hypertrophy

The NGS data was also analysed for changes in miRNA processing and 3' UTR usage between the Sham and TAC cardiomyocytes. While I found that the majority of miRNAs exist as processing variants, in general these variants did not change in a directional manner during the induction of hypertrophy. This is in contrast to the 3' end sequencing data, where 3' UTRs variants for over 400 mRNAs were found to change in proportion in the pre-hypertrophic cardiomyocytes. This suggests that miRNA processing remains fairly consistent while the 3' UTRs are changing. This could be attributable to the processing steps required to generate these two RNA types. The 3' UTR of an mRNA contains more cis-regulatory elements and thus has more scope for alterations in processing.

7.1.4 Integration of miRNA and 3' UTR data

The long-term goal of this work, beyond the scope of this PhD thesis, is to integrate the miRNA and 3' UTR data to generate a systems level interaction map of miRNAs and their targets during hypertrophy. We hypothesised that there would be examples where alterations in the 3' UTR would disrupt the interaction with a miRNA (Figure 7.1.C) or miRNA processing variants would be produced with a different targeting spectrum (Figure 7.1.D). I have already demonstrated how the isomiRs of miR-133a-3p can have different targeting properties but the function of 3' UTR variants remains unexplored. I started mining the data for examples of miRNA: 3' UTR interactions and found an example where integrating both datasets revealed potential for atered miRNA regulation that was previously hidden.

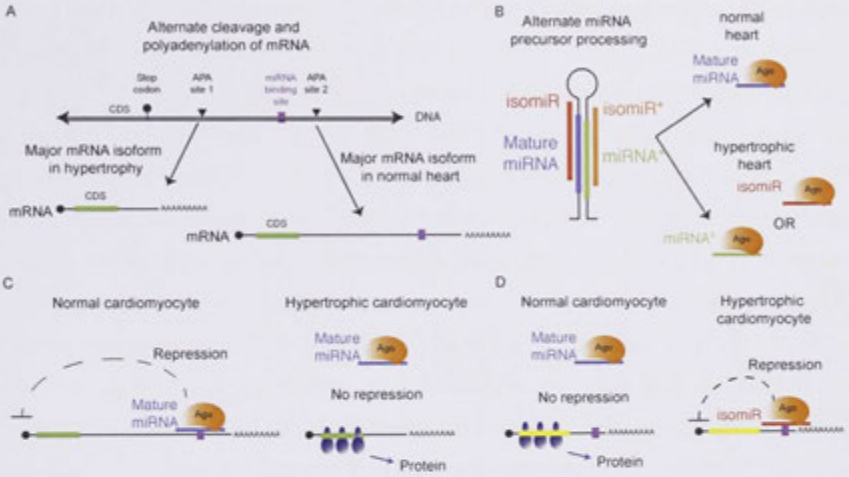


Figure 7.1. Potential role of non-conventional miRNA and mRNA processing in cardiac hypertrophy.

(A) Cleavage of an APA site results in mRNAs with different 3' UTR length and miRNA binding sites (purple box). (B) Non-canonical miRNA processing can result in the other strand or isomiRs (with different 5' start positions and targeting properties) being incorporated into RISC. (C) Hypertrophic-specific APA usage can result in disrupted miRNA-mediated repression. (D) Hypertrophic-specific miRNA processing can produce a novel miRNA species able to repress disease-relevant mRNAs.

Cyclin D2 (Ccnd2) regulates the G1-progression of the cell cycle and is necessary for the development of hypertrophy (*308*). CCnd2 levels increase during hypertrophy and CCnd2 phosphorylates retinoblastoma (Rb)(*395*). Hypertrophic growth requires new protein synthesis and transcription of tRNA genes by RNA pol III. When phosphorylated Rb becomes inactive, its association with subunits of RNA pol III-specific transcription factor B is disrupted, allowing them to activate RNA pol III (*395, 396*). Ccnd2 is expressed as three major 3' UTRs in the 3' seq data; one proximal and two distal 3' UTRs approximately 4.5 kb downstream (Figure 7.2.B.i). Both miR-133a-3p and miR-145-5p repress the expression of the Ccnd2 3' UTR in luciferase assays, indicating they directly target Ccnd2 (*279, 397*). The expression levels of miR-145-5p and miR-133a-3p do not significantly change in the pre-hypertrophic or hypertrophic cardiomyocytes (Figure 7.2.A). However, the proportion of the short 3' UTR increases at both time points, with a larger increase at 48 hours, compared to the two longer 3' UTRs (Figure 7.2.ii). The short 3' UTR is not under the regulation of miR-145-5p or miR-133a-3p, so while their levels have not changed, altering the 3' UTR of Ccnd2 changes its post-transcriptional regulation. There is no obvious correlation between this 3' UTR change and the overall levels of the Ccnd2 mRNA (Figure 7.2.B.ii). However, miRNAs can act via translational repression and thus the protein levels of Ccnd2 may increase. In agreement, it has been previously shown that the protein levels of Ccnd2 increase after TAC and after treatment with angiotensin II (*230, 308*). This illustrates an example where miRNA shows no change while the 3' UTR does, which could still result in different protein levels.

This example demonstrates how this data can be used to understand the intricacies of miRNA regulatory networks in the heart. There may be instances where the miRNA changes but the predicted binding sites may not be present, or where the levels of the miRNA remain constant but the 3' UTR changes, releasing the repression of the miRNA.

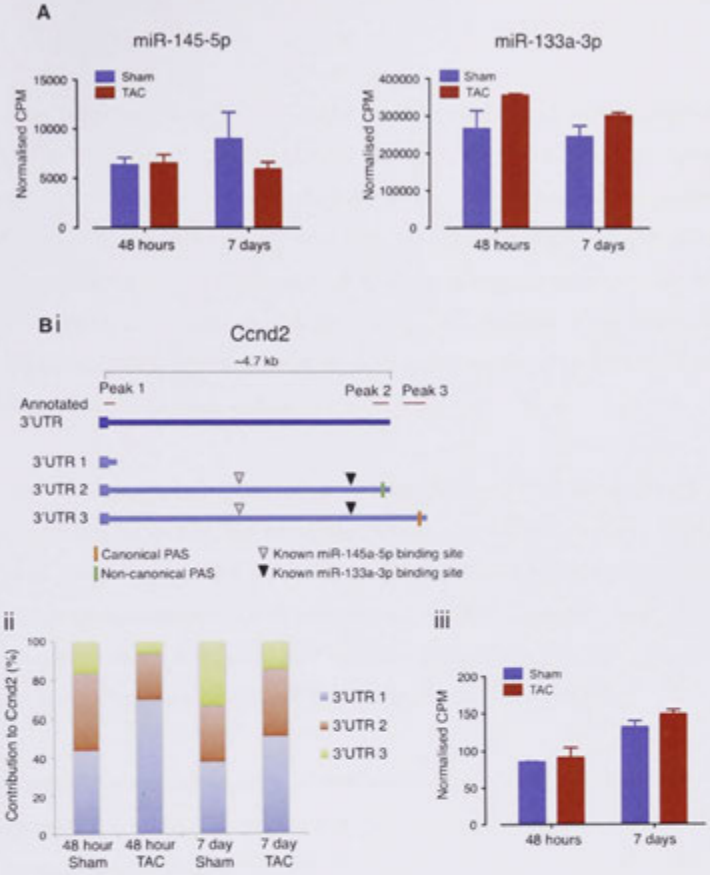


Figure 7.2. The shortening of Ccnd2 3' UTR removes miRNA binding sites.

(A) Levels of miR-145-5p and miR-133a-3p according to NGS. The data is presented as normalised CPM. (B) (i) Schematic of the Ccnd2 3' UTR. The dark blue box indicates the annotated 3' UTR, with the location of the 3' seq peaks indicated above by the red lines. Each 3' UTR represented by the 3' seq peak drawn below. The coloured bars represent the canonical or non-canonical PAS. The black triangles indicate known miR-133a-3p binding sites and the grey triangles indicate known miR-145a-5p binding sites. (ii) Change in proportion of the 3' seq peaks for Ccnd2; data shown as percentage contribution to Ccnd2. (iii) Levels of the Ccnd2 transcript according to NGS. The data is presented as normalised CPM.

7.1.5 Cardiac miRNAs as therapeutics for cardiovascular disease

The data from this thesis is important to the development of appropriate miRNA-based therapeutics for heart disease (*398*). Although the processing of miRNAs was not altered during cardiac hypertrophy, an understanding of the miRNA variants expressed in the heart is crucial when designing therapeutics that remove or introduce a specific miRNA. The current treatments for CVD primarily slow down disease progression and as a result the prevalence, mortality, and costs associated with CVD continue to increase (*399*). The ability to manipulate miRNA expression in animal models has triggered enthusiasm for miRNAs as potential therapeutic targets.

There are two main approaches for miRNA therapeutics; miRNA antagonists to reduce endogenous miRNA levels that exhibit a gain-of-function in disease or miRNA mimics to restore loss-of-function (reviewed in (*400*)). miRNA antagonists are chemically modified miRNA strands complementary to the target miRNA that bind with high affinity and inhibit its activity. miRNA mimics or "miRNA replacement therapy" reintroduces a miRNA that is expressed in a healthy cell but lost in the disease state (*401*).

The most clinically advanced example of miRNA therapeutics is against hepatitis C virus (HCV) infection. An antagonist against miR-122 blocks the replication of HCV and is currently being evaluated in Phase 2 clinical trials (*402, 403*). Thus far, no miRNA-based therapies against cardiovascular disease have entered human clinical trials but several candidates are in preclinical development. MGN-9103, an antagonist against miR-208 and miR-499, is in development with Santaris Pharma as a therapy against chronic heart failure. Deletion of the miR-208 hairpin in mice results in resistance to fibrosis, cardiac hypertrophy and heart failure after cardiac stress, suggesting that decreasing the levels of miR-208 could prevent or reduce the damage induced from heart disease (*114*). However, 55% of the miR-208a hairpin is derived from the 5p arm in our NGS data, meaning that both strands of miR-208a are available for target interactions. It is important to establish if the proportions of miR-208 strands are expressed in equivalent proportions in the human heart. The function of miR-208 has been ascertained by knocking out the miR-208 hairpin and thus decreasing expression of both strands. Therefore, to achieve the protective effects achieved in the knockout mouse model, it may be necessary to design antagomiRs against both strands of miR-208 for therapeutic use.

There are also challenges associated with the delivery of miRNA mimics or inhibitors. Currently, there is not an effective cardiac-specific delivery method for RNA therapeutics and as such, treatment would rely on systemic injection (*404*). As one miRNA can have hundreds of targets, there is the risk of side effects in other organs or cell types. However, a tissue-specific miRNA would have reduced risk of non-target effects on other organs. For example, inhibition of miR-208a, which is heart specific, by subcutaneous delivery of an antagonist was sufficient to prevent cardiac remodelling and improve cardiac function in hypertensive rats (*114, 405*). The successful use of miRNAs as therapeutics for cardiovascular disease will also rely on improvements in the delivery or targeting of the molecule to be tissue and/or cell specific.

The future of disease-associated miRNA research will need to address the complexity of mRNA and miRNA isoforms to better understand disease pathology and develop precision miRNA-based therapeutics. Therefore, this study may help to identify miRNA-based points of therapeutic intervention by understanding which miRNA isoforms support the development of hypertrophy and which mRNA isoforms they regulate to achieve this outcome.

7.2 Future Directions

Our NGS has provided a wealth of information about miRNA and mRNA 3' UTR isoforms in cardiomyocytes. Several questions arise as a result of the data presented in this thesis, some of which have been raised already in this discussion.

The next stage of analysis would be to integrate the miRNA changes with mRNA 3' UTR usage on the global scale. As mentioned above, the long-term goal of this work is to provide a systems level understanding of the miRNA interactions that occur in hypertrophy. The hypertrophy-related changes to miRNA sequences and mRNA 3' UTR usage could be bioinformatically cross-referenced to known and predicted miRNA:mRNA interactions. This analysis will answer key questions, such as: are cardiac miRNAs targeting mRNAs that are modulating their 3' UTR and thus binding sites in hypertrophy? Are the miRNA variants such as isomiRs targeting the same mRNAs? Are hypertrophy-related mRNAs changing their 3' UTR length to prevent or gain post-transcriptional gene regulation during the hypertrophic response? This would also generate a list of potential

interactions that could be validated with other methods or used for experiments investigating the functional implications of such variants.

The miRNA:mRNA target interaction network could then be further improved using methods that identify targets by their physical interaction with the miRNA. As mentioned previously, such an experiment has been conducted in the mouse heart by overexpressing miR-133a and sequencing the RNA present in the RISC complex (*76*). Overexpression of different miRNA isomiRs would identify their unique and common targets. Furthermore, if the exact binding site was determined then the interaction with 3' UTR variants could also be determined.

Questions still remain about the biological function and mode of action of miRNA variants and 3' UTR length. Apart from different targeting properties of miRNA variants, it would be interesting to see if isomiRs act through different mechanisms, for example deadenylation, exonucleolytic decay or translational repression. This could be discerned using polysome profiling, which measures the translational properties of specific mRNA molecules (*406*). This method could also be used to determine the consequences of 3' UTR site choice on gene expression. It is assumed that a shorter 3' UTR, due to use of a proximal poly(A) site, will increase translational efficiency and mRNA stability due to loss of *cis*-regulating elements, including miRNA binding sites. The translational efficiency of cardiac mRNAs identified from the sequencing data could therefore be examined, and manipulation of miRNA levels could further determine if 3' UTR variants have different susceptibility to miRNA action.

