GENETIC AND EPIGENETIC CHANGES ASSOCIATED WITH POLYGENIC LEFT VENTRICULAR HYPERTROPHY

Priscilla Ribeiro Prestes

BSc (Hons), Federal University of Rio Grande do Sul, Brazil

MSc, Federal University of Rio Grande do Sul, Brazil

A thesis submitted in total fulfilment for the requirements of the Doctor of Philosophy (PhD) degree

School of Science, Psychology and Sport

Federation University Australia

Federation University Australia
PO BOX 663
University Drive, Mount Helen
Ballarat, Victoria, 3353,
Australia

Original submission: March 2021.

Final submission: July 2021.

Brief summary

Cardiac hypertrophy (CH) is the thickening of heart muscles reducing functionality and increasing risk of cardiac disease. Commonly, pathological CH is presented as left ventricular hypertrophy (LVH) and genetic factors are known to be involved but their contribution is still poorly understood.

I used the hypertrophic heart rat (HHR), a unique normotensive polygenic model of LVH, and its control strain, the normal heart rat (NHR) to investigate genetic and epigenetic contributions to LVH independent of high blood pressure. To address this study, I used a systematic approach.

Firstly, I sequenced the whole genome of HHR and NHR to identify genes related to LVH, focusing on quantitative trait locus Cm22. I found the gene for tripartite motif-containing 55 (Trim55) was significantly downregulated and also presented decreased protein expression with the presence of one exonic missense mutation that altered the protein structure. Interestingly, Trim55 mRNA expression was reduced in idiopathic dilated cardiomyopathic hearts.

Secondly, I selected 42 genes previously described in monogenic forms of human cardiomyopathies and studied DNA variants, mRNA and micro RNA (miRNA) expression to determine their involvement in this polygenic model of LVH at five ages. This comprehensive approach identified the differential expression of 29 genes in at least one age group and two miRNAs in validated miRNA-mRNA interactions. These two miRNAs have binding sites for five of the genes studied.

Lastly, I found circular RNA (circRNA) *Hrcr* was upregulated in the hypertrophic heart. I then silenced *Hrcr* expression in human primary cardiomyocytes to investigate its miRNA downstream targets and elucidate possible regulatory mechanisms. I described four

miRNAs (miR-1-3p, miR-330, miR-27a-5p, miR-299-5p) as novel targets for *HRCR* and predicted 359 mRNA targets in the circRNA-miRNA-mRNA regulatory axis. *In silico* analysis identified 206 enriched gene ontology based on the predicted mRNA target list, including cardiomyocyte differentiation and ventricular cardiac muscle cell differentiation.

The findings in this thesis suggest that 1) *Trim55* is a novel functional candidate gene for polygenic LVH; 2) genes implicated in monogenic forms of cardiomyopathy may be involved in this condition and 3) circRNA expression is associated with changes in hypertrophic hearts and deserve further attention.

Acknowledgements

I would like to acknowledge my principal supervisor Prof Fadi Charchar. Thank you for your invaluable mentorship, friendship, advice and support. I truly appreciate your trust and the opportunities I have had since joining the Charchar lab. Most importantly, I need to thank you for believing in me when I questioned myself and was in doubt if I wanted to embark on this journey again.

To Prof Stephen Harrap, thank you for all the advice and feedback throughout my candidature. Your knowledge and patience are admirable, and your friendly smile made every early train trip to the University of Melbourne a pleasure. I strive to one day be as knowledgeable, approachable and wise as you are.

To my Federation University Australia colleagues who made time fly and laboratory work and lunchtime less lonely, thank you for the insightful discussions, company in the tearoom, any type of conversations and camaraderie.

Dr Michelle Maier, your wisdom in cell culture was invaluable to help me achieve my final results when I was uncertain I could grow cells myself. However, your friendship, care and understanding made returning to work after maternity leave more enjoyable and our many chats, about science, work and life were invaluable.

Dr Sean Byars, thank you for the last mad dash to help me analyse my sequencing experiments, discuss results and answer questions. I truly appreciate your willingness to help, patience and knowledge.

To my mentor Prof Rob Wallis, who more recently came along my journey and has always offered to read any and all piece of my work, fellowship applications or manuscript

drafts I would be working on, thank you for your feedback and constructive criticism. And most importantly, thank you for advocating for me and being my referee.

All my friends and extended family have been instrumental in this accomplishment. Thank you for your support and friendship throughout my life.

I would like to add a very special thank you to my best friend and brilliant scientist A/Prof Francine Marques. You are the reason I came to Australia, met my husband and dove into the world of cardiovascular disease. Thank you very much for teaching me so many laboratory techniques, reading many applications and cheering every accomplishment in my professional and personal life. Your guidance, patience and commitment have definitely influenced my path, my decisions and my life.

My parents have been my role models from the moment I was born. I would never be where I am if it weren't for the sacrifices you both made, love you gave me and overall support in my life and my decision. If you had not believed that I could, I certainly wouldn't. I will forever repeat: "Without you, I would be no one!".

To my loving husband who understands little of what I do but has never doubted me, enjoyed every overseas trip for conferences and cheered all my accomplishments and awards. I could not thank enough for believing in me and for having you in my life.

My cutest little girl, Sophia, you are the reason I get up in the morning, you keep me on my toes and challenge me every day to be the best person I can be. I hope one day you will realise how important you are in my life and how much you push me to be the best mum, person and scientist I can be, just so you are half as proud of me as I am of my parents.

Statement of Authorship

Except where explicit reference is made in text of the thesis, this thesis contains no material published elsewhere or extracted in whole or in part from a thesis by which I have qualified for or have been awarded another degree or diploma. No other person's work has been relied upon or used without due acknowledgement in the main text and bibliography of the thesis.

Signed: Date: 25/01/2021

(Candidate)

Signed: Date: 02/02/2021

(Prof Fadi Charchar – Principal Supervisor)

Disclosure

Priscilla Ribeiro Prestes was supported by an Australian Government Research Training Program (RTP) Stipend and RTP Fee-Offset Scholarship through Federation University Australia and an Australian Government RTP Fee-Offset Scholarship through Federation University Australia.

Table of contents

List of figures

List of tables

Publications arising from my PhD candidature and my contribution

1. **Prestes PR**, Marques FZ, Lopez-Campos G, Booth SA, McGlynn M, Lewandowski P, Delbridge LM, Harrap SB, Charchar FJ (2016). Tripartite motif-containing 55 identified as functional candidate for spontaneous cardiac hypertrophy in the rat locus cardiac mass 22. *J Hypertens*, 34:950-8.

   85% contribution

   I contributed to the experimental design, performed the experiments, analysed and interpreted data, created figures and tables, wrote and revised the manuscript.

2. **Prestes PR**, Marques FZ, Lopez-Campos G, Lewandowski P, Delbridge LMD, Charchar FJ, Harrap SB (2018). Involvement of human monogenic cardiomyopathy genes in experimental polygenic cardiac hypertrophy. *Physiol Genomics*, 50:680-687.

   90% contribution

   I contributed to the experimental design, performed the experiments, analysed and interpreted data, created figures and tables, wrote and revised the manuscript.

3. **Prestes PR**, Maier MC, Byars S, Harrap SB, Charchar FJ. Heart related circular RNA (*Hrcr*) influences polygenic cardiac hypertrophy by altering cardiac microRNAs. Submitted to the Journal of the American Heart Association.

   90% contribution

   I established the experimental design, performed the experiments, analysed and interpreted data, created figures and tables, wrote and revised the manuscript.

Additional publications arising during my PhD candidature

1. **Prestes PR**, Charchar FJ (2017). Is there a link between mitochondrial DNA and blood pressure?. *J Hum Hypertens*, 31:761-762.

2. Marques FZ, **Prestes PR**, Byars SG, Ritchie SC, Würtz P, Patel SK, Booth SA, Rana I, Minoda Y, Berzins SP, Curl CL, Bell JR, Wai B, Srivastava PM, Kangas AJ, Soininen P, Ruohonen S, Kähönen M, Lehtimäki T, Raitoharju E, Havulinna A, Perola M, Raitakari O, Salomaa V, Ala-Korpela M, Kettunen J, McGlynn M, Kelly J, Wlodek ME, Lewandowski PA, Delbridge LM, Burrell LM, Inouye M, Harrap SB, Charchar FJ (2017). Experimental and Human Evidence for Lipocalin-2 (Neutrophil Gelatinase-Associated Lipocalin [NGAL]) in the Development of Cardiac Hypertrophy and heart failure. *J Am Heart Assoc*, 6(6):e005971.

3. Griffin PC, Khadake J, LeMay KS, Lewis SE, Orchard S, Pask A, Pope B, Roessner U, Russell K, Seemann T, Treloar A, Tyagi S, Christiansen JH, Dayalan S, Gladman S, Hangartner SB, Hayden HL, Ho WWH, Keeble-Gagnère G, Korhonen PK, Neish P, **Prestes PR**, Richardson MF, Watson-Haigh NS, Wyres KL, Young ND, Schneider MV (2018). Best Practice Data Life Cycle Approaches for the Life Sciences. *F1000 Research*, 6:1618.

4. Xu X, Eales JM, Akbarov A, Guo H, Becker L, Talavera D, Ashraf F, Nawaz J, Pramanik S, Bowes J, Jiang X, Dormer J, Denniff M, Antczak A, Szulinska M, Wise I, **Prestes PR**, Glyda M, Bogdanski P, Zukowska-Szczechowska E, Berzuini C, Woolf AS, Samani NJ, Charchar FJ, Tomaszewski M (2018). Molecular insights

into genome-wide association studies of chronic kidney disease-defining traits. *Nat Commun*, 9:4800.

5. Rowland J, Akbarov A, Eales J, Xu X, Dormer JP, Guo H, Denniff M, Jiang X, Ranjzad P, Nazgiewicz A, **Prestes PR**, Antczak A, Szulinska M, Wise IA, Zukowska-Szczechowska E, Bogdanski P, Woolf AS, Samani NJ, Charchar FJ, Tomaszewski M (2019). Uncovering genetic mechanisms of kidney aging through transcriptomics, genomics, and epigenomics. *Kidney Int*, 95:624-635.

6. Eales J, Maan A, Xu X, Michoel T, Hallast P, Batini C, Zadik D, **Prestes PR**, Molina E, Denniff M, Schroeder J, Bjorkegren J, Thompson J, Maffia P, Guzik T, Keavney B, Jobling M, Samani N, Charchar FJ, Tomaszewski M (2019). Human Y chromosome exerts pleiotropic effects on susceptibility to atherosclerosis. *Arterioscler Thromb Vasc Biol*, 39:2386-2401.

7. Eales JM, Maan AA, Xu X, Michoel T, Hallast P, Batini C, Zadik D, **Prestes PR**, Molina E, Denniff M, Schroeder J, Bjorkegren JLM, Thompson J, Maffia P, Guzik TJ, Keavney B, Jobling MA, Samani NJ, Charchar FJ, Tomaszewski M (2019). Human Y Chromosome Exerts Pleiotropic Effects on Susceptibility to Atherosclerosis. *Arterioscler Thromb Vasc Biol*. 39:2386-2401.

8. **Prestes PR**, Maier MC, Charchar FJ (2020). DNA copy number variations – do these big mutations have a big effect on cardiovascular risk? *Int J Cardiol*, 298:116-117.

9. **Prestes PR**, Maier MC, Woods BA, Charchar FJ (2020). A guide to the short, long and circular RNAs in hypertension and cardiovascular disease. *Int J Mol Sci*, 21:3666.

10. Jiang X, Eales JM, Scannali D, Nazgiewicz A, **Prestes PR**, Maier MC, Denniff M, Xu X, Saluja S, Cano-Gamez E, Wystrychowski W, Szulinska M, Antczak A, Byars S, Skrypnik D, Glyda M, Król R, Zywiec J, Zukowska-Szczechowska E, Burrell LM, Woolf AS, Greenstein A, Bogdanski P, Keavney B, Morris AP, Heagerty A, Williams B, Harrap SB, Trynka G, Samani NJ, Guzik TJ, Charchar FJ, Tomaszewski M (2020). Hypertension and renin-angiotensin system blockers are not associated with expression of Angiotensin Converting Enzyme 2 (ACE2) in the kidney. *European Heart Journal*, 41:4580-4588

11. Nakai ME, Denham J, **Prestes PR**, Eikelis N, Lambert EA, Straznicky NE, Schlaich MP, Esler MD, O'Brien BJ, Charchar FJ, Lambert GW, Marques FZ. (2021) Plasma lipocalin-2/NGAL is stable over 12 weeks and is not modulated by exercise or dieting. *Scientific Reports*, 11:4056

12. Eales JM, Jiang X, Xu X, Saluja S, Akbariv A, Cano-Gamez E, McNulty M, Finan C, Guo H, Wystrychowski W, Szulinska M, Thomas H, Pramanik S, Chopade S, **Prestes PR**, Wise I, Evangelou E, Salehi M, Shakanti Y, Ekholm M, Denniff M, Nazgiewicz A, Eichinger F, Godfrey B, Antczak A, Glyda M, Król R, Eyre S, Brown J, Berzuini C, Bowes J, Caulfield M, Zukowska-Szczechowska E, Zywiec J, Bogdanski P, Kretzler M, Woolf A, Talavera D, Keavney B, Maffia P, Guzik T, O'Keefe R, Trynka G, Samani N, Hingorani A, Sampson M, Morris A, Charchar

FJ, Tomaszewski M. Uncovering novel genetic mechanisms of hypertension through multi-omic analysis of the kidney. *Accepted to Nature Genetics*

Conference abstracts arising from my PhD candidature

1. **Prestes PR**, Marques FZ, Lopez-Campos G, Curl C, Lewandowski P, Delbridge LMD, Charchar FJ, Harrap SB. Investigation on the role of monogenic human cardiomyopathies in a polygenic model of cardiac hypertrophy and heart failure, 13th Asian-Pacific Congress of Hypertension 2017, Singapore, 6-8 October 2017. Oral presentation. Young Investigator Award winner.

2. **Prestes PR**, Marques FZ, Lopez-Campos G, Curl C, Lewandowski P, Delbridge LMD, Charchar FJ, Harrap SB. Investigation on the role of monogenic human cardiomyopathies in a polygenic model of cardiac hypertrophy and heart failure, International Society of Hypertension New Investigator Symposium, Singapore, 6 October 2017. Oral presentation

3. **Prestes PR**, Marques FZ, Lewandowski P, Delbridge LMD, Charchar FJ, Harrap SB. MicroRNA signatures and gene expression of genes involved in monogenic cardiomyopathy in a polygenic model. High Blood Pressure Research Society Annual Scientific Meeting 2017, Melbourne, Australia, 29 November -1 December 2017. Oral presentation.

4. **Prestes PR**. Circular RNAs *Hrcr* and *cTtn* are upregulated in the hypertrophic heart. Federation University Australia Higher Degree by Research Conference, Ballarat, Australia, 18 July 2018. Oral presentation, second place award.

5. **Prestes PR**, Harrap SB, Charchar FJ. Circular RNAs *Hrcr* and *cTtn* are upregulated in the hypertrophic heart independent of blood pressure. 13th GeneMappers, Noosa Heads, Australia, 29-31 August 2018. Poster presentation

6. **Prestes PR**, Marques FZ, Lopez-Campos G, Curl C, Lewandowski P, Delbridge LMD, Charchar FJ, Harrap SB. Investigation on the role of monogenic human cardiomyopathies in a polygenic model of cardiac hypertrophy and heart failure. Annual Scientific Meeting of the Cardiac Society of Australia and New Zealand, Brisbane, Australia, 2-5 August 2018. Poster presentation.

7. **Prestes PR**, Harrap SB, Charchar FJ. Circular RNAs Hrcr and cTtn are upregulated in the hypertrophic heart independent of blood pressure. International Society of Hypertension 2018, Beijing, China, 20-23 September 2018. Oral presentation. New Investigator Network, second place award.

8. **Prestes PR**, Harrap SB, Charchar FJ. Circular RNAs Hrcr and cTtn are upregulated in the hypertrophic heart independent of blood pressure. High Blood Pressure Research Society Annual Scientific Meeting 2018, Adelaide, Australia, 27-30 November 2018. Poster presentation.

Additional conference abstracts arising during my PhD candidature

1. Wise IA, **Prestes PR**, Tomaszewski M, Charchar, FJ. The role of kidney DNA methylation in hypertensive gene expression. Federation University Australia Higher Degree by Research Conference, Ballarat, Australia, 27 July 2017. Oral presentation

2. Wise IA, **Prestes PR**, Tomaszewski M, Charchar, FJ. (2017). The involvement of kidney DNA methylation in blood pressure regulation. International Society of Hypertension New Investigator Symposium, Singapore, 6 October 2017. Oral presentation.

3. Wise IA, **Prestes PR**, Tomaszewski M, Charchar, FJ. (2017). The involvement of kidney DNA methylation in blood pressure regulation. 13th Asian-Pacific Congress of Hypertension 2017, Singapore, 6-8 October 2017. Oral presentation.

4. Wise IA, **Prestes PR**, Tomaszewski M, Charchar, FJ. DNA methylation pattern specificity in the hypertensive. High Blood Pressure Research Society Annual Scientific Meeting 2017, Melbourne, Australia, 29 November -1 December 2017. Oral presentation.

5. De Vries N, **Prestes PR**, Byars S, Allen AM, Harrap SB, Charchar FJ. Angiotensin converting enzyme inhibitors protect against telomere shortening in spontaneously hypertensive rat. High Blood Pressure Research Society Annual Scientific Meeting 2017, Melbourne, Australia, 29 November -1 December 2017. Oral presentation.

6. Byars S, De Vries N, **Prestes PR**, Allen A, Balding D, Charchar F, Harrap S. Genetic reprogramming following short-term renin-angiotensin system inhibition in the spontaneously hypertensive rat. High Blood Pressure Research Society Annual Scientific Meeting 2017, Melbourne, Australia, 29 November -1 December 2017. Oral presentation.

7. Lee MGY, **Prestes PR**, Porello E, Lambert G, Konstantinov IE, Charchar FJ, d'Udekem Y. Coarctation of the aorta and late hypertension in survivors may be the consequence of common gene mutations. American Heart Association Scientific Sessions 2017, Anaheim, USA, 11-15 November 2017. Oral presentation.

8. De Vries N, **Prestes PR**, Raipuria M, Byars S, Allen AM, Harrap SB, Charchar FJ. Angiotensin converting enzyme inhibitors and angiotensin receptor blockers protect against telomere shortening in spontaneously hypertensive rat. International Society of Hypertension 2018, Beijing, China, 20-23 September 2018. Poster presentation.

9. **Prestes PR**, De Vries N, Byars S, Allen AM, Harrap SB, Charchar FJ. Circular RNA expression is reset by brief antihypertensive treatment in spontaneously hypertensive rats (SHR). International Society of Hypertension 2018, Beijing, China, 20-23 September 2018. Poster presentation.

10. **Prestes PR**, De Vries N, Byars S, Allen AM, Harrap SB, Charchar FJ. Circular RNA expression is reset by brief antihypertensive treatment in spontaneously

hypertensive rats (SHR). High Blood Pressure Research Society Annual Scientific Meeting 2018, Adelaide, Australia, 27-30 November 2018. Poster presentation.

11. Woods BA, **Prestes PR**, Maier MC, Xu X, Jiang X, Eales J, Tomaszewski M, Charchar FJ. Differentially expressed circular RNAs in hypertensive patients. 15th Asian Pacific Congress of Hypertension, Brisbane, Australia, 24-27 November 2019. Oral presentation. New Investigator Network award winner.

12. Woods BA, **Prestes PR**, Maier MC, Xu X, Jiang X, Eales J, Tomaszewski M, Charchar FJ. Silencing of circular RNAs circZNF91 and circRP11-298P3.4 alters expression of micro RNAs involved in hypertension. International Society of Hypertension 2020, Glasgow, Scotland. (Postponed due to COVID-19 restrictions) Oral presentation, New Investigator Network Award Session.

Awards

2017 – Winner Young Investigator award at the Asian Pacific Congress of Hypertension, in Singapore.

2018 – Second place at the Federation University Australia Higher Degree by Research Conference.

2018 – Second place New Investigator Award at the International Society of Hypertension Conference in Beijing, China.

2020 – Recipient of the Victoria Fellowship Award.

List of abbreviations

ABC	ATP-binding cassette
Acta1	actin alpha 1
Actc1	cardiac alpha actin 1
Actn2	actinin alpha 2
AGRF	Australian Genome Research Facility
AngII	angiotensin II
Arc	activity regulated cytoskeleton associate protein
BDNF	brain derived neurotrophic factor
BN	Brown Norway
BWA	Burrows-Wheeler Aligner
Calm1	calcium binding, calmodulin
Cav3	caveolin-3
Cdc42	cell division control protein 42 homologue
CH	cardiac hypertrophy
circRNA	circular RNA
Cm22	cardiac mass 22
CNVs	copy number variations
CPM	count-per-million
CRISPR-Cas9	clustered regularly interspaced short palindromic repeats/clustered regularly interspaced short palindromic repeat–associated 9
CRN	Collaborative Research Network
CSRP3	cysteine and glycine rich protein 3
cTtn	circular titin

CVD	cardiovascular disease
CWI	cardiac weight index
DAVID	Database for Annotation, Visualization and Integrated Discovery
Des	desmin
Dsg2	desmoglein 2
ECM	extracellular matrix
ET-1	endothelin 1
FC	fold change
FDR	false discovery rate
FHL1	four-and-a-half LIM domains 1
Fhl2	four-and-a-half LIM domains 2
Gapdh	glyceraldehyde 3-phosphate dehydrogenase
GATK	Genome Analysis Tool Kit
GEO	Gene Expression Omnibus
GO	Gene Ontology
GWAS	genome-wide associated studies
HCM	hypertrophic cardiomyopathy
HF	heart failure
HHR	hypertrophic heart rat
hiPSC	human-induced pluripotent stem cells
hiPSC-CM	hiPSC-derived cardiomyocytes
HPC	human primary cardiomyocytes
Hrcr	heart-related circular RNA
IGF1	insulin-like growth factor I
IGF1R	insulin-like growth factor I receptor

InDels	insertions and deletions
INO80D	INO80 complex subunit D
KCNB1	potassium voltage-gated channel subfamily B member 1
KEGG	Kyoto Encyclopedia of Genes and Genomes
lincRNA	long intergenic ncRNAs
LV	left ventricular
LVEDD	left ventricular end diastolic diameter
LVEF	left ventricular ejection fraction
LVH	left ventricular hypertrophy
miRNA	micro RNAs
mRNA	messenger RNA
MYBPC3	myosin-binding protein C3
MYH6	β-myosin heavy chain 6
MYH7	β-myosin heavy chain 7
MYH7B	myosin heavy chain 7B
MYL2	myosin light chain 2
MYL3	myosin light chain 3
MYOZ2	myozenin 2
ncRNA	non-coding RNAs
NELFA/Whsc2	negative elongation factor complex member A
NHMRC	National Health & Medical Research Council of Australia
NHR	normal heart rat
Nppa	natriuretic peptide A
PI3K	phosphoinositide 3-kinase
piRNA	piwi-interacting RNAs

Pkp2	plakophilin 2
PLN	phospholamban
pre-mRNA	precursor-mRNA
Pwwp2a	PWWP domain containing 2A
qPCR	quantitative PCR
QTL	quantitative trait locus
Rbm20	RNA binding motif protein 20
RhoA	Ras homologue gene family, member A
rRNA	ribosomal RNA
Ryr2	ryanodine receptor 2
SHR	spontaneously hypertensive rat
siRNA	small interfering RNAs
SLC10A7	solute carrier family 10 member 7
SLC23A2	solute carrier family 23 member 2
snoRNA	small nucleolar RNAs
SNPs	single nucleotide polymorphisms
snRNA	small nuclear RNAs
SVs	structural variants
TCAP	titin-cap
TMM	trimmed mean of M values
TNNI3	cardiac troponin I3
TNNT2	cardiac troponin T2
TPM1	tropomyosin 1
TREx	targeted RNA expression
Trim55	tripartite motif containing 55

Trim63	tripartite motif containing 63
tRNA	transfer RNAs
Ttn	titin
Ttr	transthyretin
TWF1	twinfilin-1
UTR	untranslated region
Vcl	vinculin

Chapter 1. Introduction

1.1 Cardiovascular disease

Cardiovascular disease (CVD) refers to a group of heart and blood vessels diseases, including the arteries. Coronary artery disease, stroke and cardiomyopathies – including left ventricular hypertrophy – are the most common causes of CVD related to morbidity and mortality worldwide (Lozano et al., 2012). Despite significant medical advances, CVD has remained the primary cause of death for decades (Lozano et al., 2012). Latest updates from the World Health Organization estimate that in 2016, CVD was the cause of approximately 17.9 million deaths worldwide, representing 31% of all deaths (World Health Organization, 2016). The burden of CVD is expected to increase in future decades as the population ages and comorbidities such as obesity and diabetes, continue to rise (Lozano et al., 2012).

1.2 Cardiac hypertrophy

Cardiac growth is part of normal heart development in childhood occurring alongside body growth and is influenced by hormones and the number of cardiomyocytes which is fixed soon after birth (Pfaffenberger et al., 2013). Cardiac mass is a polygenic trait with normal distribution in the population (Cain et al., 2009; de Jonge et al., 2011; Petersen et al., 2017). Normal values of left ventricular mass and cardiac chamber size are used to ascertain an individual's diagnosis of abnormal cardiac growth and hypertrophy (Fuchs et al., 2016).

Cardiac hypertrophy (CH) can be concentric, characterised by the increase in cardiac muscle mass from the thickening of the heart walls and the absence of a dilated

chamber; or eccentric, with the increase of the ventricle diameter with or without the increase in wall thickness (de Simone, 2004; Marian & Braunwald, 2017). However, the terminology to describe this condition can, to some extent, be unclear (Maron & Maron, 2013). Most commonly, the term hypertrophic cardiomyopathy is used to describe inherited, monogenic forms of left ventricular hypertrophy (LVH) sometimes also known as hypertrophic obstructive cardiomyopathy. LVH refers to the specific thickening of the left ventricular walls of the heart independent of its origins (Maron & Maron, 2013). In this thesis I will explore the contribution of genetic variation to the polygenic form of LVH and this will be the basis for my use of this term.

The Framingham Heart Study identified that LVH is strongly associated with higher risks of serious CVD, including arrhythmias, high blood pressure, coronary artery disease, obesity, heart failure (HF) and cardiovascular-related death (Tsao et al., 2015). Most importantly, the study highlighted that LVH is the most important indicator of risk of cardiovascular related death, being present in around 10% of the population without an identifiable cause (Levy, Garrison, Savage, Kannel, & Castelli, 1990; Post, Larson, Myers, Galderisi, & Levy, 1997; Tsao et al., 2015). The prolonged hypertrophic response caused by the enlargement and disorganisation of cardiac myocytes leads to a maladaptive decrease of the ventricular chamber size and a reduction in the cardiac output. This results in the deterioration of cardiac function leading to HF as a consequence of energy and functional imbalance (Figure 1.1) (Marian & Braunwald, 2017; Smith, 2017). The degenerative functional changes that decrease cardiac performance and subsequently lead to HF, can cause sudden cardiac death that is especially prominent in adolescents and young adults (Marian & Braunwald, 2017). However, LVH is usually asymptomatic or has overlapping features of HF with other cardiomyopathies, resulting in a lack of or incorrect diagnosis becoming degenerative before being discovered (Mosqueira, Smith, Bhagwan,

& Denning, 2019). Currently, diagnosis of cardiomyopathies is performed by screening studies, usually after the development of symptoms (McKenna & Judge, 2020). These characteristics highlight the importance of identifying at-risk individuals and implementing preventative measures before the onset of chronic heart disease and HF (Post et al., 1997; Smith, 2017).

1.3 Types of cardiac hypertrophy

CH may occur as an adaptation to physiological or pathological conditions depending on the origin of the stressor. The main differences between physiological and pathological CH include the underlying origins of the overloading stimuli and mechanisms involved, the overall cardiac phenotype and the long-term prognosis (Kavazis, 2015; McMullen & Jennings, 2007).

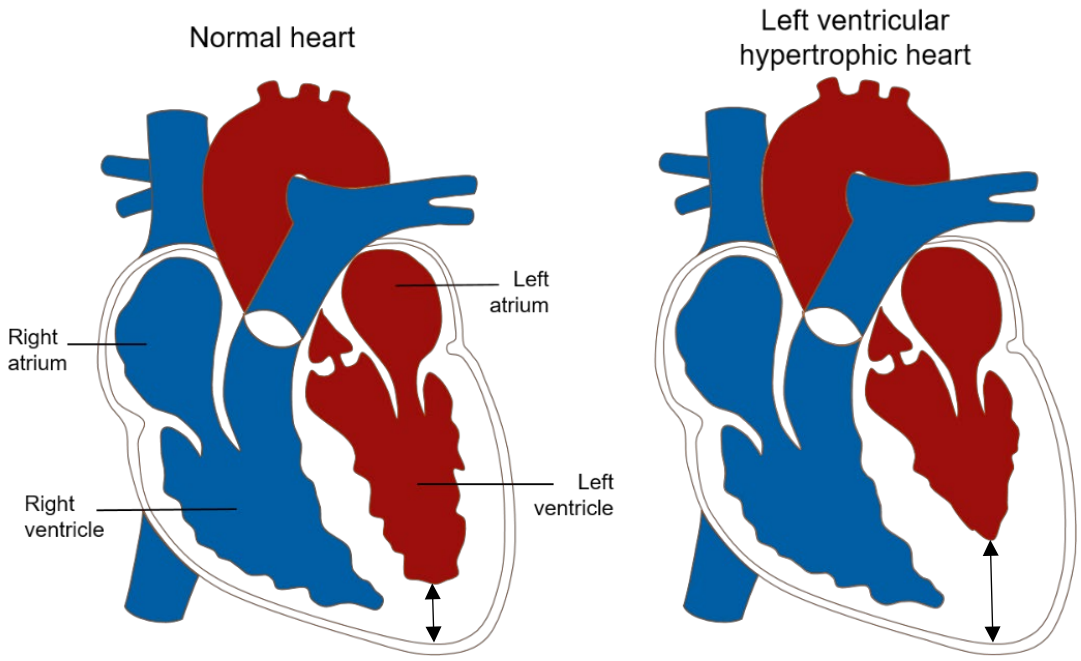

Figure 1.1: The normal heart and the hypertrophic heart. The normal heart (left) and the hypertrophic heart (right) with left (red) and right chambers (blue). The heart with left

ventricular hypertrophy presents a thickening of the left ventricular wall and a maladaptive decrease of the left ventricle represented by the arrows.

1.3.1 Physiological adaptation

CH due to physiological adaptation is characterised by the increase in cardiac size while retaining normal heart morphology and normal or improved heart function (Figure 1.2a) (McMullen & Jennings, 2007). It is represented by the adaptation of the heart usually initiated by normal development or normal overload and its progression depends on duration and intensity of the stressor. Importantly, physiological CH is transient and harmless, often reversed when the stressor is removed and it is not associated with HF or increased mortality, unlike pathological or genetic CH (Figure 1.2a) (Kavazis, 2015). Broadly, exercise training or pregnancy can induce CH in adults.

1.3.1.1 Exercise induced cardiac hypertrophy

Commonly known as athletes' heart, CH caused by exercise, is often observed in athletes and it has been proposed to enhance performance (McMullen & Jennings, 2007; Payne et al., 2006). Traditionally, athletes' heart is considered beneficial as the CH triggers increased aerobic capacity and eccentric cardiac remodelling providing a greater stroke volume and cardiac output in endurance athletes, such as long-distance runners, cyclists and swimmers (Figure 1.2a). However, concentric CH, associated with bursts of high blood pressure and pressure overload is observed to a lesser extent in strength training athletes such as weight lifters (Dorn, 2007). For example, insulin-like growth factor I (IGF1) is released as a response to exercise training and serum levels of IGF1 are elevated in professional athletes with no change to endothelin 1 (ET-1) or angiotensin II (AngII) pathways, normally associated with cardiovascular disease (McMullen & Jennings, 2007;

Shimizu & Minamino, 2016). IGF1 acts via the IGF1 receptor (IGF1R) activating phosphoinositide 3-kinase (PI3-K) (McMullen & Jennings, 2007). Transgenic mice with increased IGF1R/PI3-K activity developed CH but maintained normal life span and had normal or increased cardiac function, mimicking physiological CH and demonstrating the activation of distinct pathways when comparing physiological and pathological CH (McMullen & Jennings, 2007).

1.3.1.2 Pregnancy induced cardiac hypertrophy

During the second and third trimesters of pregnancy, the increased cardiac demand to match placental blood flow results in an enlarged heart, which gradually regresses to a normal size in the months post-partum (Figure 1.2a) (Dorn, 2007). In healthy women, it is estimated that, on average, the heart returns to its normal size 8 weeks post-partum (Maillet, van Berlo, & Molkentin, 2013). Interestingly, pathways of increased foetal gene expression, often associated with pathological CH are not present in pregnancy-induced hypertrophy in mice, reinforcing that pathways associated with pathological CH are not activated by physiological CH (Dorn, 2007).

Unfortunately, there are some rare cases when women develop peripartum cardiomyopathy following pregnancy, a type of dilated cardiomyopathy leading to HF indicating a possible change in cardiac gene expression profiles (Dorn, 2007). This condition has been associated with a mutation in the titin (*Ttn*) gene in 10-15% of affected women (Lee & Judge, 2017), highlighting possible genetic contributions to the development of pathological cardiomyopathies and HF.

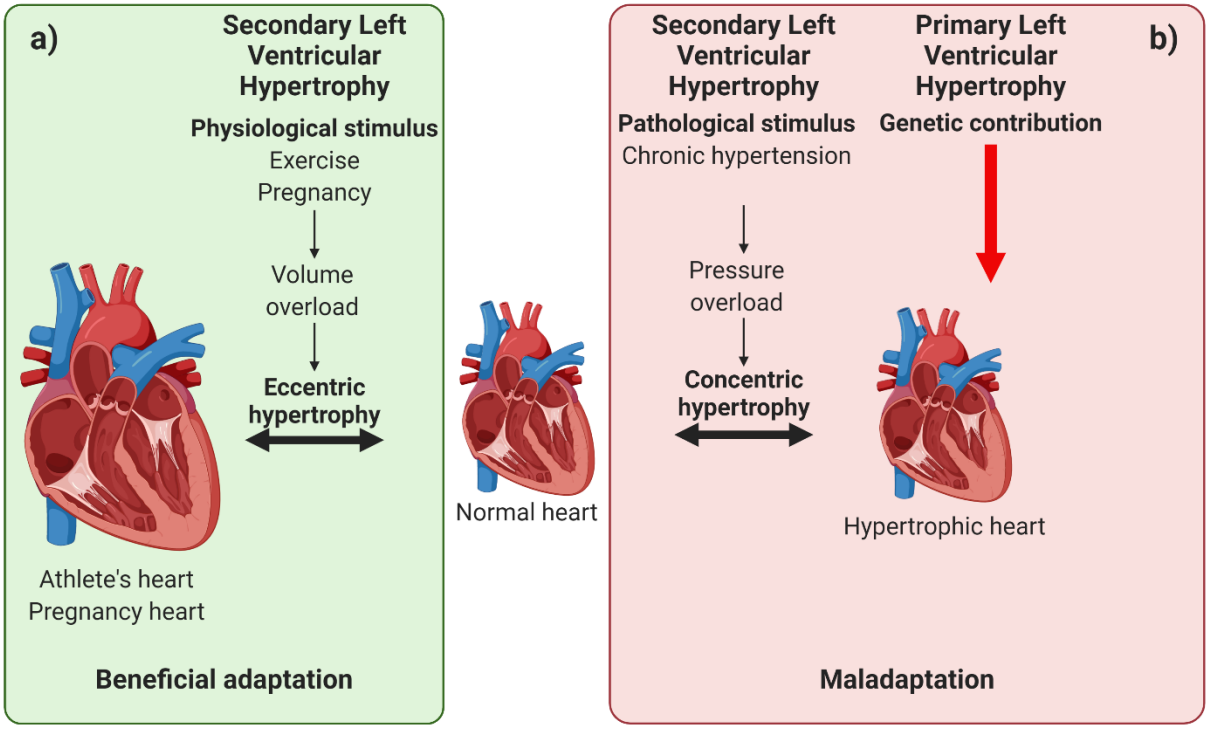

Figure 1.2: Normal heart response according to stimulus. A) Volume overload caused by physiological stimulus, i.e. exercise or pregnancy, leads to eccentric cardiac hypertrophy and beneficial adaptation of the heart. B) Pathological stimulus, such as chronic hypertension or genetic contribution, results in concentric cardiac hypertrophy and maladaptation of the heart, leading to an irreversible hypertrophic or dilated heart. The red arrow indicates the investigation undertaken in this research project. Created using BioRender.com.

1.3.2 Pathological adaptation

Pathological CH is typically characterised by the thickening of the heart walls which can lead to a decrease in the chamber size, resulting in concentric hypertrophy and a reduced capacity to pump blood around the body leading to a maladaptive response and disease (Figure 1.2b) (Shimizu & Minamino, 2016). The hypertrophy is the result of tissue growth by cell proliferation and enlargement of differentiated cardiomyocytes. As

mentioned above, CH caused by polygenic factors is most commonly recognised in the left ventricle, being called LVH and the focus of this thesis (Maron & Maron, 2013). Depending on the absence or presence of external stressors, LVH can be classified as primary or secondary, respectively, and certain risk factors such as age and weight can also increase the probability of developing LVH (Katholi & Couri, 2011).

1.3.2.1 Secondary CH: pathological adaptation to disease

Secondary CH is developed as a consequence of the presence of a primary pathological stimulus, such as chronic systemic hypertension, congenital heart disease and aortic valve stenosis (Figure 1.2b) which causes pressure overload in the heart and myocardial involvement as part of a multi-organ condition (Frey, Katus, Olson, & Hill, 2004). The majority of cases of heart hypertrophy are presented as secondary to uncontrolled high blood pressure (Katholi & Couri, 2011). Importantly, changes in life style, mainly weight loss, and the use of medication may decrease the extent of hypertrophy and thus decrease likelihood of a heart attack and sudden death (Katholi & Couri, 2011).

1.3.2.2 Primary CH: genetic contribution to disease

Primary CH is reported in the absence of pathological stressors such as mentioned above which would have otherwise been characterised as secondary CH (Figure 1.2b). Typically, the impairment burden is mainly restricted to the heart muscle (Maron et al., 2006). It is widely accepted that CH is caused by genetic factors, most commonly variants in genes responsible for the development of heart muscles, usually in protein-coding genes related to cardiac filaments and sarcomere assembly that affect proteins usually present in the left ventricle (Maron & Maron, 2013). Essentially, primary CH does not present pressure or volume overload of the heart muscle.

Hypertrophic cardiomyopathy is the most common inherited cardiovascular disease and the major contributor to HF and sudden death, being characterised by the asymmetrical enlargement of the heart walls with typical patterns of microscopic entanglement of cardiomyocyte architecture (Post et al., 1997; Tsao et al., 2015). However, more generally, cardiac hypertrophy from common causes often involves LVH. Recently, it was estimated that pathological LVH affects 1 in 200 people in the general population (Semsarian, Ingles, Maron, & Maron, 2015). Studies performed in specific groups describe that LVH occurs in 16% of European- and 43% of African-descendants and between 35% to 41% of hypertensive patients (Semsarian et al., 2015; Tsao et al., 2015). It is important to note that right ventricular hypertrophy is a less common condition, occurring with or without the presence of LVH and is often related to pulmonary complications, congenital heart disease or arrhythmogenic right ventricular cardiomyopathy (Reza, Musunuru, & Owens, 2019; van der Bruggen, Tedford, Handoko, van der Velden, & de Man, 2017).

1.4 Genetics of LVH

LVH is a common heart disease, highly inheritable, heterogeneous and with large genetic penetration. It has been described worldwide in over 50 countries, with no gender, ethnicity or racial bias and presenting similar phenotypes and genetic variants (Maron & Maron, 2013). However, studies report delayed diagnosis or under recognition in women (Olivotto et al., 2005) and African-American descendants (Maron et al., 2003).

Advances in molecular techniques, the increased understanding of genetic contributions to this disease and access to genetic testing have improved the ability to diagnose and estimate the prevalence of LVH. Recent studies suggest LVH occurs in at least 1 in 200 people in the general population but due to its manifestation occurring later

in life, this number is expected to be higher in older adults (Marian & Braunwald, 2017; Semsarian et al., 2015). However, this assessment is likely to be an underestimate as affected familial cases are excluded in such studies and many undiagnosed asymptomatic subjects are not detected (Maron & Maron, 2013; Semsarian et al., 2015). These estimates highlight that many genetic variants known to cause LVH are surprisingly common in the general population and do not have any known health consequences (Semsarian et al., 2015). Additionally, studies highlight that many undiagnosed cases remain asymptomatic throughout life without any detrimental outcomes and in other cases, LVH is only detected after a major cardiac event occurs, such as sudden death (Marian & Braunwald, 2017; Maron & Maron, 2013).

1.4.1 Monogenic LVH

Familial LVH is an autosomal dominant disorder with variable expression and penetrance and limited de novo mutations reported (Marian & Braunwald, 2017; Maron & Maron, 2013; Semsarian et al., 2015). There is debate about the proportion of the contribution of genetic variants to LVH, varying from 12% to 55% in patients with a diagnosed pathogenic mutation (Mosqueira et al., 2019; Semsarian et al., 2015; Walsh et al., 2017).

