



City Research Online

City St George's, University of London

Citation: Muggenthaler, M. M. A. (2019). Investigation of inherited cardiovascular disorders in the Anabaptist communities. (Unpublished Doctoral thesis, St. Georges, University of London)

This is the accepted version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

Permanent repository link: <https://openaccess.city.ac.uk/id/eprint/37334/>

Copyright and Reuse: Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to [City Research Online policy](#).

Investigation of inherited cardiovascular disorders in the Anabaptist communities

PhD thesis

2019

St George's, University of London

Martina Maria Auguste Muggenthaler

Declaration

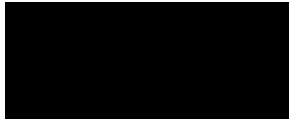
Submitted by

Martina Maria Auguste Muggenthaler

To St George's, University of London as a thesis for the degree of Doctor of Philosophy, April 2019.

This thesis is available for library use on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

I certify that all material in this thesis which is not my own work has been identified and that no material has previously been submitted and approved for the award of a degree by this or any other University.



Abstract

The work described in this thesis involves the detailed investigation of two inherited cardiovascular conditions identified amongst the Anabaptist communities of Ohio (USA) as part of a long-running clinical-genetics program, called 'Windows of Hope'.

The Anabaptist communities in North America form closed founder populations with a homogenous lifestyle, large families and extensive genealogical records enabling the construction of extensive family trees. These and other factors greatly empower the investigation of genetic disorders in the community. The first series of studies relate to a novel autosomal recessive syndrome consistent of cleft lip and palate and cor triatriatum sinister observed in multiple Amish families. The project aims were to clinically characterise the phenotype, to define the genetic basis of the condition, and to investigate the molecular function of the molecule identified. Genomic studies (genome-wide autozygosity mapping and exome sequencing) defined an Amish founder gene variant in the *HYAL2* gene, encoding hyaluronidase 2, as the likely cause of this condition. Subsequent functional and mouse model studies confirmed dramatic reductions in *HYAL2* protein levels associated with gene mutation and revealed overlapping phenotypical outcomes in *Hyal2* knock-out mice which mirror the human condition. These findings highlight the vital role hyaluronidase enzymes play in both human and mouse development and reveal a previously unrecognised pathway involved in the pathogenesis of cor triatriatum sinister and orofacial clefting, which may ultimately be amenable to treatment strategies.

The second disorder investigated is a form of hypertrophic cardiomyopathy (HCM) associated with a splice site variant (resulting in frameshift) in the *MYBPC3* gene, encoding myosin binding protein C3. HCM is the most common inherited cardiovascular disease worldwide, and affected individuals show extreme variations in severity of phenotype, even when carrying the same causative genetic variant. The mechanisms underlying the phenotypic variation are unknown, but genetic

modifying factors are likely involved impacting on expression of the phenotype. This project comprises the clinical characterisation of the largest cohort of individuals carrying the same HCM pathogenic variant described to date. Detailed molecular studies identified potential genetic modifiers of the phenotype using a combined approach of whole exome sequence analysis, genome wide SNP data analyses, and gene-splicing transcript outcomes. This study provided important insight into the expression of the phenotype, confirmed the significant phenotypic variation amongst individuals harbouring the same HCM causing genetic variant observed in other HCM cohorts, and identified putative genetic modifiers contributing to an improved understanding of the development of the variable phenotype in *MYBPC3*-associated HCM, which may ultimately enable more accurate prognostication and personalised treatment.

An additional and overarching objective of the project, in keeping with the wider aims of the Windows of Hope studies, was the translation of the beneficial findings of this research into direct clinical benefits to the families and communities involved.

Supervisors

Professor Andrew Crosby
Professor of Human Genetics
Molecular Genetics
University of Exeter Medical School
Exeter
EX1 2LU

Professor Elijah R Behr
Professor in Cardiovascular Medicine
Honorary Consultant Cardiologist and Electrophysiologist
St George's University of London
London
SW17 0RE

Professor Sanjay Sharma
Professor of Cardiology
St George's University of London
London
SW17 0RE

Dedication

Für meine Eltern.

Acknowledgements

I would like to thank the following individuals, without whom this project would have not been possible:

Prof Andrew Crosby for his continuous guidance and advice. Without his eternal optimism, inspiration and tenacity this study would have never been completed.

Prof Elijah Behr and Prof Sanjay Sharma for their unwavering support and guidance, and for teaching me everything I know about inherited cardiac conditions.

Dr Barry Chioza and Dr Gaurav Harlalka for their patience and enormous support with everything technical and molecular.

Dr Emma Baple for trying to teach me dysmorphology and her invaluable advice regarding counselling and genetic testing.

The members of the WOH team, in particular Roselyn Coblenz, Karin Wagner and Kay Gurtz. I would not have been able to complete this project without their endless motivation, kind words and resourcefulness.

Prof Barbara Triggs-Raine, Prof Lorna Harries, and Prof Mike Wheedon and their teams for their collaboration and sharing their knowledge and experience.

Dr Sam Jones for his help with the SNP association studies.

Cardiac Risk in the Young for partially sponsoring my salary and providing me with an echo machine for the clinical phenotyping.

The Medical Research Council for partially sponsoring my salary and partially funding the genetic studies in this project.

On a personal level, I would like to thank my family for their love and support, and for never doubting I could do anything I chose to.

And above all, the participants of the study and the wider Amish community for their support of the Windows of Hope project.