Within the existing 3' sequencing dataset, there is a considerable amount of data that remains unexplored. For instance, the Orphan peaks that were identified may correspond to novel protein coding transcripts or non-coding RNAs such as lincRNAs. This data could be investigated for aspects of gene transcription such as promoter elements or cross-referenced with existing RNA seq data to determine if we have detected novel transcripts. We also detected polyadenylation of lincRNAs and miRNAs. The polyadenylated miRNAs correspond to primary miRNA transcripts and could be correlated with the mature miRNA levels to gain insight into the regulation of miRNA processing. Lastly, the peaks that were upstream of the stop codon or within intronic regions were not analysed in this thesis. These peaks could correspond to the 3' UTRs of splice variants, which would also contain miRNA binding sites. The data could be reinterrogated taking into account all of the

transcript variants of a genomic loci encompassing the exon and intron peaks. In summary, both the miRNA dataset and the mRNA 3' end data offer a variety of possibilities for future cardiac biology research.

7.3 Concluding remarks

The overall aim of this thesis was to profile miRNAs and mRNA 3' ends in an established model of LV hypertrophy and then to interrogate this data for evidence of miRNA processing variants and mRNA alternative polyadenylation. These aims have been achieved and this is the first study to document both the miRNA isoform population and mRNA 3' UTRs in purified cardiomyocytes simultaneously.

Chapter Eight

Appendix

Due to the large size, the appendix files have been included digitally. Please refer to the attached CD.

8.1. Scripts for miRNA mapping

The following files used for the mapping and counting of miRNA sequencing are included:

- Adaptortrim.pl
- miRNA.bowtie.SGE.pl
- countGff
- miRNAcountsBAM.pl
- miRNAcounts.mysql.pl
- mmu.features

8.2. R script for miRNA differential expression analysis

The details for miRNA normalisation and differential expression are contained in the file in the R script file 8.1.

8.3. Scripts for 3' end sequencing mapping

The following files used for the mapping and counting of 3' end sequencing are included:

- Csqvfilter.pl
- Cspolyaextender.pl
- linkCoveragePeaks.pl
- getTranscripts.pl
- processTable.pl

8.4. R script for merging 3' seq peaks and differential expression

The details for merging 3' seq peaks associated with the same gene ID, mRNA normalisation and differential expression are contained in the file in the R script file 8.4.

8.5. GO-term enrichment for differentially expressed mRNA transcripts

The perl script 8.5.pl was used to perform GO-term enrichment of the differentially expressed mRNAs.

8.6. miRNA differential expression analysis

The excel file 8.6 contains the details for miRNA differential expression analysis, including raw counts, normalisation factors, normalised counts per million, log2-FC, spearman correlations and the miRNAs with a FDR < 30% for each pairwise comparison.

8.7. miRNA differential expression analysis using upper quartile normalisation

The excel file 8.7 contains the details for miRNA differential expression analysis after upper quartile normalisation, including raw counts, normalisation factors, normalised counts per million, log2-FC, spearman correlations and the miRNAs with a FDR < 30% for each pairwise comparison.

8.8. Processing variation of miRNAs

The excel file 8.8 contains the proportion of arm bias, 5' isomiRs, 3' isomiRs, non-templated editions and internal editing for all miRNAs that have a CPM greater than 10.

8.9. Weighted linear model for miRNA processing changes

The excel file 8.9 contains the details of the statistical model applied for all miRNA processing variants (arm bias, 5' and 3' isomiRs, non-templated additions and internal editing). The number of miRNAs with varying FDR cut-offs are listed, and if a miRNA was identified as changing proportion then further details are provided in the tabs labelled "list".

8.10. Conventional sequencing of 3'UTRs after 3' race

The conventional sequencing of mRNA 3' ends validated using 3' race are listed. The cloning site, primer sequences and the region that matches the expected genomic sequence are highlighted.

8.11. mRNA differential expression

The excel file 8.11 contains the details for mRNA differential expression analysis, including raw counts, normalisation factors, normalised counts per million, log2-FC, spearman correlations and the mRNAs with a FDR < 30% for each pairwise comparison.

8.12. GO-term enrichment for differentially expressed mRNAs

The enriched biological processes, molecular functions, cellular compartments and KEGG pathways with FDRs <30% are listed for all differentially expressed mRNAs as well as the mRNAs that are novel to this study.

8.13. Weighted linear model for changes in 3' UTR usage

The excel file 8.13 contains the details of the statistical model applied for changes in the proportion of 3' seq peaks within an mRNA. The 3'UTRs that increase or decrease in length at 48 hours and 7 days are provided.

8.14. Conventional sequencing of miR-133a luciferase constructs

The conventional sequencing for psiCHECK™-2 plasmids containing CTGF, CTGF mut, PGAM1 and PGAM1 mut sites in one, two and three copies are listed. The miRNA binding sites and restriction enzyme sites are highlighted.

8.15. Expression of snoRNAs in cardiomyocyte samples

The miRNA expression measured using RT-qPCR was normalised to the average expression of five small nucleolar RNAs (snoRNAs); snoRNA-135, snoRNA-202,

snoRNA-55, snoRNA-412 and snoRNA-234. The expression of each snoRNA across the 20 cardiomyocyte samples is shown in Figure 8.1.

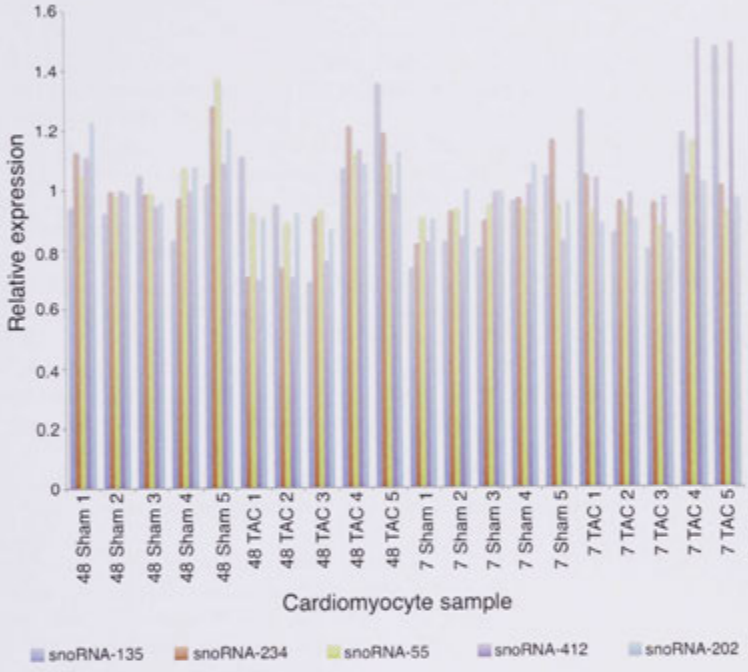


Figure 8.1. Expression of snoRNAs in cardiomyocytes.

Relative expression values of five snoRNAs (snoRNA-135, snoRNA-234, snoRNA-55, snoRNA-412 and snoRNA-202) measured using RT-qPCR on the 20 cardiomyocyte samples. Data shown as relative expression for each individual replicate.

References

1. Welfare A.I.o.H.a. (2011) Cardiovascular disease: Australian facts 2011. vol. Cardiovascular disease series Canberra: AIHW.
2. Heineke J., and Molkentin J.D. (2006) Regulation of cardiac hypertrophy by intracellular signalling pathways. *Nat Rev Mol Cell Biol* **7**, 589-600.
3. Lorell B.H., and Carabello B.A. (2000) Left ventricular hypertrophy: pathogenesis, detection, and prognosis. *Circulation* **102**, 470-479.
4. Rohini A., Agrawal N., Koyani C.N., and Singh R. (2010) Molecular targets and regulators of cardiac hypertrophy. *Pharmacol Res* **61**, 269-280.
5. Passier R., Zeng H., Frey N., Naya F.J., Nicol R.L., et al. (2000) CaM kinase signaling induces cardiac hypertrophy and activates the MEF2 transcription factor in vivo. *J Clin Invest* **105**, 1395-1406.
6. Rohini A., Agrawal N., Koyani C.N., and Singh R. (2010) Molecular targets and regulators of cardiac hypertrophy. *Pharmacol Res* **61**, 269-280.
7. Loh C., Shaw K.T., Carew J., Viola J.P., Luo C., et al. (1996) Calcineurin binds the transcription factor NFAT1 and reversibly regulates its activity. *J Biol Chem* **271**, 10884-10891.
8. Hogan P.G., Chen L., Nardone J., and Rao A. (2003) Transcriptional regulation by calcium, calcineurin, and NFAT. *Genes Dev* **17**, 2205-2232.
9. Klee C.B., Ren H., and Wang X. (1998) Regulation of the calmodulin-stimulated protein phosphatase, calcineurin. *J Biol Chem* **273**, 13367-13370.
10. Molkentin J.D., Lu J.R., Antos C.L., Markham B., Richardson J., et al. (1998) A calcineurin-dependent transcriptional pathway for cardiac hypertrophy. *Cell* **93**, 215-228.
11. McKinsey T.A., Zhang C.L., Lu J., and Olson E.N. (2000) Signal-dependent nuclear export of a histone deacetylase regulates muscle differentiation. *Nature* **408**, 106-111.
12. McKinsey T.A., Zhang C.L., and Olson E.N. (2002) MEF2: a calcium-dependent regulator of cell division, differentiation and death. *Trends Biochem Sci* **27**, 40-47.
13. Hunt S.A., Abraham W.T., Chin M.H., Feldman A.M., Francis G.S., et al. (2005) ACC/AHA 2005 Guideline Update for the Diagnosis and Management of Chronic Heart Failure in the Adult: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Writing Committee to Update the 2001 Guidelines for the Evaluation and Management of Heart Failure): developed in collaboration with the American College of Chest Physicians and the International Society for Heart and Lung Transplantation: endorsed by the Heart Rhythm Society. *Circulation* **112**, e154-235.
14. Tang C.M., and Insel P.A. (2004) GPCR expression in the heart; "new" receptors in myocytes and fibroblasts. *Trends Cardiovasc Med* **14**, 94-99.
15. Salazar N.C., Chen J., and Rockman H.A. (2007) Cardiac GPCRs: GPCR signaling in healthy and failing hearts. *Biochim Biophys Acta* **1768**, 1006-1018.
16. Ding B., Price R.L., Borg T.K., Weinberg E.O., Halloran P.F., et al. (1999) Pressure overload induces severe hypertrophy in mice treated with cyclosporine, an inhibitor of calcineurin. *Circ Res* **84**, 729-734.
17. Rockman H.A., Koch W.J., and Lefkowitz R.J. (2002) Seven-transmembrane-spanning receptors and heart function. *Nature* **415**, 206-212.
18. Neves S.R., Ram P.T., and Iyengar R. (2002) G protein pathways. *Science* **296**, 1636-1639.

19. Hubbard K.B.. and Hepler J.R. (2006) Cell signalling diversity of the Gqalpha family of heterotrimeric G proteins. *Cell Signal* **18**, 135-150.
20. Vega R.B., Harrison B.C., Meadows E., Roberts C.R., Papst P.J., et al. (2004) Protein kinases C and D mediate agonist-dependent cardiac hypertrophy through nuclear export of histone deacetylase 5. *Mol Cell Biol* **24**, 8374-8385.
21. Harrison B.C., Kim M.S., van Rooij E., Plato C.F., Papst P.J., et al. (2006) Regulation of cardiac stress signaling by protein kinase d1. *Mol Cell Biol* **26**, 3875-3888.
22. Wennerberg K., Rossman K.L., and Der C.J. (2005) The Ras superfamily at a glance. *J Cell Sci* **118**, 843-846.
23. Molkentin J.D.. and Dorn G.W., 2nd (2001) Cytoplasmic signaling pathways that regulate cardiac hypertrophy. *Annu Rev Physiol* **63**, 391-426.
24. Lezoualc'h F., Metrich M., Hmitou I., Duquesnes N., and Morel E. (2008) Small GTP-binding proteins and their regulators in cardiac hypertrophy. *J Mol Cell Cardiol* **44**, 623-632.
25. Kawamura S., Miyamoto S., and Brown J.H. (2003) Initiation and transduction of stretch-induced RhoA and Rac1 activation through caveolae: cytoskeletal regulation of ERK translocation. *J Biol Chem* **278**, 31111-31117.
26. Long C.S., Henrich C.J., and Simpson P.C. (1991) A growth factor for cardiac myocytes is produced by cardiac nonmyocytes. *Cell Regul* **2**, 1081-1095.
27. Ren J., Samson W.K., and Sowers J.R. (1999) Insulin-like growth factor I as a cardiac hormone: physiological and pathophysiological implications in heart disease. *J Mol Cell Cardiol* **31**, 2049-2061.
28. Clerk A., Pham F.H., Fuller S.J., Sahai E., Aktories K., et al. (2001) Regulation of mitogen-activated protein kinases in cardiac myocytes through the small G protein Rac1. *Mol Cell Biol* **21**, 1173-1184.
29. Garrington T.P.. and Johnson G.L. (1999) Organization and regulation of mitogen-activated protein kinase signaling pathways. *Curr Opin Cell Biol* **11**, 211-218.
30. Ruwhof C., van Wamel A.E., Egas J.M., and van der Laarse A. (2000) Cyclic stretch induces the release of growth promoting factors from cultured neonatal cardiomyocytes and cardiac fibroblasts. *Mol Cell Biochem* **208**, 89-98.
31. LaMorte V.J., Thorburn J., Absher D., Spiegel A., Brown J.H., et al. (1994) Gq- and ras-dependent pathways mediate hypertrophy of neonatal rat ventricular myocytes following alpha 1-adrenergic stimulation. *J Biol Chem* **269**, 13490-13496.
32. Wang L., Gout I., and Proud C.G. (2001) Cross-talk between the ERK and p70 S6 kinase (S6K) signaling pathways. MEK-dependent activation of S6K2 in cardiomyocytes. *J Biol Chem* **276**, 32670-32677.
33. Cook S.A., Sugden P.H., and Clerk A. (1999) Activation of c-Jun N-terminal kinases and p38-mitogen-activated protein kinases in human heart failure secondary to ischaemic heart disease. *J Mol Cell Cardiol* **31**, 1429-1434.
34. Choukroun G., Hajjar R., Fry S., del Monte F., Haq S., et al. (1999) Regulation of cardiac hypertrophy in vivo by the stress-activated protein kinases/c-Jun NH(2)-terminal kinases. *J Clin Invest* **104**, 391-398.
35. Sugden P.H.. and Clerk A. (1998) "Stress-responsive" mitogen-activated protein kinases (c-Jun N-terminal kinases and p38 mitogen-activated protein kinases) in the myocardium. *Circ Res* **83**, 345-352.
36. Liang F., Atakilit A., and Gardner D.G. (2000) Integrin dependence of brain natriuretic peptide gene promoter activation by mechanical strain. *J Biol Chem* **275**, 20355-20360.
37. Nishizuka Y. (1992) Intracellular signaling by hydrolysis of phospholipids and activation of protein kinase C. *Science* **258**, 607-614.