Mutations in one or more of over 20 genes encoding thin and thick myofilament elements related to sarcomere assembly and the adjacent Z-disc responsible for cardiomyocyte contractibility and function have been identified in patients with familial LVH (Maron & Maron, 2013; Mosqueira et al., 2019). However, individuals with the same mutation may have vastly different phenotypes and present varying degrees of cardiac impairment between and within families, which can make diagnosis on the basis of genotype challenging (Maron & Maron, 2013; Mosqueira et al., 2019).

Importantly, causative pathogenic genetic variants remain unknown in many LVH patients but would be extremely valuable in order to identify family members of diagnosed cases who may be at risk of developing the disease.

1.4.2 Polygenic LVH

The term polygenic LVH is commonly used to describe familial aggregation of LVH where single genes cannot explain the development of this condition, following a quantitative rather than qualitative effect. This is characterised by the contribution of many individual genetic factors, each with small effects, instead of the presence of specific mutations (Maron, 2002; Maron & Maron, 2013).

Despite scientific advances and the availability of genetic testing, pathogenic, likely pathogenic or non-synonymous DNA mutations account for only a portion of identified variants, whereas the remainder are benign or variants of uncertain significance (Semsarian et al., 2015). In polygenic LVH, because of the phenotypic heterogeneity observed between and within families, single DNA variants may not be the cause for the development of LVH, indicating the possible contribution of many individual genetic mutations and the role of other variants or modifier genes and environmental factors (Maron & Maron, 2013; Smith, 2017).

A study analysing a targeted 8-gene panel including genes cardiac α-actin 1 (*ACTC1*), myosin-binding protein C3 (*MYBPC3*), β-myosin heavy chain 7 (*MYH7*), myosin light chain 2 (*MYL2*), myosin light chain 3 (*MYL3*), cardiac troponin I3 (*TNNI3*), cardiac troponin T2 (*TNNT2*) and tropomyosin 1 (*TPM1*), identified 4% of subjects with multiple rare variants independent of their classification as missense, nonsense or silent (Burns, Bagnall, Lam, Semsarian, & Ingles, 2017). Interestingly, when the investigation included the comprehensive 45-gene cardiomyopathy panel, 17% of subjects presented

multiple rare variants. This study did not identify any subjects carrying more than one likely pathogenic or pathogenic variant (Burns et al., 2017). The authors also noted that patients with multiple variants were younger, with fewer comorbidities but more likely to have a family history of sudden cardiac death. They also had a worse prognosis of event-free survival from all-cause death, cardiac transplantation and cardiac arrest (Burns et al., 2017). However, other studies identified around 4% of patients with LVH have multiple pathogenic or likely pathogenic variants (Das, Ingles, Bagnall, & Semsarian, 2014; Lopes et al., 2015). These studies indicate the possible multigenic nature of some LVH cases with unidentified single pathogenic or likely pathogenic variants.

Furthermore, according to Mosqueira *et al*. (Mosqueira et al., 2019), finding a simple and single genetic cause for LVH is very difficult, as approximately 45% of patients have no known mutation and this complexity is especially pronounced in cases of non-familial LVH. Those unexplained cases with no mutations found and/or no other family members with LVH could represent a possible polygenic origin of the disease and demonstrate the likely involvement of epigenetics in pathogenic outcomes or the occurrence of unrecognised spontaneous mutations. However difficult these cases may be, they deserve attention in order to determine the genetic and epigenetic contribution to decrease the health burden and risk of sudden death, especially in young adults.

As mentioned previously, the Framingham Heart Study identified that approximately 10% of the population are affected by LVH with no known cause (Levy et al., 1990; Post et al., 1997; Tsao et al., 2015). Other studies also estimated the prevalence of LVH in the general population to be around 10% (Pedersen et al., 2020; Schirmer, Lunde, & Rasmussen, 1999) with coefficients for heritability of left ventricular mass of 0.69 for monozygotic and 0.32 for dizygotic twins and a pattern consistent with polygenic inheritance (Swan et al., 2003). These factors suggest that the 10% of unexplained cases

may be due to a polygenic effect. Furthermore, polygenic factors may also predispose and exacerbate the phenotype of those with secondary LVH leading to a sharper increase in left ventricular mass.

1.4.3 Genetic and epigenetic involvement in LVH

The genetic contributions to disease relate to changes in the DNA sequence and variations in the messenger RNA (mRNA) expression. Meanwhile, epigenetic traits are modifications that affect mRNA expression without changing the DNA sequence. They include DNA methylation, histone modification and differential expression of non-coding RNAs (ncRNA) (Egger, Liang, Aparicio, & Jones, 2004).

1.4.3.1 DNA variants

The mutation R403Q in the gene encoding *MYH7* was the first mutation associated with cardiomyopathies identified in the genetic involvement in LVH (Jarcho et al., 1989; Pare, Fraser, Pirozynski, Shanks, & Stubington, 1961). Since then, many other linkage studies identified several variants in sarcomeric and non-sarcomeric genes involved in LVH (Arnett et al., 2009; Cirino, 2008 [updated 2014]; Friedrich & Carrier, 2012; Maron & Maron, 2013). The advancement of molecular techniques, especially DNA sequencing, has improved our knowledge of specific genetic contributions to LVH with over 1400 variants identified, mostly missense, in more than 40 sarcomeric and non-sarcomeric genes (Maron & Maron, 2013; Mosqueira et al., 2019; Walsh et al., 2017).

Earlier studies used genome-wide associated studies (GWAS) to investigate variants associated with cardiac mass and LVH. One GWAS study identified a single nucleotide polymorphism in the gene encoding for potassium voltage-gated channel subfamily B member 1 (*KCNB1*) associated with left ventricular mass in a large human cohort and its potential involvement in LVH (Arnett et al., 2009). One recent large study

with almost 17000 participants using magnetic resonance imaging of left ventricles and microarrays, identified 14 loci associated with cardiac development, regulation of contractibility and abnormalities of the cardiovascular system. Eight of the 14 loci identified are unique, and 6 of those are novel, including 28 candidate genes related to left ventricular traits (Aung et al., 2019). Lastly, one meta-analysis study including over 46000 subjects from 30 studies investigated 16 traits of left ventricular structure and systolic and diastolic function. The authors describe five previously described loci and ten genome-wide relevant single nucleotide polymorphisms. These loci contained genes more frequently involved in signaling pathways or expressed in myocardial tissue amongst others (Wild et al., 2017). These findings improve our understanding of the genetic cardiac architecture and development, especially related to important left ventricle phenotypes and further functional studies may provide an insight into the development of potential novel therapeutic targets for cardiac remodelling (Aung et al., 2019; Wild et al., 2017).

Other studies investigated variants responsible for the structure of LVH phenotype. As an example of a causative genetic mutation identified, the variant R723G in gene *MYH7*, previously associated with LVH and located in the myosin head domain (Seebohm et al., 2009), was knocked-in in pigs to investigate its relevance to the development of LVH (Montag et al., 2018). Neonatal pigs developed LVH characteristics, including mild myocyte disorganisation, malformation of nuclei and overexpression of *MYH7*. Interestingly, all animals presented a change in the α/β myosin heavy chain ratio in the left ventricle, which likely caused death within 24h *post-partum* (Montag et al., 2018). Furthermore, three likely pathogenic variants were described in another gene in the myosin heavy chain family, *MYH7B*, encoding myosin heavy chain 7B (Chen et al., 2020). *Myh7b* knockout rats, with a truncated peptide missing a critical functional area, developed LVH at 12 weeks post-birth. These animals displayed differential expression of genes related to

calcium transport suggesting a link between *Myh7b* variants and the hypertrophic phenotype. This will be explored further in the next section (Chen et al., 2020).

Unfortunately, despite the discovery of many mutations associated with LVH, a recent study investigating the prevalence of genetic variants describes that around 45% of cases do not carry an identified mutation and of the 55% with an identified variant, 70% of those are found in genes encoding *MYH7* and *MYBPC3* (Figure 1.3) (Burke, Cook, Seidman, & Seidman, 2016; Mosqueira et al., 2019). In addition, variants in the genes encoding *TNNI3*, *TNNT2*, *TPM1* and a group of several other genes each account for 5% or fewer cases of LVH (Figure 1.3) (Marian, van Rooij, & Roberts, 2016; Maron & Maron, 2013; Mosqueira et al., 2019). Sequence variants in many other genes have been associated with LVH cases but their pathogenicity is still uncertain (Maron & Maron, 2013).

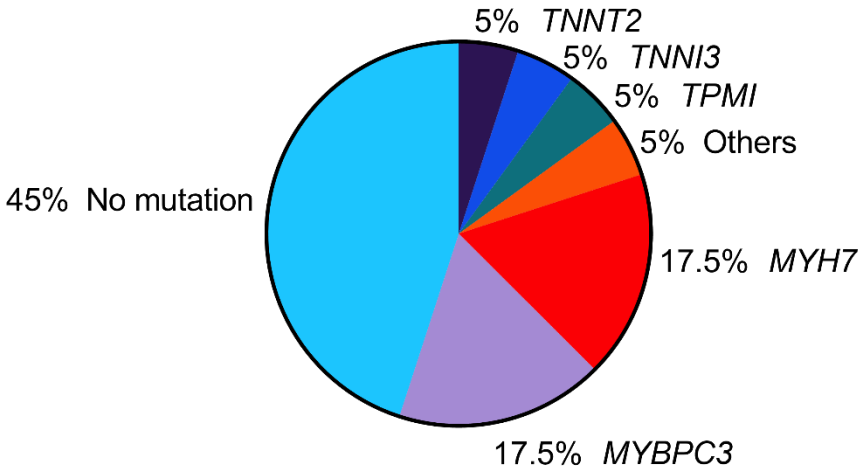

Figure 1.3: Prevalence of genetic variants in genes associated with cardiac hypertrophy. Legend: *MYH7*, β-myosin heavy chain 7; *MYBPC3*, myosin-binding protein C3; *TNNI3*, cardiac troponin I3; *TNNT2*, cardiac troponin T2; and *TPM1*, tropomyosin 1.

The prevalence of variants in genes associated with LVH is as relevant as the level of tolerance to such variations. Evidence of genetic variants causing LVH has been established in nine sarcomeric genes, encoding *ACTC1*, cysteine and glycine rich protein 3 (*CSRP3*), *MYBPC3*, *MYH7*, *MYL2*, *MYL3*, *TNNI3*, *TNNT2* and *TPM1* (Table 1.1) (Marian & Braunwald, 2017; Maron, Maron, & Semsarian, 2012b). However, these genes may have different levels of tolerance to variation of the genetic sequence. For instance, genes may be more or less impacted by a genetic variant and therefore have a lower or higher level of tolerance to a genetic variant, respectively, which may affect translation and influence phenotype (Marian & Braunwald, 2017). Linkage studies and co-segregation analysis have strongly supported the evidence of causality of genetic variants, especially in *MYH7* and *MYBPC3* (Marian & Braunwald, 2017). However, these genes along with *TNNT2*, *TNNI3*, and *TPM1* seem to be highly intolerant to missense or loss-of-function variants, but not all variants observed lead to LVH. In contrast, *ACTC1* is extremely intolerant to mutations and such genetic variants usually cause LVH (Marian & Braunwald, 2017). Furthermore, the involvement of genetic variants described in small families likely to cause LVH were identified in genes encoding four-and-a-half LIM domains 1 (*FHL1*), myozenin 2 (*MYOZ2*), phospholamban (*PLN*), titin-cap (*TCAP*), tripartite motif containing 63 (*TRIM63*) and titin (*TTN*) (Table 1.1) (Marian & Braunwald, 2017; Maron & Maron, 2013; Maron et al., 2012b). In addition, variants in a further 11 genes associated with LVH were identified in small families or have only been described as sporadic cases (Table 1.1) (Marian & Braunwald, 2017; Maron et al., 2012b).

A recent study by Walsh *et al* (Walsh et al., 2017) argues that there is no evidence that DNA variants in non-sarcomeric genes (see Table 1.1) or rare variants in sarcomeric or other heart related genes are pathogenic. The authors claim that some genes may have been incorrectly implicated in LVH diagnosis or present false positive association to the

disease in clinical genetic testing. Importantly, even though most variants identified are limited to individual or small families and may not improve LVH diagnosis in the population, they can assist in identifying asymptomatic family members at risk of developing an enlarged heart and their characterisation could avoid fatal events such as sudden death (Maron & Maron, 2013). This and other studies highlight that the cause of most LVH cases is still unknown (Marian & Braunwald, 2017; Maron & Maron, 2013; Maron, Rowin, Casey, & Maron, 2016; Mosqueira et al., 2019; Walsh et al., 2017).

Arguably, the initial testing of a targeted subset of eight genes may be sufficient to assess impact on clinical phenotype. It has been reported that an increase in the number of genes investigated from the targeted 8-cardiac gene panel to a comprehensive 45-cardiac gene panel also markedly increased the number of uninformative rare variants identified with unknown pathogenicity (Burns et al., 2017). However, expanding the number of genes investigated in the screening may be beneficial if the initial testing was inconclusive (Burns et al., 2017).

Table 1.1: Genes associated with left ventricular hypertrophy according to their level of involvement. The Z score represents the degree of deviation from the mean. In this case, a higher positive Z score indicates a gene with higher intolerance to a missense mutation. The probability of intolerance (pLI) is represented from 0 to 1 where a higher value indicates higher intolerance to loss of function (LoF). Number (n) of clinically relevant single nucleotide polymorphisms according to ClinVar database on 29th July 2020 (Landrum et al., 2018). "All" single nucleotide polymorphisms include pathogenic, likely pathogenic, variants of uncertain significance, likely benign, benign and variants with conflicting interpretations. Table adapted from Marian and Braunwald, 2017 (Marian & Braunwald, 2017).

Official gene symbol	**Protein**	**Characteristic**	**Location**	**Tolerance to variation**		**Single nucleotide polymorphism (n)**		
				Missense (*Z* Score)	**LoF (pLI)**	**Pathogenic**	**Likely pathogenic**	**All**
Established causal genes for left ventricular hypertrophy (large families)								
ACTC1	Cardiac α-actin 1	Actomyosin interaction	Sarcomere (thin filament)	5.25	0.95	6	9	366
CSRP3	Cysteine and glycine rich protein 3	Muscle LIM protein (MLP), a Z disk protein	Z-disc	−0.66	0	2	6	195

MYBPC3	Myosin-binding protein C3	Cardiac contraction	Intermediate filament	0.69	0	203	133	1,708
MYH7	β-Myosin heavy chain 7	ATPase activity, force generation	Sarcomere (thick filament)	6.54	0	118	203	2,291
MYL2	Myosin light chain 2	β-Myosin heavy chain 7–binding protein	Sarcomere (thick filament)	0.86	0.02	7	15	246
MYL3	Myosin light chain 3	β-Myosin heavy chain 7–binding protein	Sarcomere (thick filament)	0.75	0.89	3	9	189
TNNI3	Cardiac troponin I3	Inhibitor of actomyosin interaction	Sarcomere (thin filament)	1.88	0.17	32	32	362
TNNT2	Cardiac troponin T2	Regulator of actomyosin interaction	Sarcomere (thin filament)	1.54	0.01	29	38	448

TPM1	Tropomyosin 1	Places the troponin complex on cardiac actin	Sarcomere (thin filament)	3.42	0.8	12	30	429
Likely causal genes for left ventricular hypertrophy (small families)								
FHL1	Four-and-a-half LIM domains 1	Muscle development and hypertrophy	Z-disc	1.29	0.92	21	6	165
PLN	Phospholamban	Regulator of sarcoplasmic reticulum calcium	Calcium handling	0.57	0.11	6	1	73
TCAP	Titin-cap	Titin capping protein	Z-disc	0.45	0.08	3	5	135
TRIM63	Tripartite motif containing 63	E3 ligase of proteasome ubiquitin system	Other	0.02	0	0	0	12
TTN	Titin	Sarcomere function	Sarcomere (thick filament)	−5.48	0	158	500	13,081
MYOZ2	Myozenin 2	Z disk protein	Z-disc	0.03	0.02	1	0	103

Genes associated with left ventricular hypertrophy (small families and sporadic cases)

ACTN2	Actinin α2	Z disk protein	Z-disc	1.76	1	9	5	640
ANKRD1	Ankyrin repeat domain 1	Negative regulator of cardiac genes	Other	−0.01	0	0	3	146
CASQ2	Calsequestrin 2	Calcium-binding protein	Calcium handling	−1.08	0	12	12	256
CAV3	Caveolin 3	Caveolae protein	Calcium handling	1.19	0.34	11	4	213
JPH2	Junctophilin 2	Intracellular calcium signaling	Calcium handling	3.93	0.01	3	1	256
LDB3	LIM domain binding 3	Z disk protein	Z-disc	0.32	0	3	2	544
MYH6	Myosin heavy chain α	Expressed at low levels in adult human heart	Sarcomere (thick filament)	2.87	0	4	5	896

MYLK2	Myosin light chain kinase 2	Phosphorylate myosin light chain 2	Other	0.73	0.22	0	1	214
NEXN	Nexilin F-actin binding protein	Z disc protein	Z-disc	−1.32	0	2	5	234
TNNC1	Cardiac troponin C1	Calcium-sensitive regulator of myofilament function	Sarcomere (thin filament)	2.22	0.51	3	6	125
VCL	Vinculin	Z disk protein	Z-disc	3.11	0.99	0	0	489

1.4.3.2 mRNA expression and pathways

The contribution of genetic variants to the development of LVH is well established, as outlined above. However, changes in gene expression profiles and activation of signalling pathways have also been associated with LVH (Marian & Braunwald, 2017). DNA variants and mRNA expression are strongly interlinked and often a mutation in the DNA sequence in the form of a genetic variant (especially nonsense variants) will influence the mRNA expression of its gene (Figure 1.4) (Chen et al., 2020). However, the presence of DNA variants is not the exclusive source of changes in the mRNA profile of patients with LVH. Differential gene expression profiles have also been observed independently of the presence of genetic variants suggesting that epigenetics could be involved in these cases (Liu & Tang, 2019). Many global gene expression studies using microarrays and large-scale transcriptomic analyses highlight that the number of differentially expressed genes in LVH patients largely exceeds the number of genes associated with LVH (Gao, Collyer, Wang, Sun, & Xu, 2020; Rajan et al., 2006; Ren et al., 2016). Furthermore, transcriptomic studies may uncover differentially expressed genes related to LVH and warrant investigations into genetic variants leading to the discovery of novel candidate genes (Gao et al., 2020).

Independent of the origin of the differential gene expression associated with LVH, studies using large-scale genomic data investigating several genes, especially using microarrays and transcriptome sequencing, have studied genes with differential mRNA expression. These genes can be used to identify enriched biological pathways and terms related to gene ontology, biological processes and molecular functions to provide an insight into the mechanistic causes of a disease (Reimand et al., 2019).

As mentioned in the previous section, *Myh7b* knockout rats showed 22 upregulated and 21 downregulated genes (Chen et al., 2020). The Kyoto Encyclopedia of Genes and Genomes

(KEGG) pathway analysis revealed the top enriched terms in the upregulated genes were chemical homeostasis, regulation of blood circulation, cell adhesion molecule binding, and calcium ion binding. On the other hand, cytokine-receptor interaction, neuroactive ligand receptor interaction, calcium signalling pathway, and leukocyte trans-endothelial migration pathways were enriched in the downregulated genes. Interestingly, calcium-related pathways were identified in both up- and downregulated genes suggesting the triggering of a calcium-related response and its involvement in LVH (Chen et al., 2020). This study demonstrates the downstream effects that one genetic variant may have in many other genes, not only the gene carrying the original variant.

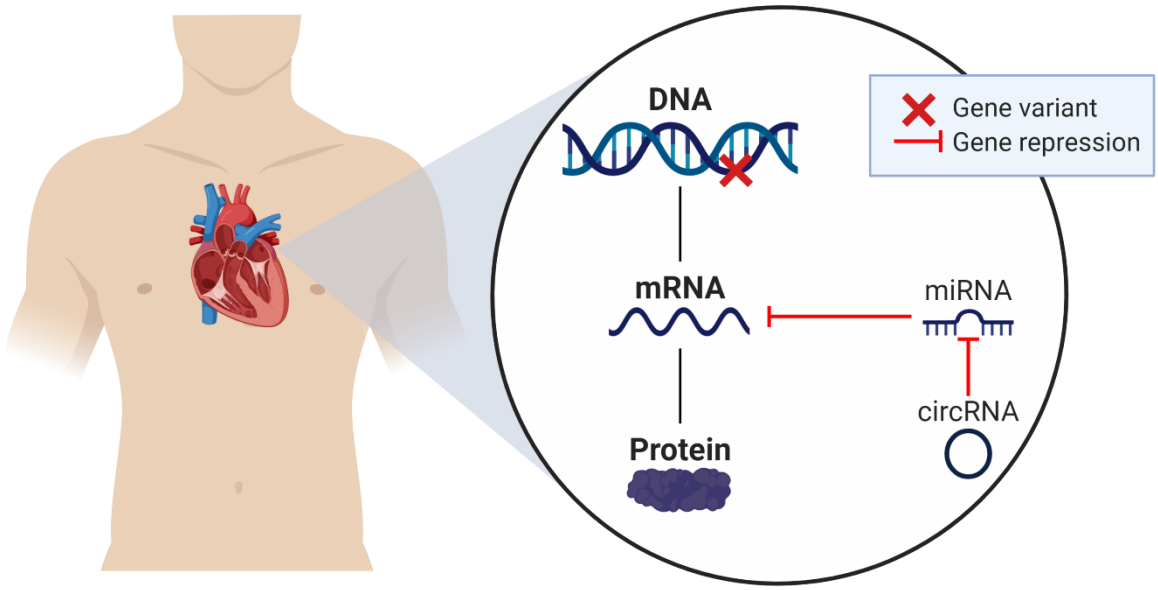

Figure 1.4: Simplified genetic and epigenetic involvement in left ventricular hypertrophy. The cross indicates a variant in the DNA sequence, which may affect mRNA expression. Genetic and epigenetic changes, including the differential expression of circular RNAs and micro RNAs, may also affect mRNA expression via repression.

The comprehensive list of DNA variants and differentially expressed genes in LVH is still largely unknown. However, there are examples of DNA variants influencing mRNA expression in LVH. Conversely, a study that identified a patient with the variant R723G in gene *MYH7*, known to cause LVH, found unchanged *MYH7* mRNA levels. This demonstrated that DNA variants are not always responsible for changes in gene expression levels (Borchert, Tripathi, Francino, Navarro-Lopez, & Kraft, 2010).

A comparison between the genetic pathways activated in pathological and physiological CH in mice to investigate transcriptome signatures using RNA sequencing identified 2014 differentially expressed genes in pathological CH, and only 245 such genes in physiological CH (Song, Hong, Kim, & Kim, 2012). The highest scoring network included genes responsible for cell cycle regulation. In pathological CH, KEGG analysis demonstrated that the top enriched pathways with downregulated genes were responsible for muscle contraction and metabolism, specifically including pathways involved in cardiac muscle contraction, oxidative phosphorylation and fatty acid metabolism; meanwhile, genes upregulated were involved in pathways associated with immune response and cell cycle, including p53 signalling pathway, extracellular matrix-receptor interaction, cell cycle and leucocyte migration (Song et al., 2012).

Another study used microarray analysis to investigate pathological and physiological CH and identified over 1650 genes differentially expressed in common between the two networks, with pathological and physiological CH displaying 1980 and 1502 differentially expressed genes, respectively (Drozdov et al., 2013). The expression networks were phenotype specific where pathological CH networks showed enrichment levels greater than physiological CH and were strongly related to negative outcomes. Pathways related to abnormal cardiovascular system morphology and physiology, pre-natal lethality and abnormal blood vessel morphology were highly enriched pathways in pathological CH (Drozdov et al., 2013).

Furthermore, the authors also suggested the presence of strong rewiring of genes differentially expressed between the two networks involved in several cellular pathways, especially related to extracellular matrix arrangements (Drozdov et al., 2013).

These are examples of a clear distinction in transcriptome signatures between pathological and physiological CH which demonstrate critical differences and the importance of characterising and investigating these conditions separately.

Many studies focusing on the transcriptome profiling of pathological LVH have identified an important and significant involvement of calcium related pathways indicating a crucial involvement in the development of LVH (Chen et al., 2020; Farrell et al., 2018; Gao et al., 2020; Kraft, Montag, Radocaj, & Brenner, 2016; Sanoudou, Vafiadaki, Arvanitis, Kranias, & Kontrogianni-Konstantopoulos, 2005).

1.4.3.3 non-coding RNAs

For many decades, ncRNA were considered part of the junk genome as albeit being transcribed from DNA, are not translated into protein. Traditionally, the main roles of ncRNAs were seen to be the messengers between the DNA sequence and the proteins. ncRNAs could act as transfer RNAs (tRNA), help decode the mRNA into protein, as ribosomal RNA (rRNA) forming ribosomes, small nuclear RNAs (snRNA) regulating RNA splicing and small nucleolar RNAs (snoRNA) acting as rRNAs modifiers (Mattick & Makunin, 2006; Raina & Ibba, 2014; Yan, Zhu, Guang, & Feng, 2019). Recently, ncRNAs groups were expanded to include short ncRNAs, such as micro RNAs (miRNA), small interfering RNAs (siRNA) and piwi-interacting RNAs (piRNA) and long ncRNAs such as long intergenic ncRNAs (lincRNA) and circular RNAs (circRNAs) (Figure 1.5) (Kung, Colognori, & Lee, 2013; Mattick & Makunin, 2006). NcRNAs are usually described as stage-, tissue- and species-specific, involved in the regulation of mRNA expression transcriptionally, post-transcriptionally and translationally and have been

described as potential biomarkers and therapeutic targets for many complex diseases (Greco & Condorelli, 2015; Kovacova, Poprach, Buchler, Cho, & Slaby, 2018).

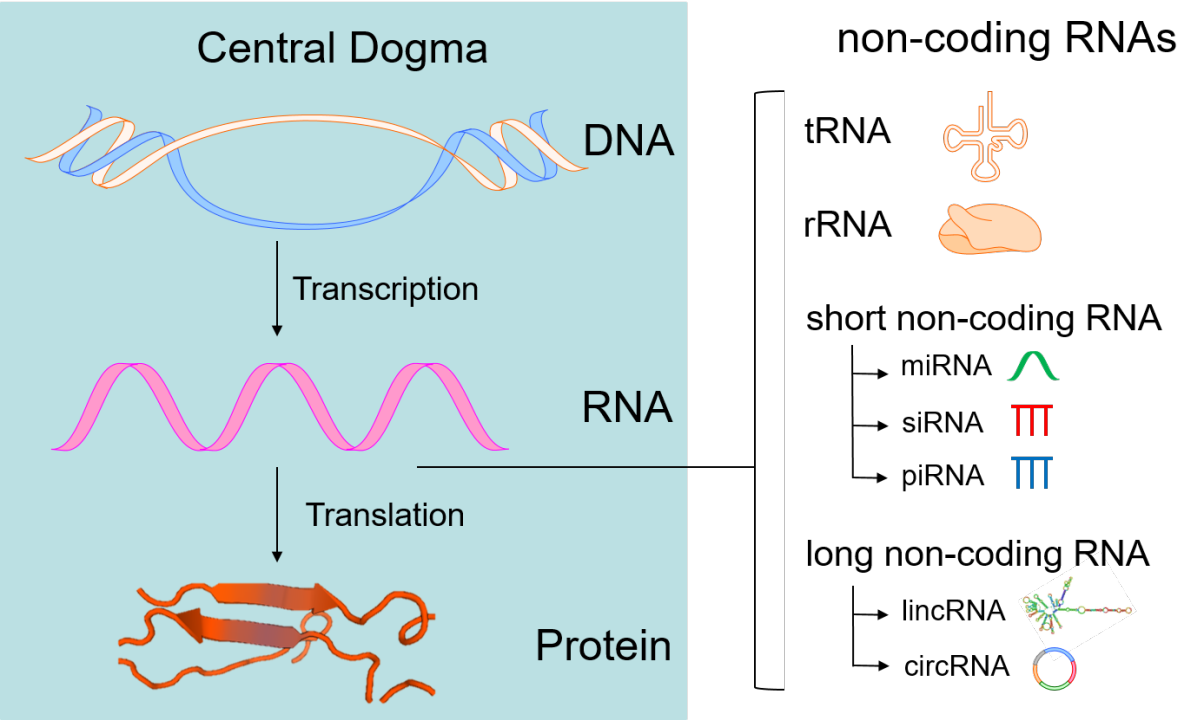

Figure 1.5: The central dogma of molecular biology and the RNA family. According to the central dogma, DNA is transcribed into RNA and translated into protein. However, non-coding RNAs (ncRNA) are not translated, being involved in protein synthesis (transporter RNA, tRNA and ribosomal RNA, rRNA) or responsible for regulation of gene expression (short or long ncRNA). Legend: miRNA - microRNA, siRNA - small interfering RNA, piRNA - piwi-interacting RNA, lincRNA - long intergenic ncRNA, circRNA - circular RNA.

1.4.3.3.1 Micro RNAs

miRNAs are endogenous, single-stranded, small regulatory RNAs approximately 21-25 nucleotides long that regulate gene expression by targeting the degradation of selective

mRNA when there is perfect matching of mRNA-miRNA sequences; or by blocking the translation of selective mRNA through the recognition of similar but not identical sequences (Figure 1.6) (Chen & Rajewsky, 2007; Gangwar, Rajagopalan, Natarajan, & Deiuliis, 2018). According to miRBase release 22.1, a miRNA database, there are over 1900 miRNA sequences described in humans and almost 500 in rat and each miRNA can target hundreds of mRNAs due to the presence of numerous binding sites throughout the genome (Kozomara, Birgaoanu, & Griffiths-Jones, 2019; Kozomara & Griffiths-Jones, 2014). Therefore, an increase in the levels of miRNA is expected to cause a decrease of its target mRNA expression (Greco & Condorelli, 2015; Kozomara & Griffiths-Jones, 2014; Mattick & Makunin, 2006).

Several miRNAs have been described as key regulators of cardiovascular biology, positively or negatively modulating mRNA levels and therefore affecting cardiac development and remodelling, endothelial function, lipid metabolism, LVH and post-infarction dysrhythmias via post-transcriptional changes (Chen et al., 2017; Ordovas & Smith, 2010; Zhao, 2018).

miRNA-1 has been identified as a major protector of heart structure and function by targeting and decreasing calcium signalling dependent LVH (Chen et al., 2017). In one study, the deletion of miRNA-1 lead to many cardiac defects and its downregulation was observed as early as day 1 during a 14 day pressure overload study (Ottaviani & da Costa Martins, 2017). The administration of miRNA-1 via injection in a rat model of LVH caused by pressure overload attenuated LVH, reduced myocardial fibrosis and apoptosis and improved calcium signalling, suggesting an important role in cardiac remodelling (Karakikes et al., 2013). Interestingly, both miRNA-1 and miRNA-133 are downregulated in human LVH and three murine models of severe LVH (Care et al., 2007).

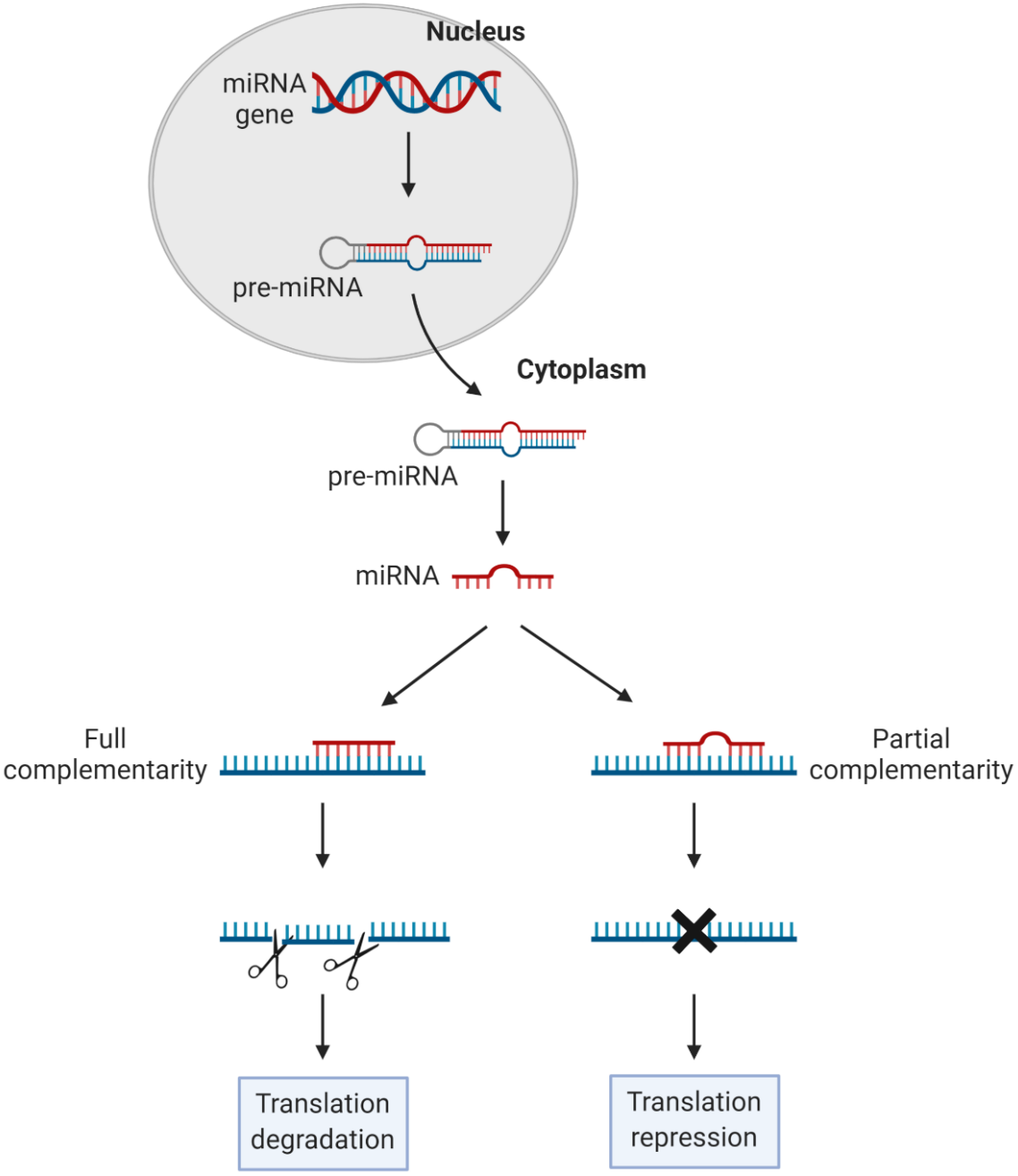

Figure 1.6: Regulation of mRNA by miRNA post-transcription. Pre-miRNA is exported from the nucleus into the cytoplasm and is processed into a mature miRNA. Full complementarity of the miRNA-mRNA complex results in the cleavage of the mRNA and translation degradation. Partial complementarity between miRNA and mRNA results in repression of translation. Created with BioRender.com.

miRNA-133 targets three proteins involved in LVH, Ras homologue gene family, member A (*RhoA*), cell division control protein 42 homologue (*Cdc42*) and negative elongation factor complex member A (*NELFA/Whsc2*) (Care et al., 2007). However, the overexpression of miRNA-133 protected against cardiac fibrosis but did not alter cardiomyocyte growth (Matkovich et al., 2010).

Moreover, miRNA-1, together with miRNA-133, miRNA-26b, miRNA-208b, miRNA-499, miRNA-21 and miRNA-672 presented distinctive expression patterns associated with LVH and have been proposed, amongst others, as prognostic markers and therapeutic targets for reduction of the hypertrophic phenotype and detrimental cardiac remodelling (Kontaraki et al., 2015; Lu & Wu, 2018).

Several studies suggest the upregulation of miRNA-195, miRNA-208, miRNA-22, miRNA-499 and others in LVH results in cardiac remodelling and HF. Some miRNAs have been identified in LVH but their specific roles have not been widely explored, while others have been described as pivotal to cardiac development and disease onset and progression (Greco & Condorelli, 2015; Ottaviani & da Costa Martins, 2017).

1.4.3.3.2 Circular RNAs

CircRNAs, an underexplored group of long ncRNAs, originate when precursor-mRNA (pre-mRNA) undergo back-splicing instead of RNA splicing. The back-splicing process promotes the circularization of pre-mRNA linking their 3' and 5' ends (Figure 1.7). This circularization protects them from degradation from most RNAses, unlike their coding counterparts and it has been proposed that circRNAs have an average half-life of 48h, around five times greater than mRNAs (Qu et al., 2015).

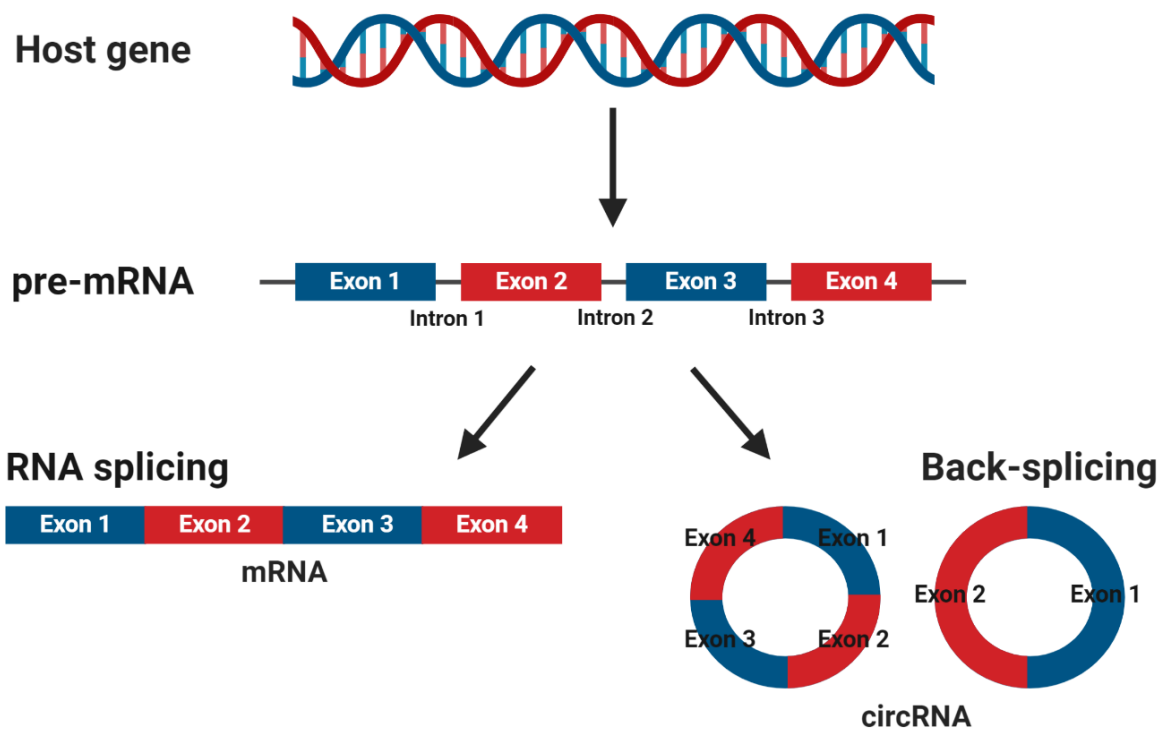

Figure 1.7: Biogenesis of messenger RNA (mRNA) and circular RNA (circRNA) from precursor-mRNA (pre-mRNA). Pre-mRNA can undergo RNA splicing producing mRNA or back-splicing generating circRNA. Each pre-mRNA can produce one or more circRNAs. Created with BioRender.com.

Most circRNAs consist of one or more protein-coding exons but limited evidence suggests that some may also retain introns (Qu et al., 2015). Importantly, some genes produce more than one circRNA (Figure 1.7) and their highly heterogeneous group of transcripts make it difficult to determine their precise function (Barrett & Salzman, 2016; Devaux et al., 2017). However, evidence suggests that circRNAs can regulate mRNA expression via two mechanisms. Firstly, directly, when circRNAs compete with mRNAs for the splicing machinery (Starke et al., 2015); and secondly, indirectly, targeting and sponging miRNAs

preventing their normal function as post-transcriptional regulators (Figure 1.8) (Chen, Lu, Yang, & Xing, 2019; Salzman, 2016).