38. Iglesias T., Waldron R.T., and Rozengurt E. (1998) Identification of in vivo phosphorylation sites required for protein kinase D activation. *J Biol Chem* **273**, 27662-27667.
39. Cantley L.C. (2002) The phosphoinositide 3-kinase pathway. *Science* **296**, 1655-1657.
40. Lawlor M.A., and Alessi D.R. (2001) PKB/Akt: a key mediator of cell proliferation, survival and insulin responses? *J Cell Sci* **114**, 2903-2910.
41. Brunet A., Datta S.R., and Greenberg M.E. (2001) Transcription-dependent and -independent control of neuronal survival by the PI3K-Akt signaling pathway. *Curr Opin Neurobiol* **11**, 297-305.
42. Proud C.G. (2005) eIF2 and the control of cell physiology. *Semin Cell Dev Biol* **16**, 3-12.
43. Dorn G.W., 2nd. and Force T. (2005) Protein kinase cascades in the regulation of cardiac hypertrophy. *J Clin Invest* **115**, 527-537.
44. Gingras A.C., Raught B., and Sonenberg N. (2001) Control of translation by the target of rapamycin proteins. *Prog Mol Subcell Biol* **27**, 143-174.
45. Da Costa Martins P.A., and De Windt L.J. (2012) MicroRNAs in control of cardiac hypertrophy. *Cardiovascular research* **93**, 563-572.
46. Han J., Lee Y., Yeom K.H., Kim Y.K., Jin H., et al. (2004) The Drosha-DGCR8 complex in primary microRNA processing. *Genes Dev* **18**, 3016-3027.
47. Lee Y., Kim M., Han J., Yeom K.H., Lee S., et al. (2004) MicroRNA genes are transcribed by RNA polymerase II. *EMBO J* **23**, 4051-4060.
48. Liu N., Williams A.H., Kim Y., McAnally J., Bezprozvannaya S., et al. (2007) An intragenic MEF2-dependent enhancer directs muscle-specific expression of microRNAs 1 and 133. *Proc Natl Acad Sci U S A* **104**, 20844-20849.
49. Bartel D.P. (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* **116**, 281-297.
50. Gregory R.I., Yan K.P., Amuthan G., Chendrimada T., Doratotaj B., et al. (2004) The Microprocessor complex mediates the genesis of microRNAs. *Nature* **432**, 235-240.
51. Han J., Lee Y., Yeom K.H., Nam J.W., Heo I., et al. (2006) Molecular basis for the recognition of primary microRNAs by the Drosha-DGCR8 complex. *Cell* **125**, 887-901.
52. Yi R., Qin Y., Macara I.G., and Cullen B.R. (2003) Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes Dev* **17**, 3011-3016.
53. Jaskiewicz L., and Filipowicz W. (2008) Role of Dicer in posttranscriptional RNA silencing. *Curr Top Microbiol Immunol* **320**, 77-97.
54. Chendrimada T.P., Gregory R.I., Kumaraswamy E., Norman J., Cooch N., et al. (2005) TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. *Nature* **436**, 740-744.
55. Humphreys D.T., Hynes C.J., Patel H.R., Wei G.H., Cannon L., et al. (2012) Complexity of Murine Cardiomyocyte miRNA Biogenesis, Sequence Variant Expression and Function. *PLoS One* **7**, e30933.
56. Warf M.B., Johnson W.E., and Bass B.L. (2011) Improved annotation of C. elegans microRNAs by deep sequencing reveals structures associated with processing by Drosha and Dicer. *Rna* **17**, 563-577.
57. Starega-Roslan J., Koscianska E., Kozlowski P., and Krzyzosiak W.J. (2011) The role of the precursor structure in the biogenesis of microRNA. *Cell Mol Life Sci* **68**, 2859-2871.
58. Frank F., Sonenberg N., and Nagar B. (2010) Structural basis for 5'-nucleotide base-specific recognition of guide RNA by human AGO2. *Nature* **465**, 818-822.
59. Khvorova A., Reynolds A., and Jayasena S.D. (2003) Functional siRNAs and miRNAs exhibit strand bias. *Cell* **115**, 209-216.
60. Schwarz D.S., Hutvagner G., Du T., Xu Z., Aronin N., et al. (2003) Asymmetry in the assembly of the RNAi enzyme complex. *Cell* **115**, 199-208.

61. Kwak P.B.. and Tomari Y. (2012) The N domain of Argonaute drives duplex unwinding during RISC assembly. *Nat Struct Mol Biol* **19**, 145-151.
62. Burroughs A.M., Ando Y., de Hoon M.J., Tomaru Y., Suzuki H., et al. (2011) Deep-sequencing of human Argonaute-associated small RNAs provides insight into miRNA sorting and reveals Argonaute association with RNA fragments of diverse origin. *RNA Biol* **8**, 158-177.
63. Meister G. (2013) Argonaute proteins: functional insights and emerging roles. *Nat Rev Genet* **14**, 447-459.
64. Liu J., Carmell M.A., Rivas F.V., Marsden C.G., Thomson J.M., et al. (2004) Argonaute2 is the catalytic engine of mammalian RNAi. *Science* **305**, 1437-1441.
65. Beilharz T.H., Humphreys D.T., and Preiss T. (2010) miRNA Effects on mRNA closed-loop formation during translation initiation. *Prog Mol Subcell Biol* **50**, 99-112.
66. Ameres S.L.. and Zamore P.D. (2013) Diversifying microRNA sequence and function. *Nat Rev Mol Cell Biol* **14**, 475-488.
67. Lewis B.P., Burge C.B., and Bartel D.P. (2005) Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* **120**, 15-20.
68. Doench J.G.. and Sharp P.A. (2004) Specificity of microRNA target selection in translational repression. *Genes Dev* **18**, 504-511.
69. Chi S.W., Hannon G.J., and Darnell R.B. (2012) An alternative mode of microRNA target recognition. *Nat Struct Mol Biol* **19**, 321-327.
70. Wen J., Parker B.J., Jacobsen A., and Krogh A. (2011) MicroRNA transfection and AGO-bound CLIP-seq data sets reveal distinct determinants of miRNA action. *Rna* **17**, 820-834.
71. Chi S.W., Zang J.B., Mele A., and Darnell R.B. (2009) Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps. *Nature* **460**, 479-486.
72. Hafner M., Landthaler M., Burger L., Khorshid M., Hausser J., et al. (2010) Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell* **141**, 129-141.
73. Chandra V., Girijadevi R., Nair A.S., Pillai S.S., and Pillai R.M. (2010) MTar: a computational microRNA target prediction architecture for human transcriptome. *BMC Bioinformatics* **11 Suppl 1**, S2.
74. Coronnello C.. and Benos P.V. (2013) ComiR: Combinatorial microRNA target prediction tool. *Nucleic Acids Res* **41**, W159-164.
75. Hammell M., Long D., Zhang L., Lee A., Carmack C.S., et al. (2008) mirWIP: microRNA target prediction based on microRNA-containing ribonucleoprotein-enriched transcripts. *Nature methods* **5**, 813-819.
76. Matkovich S.J., Van Booven D.J., Eschenbacher W.H., and Dorn G.W., 2nd (2011) RISC RNA Sequencing for Context-Specific Identification of In Vivo MicroRNA Targets. *Circ Res* **108**, 18-26.
77. Hynes C.J., Clancy J.L., and Preiss T. (2012) miRNAs in cardiac disease: sitting duck or moving target? *IUBMB life* **64**, 872-878.
78. Hou J., Lin L., Zhou W., Wang Z., Ding G., et al. (2011) Identification of miRNomes in human liver and hepatocellular carcinoma reveals miR-199a/b-3p as therapeutic target for hepatocellular carcinoma. *Cancer Cell* **19**, 232-243.
79. Chiang H.R., Schoenfeld L.W., Ruby J.G., Auyeung V.C., Spies N., et al. (2010) Mammalian microRNAs: experimental evaluation of novel and previously annotated genes. *Genes Dev* **24**, 992-1009.
80. Ro S., Park C., Young D., Sanders K.M., and Yan W. (2007) Tissue-dependent paired expression of miRNAs. *Nucleic Acids Res* **35**, 5944-5953.
81. Czech B.. and Hannon G.J. (2011) Small RNA sorting: matchmaking for Argonautes. *Nat Rev Genet* **12**, 19-31.

82. Kozomara A., and Griffiths-Jones S. (2011) miRBase: integrating microRNA annotation and deep-sequencing data. *Nucleic Acids Res* **39**, D152-157.
83. Fernandez-Valverde S.L., Taft R.J., and Mattick J.S. (2010) Dynamic isomiR regulation in Drosophila development. *Rna* **16**, 1881-1888.
84. Kuchenbauer F., Morin R.D., Argiropoulos B., Petriv O.I., Griffith M., et al. (2008) In-depth characterization of the microRNA transcriptome in a leukemia progression model. *Genome Res* **18**, 1787-1797.
85. Cloonan N., Wani S., Xu Q., Gu J., Lea K., et al. (2011) MicroRNAs and their isomiRs function cooperatively to target common biological pathways. *Genome Biol* **12**, R126.
86. Ebhardt H.A., Fedynak A., and Fahlman R.P. (2010) Naturally occurring variations in sequence length creates microRNA isoforms that differ in argonaute effector complex specificity. *Silence* **1**, 12.
87. Mi S., Cai T., Hu Y., Chen Y., Hodges E., et al. (2008) Sorting of small RNAs into Arabidopsis argonaute complexes is directed by the 5' terminal nucleotide. *Cell* **133**, 116-127.
88. Ruby J.G., Jan C., Player C., Axtell M.J., Lee W., et al. (2006) Large-scale sequencing reveals 21U-RNAs and additional microRNAs and endogenous siRNAs in C. elegans. *Cell* **127**, 1193-1207.
89. Wu H., Neilson J.R., Kumar P., Manocha M., Shankar P., et al. (2007) miRNA profiling of naive, effector and memory CD8 T cells. *PLoS One* **2**, e1020.
90. Han B.W., Hung J.H., Weng Z., Zamore P.D., and Ameres S.L. (2011) The 3'-to-5' exoribonuclease Nibbler shapes the 3' ends of microRNAs bound to Drosophila Argonaute1. *Curr Biol* **21**, 1878-1887.
91. Kim Y.K., Heo I., and Kim V.N. (2010) Modifications of small RNAs and their associated proteins. *Cell* **143**, 703-709.
92. Katoh T., Sakaguchi Y., Miyauchi K., Suzuki T., Kashiwabara S., et al. (2009) Selective stabilization of mammalian microRNAs by 3' adenylation mediated by the cytoplasmic poly(A) polymerase GLD-2. *Genes Dev* **23**, 433-438.
93. Jones M.R., Quinton L.J., Blahna M.T., Neilson J.R., Fu S., et al. (2009) Zcchc11-dependent uridylation of microRNA directs cytokine expression. *Nat Cell Biol* **11**, 1157-1163.
94. Bass B.L. (2002) RNA editing by adenosine deaminases that act on RNA. *Annu Rev Biochem* **71**, 817-846.
95. Kawahara Y., Megraw M., Kreider E., Iizasa H., Valente L., et al. (2008) Frequency and fate of microRNA editing in human brain. *Nucleic Acids Res* **36**, 5270-5280.
96. Kawahara Y., Zinshteyn B., Sethupathy P., Iizasa H., Hatzigeorgiou A.G., et al. (2007) Redirection of silencing targets by adenosine-to-inosine editing of miRNAs. *Science* **315**, 1137-1140.
97. Agranat L., Sperling J., and Sperling R. (2010) A novel tissue-specific alternatively spliced form of the A-to-I RNA editing enzyme ADAR2. *RNA Biol* **7**, 253-262.
98. Ozsolak F., Poling L.L., Wang Z., Liu H., Liu X.S., et al. (2008) Chromatin structure analyses identify miRNA promoters. *Genes Dev* **22**, 3172-3183.
99. Tsai K.W., Liao Y.L., Wu C.W., Hu L.Y., Li S.C., et al. (2011) Aberrant hypermethylation of miR-9 genes in gastric cancer. *Epigenetics* **6**, 1189-1197.
100. Thomson J.M., Newman M., Parker J.S., Morin-Kensicki E.M., Wright T., et al. (2006) Extensive post-transcriptional regulation of microRNAs and its implications for cancer. *Genes Dev* **20**, 2202-2207.
101. Yamagata K., Fujiyama S., Ito S., Ueda T., Murata T., et al. (2009) Maturation of microRNA is hormonally regulated by a nuclear receptor. *Mol Cell* **36**, 340-347.
102. Bates G.J., Nicol S.M., Wilson B.J., Jacobs A.M., Bourdon J.C., et al. (2005) The DEAD box protein p68: a novel transcriptional coactivator of the p53 tumour suppressor. *The EMBO journal* **24**, 543-553.