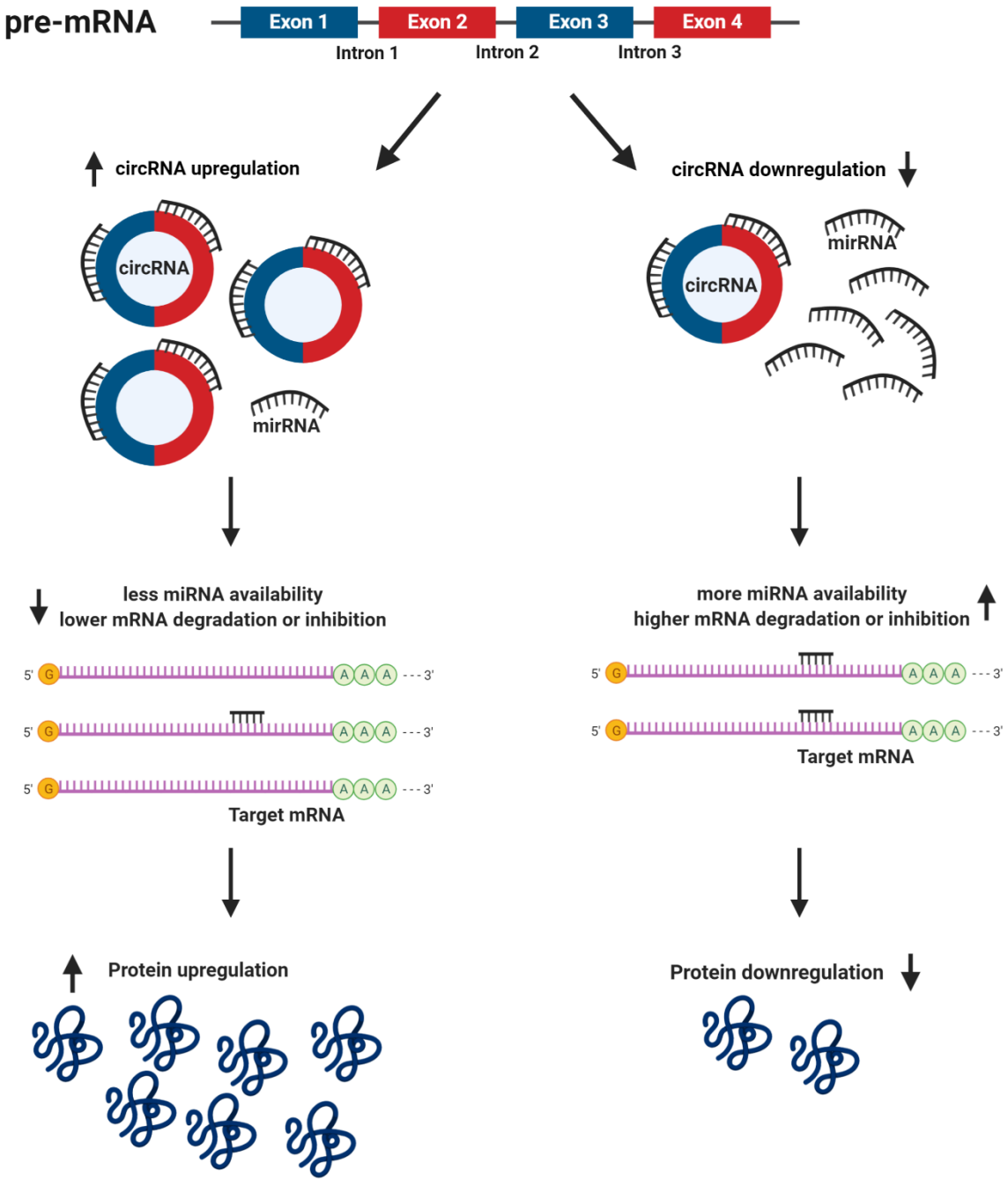

Figure 1.8: circRNA-miRNA-mRNA regulation. The increase of circular RNA (circRNA) sequesters more micro RNAs (miRNAs), decreasing availability and lowering messenger RNA

(mRNA) degradation or inhibition resulting the upregulation of protein levels. Conversely, the downregulation of circRNA results in more miRNA availability and higher levels of mRNA degradation or inhibition leading to the downregulation of protein expression. Created with BioRender.com.

Only recently, the involvement of circRNAs has been described in cardiovascular disease, in an array of conditions and more specifically in LVH (Fan et al., 2017; Li, Ding, et al., 2018; Tan et al., 2017). Currently, a small number of studies have explored the role of circRNAs in LVH and their function and involvement as regulators of disease remain largely unknown (Ottaviani & da Costa Martins, 2017). Thus far, heart-related circular RNA (*Hrcr*) is the most promising circRNA related to pathological LVH and HF. It is upregulated in pathological LVH induced with isoproterenol in mice and acts as a sponge, downregulating miRNA-223 expression, which in turn upregulates the gene for activity regulated cytoskeleton associated protein (*Arc*) mRNA expression attenuating hypertrophy (Wang et al., 2016).

Furthermore, circSLC8A1 and circRNA-000203 are also upregulated in induced LVH, in the heart of patients with pressure overload induced LVH and angiotensin-II infused mice, respectively. However, while the possible targets for circSLC8A1 have not been established, circRNA-000203 downregulated miRNA-26b-5p and miRNA-140-3p, which target *Gata4* levels (Li et al., 2020; Lim et al., 2019). circSLC8A1, along with circCACNA1D, circSPHKAP and circALPK2 are differentially expressed during cardiac development and show human heart specific enrichment in foetal tissues (Lei et al., 2018). These studies demonstrate that circSLC8A1 may be a promising biomarker of LVH or a possible therapeutic target to be explored (Lei et al., 2018; Lim et al., 2019; Tan et al., 2017).

Finally, over 5600 circRNAs have been identified in human-induced pluripotent stem cells (hiPSC) and hiPSC-derived cardiomyocytes (hiPSC-CM) (Lei et al., 2018). This warrants further investigation and reinforces the potential of circRNAs in pathological LVH as an emerging field.

1.5 The hypertrophic heart rat (HHR), a unique animal model of polygenic cardiac hypertrophy

Animal models of heart disease are relevant to study LVH as the remodelling of cardiac tissue demonstrates good agreement with LVH in humans (Lorell & Carabello, 2000). Commonly, studies into the genetics of LVH use experimental models of pathological hypertrophy induced by either pressure or volume overload using mini-pump infusions of vasoactive substances (Lorell & Carabello, 2000).

In order to investigate polygenic LVH, I used the hypertrophic heart rat (HHR) – a unique polygenic, normotensive model of spontaneous left ventricular hypertrophy, cardiac failure and premature death and its genetic control strain, the normal heart rat (NHR) – a normotensive rat with a normal size heart. These strains were developed through selectively crossbreeding the spontaneously hypertensive rat (SHR) which has spontaneous high blood pressure and secondary left ventricular hypertrophy with Fischer 344 rats which have smaller hearts and normal blood pressure (Harrap et al., 2002). The progenitors of the HHR and NHR were animals with big hearts and normal blood pressure; and animals with normal heart size and normal blood pressure, respectively. The progenies have been inbred for over 60 generations to ensure the establishment of these new strains for studies on LVH independent of confounding factors such as high blood pressure (Harrap et al., 2002).

Many studies have established that the HHR develops primary LVH characterised by the gross enlargement of cardiomyocytes in the adult animals (Bell, Porrello, Huggins, Harrap, & Delbridge, 2008; Dwyer et al., 2008; Harrap et al., 2002; Porrello, Bell, et al., 2009; Porrello, D'Amore, et al., 2009). Interestingly, neonatal HHR have smaller hearts with fewer and smaller cardiomyocytes when compared to the NHR (Figure 1.9) (Porrello, Bell, et al., 2009; Prestes et al., 2018).

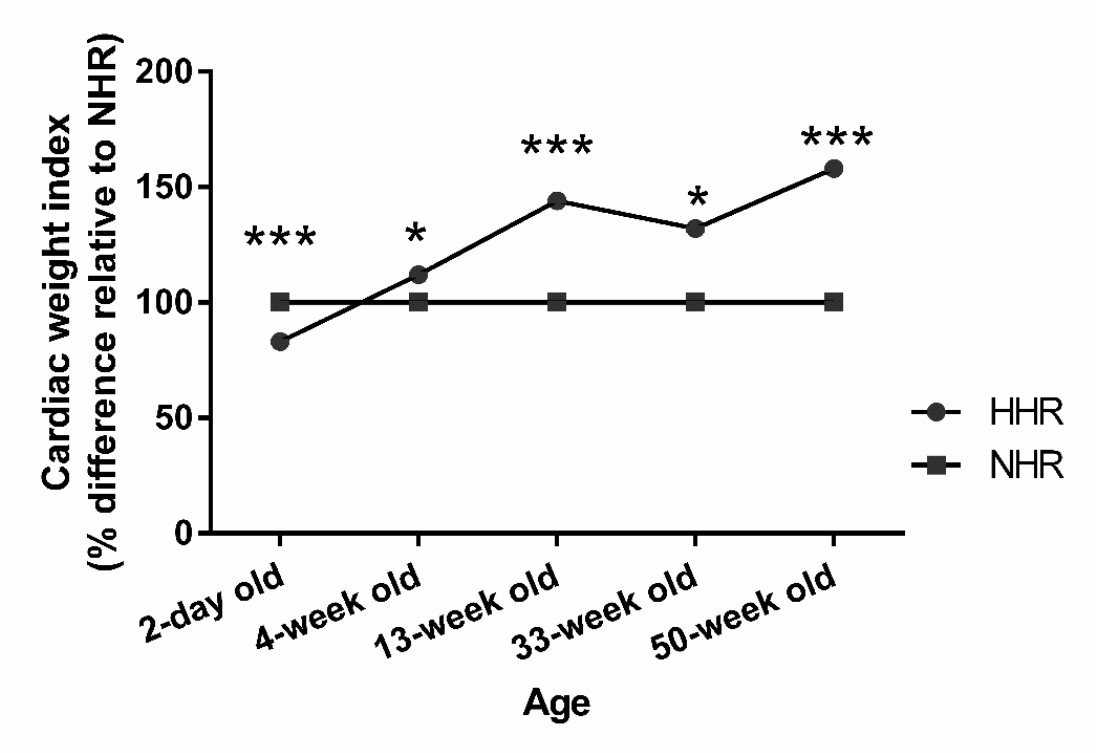

Figure 1.9: Cardiac weight index (CWI, mg/g) of hypertrophic heart rat (HHR) relative to normal heart rat (NHR). CWI is represented as a percentage difference to NHR (normalized to 100%) at 2-days old, 4-, 13-, 33- and 50-weeks old (n=7-12 per group per age). *P<0.05; **P<0.01; ***P<0.001 (Prestes et al., 2018).

However, between 6 to 10 weeks of age (coinciding with reproductive maturity in rats), cardiomyocytes in the HHR are enlarged and the left ventricle chamber is smaller (Figure 1.10), resulting in established LVH which leads to HF and premature death (Porrello, Bell, et al., 2009). Given the possible number of polygenic cases of LVH, the HHR is especially relevant to study the hypertrophic heart.

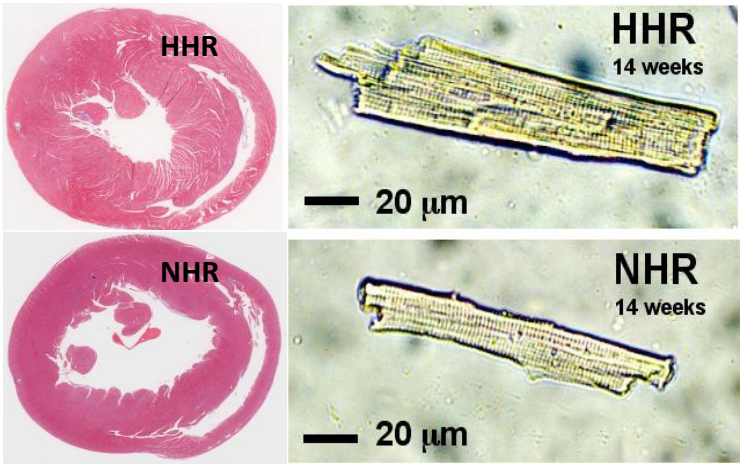

Figure 1.10: Differences between hypertrophic heart rat (HHR) and normal heart rat (NHR). Cross-section of left ventricle chamber showing a decrease in size and isolated cardiomyocytes from adult (14 weeks old HHR and NHR) (Porrello, Bell, et al., 2009).

1.6 Current treatments

The strategies for the management of patients with LVH can include the use of medications or they can be non-pharmacological. Traditionally, most non-pharmacological approaches are more effective but likely to be extremely invasive, therefore are used as a last option, such as implantation of a defibrillator, heart transplant and surgical myectomy. Some non-pharmacological techniques can be less invasive but are often less efficient, such as alcohol abstinence or external defibrillation/therapeutic hypothermia (Maron et al., 2016).

These approaches have ameliorated symptoms, decreasing mortality and improving quality of life and life expectancy of people at risk or who experience disease-related complications (Force et al., 2010; Maron et al., 2016).

The pharmacological approaches include the use of beta-blockers, anticoagulants, calcium channel blockers and antiarrhythmics, amongst others (Marian & Braunwald, 2017). Whilst these strategies are less invasive, they are also much less effective. A retrospective study investigating the efficacy of medication use for the prevention of sudden death in LVH patients found that 10% of patients died suddenly or had an aborted cardiac arrest while on medication for an average of 62 months prior to the cardiac event (Melacini et al., 2007). There were no statistically significant differences between patients on or without medication treatment. This study suggests that pharmaceutical treatment is not an effective approach for the treatment of LVH or the prevention of sudden death in patients with LVH.

1.7 New biomarkers, diagnostic tools and treatment targets

Some genetic and epigenetic changes involved in LVH have been well described. However, the origins of primary LVH in a large proportion of patients are still unknown. Therefore, investigations into new biomarkers and diagnostic tools for LVH patients and people at risk of developing LVH are still relevant and necessary.

Moreover, as highlighted above, the current prevention and treatment options available for LVH patients are predominantly invasive or ineffective. However, scientific advances into the molecular basis of LVH have provided new diagnostic tools and treatment avenues to be explored. Novel pharmacological approaches have targeted the molecular basis of LVH and new non-pharmacological techniques may potentially offer minimally invasive alternatives to the current invasive options (Tuohy, Kaul, Song, Nazer, & Heitner, 2020).

The advance in targeted gene editing techniques, mainly using CRISPR-Cas9 (clustered regularly interspaced short palindromic repeats/clustered regularly interspaced short palindromic repeat–associated 9), could potentially correct pathogenic variants in the embryo to avoid the development of the LVH phenotype (Repetti, Toepfer, Seidman, & Seidman, 2019; Tuohy et al., 2020). A recent study performed prezygotic correction of a pathogenic variant by manipulating the DNA sequence of sperm carrying a 4-base pair pathogenic deletion in the *MYPBC3* gene before fertilisation of an egg with normal *MYPBC3* sequence. This study used the CRISPR-Cas9 technology as a proof-of-concept to repair the deletion in the parental gene with 72% of embryos carrying the normal biallelic *MYPBC3* gene (Ma et al., 2017). This technique seems promising for correcting genetic malignancies but potential off-targets can be created that need to be considered to decrease the chances of mutagenesis (Ma et al., 2017).

Other genetic approaches such as small interfering RNA therapy (siRNA), miRNA/anti-miR and gene therapy are alternative options to target genetic variants (Repetti et al., 2019). As an example, animal studies identified multiple miRNAs (miR-16, miR-20b,miR-93, miR-106b, miR-223, and miR-423-5p) upregulated in plasma of hypertension-induced HF. Animals treated with anti-miR-208a had miRNA levels partly restored or totally normalised indicating that miRNAs levels in the circulation can respond to therapy and indicate treatment efficacy (Dickinson et al., 2013). Similarly, as mentioned previously, miRNA-1 is an important cardioprotective miRNA and downregulated in LVH. The overexpression of miRNA-1 inhibits LVH, regulating genes related to calcium handling pathways demonstrating its potential as therapeutic target (Ikeda et al., 2009; Karakikes et al., 2013). Therefore, the characterisation of miRNAs related to LVH could offer novel therapeutic targets to prevent or attenuate the condition (Lei, Hu, Sun, & Xu, 2020; Vegter, van der Meer, de Windt, Pinto, & Voors, 2016; Wehbe et al., 2019).

Lastly, the lack of studies on circRNAs involved in cardiovascular health, their expression patterns in different tissues and developmental stages, pathophysiological role, biomarker and therapeutic uses highlight the relevance of exploring the involvement of circRNA in LVH (Lei et al., 2020; Lim et al., 2019; Zhao, 2018).

As mentioned previously, the development of LVH increases the risk of stroke, leading to HF and sudden death, highlighting the importance of discovering novel biomarkers and diagnostic markers for early detection, new therapeutic targets and improved prognostic methods (Shimizu & Minamino, 2016).

1.8 Aims and hypothesis

Currently, there are few studies investigating genetic and epigenetic contributions in a model of established polygenic LVH. The polygenic model of LVH introduced above will be used to investigate both genetic and epigenetic contributions to LVH. To achieve this, I will employ a systematic approach:

(i) conducting whole genome sequencing to investigate the presence of genetic variants in the HHR and NHR models and explore potential relationships with the hypertrophic phenotype;

(ii) investigating mRNA expression of genes with known association with human monogenic forms of LVH during the developmental stages of the HHR and NHR. The primary purpose of this is to characterise the potential influence of genes associated with monogenic LVH in the polygenic context, especially in stage-specific changes;

(iii) comparing epigenetic profiles observed in the HHR and NHR using the emerging field of circRNAs to investigate their feasibility as potential biomarkers to identify epigenetic changes that may modulate mRNA expression and predispose to LVH.

The aims of this thesis are to:

(1) employ whole genome sequencing to investigate genetic variants involved in LVH independent of blood pressure, focusing on quantitative trait locus cardiac mass 22 (Cm22), with the objective of identifying candidate genes of polygenic LVH;

(2) use a combined approach of DNA sequence variants, mRNA and miRNA expression of genes previously associated with human monogenic forms of cardiomyopathies with the purpose of investigating their potential involvement in polygenic LVH; and

(3) establish circRNA expression profiles in the HHR and NHR to determine their potential influence in polygenic LVH.

I hypothesise that changes in the genomic sequence, mRNA expression and epigenome play a crucial and important role in the development of LVH.

1.9. This thesis

In chapter 2, I present an article in which I was lead author entitled "Tripartite motif-containing 55 identified as functional candidate for spontaneous cardiac hypertrophy in the rat locus cardiac mass 22" published in the *Journal of Hypertension* (DOI: 10.1097/HJH.0000000000000875). Here I demonstrated that a missense genetic variant in the gene encoding tripartite motif containing 55 (*Trim55*) results in the downregulation of its mRNA expression in a unique rat model of polygenic LVH, the HHR, and in hearts of subjects with idiopathic cardiomyopathy (Prestes et al., 2016). I investigated rat hearts at 2 days, 13

weeks and 33 weeks-old and found that this gene was downregulated throughout life suggesting *Trim55* as a candidate gene related to the neonatal programming that controls LVH in adulthood in rat with LVH (Prestes et al., 2016).

In Chapter 3, I present an article in which I was lead author entitled "Involvement of human monogenic cardiomyopathy genes in experimental polygenic cardiac hypertrophy" published in the journal *Physiological Genomics* (DOI: 10.1152/physiolgenomics.00143.2017). In this publication, I demonstrated that genes associated with monogenic LVH may be involved in polygenic LVH using a combined approach investigating gene variants data and mRNA expression changes (Prestes et al., 2018). I used the unique rat model of polygenic LVH, the HHR, to study 42 genes previously described in monogenic forms of LVH in five age groups (2 days, 4 weeks, 13 weeks, 33 weeks and 55 weeks of age). I found that 41 of the 42 genes investigated had genetic variants in either the HHR or the NHR which may contribute to the differential expression of such genes. Furthermore, I found that 39 of the 42 genes had at least one gene variant and 29 genes showed differential expression in at least one age group. These findings demonstrate the involvement of monogenic genes in polygenic LVH (Prestes et al., 2018).

In Chapter 4, I present a submitted article in which I was lead author entitled "Heart related circular RNA (*Hrcr*) influences polygenic cardiac hypertrophy by altering cardiac microRNAs" submitted to the *Journal of the American Heart Association*. In this article, I characterised the expression of circRNA *Hrcr* in the animal model of polygenic LVH, the HHR; silenced this circRNA in human primary cardiomyocytes to discover novel downstream targets; and performed *in silico* investigations to predict miRNA-mRNA targets. I found that circRNA *Hrcr* is upregulated in the hypertrophic hearts of rats and revealed four novel miRNA targets (miR-1-3p, miR-330, miR-27a-5p and miR-299-5p) and two noncoding RNA targets (ribosomal RNA RNA5-8SP2 and piwi-interacting RNA piR-17153).

1.10 References

Arnett, D. K., Li, N., Tang, W., Rao, D. C., Devereux, R. B., Claas, S. A., . . . Broeckel, U. (2009). Genome-wide association study identifies single-nucleotide polymorphism in KCNB1 associated with left ventricular mass in humans: the HyperGEN Study. BMC Med Genet, 10, 43.

Aung, N., Vargas, J. D., Yang, C., Cabrera, C. P., Warren, H. R., Fung, K., . . . Petersen, S. E. (2019). Genome-Wide Analysis of Left Ventricular Image-Derived Phenotypes Identifies Fourteen Loci Associated With Cardiac Morphogenesis and Heart Failure Development. Circulation, 140(16), 1318-1330.

Barrett, S. P., & Salzman, J. (2016). Circular RNAs: analysis, expression and potential functions. Development, 143(11), 1838-1847.

Bell, J. R., Porrello, E. R., Huggins, C. E., Harrap, S. B., & Delbridge, L. M. (2008). The intrinsic resistance of female hearts to an ischemic insult is abrogated in primary cardiac hypertrophy. Am J Physiol Heart Circ Physiol, 294(4), H1514-1522.

Borchert, B., Tripathi, S., Francino, A., Navarro-Lopez, F., & Kraft, T. (2010). The left and right ventricle of a patient with a R723G mutation of the beta-myosin heavy chain and severe hypertrophic cardiomyopathy show no differences in the expression of myosin mRNA. Cardiol J, 17(5), 518-522.

Burke, M. A., Cook, S. A., Seidman, J. G., & Seidman, C. E. (2016). Clinical and Mechanistic Insights Into the Genetics of Cardiomyopathy. J Am Coll Cardiol, 68(25), 2871-2886.

Burns, C., Bagnall, R. D., Lam, L., Semsarian, C., & Ingles, J. (2017). Multiple Gene Variants in Hypertrophic Cardiomyopathy in the Era of Next-Generation Sequencing. Circ Cardiovasc Genet, 10(4).

Cain, P. A., Ahl, R., Hedstrom, E., Ugander, M., Allansdotter-Johnsson, A., Friberg, P., & Arheden, H. (2009). Age and gender specific normal values of left ventricular mass,

volume and function for gradient echo magnetic resonance imaging: a cross sectional study. BMC Med Imaging, 9, 2.

Care, A., Catalucci, D., Felicetti, F., Bonci, D., Addario, A., Gallo, P., . . . Condorelli, G. (2007). MicroRNA-133 controls cardiac hypertrophy. Nat Med, 13(5), 613-618.

Chen, C., Ponnusamy, M., Liu, C., Gao, J., Wang, K., & Li, P. (2017). MicroRNA as a Therapeutic Target in Cardiac Remodeling. Biomed Res Int, 2017, 1278436.

Chen, D., Lu, X., Yang, F., & Xing, N. (2019). Circular RNA circHIPK3 promotes cell proliferation and invasion of prostate cancer by sponging miR-193a-3p and regulating MCL1 expression. Cancer Manag Res, 11, 1415-1423.

Chen, K., & Rajewsky, N. (2007). The evolution of gene regulation by transcription factors and microRNAs. Nat Rev Genet, 8(2), 93-103.

Chen, P., Li, Z., Nie, J., Wang, H., Yu, B., Wen, Z., . . . Wang, D. W. (2020). MYH7B variants cause hypertrophic cardiomyopathy by activating the CaMK-signaling pathway. Sci China Life Sci.

Cirino, A. L. (2008 [updated 2014]). Hypertrophic cardiomyopathy overview. Seattle: University of Washington, Seattle.

Das, K. J., Ingles, J., Bagnall, R. D., & Semsarian, C. (2014). Determining pathogenicity of genetic variants in hypertrophic cardiomyopathy: importance of periodic reassessment. Genet Med, 16(4), 286-293.

de Jonge, L. L., van Osch-Gevers, L., Willemsen, S. P., Steegers, E. A., Hofman, A., Helbing, W. A., & Jaddoe, V. W. (2011). Growth, obesity, and cardiac structures in early childhood: the Generation R Study. Hypertension, 57(5), 934-940.

de Simone, G. (2004). Concentric or eccentric hypertrophy: how clinically relevant is the difference? Hypertension, 43(4), 714-715.

Devaux, Y., Creemers, E. E., Boon, R. A., Werfel, S., Thum, T., Engelhardt, S., . . . Cardiolinc, n. (2017). Circular RNAs in heart failure. Eur J Heart Fail, 19(6), 701-709.

Dickinson, B. A., Semus, H. M., Montgomery, R. L., Stack, C., Latimer, P. A., Lewton, S. M., . . . van Rooij, E. (2013). Plasma microRNAs serve as biomarkers of therapeutic efficacy and disease progression in hypertension-induced heart failure. Eur J Heart Fail, 15(6), 650-659.

Dorn, G. W., 2nd. (2007). The fuzzy logic of physiological cardiac hypertrophy. Hypertension, 49(5), 962-970.

Drozdov, I., Didangelos, A., Yin, X., Zampetaki, A., Abonnenc, M., Murdoch, C., . . . Shah, A. M. (2013). Gene network and proteomic analyses of cardiac responses to pathological and physiological stress. Circ Cardiovasc Genet, 6(6), 588-597.

Dwyer, J. P., Ritchie, M. E., Smyth, G. K., Harrap, S. B., Delbridge, L., Domenighetti, A. A., & Di Nicolantonio, R. (2008). Myocardial gene expression associated with genetic cardiac hypertrophy in the absence of hypertension. Hypertens Res, 31(5), 941-955.

Egger, G., Liang, G., Aparicio, A., & Jones, P. A. (2004). Epigenetics in human disease and prospects for epigenetic therapy. Nature, 429(6990), 457-463.

Fan, X., Weng, X., Zhao, Y., Chen, W., Gan, T., & Xu, D. (2017). Circular RNAs in Cardiovascular Disease: An Overview. Biomed Res Int, 2017, 5135781.

Farrell, E., Armstrong, A. E., Grimes, A. C., Naya, F. J., de Lange, W. J., & Ralphe, J. C. (2018). Transcriptome Analysis of Cardiac Hypertrophic Growth in MYBPC3-Null Mice Suggests Early Responders in Hypertrophic Remodeling. Front Physiol, 9, 1442.

Force, T., Bonow, R. O., Houser, S. R., Solaro, R. J., Hershberger, R. E., Adhikari, B., . . . Seidman, C. E. (2010). Research priorities in hypertrophic cardiomyopathy: report of a Working Group of the National Heart, Lung, and Blood Institute. Circulation, 122(11), 1130-1133.

Frey, N., Katus, H. A., Olson, E. N., & Hill, J. A. (2004). Hypertrophy of the heart: a new therapeutic target? Circulation, 109(13), 1580-1589.

Friedrich, F. W., & Carrier, L. (2012). Genetics of hypertrophic and dilated cardiomyopathy. Curr Pharm Biotechnol, 13(13), 2467-2476.

Fuchs, A., Mejdahl, M. R., Kuhl, J. T., Stisen, Z. R., Nilsson, E. J., Kober, L. V., . . . Kofoed, K. F. (2016). Normal values of left ventricular mass and cardiac chamber volumes assessed by 320-detector computed tomography angiography in the Copenhagen General Population Study. Eur Heart J Cardiovasc Imaging, 17(9), 1009-1017.

Gangwar, R. S., Rajagopalan, S., Natarajan, R., & Deiuliis, J. A. (2018). Noncoding RNAs in Cardiovascular Disease: Pathological Relevance and Emerging Role as Biomarkers and Therapeutics. Am J Hypertens, 31(2), 150-165.

Gao, J., Collyer, J., Wang, M., Sun, F., & Xu, F. (2020). Genetic Dissection of Hypertrophic Cardiomyopathy with Myocardial RNA-Seq. Int J Mol Sci, 21(9).

Greco, C. M., & Condorelli, G. (2015). Epigenetic modifications and noncoding RNAs in cardiac hypertrophy and failure. Nat Rev Cardiol, 12(8), 488-497.

Harrap, S. B., Danes, V. R., Ellis, J. A., Griffiths, C. D., Jones, E. F., & Delbridge, L. M. (2002). The hypertrophic heart rat: a new normotensive model of genetic cardiac and cardiomyocyte hypertrophy. Physiol Genomics, 9(1), 43-48.

Ikeda, S., He, A., Kong, S. W., Lu, J., Bejar, R., Bodyak, N., . . . Pu, W. T. (2009). MicroRNA-1 negatively regulates expression of the hypertrophy-associated calmodulin and Mef2a genes. Mol Cell Biol, 29(8), 2193-2204.

Jarcho, J. A., McKenna, W., Pare, J. A., Solomon, S. D., Holcombe, R. F., Dickie, S., . . . Seidman, C. E. (1989). Mapping a gene for familial hypertrophic cardiomyopathy to chromosome 14q1. N Engl J Med, 321(20), 1372-1378.

Karakikes, I., Chaanine, A. H., Kang, S., Mukete, B. N., Jeong, D., Zhang, S., . . . Lebeche, D. (2013). Therapeutic cardiac-targeted delivery of miR-1 reverses pressure overload-induced cardiac hypertrophy and attenuates pathological remodeling. J Am Heart Assoc, 2(2), e000078.

Katholi, R. E., & Couri, D. M. (2011). Left ventricular hypertrophy: major risk factor in patients with hypertension: update and practical clinical applications. Int J Hypertens, 2011, 495349

Kavazis, A. N. (2015). Pathological vs. physiological cardiac hypertrophy. J Physiol, 593(17), 3767.

Kontaraki, J. E., Marketou, M. E., Parthenakis, F. I., Maragkoudakis, S., Zacharis, E. A., Petousis, S., . . . Vardas, P. E. (2015). Hypertrophic and antihypertrophic microRNA levels in peripheral blood mononuclear cells and their relationship to left ventricular hypertrophy in patients with essential hypertension. J Am Soc Hypertens, 9(10), 802-810.

Kovacova, J., Poprach, A., Buchler, T., Cho, W. C., & Slaby, O. (2018). MicroRNAs as predictive biomarkers of response to tyrosine kinase inhibitor therapy in metastatic renal cell carcinoma. Clin Chem Lab Med, 56(9):1426-1431.

Kozomara, A., Birgaoanu, M., & Griffiths-Jones, S. (2019). miRBase: from microRNA sequences to function. Nucleic Acids Res, 47(D1), D155-D162.

Kozomara, A., & Griffiths-Jones, S. (2014). miRBase: annotating high confidence microRNAs using deep sequencing data. Nucleic Acids Res, 42(Database issue), D68-73.

Kraft, T., Montag, J., Radocaj, A., & Brenner, B. (2016). Hypertrophic Cardiomyopathy: Cell-to-Cell Imbalance in Gene Expression and Contraction Force as Trigger for Disease Phenotype Development. Circ Res, 119(9), 992-995.

Kung, J. T., Colognori, D., & Lee, J. T. (2013). Long noncoding RNAs: past, present, and future. Genetics, 193(3), 651-669.

Landrum, M. J., Lee, J. M., Benson, M., Brown, G. R., Chao, C., Chitipiralla, S., . . . Maglott, D. R. (2018). ClinVar: improving access to variant interpretations and supporting evidence. Nucleic Acids Res, 46(D1), D1062-D1067.

Lee, Y. Z. J., & Judge, D. P. (2017). The Role of Genetics in Peripartum Cardiomyopathy. J Cardiovasc Transl Res, 10(5-6), 437-445.

Lei, H., Hu, J., Sun, K., & Xu, D. (2020). The role and molecular mechanism of epigenetics in cardiac hypertrophy. Heart Fail Rev. Online ahead of print.

Lei, W., Feng, T., Fang, X., Yu, Y., Yang, J., Zhao, Z. A., . . . Hu, S. (2018). Signature of circular RNAs in human induced pluripotent stem cells and derived cardiomyocytes. Stem Cell Res Ther, 9(1), 56.

Levy, D., Garrison, R. J., Savage, D. D., Kannel, W. B., & Castelli, W. P. (1990). Prognostic implications of echocardiographically determined left ventricular mass in the Framingham Heart Study. N Engl J Med, 322(22), 1561-1566.

Li, H., Xu, J. D., Fang, X. H., Zhu, J. N., Yang, J., Pan, R., . . . Shan, Z. X. (2020). Circular RNA circRNA_000203 aggravates cardiac hypertrophy via suppressing miR-26b-5p and miR-140-3p binding to Gata4. Cardiovasc Res, 116(7), 1323-1334.

Li, M., Ding, W., Sun, T., Tariq, M. A., Xu, T., Li, P., & Wang, J. (2018). Biogenesis of circular RNAs and their roles in cardiovascular development and pathology. FEBS J, 285(2), 220-232.

Lim, T. B., Aliwarga, E., Luu, T. D. A., Li, Y. P., Ng, S. L., Annadoray, L., . . . Foo, R. S. (2019). Targeting the highly abundant circular RNA circSlc8a1 in cardiomyocytes attenuates pressure overload induced hypertrophy. Cardiovasc Res, 115(14), 1998-2007.

Liu, C. F., & Tang, W. H. W. (2019). Epigenetics in Cardiac Hypertrophy and Heart Failure. JACC Basic Transl Sci, 4(8), 976-993.

Lopes, L. R., Syrris, P., Guttmann, O. P., O'Mahony, C., Tang, H. C., Dalageorgou, C., . . . Elliott, P. M. (2015). Novel genotype-phenotype associations demonstrated by high-throughput sequencing in patients with hypertrophic cardiomyopathy. Heart, 101(4), 294-301.

Lorell, B. H., & Carabello, B. A. (2000). Left ventricular hypertrophy: pathogenesis, detection, and prognosis. Circulation, 102(4), 470-479.

Lozano, R., Naghavi, M., Foreman, K., Lim, S., Shibuya, K., Aboyans, V., . . . Memish, Z. A. (2012). Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet, 380(9859), 2095-2128.

Lu, Y., & Wu, F. (2018). A new miRNA regulator, miR-672, reduces cardiac hypertrophy by inhibiting JUN expression. Gene, 648, 21-30.

Ma, H., Marti-Gutierrez, N., Park, S. W., Wu, J., Lee, Y., Suzuki, K., . . . Mitalipov, S. (2017). Correction of a pathogenic gene mutation in human embryos. Nature, 548(7668), 413-419.

Maillet, M., van Berlo, J. H., & Molkentin, J. D. (2013). Molecular basis of physiological heart growth: fundamental concepts and new players. Nat Rev Mol Cell Biol, 14(1), 38-48.

Marian, A. J., & Braunwald, E. (2017). Hypertrophic Cardiomyopathy: Genetics, Pathogenesis, Clinical Manifestations, Diagnosis, and Therapy. Circ Res, 121(7), 749-770.

Marian, A. J., van Rooij, E., & Roberts, R. (2016). Genetics and Genomics of Single-Gene Cardiovascular Diseases: Common Hereditary Cardiomyopathies as Prototypes of Single-Gene Disorders. J Am Coll Cardiol, 68(25), 2831-2849.

Maron, B. J. (2002). Hypertrophic cardiomyopathy: A systematic review. JAMA, 287(10), 1308-1320.

Maron, B. J., Carney, K. P., Lever, H. M., Lewis, J. F., Barac, I., Casey, S. A., & Sherrid, M. V. (2003). Relationship of race to sudden cardiac death in competitive athletes with hypertrophic cardiomyopathy. J Am Coll Cardiol, 41(6), 974-980.

Maron, B. J., & Maron, M. S. (2013). Hypertrophic cardiomyopathy. Lancet, 381(9862), 242-255.

Maron, B. J., Maron, M. S., & Semsarian, C. (2012). Genetics of hypertrophic cardiomyopathy after 20 years: clinical perspectives. J Am Coll Cardiol, 60(8), 705-715.

Maron, B. J., Rowin, E. J., Casey, S. A., & Maron, M. S. (2016). How Hypertrophic Cardiomyopathy Became a Contemporary Treatable Genetic Disease With Low Mortality: Shaped by 50 Years of Clinical Research and Practice. JAMA Cardiol, 1(1), 98-105.

Maron, B. J., Towbin, J. A., Thiene, G., Antzelevitch, C., Corrado, D., Arnett, D., . . . Prevention. (2006). Contemporary definitions and classification of the cardiomyopathies: an American Heart Association Scientific Statement from the Council on Clinical Cardiology, Heart Failure and Transplantation Committee; Quality of Care and Outcomes Research and Functional Genomics and Translational Biology Interdisciplinary Working Groups; and Council on Epidemiology and Prevention. Circulation, 113(14), 1807-1816.

Matkovich, S. J., Wang, W., Tu, Y., Eschenbacher, W. H., Dorn, L. E., Condorelli, G., . . . Dorn, G. W., 2nd. (2010). MicroRNA-133a protects against myocardial fibrosis and modulates electrical repolarization without affecting hypertrophy in pressure-overloaded adult hearts. Circ Res, 106(1), 166-175.

Mattick, J. S., & Makunin, I. V. (2006). Non-coding RNA. Hum Mol Genet, 15 Spec No 1, R17-29.

McKenna, W. J., & Judge, D. P. (2020). Epidemiology of the inherited cardiomyopathies. Nat Rev Cardiol.

McMullen, J. R., & Jennings, G. L. (2007). Differences between pathological and physiological cardiac hypertrophy: novel therapeutic strategies to treat heart failure. Clin Exp Pharmacol Physiol, 34(4), 255-262.

Melacini, P., Maron, B. J., Bobbo, F., Basso, C., Tokajuk, B., Zucchetto, M., . . . Iliceto, S. (2007). Evidence that pharmacological strategies lack efficacy for the prevention of sudden death in hypertrophic cardiomyopathy. Heart, 93(6), 708-710.

Montag, J., Petersen, B., Flogel, A. K., Becker, E., Lucas-Hahn, A., Cost, G. J., . . . Brenner, B. (2018). Successful knock-in of Hypertrophic Cardiomyopathy-mutation R723G into the MYH7 gene mimics HCM pathology in pigs. Sci Rep, 8(1), 4786.

Mosqueira, D., Smith, J. G. W., Bhagwan, J. R., & Denning, C. (2019). Modeling Hypertrophic Cardiomyopathy: Mechanistic Insights and Pharmacological Intervention. Trends Mol Med, 25(9), 775-790.

Olivotto, I., Maron, M. S., Adabag, A. S., Casey, S. A., Vargiu, D., Link, M. S., . . . Maron, B. J. (2005). Gender-related differences in the clinical presentation and outcome of hypertrophic cardiomyopathy. J Am Coll Cardiol, 46(3), 480-487.

Ordovas, J. M., & Smith, C. E. (2010). Epigenetics and cardiovascular disease. Nat Rev Cardiol, 7(9), 510-519.

Ottaviani, L., & da Costa Martins, P. A. (2017). Non-coding RNAs in cardiac hypertrophy. J Physiol, 595(12), 4037-4050.

Pare, J. A., Fraser, R. G., Pirozynski, W. J., Shanks, J. A., & Stubington, D. (1961). Hereditary cardiovascular dysplasia. A form of familial cardiomyopathy. Am J Med, 31, 37-62.

Payne, J. R., Eleftheriou, K. I., James, L. E., Hawe, E., Mann, J., Stronge, A., . . . Montgomery, H. E. (2006). Left ventricular growth response to exercise and cigarette smoking: data from LARGE Heart. Heart, 92(12), 1784-1788.

Pedersen, L. R., Kristensen, A. M. D., Petersen, S. S., Vaduganathan, M., Bhatt, D. L., Juel, J., . . . Pareek, M. (2020). Prognostic implications of left ventricular hypertrophy diagnosed on electrocardiogram vs echocardiography. J Clin Hypertens (Greenwich), 22(9), 1647-1658.

Petersen, S. E., Aung, N., Sanghvi, M. M., Zemrak, F., Fung, K., Paiva, J. M., . . . Neubauer, S. (2017). Reference ranges for cardiac structure and function using cardiovascular magnetic resonance (CMR) in Caucasians from the UK Biobank population cohort. J Cardiovasc Magn Reson, 19(1), 18.

Pfaffenberger, S., Bartko, P., Graf, A., Pernicka, E., Babayev, J., Lolic, E., . . . Mascherbauer, J. (2013). Size matters! Impact of age, sex, height, and weight on the normal heart size. Circ Cardiovasc Imaging, 6(6), 1073-1079.

Porrello, E. R., Bell, J. R., Schertzer, J. D., Curl, C. L., McMullen, J. R., Mellor, K. M., . . . Delbridge, L. M. (2009). Heritable pathologic cardiac hypertrophy in adulthood is preceded by neonatal cardiac growth restriction. Am J Physiol Regul Integr Comp Physiol, 296(3), R672-680.

Porrello, E. R., D'Amore, A., Curl, C. L., Allen, A. M., Harrap, S. B., Thomas, W. G., & Delbridge, L. M. (2009). Angiotensin II type 2 receptor antagonizes angiotensin II type 1 receptor-mediated cardiomyocyte autophagy. Hypertension, 53(6), 1032-1040.

Post, W. S., Larson, M. G., Myers, R. H., Galderisi, M., & Levy, D. (1997). Heritability of left ventricular mass: the Framingham Heart Study. Hypertension, 30(5), 1025-1028.

Prestes, P. R., Marques, F. Z., Lopez-Campos, G., Booth, S. A., McGlynn, M., Lewandowski, P., . . . Charchar, F. J. (2016). Tripartite motif-containing 55 identified as functional candidate for spontaneous cardiac hypertrophy in the rat locus cardiac mass 22. J Hypertens, 34(5), 950-958.

Prestes, P. R., Marques, F. Z., Lopez-Campos, G., Lewandowski, P., Delbridge, L. M. D., Charchar, F. J., & Harrap, S. B. (2018). Involvement of human monogenic cardiomyopathy genes in experimental polygenic cardiac hypertrophy. Physiol Genomics, 50(9), 680-687.

Qu, S., Yang, X., Li, X., Wang, J., Gao, Y., Shang, R., . . . Li, H. (2015). Circular RNA: A new star of noncoding RNAs. Cancer Lett, 365(2), 141-148.

Raina, M., & Ibba, M. (2014). tRNAs as regulators of biological processes. Front Genet, 5, 171.

Rajan, S., Williams, S. S., Jagatheesan, G., Ahmed, R. P., Fuller-Bicer, G., Schwartz, A., . . . Wieczorek, D. F. (2006). Microarray analysis of gene expression during early stages of mild and severe cardiac hypertrophy. Physiol Genomics, 27(3), 309-317.

Reimand, J., Isserlin, R., Voisin, V., Kucera, M., Tannus-Lopes, C., Rostamianfar, A., . . . Bader, G. D. (2019). Pathway enrichment analysis and visualization of omics data using g:Profiler, GSEA, Cytoscape and EnrichmentMap. Nat Protoc, 14(2), 482-517.

Ren, C. W., Liu, J. J., Li, J. H., Li, J. W., Dai, J., & Lai, Y. Q. (2016). RNAseq profiling of mRNA associated with hypertrophic cardiomyopathy. Mol Med Rep, 14(6), 5573-5586.

Repetti, G. G., Toepfer, C. N., Seidman, J. G., & Seidman, C. E. (2019). Novel Therapies for Prevention and Early Treatment of Cardiomyopathies. Circ Res, 124(11), 1536-1550.

Reza, N., Musunuru, K., & Owens, A. T. (2019). From Hypertrophy to Heart Failure: What Is New in Genetic Cardiomyopathies. Curr Heart Fail Rep, 16(5), 157-167.

Salzman, J. (2016). Circular RNA Expression: Its Potential Regulation and Function. Trends Genet, 32(5), 309-316.

Sanoudou, D., Vafiadaki, E., Arvanitis, D. A., Kranias, E., & Kontrogianni-Konstantopoulos, A. (2005). Array lessons from the heart: focus on the genome and transcriptome of cardiomyopathies. Physiol Genomics, 21(2), 131-143.

Schirmer, H., Lunde, P., & Rasmussen, K. (1999). Prevalence of left ventricular hypertrophy in a general population; The Tromso Study. Eur Heart J, 20(6), 429-438.

Seebohm, B., Matinmehr, F., Kohler, J., Francino, A., Navarro-Lopez, F., Perrot, A., . . . Kraft, T. (2009). Cardiomyopathy mutations reveal variable region of myosin converter as major element of cross-bridge compliance. Biophys J, 97(3), 806-824.

Semsarian, C., Ingles, J., Maron, M. S., & Maron, B. J. (2015). New perspectives on the prevalence of hypertrophic cardiomyopathy. J Am Coll Cardiol, 65(12), 1249-1254.

Shimizu, I., & Minamino, T. (2016). Physiological and pathological cardiac hypertrophy. J Mol Cell Cardiol, 97, 245-262.

Smith, J. G. (2017). Molecular Epidemiology of Heart Failure: Translational Challenges and Opportunities. JACC Basic Transl Sci, 2(6), 757-769.

Song, H. K., Hong, S. E., Kim, T., & Kim, D. H. (2012). Deep RNA sequencing reveals novel cardiac transcriptomic signatures for physiological and pathological hypertrophy. PLoS One, 7(4), e35552.

Starke, S., Jost, I., Rossbach, O., Schneider, T., Schreiner, S., Hung, L. H., & Bindereif, A. (2015). Exon circularization requires canonical splice signals. Cell Rep, 10(1), 103-111.

Swan, L., Birnie, D. H., Padmanabhan, S., Inglis, G., Connell, J. M., & Hillis, W. S. (2003). The genetic determination of left ventricular mass in healthy adults. Eur Heart J, 24(6), 577-582.

Tan, W. L., Lim, B. T., Anene-Nzelu, C. G., Ackers-Johnson, M., Dashi, A., See, K., . . . Foo, R. S. (2017). A landscape of circular RNA expression in the human heart. Cardiovasc Res, 113(3), 298-309.

Tsao, C. W., Gona, P. N., Salton, C. J., Chuang, M. L., Levy, D., Manning, W. J., & O'Donnell, C. J. (2015). Left Ventricular Structure and Risk of Cardiovascular Events: A Framingham Heart Study Cardiac Magnetic Resonance Study. J Am Heart Assoc, 4(9), e002188.

Tuohy, C. V., Kaul, S., Song, H. K., Nazer, B., & Heitner, S. B. (2020). Hypertrophic cardiomyopathy: the future of treatment. Eur J Heart Fail, 22(2), 228-240.

van der Bruggen, C. E. E., Tedford, R. J., Handoko, M. L., van der Velden, J., & de Man, F. S. (2017). RV pressure overload: from hypertrophy to failure. Cardiovasc Res, 113(12), 1423-1432.

Vegter, E. L., van der Meer, P., de Windt, L. J., Pinto, Y. M., & Voors, A. A. (2016). MicroRNAs in heart failure: from biomarker to target for therapy. Eur J Heart Fail, 18(5), 457-468.

Walsh, R., Buchan, R., Wilk, A., John, S., Felkin, L. E., Thomson, K. L., . . . Cook, S. A. (2017). Defining the genetic architecture of hypertrophic cardiomyopathy: re-evaluating the role of non-sarcomeric genes. Eur Heart J, 38(46), 3461-3468.

Wang, K., Long, B., Liu, F., Wang, J. X., Liu, C. Y., Zhao, B., . . . Li, P. F. (2016). A circular RNA protects the heart from pathological hypertrophy and heart failure by targeting miR-223. Eur Heart J, 37(33), 2602-2611.

Wehbe, N., Nasser, S. A., Pintus, G., Badran, A., Eid, A. H., & Baydoun, E. (2019). MicroRNAs in Cardiac Hypertrophy. Int J Mol Sci, 20(19).

Wild, P. S., Felix, J. F., Schillert, A., Teumer, A., Chen, M. H., Leening, M. J. G., . . . Dorr, M. (2017). Large-scale genome-wide analysis identifies genetic variants associated with cardiac structure and function. J Clin Invest, 127(5), 1798-1812.

World Health Organization. (September 2016). Cardiovascular diseases (CVDs) - fact sheet. Retrieved from http://www.who.int/mediacentre/factsheets/fs317/en/

Yan, Q., Zhu, C., Guang, S., & Feng, X. (2019). The Functions of Non-coding RNAs in rRNA Regulation. Front Genet, 10, 290.

Zhao, G. (2018). Significance of non-coding circular RNAs and micro RNAs in the pathogenesis of cardiovascular diseases. J Med Genet, 55(11), 713-720.

Chapter 2. *Trim55* identified as functional candidate for spontaneous cardiac hypertrophy in the rat locus Cm22

"This is a non-final version of an article published in final form in: Prestes PR, Marques FZ, Lopez-Campos G, Booth SA, McGlynn M, Lewandowski P, Delbridge LM, Harrap SB, Charchar FJ (2016). Tripartite motif-containing 55 identified as functional candidate for spontaneous cardiac hypertrophy in the rat locus cardiac mass 22. J Hypertens. 34(5):950-8. doi: 10.1097/HJH.0000000000000875."

Total word count: 4711

Number of tables: 4

Number of figures: 5

Number of supplementary tables: 6

Number of supplementary figures: 3

Trim55 identified as functional candidate for spontaneous cardiac hypertrophy in the rat locus Cm22

Short title: *Trim55* in polygenic cardiac hypertrophy

Priscilla R PRESTES[a], Francine Z MARQUES[a], Guillermo LOPEZ-CAMPOS[b], Scott A BOOTH[a], Maree McGLYNN[c], Paul LEWANDOWSKI[c], Lea MD DELBRIDGE[d], Stephen B HARRAP[d], Fadi J CHARCHAR[a]

[a] School of Applied and Biomedical Sciences, Faculty of Science and Technology, Federation University Australia, VIC, Australia; [b] Health and Biomedical Informatics Centre, Melbourne Medical School, University of Melbourne, VIC, Australia; [c] School of Medicine, Deakin University, VIC, Australia; [d] Department of Physiology, University of Melbourne, VIC, Australia.

Conflict of interest

None.

Funding sources

This work was supported by grants from the National Health & Medical Research Council of Australia (NHMRC) (project grant APP1034371 and 628643), the National Heart Foundation (grant G10M5155) and the Federation University Australia 'Self-sustaining Regions Research and Innovation Initiative', an Australian Government Collaborative Research Network (CRN). P.R.P. is supported by a Robert HT Smith Fellowship from Federation University. F.Z.M. is supported by NHMRC (APP1052659) and National Heart Foundation (PF12M6785) co-shared Early Career Fellowships.

Corresponding author: Fadi Charchar, Room 117, Y Building, University Drive, Federation University, Mt Helen, 3350. P: (03) 5327 6098, Fax: (03) 5327 9602

Email: f.charchar@federation.edu.au

2.1 Abstract

Background: Left ventricular (LV) hypertrophy is a risk factor for cardiovascular death, but the genetic factors determining LV size and predisposition to hypertrophy are not well understood. We have previously linked the quantitative trait locus (QTL) Cm22 on chromosome 2 with cardiac hypertrophy independent of blood pressure in the spontaneously hypertensive rat (SHR). From an original cross of SHR with F344 rats, we derived a normotensive polygenic model of spontaneous cardiac hypertrophy, the hypertrophic heart rat (HHR) and its control strain, the normal heart rat (NHR).

Methods and Results: To identify the genes and molecular mechanisms underlying spontaneous LV hypertrophy we sequenced the HHR genome with special focus on QTL-Cm22. For correlative analyses of function, we measured global RNA transcripts in LV of neonatal HHR and NHR and 198 neonatal rats of an HHRxNHR F2 crossbred population. Only one gene within locus Cm22 was differentially expressed in the parental generation: tripartite motif-containing 55 (*Trim55*), with mRNA down-regulation in HHR ($P<0.05$) and reduced protein expression. *Trim55* mRNA levels were negatively correlated with LV mass in the F2 cross (r=-0.16, $P=0.025$). In exon 9 of *Trim55* in HHR we found one missense mutation that functionally alters protein structure. This mutation was strongly associated with *Trim55* mRNA expression in F2 rats (F=10.35, $P<0.0001$). Similarly, in humans we found reduced *Trim55* expression in hearts of subjects with idiopathic dilated cardiomyopathy.

Conclusions: Our study suggests that the *Trim55* gene, located in Cm22, is a novel candidate gene for polygenic LV hypertrophy independent of blood pressure.

2.2 Condensed abstract

We found one gene within locus Cm22 was differentially expressed in the parental generation: tripartite motif-containing 55 (*Trim55*), with mRNA down-regulation in HHR ($P<0.05$) and reduced protein expression. *Trim55* mRNA levels were negatively correlated with LV mass in the F2 cross (r=-0.16, $P=0.025$). In exon 9 of *Trim55* in HHR we found one missense mutation that functionally alters protein structure. This mutation was strongly associated with *Trim*55 mRNA expression in F2 rats (F=10.35, $P<0.0001$). Similarly, in humans we found reduced *Trim55* expression in hearts of subjects with idiopathic dilated cardiomyopathy.

2.3 Keywords

Cardiomyopathy; microarray; animal models; functional genomics; transcriptome.

2.4 Introduction

Evidence for genetic determination of cardiac size comes from numerous twin and family studies in which the heritability for cardiac mass has been estimated to be 30–84%. (Busjahn et al., 2009; Mayosi et al., 2002; Sharma et al., 2006) Cardiac hypertrophy, in particular left ventricular (LV) hypertrophy, is the main predisposition factor to cardiovascular disease outcomes after age. (Levy et al., 1990) The molecular processes leading to an increase in cardiac size independent of blood pressure are still unknown.

The hypertrophic heart rat (HHR), a normotensive polygenic model of cardiac hypertrophy, is a unique model because it permits the study of cardiac mass without the confounding effects of high blood pressure. This model and its control, the normal heart rat (NHR), were derived from selectively crossing the spontaneously hypertensive rat (SHR) and Fisher 344 rats for over 13 generations. (Harrap et al., 2002) At 2 days old, the HHR have smaller hearts with fewer and smaller cells when compared to the NHR defining hypertrophy later in life. By 12 weeks of age, however, the HHR presents established LV hypertrophy, which leads to premature death due to heart failure as early as 48 weeks of age. (Harrap et al., 2002; Porrello, Bell, et al., 2009)

Studies that identify genetic factors contributing to LV hypertrophy independent of pressure overload are crucial to understand the molecular mechanisms involved. The whole-genome sequencing of rat strains displaying diverse phenotypes (Atanur et al., 2010; Atanur et al., 2013) and the creation of the Rat Genome Database (Laulederkind et al., 2013) (RGD) permitted the discovery of a greater number of allelic variants amongst rat strains. These data, associated with gene expression studies, could identify disease-associated candidate genes. (Nabika, Ohara, Kato, & Isomura, 2012) Here we aimed to take advantage of the HHR as a unique model of polygenic cardiac hypertrophy to investigate mutations in genes contributing to this condition.

In this study, we compared the whole-genome sequence of the HHR with the NHR and the reference laboratory rat, *Rattus norvegicus*. We focused particularly on a region of rat Chromosome 2 that we identified in previous crosses of the SHR and Wistar Kyoto rats as linked with LV mass independent of blood pressure. (Di Nicolantonio et al., 2006; Innes, McLaughlin, Kapuscinski, Jacob, & Harrap, 1998) We originally designated this quantitative trait locus (QTL) as Lvm1, but it has been subsequently labelled as Cm22. In order to identify functional candidates in Cm22 we correlated sequence and expression data in HHR and NHR and tested their linkage with the LV mass phenotype in an F2 cross of the HHR and NHR.

2.5 Materials and methods

2.5.1 Samples and tissue collection

2-day old HHR (n=8 for arrays and n=8 for follow up studies) and NHR (n=8 for arrays and n=9 for follow up studies) were euthanized by decapitation. 13-; (n=10 HHR and n=11 NHR) and 33-week old (n=7 HHR and n=9 NHR) were euthanized with an overdose of pentobarbitone (Lethabarb) (Vet N-Pet Direct, Brisbane, Australia). We crossed 4 male HHRs with 4 female NHRs to produce an F1 population (n=9). By crossing male and female F1 rats we derived an HHRxNHR F2 population (n=198). F2 animals were euthanized by decapitation at 2-day of age. Hearts were immediately removed, and LV dissected from the atria. Cardiac weight index (CWI, mg/g) was calculated from the total heart weight (mg) relative to total body weight (g) of the animal (Table S2.1, Figure S2.1). Tissues were first preserved in liquid nitrogen and later transferred to a –80^{o}C freezer. The study was approved by the Animal Ethics Committees of the University of Melbourne and Deakin University, and ratified at Federation University Australia.

2.5.2 DNA and RNA extraction

DNA from LV tissue was extracted using the PureLink® Genomic Extraction kit (Life Technologies, Carlsbad, USA). RNA was extracted using the miRNeasy kit (Qiagen, Hilden, Germany). Both RNA and DNA were quantified by spectrophotometry using a NanoDrop® ND-100 spectrophotometer (Thermo Scientific, Wilmington, USA). RNA used in microarrays was also assessed for quality based on a RNA integrity number (RIN) higher than 8 by electrophoresis in a 2100 Bioanalyzer (Agilent, Waldbronn, Germany).

Figure S2.2 outlines the data analysis we used to identify a new candidate gene for polygenic LV hypertrophy.

2.5.3 Whole-genome sequencing and analysis

Three male 13-week old NHR and three age-matched HHR (with established LV hypertrophy) were selected for whole-genome sequencing. DNA concentration and quantity were determined by Qubit Fluorometer (Life Technologies, Carlsbad, USA) and integrity was determined by agarose gel electrophoresis. The Illumina Cluster Station and Illumina HiSeq 2000 (San Diego, USA) were used for preparation of paired-end libraries with insert size of 500bp. The whole-genome sequencing was performed on the Illumina HiSeq 2000 platform, where the sequences of each individual were generated as 90bp paired-end reads. Quality filtered FASTQ paired-end sequences were mapped to the Brown Norway (BN) reference genome Rnor_5.0 from March 2012 (Rat Genome Sequencing Consortium, RGSC v5.0) using the Burrows-Wheeler Aligner (BWA-0.7.5a) software. (Li & Durbin, 2009) Pre-processing for sorting and removal of duplicates was performed using Picard tools. (Picard tools) We used default parameters for the BWA-MEM algorithm to map reads and the Genome Analysis Tool Kit (GATK 3.1_1) (McKenna et al., 2010) to analyse mapped sequences. (Atanur et al., 2013) The sequences were realigned around insertions and deletions (InDels) and base quality scores were recalibrated for accurate base quality. We used GATK best practices (Van der Auwera et

al., 2013) for variant calling analysis, to identify single nucleotide polymorphisms (SNPs) and SnpEff software (Cingolani et al., 2012) for functional annotation. Copy number variations (CNVs) were identified using Varscan (Koboldt et al., 2009) by simultaneously comparing reads counts, base quality and allele frequency and structural variants (SVs) were identified using Seeksv to detect large insertions, deletions and inversions.

We developed an in-house database for the storage and analysis of the variants identified in the genome sequencing. This database was used to identify the unique variants in HHR and NHR. Change rate, defined as the frequency of a substitution occurring in the genome, was calculated as the ratio between the number of changes and the genome region.

2.5.4 Whole transcriptome profiling and analysis

Affymetrix GeneChip® Rat Gene 1.0 ST Arrays (Tokyo, Japan) were used for transcriptome-wide gene expression analysis. We analysed 8 neonatal HHR (4 male, 4 female) and 8 NHR (4 male, 4 female). Each animal was considered an individual sample and no pooling was performed. Briefly, mRNA was converted to ssDNA, labelled and hybridized to GeneChip® Rat Gene 1.0 ST Arrays, which analyse 27,342 gene transcripts using 722,254 probe sets (on average 26 probes per gene), according to the manufacturer's instructions, and with the assistance of the Ramaciotti Centre for Gene Function Analysis (University of New South Wales, Sydney, Australia). The data set obtained has been deposited in the NCBI Gene Expression Omnibus (GEO) database according to MIAME guidelines with series accession number GSE38607 (http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?token=bjsxbiiaowmwozy&acc=GSE38607).

Results from arrays were normalized using robust-multi-array analysis. Differentially expressed genes were identified using a two-sample t-test in the Partek® Genomics Suite™ (version 6.6) and selected based on their false discovery rate (FDR) q-value <0.05.

2.5.5 Gene functional annotation analysis

Gene functional annotation analysis was performed using the Database for Annotation, Visualization and Integrated Discovery (DAVID) (Huang, Sherman, & Lempicki, 2009a, 2009b) to ask which Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways were enriched within the 1231 genes that contained nonsynonymous unique SNPs in the hypertrophic strain.

2.5.6 Quantitative trait loci characterisation

We previously described the QTL cardiac mass 22 (Cm22), also known as Lvm1, to be associated with LV hypertrophy independent of blood pressure. (Di Nicolantonio et al., 2006; Innes et al., 1998) QTL Cm22 is located on rat chromosome 2, nucleotide 98,037,122 to 177,517,146 in RGSC v5.0. (Rat Genome Database Website, 2015)

2.5.7 Real-time quantitative PCR (qPCR) for gene expression analysis

First-strand complementary synthesis reaction was performed using the High Capacity cDNA Reverse Transcription Kit (Life Technologies, Carlsbad, USA). Primers were specifically designed around the most differentially expressed probe in the transcript cluster of each gene using Primer3 (Rozen & Skaletsky, 2000) and NCBI online tool Primer Blast. Primers were designed to flank an exon-exon junction. Primers and conditions used are shown in Table S2.2. Amplification reactions used the SensiFast^{TM} SYBR Low-ROX Kit qPCR reagent system (Bioline, Foster City, USA) in a Viia7 qPCR instrument (Life Technologies). Samples were run in duplicates. The specificity of the qPCR was ensured by melting curve analysis and electrophoresis in agarose gels (data not shown). The glyceraldehyde 3-phosphate dehydrogenase (*Gapdh*) was used as reference transcript. Significance was assessed by $2^{-\Delta\Delta CT}$. (Schmittgen & Livak, 2008)

2.5.8 Protein measurements of Trim55 by Western Blot

Whole soluble protein was extracted from left ventricle of 13 week old HHR and NHR using RIPA buffer (Sigma-Aldrich, Saint Louis, USA) and 1% Halt Protease & Phosphatase Inhibitor Cocktail (Life Technologies, Rockford, USA). One hundred micrograms of extracted rat protein were resolved by 4–15% Mini-PROTEAN® TGX™ Precast Gel (BioRad, Hercules, USA) then electro blotted on to a Nitrocellulose Membrane (Life Technologies, Rockford, USA). Membranes were blocked for one hour at room temperature in 5% skim milk in TBS-Tween 20, then incubated overnight at 4ºC with anti-TRIM55 (N-terminal) Rabbit Polyclonal antibody (1:2500 dilution, ABIN405554, Antibodies-online Inc., Atlanta, USA) in 0.5% skim milk/TBS-Tween 20. Membranes were stripped using Restore^{TM} Western Blot Stripping Buffer (Life Technologies, Rockford, USA) and probed with β-actin (1:5000 dilution, 3700, Cell Signalling, Danvers, USA) in 0.5% skim milk/TBS-Tween 20 solution for one hour at room temperature, followed by four washing steps for 15 minutes in TBS-Tween 20 before detection using enhanced chemiluminescence SuperSignal West Pico Substrate (Life Technologies, Rockford, USA). Images were captured with a UVITEC Alliance digital imaging system (Thermo Scientific).

2.5.9 3D protein prediction

The 3D protein prediction of *Trim55* of HHR and NHR were generated after uploading the amino acid sequences separately into online software Phyre, (Kelley & Sternberg, 2009) using the normal modelling mode.

2.5.10 Statistical analyses

GraphPad Prism (version 6) package was used for graphing and statistical analyses. Data sets were tested for normal distribution using the D'Agostino & Pearson normality test, and equal variances were analysed using the F test. An independent sample *t*-test, Welch's test,

or Mann-Whitney test were used to compare the data between the groups. Significance was set at $P<0.05$. A one-way ANOVA test was used to compare *Trim55* abundance and exonic variant genotypes. A linear regression test was performed to compare CWI and *Trim55* abundance (Figure S2.3).

2.5.11 Gene expression in human hearts

We used data in the repository GEO series GSE1145 to investigate the levels of *TRIM55* in human idiopathic dilated hearts (n=11 control hearts and n=15 idiopathic dilated hearts). We then performed a whole-genome analysis using the GEO tools to determine whether *TRIM55* was over-expressed in human idiopathic dilated hearts.

2.6 Results

2.6.1 Whole-genome sequencing

We sequenced the whole-genome of the HHR and NHR and separately compared each to the BN reference genome (RGSC version 5.0). Post-quality filtering was performed and duplicated sequences produced by PCR amplification were removed. Approximately 1.05 billion reads in the HHR and 977 million reads in the NHR were mapped to the BN genome, resulting in over 77% genome coverage and mean depth of whole-genome exceeding 25% for each strain.

The comparative analysis revealed over 8 million variants in the HHR and 7 million in the NHR (Figure 2.1a and b) relative to the BN reference genome. The majority of these variants were SNPs – over 5.6 million and 4.7 million in the HHR and NHR, respectively. To enrich our comparison between the two genomes under investigation, a comparison of the variants of the HHR to the NHR was performed to identify those unique to either strain. Unique variants represent 2.8 million and 1.7 million of the HHR and NHR genome respectively

(Figure 2.1c and d). In contrast to the overall variant comparison, CNVs were the most prevalent type of variants between HHR and NHR strains.

We also investigated change rates in the HHR and NHR genomes according to analyses of SNPs and InDels. Overall, one variant is present in every 398 and 473 nucleotides in the HHR and NHR, respectively. According to our analysis, chromosome 20 and 12 have the highest change rate in the HHR and NHR, respectively. Meanwhile, the X chromosome has the lowest change rate in both genomes (Table S2.3).

We classified SNPs, insertions and deletions according to zygosity (Table S2.4). As expected, due to the inbred nature of both strains, most SNPs and InDels are homozygous. In regards to location, the majority of SNPs identified were intergenic, and the SNPs located in genomic regions were mostly present in introns in both genomes. Remarkably, no SNPs caused a frameshift in a coding region or loss of stop codon at a splice site (Table 2.1).

The SNPs located in coding regions were classified as synonymous or nonsynonymous. In our study, these SNPs were mostly synonymous not resulting in a change in the amino acid sequence. Furthermore, we classified the nonsynonymous SNPs into missense and nonsense. As expected, most SNPs were missense in both strains (Table S2.5).

2.6.2 Pathways associated with LV hypertrophy

Among the genes with unique SNPs in the HHR, there was a significant overrepresentation of genes involved in adenosine triphosphate-binding cassette (ABC) transporters (rno02010), extracellular matrix (ECM)-receptor interaction (rno04512) and tryptophan metabolism (rno00380), among others (Table S2.6).

2.6.3 QTL Cm22

We focused our gene expression and specific variants analysis present within the QTL Cm22. The overall number of variants and the unique variants were identified for each strain

in this QTL (Table 2.2) and those variants were classified according to their type and region in the genome (Table 2.3).

The microarray data analysis showed that 109 genes were significantly differentially expressed (fold change >1.5) in the LV of neonatal HHR compared with neonatal NHR. The only gene differentially expressed in Cm22 was the gene for tripartite motif 55 (*Trim55*), being significantly down-regulated in our data set (-1.52, adjusted *P*=0.0045). Array results were validated using qPCR in LV from 2-day, 13- and 33-weeks old animals confirming that *Trim55* was down-regulated in HHR from neonatal to late adulthood (Figure 2.2). Protein levels of *Trim55*, measured by Western blot, were also lower in HHR than NHR (Figure 2.3) in LV of 13-week old animals.

In order to identify possible variant targets that could contribute to the hypertrophic phenotype, we explored unique variants found in the *Trim55* gene region (Chr2: 124,140,061-124,182,032) in addition to its upstream and downstream regions (Table 2.4). Our most significant finding was the exonic SNP present in the HHR. This exonic SNP is found in nucleotide position 124,181,039 of chromosome 2 (RGSC v5.0). This nonsynonymous variant leads to a missense mutation (C>T), resulting in a codon change (Ccc>Tcc) and, therefore, a change from a proline to serine in the amino-acid position 513 (P513S). *In silico* protein modelling suggested this variant would lead to a change in the predicted 3D protein structure (Figure 2.4).

Trim55 mRNA levels were negatively correlated with LV mass in the F2 cross (r=-0.16, *P*=0.025) at neonatal age. Although not statistically significant, *Trim55* mRNA abundance was also negatively correlated with CWI at 2 days (r=-0.13, *P*=0.05), 13 weeks (r=-1.04, *P*=0.09) and 33 weeks old (r=-0.74, *P*=0.07) (Figure S2.3). We also found that the DNA sequence variant (P513S) was strongly associated with *Trim55* abundance (F=10.35,

P<0.0001) (Figure 2.5). The allele from the HHR was associated with a lower expression of *Trim55* in the F2 rats, consistent with the association observed in the parental strains.

Finally, we investigated the expression of *TRIM55* mRNA in human idiopathic dilated hearts. In these hearts, we found that the expression of TRIM55 was significantly down-regulated compared to control hearts (-1.5, *P*=0.02, Figure 2.2d).

2.7 Discussion

Our combined sequencing and expression analyses of the QTL Cm22, which we previously discovered on chromosome 2, identified the gene encoding *Trim55* (alias *MuRF2*) as a functional candidate for cardiac hypertrophy. We discovered an exonic variant in *Trim55* that is strongly associated with *Trim55* gene expression in the HHR and the HHRxNHR F2 population. This variant was also associated with reduced Trim55 protein levels in the heart and the *in silico* analysis predicted that the resultant amino acid change would alter protein folding. Furthermore, we have found that *Trim55* mRNA levels were reduced in the hearts from patients with human idiopathic dilated cardiomyopathy.

Our primary goal in this study was to delineate the genetic factors that program the characteristics of the heart in neonatal HHR that predisposes to adult hypertrophy. The clue to the origins of the hypertrophy was our previous discovery of reduced numbers of cardiomyocytes in the adult HHR. (Porrello, Bell, et al., 2009) Importantly we found that soon after birth (neonates) HHR hearts tend to be smaller than NHR and possess fewer cardiomyocytes. (Porrello, Bell, et al., 2009) Around 6 to 10 weeks of age (at the age of reproductive maturity) cardiac hypertrophy begins to become apparent in the HHR hearts to become clearly established in adulthood. The implications of starting life with fewer cardiomyocytes are far-reaching, as the loss of ability to replicate after the neonatal period (switch from proliferative to hypertrophic cellular growth) means that the deficit in cell

numbers cannot be restored. (Ahuja, Sdek, & MacLellan, 2007) HHR hearts entering adolescence with fewer cardiomyocytes are less able to cope with the considerable physiological stresses of body growth (volume and pressure loads). Individual cells must hypertrophy rapidly not only to make up for missing cells but also for the relative reduction in contractile efficiency that often accompanies such hypertrophy. This exaggerated hypertrophy results in a bigger heart than normal in relation to body size after sexual maturity. In our study we find *Trim55* as a primary candidate for the neonatal programming that predetermines adult cardiac hypertrophy and eventual failure in HHR.

Trim55 and the cognate gene *Trim63* (alias *MuRF1*) encode striated muscle specific proteins which are localized at the sarcomere M-line (Su et al., 2014) – the site of myosin crosslinking attachment at the titin C-terminal. This sarcomere region is understood to be of key importance in the transduction of sarcomeric mechanical stress. *Trim55* and *Trim63* are members of a subfamily of the RING-finger E3 ubiquitin ligases involved in mediating replacement turnover of sarcomeric proteins via proteasomal degradation. (Vasiliou, Vasiliou, & Nebert, 2009) Thus, altered regulation of *Trim55* coding could be implicated in impaired sarcomeric protein turnover as a trigger for the hypertrophic growth response.(Willis et al., 2009) Trim55 knockout mice display severe diabetic cardiomyopathy characterised by early onset systolic dysfunction, higher left ventricular mass and increased heart weight when compared to wild-type littermates. (He et al., 2015) Moreover, neonatal *Trim55/Trim63* double knockout mice have severe LV hypertrophy and high mortality rate after birth. (Willis et al., 2014) *Trim63* was not differentially expressed in our data set and variants were not present (data not shown). However, the presence of a variant in the *Trim55* coding region leading to a change in the amino-acid and protein sequence causing a change in the 3D protein structure could influence the protein-protein interaction of *Trim55* and *Trim63* causing a disarray in the sarcomere assembly of cardiac muscle cells, as previously suggested. (Willis et al., 2009)

Furthermore, the prevalence of variants in TRIM55 and TRIM63 is higher in patients with LV hypertrophy than in healthy controls and could contribute to increased risk of LV hypertrophy. (Su et al., 2014)

This study is the first to describe genomic variation in a model of spontaneous cardiac hypertrophy, the HHR, and revealed many potential variants associated with cardiac mass independent of blood pressure that included SNPs, CNVs and InDels.

The HHR strain carried a larger number of variants relative to the NHR strain. Both genomes, however, carried more SNPs and InDels than either of their parental strains, the SHR and Fischer 344 (approximately 3.8 million SNPs and 1.2 million InDels, and 3.4 million SNPs and 1.1 million InDels, respectively). (Atanur et al., 2010; Atanur et al., 2013) In this context, the identification of unique variants to the HHR strain could assist the discovery of genetic variants and, therefore, candidate genes leading to the hypertrophic phenotype. Moreover, variants unique to NHR could present a protection or compensatory mechanism maintaining the normal phenotype.

Surprisingly, the frequency of variants is not uniform throughout the genome, and appeared to show no relation to chromosome size (Table S2.3). Chromosome 1, the largest rat chromosome, had the third largest change rate in the HHR and the seventh in the NHR. In contrast, the seventh largest rat chromosome, the X chromosome, had the largest change rate in both strains. These findings indicate a possible hotspot for mutations in the HHR that could contribute to the development of cardiac disease.

SNPs comprise the largest source of genetic variability in both humans and rats. The location and type of a SNP can differently affect gene expression and lead to diseases and phenotypical changes (Ng & Henikoff, 2006). Nonsynonymous SNPs in an exon affect the amino-acid and, consequently, the protein sequence and are believed to lead to the highest phenotypical changes in any genome. (Ng & Henikoff, 2006; Ramensky, Bork, & Sunyaev,

2002) Meanwhile, intronic and synonymous SNPs do not affect the amino-acid sequence of a gene but could interfere with the splicing site region making site recognition difficult (Harland, Mistry, Bishop, & Bishop, 2001; Pagani & Baralle, 2004) or affect gene and protein expression by changing the shape of the messenger RNA hindering its interaction with the ribosome. (Sauna & Kimchi-Sarfaty, 2011) The intronic SNPs present in the *Trim55* could be hindering the recognition of a splice site, which may influence mRNA levels of this gene.

Besides our *Trim55* findings, our gene functional annotation analysis identified the ATP-binding cassette (ABC) transporters pathway with the highest fold enrichment in our gene set. ABC transporters are membrane proteins involved in energy-dependant transport of substrates across biologic membranes. (Vasiliou et al., 2009) Mutations in its components have been previously described to cause dilated cardiomyopathy in humans. (Solbach, Konig, Fromm, & Zolk, 2006) Moreover, other enriched pathways in our dataset (e.g. extracellular matrix remodelling, tryptophan metabolism and lysosomal compartments) have also been associated with cardiac disease and remodelling. (Mangge et al., 2014; Terman, Kurz, Gustafsson, & Brunk, 2008; Weber, 1997) Nonsynonymous mutations in genes from these pathways may be influencing the LVH observed in the HHR.

This study has some limitations. It would be preferable to sequence a larger number of samples of each strain. Other studies with similar design, however, have also analysed only one sample per strain. (Guo et al., 2013) Moreover, microarrays were analysed in neonatal heart, before LVH is established. The analyses of microarrays in hearts from animals after the development of LVH could identify other genes in the QTL Cm22 differentially expressed with relevant function in this condition.

In conclusion, we have characterised unique variants present in the genome and identified *Trim55* as a candidate gene contributing to LVH. Furthermore, we identified *Trim55*

in the QTL Cm22 as differentially expressed in LVH in both model and humans rendering a novel candidate gene for polygenic LVH independent of blood pressure.

2.8 Acknowledgements

We would like to thank the Ramaciotti Centre for Functional Analyses and their help with the microarrays.

2.9 Funding sources

This work was supported by grants from the National Health & Medical Research Council of Australia (NHMRC) (project grant APP1034371 and 628643), the National Heart Foundation (grant G10M5155) and the Federation University Australia 'Self-sustaining Regions Research and Innovation Initiative', an Australian Government Collaborative Research Network (CRN). P.R.P. is supported by a Robert HT Smith Fellowship from Federation University. F.Z.M. is supported by NHMRC (APP1052659) and National Heart Foundation (PF12M6785) co-shared Early Career Fellowships.

2.10 References

Ahuja, P., Sdek, P., & MacLellan, W. R. (2007). Cardiac myocyte cell cycle control in development, disease, and regeneration. Physiol Rev, 87(2), 521-544.

Atanur, S. S., Birol, I., Guryev, V., Hirst, M., Hummel, O., Morrissey, C., . . . Aitman, T. J. (2010). The genome sequence of the spontaneously hypertensive rat: Analysis and functional significance. Genome Res, 20(6), 791-803.

Atanur, S. S., Diaz, A. G., Maratou, K., Sarkis, A., Rotival, M., Game, L., . . . Aitman, T. J. (2013). Genome sequencing reveals loci under artificial selection that underlie disease phenotypes in the laboratory rat. Cell, 154(3), 691-703.

Busjahn, C. A., Schulz-Menger, J., Abdel-Aty, H., Rudolph, A., Jordan, J., Luft, F. C., & Busjahn, A. (2009). Heritability of left ventricular and papillary muscle heart size: a twin study with cardiac magnetic resonance imaging. Eur Heart J, 30(13), 1643-1647.

Cingolani, P., Platts, A., Wang le, L., Coon, M., Nguyen, T., Wang, L., . . . Ruden, D. M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly (Austin), 6(2), 80-92.

Di Nicolantonio, R., Kostka, V., Kwitek, A., Jacob, H., Thomas, W. G., & Harrap, S. B. (2006). Fine mapping of Lvm1: a quantitative trait locus controlling heart size independently of blood pressure. Pulm Pharmacol Ther, 19(1), 70-73.

Guo, X., Brenner, M., Zhang, X., Laragione, T., Tai, S., Li, Y., . . . Gulko, P. S. (2013). Whole-genome sequences of DA and F344 rats with different susceptibilities to arthritis, autoimmunity, inflammation and cancer. Genetics, 194(4), 1017-1028.

Harland, M., Mistry, S., Bishop, D. T., & Bishop, J. A. (2001). A deep intronic mutation in CDKN2A is associated with disease in a subset of melanoma pedigrees. Hum Mol Genet, 10(23), 2679-2686.

Harrap, S. B., Danes, V. R., Ellis, J. A., Griffiths, C. D., Jones, E. F., & Delbridge, L. M. (2002). The hypertrophic heart rat: a new normotensive model of genetic cardiac and cardiomyocyte hypertrophy. Physiol Genomics, 9(1), 43-48.

He, J., Quintana, M. T., Sullivan, J., T, L. P., T, J. G., Schisler, J. C., . . . Willis, M. S. (2015). MuRF2 regulates PPARgamma1 activity to protect against diabetic cardiomyopathy and enhance weight gain induced by a high fat diet. Cardiovasc Diabetol, 14, 97.

Huang da, W., Sherman, B. T., & Lempicki, R. A. (2009). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res, 37(1), 1-13.

Huang, D. W., Sherman, B. T., & Lempicki, R. A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nature Protocols, 4(1), 44-57.

Innes, B. A., McLaughlin, M. G., Kapuscinski, M. K., Jacob, H. J., & Harrap, S. B. (1998). Independent genetic susceptibility to cardiac hypertrophy in inherited hypertension. Hypertension, 31(3), 741-746.

Kelley, L. A., & Sternberg, M. J. (2009). Protein structure prediction on the web: a case study using the Phyre server. Nat Protoc, 4(3), 363-371.

Koboldt, D. C., Chen, K., Wylie, T., Larson, D. E., McLellan, M. D., Mardis, E. R., . . . Ding, L. (2009). VarScan: variant detection in massively parallel sequencing of individual and pooled samples. Bioinformatics, 25(17), 2283-2285.

Laulederkind, S. J., Hayman, G. T., Wang, S. J., Smith, J. R., Lowry, T. F., Nigam, R., . . . Jacob, H. J. (2013). The Rat Genome Database 2013--data, tools and users. Brief Bioinform, 14(4), 520-526.

Levy, D., Garrison, R. J., Savage, D. D., Kannel, W. B., & Castelli, W. P. (1990). Prognostic implications of echocardiographically determined left ventricular mass in the Framingham Heart Study. N Engl J Med, 322(22), 1561-1566.

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics, 25(14), 1754-1760.

Mangge, H., Stelzer, I., Reininghaus, E. Z., Weghuber, D., Postolache, T. T., & Fuchs, D. (2014). Disturbed tryptophan metabolism in cardiovascular disease. Curr Med Chem, 21(17), 1931-1937.

Mayosi, B. M., Keavney, B., Kardos, A., Davies, C. H., Ratcliffe, P. J., Farrall, M., & Watkins, H. (2002). Electrocardiographic measures of left ventricular hypertrophy show greater heritability than echocardiographic left ventricular mass. Eur Heart J, 23(24), 1963-1971.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., . . . DePristo, M. A. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res, 20(9), 1297-1303.

Nabika, T., Ohara, H., Kato, N., & Isomura, M. (2012). The stroke-prone spontaneously hypertensive rat: still a useful model for post-GWAS genetic studies? Hypertens Res, 35(5), 477-484.

Ng, P. C., & Henikoff, S. (2006). Predicting the effects of amino acid substitutions on protein function. Annu Rev Genomics Hum Genet, 7, 61-80.

Pagani, F., & Baralle, F. E. (2004). Genomic variants in exons and introns: identifying the splicing spoilers. Nat Rev Genet, 5(5), 389-396.

Picard tools. Retrieved from http://sourceforge.net/projects/picard/. http://sourceforge.net/projects/picard/

Porrello, E. R., Bell, J. R., Schertzer, J. D., Curl, C. L., McMullen, J. R., Mellor, K. M., . . . Delbridge, L. M. (2009). Heritable pathologic cardiac hypertrophy in adulthood is preceded by neonatal cardiac growth restriction. Am J Physiol Regul Integr Comp Physiol, 296(3), R672-680.

Ramensky, V., Bork, P., & Sunyaev, S. (2002). Human non-synonymous SNPs: server and survey. Nucleic Acids Res, 30(17), 3894-3900.