103. Davis B.N., Hilyard A.C., Lagna G., and Hata A. (2008) SMAD proteins control DROSHA-mediated microRNA maturation. *Nature* **454**, 56-61.
104. Dobaczewski M., Chen W., and Frangogiannis N.G. (2011) Transforming growth factor (TGF)-beta signaling in cardiac remodeling. *J Mol Cell Cardiol* **51**, 600-606.
105. Rau F., Freyermuth F., Fugier C., Villemin J.P., Fischer M.C., et al. (2011) Misregulation of miR-1 processing is associated with heart defects in myotonic dystrophy. *Nat Struct Mol Biol* **18**, 840-845.
106. Michlewski G., and Caceres J.F. (2010) Antagonistic role of hnRNP A1 and KSRP in the regulation of let-7a biogenesis. *Nat Struct Mol Biol* **17**, 1011-1018.
107. Piskounova E., Polytarchou C., Thornton J.E., LaPierre R.J., Pothoulakis C., et al. (2011) Lin28A and Lin28B inhibit let-7 microRNA biogenesis by distinct mechanisms. *Cell* **147**, 1066-1079.
108. Guil S., and Caceres J.F. (2007) The multifunctional RNA-binding protein hnRNP A1 is required for processing of miR-18a. *Nat Struct Mol Biol* **14**, 591-596.
109. van Almen G.C., Verhesen W., van Leeuwen R.E., van de Vrie M., Eurlings C., et al. (2011) MicroRNA-18 and microRNA-19 regulate CTGF and TSP-1 expression in age-related heart failure. *Aging Cell* **10**, 769-779.
110. Lai E.C. (2005) miRNAs: whys and wherefores of miRNA-mediated regulation. *Curr Biol* **15**, R458-460.
111. Rao P.K., Toyama Y., Chiang H.R., Gupta S., Bauer M., et al. (2009) Loss of cardiac microRNA-mediated regulation leads to dilated cardiomyopathy and heart failure. *Circ Res* **105**, 585-594.
112. Zhao Y., Samal E., and Srivastava D. (2005) Serum response factor regulates a muscle-specific microRNA that targets Hand2 during cardiogenesis. *Nature* **436**, 214-220.
113. McCarthy J.J. (2008) MicroRNA-206: the skeletal muscle-specific myomiR. *Biochim Biophys Acta* **1779**, 682-691.
114. van Rooij E., Sutherland L.B., Qi X., Richardson J.A., Hill J., et al. (2007) Control of stress-dependent cardiac growth and gene expression by a microRNA. *Science* **316**, 575-579.
115. Callis T.E., Pandya K., Seok H.Y., Tang R.H., Tatsuguchi M., et al. (2009) MicroRNA-208a is a regulator of cardiac hypertrophy and conduction in mice. *J Clin Invest* **119**, 2772-2786.
116. Ding S.L., Zhou L.Y., and Li P.F. (2011) MicroRNAs in cardiac hypertrophy: angels or devils. *Wiley Interdiscip Rev RNA* **2**, 124-134.
117. Latronico M.V., and Condorelli G. (2011) microRNAs in hypertrophy and heart failure. *Exp Biol Med (Maywood)* **236**, 125-131.
118. Gladka M.M., da Costa Martins P.A., and De Windt L.J. (2012) Small changes can make a big difference - microRNA regulation of cardiac hypertrophy. *J Mol Cell Cardiol* **52**, 74-82.
119. Small E.M., and Olson E.N. (2011) Pervasive roles of microRNAs in cardiovascular biology. *Nature* **469**, 336-342.
120. van Rooij E., Sutherland L.B., Liu N., Williams A.H., McAnally J., et al. (2006) A signature pattern of stress-responsive microRNAs that can evoke cardiac hypertrophy and heart failure. *Proc Natl Acad Sci U S A* **103**, 18255-18260.
121. Tatsuguchi M., Seok H.Y., Callis T.E., Thomson J.M., Chen J.F., et al. (2007) Expression of microRNAs is dynamically regulated during cardiomyocyte hypertrophy. *J Mol Cell Cardiol* **42**, 1137-1141.
122. Thum T., Galuppo P., Wolf C., Fiedler J., Kneitz S., et al. (2007) MicroRNAs in the human heart: a clue to fetal gene reprogramming in heart failure. *Circulation* **116**, 258-267.
123. Sucharov C., Bristow M.R., and Port J.D. (2008) miRNA expression in the failing human heart: functional correlates. *J Mol Cell Cardiol* **45**, 185-192.

124. Ikeda S., Kong S.W., Lu J., Bisping E., Zhang H., et al. (2007) Altered microRNA expression in human heart disease. *Physiol Genomics* **31**, 367-373.
125. Dec G.W., and Fuster V. (1994) Idiopathic dilated cardiomyopathy. *N Engl J Med* **331**, 1564-1575.
126. Alla F., Briancon S., Juilliere Y., Mertes P.M., Villemot J.P., et al. (2000) Differential clinical prognostic classifications in dilated and ischemic advanced heart failure: the EPICAL study. *Am Heart J* **139**, 895-904.
127. Sayed D., Hong C., Chen I.Y., Lypowy J., and Abdellatif M. (2007) MicroRNAs play an essential role in the development of cardiac hypertrophy. *Circ Res* **100**, 416-424.
128. Cheng Y., Ji R., Yue J., Yang J., Liu X., et al. (2007) MicroRNAs are aberrantly expressed in hypertrophic heart: do they play a role in cardiac hypertrophy? *The American journal of pathology* **170**, 1831-1840.
129. Bagnall R.D., Tsoutsman T., Shephard R.E., Ritchie W., and Semsarian C. (2012) Global microRNA profiling of the mouse ventricles during development of severe hypertrophic cardiomyopathy and heart failure. *PLoS One* **7**, e44744.
130. Yang K.C., Ku Y.C., Lovett M., and Nerbonne J.M. (2012) Combined deep microRNA and mRNA sequencing identifies protective transcriptomal signature of enhanced PI3Kalpha signaling in cardiac hypertrophy. *J Mol Cell Cardiol* **53**, 101-112.
131. Hu Y., Matkovich S.J., Hecker P.A., Zhang Y., Edwards J.R., et al. (2012) Epitranscriptional orchestration of genetic reprogramming is an emergent property of stress-regulated cardiac microRNAs. *Proc Natl Acad Sci U S A* **109**, 19864-19869.
132. Lin D., Hollander Z., Meredith A., Stadnick E., Sasaki M., et al. (2011) Molecular signatures of end-stage heart failure. *J Card Fail* **17**, 867-874.
133. Towbin J.A., and Bowles N.E. (2001) Molecular genetics of left ventricular dysfunction. *Curr Mol Med* **1**, 81-90.
134. Porrello E.R., Johnson B.A., Aurora A.B., Simpson E., Nam Y.J., et al. (2011) MiR-15 family regulates postnatal mitotic arrest of cardiomyocytes. *Circ Res* **109**, 670-679.
135. van Rooij E., and Olson E.N. (2007) MicroRNAs: powerful new regulators of heart disease and provocative therapeutic targets. *J Clin Invest* **117**, 2369-2376.
136. McCrohon J.A., Moon J.C., Prasad S.K., McKenna W.J., Lorenz C.H., et al. (2003) Differentiation of heart failure related to dilated cardiomyopathy and coronary artery disease using gadolinium-enhanced cardiovascular magnetic resonance. *Circulation* **108**, 54-59.
137. Ai J., Zhang R., Li Y., Pu J., Lu Y., et al. (2010) Circulating microRNA-1 as a potential novel biomarker for acute myocardial infarction. *Biochem Biophys Res Commun* **391**, 73-77.
138. Ji X., Takahashi R., Hiura Y., Hirokawa G., Fukushima Y., et al. (2009) Plasma miR-208 as a biomarker of myocardial injury. *Clin Chem* **55**, 1944-1949.
139. Tijsen A.J., Creemers E.E., Moerland P.D., de Windt L.J., van der Wal A.C., et al. (2010) MiR423-5p as a circulating biomarker for heart failure. *Circ Res* **106**, 1035-1039.
140. Care A., Catalucci D., Felicetti F., Bonci D., Addario A., et al. (2007) MicroRNA-133 controls cardiac hypertrophy. *Nat Med* **13**, 613-618.
141. Wang K., Long B., Zhou J., and Li P.F. (2010) miR-9 and NFATc3 regulate myocardin in cardiac hypertrophy. *J Biol Chem* **285**, 11903-11912.
142. Xing W., Zhang T.C., Cao D., Wang Z., Antos C.L., et al. (2006) Myocardin induces cardiomyocyte hypertrophy. *Circ Res* **98**, 1089-1097.
143. Morissette M., Cook S., Foo S., McKoy G., Ashida N., et al. (2006) Myostatin regulates cardiomyocyte growth through modulation of Akt signaling. *Circ Res* **99**, 15-39.

144. Wang J., Song Y., Zhang Y., Xiao H., Sun Q., et al. (2011) Cardiomyocyte overexpression of miR-27b induces cardiac hypertrophy and dysfunction in mice. *Cell Res.*

145. Brown C.E.. and Sachs A.B. (1998) Poly(A) tail length control in Saccharomyces cerevisiae occurs by message-specific deadenylation. *Mol Cell Biol* **18**, 6548-6559.

146. Zhao J., Hyman L., and Moore C. (1999) Formation of mRNA 3' ends in eukaryotes: mechanism, regulation, and interrelationships with other steps in mRNA synthesis. *Microbiol Mol Biol Rev* **63**, 405-445.

147. Shi Y., Di Giammartino D.C., Taylor D., Sarkeshik A., Rice W.J., et al. (2009) Molecular architecture of the human pre-mRNA 3' processing complex. *Mol Cell* **33**, 365-376.

148. Danckwardt S., Hentze M.W., and Kulozik A.E. (2008) 3' end mRNA processing: molecular mechanisms and implications for health and disease. *The EMBO journal* **27**, 482-498.

149. Murthy K.G.. and Manley J.L. (1995) The 160-kD subunit of human cleavage-polyadenylation specificity factor coordinates pre-mRNA 3'-end formation. *Genes Dev* **9**, 2672-2683.

150. Ryan K., Calvo O., and Manley J.L. (2004) Evidence that polyadenylation factor CPSF-73 is the mRNA 3' processing endonuclease. *Rna* **10**, 565-573.

151. Kaufmann I., Martin G., Friedlein A., Langen H., and Keller W. (2004) Human Fip1 is a subunit of CPSF that binds to U-rich RNA elements and stimulates poly(A) polymerase. *The EMBO journal* **23**, 616-626.

152. Barabino S.M., Hubner W., Jenny A., Minvielle-Sebastia L., and Keller W. (1997) The 30-kD subunit of mammalian cleavage and polyadenylation specificity factor and its yeast homolog are RNA-binding zinc finger proteins. *Genes Dev* **11**, 1703-1716.

153. Perez Canadillas J.M.. and Varani G. (2003) Recognition of GU-rich polyadenylation regulatory elements by human CstF-64 protein. *The EMBO journal* **22**, 2821-2830.

154. Venkataraman K., Brown K.M., and Gilmartin G.M. (2005) Analysis of a noncanonical poly(A) site reveals a tripartite mechanism for vertebrate poly(A) site recognition. *Genes Dev* **19**, 1315-1327.

155. Millevoi S., Decorsiere A., Loulergue C., Iacovoni J., Bernat S., et al. (2009) A physical and functional link between splicing factors promotes pre-mRNA 3' end processing. *Nucleic Acids Res* **37**, 4672-4683.

156. Arhin G.K., Boots M., Bagga P.S., Milcarek C., and Wilusz J. (2002) Downstream sequence elements with different affinities for the hnRNP H/H' protein influence the processing efficiency of mammalian polyadenylation signals. *Nucleic Acids Res* **30**, 1842-1850.

157. Chen F.. and Wilusz J. (1998) Auxiliary downstream elements are required for efficient polyadenylation of mammalian pre-mRNAs. *Nucleic Acids Res* **26**, 2891-2898.

158. Millevoi S.. and Vagner S. (2010) Molecular mechanisms of eukaryotic pre-mRNA 3' end processing regulation. *Nucleic Acids Res* **38**, 2757-2774.

159. Tian B.. and Graber J.H. (2012) Signals for pre-mRNA cleavage and polyadenylation. *Wiley Interdiscip Rev RNA* **3**, 385-396.

160. Tian B., Hu J., Zhang H., and Lutz C. (2005) A large-scale analysis of mRNA polyadenylation of human and mouse genes. *Nucleic Acids Res* **33**, 201-212.

161. Ji Z., Lee J.Y., Pan Z., Jiang B., and Tian B. (2009) Progressive lengthening of 3' untranslated regions of mRNAs by alternative polyadenylation during mouse embryonic development. *Proc Natl Acad Sci U S A* **106**, 7028-7033.

162. Legendre M., Ritchie W., Lopez F., and Gautheret D. (2006) Differential repression of alternative transcripts: a screen for miRNA targets. *PLoS computational biology* **2**.