Rat Genome Database Website. (2015). Quantitative trait loci cardiac mass 22 location (Publication no. World Wide Web (URL: http://rgd.mcw.edu/)). Retrieved January 2015, from Medical College of Wisconsin, Milwaukee, Wisconsin

Rozen, S., & Skaletsky, H. (Eds.). (2000). Primer3 on the WWW for general users and for biologist programmers. Totowa, NJ: Humana Press.

Sauna, Z. E., & Kimchi-Sarfaty, C. (2011). Understanding the contribution of synonymous mutations to human disease. Nat Rev Genet, 12(10), 683-691.

Schmittgen, T. D., & Livak, K. J. (2008). Analyzing real-time PCR data by the comparative CT method. Nat Protoc, 3(6), 1101-1108.

Sharma, P., Middelberg, R. P., Andrew, T., Johnson, M. R., Christley, H., & Brown, M. J. (2006). Heritability of left ventricular mass in a large cohort of twins. J Hypertens, 24(2), 321-324.

Solbach, T. F., Konig, J., Fromm, M. F., & Zolk, O. (2006). ATP-binding cassette transporters in the heart. Trends Cardiovasc Med, 16(1), 7-15.

Su, M., Wang, J., Kang, L., Wang, Y., Zou, Y., Feng, X., . . . Song, L. (2014). Rare variants in genes encoding MuRF1 and MuRF2 are modifiers of hypertrophic cardiomyopathy. Int J Mol Sci, 15(6), 9302-9313.

Terman, A., Kurz, T., Gustafsson, B., & Brunk, U. T. (2008). The involvement of lysosomes in myocardial aging and disease. Curr Cardiol Rev, 4(2), 107-115.

Van der Auwera, G. A., Carneiro, M. O., Hartl, C., Poplin, R., Del Angel, G., Levy-Moonshine, A., . . . DePristo, M. A. (2013). From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. Curr Protoc Bioinformatics, 11(1110), 11 10 11-11 10 33.

Vasiliou, V., Vasiliou, K., & Nebert, D. W. (2009). Human ATP-binding cassette (ABC) transporter family. Hum Genomics, 3(3), 281-290.

Weber, K. T. (1997). Extracellular matrix remodeling in heart failure: a role for de novo angiotensin II generation. Circulation, 96(11), 4065-4082.

Willis, M. S., Rojas, M., Li, L., Selzman, C. H., Tang, R. H., Stansfield, W. E., . . . Patterson, C. (2009). Muscle ring finger 1 mediates cardiac atrophy in vivo. Am J Physiol Heart Circ Physiol, 296(4), H997-H1006.

Willis, M. S., Wadosky, K. M., Rodriguez, J. E., Schisler, J. C., Lockyer, P., Hilliard, E. G., . . . Patterson, C. (2014). Muscle ring finger 1 and muscle ring finger 2 are necessary but functionally redundant during developmental cardiac growth and regulate E2F1-mediated gene expression in vivo. Cell Biochem Funct, 32(1), 39-50.

2.11 Tables

Table 2.1. Classification of all single nucleotide polymorphisms according to type and region in the genome.

Genome region	HHR	NHR
3' UTR	14,842	12,251
5' UTR	2,295	2,031
Downstream	241,403	203,590
Frameshift coding	-	-
Intergenic	4,253,768	3,554,880
Intronic	1,352,486	1,128,853
Processed pseudogene	6,066	5,133
Pseudogene	39,185	33,264
Splice site, 3' UTR	21	16
Splice site, 5' UTR	34	25
Splice site, intronic	2,731	2,303
Splice site, nonsynonymous coding	410	369
Splice site, synonymous coding	484	411
Start codon gained	356	297
Start codon lost	16	14
Stop gained	319	258
Stop gained at splice site	11	5
Stop lost	19	19
Stop lost at splice site	-	-
Upstream	247,516	209,474

Within mature microRNA	18,857	16,332
Synonymous coding	22,826	18,781
Nonsynonymous coding	19,033	16,002

Legend: UTR, untranslated region.

Table 2.2. Variants identified in the quantitative trait locus cardiac mass 22.

Type of variant	All		Unique	
	HHR	NHR	HHR	NHR
SNPs	182,404	153,527	45,731	16,854
InDels	46,787	10,038	15,954	10,038

Legend: SNP, single nucleotide polymorphisms; InDel, insertions and deletions.

Table 2.3. Single nucleotide polymorphisms in the quantitative trait locus cardiac mass 22 classified by type and region of the genome.

Genome region	All		Unique	
	HHR	NHR	HHR	NHR
3' UTR	305	254	60	9
5' UTR	46	40	11	5
Downstream	3,815	3,296	853	334
Frameshift coding	-	-	-	-
Intergenic	148,512	124,158	38,712	14,358
Intronic	32,686	28,288	5,896	2,416
Processed pseudogene	62	47	21	6
Pseudogene	948	816	304	172
Splice site, 3' UTR	-	-	-	-
Splice site, 5' UTR	1	1	-	-
Splice site, intronic	43	38	6	1
Splice site, nonsynonymous coding	8	6	2	-
Splice site, synonymous coding	6	5	1	-
Start codon gained	8	7	2	1
Start codon lost	-	-	-	-
Stop gained	6	4	2	-
Stop gained at splice site	-	-	-	-
Stop lost	-	-	-	-
Stop lost at splice site	-	-	-	-
Upstream	3,979	3,533	871	425
Within mature microRNA	372	256	146	30

Synonymous coding	384	328	79	23
Nonsynonymous coding	404	372	73	41

Legend: UTR, untranslated region.

Table 2.4. Genomic region of variants in *Trim55* unique to each rat strain.

Type of variant	Genomic region	HHR	NHR
SNP	Upstream	0	0
	Exonic	1	0
	Intronic	8	0
	Downstream	1	0
InDel	Upstream	3	1
	Exonic	0	0
	Intronic	7	12
	Downstream	0	2

Legend: SNP, single nucleotide polymorphisms; InDel, insertions and deletions.

2.12 Figures

Figure 2.1. Whole-genome sequencing analysis of variants. **a**, all variants in the hypertrophic heart rat (HHR) and **b**, the normal heart rat (NHR). **c**, unique variants in the HHR and **d**, NHR. SNP: single nucleotide polymorphism, InDel: insertions and deletions, CNV: copy number variation and SV: structural variants.

Figure 2.2. *Trim55* mRNA expression relative to *Gapdh* in the hypertrophic heart rat (HHR), normal heart rat (NHR) and human hearts. *Trim55* is down-regulated in **a**, neonatal, **b**, 13-week old, **c**, 33-week old animals and **d**, human idiopathic dilated hearts. Graphs represent mean, error bars represent standard error. * indicates $P<0.05$, ** $P<0.01$, *** $P<0.001$.

Figure 2.3. Representative Western blot analysis showing protein levels of *Trim55* in the left ventricle of the hypertrophic heart rat (HHR) and the normal heart rat (NHR) at 13 weeks of age.

Figure 2.4. *In silico* 3D modelling of *Trim55* protein structure. **a**, the predicted structure in the normal protein sequence and **b**, including exonic SNP variant present in the HHR sequence using online software Phyre. A predicted change in the protein domain is circled in red. HHR: hypertrophic heart rat; NHR: normal heart rat.

Figure 2.5. *Trim55* abundance relative to *Gapdh* in 2-day old rats from the HHRxNHR F2 population according to exonic variant genotype. Graphs represent mean, error bars represent standard error. $F=10.35$, $P<0.0001$.

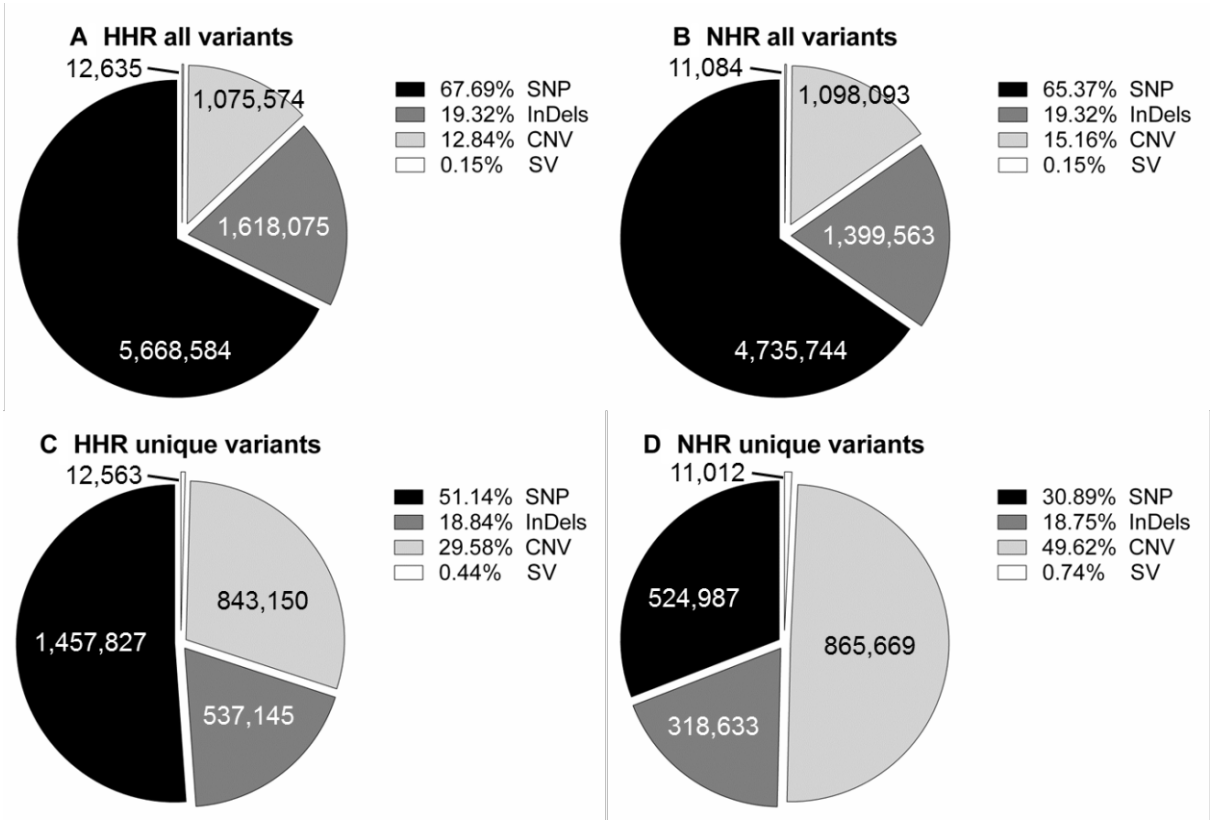
A HHR all variants
12,635
1,075,574
1,618,075
5,668,584
67.69% SNP
19.32% InDels
12.84% CNV
0.15% SV
B NHR all variants
11,084
1,098,093
1,399,563
4,735,744
65.37% SNP
19.32% InDels
15.16% CNV
0.15% SV
C HHR unique variants
12,563
843,150
1,457,827
537,145
51.14% SNP
18.84% InDels
29.58% CNV
0.44% SV
D NHR unique variants
11,012
524,987
865,669
318,633
30.89% SNP
18.75% InDels
49.62% CNV
0.74% SV

Figure 2.1

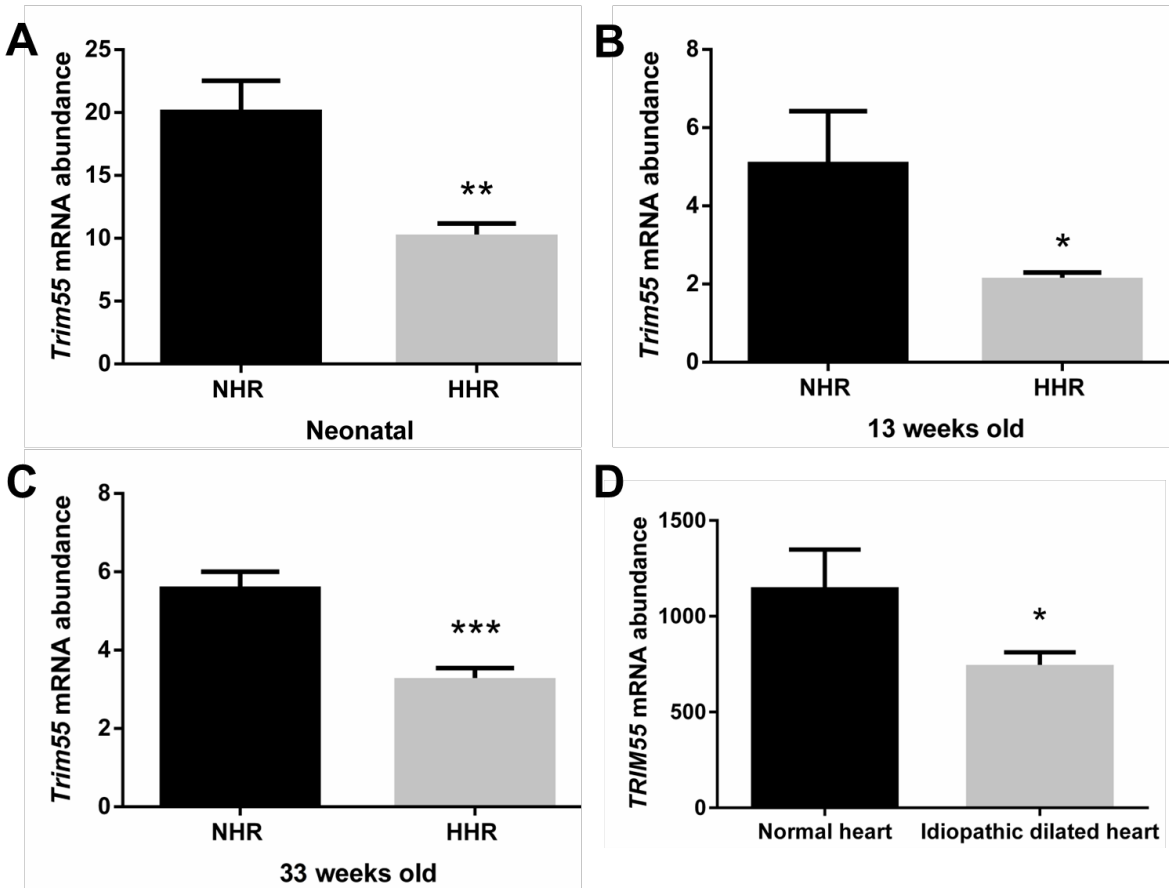
A
Trim55 mRNA abundance
25
20
15
10
5
0
**
NHR
HHR
Neonatal
B
Trim55 mRNA abundance
8
6
4
2
0
*
NHR
HHR
13 weeks old
C
Trim55 mRNA abundance
8
6
4
2
0
***
NHR
HHR
33 weeks old
D
TRIM55 mRNA abundance
1500
1000
500
0
*
Normal heart
Idiopathic dilated heart

Figure 2.2

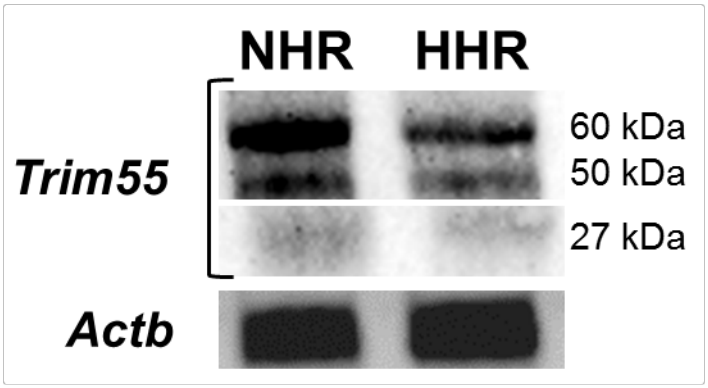

Figure 2.3

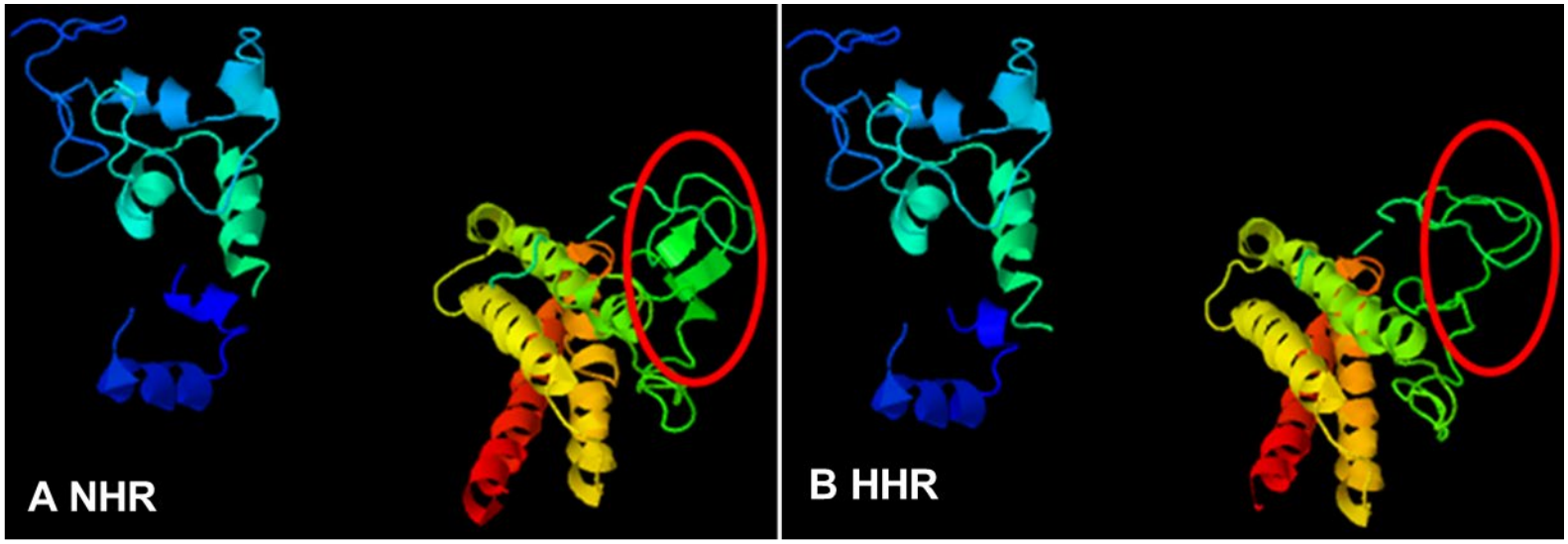

Figure 2.4

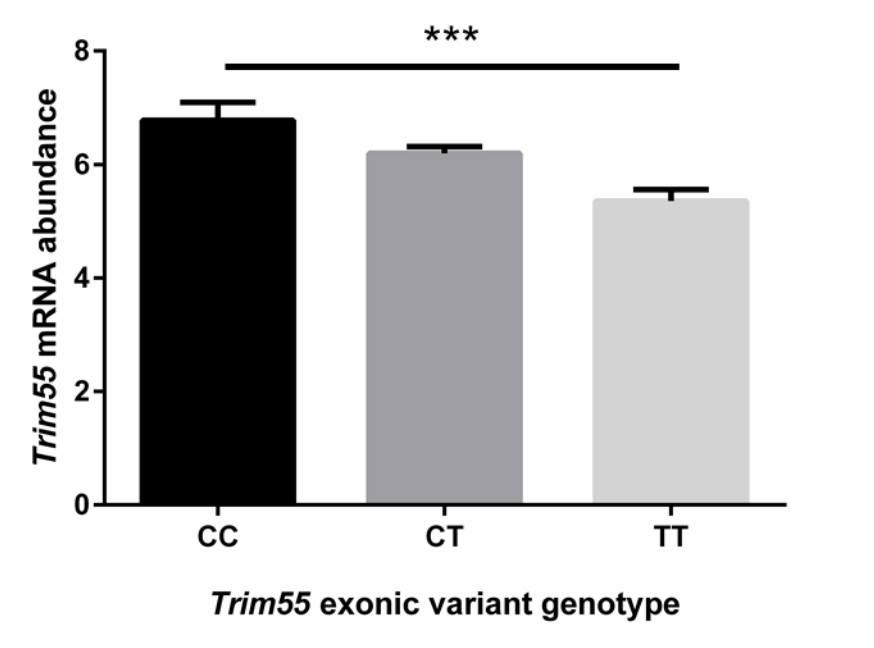
***
8
6
4
2
0
Trim55 mRNA abundance
CC
CT
TT
Trim55 exonic variant genotype

Figure 2.5

Chapter 3. Involvement of human monogenic cardiomyopathy genes in experimental polygenic cardiac hypertrophy

Published in: Prestes PR, Marques FZ, Lopez-Campos G, Lewandowski P, Delbridge LMD, Charchar FJ, Harrap SB (2018). Involvement of human monogenic cardiomyopathy genes in experimental polygenic cardiac hypertrophy. Physiol Genomics; 50(9):680-687. doi: 10.1152/physiolgenomics.00143.2017.

INVOLVEMENT OF HUMAN MONOGENIC CARDIOMYOPATHY GENES IN EXPERIMENTAL POLYGENIC CARDIAC HYPERTROPHY

Short title: CARDIOMYOPATHY GENES IN POLYGENIC CARDIAC HYPERTROPHY

Prestes PR[a], Marques FZ[a,b], Lopez-Campos G[c,d], Lewandowski P[e], Delbridge LMD[f], Charchar FJ[a], Harrap SB[f]

[a]School of Applied and Biomedical Sciences, Faculty of Science and Technology, Federation University Australia, Ballarat, VIC; [b]Heart Failure Research Laboratory, Baker Heart and Diabetes Research Institute, Melbourne, VIC; [c]Wellcome-Wolfson Institute for Experimental Medicine, Queen's University of Belfast; [d]Health and Biomedical Informatics Centre, University of Melbourne, Melbourne, VIC; [e]School of Medicine, Deakin University, Geelong, VIC; [f]Department of Physiology, University of Melbourne, Melbourne, VIC.

Corresponding author: Stephen Harrap, Department of Physiology, University of Melbourne, N421, Level 4, Medical Building, Cnr Grattan Street and Royal Parade, Parkville, 3052, Victoria, Australia. Tel: +61 3 8344 5837, Fax: +61 3 9349 4519; e-mail: s.harrap@unimelb.edu.au

Author contribution

PRP, LMDD, FJC and SBH conception and study design; PRP, FZM, GL and PL performed experiments; PRP analyzed data; PRP interpreted results; PRP prepared figures; PRP drafted

manuscript; PRP, FZM, GL, PL, LMDD, FJC and SBH edited and revised manuscript; PRP, FZM, GL, PL, LMDD, FJC and SBH approved final version of manuscript.

3.1 Abstract

Hypertrophic cardiomyopathy thickens heart muscles reducing functionality and increasing risk of cardiac disease and morbidity. Genetic factors are involved, but their contribution is poorly understood. We used the hypertrophic heart rat (HHR), a unique normotensive polygenic model of cardiac hypertrophy and heart failure to investigate the role of genes associated with monogenic human cardiomyopathy. We selected 42 genes involved in monogenic human cardiomyopathies to study: 1) DNA variants, by sequencing the whole-genome of 13-week old HHR and age-matched normal heart rat (NHR), its genetic control strain; 2) mRNA expression, by targeted RNA-sequencing in left ventricles of HHR and NHR at five ages (2-days old, 4-, 13-, 33- and 50-weeks old) compared to human idiopathic dilated data; and 3) microRNA expression, with rat microRNA microarrays in left ventricles of 2-days old HHR and age-matched NHR. We also investigated experimentally validated microRNA-mRNA interactions. Whole-genome sequencing revealed unique variants mostly located in non-coding regions of HHR and NHR. We found 29 genes differentially expressed in at least one age. Genes encoding desmoglein 2 (*Dsg2*) and transthyretin (*Ttr*) were significantly differentially expressed at all ages in the HHR, but only *Ttr* was also differentially expressed in human idiopathic cardiomyopathy. Lastly, only two microRNAs differentially expressed in the HHR were present in our comparison of validated microRNA-mRNA interactions. These two microRNAs interact with five of the genes studied. Our study shows that genes involved in monogenic forms of human cardiomyopathies may also influence polygenic forms of the disease.

3.2 Keywords

DNA sequencing, gene expression, microRNA, cardiac hypertrophy, cardiomyopathy

3.3 Introduction

Cardiovascular disease (CVD) is the main cause of death and morbidity worldwide, killing approximately 17.5 million people in 2012 (Lim et al., 2012; World Health Organization, 2016). Both genetic and environmental factors contribute to CVD. The most common genetic contributions are considered to be polygenic, although the exact number and nature of genes involved has been difficult to determine as new genes are yet to be identified with the advancement of technology (Hwang, Dempsey, Lee, & Liew, 2000). However, less common monogenic causes of CH have been well characterized in terms of the causative DNA variants and pathophysiology (Friedrich & Carrier, 2012; Maron & Maron, 2013). Variation of the expression of genes involved in monogenic CH might provide clues to the causes of polygenic etiology of CH.

Hypertrophic cardiomyopathy (HCM) is the most common inherited form of CVD, affecting one in 500 adults and is the major cause of heart failure and sudden death in young people (Kimura, 2016; Maron, 2002; Maron & Maron, 2013). The condition is characterized by the asymmetric thickening of the cardiac wall, heart failure and risk of sudden death. A variety of mutations in genes coding sarcomere and cardiac filament proteins account for over 88% of familial HCM (Cirino, 2008 [updated 2014]; Maron & Maron, 2013; Richard et al., 2003). Mutations in two genes that encode myosin heavy chain 7 (*MYH7*) and cardiac myosin binding protein C (*MYBPC3*) are the most common causes of monogenic HCM (Richard et al., 2003; Xu, Dewey, Nguyen, & Gomes, 2010). Familial dilated cardiomyopathies form another important group of monogenic CVD characterized by CH and heart failure, for which mutations in genes encoding sarcomeric proteins account for almost half of the known forms (Hershberger & Morales, 2007 [updated 2015]).

The study of genes involved in monogenic hypertrophy and failure have led to an understanding of disease mechanisms and might also provide explanations for more common polygenic forms of heart failure.

We have developed and characterized the hypertrophic heart rat (HHR), a unique polygenic normotensive model of spontaneous ventricular hypertrophy, cardiac failure and premature death (Harrap et al., 2002). Compared with their genetic control strain, the normal heart rat (NHR), HHR begin life with fewer cardiomyocytes that develop cellular hypertrophy leading to cardiac enlargement and heart failure (Porrello, Bell, et al., 2009).

Our aim was to study genes previously associated with human monogenic forms of dilated and hypertrophic cardiomyopathies in the polygenic etiology of cardiac hypertrophy in the HHR. We combined analyses of RNA expression and DNA sequence variation to identify those genes that might be of importance in the polygenic setting.

3.4 Materials and methods

3.4.1 Sample collection

The HHR and NHR strains have been described in detail elsewhere (Harrap et al., 2002). Two-day-old HHR and NHR (n=8 HHR, n=9 NHR) were euthanized by decapitation. At 4-, (n=10 HHR, n=10 NHR); 13-, (n=10 HHR, n=11 NHR); 33- (n=7 HHR, n=9 NHR) and 50-weeks of age (n=12 HHR, n=10 NHR) rats were euthanized using a lethal dose of pentobarbitone (Lethobarb). Hearts were removed and left ventricles (LV) were immediately dissected from atria. Cardiac weight indexes (CWI, mg/g) were calculated as the total heart weight (mg) relative to total body weight (g) of each animal (ref 28 and Figure 3.1).

The five age groups investigated represent different developmental stages in their life. Prior to hypertrophy (2-day-old), during the development of hypertrophy (4-weeks), early hypertrophy (13-weeks), established hypertrophy (33-weeks), hypertrophy complicated by heart failure (50-weeks-old). This study was approved by the Animal Ethics Committees of the University of Melbourne and Deakin University, and ratified at Federation University Australia.

3.4.2 DNA and RNA extraction

DNA from LV was extracted using PureLink® Genomic Extraction kit (ThermoFisher Scientific). RNA from LV was extracted using miRNeasy kit (Qiagen). DNA was quantified by spectrophotometry using NanoDrop® 2000 and RNA was quantified by fluorescence using Qubit™ 3.0 Fluorometer and the RNA high sensitivity assay kit (ThermoFisher Scientific).

3.4.3 Genes investigated

For these focused studies we selected 42 genes (Table 3.1) involved in monogenic forms of familial cardiomyopathies (Cirino, 2008 [updated 2014]; Hershberger & Morales, 2007 [updated 2015]; Richard et al., 2003; Xu et al., 2010). Physiologically, most of these genes are involved in growth and contractility, regulation of mechanical-stress, calcium channels and a variety of muscle development pathways, mainly for cardiac filaments and sarcomere assembly (Cirino, 2008 [updated 2014]; Richard et al., 2003).

DNA sequence variants in HHR were identified according to methods detailed previously (Prestes et al., 2016). Briefly, we sequenced the whole-genome of one male 13-week-old NHR and one age-matched HHR. Variants were analyzed according to GATK best practices (Van der Auwera et al., 2013) and functional annotation was performed using SnpEff software (Cingolani et al., 2012). Results were stored in a database developed "in-house" and then we identified unique single nucleotide polymorphisms (SNPs) and insertions/deletions (InDels) in the NHR and HHR (Prestes et al., 2016).

mRNA expression of the genes listed in Table 3.1 was measured using Targeted RNA Expression (TREx) custom panel in the MiSeq Desktop sequencer (Illumina®) and analyzed using MSR: Targeted RNA v2.4.60.8 on Illumina® BaseSpace. False discovery rate (FDR) was set as <0.1.

microRNA (miRNA) arrays were conducted using the Agilent rat microRNA microarray kit 16.0 in left ventricles of 2 days old male HHR and age-matched NHR (n=4/group). The data obtained has been deposited in the National Center for Biotechnology

Information Gene Expression Omnibus database with series accession number GSE38710. Differentially expressed miRNAs were identified using Partek Genomics Suit v6.6 with FDR set as <0.05.

3.4.4 *In silico* investigations

We combined *in silico* approaches to explore a link between gene expression and DNA sequencing data from the HHR and NHR and human studies available online.

We investigated the possibility that miRNAs might be involved in regulating mRNA expression pre- and post-transcriptionally using three algorithms individually (miRWalk 2.0, (Dweep & Gretz, 2015; Dweep, Sticht, Pandey, & Gretz, 2011) miRanda (Betel, Koppal, Agius, Sander, & Leslie, 2010) and TargetScan (Garcia et al., 2011)) as a comparative platform to predict possible miRNA binding sites within the sequence in and around each gene of interest. We also used miRWalk 2.0 to investigate experimentally validated miRNA-mRNA interactions and evaluate which miRNAs targeted the genes under investigation.

As DNA methylation can modulate gene expression by compacting DNA sequences, we investigated if any of the DNA sites were differentially methylated in CVD using the Disease Meth database 2.0 (Xiong et al., 2017).

3.4.5 Gene expression in human hearts

We also investigated the mRNA expression of the genes listed identified in HHR in human cardiac samples from the dataset “heart failure arising from different etiologies” in the repository Gene Expression Omnibus (GEO) reference series GSE1145 (Barrett et al., 2013; Edgar, Domrachev, & Lash, 2002) (n=11 control hearts and n=15 idiopathic dilated hearts). We then determined the expression of the genes investigated using the GEO tool GEO2R.

3.5 Results

3.5.1 DNA Sequence Analyses

DNA sequencing in and around the 42 genes of interest revealed greater number of unique DNA variants discovered in the HHR (compared to the rat reference genome) with 851 SNPs and 491 InDels, as opposed to 383 SNPs and 316 InDels in the NHR (Table 3.1). We found no evidence of DNA sequence variation in 6 genes in the NHR (namely *Csrp3, Emd*, *Gla*, *Mylk2*, *Pln* and *Taz*) and 3 genes in the HHR (*Emd*, *Myoz2* and *Tpm1***,** Table 3.1).

In both the HHR and NHR, most unique variants were located in intergenic or intronic regions. However, we found 11 unique synonymous SNPs in exonic regions in the HHR and 2 in the NHR. These SNPs were located in genes encoding actinin alpha 2 (*Actn2*), desmin (*Des*), ryanodine receptor 2 (*Ryr2*), troponin T type 2 (*Tnnt2*) and vinculin (*Vcl*) in the HHR; and *Ryr2* and *Vcl* in the NHR. Interestingly, only one unique non-synonymous missense variant (gGt>gTt) in an exonic region was found in the *Vcl* gene in the NHR changing the aminoacid from glycine to valine.

3.5.2 Cardiac RNA Expression Analyses

The relative fold changes (FC) in mRNA expression in the HHR relative to NHR were never greater than 4-fold for any of the 42 genes under investigation. We found that 29 of the 42 genes showed significant differential expression for at least one age (Figure 3.2, 3.3 and 3.4).

Genes differentially expressed were not consistent throughout the age groups, possibly reflecting developmental stage-specific regulation. At 2 days old, only four genes were differentially expressed (Figure 3.2a). In contrast, at 4 weeks of age, we observed 50% of genes differentially expressed, the highest prevalence of all the age groups (Figure 3.2b). Most of the differences at this age presented increased expression. For some genes (such as *Gla*, *Jup*, *Lamp2* and *Pln*) differential expression was evident only at 4 weeks of age. The differential

expression of other genes (such as *Actc1*, *Cav3* and *Fhl2*) became first evident at 4 weeks of age and then persisted throughout adulthood. Differences in the expression of other genes (such as *Tnnt2*, *Myoz2* and *Myl3*) appeared at 4 weeks of age, but disappeared in later adulthood. Still other genes (such as *Ankrd1*) did not show differential expression until later adulthood (Figure 3.2, 3.3 and 3.4).

Two genes, those encoding desmoglein 2 (*Dsg2*) and transthyretin (*Ttr*) were significantly differentially expressed at all ages in the HHR compared with NHR. *Dsg2* was underexpressed in the hearts of HHR at all ages sampled, whereas *Ttr* was significantly downregulated at day 2 in the neonatal period but significantly upregulated at all subsequent ages corresponding to the development of cardiomyocyte hypertrophy (Figure 3.2). *Ttr* also showed greater expression in adult human idiopathic cardiomyopathy (Figure 3.4).

The analysis of human idiopathic cardiomyopathy revealed that 16 orthologous genes (supplementary table 3.1) differentially expressed of the 42 genes investigated. We found only 9 of those 16 genes were also differentially expressed in at least one rat age.

3.5.3 Predicted miRNA binding sites

We also predicted possible miRNA binding sites within and around each gene region. Potentially, there are over 220,000 miRNA binding sites in the gene coding regions alone and almost 54,000 in the promoter regions (Table 3.2). However, our comparison of validated miRNA-mRNA interactions to miRNAs differentially expressed in HHR compared to NHR in our microarray data found two miRNAs (miR-34a-5p and miR-17-5p) upregulated in the HHR (Table 3.3). Interestingly, those two miRNAs interact with five of the genes investigated, plakophilin 2 (*Pkp2*), RNA binding motif protein 20 (*Rbm20*), ryanodine receptor 2 (*Ryr2*), tropomyosin 1 (*Tpm1*) and vinculin (*Vcl*) (Helwak, Kudla, Dudnakova, & Tollervey, 2013; Kaller et al., 2011; Lipchina et al., 2011). Although not statistically significant, the gene expression of *Pkp2*, *Ryr2*, *Tpm1* and *Vcl* is downregulated in 2 days old HHR (FC=0.93, 0.95, 0.89 and 0.88, respectively).

3.5.4 Other analyses

Scans of published GWAS (using GWASdb v2 (M. J. Li et al., 2016)) data also indicated the presence of SNPs found in the 42 genes investigated that are associated with CVD traits in humans (Table 3.4).

Surprisingly no methylation profiles have been reported in human heart related diseases for any of the genes in our dataset.

3.6 Discussion

The HHR and their NHR genetic control are derived from the spontaneously hypertensive rat (SHR) and Fisher 344 rat and provide a unique model of left ventricular hypertrophy independent of high blood pressure and heart failure. Although HHR and NHR have a polygenic background, it is not unreasonable to presume that genes best known for their major mutations causing human cardiac and failure might also encompass DNA variants with a more subtle quantitative impact on cardiac structure and function. Here we sought to determine whether any of the 42 genes implicated in Mendelian human hypertrophic and dilated cardiomyopathies might also be relevant to the polygenic hypertrophy of HHR – either as DNA sequence variants or abnormal gene expression patterns. We had the advantage with our life course approach to also examine the ontogeny of the expressions of these genes and their relationships with the developmental stages of hypertrophy in the HHR.

Molecular genetic studies have identified over 1400 mutations responsible for inherited cardiomyopathies and genetic testing is offered to identify genetic causes in families or assess family members at risk (Maron & Maron, 2013; Towbin, 2014). However, emerging data suggests that compound mutations have an additive role causing a gene dosage effect influencing the severity and progression of HCM (Ingles et al., 2005; Kelly & Semsarian, 2009; Maron, Maron, & Semsarian, 2012a). Such interactions at the level of monogenic

cardiomyopathic disease provide general support for the hypothesis that interaction between other DNA variants in the same genes might influence polygenic cardiac hypertrophy.

Among the sequence variants identified in our sequencing analyses we found 11 synonymous SNPs in exonic regions in the HHR and one nonsynonymous SNP in the NHR. It is well established that nonsynonymous mutations may impact aminoacid sequences and therefore human health (Jubb et al., 2016). However, synonymous variants which had previously been regarded as "silent" mutations and thought to have no effect on disease phenotype, have been reported to also cause changes in protein-protein interactions. Importantly, studies suggest that these synonymous mutations contribute to human disease risk and other complex traits (Sauna & Kimchi-Sarfaty, 2011). The synonymous variants we discovered (Table 3.1) might result in changes in protein-protein interactions and explain changes in gene expression (Figure 3.1) due to imbalanced availability of tRNA caused by codon bias (Gustilo, Vendeix, & Agris, 2008).

Interestingly, HHR are born with smaller and fewer cardiomyocytes than NHR. We believe that this reduced endowment is important for the subsequent development of cardiomyocyte hypertrophy, which becomes evident in early adolescence (4 weeks, Figure 3.1), when increased pressure and volume loads in the growing animals place proportionately greater stress on the fewer individual myocytes (Porrello, Bell, et al., 2009). By 13 weeks of age, both cardiac and cardiomyocyte hypertrophy are established (ref 13 and Figure 3.1), but in the end this is counterproductive, as the enlarged cells do not function efficiently and gradually deteriorate, leading to heart failure (evident as early as 30 weeks of age) and premature death (towards 50 weeks of age) (Harrap et al., 2002; Porrello, Bell, et al., 2009). These pathophysiological phenotypic changes can provide context for the developmental stage-specific changes in gene expression we observed.

Our analysis of gene expression showed that at 2 days old, when HHR hearts are smaller than NHR (Porrello, Bell, et al., 2009), the genes encoding cardiac filaments and sarcomeric

proteins are generally underexpressed (Figure 3.2). Whether this is a cause or effect of the fewer, smaller cells is not possible to say from these data. As the cells begin to hypertrophy and throughout the rest of their lives, most of these genes are overexpressed. These findings demonstrate that the left ventricle undergoes changes in gene expression that could be related to the pathophysiology of the hypertrophy (Wagner, Tabibiazar, Powers, Bernstein, & Quertermous, 2004). Interestingly, the switch of the *Ttr* gene from underexpressed in neonates to overexpressed in adults suggest stage-specific regulation. *Ttr* encodes a protein for exosome production and high protein levels in serum have been previously associated with lung cancer (Ding et al., 2014) and heart failure caused by accumulation of transthyretin amyloid fibrils in the heart (Ton, Mukherjee, & Judge, 2014). Additionally, the genes for cardiac alpha actin 1 (*Actc1*), caveolin-3 (*Cav3*) and four and a half LIM domains 2 (*Fhl2*) are upregulated in adult HHR and have been associated with cardiomyopathies in mice and humans (Aravamudan et al., 2003; Morita et al., 2008). Conversely, the persistent underexpression of *Dsg2* suggests a constitutive difference between HHR and NHR that appears independent of the pathophysiological changes with age.

Interestingly, the majority of genes identified in HHR were not differentially expressed in the human samples (Figure 3.4B). This might reflect the nature of the diseases in the human repository that is composed primarily of samples from human idiopathic dilated cardiomyopathy rather than Mendelian human hypertrophic cardiomyopathy. Furthermore, as we were unable to establish the age of the human samples, we cannot make direct comparisons to our rat data to further our understanding of gene expression patterns with the progression of HCM.

In order to provide a comprehensive perspective about the processes involved in this pathology, we investigated DNA methylation profiles and SNPs published in GWAS data of the genes investigated. We were unable to find any relevant methylation data supporting further analyses as the majority of the studies available was performed in a variety of cancers and only

few and individual studies are available in CVD. However, two SNPs found in the GWAS data located in the cardiac troponin T2 (*TNNT2*) gene are associated with cardiac troponin-T levels and directly correlated to heart failure (Yu et al., 2013). These challenges in merging DNA, RNA and epigenome data highlight the importance of comprehensive studies publicly available.

Our findings provide evidence of the involvement of monogenic genes in polygenic hypertrophic cardiomyopathy but might only account for part of the genes responsible for this disease. The prevalence of rare variants and unidentified genes responsible for HCM are yet to be elucidated. Furthermore, the interaction mechanisms used by these genes deserve more attention and might aid professionals in determining diagnostics and prognosis of the disease.