163. Ji Z., and Tian B. (2009) Reprogramming of 3' untranslated regions of mRNAs by alternative polyadenylation in generation of pluripotent stem cells from different cell types. *PLoS One* **4**.
164. Mayr C., and Bartel D. (2009) Widespread shortening of 3'UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells. *Cell* **138**, 673-684.
165. Liu D., Brockman J.M., Dass B., Hutchins L.N., Singh P., et al. (2007) Systematic variation in mRNA 3'-processing signals during mouse spermatogenesis. *Nucleic Acids Res* **35**, 234-246.
166. Zhang H., Lee J.Y., and Tian B. (2005) Biased alternative polyadenylation in human tissues. *Genome Biol* **6**, R100.
167. Brockman J., Singh P., Liu D., Quinlan S., Salisbury J., et al. (2005) PACdb: PolyA Cleavage Site and 3'-UTR Database. *Bioinformatics (Oxford, England)* **21**, 3691-3693.
168. Park J., Li W., Zheng D., Zhai P., Zhao Y., et al. (2011) Comparative analysis of mRNA isoform expression in cardiac hypertrophy and development reveals multiple post-transcriptional regulatory modules. *PLoS One* **6**.
169. Wang D., Oparil S., Feng J.A., Li P., Perry G., et al. (2003) Effects of pressure overload on extracellular matrix expression in the heart of the atrial natriuretic peptide-null mouse. *Hypertension* **42**, 88-95.
170. Clancy J., Wei G., Echner N., Humphreys D., Beilharz T., et al. (2011) mRNA isoform diversity can obscure detection of miRNA-mediated control of translation. *RNA (New York, N.Y.)* **17**, 1025-1031.
171. Ulitsky I., Shkumatava A., Jan C., Subtelny A., Koppstein D., et al. (2012) Extensive alternative polyadenylation during zebrafish development. *Genome Res* **22**, 2054-2066.
172. Vickers T.A., Wyatt J.R., Burckin T., Bennett C.F., and Freier S.M. (2001) Fully modified 2' MOE oligonucleotides redirect polyadenylation. *Nucleic Acids Res* **29**, 1293-1299.
173. Djikeng A., Halpin R., Kuzmickas R., Depasse J., Feldblyum J., et al. (2008) Viral genome sequencing by random priming methods. *BMC genomics* **9**, 5.
174. Lin F., Owens W.A., Chen S., Stevens M.E., Kesteven S., et al. (2001) Targeted alpha(1A)-adrenergic receptor overexpression induces enhanced cardiac contractility but not hypertrophy. *Circ Res* **89**, 343-350.
175. Georgakopoulos D., Mitzner W.A., Chen C.H., Byrne B.J., Millar H.D., et al. (1998) In vivo murine left ventricular pressure-volume relations by miniaturized conductance micromanometry. *The American journal of physiology* **274**, H1416-1422.
176. O'Connell T.D., Rodrigo M.C., and Simpson P.C. (2007) Isolation and culture of adult mouse cardiac myocytes. *Methods in molecular biology* **357**, 271-296.
177. Janicke A., Vancuylenberg J., Boag P.R., Traven A., and Beilharz T.H. (2012) ePAT: a simple method to tag adenylated RNA to measure poly(A)-tail length and other 3' RACE applications. *Rna* **18**, 1289-1295.
178. Langmead B., Trapnell C., Pop M., and Salzberg S.L. (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10**, R25.
179. Sai Lakshmi S., and Agrawal S. (2008) piRNABank: a web resource on classified and clustered Piwi-interacting RNAs. *Nucleic Acids Res* **36**, D173-177.
180. Rosenkranz D., and Zischler H. (2012) proTRAC--a software for probabilistic piRNA cluster detection, visualization and analysis. *BMC Bioinformatics* **13**, 5.
181. Barnett D.W., Garrison E.K., Quinlan A.R., Stromberg M.P., and Marth G.T. (2011) BamTools: a C++ API and toolkit for analyzing and managing BAM files. *Bioinformatics* **27**, 1691-1692.

182. Robinson M.D., McCarthy D.J., and Smyth G.K. (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139-140.
183. Robinson M.D.. and Oshlack A. (2010) A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* **11**, R25.
184. Smyth G.K. (2005) Limma: Linear Models for Microarray Data. In: Bioinformatics and Computational Biology Solutions using R and Bioconductor (Gentleman, R., Carey, V., Dudoit, S., Izarry, R., and Huber, W., eds), pp 397-420 New York: Springer.
185. Dennis G., Jr., Sherman B.T., Hosack D.A., Yang J., Gao W., et al. (2003) DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol* **4**, P3.
186. Rockman H.A., Ross R.S., Harris A.N., Knowlton K.U., Steinhelper M.E., et al. (1991) Segregation of atrial-specific and inducible expression of an atrial natriuretic factor transgene in an in vivo murine model of cardiac hypertrophy. *Proc Natl Acad Sci U S A* **88**, 8277-8281.
187. Zhao M., Chow A., Powers J., Fajardo G., and Bernstein D. (2004) Microarray analysis of gene expression after transverse aortic constriction in mice. *Physiol Genomics* **19**, 93-105.
188. van Nierop B.J., van Assen H.C., van Deel E.D., Niesen L.B., Duncker D.J., et al. (2013) Phenotyping of left and right ventricular function in mouse models of compensated hypertrophy and heart failure with cardiac MRI. *PLoS One* **8**, e55424.
189. Nakamura A., Rokosh D.G., Paccanaro M., Yee R.R., Simpson P.C., et al. (2001) LV systolic performance improves with development of hypertrophy after transverse aortic constriction in mice. *American journal of physiology. Heart and circulatory physiology* **281**, H1104-1112.
190. Lee J.H., Gao C., Peng G., Greer C., Ren S., et al. (2011) Analysis of transcriptome complexity through RNA sequencing in normal and failing murine hearts. *Circ Res* **109**, 1332-1341.
191. Camelliti P., Borg T.K., and Kohl P. (2005) Structural and functional characterisation of cardiac fibroblasts. *Cardiovascular research* **65**, 40-51.
192. Porter K.E.. and Turner N.A. (2009) Cardiac fibroblasts: at the heart of myocardial remodeling. *Pharmacology & therapeutics* **123**, 255-278.
193. Gao S., Ho D., Vatner D.E., and Vatner S.F. (2011) Echocardiography in Mice. *Current protocols in mouse biology* **1**, 71-83.
194. Knowles J.W., Esposito G., Mao L., Hagaman J.R., Fox J.E., et al. (2001) Pressure-independent enhancement of cardiac hypertrophy in natriuretic peptide receptor A-deficient mice. *J Clin Invest* **107**, 975-984.
195. Deckmann A.C., Theizen T.H., Medrano F.J., Franchini K.G., and Pereira G.A. (2010) Immediate response of myocardium to pressure overload includes transient regulation of genes associated with mitochondrial bioenergetics and calcium availability. *Genetics and molecular biology* **33**, 12-16.
196. Mielniczuk L.M., Lamas G.A., Flaker G.C., Mitchell G., Smith S.C., et al. (2007) Left ventricular end-diastolic pressure and risk of subsequent heart failure in patients following an acute myocardial infarction. *Congestive heart failure* **13**, 209-214.
197. Bargiggia G.S., Bertucci C., Recusani F., Raisaro A., de Servi S., et al. (1989) A new method for estimating left ventricular dP/dt by continuous wave Doppler-echocardiography. Validation studies at cardiac catheterization. *Circulation* **80**, 1287-1292.
198. Passier R., Zeng H., Frey N., Naya F.J., Nicol R.L., et al. (2000) CaM kinase signaling induces cardiac hypertrophy and activates the MEF2 transcription factor in vivo. *J Clin Invest* **105**, 1395-1406.

199. Meerson F.Z. (1969) The myocardium in hyperfunction, hypertrophy and heart failure. *Circ Res* **25**, Suppl 2:1-163.
200. Hein S., Arnon E., Kostin S., Schonburg M., Elsasser A., et al. (2003) Progression from compensated hypertrophy to failure in the pressure-overloaded human heart: structural deterioration and compensatory mechanisms. *Circulation* **107**, 984-991.
201. Nishikimi T., Maeda N., and Matsuoka H. (2006) The role of natriuretic peptides in cardioprotection. *Cardiovascular research* **69**, 318-328.
202. Bubikat A., De Windt L.J., Zetsche B., Fabritz L., Sickler H., et al. (2005) Local atrial natriuretic peptide signaling prevents hypertensive cardiac hypertrophy in endothelial nitric-oxide synthase-deficient mice. *J Biol Chem* **280**, 21594-21599.
203. Tokudome T., Horio T., Kishimoto I., Soeki T., Mori K., et al. (2005) Calcineurin-nuclear factor of activated T cells pathway-dependent cardiac remodeling in mice deficient in guanylyl cyclase A, a receptor for atrial and brain natriuretic peptides. *Circulation* **111**, 3095-3104.
204. Morita E., Yasue H., Yoshimura M., Ogawa H., Jougasaki M., et al. (1993) Increased plasma levels of brain natriuretic peptide in patients with acute myocardial infarction. *Circulation* **88**, 82-91.
205. Nakagawa O., Ogawa Y., Itoh H., Suga S., Komatsu Y., et al. (1995) Rapid transcriptional activation and early mRNA turnover of brain natriuretic peptide in cardiocyte hypertrophy. Evidence for brain natriuretic peptide as an "emergency" cardiac hormone against ventricular overload. *J Clin Invest* **96**, 1280-1287.
206. Klinge R., Hystad M., Kjekshus J., Karlberg B.E., Djoseland O., et al. (1998) An experimental study of cardiac natriuretic peptides as markers of development of congestive heart failure. *Scandinavian journal of clinical and laboratory investigation* **58**, 683-691.
207. Langenickel T., Pagel I., Hohnel K., Dietz R., and Willenbrock R. (2000) Differential regulation of cardiac ANP and BNP mRNA in different stages of experimental heart failure. *American journal of physiology. Heart and circulatory physiology* **278**, H1500-1506.
208. Izumo S., Nadal-Ginard B., and Mahdavi V. (1988) Protooncogene induction and reprogramming of cardiac gene expression produced by pressure overload. *Proc Natl Acad Sci U S A* **85**, 339-343.
209. Tardiff J.C., Hewett T.E., Factor S.M., Vikstrom K.L., Robbins J., et al. (2000) Expression of the beta (slow)-isoform of MHC in the adult mouse heart causes dominant-negative functional effects. *American journal of physiology. Heart and circulatory physiology* **278**, H412-419.
210. Alpert N.R.. and Mulieri L.A. (1982) Increased myothermal economy of isometric force generation in compensated cardiac hypertrophy induced by pulmonary artery constriction in the rabbit. A characterization of heat liberation in normal and hypertrophied right ventricular papillary muscles. *Circ Res* **50**, 491-500.
211. Krenz M.. and Robbins J. (2004) Impact of beta-myosin heavy chain expression on cardiac function during stress. *Journal of the American College of Cardiology* **44**, 2390-2397.
212. Pope B., Hoh J.F., and Weeds A. (1980) The ATPase activities of rat cardiac myosin isoenzymes. *FEBS Lett* **118**, 205-208.
213. VanBuren P., Harris D.E., Alpert N.R., and Warshaw D.M. (1995) Cardiac V1 and V3 myosins differ in their hydrolytic and mechanical activities in vitro. *Circ Res* **77**, 439-444.
214. Stilli D., Bocchi L., Berni R., Zaniboni M., Cacciani F., et al. (2006) Correlation of alpha-skeletal actin expression, ventricular fibrosis and heart function with the degree of pressure overload cardiac hypertrophy in rats. *Experimental physiology* **91**, 571-580.

215. Gantier M.P., McCoy C.E., Rusinova I., Saulep D., Wang D., et al. (2011) Analysis of microRNA turnover in mammalian cells following Dicer1 ablation. *Nucleic Acids Res* **39**, 5692-5703.
216. Kuhn D.E., Roy S., Radtke J., Khanna S., and Sen C.K. (2007) Laser microdissection and capture of pure cardiomyocytes and fibroblasts from infarcted heart regions: perceived hyperoxia induces p21 in peri-infarct myocytes. *American journal of physiology. Heart and circulatory physiology* **292**, H1245-1253.
217. Lian X., Hsiao C., Wilson G., Zhu K., Hazeltine L.B., et al. (2012) Robust cardiomyocyte differentiation from human pluripotent stem cells via temporal modulation of canonical Wnt signaling. *Proc Natl Acad Sci U S A* **109**, E1848-1857.
218. McCormick K.P., Willmann M.R., and Meyers B.C. (2011) Experimental design, preprocessing, normalization and differential expression analysis of small RNA sequencing experiments. *Silence* **2**, 2.
219. Robinson M.D.. and Oshlack A. (2010) A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol* **11**, R25.
220. Bullard J.H., Purdom E., Hansen K.D., and Dudoit S. (2010) Evaluation of statistical methods for normalization and differential expression in mRNA-Seq experiments. *BMC Bioinformatics* **11**, 94.
221. Farazi T.A., Horlings H.M., Ten Hoeve J.J., Mihailovic A., Halfwerk H., et al. (2011) MicroRNA sequence and expression analysis in breast tumors by deep sequencing. *Cancer research* **71**, 4443-4453.
222. Hamfjord J., Stangeland A.M., Hughes T., Skrede M.L., Tveit K.M., et al. (2012) Differential expression of miRNAs in colorectal cancer: comparison of paired tumor tissue and adjacent normal mucosa using high-throughput sequencing. *PLoS One* **7**, e34150.
223. Juhila J., Sipila T., Icay K., Nicorici D., Ellonen P., et al. (2011) MicroRNA expression profiling reveals miRNA families regulating specific biological pathways in mouse frontal cortex and hippocampus. *PLoS One* **6**, e21495.
224. Benjamini Y.. and Hochberg Y. (1995) Controlling the False Discovery Rate: a Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B* **57**, 289-300.
225. Bernardo B.C., Gao X.M., Winbanks C.E., Boey E.J., Tham Y.K., et al. (2012) Therapeutic inhibition of the miR-34 family attenuates pathological cardiac remodeling and improves heart function. *Proc Natl Acad Sci U S A* **109**, 17615-17620.
226. Thum T., Gross C., Fiedler J., Fischer T., Kissler S., et al. (2008) MicroRNA-21 contributes to myocardial disease by stimulating MAP kinase signalling in fibroblasts. *Nature* **456**, 980-984.
227. da Costa Martins P.A., Salic K., Gladka M.M., Armand A.S., Leptidis S., et al. (2010) MicroRNA-199b targets the nuclear kinase Dyrk1a in an auto-amplification loop promoting calcineurin/NFAT signalling. *Nat Cell Biol* **12**, 1220-1227.
228. Chen C., Ridzon D.A., Broomer A.J., Zhou Z., Lee D.H., et al. (2005) Real-time quantification of microRNAs by stem-loop RT-PCR. *Nucleic Acids Res* **33**, e179.
229. Biosystems A. Endogenous Controls for Real-Time Quantitation of miRNA Using TaqMan MicroRNA Assays. vol. TaqMan MicroRNA Assays.
230. Yang Y., Ago T., Zhai P., Abdellatif M., and Sadoshima J. (2011) Thioredoxin 1 negatively regulates angiotensin II-induced cardiac hypertrophy through upregulation of miR-98/let-7. *Circ Res* **108**, 305-313.
231. Zhou H., Arcila M.L., Li Z., Lee E.J., Henzler C., et al. (2012) Deep annotation of mouse iso-miR and iso-moR variation. *Nucleic Acids Res* **40**, 5864-5875.

232. Landgraf P., Rusu M., Sheridan R., Sewer A., Iovino N., et al. (2007) A mammalian microRNA expression atlas based on small RNA library sequencing. *Cell* **129**, 1401-1414.
233. Wyman S.K., Knouf E.C., Parkin R.K., Fritz B.R., Lin D.W., et al. (2011) Post-transcriptional generation of miRNA variants by multiple nucleotidyl transferases contributes to miRNA transcriptome complexity. *Genome Res* **21**, 1450-1461.
234. Jones M.R., Quinton L.J., Blahna M.T., Neilson J.R., Fu S., et al. (2009) Zcchc11-dependent uridylation of microRNA directs cytokine expression. *Nat Cell Biol* **11**, 1157-1163.
235. Katoh T., Sakaguchi Y., Miyauchi K., Suzuki T., Kashiwabara S., et al. (2009) Selective stabilization of mammalian microRNAs by 3' adenylation mediated by the cytoplasmic poly(A) polymerase GLD-2. *Genes Dev* **23**, 433-438.
236. Castellano L., and Stebbing J. (2013) Deep sequencing of small RNAs identifies canonical and non-canonical miRNA and endogenous siRNAs in mammalian somatic tissues. *Nucleic Acids Res* **41**, 3339-3351.
237. Maute R.L., Schneider C., Sumazin P., Holmes A., Califano A., et al. (2013) tRNA-derived microRNA modulates proliferation and the DNA damage response and is down-regulated in B cell lymphoma. *Proc Natl Acad Sci U S A* **110**, 1404-1409.
238. Lee Y.S., Shibata Y., Malhotra A., and Dutta A. (2009) A novel class of small RNAs: tRNA-derived RNA fragments (tRFs). *Genes Dev* **23**, 2639-2649.
239. Kawaji H., Nakamura M., Takahashi Y., Sandelin A., Katayama S., et al. (2008) Hidden layers of human small RNAs. *BMC genomics* **9**, 157.
240. Zaragosi L.E., Wdziekonski B., Brigand K.L., Villageois P., Mari B., et al. (2011) Small RNA sequencing reveals miR-642a-3p as a novel adipocyte-specific microRNA and miR-30 as a key regulator of human adipogenesis. *Genome Biol* **12**, R64.
241. Toedling J., Servant N., Ciaudo C., Farinelli L., Voinnet O., et al. (2012) Deep-sequencing protocols influence the results obtained in small-RNA sequencing. *PLoS One* **7**, e32724.
242. Linsen S.E., de Wit E., Janssens G., Heater S., Chapman L., et al. (2009) Limitations and possibilities of small RNA digital gene expression profiling. *Nature methods* **6**, 474-476.
243. Marioni J.C., Mason C.E., Mane S.M., Stephens M., and Gilad Y. (2008) RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. *Genome Res* **18**, 1509-1517.
244. Zhou X., Oshlack A., and Robinson M.D. (2013) miRNA-Seq normalization comparisons need improvement. *Rna* **19**, 733-734.
245. Garmire L.X., and Subramaniam S. (2012) Evaluation of normalization methods in mammalian microRNA-Seq data. *Rna* **18**, 1279-1288.
246. Git A., Dvinge H., Salmon-Divon M., Osborne M., Kutter C., et al. (2010) Systematic comparison of microarray profiling, real-time PCR, and next-generation sequencing technologies for measuring differential microRNA expression. *Rna* **16**, 991-1006.
247. Fahlgren N., Sullivan C.M., Kasschau K.D., Chapman E.J., Cumbie J.S., et al. (2009) Computational and analytical framework for small RNA profiling by high-throughput sequencing. *Rna* **15**, 992-1002.
248. van Rooij E., Quiat D., Johnson B.A., Sutherland L.B., Qi X., et al. (2009) A family of microRNAs encoded by myosin genes governs myosin expression and muscle performance. *Developmental cell* **17**, 662-673.
249. Zhang X., Wang X., Zhu H., Zhu C., Wang Y., et al. (2010) Synergistic effects of the GATA-4-mediated miR-144/451 cluster in protection against simulated ischemia/reperfusion-induced cardiomyocyte death. *J Mol Cell Cardiol* **49**, 841-850.

250. Liang Q., and Molkentin J.D. (2002) Divergent signaling pathways converge on GATA4 to regulate cardiac hypertrophic gene expression. *J Mol Cell Cardiol* **34**, 611-616.
251. Hautala N., Tokola H., Luodonpaa M., Puhakka J., Romppanen H., et al. (2001) Pressure overload increases GATA4 binding activity via endothelin-1. *Circulation* **103**, 730-735.
252. Lee E.J., Baek M., Gusev Y., Brackett D.J., Nuovo G.J., et al. (2008) Systematic evaluation of microRNA processing patterns in tissues, cell lines, and tumors. *Rna* **14**, 35-42.
253. Bail S., Swerdel M., Liu H., Jiao X., Goff L.A., et al. (2010) Differential regulation of microRNA stability. *Rna* **16**, 1032-1039.
254. Thomas M.F., Abdul-Wajid S., Panduro M., Babiarz J.E., Rajaram M., et al. (2012) Eri1 regulates microRNA homeostasis and mouse lymphocyte development and antiviral function. *Blood* **120**, 130-142.
255. Yang D.K., Choi B.Y., Lee Y.H., Kim Y.G., Cho M.C., et al. (2007) Gene profiling during regression of pressure overload-induced cardiac hypertrophy. *Physiol Genomics* **30**, 1-7.
256. Song X.W., Li Q., Lin L., Wang X.C., Li D.F., et al. (2010) MicroRNAs are dynamically regulated in hypertrophic hearts, and miR-199a is essential for the maintenance of cell size in cardiomyocytes. *Journal of cellular physiology* **225**, 437-443.
257. Lu H., Buchan R.J., and Cook S.A. (2010) MicroRNA-223 regulates Glut4 expression and cardiomyocyte glucose metabolism. *Cardiovascular research* **86**, 410-420.
258. Tatsuguchi M., Seok H.Y., Callis T.E., Thomson J.M., Chen J.F., et al. (2007) Expression of microRNAs is dynamically regulated during cardiomyocyte hypertrophy. *J Mol Cell Cardiol* **42**, 1137-1141.
259. Suarez Y., Fernandez-Hernando C., Pober J.S., and Sessa W.C. (2007) Dicer dependent microRNAs regulate gene expression and functions in human endothelial cells. *Circ Res* **100**, 1164-1173.
260. Elia L., Quintavalle M., Zhang J., Contu R., Cossu L., et al. (2009) The knockout of miR-143 and -145 alters smooth muscle cell maintenance and vascular homeostasis in mice: correlates with human disease. *Cell death and differentiation* **16**, 1590-1598.
261. Pan Z., Sun X., Shan H., Wang N., Wang J., et al. (2012) MicroRNA-101 inhibited postinfarct cardiac fibrosis and improved left ventricular compliance via the FBJ osteosarcoma oncogene/transforming growth factor-beta1 pathway. *Circulation* **126**, 840-850.
262. Duan Y., Zhou B., Su H., Liu Y., and Du C. (2013) miR-150 regulates high glucose-induced cardiomyocyte hypertrophy by targeting the transcriptional co-activator p300. *Experimental cell research* **319**, 173-184.
263. Wang X., Zhu H., Zhang X., Liu Y., Chen J., et al. (2012) Loss of the miR-144/451 cluster impairs ischaemic preconditioning-mediated cardioprotection by targeting Rac-1. *Cardiovascular research* **94**, 379-390.
264. Cordes K.R., Sheehy N.T., White M.P., Berry E.C., Morton S.U., et al. (2009) miR-145 and miR-143 regulate smooth muscle cell fate and plasticity. *Nature* **460**, 705-710.
265. Wang Y., Lee A.T., Ma J.Z., Wang J., Ren J., et al. (2008) Profiling microRNA expression in hepatocellular carcinoma reveals microRNA-224 up-regulation and apoptosis inhibitor-5 as a microRNA-224-specific target. *J Biol Chem* **283**, 13205-13215.
266. Yao G., Yin M., Lian J., Tian H., Liu L., et al. (2010) MicroRNA-224 is involved in transforming growth factor-beta-mediated mouse granulosa cell proliferation and granulosa cell function by targeting Smad4. *Molecular endocrinology* **24**, 540-551.

267. Mees S.T., Mardin W.A., Sielker S., Willscher E., Senninger N., et al. (2009) Involvement of CD40 targeting miR-224 and miR-486 on the progression of pancreatic ductal adenocarcinomas. *Annals of surgical oncology* **16**, 2339-2350.
268. MacLellan W.R., Brand T., and Schneider M.D. (1993) Transforming growth factor-beta in cardiac ontogeny and adaptation. *Circ Res* **73**, 783-791.
269. Tenedini E., Roncaglia E., Ferrari F., Orlandi C., Bianchi E., et al. (2010) Integrated analysis of microRNA and mRNA expression profiles in physiological myelopoiesis: role of hsa-mir-299-5p in CD34+ progenitor cells commitment. *Cell death & disease* **1**, e28.
270. Zhou Q., Gallagher R., Ufret-Vincenty R., Li X., Olson E.N., et al. (2011) Regulation of angiogenesis and choroidal neovascularization by members of microRNA-23~27~24 clusters. *Proc Natl Acad Sci U S A* **108**, 8287-8292.
271. Wang S.. and Olson E.N. (2009) AngiomiRs--key regulators of angiogenesis. *Current opinion in genetics & development* **19**, 205-211.
272. Tijsen A.J., Pinto Y.M., and Creemers E.E. (2012) Non-cardiomyocyte microRNAs in heart failure. *Cardiovascular research* **93**, 573-582.
273. Pawitan Y., Michiels S., Koscielny S., Gusnanto A., and Ploner A. (2005) False discovery rate, sensitivity and sample size for microarray studies. *Bioinformatics* **21**, 3017-3024.
274. Krol J., Sobczak K., Wilczynska U., Drath M., Jasinska A., et al. (2004) Structural features of microRNA (miRNA) precursors and their relevance to miRNA biogenesis and small interfering RNA/short hairpin RNA design. *J Biol Chem* **279**, 42230-42239.
275. Hu H.Y., Yan Z., Xu Y., Hu H., Menzel C., et al. (2009) Sequence features associated with microRNA strand selection in humans and flies. *BMC genomics* **10**, 413.
276. Morin R.D., O'Connor M.D., Griffith M., Kuchenbauer F., Delaney A., et al. (2008) Application of massively parallel sequencing to microRNA profiling and discovery in human embryonic stem cells. *Genome Res* **18**, 610-621.
277. Wu H., Ye C., Ramirez D., and Manjunath N. (2009) Alternative processing of primary microRNA transcripts by Drosha generates 5' end variation of mature microRNA. *PLoS One* **4**, e7566.
278. Liu N., Abe M., Sabin L.R., Hendriks G.J., Naqvi A.S., et al. (2011) The exoribonuclease Nibbler controls 3' end processing of microRNAs in Drosophila. *Curr Biol* **21**, 1888-1893.
279. Liu N., Bezprozvannaya S., Williams A.H., Qi X., Richardson J.A., et al. (2008) microRNA-133a regulates cardiomyocyte proliferation and suppresses smooth muscle gene expression in the heart. *Genes Dev* **22**, 3242-3254.
280. Xiao J., Luo X., Lin H., Zhang Y., Lu Y., et al. (2007) MicroRNA miR-133 represses HERG K+ channel expression contributing to QT prolongation in diabetic hearts. *J Biol Chem* **282**, 12363-12367.
281. Burroughs A.M., Ando Y., de Hoon M.J., Tomaru Y., Nishibu T., et al. (2010) A comprehensive survey of 3' animal miRNA modification events and a possible role for 3' adenylation in modulating miRNA targeting effectiveness. *Genome Res* **20**, 1398-1410.
282. Westholm J.O., Ladewig E., Okamura K., Robine N., and Lai E.C. (2012) Common and distinct patterns of terminal modifications to mirtrons and canonical microRNAs. *Rna* **18**, 177-192.
283. Knouf E.C., Wyman S.K., and Tewari M. (2013) The human TUT1 nucleotidyl transferase as a global regulator of microRNA abundance. *PLoS One* **8**, e69630.
284. Diederichs S., Jung S., Rothenberg S.M., Smolen G.A., Mlody B.G., et al. (2008) Coexpression of Argonaute-2 enhances RNA interference toward perfect match binding sites. *Proc Natl Acad Sci U S A* **105**, 9284-9289.