3.7 Funding sources

This work was supported by grants from the National Health & Medical Research Council of Australia (project grant APP1034371, APP509252), the National Heart Foundation (project grant G10M5155, GM6368), and the Federation University Australia "Self-sustaining Regions Research and Innovation Initiative," an Australian Government Collaborative Research Network (CRN). F.Z.M is supported by a National Heart Foundation Future Leader and Baker co-shared Fellowships. Prestes is supported by a Robert HT Smith Fellowship from the Federation University Australia.

3.8 Disclosures

None.

3.9 References

Aravamudan, B., Volonte, D., Ramani, R., Gursoy, E., Lisanti, M. P., London, B., & Galbiati, F. (2003). Transgenic overexpression of caveolin-3 in the heart induces a cardiomyopathic phenotype. Hum Mol Genet, 12(21), 2777-2788.

Barrett, Wilhite, S. E., Ledoux, P., Evangelista, C., Kim, I. F., Tomashevsky, M., . . . Soboleva, A. (2013). NCBI GEO: archive for functional genomics data sets--update. Nucleic Acids Res, 41(Database issue), D991-995.

Betel, D., Koppal, A., Agius, P., Sander, C., & Leslie, C. (2010). Comprehensive modeling of microRNA targets predicts functional non-conserved and non-canonical sites. Genome Biol, 11(8), R90.

Cingolani, P., Platts, A., Wang le, L., Coon, M., Nguyen, T., Wang, L., . . . Ruden, D. M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly (Austin), 6(2), 80-92.

Cirino, A. L. (2008 [updated 2014]). Hypertrophic cardiomyopathy overview. Seattle: University of Washington, Seattle.

Ding, H., Liu, J., Xue, R., Zhao, P., Qin, Y., Zheng, F., & Sun, X. (2014). Transthyretin as a potential biomarker for the differential diagnosis between lung cancer and lung infection. Biomed Rep, 2(5), 765-769.

Dweep, H., & Gretz, N. (2015). miRWalk2.0: a comprehensive atlas of microRNA-target interactions. Nat Methods, 12(8), 697.

Dweep, H., Sticht, C., Pandey, P., & Gretz, N. (2011). miRWalk-database: prediction of possible miRNA binding sites by "walking" the genes of three genomes. J Biomed Inform, 44(5), 839-847.

Edgar, R., Domrachev, M., & Lash, A. E. (2002). Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res, 30(1), 207-210.

Friedrich, F. W., & Carrier, L. (2012). Genetics of hypertrophic and dilated cardiomyopathy. Curr Pharm Biotechnol, 13(13), 2467-2476.

Garcia, D. M., Baek, D., Shin, C., Bell, G. W., Grimson, A., & Bartel, D. P. (2011). Weak seed-pairing stability and high target-site abundance decrease the proficiency of lsy-6 and other microRNAs. Nat Struct Mol Biol, 18(10), 1139-1146.

Gustilo, E. M., Vendeix, F. A., & Agris, P. F. (2008). tRNA's modifications bring order to gene expression. Curr Opin Microbiol, 11(2), 134-140.

Harrap, S. B., Danes, V. R., Ellis, J. A., Griffiths, C. D., Jones, E. F., & Delbridge, L. M. (2002). The hypertrophic heart rat: a new normotensive model of genetic cardiac and cardiomyocyte hypertrophy. Physiol Genomics, 9(1), 43-48.

Helwak, A., Kudla, G., Dudnakova, T., & Tollervey, D. (2013). Mapping the human miRNA interactome by CLASH reveals frequent noncanonical binding. Cell, 153(3), 654-665.

Hershberger, R. E., & Morales, A. (2007 [updated 2015]). Dilated cardiomyopathy overview. Seattle: University of Washington, Seattle.

Hwang, D. M., Dempsey, A. A., Lee, C. Y., & Liew, C. C. (2000). Identification of differentially expressed genes in cardiac hypertrophy by analysis of expressed sequence tags. Genomics, 66(1), 1-14.

Ingles, J., Doolan, A., Chiu, C., Seidman, J., Seidman, C., & Semsarian, C. (2005). Compound and double mutations in patients with hypertrophic cardiomyopathy: implications for genetic testing and counselling. J Med Genet, 42(10), e59.

Jubb, H. C., Pandurangan, A. P., Turner, M. A., Ochoa-Montaño, B., Blundell, T. L., & Ascher, D. B. (2016). Mutations at protein-protein interfaces: Small changes over big surfaces have large impacts on human health. Progress in Biophysics and Molecular Biology.

Kaller, M., Liffers, S. T., Oeljeklaus, S., Kuhlmann, K., Roh, S., Hoffmann, R., . . . Hermeking, H. (2011). Genome-wide characterization of miR-34a induced changes in protein and mRNA expression by a combined pulsed SILAC and microarray analysis. Mol Cell Proteomics, 10(8), M111 010462.

Kelly, M., & Semsarian, C. (2009). Multiple mutations in genetic cardiovascular disease: a marker of disease severity? Circ Cardiovasc Genet, 2(2), 182-190.

Kimura, A. (2016). Molecular genetics and pathogenesis of cardiomyopathy. J Hum Genet, 61(1), 41-50.

Li, M. J., Liu, Z., Wang, P., Wong, M. P., Nelson, M. R., Kocher, J. P., . . . Wang, J. (2016). GWASdb v2: an update database for human genetic variants identified by genome-wide association studies. Nucleic Acids Res, 44(D1), D869-876.

Lim, Vos, T., Flaxman, A. D., Danaei, G., Shibuya, K., Adair-Rohani, H., . . . Memish, Z. A. (2012). A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet, 380(9859), 2224-2260.

Lipchina, I., Elkabetz, Y., Hafner, M., Sheridan, R., Mihailovic, A., Tuschl, T., . . . Betel, D. (2011). Genome-wide identification of microRNA targets in human ES cells reveals a role for miR-302 in modulating BMP response. Genes Dev, 25(20), 2173-2186.

Maron, B. J. (2002). Hypertrophic cardiomyopathy: A systematic review. JAMA, 287(10), 1308-1320.

Maron, B. J., & Maron, M. S. (2013). Hypertrophic cardiomyopathy. Lancet, 381(9862), 242-255.

Maron, B. J., Maron, M. S., & Semsarian, C. (2012). Double or compound sarcomere mutations in hypertrophic cardiomyopathy: a potential link to sudden death in the absence of conventional risk factors. Heart Rhythm, 9(1), 57-63.

Morita, H., Rehm, H. L., Menesses, A., McDonough, B., Roberts, A. E., Kucherlapati, R., . . . Seidman, C. E. (2008). Shared genetic causes of cardiac hypertrophy in children and adults. N Engl J Med, 358(18), 1899-1908.

Porrello, E. R., Bell, J. R., Schertzer, J. D., Curl, C. L., McMullen, J. R., Mellor, K. M., . . . Delbridge, L. M. (2009). Heritable pathologic cardiac hypertrophy in adulthood is preceded by neonatal cardiac growth restriction. Am J Physiol Regul Integr Comp Physiol, 296(3), R672-680.

Prestes, P. R., Marques, F. Z., Lopez-Campos, G., Booth, S. A., McGlynn, M., Lewandowski, P., . . . Charchar, F. J. (2016). Tripartite motif-containing 55 identified as functional candidate for spontaneous cardiac hypertrophy in the rat locus cardiac mass 22. J Hypertens, 34(5), 950-958.

Richard, P., Charron, P., Carrier, L., Ledeuil, C., Cheav, T., Pichereau, C., . . . Project, E. H. F. (2003). Hypertrophic cardiomyopathy: distribution of disease genes, spectrum of mutations, and implications for a molecular diagnosis strategy. Circulation, 107(17), 2227-2232.

Sauna, Z. E., & Kimchi-Sarfaty, C. (2011). Understanding the contribution of synonymous mutations to human disease. Nat Rev Genet, 12(10), 683-691.

Ton, V. K., Mukherjee, M., & Judge, D. P. (2014). Transthyretin cardiac amyloidosis: pathogenesis, treatments, and emerging role in heart failure with preserved ejection fraction. Clin Med Insights Cardiol, 8(Suppl 1), 39-44.

Towbin, J. A. (2014). Inherited cardiomyopathies. Circ J, 78(10), 2347-2356.

Van der Auwera, G. A., Carneiro, M. O., Hartl, C., Poplin, R., Del Angel, G., Levy-Moonshine, A., . . . DePristo, M. A. (2013). From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. Curr Protoc Bioinformatics, 43(1110):11.10.1-11.10.33.

Wagner, R. A., Tabibiazar, R., Powers, J., Bernstein, D., & Quertermous, T. (2004). Genome-wide expression profiling of a cardiac pressure overload model identifies major metabolic and signaling pathway responses. J Mol Cell Cardiol, 37(6), 1159-1170.

World Health Organization. (2016, September 2016). Cardiovascular diseases (CVDs) - fact sheet. Retrieved from http://www.who.int/mediacentre/factsheets/fs317/en/

Xiong, Y., Wei, Y., Gu, Y., Zhang, S., Lyu, J., Zhang, B., . . . Zhang, Y. (2017). DiseaseMeth version 2.0: a major expansion and update of the human disease methylation database. Nucleic Acids Res, 45(D1), D888-D895.

Xu, Dewey, S., Nguyen, S., & Gomes, A. V. (2010). Malignant and benign mutations in familial cardiomyopathies: insights into mutations linked to complex cardiovascular phenotypes. J Mol Cell Cardiol, 48(5), 899-909.

Yu, B., Barbalic, M., Brautbar, A., Nambi, V., Hoogeveen, R. C., Tang, W., . . . Consortium, C. A. (2013). Association of genome-wide variation with highly sensitive cardiac troponin-T levels in European Americans and Blacks: a meta-analysis from atherosclerosis risk in communities and cardiovascular health studies. Circ Cardiovasc Genet, 6(1), 82-88.

3.10 Figures

Figure 3.1. The hypertrophic heart rat (HHR) has an enlarged heart when compared to its genetic control, the normal heart rat (NHR). Cardiac weight index (CWI, mg/g) of HHR is represented as a percentage difference to NHR (normalized to 100%) at 2-day old, 4-, 13-, 33- and 50-week old (n=7-12 per group per age). *P<0.05; **P<0.01; ***P<0.001.

Figure 3.2. Heart mRNA expression of genes differentially expressed in at least one age group. There are four and 21 genes differentially expressed in A) neonatal and B) 4 weeks old hypertrophic heart rats (HHR), respectively. HHR fold change relative to normal heart rat (NHR) is shown. Genes not differentially expressed are shown in open boxes. *P<0.05; **P<0.01; ***P<0.001.

Figure 3.3. Heart mRNA expression of genes differentially expressed in at least one age group. There are 14 and 13 genes differentially expressed in A) 13 weeks old and B) 33 weeks old hypertrophic heart rats (HHR), respectively. HHR fold change relative to normal heart rat (NHR) is shown. Genes not differentially expressed are shown in open boxes. *P<0.05; **P<0.01; ***P<0.001.

Figure 3.4. Heart mRNA expression of genes differentially expressed in at least one age group. There are ten and nine genes differentially expressed in A) 50 weeks old hypertrophic heart rats (HHR) and B) human idiopathic dilated cardiomyopathy (DCM), respectively. HHR fold change relative to normal heart rat (NHR) is shown. DCM patients fold change relative to healthy is shown in humans. Genes not differentially expressed are shown in open boxes. *P<0.05; **P<0.01; ***P<0.001.

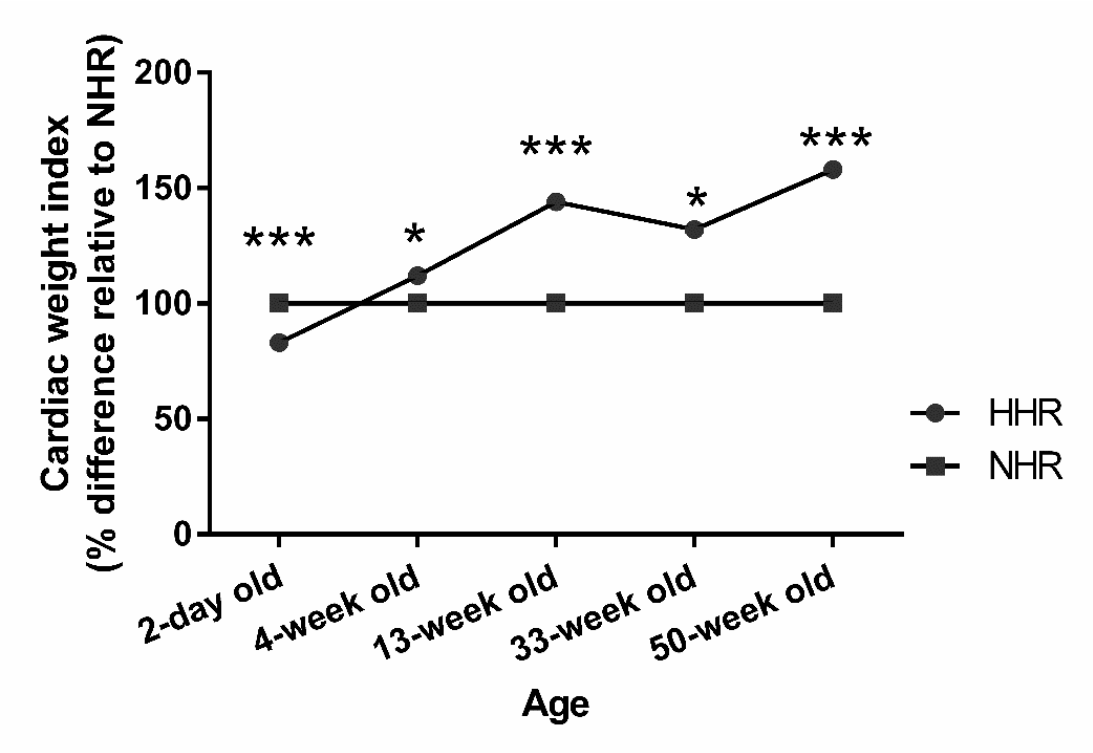
Cardiac weight index
(% difference relative to NHR)
200
150
100
50
0
***
*
***
*
***
2-day old
4-week old
13-week old
33-week old
50-week old
Age
HHR
NHR

Figure 3.1

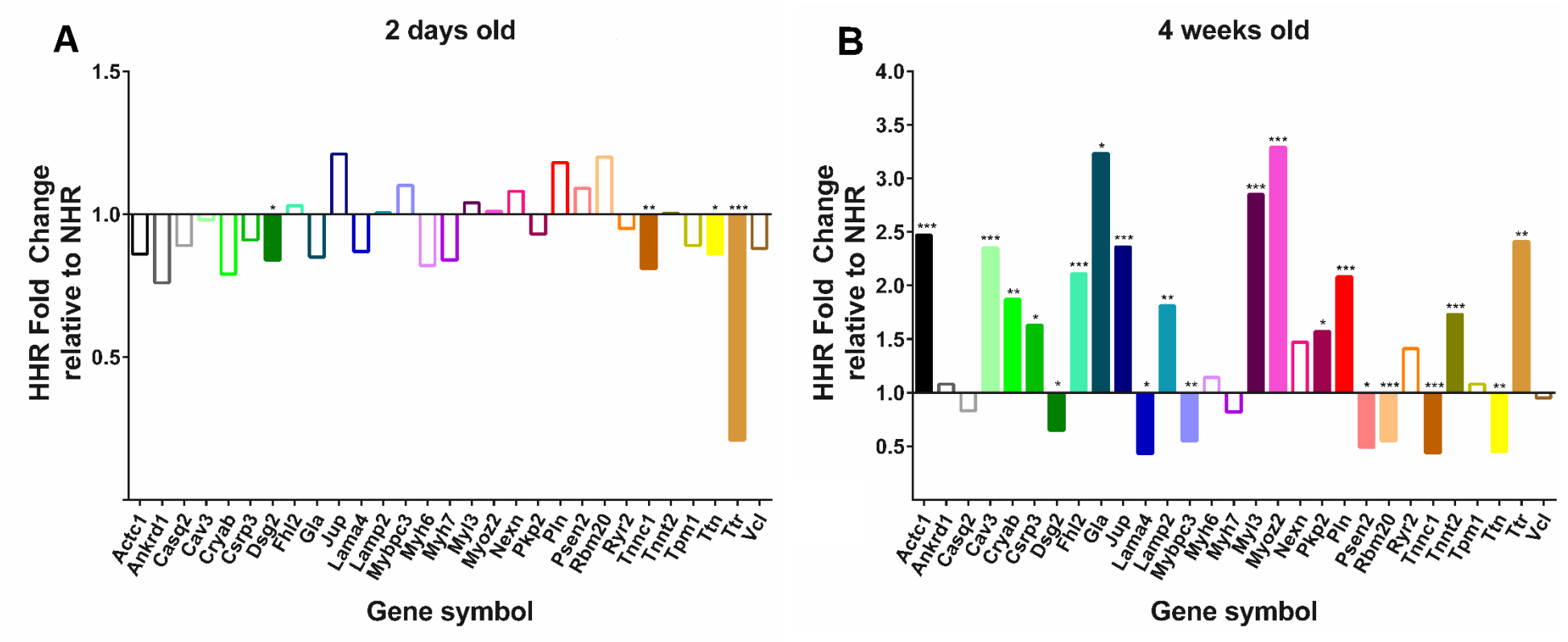
A
2 days old
HHR Fold Change relative to NHR
1.5
1.0
0.5
Actc1 Ankrd1 Casq2 Cav3 Cryab Csrp3 Dsg2 Fhl2 Gla Jup Lama4 Lamp2 Mybpc3 Myh6 Myh7 Myl3 Myoz2 Nexn Pkp2 Pln Psen2 Rbm20 Ryr2 Tnnc1 Tnnt2 Tpm1 Ttn Ttr Vcl
Gene symbol
B
4 weeks old
HHR Fold Change relative to NHR
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
Actc1 Ankrd1 Casq2 Cav3 Cryab Csrp3 Dsg2 Fhl2 Gla Jup Lama4 Lamp2 Mybpc3 Myh6 Myh7 Myl3 Myoz2 Nexn Pkp2 Pln Psen2 Rbm20 Ryr2 Tnnc1 Tnnt2 Tpm1 Ttn Ttr Vcl
Gene symbol

Figure 3.2

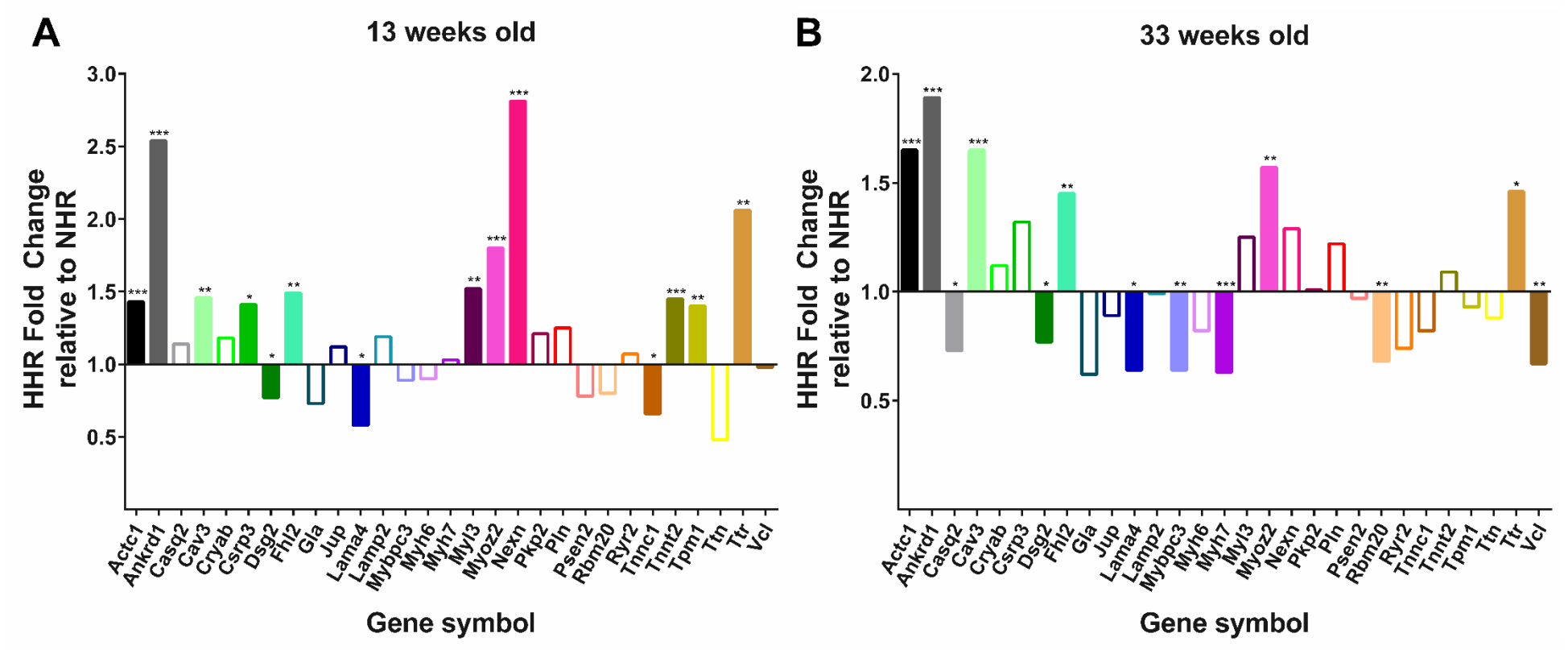
A
13 weeks old
HHR Fold Change relative to NHR
3.0
2.5
2.0
1.5
1.0
0.5
Actc1 Ankrd1 Casq2 Cav3 Cryab Csrp3 Dsg2 Fhl2 Gla Jup Lama4 Lamp2 Mybpc3 Myh6 Myh7 Myl3 Myoz2 Nexn Pkp2 Pln Psen2 Rbm20 Ryr2 Tnnc1 Tnnt2 Tpm1 Ttn Ttr Vcl
Gene symbol
B
33 weeks old
HHR Fold Change relative to NHR
2.0
1.5
1.0
0.5
Actc1 Ankrd1 Casq2 Cav3 Cryab Csrp3 Dsg2 Fhl2 Gla Jup Lama4 Lamp2 Mybpc3 Myh6 Myh7 Myl3 Myoz2 Nexn Pkp2 Pln Psen2 Rbm20 Ryr2 Tnnc1 Tnnt2 Tpm1 Ttn Ttr Vcl
Gene symbol

Figure 3.3

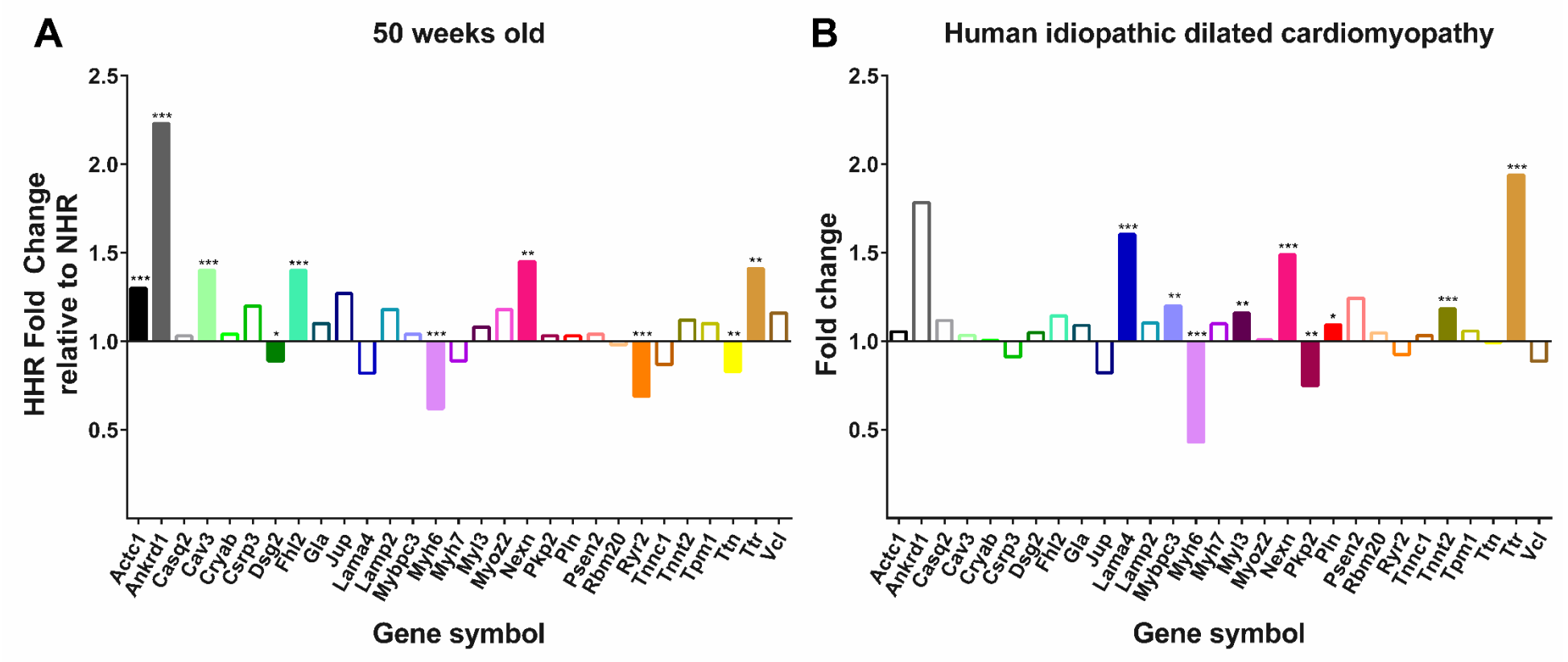
A
50 weeks old
HHR Fold Change relative to NHR
2.5
2.0
1.5
1.0
0.5
Actc1 Ankrd1 Casq2 Cav3 Cryab Csrp3 Dsg2 Fhl2 Gla Jup Lama4 Lamp2 Mybpc3 Myh6 Myh7 Myl3 Myoz2 Nexn Pkp2 Pln Psen2 Rbm20 Ryr2 Tnnc1 Tnnt2 Tpm1 Ttn Ttr Vcl
Gene symbol
B
Human idiopathic dilated cardiomyopathy
Fold change
2.5
2.0
1.5
1.0
0.5
Actc1 Ankrd1 Casq2 Cav3 Cryab Csrp3 Dsg2 Fhl2 Gla Jup Lama4 Lamp2 Mybpc3 Myh6 Myh7 Myl3 Myoz2 Nexn Pkp2 Pln Psen2 Rbm20 Ryr2 Tnnc1 Tnnt2 Tpm1 Ttn Ttr Vcl
Gene symbol

Figure 3.4

3.11 Tables

Table 3.1. Variants found in each gene investigated unique to each strain classified according to type.

Gene ID[a]	Gene name	Frequency in patients		Rat Genome			HHR[c]		NHR[d]	
		Hypertrophic	Dilated	Chr[b]	Location	Strand	SNP[e]	InDel[f]	SNP[e]	InDel[f]
Abcc9	ATP-binding cassette, subfamily C member 9	n/d[g]	0.6%	4	241,019,980 - 241,139,051	-	4	10	2	9
Actc1	cardiac actin alpha 1	<1%	rare	3	112,080,853 - 112,086,389	-	3	2	1	1
Actn2	actinin alpha 2	rare	0.9%	17	68,050,946 - 68,143,522	-	61	42	58	39
Ankrd1	ankyrin repeat domain 1	rare	1.9%	1	262,038,137 - 262,046,691	+	none	5	2	2
Casq2	calsequestrin 2	rare	n/d	2	223,945,611 - 224,001,893	+	11	8	5	12
Cav3	caveolin 3	0.6%	n/d	4	207,683,202 - 207,699,176	+	none	2	none	1
Cryab	crystallin, alpha B	n/d[g]	0.7%	8	53,776,100 - 53,779,780	+	3	6	2	8
Csrp3	cysteine and glycine-rich protein 3	rare	0.3%	1	105,206,459 - 105,225,635	-	1	1	none	none
Ctf1	cardiotrophin 1	n/d[g]	n/d[g]	1	206,186,511 - 206,191,930	+	2	3	none	3

Des	desmin	n/d[g]	0.3-2.1%	9	82,325,835 - 82,333,549	+	34	7	none	1
Dsc2	desmocollin 2	n/d[g]	n/d[g]	18	11,625,847 - 11,658,036	-	9	4	2	3
Dsg2	desmoglein 2	n/d[g]	2.3%	18	15,353,348 - 15,411,825	-	5	3	24	4
Emd	emerin	n/d[g]	n/d[g]	1	152,192,993 - 152,196,004	-	none	none	none	none
Fhl2	four and a half LIM domains 2	n/d[g]	2.1%	9	49,591,185 - 49,664,022	-	1	5	none	3
Gla	galactosidase, alpha	1-2% in men	n/d[g]	X	105,295,029 - 105,306,686	-	none	1	none	none
Jup	junction plakoglobin	n/d[g]	n/d	10	88,073,764 - 88,100,700	-	1	10	1	6
Lama4	laminin, alpha 4	n/d[g]	1.1%	20	45,786,892 - 45,926,468	+	37	24	101	17
Lamp2	lysosomal-associated membrane protein 2	rare	n/d[g]	X	124,809,053 - 124,852,509	-	2	4	1	3
Lmna	lamin A/C	n/d[g]	6-7.5%	2	207,245,237 - 207,265,928	-	24	11	3	11
Mybpc3	myosin binding protein C	20-42%	0.2-4%	3	86,649,264 - 86,667,484	+	none	3	none	3
Myh6	myosin heavy chain 6	rare	4.3%	15	37,492,599 - 37,516,786	-	1	4	none	3

Myh7	myosin heavy chain 7	20-40%	4.2-6.3%	15	37,512,803 - 37,544,317	-	2	none	none	1
Myl2	myosin light chain 2	<5%	n/d[g]	12	41,831,137 - 41,835,806	-	6	13	5	7
Myl3	myosin light chain 3	1-2%	n/d[g]	8	118,370,030 - 118,376,218	+	none	4	3	4
Mylk2	myosin light chain kinase 2	rare	n/d[g]	3	154,789,177 - 154,800,843	+	none	1	none	none
Myoz2	myozenin 2	rare	n/d[g]	2	246,542,267 - 246,569,008	-	none	none	2	none
Nexn	nexilin	n/d[g]	1%	2	276,129,684 - 276,161,434	-	2	5	2	2
Pkp2	plakophilin 2	n/d[g]	n/d[g]	11	91,966,047 - 92,031,277	-	2	4	2	4
Pln	phospholamban	rare	rare	20	36,390,879 - 36,400,626	+	1	none	none	none
Prkag2	protein kinase, AMP-activated, non-catalytic subunit gamma 2	<1%	n/d[g]	4	6,577,007 - 6,816,813	+	2	4	none	5
Psen2	presenilin 2	n/d[g]	1%	13	103,521,460 - 103,547,174	-	1	0	2	2
Rbm20	RNA binding motif protein 20	n/d[g]	1.9%	1	281,783,646 - 282,003,053	+	12	21	3	12

Ryr2	ryanodine receptor 2	rare	n/d[g]	17	67,285,205 - 67,704,766	-	518	218	122	106
Sgcd	sarcoglycan delta	n/d[g]	rare	10	31,878,904 - 32,285,036	-	26	34	15	24
Taz	tafazzin	n/d[g]	n/d[g]	1	152,161,153 - 152,169,569	-	1	none	none	none
Tmem43	transmembrane protein 43	n/d[g]	n/d[g]	4	187,412,090 - 187,427,232	-	3	2	2	1
Tnnc1	troponin C type 1	rare	0.4%	16	7,220,777 - 7,223,730	+	0	0	0	0
Tnnt2	troponin T type 2	3-10%	2.9%	13	57,711,369 - 57,729,182	+	48	13	3	4
Tpm1	tropomyosin 1	<5%	0.6-1.9%	8	77,147,580 - 77,174,392	-	none	none	none	2
Ttn	titin	rare	14.1-20%	3	70,138,896 - 70,408,647	-	24	12	6	10
Ttr	transthyretin	1-10%	n/d[g]	18	15,307,563 - 15,316,780	-	none	1	8	1
Vcl	vinculin	rare	rare	15	3,433,521 - 3,522,441	-	4	4	6	2
Total							851	491	383	316

Legend: [a]ID, identification; [b]Chr, chromosome; [c]HHR, hypertrophic heart rat; [d]NHR, normal heart rat; [e]SNP, single nucleotide polymorphism; [f]InDel, insertion/deletion; [g]n/d, not described.

Table 3.2. Predicted number of microRNA binding sites in or around the genes of interest.

	Region			
Gene	Promoter	5'UTR*	CDS†	3'UTR*
Abcc9	1255	3618	19131	8848
Actc1	1394	290	1126	1721
Actn2	1195	1296	6071	3848
Ankrd1	1265	407	1016	625
Casq2	1332	835	2678	2594
Cav3	1443	240	1103	1411
Cryab	1356	487	871	116
Csrp3	1260	332	772	620
Ctf1	1371	106	1231	1607
Des	1379	137	1194	907
Dsc2	1161	1345	4605	2698
Dsg2	1289	222	1982	1224
Emd	1249	209	946	458
Fhl2	1319	2240	7695	3197
Gla	1271	89	1354	6
Jup	1280	2251	25961	14746
Lama4	1233	2439	16253	3119
Lamp2	1202	732	3485	4589
Lmna	1279	976	5076	1977
Mybpc3	1362	137	2025	592
Myh6	1314	118	5753	234
Myh7	1341	461	5832	427

Myl2	1330	205	646	472
Myl3	1197	293	819	482
Mylk2	1403	256	1552	858
Myoz2	1235	344	844	933
Nexn	1242	1939	7475	7584
Pkp2	1345	234	3398	2367
Pln	1024	254	215	1062
Prkag2	1293	3224	12154	12185
Psen2	1212	1818	4518	2407
Rbm20	1262	41	2014	1835
Ryr2	1270	565	5045	1941
Sgcd	1216	1623	3963	9768
Taz	1191	1139	3286	3161
Tmem43	1362	286	1218	1263
Tnnc1	1370	41	567	347
Tnnt2	1299	2222	5604	2586
Tpm1	1302	4249	15719	14114
Ttn	1117	648	28263	4702
Ttr	1287	224	750	435
Vcl	1295	396	7084	6209
Total	53802	38968	221294	130275

Legend: UTR*, untranslated region; CDS†, coding DNA sequence.

Table 3.3. Differentially expressed microRNAs experimentally validated to interact with genes under investigation.

miRNA*	MIMATid	Count	Genes	p-value	FDR†	Mean Ratio (HHR‡/NHR§)
hsa-miR-34a-5p	MIMAT0000255	3	*PKP2, TPM1, VCL*	0.000117868	0.00977691	1.83
hsa-miR-17-5p	MIMAT0000070	2	*RBM20, RYR2*	0.0000029	0.00076469	1.10

Legend: miRNA*, microRNA; FDR†, false discovery rate; HHR‡, hypertrophic heart rat; NHR§, normal heart rat.

Table 3.4. Single nucleotide polymorphisms found in genes investigated associated with traits related to cardiovascular disease in human genome wide association studies.

Trait	SNP count	Genes
Aortic root size	1	*VCL*
Atrial fibrillation	2	*RYR2, TTN*
Blood pressure	5	*RBM20, SGCD, RYR2, PRKAG2*
Blood pressure (response to angiotensin II receptor blocker)	3	*ABCC9, RYR2*
Blood pressure, CVD* RF† and other traits (body mass index, waist to hip ratio, renin activity and aldosterone concentration in plasma, BNP‡ levels in plasma, alcohol consumption)	9	*TTN, PRKAG2, MYH6, LAMA4*
Body mass index	4	*RYR2, PKP2, ACTN2*
Cardiac troponin-T levels	2	*TNNT2*
Cardiovascular disease	1	*RYR2*
Chronic kidney disease	1	*PRKAG2*
Coronary artery calcification	8	*PRKAG, SGCD, RYR2, MYBPC3, MYH6, RYR2*
Coronary artery disease	1	*VCL*

Coronary heart disease	2	*VCL, TNNT2*
ECG§ dimensions, brachial artery endothelial function, treadmill exercise responses	3	*PRKAG2, RYR2*
Electrocardiographic conduction measures	1	*RYR2*
Electrocardiographic traits	1	*MYH6*
Glomerular filtration rate	1	*PRKAG2*
Health and aging, CVD and cancer age of onset	1	*TTN*
Heart failure	1	*SGCD*
Heart rate	2	*CRYAB, MYH6*
Height	1	*DSC2*
Hypertension	2	*SGCD*
Hypertension (early onset hypertension)	2	*ACTN2*
Multiple complex diseases	26	*FHL2, RYR2, LMNA, CSRP3, PRKAG2, SGCD, RBM20, DSC2, ACTN2, LAMA4, MYLK2, MYOZ2*
Myocardial infarction	2	*CASQ2*
Obesity-related traits	7	*ABCC9, RYR2, PKP2*

Red blood cell traits	13	*PRKAG2, CTF1*
Resting heart rate	1	*MYH6*
Sudden cardiac arrest	1	*RBM20*
Triglycerides	5	*MYBPC3, VCL, RYR2, TNNT2, CSRP3*
Ventricular conduction	1	*CASQ2*
Waist circumference	1	*RYR2*

Legend: CVD*, cardiovascular disease; RF†, risk factors; BNP‡, brain natriuretic peptide; ECG§, electrocardiogram.

Chapter 5. Discussion

5.1 Overview

Although relatively common, primary LVH remains underdiagnosed, especially the polygenic, familial condition. Typically, previous studies have investigated genetic and epigenetic contributions to LVH separately (Burke et al., 2016; Farrell et al., 2018; Friedrich & Carrier, 2012; Khan et al., 2016; Marian & Braunwald, 2017; Sabater-Molina, Perez-Sanchez, Hernandez Del Rincon, & Gimeno, 2018). In this thesis, I present a comprehensive study incorporating DNA variants, RNA expression and epigenetic – miRNA and circRNA – data to elucidate contributions to LVH. To address this study, I used the HHR, a unique animal model of polygenic LVH, independent of confounding factors such as high blood pressure, and its control strain, the NHR (Harrap et al., 2002). Soon after birth, the HHR has fewer and smaller cardiomyocytes than the NHR, developing LVH around reproductive maturity between 10-12 weeks of age, leading to premature death around 55-60 weeks of age from cardiac failure (Porrello, Bell, et al., 2009). Starting life with fewer cardiomyocytes has far reaching consequences as the hypertrophy observed seems to be a compensatory mechanism, resulting from the stresses of body growth caused by the normal pressure and volume overloads of physiological changes (Bernardo, Weeks, Pretorius, & McMullen, 2010; de Jonge et al., 2011; Sharma et al., 2006). Furthermore, cardiomyocytes become enlarged to account for a decrease on contractile efficiency due to the overall LVH (Porrello, Bell, et al., 2009). Even though there are animal models of monogenic LVH, the use of the HHR as a model of polygenic LVH is more relevant to the investigation of common human LVH and crucial to understanding its primary origins as it develops this condition without pressure or volume overload, similar to the condition observed in humans with primary LVH (Harrap et al., 2002; Porrello, Bell, et al., 2009; Tsao et al., 2015).

During my PhD candidature, I firstly explored DNA variants using whole-genome sequencing while focusing on quantitative trait locus Cm22, previously identified as involved in cardiac mass and hypertrophy; secondly, I studied genes previously described in human monogenic forms of cardiomyopathies using a comprehensive approach to explore genetic and epigenetic links combining DNA sequence variants, mRNA and miRNA expression; and lastly, I extended my epigenetic investigation by measuring circRNA expression in LVH, that can modulate mRNA levels by targeting miRNA. I silenced circRNA expression of *HRCR* in human primary cardiomyocytes (HPC) to explore its novel miRNA targets and investigate if those changes are potentially related to LVH indicating its potential as a biomarker, diagnostic or prognostic tool.

5.2 Tripartite motif-containing 55 (*Trim55*), a novel candidate gene related to LVH

I characterised unique variants present in the HHR and NHR using whole-genome sequencing compared to the Brown Norway reference genomic sequence. I explored DNA variants between the two strains and combined these findings with microarray data to identify genes carrying DNA variants which could lead to the differential expression of its mRNA, focusing on quantitative trait locus Cm22 due to its relevance to cardiac mass determination and hypertrophy (Di Nicolantonio et al., 2006). *Trim55* was the only gene differentially expressed in Cm22, being significantly downregulated and carrying a nonsynonymous exonic single nucleotide polymorphism (SNP) in the HHR. This SNP results in a missense mutation, causing a codon and amino acid change which will change the protein structure, according to *in silico* modelling, resulting in the loss of a protein domain and impaired protein folding. Importantly, *Trim55* was downregulated throughout life in the HHR and was also

downregulated in human idiopathic dilated cardiomyopathy indicating its possible relevance as functional candidate gene in polygenic LVH.