285. Yang W., Chendrimada T.P., Wang Q., Higuchi M., Seeburg P.H., et al. (2006) Modulation of microRNA processing and expression through RNA editing by ADAR deaminases. *Nat Struct Mol Biol* **13**, 13-21.
286. Shin C., Nam J.W., Farh K.K., Chiang H.R., Shkumatava A., et al. (2010) Expanding the microRNA targeting code: functional sites with centered pairing. *Mol Cell* **38**, 789-802.
287. Sandberg R., Neilson J.R., Sarma A., Sharp P.A., and Burge C.B. (2008) Proliferating cells express mRNAs with shortened 3' untranslated regions and fewer microRNA target sites. *Science* **320**, 1643-1647.
288. Yoon O.K., and Brem R.B. (2010) Noncanonical transcript forms in yeast and their regulation during environmental stress. *Rna* **16**, 1256-1267.
289. Derti A., Garrett-Engele P., Macisaac K.D., Stevens R.C., Sriram S., et al. (2012) A quantitative atlas of polyadenylation in five mammals. *Genome Res* **22**, 1173-1183.
290. Martin G., Gruber A.R., Keller W., and Zavolan M. (2012) Genome-wide analysis of pre-mRNA 3' end processing reveals a decisive role of human cleavage factor I in the regulation of 3' UTR length. *Cell reports* **1**, 753-763.
291. Shepard P.J., Choi E.A., Lu J., Flanagan L.A., Hertel K.J., et al. (2011) Complex and dynamic landscape of RNA polyadenylation revealed by PAS-Seq. *Rna* **17**, 761-772.
292. Jenal M., Elkon R., Loayza-Puch F., van Haaften G., Kuhn U., et al. (2012) The poly(A)-binding protein nuclear 1 suppresses alternative cleavage and polyadenylation sites. *Cell* **149**, 538-553.
293. Jan C.H., Friedman R.C., Ruby J.G., and Bartel D.P. (2011) Formation, regulation and evolution of Caenorhabditis elegans 3'UTRs. *Nature* **469**, 97-101.
294. Cai X., Hagedorn C.H., and Cullen B.R. (2004) Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. *Rna* **10**, 1957-1966.
295. Mercer T.R., Dinger M.E., and Mattick J.S. (2009) Long non-coding RNAs: insights into functions. *Nat Rev Genet* **10**, 155-159.
296. Safran M., Dalah I., Alexander J., Rosen N., Iny Stein T., et al. (2010) GeneCards Version 3: the human gene integrator. *Database : the journal of biological databases and curation* **2010**, baq020.
297. Poelchau M.F., Reynolds J.A., Elsik C.G., Denlinger D.L., and Armbruster P.A. (2013) Deep sequencing reveals complex mechanisms of diapause preparation in the invasive mosquito, Aedes albopictus. *Proceedings. Biological sciences / The Royal Society* **280**, 20130143.
298. Ross-Innes C.S., Stark R., Teschendorff A.E., Holmes K.A., Ali H.R., et al. (2012) Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. *Nature* **481**, 389-393.
299. Weiss Y., Foret S., Hayward D.C., Ainsworth T., King R., et al. (2013) The acute transcriptional response of the coral Acropora millepora to immune challenge: expression of GiMAP/IAN genes links the innate immune responses of corals with those of mammals and plants. *BMC genomics* **14**, 400.
300. Ashburner M., Ball C.A., Blake J.A., Botstein D., Butler H., et al. (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nature genetics* **25**, 25-29.
301. van den Bosch B.J., Lindsey P.J., van den Burg C.M., van der Vlies S.A., Lips D.J., et al. (2006) Early and transient gene expression changes in pressure overload-induced cardiac hypertrophy in mice. *Genomics* **88**, 480-488.
302. Friddle C.J., Koga T., Rubin E.M., and Bristow J. (2000) Expression profiling reveals distinct sets of genes altered during induction and regression of cardiac hypertrophy. *Proc Natl Acad Sci U S A* **97**, 6745-6750.

303. Pathak A., del Monte F., Zhao W., Schultz J.E., Lorenz J.N., et al. (2005) Enhancement of cardiac function and suppression of heart failure progression by inhibition of protein phosphatase 1. *Circ Res* **96**, 756-766.
304. Sugden P.H., Fuller S.J., Weiss S.C., and Clerk A. (2008) Glycogen synthase kinase 3 (GSK3) in the heart: a point of integration in hypertrophic signalling and a therapeutic target? A critical analysis. *British journal of pharmacology* **153 Suppl 1**, S137-153.
305. Zhang T., and Brown J.H. (2004) Role of Ca2+/calmodulin-dependent protein kinase II in cardiac hypertrophy and heart failure. *Cardiovascular research* **63**, 476-486.
306. Zuern C., Krenacs L., Starke S., Heimrich J., Palmetshofer A., et al. (2012) Microtubule associated tumor suppressor 1 deficient mice develop spontaneous heart hypertrophy and SLE-like lymphoproliferative disease. *International journal of oncology* **40**, 1079-1088.
307. Busk P.K., Hinrichsen R., Bartkova J., Hansen A.H., Christoffersen T.E., et al. (2005) Cyclin D2 induces proliferation of cardiac myocytes and represses hypertrophy. *Experimental cell research* **304**, 149-161.
308. Busk P.K., Bartkova J., Strom C.C., Wulf-Andersen L., Hinrichsen R., et al. (2002) Involvement of cyclin D activity in left ventricle hypertrophy in vivo and in vitro. *Cardiovascular research* **56**, 64-75.
309. Coso O.A., Chiariello M., Yu J.C., Teramoto H., Crespo P., et al. (1995) The small GTP-binding proteins Rac1 and Cdc42 regulate the activity of the JNK/SAPK signaling pathway. *Cell* **81**, 1137-1146.
310. Goberdhan D.C., Paricio N., Goodman E.C., Mlodzik M., and Wilson C. (1999) Drosophila tumor suppressor PTEN controls cell size and number by antagonizing the Chico/PI3-kinase signaling pathway. *Genes Dev* **13**, 3244-3258.
311. Handschin C., and Spiegelman B.M. (2006) Peroxisome proliferator-activated receptor gamma coactivator 1 coactivators, energy homeostasis, and metabolism. *Endocrine reviews* **27**, 728-735.
312. Watanabe A., Arai M., Koitabashi N., Niwano K., Ohyama Y., et al. (2011) Mitochondrial transcription factors TFAM and TFB2M regulate Serca2 gene transcription. *Cardiovascular research* **90**, 57-67.
313. Lianoglou S., Garg V., Yang J.L., Leslie C.S., and Mayr C. (2013) Ubiquitously transcribed genes use alternative polyadenylation to achieve tissue-specific expression. *Genes Dev* **27**, 2380-2396.
314. Miura P., Shenker S., Andreu-Agullo C., Westholm J.O., and Lai E.C. (2013) Widespread and extensive lengthening of 3' UTRs in the mammalian brain. *Genome Res* **23**, 812-825.
315. Djebali S., Davis C.A., Merkel A., Dobin A., Lassmann T., et al. (2012) Landscape of transcription in human cells. *Nature* **489**, 101-108.
316. Hangauer M.J., Vaughn I.W., and McManus M.T. (2013) Pervasive transcription of the human genome produces thousands of previously unidentified long intergenic noncoding RNAs. *PLoS genetics* **9**, e1003569.
317. Elkon R., Ugalde A.P., and Agami R. (2013) Alternative cleavage and polyadenylation: extent, regulation and function. *Nat Rev Genet* **14**, 496-506.
318. Lin Y., Li Z., Ozsolak F., Kim S.W., Arango-Argoty G., et al. (2012) An in-depth map of polyadenylation sites in cancer. *Nucleic Acids Res* **40**, 8460-8471.
319. Beaudoing E., Freier S., Wyatt J.R., Claverie J.M., and Gautheret D. (2000) Patterns of variant polyadenylation signal usage in human genes. *Genome Res* **10**, 1001-1010.
320. Nunes N.M., Li W., Tian B., and Furger A. (2010) A functional human Poly(A) site requires only a potent DSE and an A-rich upstream sequence. *EMBO J* **29**, 1523-1536.

321. Hogg J.R., and Goff S.P. (2010) Upf1 senses 3'UTR length to potentiate mRNA decay. *Cell* **143**, 379-389.
322. Chen C.Y., Gherzi R., Ong S.E., Chan E.L., Raijmakers R., et al. (2001) AU binding proteins recruit the exosome to degrade ARE-containing mRNAs. *Cell* **107**, 451-464.
323. Mbella E.G., Bertrand S., Huez G., and Octave J.N. (2000) A GG nucleotide sequence of the 3' untranslated region of amyloid precursor protein mRNA plays a key role in the regulation of translation and the binding of proteins. *Mol Cell Biol* **20**, 4572-4579.
324. Blichenberg A., Rehbein M., Muller R., Garner C.C., Richter D., et al. (2001) Identification of a cis-acting dendritic targeting element in the mRNA encoding the alpha subunit of Ca2+/calmodulin-dependent protein kinase II. *The European journal of neuroscience* **13**, 1881-1888.
325. Dickson J.R., Kruse C., Montagna D.R., Finsen B., and Wolfe M.S. (2013) Alternative polyadenylation and miR-34 family members regulate tau expression. *Journal of neurochemistry* **127**, 739-749.
326. Young R.M., Shull G.E., and Lingrel J.B. (1987) Multiple mRNAs from rat kidney and brain encode a single Na+,K+-ATPase beta subunit protein. *J Biol Chem* **262**, 4905-4910.
327. Shao Y., and Ismail-Beigi F. (2001) Different Na, K-ATPase mRNA(beta1) species exhibit unique translational efficiencies. *Archives of biochemistry and biophysics* **390**, 78-86.
328. Ranganathan G., Ong J.M., Yukht A., Saghizadeh M., Simsolo R.B., et al. (1995) Tissue-specific expression of human lipoprotein lipase. Effect of the 3'-untranslated region on translation. *J Biol Chem* **270**, 7149-7155.
329. An J.J., Gharami K., Liao G.Y., Woo N.H., Lau A.G., et al. (2008) Distinct role of long 3' UTR BDNF mRNA in spine morphology and synaptic plasticity in hippocampal neurons. *Cell* **134**, 175-187.
330. Ranganathan G., Vu D., and Kern P.A. (1997) Translational regulation of lipoprotein lipase by epinephrine involves a trans-acting binding protein interacting with the 3' untranslated region. *J Biol Chem* **272**, 2515-2519.
331. Wang L., Dowell R.D., and Yi R. (2013) Genome-wide maps of polyadenylation reveal dynamic mRNA 3'-end formation in mammalian cell lineages. *Rna* **19**, 413-425.
332. Beck A.H., Weng Z., Witten D.M., Zhu S., Foley J.W., et al. (2010) 3'-end sequencing for expression quantification (3SEQ) from archival tumor samples. *PLoS One* **5**, e8768.
333. Abdellatif M., Packer S.E., Michael L.H., Zhang D., Charng M.J., et al. (1998) A Ras-dependent pathway regulates RNA polymerase II phosphorylation in cardiac myocytes: implications for cardiac hypertrophy. *Mol Cell Biol* **18**, 6729-6736.
334. McDermott P.J., Carl L.L., Conner K.J., and Allo S.N. (1991) Transcriptional regulation of ribosomal RNA synthesis during growth of cardiac myocytes in culture. *J Biol Chem* **266**, 4409-4416.
335. Cutilletta A.F., Rudnik M., and Zak R. (1978) Muscle and non-muscle cell RNA polymerase activity during the development of myocardial hypertrophy. *J Mol Cell Cardiol* **10**, 677-687.
336. Feldman A.M., Weinberg E.O., Ray P.E., and Lorell B.H. (1993) Selective changes in cardiac gene expression during compensated hypertrophy and the transition to cardiac decompensation in rats with chronic aortic banding. *Circ Res* **73**, 184-192.
337. Hein S., Kostin S., Heling A., Maeno Y., and Schaper J. (2000) The role of the cytoskeleton in heart failure. *Cardiovascular research* **45**, 273-278.

338. Skwarek-Maruszewska A., Hotulainen P., Mattila P.K., and Lappalainen P. (2009) Contractility-dependent actin dynamics in cardiomyocyte sarcomeres. *J Cell Sci* **122**, 2119-2126.
339. Vandekerckhove J., Bugaisky G., and Buckingham M. (1986) Simultaneous expression of skeletal muscle and heart actin proteins in various striated muscle tissues and cells. A quantitative determination of the two actin isoforms. *J Biol Chem* **261**, 1838-1843.
340. Ingber D. (1991) Integrins as mechanochemical transducers. *Curr Opin Cell Biol* **3**, 841-848.
341. Stossel T.P., Condeelis J., Cooley L., Hartwig J.H., Noegel A., et al. (2001) Filamins as integrators of cell mechanics and signalling. *Nat Rev Mol Cell Biol* **2**, 138-145.
342. Gardel M.L., Schneider I.C., Aratyn-Schaus Y., and Waterman C.M. (2010) Mechanical integration of actin and adhesion dynamics in cell migration. *Annual review of cell and developmental biology* **26**, 315-333.
343. Ruperez M., Lorenzo O., Blanco-Colio L.M., Esteban V., Egido J., et al. (2003) Connective tissue growth factor is a mediator of angiotensin II-induced fibrosis. *Circulation* **108**, 1499-1505.
344. Hannan R.D., Jenkins A., Jenkins A.K., and Brandenburger Y. (2003) Cardiac hypertrophy: a matter of translation. *Clinical and experimental pharmacology & physiology* **30**, 517-527.
345. Brandenburger Y., Jenkins A., Autelitano D.J., and Hannan R.D. (2001) Increased expression of UBF is a critical determinant for rRNA synthesis and hypertrophic growth of cardiac myocytes. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* **15**, 2051-2053.
346. Lopaschuk G.D., Belke D.D., Gamble J., Itoi T., and Schonekess B.O. (1994) Regulation of fatty acid oxidation in the mammalian heart in health and disease. *Biochim Biophys Acta* **1213**, 263-276.
347. Barger P.M., and Kelly D.P. (1999) Fatty acid utilization in the hypertrophied and failing heart: molecular regulatory mechanisms. *The American journal of the medical sciences* **318**, 36-42.
348. Allard M.F., Schonekess B.O., Henning S.L., English D.R., and Lopaschuk G.D. (1994) Contribution of oxidative metabolism and glycolysis to ATP production in hypertrophied hearts. *The American journal of physiology* **267**, H742-750.
349. Bishop S.P., and Altschuld R.A. (1970) Increased glycolytic metabolism in cardiac hypertrophy and congestive failure. *The American journal of physiology* **218**, 153-159.
350. Christe M.E., and Rodgers R.L. (1994) Altered glucose and fatty acid oxidation in hearts of the spontaneously hypertensive rat. *J Mol Cell Cardiol* **26**, 1371-1375.
351. Wittels B., and Spann J.F., Jr. (1968) Defective lipid metabolism in the failing heart. *J Clin Invest* **47**, 1787-1794.
352. Sack M.N., Disch D.L., Rockman H.A., and Kelly D.P. (1997) A role for Sp and nuclear receptor transcription factors in a cardiac hypertrophic growth program. *Proc Natl Acad Sci U S A* **94**, 6438-6443.
353. Sack M.N., Rader T.A., Park S., Bastin J., McCune S.A., et al. (1996) Fatty acid oxidation enzyme gene expression is downregulated in the failing heart. *Circulation* **94**, 2837-2842.
354. Nascimben L., Ingwall J.S., Lorell B.H., Pinz I., Schultz V., et al. (2004) Mechanisms for increased glycolysis in the hypertrophied rat heart. *Hypertension* **44**, 662-667.
355. Razeghi P., Young M.E., Alcorn J.L., Moravec C.S., Frazier O.H., et al. (2001) Metabolic gene expression in fetal and failing human heart. *Circulation* **104**, 2923-2931.

356. Allard M.F., Wambolt R.B., Longnus S.L., Grist M., Lydell C.P., et al. (2000) Hypertrophied rat hearts are less responsive to the metabolic and functional effects of insulin. *American journal of physiology. Endocrinology and metabolism* **279**, E487-493.
357. Kolwicz S.C., Jr., and Tian R. (2011) Glucose metabolism and cardiac hypertrophy. *Cardiovascular research* **90**, 194-201.
358. Morisco C., Marrone C., Galeotti J., Shao D., Vatner D.E., et al. (2008) Endocytosis machinery is required for beta1-adrenergic receptor-induced hypertrophy in neonatal rat cardiac myocytes. *Cardiovascular research* **78**, 36-44.
359. Chuang S.M., and Madura K. (2005) Saccharomyces cerevisiae Ub-conjugating enzyme Ubc4 binds the proteasome in the presence of translationally damaged proteins. *Genetics* **171**, 1477-1484.
360. Hedhli N., and Depre C. (2010) Proteasome inhibitors and cardiac cell growth. *Cardiovascular research* **85**, 321-329.
361. Friehs I., and del Nido P.J. (2003) Increased susceptibility of hypertrophied hearts to ischemic injury. *The Annals of thoracic surgery* **75**, S678-684.
362. Shiojima I., and Walsh K. (2006) Regulation of cardiac growth and coronary angiogenesis by the Akt/PKB signaling pathway. *Genes Dev* **20**, 3347-3365.
363. Yoshioka J., Prince R.N., Huang H., Perkins S.B., Cruz F.U., et al. (2005) Cardiomyocyte hypertrophy and degradation of connexin43 through spatially restricted autocrine/paracrine heparin-binding EGF. *Proc Natl Acad Sci U S A* **102**, 10622-10627.
364. Panek A.N., Posch M.G., Alenina N., Ghadge S.K., Erdmann B., et al. (2009) Connective tissue growth factor overexpression in cardiomyocytes promotes cardiac hypertrophy and protection against pressure overload. *PLoS One* **4**, e6743.
365. Fisch S., Gray S., Heymans S., Haldar S.M., Wang B., et al. (2007) Kruppel-like factor 15 is a regulator of cardiomyocyte hypertrophy. *Proc Natl Acad Sci U S A* **104**, 7074-7079.
366. Leenders J.J., Wijnen W.J., Hiller M., van der Made I., Lentink V., et al. (2010) Regulation of cardiac gene expression by KLF15, a repressor of myocardin activity. *J Biol Chem* **285**, 27449-27456.
367. Windak R., Muller J., Felley A., Akhmedov A., Wagner E.F., et al. (2013) The AP-1 transcription factor c-Jun prevents stress-imposed maladaptive remodeling of the heart. *PLoS One* **8**, e73294.
368. Khatri P., Sirota M., and Butte A.J. (2012) Ten years of pathway analysis: current approaches and outstanding challenges. *PLoS Comput Biol* **8**, e1002375.
369. Takagaki Y., Seipelt R.L., Peterson M.L., and Manley J.L. (1996) The polyadenylation factor CstF-64 regulates alternative processing of IgM heavy chain pre-mRNA during B cell differentiation. *Cell* **87**, 941-952.
370. Tian B., and Manley J.L. (2013) Alternative cleavage and polyadenylation: the long and short of it. *Trends Biochem Sci* **38**, 312-320.
371. Bava F.A., Eliscovich C., Ferreira P.G., Minana B., Ben-Dov C., et al. (2013) CPEB1 coordinates alternative 3'-UTR formation with translational regulation. *Nature* **495**, 121-125.
372. Perales R., and Bentley D. (2009) "Cotranscriptionality": the transcription elongation complex as a nexus for nuclear transactions. *Mol Cell* **36**, 178-191.
373. Di Giammartino D.C., Nishida K., and Manley J.L. (2011) Mechanisms and consequences of alternative polyadenylation. *Mol Cell* **43**, 853-866.
374. Ji Z., Luo W., Li W., Hoque M., Pan Z., et al. (2011) Transcriptional activity regulates alternative cleavage and polyadenylation. *Molecular systems biology* **7**, 534.
375. Pinto P.A., Henriques T., Freitas M.O., Martins T., Domingues R.G., et al. (2011) RNA polymerase II kinetics in polo polyadenylation signal selection. *EMBO J* **30**, 2431-2444.

376. Elkon R., Drost J., van Haaften G., Jenal M., Schrier M., et al. (2012) E2F mediates enhanced alternative polyadenylation in proliferation. *Genome Biol* **13**, R59.
377. Nelson T.J., Balza R., Jr., Xiao Q., and Misra R.P. (2005) SRF-dependent gene expression in isolated cardiomyocytes: regulation of genes involved in cardiac hypertrophy. *J Mol Cell Cardiol* **39**, 479-489.
378. Chen J.F., Mandel E.M., Thomson J.M., Wu Q., Callis T.E., et al. (2006) The role of microRNA-1 and microRNA-133 in skeletal muscle proliferation and differentiation. *Nature genetics* **38**, 228-233.
379. Abdellatif M. (2010) The role of microRNA-133 in cardiac hypertrophy uncovered. *Circ Res* **106**, 16-18.
380. Li Q., Lin X., Yang X., and Chang J. (2010) NFATc4 is negatively regulated in miR-133a-mediated cardiomyocyte hypertrophic repression. *American journal of physiology. Heart and circulatory physiology* **298**, H1340-1347.
381. Xu C., Lu Y., Pan Z., Chu W., Luo X., et al. (2007) The muscle-specific microRNAs miR-1 and miR-133 produce opposing effects on apoptosis by targeting HSP60, HSP70 and caspase-9 in cardiomyocytes. *J Cell Sci* **120**, 3045-3052.
382. Li Z., Hassan M.Q., Volinia S., van Wijnen A.J., Stein J.L., et al. (2008) A microRNA signature for a BMP2-induced osteoblast lineage commitment program. *Proc Natl Acad Sci U S A* **105**, 13906-13911.
383. Zhang Y.. and Verbeek F.J. (2010) Comparison and integration of target prediction algorithms for microRNA studies. *Journal of integrative bioinformatics* **7**.
384. John B., Enright A.J., Aravin A., Tuschl T., Sander C., et al. (2004) Human MicroRNA targets. *PLoS biology* **2**, e363.
385. Duisters R.F., Tijsen A.J., Schroen B., Leenders J.J., Lentink V., et al. (2009) miR-133 and miR-30 regulate connective tissue growth factor: implications for a role of microRNAs in myocardial matrix remodeling. *Circ Res* **104**, 170-178, 176p following 178.
386. Baek D., Villen J., Shin C., Camargo F.D., Gygi S.P., et al. (2008) The impact of microRNAs on protein output. *Nature* **455**, 64-71.
387. Elefant N., Altuvia Y., and Margalit H. (2011) A wide repertoire of miRNA binding sites: prediction and functional implications. *Bioinformatics* **27**, 3093-3101.
388. Brennecke J., Stark A., Russell R.B., and Cohen S.M. (2005) Principles of microRNA-target recognition. *PLoS biology* **3**, e85.
389. McGahon M.K., Yarham J.M., Daly A., Guduric-Fuchs J., Ferguson L.J., et al. (2013) Distinctive profile of IsomiR expression and novel microRNAs in rat heart left ventricle. *PLoS One* **8**, e65809.
390. Maragkakis M., Reczko M., Simossis V.A., Alexiou P., Papadopoulos G.L., et al. (2009) DIANA-microT web server: elucidating microRNA functions through target prediction. *Nucleic Acids Res* **37**, W273-276.
391. Schmitt M.J., Philippidou D., Reinsbach S.E., Margue C., Wienecke-Baldacchino A., et al. (2012) Interferon-gamma-induced activation of Signal Transducer and Activator of Transcription 1 (STAT1) up-regulates the tumor suppressing microRNA-29 family in melanoma cells. *Cell communication and signaling : CCS* **10**, 41.
392. Nazarov P.V., Reinsbach S.E., Muller A., Nicot N., Philippidou D., et al. (2013) Interplay of microRNAs, transcription factors and target genes: linking dynamic expression changes to function. *Nucleic Acids Res* **41**, 2817-2831.
393. Pruitt K.D., Tatusova T., and Maglott D.R. (2005) NCBI Reference Sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Res* **33**, D501-504.
394. Majoros W.H.. and Ohler U. (2007) Spatial preferences of microRNA targets in 3' untranslated regions. *BMC genomics* **8**, 152.

395. Angelis E., Garcia A., Chan S.S., Schenke-Layland K., Ren S., et al. (2008) A cyclin D2-Rb pathway regulates cardiac myocyte size and RNA polymerase III after biomechanical stress in adult myocardium. *Circ Res* **102**, 1222-1229.
396. White R.J., Trouche D., Martin K., Jackson S.P., and Kouzarides T. (1996) Repression of RNA polymerase III transcription by the retinoblastoma protein. *Nature* **382**, 88-90.
397. Yan G., Zhang L., Fang T., Zhang Q., Wu S., et al. (2012) MicroRNA-145 suppresses mouse granulosa cell proliferation by targeting activin receptor IB. *FEBS Lett* **586**, 3263-3270.
398. Wahid F., Shehzad A., Khan T., and Kim Y.Y. (2010) MicroRNAs: synthesis, mechanism, function, and recent clinical trials. *Biochim Biophys Acta* **1803**, 1231-1243.
399. Go A.S., Mozaffarian D., Roger V.L., Benjamin E.J., Berry J.D., et al. (2013) Heart disease and stroke statistics--2013 update: a report from the American Heart Association. *Circulation* **127**, e6-e245.
400. Dangwal S.. and Thum T. (2014) microRNA Therapeutics in Cardiovascular Disease Models. *Annual review of pharmacology and toxicology* **54**, 185-203.
401. Bader A.G., Brown D., and Winkler M. (2010) The promise of microRNA replacement therapy. *Cancer research* **70**, 7027-7030.
402. Lanford R.E., Hildebrandt-Eriksen E.S., Petri A., Persson R., Lindow M., et al. (2010) Therapeutic silencing of microRNA-122 in primates with chronic hepatitis C virus infection. *Science* **327**, 198-201.
403. Pharma S. (2008) Santaris Pharma begins human clinical testing of world's first medicine targeted at a human microRNA. In: News Release, pp 1-4.
404. Kwekkeboom R.F., Lei Z., Doevendans P.A., Musters R.J., and Sluijter J.P. (2014) Targeted delivery of miRNA therapeutics for cardiovascular diseases: opportunities and challenges. *Clinical science* **127**, 351-365.
405. Montgomery R.L., Hullinger T.G., Semus H.M., Dickinson B.A., Seto A.G., et al. (2011) Therapeutic inhibition of miR-208a improves cardiac function and survival during heart failure. *Circulation* **124**, 1537-1547.
406. Clancy J.L., Nousch M., Humphreys D.T., Westman B.J., Beilharz T.H., et al. (2007) Methods to analyze microRNA-mediated control of mRNA translation. *Methods in enzymology* **431**, 83-111.