Trim55 and its related gene tripartite motif-containing 63 (*Trim63*) encode proteins containing RING zinc fingers, involved in protein-protein interactions and associated with microtubules, myosin and titin during sarcomere assembly (Su et al., 2014). It has been proposed that these proteins regulate cardiac mass and function and the presence of rare variants is involved in LVH (Su et al., 2014). The authors also suggest that these rare variants increase the penetrance and severity of clinical manifestations of LVH, reinforcing the theory that variants in sarcomeric proteins have a greater impact in disease outcomes when compared to non-sarcomeric variants and highlighting the relevance of *Trim55* in sarcomere assembly and its protein-protein interactions (Friedrich & Carrier, 2012; Maron & Maron, 2013; Su et al., 2014). Furthermore, *Trim55* mutations can contribute to hypertrophic cardiomyopathy and heart failure and *Trim55* knockout mice have been described to present severe diabetic cardiomyopathy (He et al., 2015; Heliste et al., 2020).

These data support *Trim55* as a novel functional candidate gene involved in primary LVH. Variants present in this gene could assist diagnosis of patients and families that do not present variants in the typical sarcomeric genes currently investigated in routine diagnostic DNA testing.

5.3 Monogenic genes partially involved in polygenic LVH

Many genes have been described in human monogenic LVH (Marian & Braunwald, 2017; Maron et al., 2012b; Walsh et al., 2017) but their involvement in primary polygenic LVH remained unexplored. Based on the literature, I investigated 42 genes described in human monogenic hypertrophic and dilated cardiomyopathies (Friedrich & Carrier, 2012; Guo et al.,

2012; Hershberger & Morales, 2007 [updated 2015]; Marian & Braunwald, 2017; Maron et al., 2012b; McKenna & Judge, 2020; Reza et al., 2019; Sanoudou et al., 2005; Towbin, 2014; Walsh et al., 2017) at five age groups to understand the relationship between the expression of these genes and the developmental stages of LVH. Most of these genes are responsible for sarcomere assembly and cardiomyocyte contractibility, where the presence of mutations has a well-established link with LVH (Cirino, 2008 [updated 2014]; Richard et al., 2003).

This study investigated if the genes selected had a quantitative impact on cardiac development, structure and function, evaluating if compound mutations played an additive role influencing gene expression and impacting disease outcomes. This thesis shows that genes encoding cardiac filaments and sarcomeric proteins are mostly downregulated in neonatal HHR, when their hearts are smaller in comparison to their control strain, NHR. However, as LVH starts developing (around 4 weeks of age) and becoming established (after 13 weeks of age), most genes are upregulated. These changes suggest a link between LVH development and changes in gene expression (Wagner et al., 2004).

Amongst the wealth of data generated in this study, I would like to highlight the change in expression of the gene encoding transthyretin (*Ttr*), as it shifts from underexpressed in neonatal hearts to overexpressed in adults in HHR when compared to NHR suggesting its relevance in age-specific LVH. *Ttr* is mainly synthesised in the liver and a mutation in this gene causes the misfolding of the protein leading to transthyretin amyloidosis, a condition that mimics hypertrophic cardiomyopathy, causing heart failure (Morner, Hellman, Suhr, Kazzam, & Waldenstrom, 2005; Vermeer et al., 2017). Furthermore, an additional three genes are upregulated in adult HHR, encoding cardiac alpha actin 1 (*Actc1*), caveolin-3 (*Cav3*), and four and a half LIM domains 2 (*Fhl2*) and have been described in cardiomyopathies and arrhythmias (Despond & Dawson, 2018; Frank et al., 2019; Friedrich et al., 2014; Gao et al., 2020; Vaidyanathan, Reilly, & Eckhardt, 2018).

The lack of strong association between any specific gene and LVH, supports the polygenic nature of this animal model and highlighted the additive role of genes related to this condition. Furthermore, this study improved the understanding of the contribution of genes previously described in monogenic LVH in the polygenic context and provided insights into possible stage-specific changes related to its pathophysiology in HHR. Importantly, it has demonstrated the different aetiology between monogenic and polygenic LVH. Importantly, this study may provide valuable information for the development of targeted treatment for polygenic LVH, instead of the one size fits most currently available.

5.4 Heart related circular RNAis a potential epigenetic regulator in polygenic LVH

Epigenetic factors are heritable or environmental elements that influence gene expression without changing the DNA sequence, such as DNA methylation, histone modifications, noncoding RNAs, exercise or diet (Cavalli & Heard, 2019; Furrow, Christiansen, & Feldman, 2011). Importantly, epigenetics may explain the "missing heritability" observed in complex diseases, especially related to CVD. Thus far, the use of noncoding RNAs as contributors to epigenetic change has focused on miRNA expression and its role in post-transcriptional regulation (Zhao, 2018). Therefore, the involvement of miRNAs in disease is better understood and their potential use as biomarkers and therapeutic target has been promising (Navickas et al., 2016; Scolari, Faganello, Garbin, Piva, & Biolo, 2020; Zhang, Shen, Shi, & Yu, 2020). Other noncoding RNAs, including circRNAs, have been overlooked as epigenetic contributors until recently but are emerging as a new and promising type of regulatory molecules (Furrow et al., 2011; Qu et al., 2015).

CircRNAs have been identified as post-transcription modifiers, targeting miRNAs and acting as sponges (Fan et al., 2017; Qu et al., 2015). Recently, *Hrcr* has been the most

promising circRNA in LVH and is upregulated in mice with pathological LVH. It has proposed that *Hrcr* sponges miR-223 which then targets the *Arc* gene (Wang et al., 2016). During my candidature, I measured *Hrcr* levels and its proposed downstream targets in physiological LVH, using the HHR and silenced *HRCR* in HPC to investigate novel miRNA targets.

I found *Hrcr* upregulated in the HHR after LVH is established. Unfortunately, I did not find an association between levels of *Hrcr* and its proposed downstream targets which may be explained by the distinct aetiology between the polygenic model of LVH used in my studies and the pathologically induced cardiac hypertrophy used in the previous study (Wang et al., 2016). Interestingly, my investigation of novel targets for *HRCR* following silencing revealed four miRNAs differentially expressed (miR-1-3p, miR-330, miR-27a-5p, miR-299-5p). Recent studies describe an association between these miRNAs and several CVD, including LVH and heart failure (Cakmak et al., 2015; Condorelli et al., 2010; D'Alessandra et al., 2020; Li et al., 2016; Zhang et al., 2019). Furthermore, I also identified 359 possible mRNA targets and 206 enriched gene ontology terms in this dataset.

Importantly, this study highlights the importance of understanding the role of circRNAs in epigenetics and their potential use as biomarkers, diagnostic and/or prognostic tools.

5.5 Conclusion

This thesis explores the benefits of using genetic and epigenetic markers to understand biological impact in LVH. This combined approach could be key to explain the "missing heritability" observed in complex diseases and may represent a more efficient method to identify individuals at risk of developing LVH, determine diagnosis and prognosis of disease and develop targeted treatments.

5.6 References

Bernardo, B. C., Weeks, K. L., Pretorius, L., & McMullen, J. R. (2010). Molecular distinction between physiological and pathological cardiac hypertrophy: experimental findings and therapeutic strategies. Pharmacol Ther, 128(1), 191-227.

Burke, M. A., Cook, S. A., Seidman, J. G., & Seidman, C. E. (2016). Clinical and mechanistic insights into the genetics of cardiomyopathy. J Am Coll Cardiol, 68(25), 2871-2886.

Cakmak, H. A., Coskunpinar, E., Ikitimur, B., Barman, H. A., Karadag, B., Tiryakioglu, N. O., . . . Vural, V. A. (2015). The prognostic value of circulating microRNAs in heart failure: preliminary results from a genome-wide expression study. J Cardiovasc Med (Hagerstown), 16(6), 431-437.

Cavalli, G., & Heard, E. (2019). Advances in epigenetics link genetics to the environment and disease. Nature, 571(7766), 489-499.

Cirino, A. L. (2008 [updated 2014]). Hypertrophic cardiomyopathy overview. Seattle: University of Washington, Seattle.

Condorelli, G., Latronico, M. V., & Dorn, G. W., 2nd. (2010). microRNAs in heart disease: putative novel therapeutic targets? Eur Heart J, 31(6), 649-658.

D'Alessandra, Y., Chiesa, M., Carena, M. C., Beltrami, A. P., Rizzo, P., Buzzetti, M., . . . Colombo, G. I. (2020). Differential Role of Circulating microRNAs to track progression and pre-symptomatic stage of chronic heart failure: a pilot study. Biomedicines, 8(12).

de Jonge, L. L., van Osch-Gevers, L., Willemsen, S. P., Steegers, E. A., Hofman, A., Helbing, W. A., & Jaddoe, V. W. (2011). Growth, obesity, and cardiac structures in early childhood: the Generation R Study. Hypertension, 57(5), 934-940.

Despond, E. A., & Dawson, J. F. (2018). Classifying cardiac actin mutations associated with hypertrophic cardiomyopathy. Front Physiol, 9, 405.

Di Nicolantonio, R., Kostka, V., Kwitek, A., Jacob, H., Thomas, W. G., & Harrap, S. B. (2006). Fine mapping of Lvm1: a quantitative trait locus controlling heart size independently of blood pressure. Pulm Pharmacol Ther, 19(1), 70-73.

Fan, X., Weng, X., Zhao, Y., Chen, W., Gan, T., & Xu, D. (2017). Circular RNAs in cardiovascular disease: an overview. Biomed Res Int, 2017, 5135781.

Farrell, E., Armstrong, A. E., Grimes, A. C., Naya, F. J., de Lange, W. J., & Ralphe, J. C. (2018). Transcriptome analysis of cardiac hypertrophic growth in MYBPC3-null mice suggests early responders in hypertrophic hemodeling. Front Physiol, 9, 1442.

Frank, D., Yusuf Rangrez, A., Friedrich, C., Dittmann, S., Stallmeyer, B., Yadav, P., . . . Schulze-Bahr, E. (2019). Cardiac alpha-actin (ACTC1) gene mutation causes atrial-septal defects associated with late-onset dilated cardiomyopathy. Circ Genom Precis Med, 12(8), e002491.

Friedrich, F. W., & Carrier, L. (2012). Genetics of hypertrophic and dilated cardiomyopathy. Curr Pharm Biotechnol, 13(13), 2467-2476.

Friedrich, F. W., Reischmann, S., Schwalm, A., Unger, A., Ramanujam, D., Munch, J., . . . Carrier, L. (2014). FHL2 expression and variants in hypertrophic cardiomyopathy. Basic Res Cardiol, 109(6), 451.

Furrow, R. E., Christiansen, F. B., & Feldman, M. W. (2011). Environment-sensitive epigenetics and the heritability of complex diseases. Genetics, 189(4), 1377-1387.

Gao, J., Collyer, J., Wang, M., Sun, F., & Xu, F. (2020). Genetic Dissection of Hypertrophic Cardiomyopathy with Myocardial RNA-Seq. Int J Mol Sci, 21(9).

Guo, Schafer, S., Greaser, M. L., Radke, M. H., Liss, M., Govindarajan, T., . . . Gotthardt, M. (2012). RBM20, a gene for hereditary cardiomyopathy, regulates titin splicing. Nat Med, 18(5), 766-773.

Harrap, S. B., Danes, V. R., Ellis, J. A., Griffiths, C. D., Jones, E. F., & Delbridge, L. M. (2002). The hypertrophic heart rat: a new normotensive model of genetic cardiac and cardiomyocyte hypertrophy. Physiol Genomics, 9(1), 43-48.

He, J., Quintana, M. T., Sullivan, J., T, L. P., T, J. G., Schisler, J. C., . . . Willis, M. S. (2015). MuRF2 regulates PPARgamma1 activity to protect against diabetic cardiomyopathy and enhance weight gain induced by a high fat diet. Cardiovasc Diabetol, 14, 97.

Heliste, J., Chheda, H., Paatero, I., Salminen, T. A., Akimov, Y., Paavola, J., . . . Aittokallio, T. (2020). Genetic and functional implications of an exonic TRIM55 variant in heart failure. J Mol Cell Cardiol, 138, 222-233.

Hershberger, R. E., & Morales, A. (2007 [updated 2015]). Dilated cardiomyopathy overview. Seattle: University of Washington, Seattle.

Khan, M. A., Reckman, Y. J., Aufiero, S., van den Hoogenhof, M. M., van der Made, I., Beqqali, A., . . . Pinto, Y. M. (2016). RBM20 Regulates Circular RNA Production From the Titin Gene. Circ Res, 119(9), 996-1003.

Li, Fan, J., Yin, Z., Wang, F., Chen, C., & Wang, D. W. (2016). Identification of cardiac-related circulating microRNA profile in human chronic heart failure. Oncotarget, 7(1), 33-45.

Marian, A. J., & Braunwald, E. (2017). Hypertrophic Cardiomyopathy: Genetics, Pathogenesis, Clinical Manifestations, Diagnosis, and Therapy. Circ Res, 121(7), 749-770.

Maron, B. J., & Maron, M. S. (2013). Hypertrophic cardiomyopathy. Lancet, 381(9862), 242-255.

Maron, B. J., Maron, M. S., & Semsarian, C. (2012). Genetics of hypertrophic cardiomyopathy after 20 years: clinical perspectives. J Am Coll Cardiol, 60(8), 705-715.

McKenna, & Judge, D. (2020). Epidemiology of the inherited cardiomyopathies. Nat Rev Cardiol.

Morner, S., Hellman, U., Suhr, O. B., Kazzam, E., & Waldenstrom, A. (2005). Amyloid heart disease mimicking hypertrophic cardiomyopathy. J Intern Med, 258(3), 225-230.

Navickas, R., Gal, D., Laucevicius, A., Taparauskaite, A., Zdanyte, M., & Holvoet, P. (2016). Identifying circulating microRNAs as biomarkers of cardiovascular disease: a systematic review. Cardiovasc Res, 111(4), 322-337.

Porrello, E. R., Bell, J. R., Schertzer, J. D., Curl, C. L., McMullen, J. R., Mellor, K. M., . . . Delbridge, L. M. (2009). Heritable pathologic cardiac hypertrophy in adulthood is preceded by neonatal cardiac growth restriction. Am J Physiol Regul Integr Comp Physiol, 296(3), R672-680.

Qu, S., Yang, X., Li, X., Wang, J., Gao, Y., Shang, R., . . . Li, H. (2015). Circular RNA: A new star of noncoding RNAs. Cancer Lett, 365(2), 141-148.

Reza, N., Musunuru, K., & Owens, A. T. (2019). From Hypertrophy to Heart Failure: What Is New in Genetic Cardiomyopathies. Curr Heart Fail Rep, 16(5), 157-167.

Richard, P., Charron, P., Carrier, L., Ledeuil, C., Cheav, T., Pichereau, C., . . . Project, E. H. F. (2003). Hypertrophic cardiomyopathy: distribution of disease genes, spectrum of mutations, and implications for a molecular diagnosis strategy. Circulation, 107(17), 2227-2232.

Sabater-Molina, M., Perez-Sanchez, I., Hernandez Del Rincon, J. P., & Gimeno, J. R. (2018). Genetics of hypertrophic cardiomyopathy: A review of current state. Clin Genet, 93(1), 3-14.

Sanoudou, D., Vafiadaki, E., Arvanitis, D. A., Kranias, E., & Kontrogianni-Konstantopoulos, A. (2005). Array lessons from the heart: focus on the genome and transcriptome of cardiomyopathies. Physiol Genomics, 21(2), 131-143.

Scolari, F. L., Faganello, L. S., Garbin, H. I., Piva, E. M. B., & Biolo, A. (2020). A systematic review of microRNAs in patients with hypertrophic cardiomyopathy. Int J Cardiol. 327:146-154

Sharma, P., Middelberg, R. P., Andrew, T., Johnson, M. R., Christley, H., & Brown, M. J. (2006). Heritability of left ventricular mass in a large cohort of twins. J Hypertens, 24(2), 321-324.

Su, M., Wang, J., Kang, L., Wang, Y., Zou, Y., Feng, X., . . . Song, L. (2014). Rare variants in genes encoding MuRF1 and MuRF2 are modifiers of hypertrophic cardiomyopathy. Int J Mol Sci, 15(6), 9302-9313.

Towbin, J. A. (2014). Inherited cardiomyopathies. Circ J, 78(10), 2347-2356.

Tsao, C. W., Gona, P. N., Salton, C. J., Chuang, M. L., Levy, D., Manning, W. J., & O'Donnell, C. J. (2015). Left Ventricular Structure and Risk of Cardiovascular Events: A Framingham Heart Study Cardiac Magnetic Resonance Study. J Am Heart Assoc, 4(9), e002188.

Vaidyanathan, R., Reilly, L., & Eckhardt, L. L. (2018). Caveolin-3 Microdomain: Arrhythmia Implications for Potassium Inward Rectifier and Cardiac Sodium Channel. Front Physiol, 9, 1548.

Vermeer, A. M. C., Janssen, A., Boorsma, P. C., Mannens, M., Wilde, A. A. M., & Christiaans, I. (2017). Transthyretin amyloidosis: a phenocopy of hypertrophic cardiomyopathy. Amyloid, 24(2), 87-91.

Wagner, R. A., Tabibiazar, R., Powers, J., Bernstein, D., & Quertermous, T. (2004). Genome-wide expression profiling of a cardiac pressure overload model identifies major metabolic and signaling pathway responses. J Mol Cell Cardiol, 37(6), 1159-1170.

Walsh, R., Buchan, R., Wilk, A., John, S., Felkin, L. E., Thomson, K. L., . . . Cook, S. A. (2017). Defining the genetic architecture of hypertrophic cardiomyopathy: re-evaluating the role of non-sarcomeric genes. Eur Heart J, 38(46), 3461-3468.

Wang, K., Long, B., Liu, F., Wang, J. X., Liu, C. Y., Zhao, B., . . . Li, P. F. (2016). A circular RNA protects the heart from pathological hypertrophy and heart failure by targeting miR-223. Eur Heart J, 37(33), 2602-2611.

Zhang, Qiu, W., Ma, J., Wang, Y., Hu, Z., Long, K., . . . Li, M. (2019). miR-27a-5p Attenuates Hypoxia-induced Rat Cardiomyocyte Injury by Inhibiting Atg7. Int J Mol Sci, 20(10).

Zhang, Shen, Y. J., Shi, J., & Yu, J. G. (2020). MiR-223-3p in Cardiovascular Diseases: A Biomarker and Potential Therapeutic Target. Front Cardiovasc Med, 7, 610561.

Zhao. (2018). Significance of non-coding circular RNAs and micro RNAs in the pathogenesis of cardiovascular diseases. J Med Genet, 55(11), 713-720.

Appendices

Appendix A

ONLINE SUPPLEMENT to *Trim55* identified as functional candidate for spontaneous cardiac hypertrophy in the rat locus Cm22 – published in the *Journal of Hypertension*

Trim55 identified as functional candidate for spontaneous cardiac hypertrophy in the rat locus Cm22

Running title: *Trim55* in polygenic cardiac hypertrophy

Priscilla R PRESTES[a], Francine Z MARQUES[a], Guillermo LOPEZ-CAMPOS[b], Scott A BOOTH[a], Maree McGLYNN[c], Paul LEWANDOWSKI[c], Lea MD DELBRIDGE[d], Stephen B HARRAP[d], Fadi J CHARCHAR[a]

[a] School of Applied and Biomedical Sciences, Faculty of Science and Technology, Federation University Australia, VIC, Australia; [b] Health and Biomedical Informatics Centre, Melbourne Medical School, University of Melbourne, VIC, Australia; [c] School of Medicine, Deakin University, VIC, Australia; [d] Department of Physiology, University of Melbourne, VIC, Australia.

Corresponding author: Fadi Charchar, Room 117, Y Building, University Drive, Federation University, Mt Helen, 3350. P: (03) 5327 6098, Fax: (03) 5327 9602

Email: f.charchar@federation.edu.au

Table S2.1. Characteristics of hypertrophic heart rat (HHR) and normal heart rat (NHR) samples used in this study.

Age	NHR				HHR			
	n	Body weight (g)	Heart weight (mg)	CWI (mg/g)	n	Body weight (g)	Heart weight (mg)	CWI (mg/g)
2-day old	9	5.9±0.2	44.1±2.0	7.4±0.1	8	5.3±0.3	34.3±2.4**	6.6±0.3**
13-week old	11	205.2±15.9	697.0±44.8	3.4± 0.1	10	186.2±13.0	890.8±46.8**	4.9±0.3***
33-week old	9	355.0±36.6	1188.6±86.2	3.4±0.1	7	257.6±29.3*	1069.7±104.1	4.3±0.5*

Footnote: CWI, cardiac weight index. Values are represented as mean ± standard error of mean. * indicates $P<0.05$, ** $P<0.01$, *** $P<0.001$.

Table S2.2. Primer information for real-time quantitative PCR for expression of *Gapdh* and *Trim55* mRNA.

Official gene symbol	Transcript name	Primers sequence (5'-3')	Amplicon size	Primer final concentration (nM)	Annealing temperature
Gapdh	NM_017008.4	F: GGGGCTCTCTGCTCCTCCCTG	108 bp	200	58ºC
		R: ACGGCCAAATCCGTTCACACCG			
Trim55	NM_001012218.1	F: CATGTGTTCCAGAGGCAGAAG	129 bp	200	58ºC
		R: TCTGCAGCACTCCTCAATAGT			

Legend: F, forward; R, reverse; bp, base pairs.

Table S2.3. Change rate in the HHR and NHR. Length of rat chromosomes, number of variants (single nucleotide polymorphisms and insertions/deletions only) and change rate per chromosome in the hypertrophic heart rat (HHR) and normal heart rat (NHR).

		HHR		NHR	
	Length (in base pairs)	Number of variants	Change rate	Number of variants	Change rate
Chr 01	290,094,216	654,017	444	574,868	505
Chr 02	285,068,071	849,235	336	708,630	402
Chr 03	183,740,530	477,212	385	397,786	462
Chr 04	248,343,840	481,334	516	418,327	594
Chr 05	177,180,328	537,974	329	428,648	413
Chr 06	156,897,508	373,813	420	318,855	492
Chr 07	143,501,887	416,114	345	355,662	403
Chr 08	132,457,389	340,779	389	293,217	452
Chr 09	121,549,591	321,186	378	240,380	506
Chr 10	112,200,500	260,307	431	215,293	521
Chr 11	93,518,069	220,565	424	182,893	511
Chr 12	54,450,796	165,715	329	148,108	368
Chr 13	118,718,031	286,972	414	235,617	504
Chr 14	115,151,701	343,102	336	289,967	397
Chr 15	114,627,140	279,691	410	230,448	497
Chr 16	90,051,983	227,189	396	198,242	454
Chr 17	92,503,511	269,530	343	232,677	398
Chr 18	87,229,863	207,221	421	161,305	541
Chr 19	72,914,587	181,860	401	162,757	448

Chr 20	57,791,882	184,735	313	144,526	400
Chr X	154,597,545	208,108	743	197,101	784
mtDNA	16,313	86	190	92	177
Total	**2,902,605,281**	**7,286,745**	**398**	**6,135,399**	**473**

Legend: chr, chromosome, mtDNA, mitochondrial DNA.

Table S2.4. Classification of variants by zygosity (homozygotes and heterozygotes).

	HHR			NHR		
Type	Homo	Hetero	Total	Homo	Hetero	Total
SNP	3,116,825	2,551,840	5,668,665	3,896,877	838,953	4,735,830
Insertion	491,187	293,324	784,511	590,441	88,347	678,788
Deletion	504,416	329,153	833,569	623,014	97,767	720,781

Legend: SNP, single nucleotide polymorphism.

Table S2.5. Classification of the nonsynonymous single nucleotide polymorphisms.

Type of SNP	HHR	NHR
Missense	19,196	16,111
Nonsense	330	254

Legend: SNP, single nucleotide polymorphism.

Table S2.6. Pathways enriched within the 1231 genes that contained nonsynonymous unique single nucleotide polymorphisms in the hypertrophic heart rat using the Kyoto Encyclopedia of Genes and Genomes.

Term	Count	%	*P* value	Genes	Fold Enrichment
rno02010: ABC transporters	14	1.4	0.0000002	*ABCA9, ABCA8A, ABCC10, ABCA2, ABCA4, ABCA6, ABCA5, ABCG5, ABCB1B, TAP1, ABCD2, ABCC2, ABCA13, ABCA12*	6.2
rno04512: ECM-receptor interaction	11	1.1	0.008	*IBSP, LAMA3, CD36, COL3A1, ITGB4, ITGA2, ITGA10, RELN, COL1A1, COL11A2, COL5A2*	2.6
rno00380: Tryptophan metabolism	7	0.7	0.01	*ABP1, CCBL1, HAAO, WARS2, CYP1A2, IDO1, AFMID*	3.5
rno03440: Homologous recombination	5	0.5	0.04	*NBN, POLD2, MUS81, RAD52, RAD54L*	3.7
rno04142: Lysosome	11	1.1	0.08	*LAMP1, AP1B1, PSAP, CTSE, ATP6V1H, CTSD, ABCA2, CD63, MANBA, ATP6V0A2, GGA3*	1.8
rno04120: Ubiquitin mediated proteolysis	12	1.2	0.08	*UBE2N, FANCL, UBE2O, ANAPC5, UBA1, PIAS3, UBE2G1, AIRE, PML, UBE2Q2, UBE2S, UBE2U*	1.7

rno00230: Purine metabolism	13	1.3	0.10	*PDE4D, POLA2, CANT1, GART, ITPA, PRUNE, POLE3, PKLR, PDE5A, POLD2, PAPSS1, PRPS1, NT5C*	1.6

Legend: ABC, adenosine triphosphate-binding cassette; ECM, extracellular matrix.

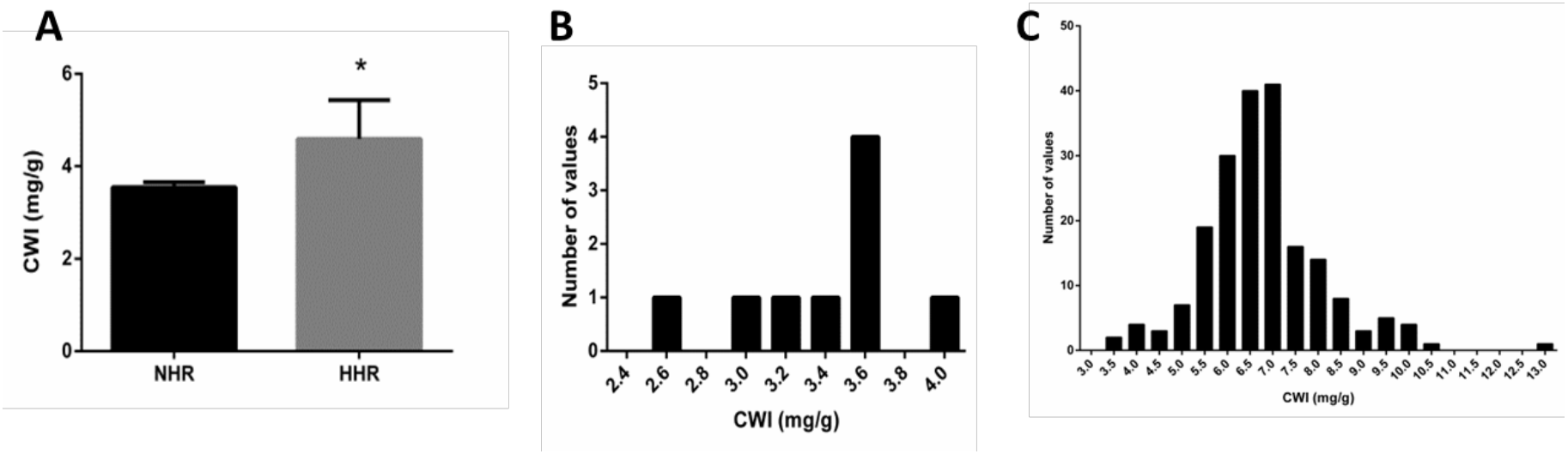

Figure S2.1: Cardiac weight index (mg/g) data from HHRxNHR F2 population. a, NHR and HHR parentals, b, F1 and c, F2 distribution.

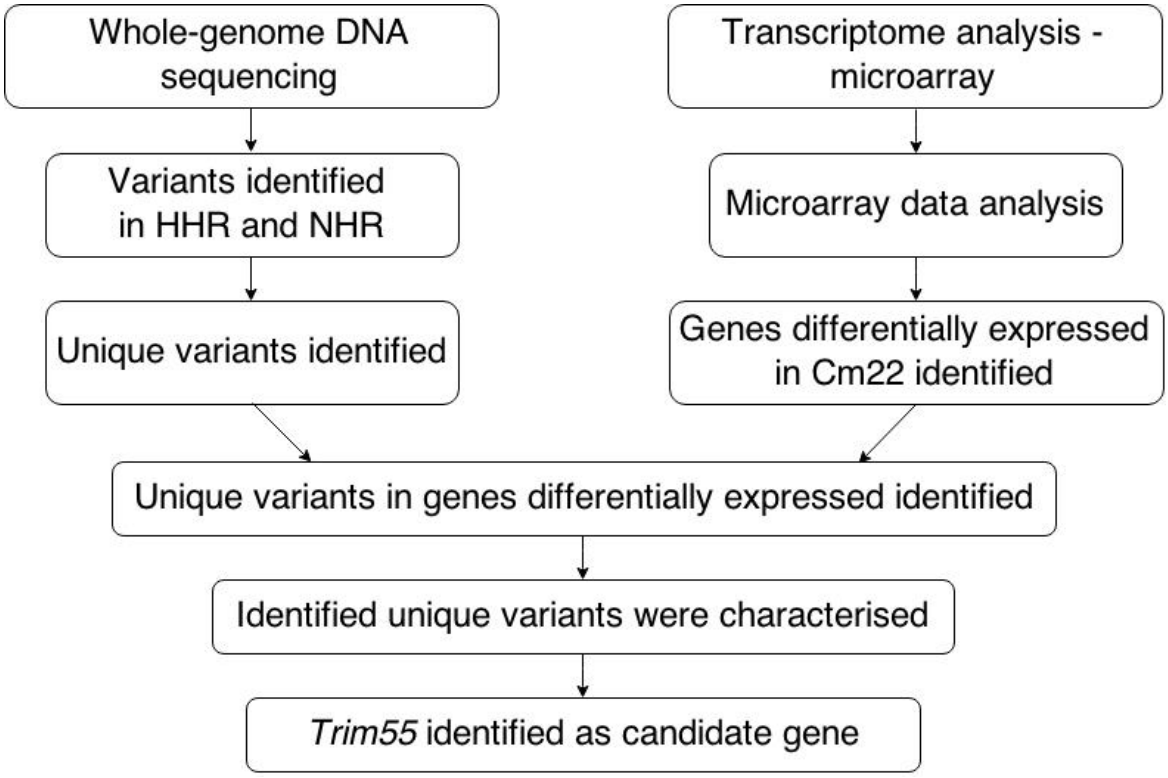

Figure S2.2: Flowchart outlining the combined data analysis of whole-genome sequencing and microarrays. HHR: hypertrophic heart rat; NHR: normal heart rat.

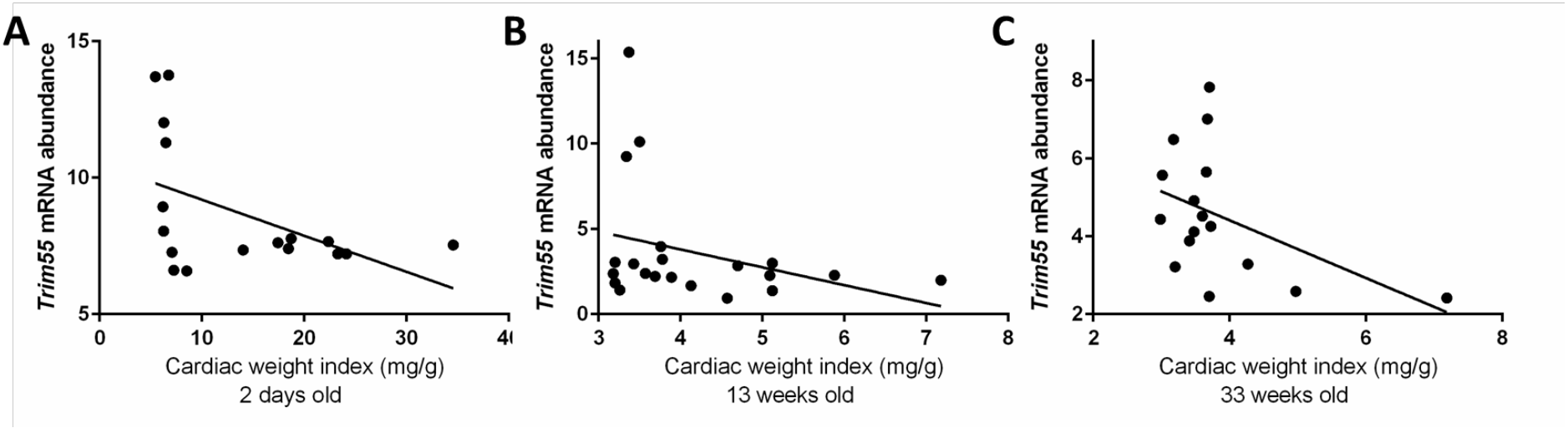

Figure S2.3: Linear regression of cardiac weight index (mg/g) data and *Trim55* mRNA abundance in the F2 a) 2 days old, b) 13 weeks old and c) 33 weeks old animals.

Appendix B

Supplementary table 3.1. Gene orthology comparison between rat and human genes. Percentages indicate coverage similarity between the genomes.

Gene ID	Gene name	Coverage %
Actc1	cardiac actin alpha 1	100
Ankrd1	ankyrin repeat domain 1	91
Casq2	calsequestrin 2	91
Cav3	caveolin 3	95
Cryab	crystallin, alpha B	97
Csrp3	cysteine and glycine-rich protein 3	96
Fhl2	four and a half LIM domains 2	91
Jup	junction plakoglobin	98
Myh6	myosin heavy chain 6	97
Myh7	myosin heavy chain 7	98
Myl3	myosin light chain 3	94
Pln	phospholamban	98
Psen2	presenilin 2	95
Ryr2	ryanodine receptor 2	95

Tnnc1	troponin C type 1	99
Vcl	vinculin	93

Legend: ID, identification

Appendix C

ONLINE SUPPLEMENT to Heart related circular RNA (*Hrcr*) may be involved in polygenic cardiac hypertrophy – Submitted to the Journal of the American Heart Association

Heart related circular RNA (*Hrcr*) may be involved in polygenic cardiac hypertrophy

Short title: *Hrcr* in polygenic cardiac hypertrophy

Prestes PR[a], Maier MC[a], Byars S[b,c], Harrap SB[d], Charchar FJ[a]

[a]School of Science, Psychology and Sport, Federation University Australia, Ballarat, Australia; [b]Centre for Systems Genomics, School of BioSciences, The University of Melbourne, Parkville, Australia; [c]Department of Pathology, The University of Melbourne, Parkville, Australia; [d]Department of Physiology, University of Melbourne, Melbourne, Australia

Corresponding author: Fadi Charchar, Room 117, Y Building, University Drive, Federation University, Mt Helen, 3350. P: (03) 5327 6098, Fax: (03) 5327 9602

Email: f.charchar@federation.edu.au

Supplementary table 4.1. Circular RNA primer information for quantitative real-time PCR (qPCR).

Circular RNA symbol	Circular RNA name	Primer sequence	Primer concentration (nM)	Annealing temperature	Cycles	Organism
Hrcr	Heart related circular RNA	F: TCCGGTTTGTCCTTACATTC	200	58	45	Rat
		R: ACTGGAGAAAATTCGGAGT	200			
Gapdh	Glyceraldehyde 3-phosphate dehydrogenase	F: GGTGGACCTCATGGCCTACA	200	58	45	Rat
		R: CTCTTGCTCTCAGTATCCTTGCT	200			
HRCR	Heart related circular RNA	F: CAGTCTCGCTGCACCTCTAC	200	58	45	Human
		R: GGAATGTATTACTTTGGAGCGGC	200			
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase	F: ACCCACTCCTCCACCTTTGAC	200	58	45	Human
		R: ACCCTGTTGCTGTAGCCAAATT	200			

Supplementary table 4.2. mRNA primer information for quantitative real-time PCR (qPCR).

Official gene symbol	Official gene name	Primer sequence	Primer concentration	Annealing temperature	Cycles	Product size
Arc	Activity regulated cytoskeleton associate protein	F: GCCAGTCTTGGGCAGCATAG	200nM	58°C	40	97
		R: CTGGTATGAATCACTGCTGGGG	200nM			
Gapdh	Glyceraldehyde 3-phosphate dehydrogenase	F: GGGGCTCTCTGCTCCTCCCTG	200nM	58°C	40	108
		R: ACGGCCAAATCCGTTCACACCG	200nM			

Supplementary table 4.3. Small interfering RNA (siRNA) sequences used for *HRCR* silencing.

siRNA name	Type	Sequence	Concentration (pmol)
HRCR	Duplexed	F: rUrCrUrArCrCrCrGrCrUrCrArGrCrArGrGrUrUrUrG	25
		R: rCrArArArArCrCrUrGrCrUrGrArGrCrGrGrGrUrArGrA	25
Scramble	Single-stranded	F: rGrCrCrCrArUrArCrCrGrGrGrUrUrCrArCrUrUrGrU	25

Supplementary table 4.4. MiRNA-sequencing alignment rates provided by the Australian Genome Research Facility (AGRF). Legend: SD - standard deviation, SEM - standard error of the mean.

	sample #	miRNA-seq sample #	initial reads	reads mapped	%
Control	1	scHRCR_2_HY3JYDRXX_CAAAAG_L001_R1.fastq	18832012	12465679	66.2
	2	scHRCR_3_HY3JYDRXX_CACCGG_L001_R1.fastq	18688425	11909556	63.7
	3	scHRCR_4_HY3JYDRXX_CACTCA_L001_R1.fastq	18827692	12333529	65.5
Experimental	4	siHRCR_2_HY3JYDRXX_ATTCCT_L001_R1.fastq	16812209	10061254	60.0
	5	siHRCR_3_HY3JYDRXX_CAACTA_L001_R1.fastq	18922909	12435582	65.7
	6	siHRCR_4_HY3JYDRXX_CACGAT_L001_R1.fastq	24310607	14737583	60.6
				Mean	63.6
				SD	2.7
				SEM	1.1

Supplementary table 4.5. mRNA-miRNA target prediction.

Gene symbol	Gene name	miRNA name
ABCC4	ATP binding cassette subfamily C member 4	hsa-miR-299-5p
ABHD17C	abhydrolase domain containing 17C	hsa-miR-330-3p
ACBD5	acyl-CoA binding domain containing 5	hsa-miR-330-3p
ADAR	adenosine deaminase, RNA specific	hsa-miR-1-3p
ADCY1	adenylate cyclase 1	hsa-miR-27a-5p
ADPGK	ADP dependent glucokinase	hsa-miR-1-3p
AGAP2	ArfGAP with GTPase domain, ankyrin repeat and PH domain 2	hsa-miR-330-3p
AGTR2	angiotensin II receptor type 2	hsa-miR-330-3p
AK7	adenylate kinase 7	hsa-miR-330-3p

ANKRD29	ankyrin repeat domain 29	hsa-miR-1-3p
ANTXR1	ANTXR cell adhesion molecule 1	hsa-miR-27a-5p
AP1G1	adaptor related protein complex 1 subunit gamma 1	hsa-miR-1-3p
API5	apoptosis inhibitor 5	hsa-miR-1-3p
AQP9	aquaporin 9	hsa-miR-330-3p
ARF3	ADP ribosylation factor 3	hsa-miR-1-3p
ASXL3	ASXL transcriptional regulator 3	hsa-miR-1-3p
ATF2	activating transcription factor 2	hsa-miR-1-3p
ATP2B1	ATPase plasma membrane Ca2+ transporting 1	hsa-miR-330-3p
ATP6V1A	ATPase H+ transporting V1 subunit A	hsa-miR-1-3p
AZIN1	antizyme inhibitor 1	hsa-miR-330-3p

BCL11B	BCL11B, BAF complex component	hsa-miR-330-3p
BDNF	brain derived neurotrophic factor	hsa-miR-1-3p; hsa-miR-330-3p
BFSP1	beaded filament structural protein 1	hsa-miR-330-3p
BMI1	BMI1 proto-oncogene, polycomb ring finger	hsa-miR-330-3p
BTAF1	B-TFIID TATA-box binding protein associated factor 1	hsa-miR-1-3p
BTF3	basic transcription factor 3	hsa-miR-27a-5p
BTRC	beta-transducin repeat containing E3 ubiquitin protein ligase	hsa-miR-330-3p
C2CD2	C2 calcium dependent domain containing 2	hsa-miR-330-3p
C2orf69	chromosome 2 open reading frame 69	hsa-miR-1-3p
C3orf62	chromosome 3 open reading frame 62	hsa-miR-330-3p
C5orf51	chromosome 5 open reading frame 51	hsa-miR-1-3p

CAAP1	caspase activity and apoptosis inhibitor 1	hsa-miR-1-3p
CALCR	calcitonin receptor	hsa-miR-299-5p
CAVIN2	caveolae associated protein 2	hsa-miR-1-3p
CBL	Cbl proto-oncogene	hsa-miR-1-3p
CCDC32	coiled-coil domain containing 32	hsa-miR-1-3p
CCN2	cellular communication network factor 2	hsa-miR-330-3p
CCND3	cyclin D3	hsa-miR-330-3p
CCNG1	cyclin G1	hsa-miR-299-5p
CCSAP	centriole, cilia and spindle associated protein	hsa-miR-1-3p
CD2AP	CD2 associated protein	hsa-miR-1-3p
CD83	CD83 molecule	hsa-miR-299-5p

CDCA7	cell division cycle associated 7	hsa-miR-299-5p
CDK14	cyclin dependent kinase 14	hsa-miR-1-3p
CDV3	CDV3 homolog	hsa-miR-330-3p
CEBPZ	CCAAT enhancer binding protein zeta	hsa-miR-1-3p
CEP55	centrosomal protein 55	hsa-miR-299-5p
CERS6	ceramide synthase 6	hsa-miR-330-3p
CLCN3	chloride voltage-gated channel 3	hsa-miR-1-3p
CLOCK	clock circadian regulator	hsa-miR-1-3p
CLTC	clathrin heavy chain	hsa-miR-1-3p
CMPK1	cytidine/uridine monophosphate kinase 1	hsa-miR-330-3p
CNN3	calponin 3	hsa-miR-1-3p

COL6A3	collagen type VI alpha 3 chain	hsa-miR-330-3p
COMMD3-BMI1	COMMD3-BMI1 readthrough	hsa-miR-330-3p
CORO1C	coronin 1C	hsa-miR-1-3p
CPEB2	cytoplasmic polyadenylation element binding protein 2	hsa-miR-299-5p
CPED1	cadherin like and PC-esterase domain containing 1	hsa-miR-1-3p
CREB5	cAMP responsive element binding protein 5	hsa-miR-1-3p
CREBL2	cAMP responsive element binding protein like 2	hsa-miR-1-3p
CSNK1G3	casein kinase 1 gamma 3	hsa-miR-330-3p
CTTNBP2NL	CTTNBP2 N-terminal like	hsa-miR-1-3p
DCAF7	DDB1 and CUL4 associated factor 7	hsa-miR-330-3p

DDHD2	DDHD domain containing 2	hsa-miR-299-5p
DDX5	DEAD-box helicase 5	hsa-miR-1-3p
DEFA4	defensin alpha 4	hsa-miR-330-3p
DICER1	dicer 1, ribonuclease III	hsa-miR-330-3p
DIP2B	disco interacting protein 2 homolog B	hsa-miR-330-3p
DIRAS2	DIRAS family GTPase 2	hsa-miR-330-3p
DLX1	distal-less homeobox 1	hsa-miR-330-3p
DLX6	distal-less homeobox 6	hsa-miR-330-3p
DST	dystonin	hsa-miR-299-5p
EBF2	EBF transcription factor 2	hsa-miR-330-3p
EBPL	EBP like	hsa-miR-1-3p

EDN1	endothelin 1	hsa-miR-1-3p
EIF1AX	eukaryotic translation initiation factor 1A X-linked	hsa-miR-1-3p
EIF4E	eukaryotic translation initiation factor 4E	hsa-miR-1-3p
EIF5	eukaryotic translation initiation factor 5	hsa-miR-27a-5p
ELK4	ELK4, ETS transcription factor	hsa-miR-330-3p
ERAP1	endoplasmic reticulum aminopeptidase 1	hsa-miR-330-3p
ERBB4	erb-b2 receptor tyrosine kinase 4	hsa-miR-330-3p
ETS1	ETS proto-oncogene 1, transcription factor	hsa-miR-1-3p
FAM102A	family with sequence similarity 102 member A	hsa-miR-1-3p
FAM107B	family with sequence similarity 107 member B	hsa-miR-330-3p
FAM114A1	family with sequence similarity 114 member A1	hsa-miR-330-3p

FAM155A	family with sequence similarity 155 member A	hsa-miR-1-3p
FAM168A	family with sequence similarity 168 member A	hsa-miR-1-3p
FAM206A	family with sequence similarity 206 member A	hsa-miR-299-5p
FAM3C	family with sequence similarity 3 member C	hsa-miR-299-5p
FBXO33	F-box protein 33	hsa-miR-1-3p
FGFR1	fibroblast growth factor receptor 1	hsa-miR-330-3p
FNDC3A	fibronectin type III domain containing 3A	hsa-miR-1-3p
FNTB	farnesyltransferase, CAAX box, beta	hsa-miR-330-3p
FOXJ2	forkhead box J2	hsa-miR-330-3p
FOXP1	forkhead box P1	hsa-miR-1-3p
FREM1	FRAS1 related extracellular matrix 1	hsa-miR-299-5p

FRS2	fibroblast growth factor receptor substrate 2	hsa-miR-1-3p
FST	follistatin	hsa-miR-299-5p
G6PD	glucose-6-phosphate dehydrogenase	hsa-miR-1-3p
GCH1	GTP cyclohydrolase 1	hsa-miR-1-3p
GCLC	glutamate-cysteine ligase catalytic subunit	hsa-miR-1-3p
GCN1	GCN1, eIF2 alpha kinase activator homolog	hsa-miR-330-3p
GJA1	gap junction protein alpha 1	hsa-miR-1-3p
GLCCI1	glucocorticoid induced 1	hsa-miR-1-3p
GLCE	glucuronic acid epimerase	hsa-miR-1-3p
GNPDA2	glucosamine-6-phosphate deaminase 2	hsa-miR-1-3p
GNPTG	N-acetylglucosamine-1-phosphate transferase subunit gamma	hsa-miR-299-5p

GPD2	glycerol-3-phosphate dehydrogenase 2	hsa-miR-1-3p
GPR137C	G protein-coupled receptor 137C	hsa-miR-1-3p
GPR34	G protein-coupled receptor 34	hsa-miR-299-5p
GRIA3	glutamate ionotropic receptor AMPA type subunit 3	hsa-miR-330-3p
GRM5	glutamate metabotropic receptor 5	hsa-miR-330-3p
GSDMA	gasdermin A	hsa-miR-27a-5p
GSKIP	GSK3B interacting protein	hsa-miR-330-3p
GUCY1A1	guanylate cyclase 1 soluble subunit alpha 1	hsa-miR-330-3p
GUCY1A2	guanylate cyclase 1 soluble subunit alpha 2	hsa-miR-330-3p
HACD3	3-hydroxyacyl-CoA dehydratase 3	hsa-miR-1-3p
HEATR1	HEAT repeat containing 1	hsa-miR-330-3p

HECW2	HECT, C2 and WW domain containing E3 ubiquitin protein ligase 2	hsa-miR-27a-5p
HELZ2	helicase with zinc finger 2	hsa-miR-1-3p
HIGD1A	HIG1 hypoxia inducible domain family member 1A	hsa-miR-1-3p
HIVEP3	human immunodeficiency virus type I enhancer binding protein 3	hsa-miR-1-3p
HMBOX1	homeobox containing 1	hsa-miR-1-3p
HNRNPU	heterogeneous nuclear ribonucleoprotein U	hsa-miR-1-3p
HP1BP3	heterochromatin protein 1 binding protein 3	hsa-miR-1-3p
HS3ST3B1	heparan sulfate-glucosamine 3-sulfotransferase 3B1	hsa-miR-1-3p
HSP90B1	heat shock protein 90 beta family member 1	hsa-miR-1-3p
HSPD1	heat shock protein family D (Hsp60) member 1	hsa-miR-1-3p
HSPH1	heat shock protein family H (Hsp110) member 1	hsa-miR-330-3p

IFI30	IFI30, lysosomal thiol reductase	hsa-miR-27a-5p
IFNE	interferon epsilon	hsa-miR-330-3p
IFT52	intraflagellar transport 52	hsa-miR-1-3p
IL2	interleukin 2	hsa-miR-27a-5p
IMPACT	impact RWD domain protein	hsa-miR-1-3p
INO80D	INO80 complex subunit D	hsa-miR-330-3p; hsa-miR-27a-5p
INSL5	insulin like 5	hsa-miR-330-3p
IP6K2	inositol hexakisphosphate kinase 2	hsa-miR-1-3p
IRX2	iroquois homeobox 2	hsa-miR-330-3p
ITM2C	integral membrane protein 2C	hsa-miR-330-3p
JMY	junction mediating and regulatory protein, p53 cofactor	hsa-miR-330-3p

JPH1	junctophilin 1	hsa-miR-330-3p
KAT6A	lysine acetyltransferase 6A	hsa-miR-1-3p
KCNC2	potassium voltage-gated channel subfamily C member 2	hsa-miR-299-5p
KCNIP3	potassium voltage-gated channel interacting protein 3	hsa-miR-1-3p
KCNJ16	potassium voltage-gated channel subfamily J member 16	hsa-miR-330-3p
KCTD10	potassium channel tetramerization domain containing 10	hsa-miR-1-3p
KIF2A	kinesin family member 2A	hsa-miR-1-3p
KITLG	KIT ligand	hsa-miR-299-5p
KLF10	Kruppel like factor 10	hsa-miR-330-3p
KLF9	Kruppel like factor 9	hsa-miR-330-3p
KNOP1	lysine rich nucleolar protein 1	hsa-miR-330-3p

KRCC1	lysine rich coiled-coil 1	hsa-miR-330-3p
KSR1	kinase suppressor of ras 1	hsa-miR-330-3p
KTN1	kinectin 1	hsa-miR-1-3p
LAMP3	lysosomal associated membrane protein 3	hsa-miR-299-5p
LAPTM5	lysosomal protein transmembrane 5	hsa-miR-330-3p
LASP1	LIM and SH3 protein 1	hsa-miR-1-3p
LNX2	ligand of numb-protein X 2	hsa-miR-330-3p
LRCH1	leucine rich repeats and calponin homology domain containing 1	hsa-miR-1-3p
LRP2BP	LRP2 binding protein	hsa-miR-330-3p
LRRC8A	leucine rich repeat containing 8 VRAC subunit A	hsa-miR-1-3p
LRRTM3	leucine rich repeat transmembrane neuronal 3	hsa-miR-299-5p

LTBP1	latent transforming growth factor beta binding protein 1	hsa-miR-27a-5p
LTN1	listerin E3 ubiquitin protein ligase 1	hsa-miR-330-3p
LYRM1	LYR motif containing 1	hsa-miR-299-5p
MAB21L1	mab-21 like 1	hsa-miR-1-3p
MAEA	macrophage erythroblast attacher	hsa-miR-299-5p
MAML1	mastermind like transcriptional coactivator 1	hsa-miR-330-3p
MAP3K8	mitogen-activated protein kinase 8	hsa-miR-299-5p
MAP4K3	mitogen-activated protein kinase kinase kinase kinase 3	hsa-miR-1-3p
MARCHF9	membrane associated ring-CH-type finger 9	hsa-miR-299-5p
MAT2A	methionine adenosyltransferase 2A	hsa-miR-299-5p
MATR3	matrin 3	hsa-miR-1-3p

MBNL3	muscleblind like splicing regulator 3	hsa-miR-330-3p
MCHR1	melanin concentrating hormone receptor 1	hsa-miR-1-3p
MECOM	MDS1 and EVI1 complex locus	hsa-miR-1-3p
MEF2A	myocyte enhancer factor 2A	hsa-miR-330-3p
METTL21A	methyltransferase like 21A	hsa-miR-1-3p
MEX3C	mex-3 RNA binding family member C	hsa-miR-1-3p
MFSD14A	major facilitator superfamily domain containing 14A	hsa-miR-1-3p
MFSD5	major facilitator superfamily domain containing 5	hsa-miR-330-3p
MGAT4A	alpha-1,3-mannosyl-glycoprotein 4-beta-N-acetylglucosaminyltransferase A	hsa-miR-1-3p
MLLT10	MLLT10, histone lysine methyltransferase DOT1L cofactor	hsa-miR-330-3p
MMD	monocyte to macrophage differentiation associated	hsa-miR-1-3p

MMD2	monocyte to macrophage differentiation associated 2	hsa-miR-1-3p
MMP12	matrix metallopeptidase 12	hsa-miR-299-5p
MON2	MON2 homolog, regulator of endosome-to-Golgi trafficking	hsa-miR-1-3p
MOSMO	modulator of smoothened	hsa-miR-330-3p
MPP5	membrane palmitoylated protein 5	hsa-miR-1-3p
MRPS6	mitochondrial ribosomal protein S6	hsa-miR-330-3p
MTMR9	myotubularin related protein 9	hsa-miR-330-3p
MXD1	MAX dimerization protein 1	hsa-miR-1-3p
MYEF2	myelin expression factor 2	hsa-miR-299-5p
MYOCD	myocardin	hsa-miR-1-3p
NAV3	neuron navigator 3	hsa-miR-299-5p

NBEAL1	neurobeachin like 1	hsa-miR-330-3p
NCL	nucleolin	hsa-miR-1-3p
NCOA2	nuclear receptor coactivator 2	hsa-miR-330-3p
NEUROD1	neuronal differentiation 1	hsa-miR-299-5p
NFAT5	nuclear factor of activated T cells 5	hsa-miR-1-3p
NFATC2	nuclear factor of activated T cells 2	hsa-miR-1-3p
NFATC3	nuclear factor of activated T cells 3	hsa-miR-1-3p
NFKB1	nuclear factor kappa B subunit 1	hsa-miR-299-5p
NPM1	nucleophosmin 1	hsa-miR-27a-5p
NPTX2	neuronal pentraxin 2	hsa-miR-330-3p
NRP1	neuropilin 1	hsa-miR-1-3p

NSMAF	neutral sphingomyelinase activation associated factor	hsa-miR-299-5p
NUPL2	nucleoporin like 2	hsa-miR-330-3p
NXT2	nuclear transport factor 2 like export factor 2	hsa-miR-1-3p
OSBPL11	oxysterol binding protein like 11	hsa-miR-299-5p
PABPC1L2B	poly(A) binding protein cytoplasmic 1 like 2B	hsa-miR-1-3p
PAK3	p21 (RAC1) activated kinase 3	hsa-miR-299-5p
PAX3	paired box 3	hsa-miR-1-3p
PAX7	paired box 7	hsa-miR-1-3p
PBX1	PBX homeobox 1	hsa-miR-1-3p
PCDHB13	protocadherin beta 13	hsa-miR-1-3p
PDCD10	programmed cell death 10	hsa-miR-1-3p

PDCD4	programmed cell death 4	hsa-miR-1-3p
PDGFC	platelet derived growth factor C	hsa-miR-299-5p
PDIK1L	PDLIM1 interacting kinase 1 like	hsa-miR-1-3p
PDZRN3	PDZ domain containing ring finger 3	hsa-miR-330-3p
PHAX	phosphorylated adaptor for RNA export	hsa-miR-1-3p
PHYHIPL	phytanoyl-CoA 2-hydroxylase interacting protein like	hsa-miR-299-5p
PIK3C2A	phosphatidylinositol-4-phosphate 3-kinase catalytic subunit type 2 alpha	hsa-miR-1-3p
PLPPR4	phospholipid phosphatase related 4	hsa-miR-1-3p
PLXNA4	plexin A4	hsa-miR-1-3p
POGK	pogo transposable element derived with KRAB domain	hsa-miR-1-3p
PPIB	peptidylprolyl isomerase B	hsa-miR-1-3p

PPM1B	protein phosphatase, Mg2+/Mn2+ dependent 1B	hsa-miR-330-3p
PPM1L	protein phosphatase, Mg2+/Mn2+ dependent 1L	hsa-miR-330-3p
PRDM6	PR/SET domain 6	hsa-miR-330-3p
PRIMPOL	primase and DNA directed polymerase	hsa-miR-299-5p
PROX1	prospero homeobox 1	hsa-miR-330-3p
PRR23B	proline rich 23B	hsa-miR-299-5p
PSD3	pleckstrin and Sec7 domain containing 3	hsa-miR-330-3p
PTPRG	protein tyrosine phosphatase, receptor type G	hsa-miR-1-3p
PTPRM	protein tyrosine phosphatase, receptor type M	hsa-miR-330-3p
PURA	purine rich element binding protein A	hsa-miR-330-3p
RABGAP1L	RAB GTPase activating protein 1 like	hsa-miR-1-3p

RAP2A	RAP2A, member of RAS oncogene family	hsa-miR-330-3p
RARB	retinoic acid receptor beta	hsa-miR-1-3p
RBM12	RNA binding motif protein 12	hsa-miR-330-3p
RBPJ	recombination signal binding protein for immunoglobulin kappa J region	hsa-miR-299-5p
RCC2	regulator of chromosome condensation 2	hsa-miR-330-3p
RFK	riboflavin kinase	hsa-miR-27a-5p
RGS10	regulator of G protein signaling 10	hsa-miR-330-3p
RGS7	regulator of G protein signaling 7	hsa-miR-1-3p
RIT2	Ras like without CAAX 2	hsa-miR-1-3p
RND3	Rho family GTPase 3	hsa-miR-330-3p
RNF126	ring finger protein 126	hsa-miR-330-3p

RNF145	ring finger protein 145	hsa-miR-1-3p
RO60	Ro60, Y RNA binding protein	hsa-miR-330-3p
ROBO1	roundabout guidance receptor 1	hsa-miR-299-5p
RSBN1	round spermatid basic protein 1	hsa-miR-299-5p
RSPO2	R-spondin 2	hsa-miR-330-3p
RTL6	retrotransposon Gag like 6	hsa-miR-299-5p
RTN4IP1	reticulon 4 interacting protein 1	hsa-miR-1-3p
S100A7A	S100 calcium binding protein A7A	hsa-miR-1-3p
SAMSN1	SAM domain, SH3 domain and nuclear localization signals 1	hsa-miR-1-3p
SCN7A	sodium voltage-gated channel alpha subunit 7	hsa-miR-299-5p
SCP2	sterol carrier protein 2	hsa-miR-330-3p

SEC22B	SEC22 homolog B, vesicle trafficking protein (gene/pseudogene)	hsa-miR-1-3p
SEC61A1	Sec61 translocon alpha 1 subunit	hsa-miR-1-3p
SERINC3	serine incorporator 3	hsa-miR-330-3p
SERP1	stress associated endoplasmic reticulum protein 1	hsa-miR-1-3p
SGK1	serum/glucocorticoid regulated kinase 1	hsa-miR-299-5p
SIK3	SIK family kinase 3	hsa-miR-330-3p
SLAIN1	SLAIN motif family member 1	hsa-miR-330-3p
SLC10A7	solute carrier family 10 member 7	hsa-miR-1-3p; hsa-miR-330-3p
SLC17A6	solute carrier family 17 member 6	hsa-miR-330-3p
SLC23A2	solute carrier family 23 member 2	hsa-miR-27a-5p; hsa-miR-330-3p
SLC25A22	solute carrier family 25 member 22	hsa-miR-1-3p

SLC25A30	solute carrier family 25 member 30	hsa-miR-1-3p
SLC25A53	solute carrier family 25 member 53	hsa-miR-1-3p
SLC2A2	solute carrier family 2 member 2	hsa-miR-299-5p
SLC39A10	solute carrier family 39 member 10	hsa-miR-1-3p
SLC39A9	solute carrier family 39 member 9	hsa-miR-1-3p
SLC44A1	solute carrier family 44 member 1	hsa-miR-1-3p
SLCO3A1	solute carrier organic anion transporter family member 3A1	hsa-miR-330-3p
SLITRK2	SLIT and NTRK like family member 2	hsa-miR-330-3p
SMARCC1	SWI/SNF related, matrix associated, actin dependent regulator of chromatin subfamily c member 1	hsa-miR-1-3p
SMG7	SMG7, nonsense mediated mRNA decay factor	hsa-miR-299-5p
SMIM14	small integral membrane protein 14	hsa-miR-1-3p

SNX2	sorting nexin 2	hsa-miR-1-3p
SORL1	sortilin related receptor 1	hsa-miR-330-3p
SOSTDC1	sclerostin domain containing 1	hsa-miR-330-3p
SP1	Sp1 transcription factor	hsa-miR-299-5p
SPRED1	sprouty related EVH1 domain containing 1	hsa-miR-1-3p
SPRYD7	SPRY domain containing 7	hsa-miR-330-3p
SRGAP2	SLIT-ROBO Rho GTPase activating protein 2	hsa-miR-1-3p
SRI	sorcin	hsa-miR-1-3p
SRPK2	SRSF protein kinase 2	hsa-miR-330-3p
SRSF9	serine and arginine rich splicing factor 9	hsa-miR-1-3p
SS18	SS18, nBAF chromatin remodeling complex subunit	hsa-miR-1-3p

STAG2	stromal antigen 2	hsa-miR-1-3p
STC2	stanniocalcin 2	hsa-miR-1-3p
STK35	serine/threonine kinase 35	hsa-miR-299-5p
STX16	syntaxin 16	hsa-miR-299-5p
STYX	serine/threonine/tyrosine interacting protein	hsa-miR-299-5p
SULF1	sulfatase 1	hsa-miR-1-3p
SYK	spleen associated tyrosine kinase	hsa-miR-330-3p
SYT1	synaptotagmin 1	hsa-miR-1-3p
SYT16	synaptotagmin 16	hsa-miR-299-5p
TAGLN2	transgelin 2	hsa-miR-1-3p
TARDBP	TAR DNA binding protein	hsa-miR-330-3p

TBC1D7	TBC1 domain family member 7	hsa-miR-330-3p
TBL1XR1	transducin beta like 1 X-linked receptor 1	hsa-miR-330-3p
TCERG1	transcription elongation regulator 1	hsa-miR-330-3p
TFE3	transcription factor binding to IGHM enhancer 3	hsa-miR-1-3p
THADA	THADA, armadillo repeat containing	hsa-miR-330-3p
THAP12	THAP domain containing 12	hsa-miR-1-3p
THUMPD2	THUMP domain containing 2	hsa-miR-299-5p
TIMP3	TIMP metallopeptidase inhibitor 3	hsa-miR-1-3p
TIPARP	TCDD inducible poly(ADP-ribose) polymerase	hsa-miR-299-5p
TMCC1	transmembrane and coiled-coil domain family 1	hsa-miR-1-3p
TMEM178A	transmembrane protein 178A	hsa-miR-1-3p

TMEM243	transmembrane protein 243	hsa-miR-1-3p
TMSB4X	thymosin beta 4 X-linked	hsa-miR-1-3p
TNFAIP8	TNF alpha induced protein 8	hsa-miR-299-5p
TNKS	tankyrase	hsa-miR-330-3p
TNKS2	tankyrase 2	hsa-miR-1-3p
TNPO2	transportin 2	hsa-miR-1-3p
TOX	thymocyte selection associated high mobility group box	hsa-miR-330-3p
TPPP	tubulin polymerization promoting protein	hsa-miR-1-3p
TRA2B	transformer 2 beta homolog	hsa-miR-1-3p
TRAPPC3	trafficking protein particle complex 3	hsa-miR-1-3p
TRAPPC8	trafficking protein particle complex 8	hsa-miR-330-3p

TRIB2	tribbles pseudokinase 2	hsa-miR-330-3p
TRIM2	tripartite motif containing 2	hsa-miR-330-3p
TRPS1	transcriptional repressor GATA binding 1	hsa-miR-330-3p
TSPAN2	tetraspanin 2	hsa-miR-299-5p
TSPAN4	tetraspanin 4	hsa-miR-1-3p
TTC37	tetratricopeptide repeat domain 37	hsa-miR-299-5p
TWF1	twinfilin actin binding protein 1	hsa-miR-1-3p
UBE2Q2	ubiquitin conjugating enzyme E2 Q2	hsa-miR-299-5p
UBL3	ubiquitin like 3	hsa-miR-330-3p
UNC119B	unc-119 lipid binding chaperone B	hsa-miR-1-3p
UST	uronyl 2-sulfotransferase	hsa-miR-1-3p

UTRN	utrophin	hsa-miR-1-3p
VANGL1	VANGL planar cell polarity protein 1	hsa-miR-299-5p
VGLL3	vestigial like family member 3	hsa-miR-330-3p
WBP1L	WW domain binding protein 1 like	hsa-miR-1-3p
WDFY3	WD repeat and FYVE domain containing 3	hsa-miR-330-3p
WDR48	WD repeat domain 48	hsa-miR-1-3p
WDR76	WD repeat domain 76	hsa-miR-330-3p
WWC1	WW and C2 domain containing 1	hsa-miR-1-3p
XPO6	exportin 6	hsa-miR-1-3p
YLPM1	YLP motif containing 1	hsa-miR-330-3p
ZBTB34	zinc finger and BTB domain containing 34	hsa-miR-330-3p

ZBTB41	zinc finger and BTB domain containing 41	hsa-miR-1-3p
ZDHHC5	zinc finger DHHC-type containing 5	hsa-miR-299-5p
ZFC3H1	zinc finger C3H1-type containing	hsa-miR-330-3p
ZFP36L1	ZFP36 ring finger protein like 1	hsa-miR-1-3p
ZFP36L2	ZFP36 ring finger protein like 2	hsa-miR-1-3p
ZFR	zinc finger RNA binding protein	hsa-miR-330-3p
ZMAT3	zinc finger matrin-type 3	hsa-miR-330-3p
ZNF281	zinc finger protein 281	hsa-miR-1-3p
ZNF423	zinc finger protein 423	hsa-miR-330-3p
ZNF706	zinc finger protein 706	hsa-miR-330-3p

Supplementary table 4.6. Gene ontology of predicted mRNA targets. Legend: GO - gene ontology, ID - identification, FDR - false discovery rate, BP - biological process, MF - molecular function, CC - cellular component.

GO term	GO term ID	Fold enrichment	P-value	FDR	Type
cellular response to mineralocorticoid stimulus	71389	24.25	0.00057	0.04930	BP
epithelial cell fate commitment	72148	15.09	0.00028	0.02950	BP
regulation of vascular associated smooth muscle cell differentiation	1905063	15.09	0.00028	0.02930	BP
ventricular cardiac muscle cell differentiation	55012	15.09	0.00028	0.02910	BP
negative regulation of smooth muscle cell differentiation	51151	13.31	0.00042	0.03930	BP
positive regulation of transcription of Notch receptor target	7221	12.57	0.00051	0.04530	BP
regulation of smooth muscle cell differentiation	51150	8.84	0.00042	0.03910	BP
negative regulation of muscle cell differentiation	51148	8.23	0.00001	0.00224	BP
cellular response to corticosteroid stimulus	71384	8.21	0.00000	0.00075	BP
cellular response to glucocorticoid stimulus	71385	7.94	0.00002	0.00277	BP
maintenance of protein location	45185	5.25	0.00010	0.01220	BP
cardiac septum development	3279	5.04	0.00013	0.01540	BP

cardiomyocyte differentiation	35051	4.99	0.00014	0.01610	BP
regulation of blood vessel endothelial cell migration	43535	4.97	0.00033	0.03270	BP
regulation of organ growth	46620	4.97	0.00033	0.03250	BP
regulation of muscle cell differentiation	51147	4.92	0.00001	0.00222	BP
maintenance of location in cell	51651	4.82	0.00041	0.03800	BP
positive regulation of neuron differentiation	45666	4.76	0.00043	0.03960	BP
positive regulation of endothelial cell migration	10595	4.71	0.00021	0.02270	BP
regulation of endothelial cell migration	10594	4.33	0.00002	0.00330	BP
response to corticosteroid	31960	4.22	0.00005	0.00755	BP
cellular response to steroid hormone stimulus	71383	3.98	0.00033	0.03230	BP
response to glucocorticoid	51384	3.96	0.00035	0.03380	BP
regulation of neuron differentiation	45664	3.79	0.00007	0.00983	BP
monocarboxylic acid transport	15718	3.72	0.00055	0.04820	BP
regulation of epithelial cell migration	10632	3.67	0.00003	0.00445	BP
negative regulation of cell migration	30336	3.39	0.00004	0.00596	BP

negative regulation of cell motility	2000146	3.21	0.00007	0.00978	BP
positive regulation of MAP kinase activity	43406	3.16	0.00039	0.03690	BP
regulation of epithelial cell proliferation	50678	3.14	0.00002	0.00333	BP
negative regulation of cellular component movement	51271	3.12	0.00010	0.01210	BP
regulation of MAP kinase activity	43405	3.09	0.00007	0.00930	BP
negative regulation of locomotion	40013	3.06	0.00007	0.00995	BP
nucleocytoplasmic transport	6913	3.06	0.00032	0.03190	BP
MAPK cascade	165	3.05	0.00001	0.00136	BP
nuclear transport	51169	3.02	0.00036	0.03440	BP
regulation of cellular response to growth factor stimulus	90287	2.91	0.00033	0.03250	BP
regulation of developmental growth	48638	2.85	0.00027	0.02810	BP
negative regulation of cell differentiation	45596	2.66	0.00000	0.00049	BP
regulation of cellular amide metabolic process	34248	2.57	0.00017	0.01940	BP
cell leading edge	31252	2.55	0.00028	0.03140	CC
skeletal system development	1501	2.52	0.00016	0.01760	BP

negative regulation of multicellular organismal process	51241	2.48	0.00000	0.00003	BP
regulation of protein kinase activity	45859	2.48	0.00000	0.00041	BP
negative regulation of developmental process	51093	2.47	0.00000	0.00018	BP
positive regulation of transferase activity	51347	2.47	0.00001	0.00214	BP
heart development	7507	2.45	0.00024	0.02620	BP
regulation of protein serine/threonine kinase activity	71900	2.41	0.00039	0.03640	BP
regulation of kinase activity	43549	2.36	0.00000	0.00049	BP
positive regulation of protein kinase activity	45860	2.33	0.00037	0.03540	BP
positive regulation of kinase activity	33674	2.32	0.00014	0.01650	BP
regulation of transferase activity	51338	2.31	0.00000	0.00022	BP
positive regulation of developmental process	51094	2.29	0.00000	0.00003	BP
regulation of protein phosphorylation	1932	2.29	0.00000	0.00006	BP
positive regulation of cell differentiation	45597	2.26	0.00002	0.00363	BP
regulation of MAPK cascade	43408	2.19	0.00021	0.02260	BP
negative regulation of transcription by RNA polymerase II	122	2.19	0.00003	0.00432	BP

regulation of cell differentiation	45595	2.19	0.00000	0.00001	BP
positive regulation of cell population proliferation	8284	2.19	0.00002	0.00334	BP
positive regulation of transcription by RNA polymerase II	45944	2.16	0.00000	0.00039	BP
circulatory system development	72359	2.13	0.00010	0.01210	BP
positive regulation of protein phosphorylation	1934	2.12	0.00014	0.01610	BP
regulation of cell migration	30334	2.11	0.00008	0.01050	BP
regulation of protein modification process	31399	2.11	0.00000	0.00003	BP
response to organic cyclic compound	14070	2.11	0.00008	0.01060	BP
animal organ morphogenesis	9887	2.10	0.00007	0.00993	BP
regulation of developmental process	50793	2.09	0.00000	0.00000	BP
positive regulation of catalytic activity	43085	2.08	0.00000	0.00030	BP
positive regulation of protein modification process	31401	2.08	0.00002	0.00388	BP
chromatin	785	2.07	0.00001	0.00181	CC
regulation of phosphorylation	42325	2.06	0.00000	0.00041	BP
positive regulation of phosphorus metabolic process	10562	2.06	0.00007	0.00928	BP

positive regulation of phosphate metabolic process	45937	2.06	0.00007	0.00920	BP
positive regulation of RNA metabolic process	51254	2.04	0.00000	0.00005	BP
negative regulation of cell death	60548	2.04	0.00007	0.00991	BP
positive regulation of nitrogen compound metabolic process	51173	2.03	0.00000	0.00000	BP
regulation of phosphate metabolic process	19220	2.03	0.00000	0.00014	BP
regulation of phosphorus metabolic process	51174	2.03	0.00000	0.00014	BP
positive regulation of nucleobase-containing compound metabolic process	45935	2.03	0.00000	0.00002	BP
regulation of cell motility	2000145	2.03	0.00013	0.01580	BP
negative regulation of biosynthetic process	9890	2.03	0.00000	0.00017	BP
tube development	35295	2.02	0.00044	0.04010	BP
positive regulation of molecular function	44093	2.02	0.00000	0.00006	BP
positive regulation of signal transduction	9967	2.01	0.00000	0.00030	BP
negative regulation of macromolecule biosynthetic process	10558	2.00	0.00000	0.00050	BP
regulation of protein metabolic process	51246	2.00	0.00000	0.00000	BP
negative regulation of programmed cell death	43069	2.00	0.00028	0.02890	BP

positive regulation of nucleic acid-templated transcription	1903508	1.99	0.00000	0.00031	BP
positive regulation of transcription, DNA-templated	45893	1.99	0.00000	0.00030	BP
positive regulation of RNA biosynthetic process	1902680	1.99	0.00000	0.00030	BP
regulation of cell death	10941	1.98	0.00000	0.00024	BP
positive regulation of phosphorylation	42327	1.98	0.00031	0.03110	BP
negative regulation of protein metabolic process	51248	1.98	0.00011	0.01340	BP
positive regulation of cell communication	10647	1.98	0.00000	0.00017	BP
negative regulation of cellular macromolecule biosynthetic process	2000113	1.98	0.00000	0.00075	BP
regulation of programmed cell death	43067	1.98	0.00000	0.00059	BP
positive regulation of signaling	23056	1.97	0.00000	0.00017	BP
negative regulation of cellular biosynthetic process	31327	1.96	0.00000	0.00071	BP
regulation of cell population proliferation	42127	1.95	0.00000	0.00038	BP
negative regulation of cellular protein metabolic process	32269	1.95	0.00029	0.02980	BP
positive regulation of protein metabolic process	51247	1.95	0.00001	0.00099	BP
regulation of locomotion	40012	1.94	0.00031	0.03110	BP

negative regulation of transcription, DNA-templated	45892	1.93	0.00005	0.00707	BP
regulation of cellular component movement	51270	1.93	0.00026	0.02760	BP
negative regulation of nucleic acid-templated transcription	1903507	1.93	0.00005	0.00705	BP
cellular response to endogenous stimulus	71495	1.93	0.00013	0.01530	BP
negative regulation of RNA biosynthetic process	1902679	1.93	0.00005	0.00711	BP
regulation of cellular protein metabolic process	32268	1.92	0.00000	0.00001	BP
regulation of apoptotic process	42981	1.91	0.00002	0.00256	BP
positive regulation of biosynthetic process	9891	1.90	0.00000	0.00015	BP
positive regulation of macromolecule biosynthetic process	10557	1.89	0.00000	0.00046	BP
negative regulation of gene expression	10629	1.89	0.00000	0.00017	BP
regulation of multicellular organismal development	2000026	1.89	0.00009	0.01140	BP
positive regulation of cellular metabolic process	31325	1.89	0.00000	0.00000	BP
regulation of multicellular organismal process	51239	1.88	0.00000	0.00001	BP
positive regulation of macromolecule metabolic process	10604	1.88	0.00000	0.00000	BP
negative regulation of nitrogen compound metabolic process	51172	1.87	0.00000	0.00003	BP

negative regulation of RNA metabolic process	51253	1.87	0.00006	0.00858	BP
negative regulation of macromolecule metabolic process	10605	1.86	0.00000	0.00000	BP
positive regulation of cellular biosynthetic process	31328	1.85	0.00000	0.00050	BP
negative regulation of metabolic process	9892	1.84	0.00000	0.00000	BP
negative regulation of nucleobase-containing compound metabolic process	45934	1.83	0.00007	0.00975	BP
positive regulation of metabolic process	9893	1.83	0.00000	0.00000	BP
positive regulation of gene expression	10628	1.81	0.00000	0.00017	BP
response to endogenous stimulus	9719	1.80	0.00017	0.01850	BP
positive regulation of multicellular organismal process	51240	1.80	0.00021	0.02280	BP
positive regulation of cellular protein metabolic process	32270	1.79	0.00019	0.02090	BP
regulation of catalytic activity	50790	1.78	0.00000	0.00016	BP
negative regulation of cellular metabolic process	31324	1.77	0.00000	0.00014	BP
anatomical structure morphogenesis	9653	1.75	0.00001	0.00171	BP
regulation of molecular function	65009	1.73	0.00000	0.00003	BP
regulation of signal transduction	9966	1.71	0.00000	0.00009	BP

establishment of protein localization	45184	1.71	0.00045	0.04010	BP
regulation of nitrogen compound metabolic process	51171	1.70	0.00000	0.00000	BP
animal organ development	48513	1.69	0.00000	0.00010	BP
chromosome	5694	1.69	0.00028	0.02980	CC
regulation of transcription by RNA polymerase II	6357	1.68	0.00001	0.00100	BP
regulation of primary metabolic process	80090	1.68	0.00000	0.00000	BP
regulation of cell communication	10646	1.67	0.00000	0.00006	BP
regulation of signaling	23051	1.67	0.00000	0.00006	BP
positive regulation of response to stimulus	48584	1.66	0.00005	0.00710	BP
regulation of macromolecule metabolic process	60255	1.66	0.00000	0.00000	BP
negative regulation of cellular process	48523	1.66	0.00000	0.00000	BP
regulation of intracellular signal transduction	1902531	1.66	0.00047	0.04180	BP
regulation of cellular metabolic process	31323	1.64	0.00000	0.00000	BP
regulation of localization	32879	1.63	0.00001	0.00215	BP
negative regulation of biological process	48519	1.63	0.00000	0.00000	BP

nitrogen compound transport	71705	1.63	0.00053	0.04690	BP
cellular response to organic substance	71310	1.63	0.00008	0.01080	BP
system development	48731	1.61	0.00000	0.00002	BP
response to organic substance	10033	1.61	0.00001	0.00175	BP
regulation of biosynthetic process	9889	1.60	0.00000	0.00003	BP
regulation of macromolecule biosynthetic process	10556	1.60	0.00000	0.00008	BP
cellular response to chemical stimulus	70887	1.60	0.00002	0.00324	BP
positive regulation of cellular process	48522	1.60	0.00000	0.00000	BP
positive regulation of biological process	48518	1.59	0.00000	0.00000	BP
regulation of metabolic process	19222	1.58	0.00000	0.00000	BP
regulation of cellular biosynthetic process	31326	1.58	0.00000	0.00008	BP
regulation of gene expression	10468	1.58	0.00000	0.00001	BP
regulation of RNA metabolic process	51252	1.57	0.00000	0.00035	BP
macromolecule localization	33036	1.57	0.00016	0.01760	BP
regulation of cellular macromolecule biosynthetic process	2000112	1.57	0.00000	0.00022	BP

protein localization	8104	1.57	0.00054	0.04770	BP
regulation of nucleobase-containing compound metabolic process	19219	1.57	0.00000	0.00020	BP
multicellular organism development	7275	1.57	0.00000	0.00001	BP
nucleoplasm	5654	1.57	0.00000	0.00013	CC
regulation of transcription, DNA-templated	6355	1.55	0.00001	0.00203	BP
regulation of nucleic acid-templated transcription	1903506	1.55	0.00001	0.00201	BP
regulation of RNA biosynthetic process	2001141	1.55	0.00001	0.00202	BP
nuclear lumen	31981	1.55	0.00000	0.00008	CC
nucleic acid binding	3676	1.54	0.00000	0.00244	MF
regulation of response to stimulus	48583	1.53	0.00000	0.00031	BP
anatomical structure development	48856	1.50	0.00000	0.00004	BP
developmental process	32502	1.50	0.00000	0.00001	BP
heterocyclic compound binding	1901363	1.47	0.00000	0.00012	MF
organic cyclic compound binding	97159	1.46	0.00000	0.00012	MF
cellular developmental process	48869	1.44	0.00027	0.02860	BP

cell differentiation	30154	1.43	0.00043	0.03940	BP
intracellular organelle lumen	70013	1.43	0.00000	0.00068	CC
organelle lumen	43233	1.42	0.00000	0.00075	CC
membrane-enclosed lumen	31974	1.42	0.00000	0.00069	CC
nucleus	5634	1.42	0.00000	0.00000	CC
cytosol	5829	1.38	0.00004	0.00584	CC
cell communication	7154	1.35	0.00009	0.01090	BP
cellular response to stimulus	51716	1.32	0.00004	0.00588	BP
intracellular membrane-bounded organelle	43231	1.30	0.00000	0.00000	CC
multicellular organismal process	32501	1.29	0.00011	0.01320	BP
regulation of cellular process	50794	1.29	0.00000	0.00000	BP
regulation of biological process	50789	1.25	0.00000	0.00001	BP
cytoplasm	5737	1.24	0.00000	0.00002	CC
intracellular organelle	43229	1.24	0.00000	0.00000	CC
membrane-bounded organelle	43227	1.23	0.00000	0.00000	CC

biological regulation	65007	1.21	0.00000	0.00014	BP
organelle	43226	1.21	0.00000	0.00000	CC
intracellular anatomical structure	5622	1.20	0.00000	0.00000	CC
protein binding	5515	1.19	0.00000	0.00013	MF
binding	5488	1.13	0.00000	0.00021	MF
cellular process	9987	1.11	0.00011	0.01280	BP
molecular function	3674	1.08	0.00001	0.00954	MF
cellular anatomical entity	110165	1.06	0.00011	0.01470	CC
cellular component	5575	1.05	0.00027	0.03210	CC
sensory perception	7600	0.17	0.00005	0.00708	BP
detection of chemical stimulus involved in sensory perception	50907	< 0.01	0.00031	0.03120	BP
detection of stimulus involved in sensory perception	50906	< 0.01	0.00010	0.01210	BP
sensory perception of smell	7608	< 0.01	0.00047	0.04200	BP
sensory perception of chemical stimulus	7606	< 0.01	0.00015	0.01710	BP