Table of Contents

1	INTRODUCTION.....	16
1.1	CARDIAC EMBRYOLOGICAL DEVELOPMENT	16
1.1.1	FORMATION OF THE LINEAR HEART TUBE	16
1.1.2	DEVELOPMENT OF THE FOUR HEART CHAMBERS	19
1.1.3	SEPTATION AND VALVE FORMATION	21
1.2	CONGENITAL HEART DISEASE	25
1.2.1	EPIDEMIOLOGY	25
1.2.2	GENETICS OF CONGENITAL HEART DISEASE.....	25
1.3	HYPERTROPHIC CARDIOMYOPATHY	30
1.3.1	EPIDEMIOLOGY, MORPHOLOGY AND CLINICAL PRESENTATION	30
1.3.2	GENETIC BASIS OF HYPERTROPHIC CARDIOMYOPATHY.....	32
1.3.3	VARIABLE PHENOTYPIC EXPRESSION AND GENETIC MODIFIERS	36
1.4	THE AMISH POPULATION AND THE WINDOWS OF HOPE PROJECT	42
1.4.1	AMISH HISTORY AND CULTURE.....	42
1.4.2	THE WINDOWS OF HOPE PROJECT	48
1.5	CARDIOVASCULAR DISORDERS IN THE ANABAPTIST COMMUNITIES	49
1.5.1	HYPERTROPHIC CARDIOMYOPATHY IN THE AMISH.....	49
1.5.2	OTHER CARDIOVASCULAR DISORDERS IN ANABAPTIST COMMUNITIES	50
2	ORIGINAL HYPOTHESES AND AIMS	54
2.1	SYNDROME OF OROFACIAL CLEFTING AND CARDIAC ANOMALIES	54
2.2	HYPERTROPHIC CARDIOMYOPATHY	54
3	METHODS	56
3.1	ETHICAL APPROVAL	56
3.2	RECRUITMENT AND PHENOTYPIC ASSESSMENT	56
3.2.1	OROFACIAL CLEFTING AND CARDIAC ANOMALIES PROJECT	56
3.2.2	HYPERTROPHIC CARDIOMYOPATHY PROJECT	57

3.3 DATA MANAGEMENT.....	63
3.4 MOLECULAR METHODS	64
3.4.1 DNA EXTRACTION	64
3.4.2 RNA EXTRACTION	65
3.4.3 PRIMER DESIGN	67
3.4.4 POLYMERASE CHAIN REACTION	67
3.4.5 OPTIMISATION OF AMPLIFICATION CONDITIONS.....	69
3.4.6 AGAROSE GEL ELECTROPHORESIS	70
3.4.7 PURIFICATION OF PCR PRODUCTS.....	70
3.4.8 SEQUENCING REACTION	71
3.4.9 SEQUENCING REACTION PURIFICATION.....	71
3.4.10 SINGLE NUCLEOTIDE POLYMORPHISM (SNP) GENOTYPING	72
3.4.11 EXOME SEQUENCING	72
3.4.12 REVERSE TRANSCRIPTION PCR.....	74
3.4.13 OTHER METHODS USED IN THE <i>HYAL2</i> PROJECT	76
 4 <u>MUTATIONS IN <i>HYAL2</i>, ENCODING HYALURONIDASE 2, CAUSE A SYNDROME OF COR TRIATRIATUM SINISTER AND OROFACIAL CLEFTING IN HUMANS AND MICE</u>	 79
4.1 SUMMARY.....	79
4.2 INTRODUCTION.....	80
4.2.1 COR TRIATRIATUM SINISTER	80
4.2.2 OROFACIAL CLEFTING.....	82
4.2.3 HYALURONAN AND HYALURONIDASE 2	83
4.3 STUDY-SPECIFIC METHODS	85
4.3.1 EXOME SEQUENCING	85
4.3.2 FUNCTIONAL STUDIES	85
4.3.3 MOUSE STUDIES	85
4.4 RESULTS.....	86
4.4.1 SUMMARY OF PHENOTYPE DATA	86
4.4.2 IDENTIFICATION OF PUTATIVE PATHOGENIC VARIANTS IN <i>HYAL2</i>	91
4.4.3 <i>HYAL2</i> P.K148R AND P.P250L SIGNIFICANTLY REDUCE <i>HYAL2</i> PROTEIN LEVELS	95

4.4.4	<i>HYAL2</i> P.K148R SIGNIFICANTLY REDUCES ENZYME ACTIVITY IN <i>HYAL1</i>	95
4.4.5	<i>HYAL2</i> ^{-/-} MICE EXHIBIT A SIMILAR PHENOTYPE TO HUMANS WITH <i>HYAL2</i> MUTATIONS	99
4.5	DISCUSSION	103

5 INVESTIGATION OF THE PHENOTYPE AND GENETIC MODIFIERS IN HYPERTROPHIC CARDIOMYOPATHY CAUSED BY THE C.3330+2T>G VARIANT IN *MYBPC3* 109

5.1	SUMMARY	109
5.2	INTRODUCTION	111
5.3	INVESTIGATION OF THE PHENOTYPE OF CARRIERS OF THE <i>MYBPC3</i> C.3330+2T>G VARIANT	113
5.3.1	SPECIFIC METHODS	113
5.3.2	RESULTS	114
5.3.3	DISCUSSION	138
5.4	INVESTIGATION OF THE TRANSCRIPTIONAL OUTCOMES DERIVING FROM THE PATHOGENIC C.3330+2T>G VARIANT IN <i>MYBPC3</i>	147
5.4.1	SPECIFIC METHODS	147
5.4.2	RESULTS	148
5.4.3	DISCUSSION	152
5.5	EXOME SEQUENCE ANALYSIS OF CARRIERS OF THE <i>MYBPC3</i> C.3330+2T>G VARIANT WITH SEVERE PHENOTYPE	157
5.5.1	SPECIFIC METHODS	157
5.5.2	RESULTS	158
5.5.3	DISCUSSION	168
5.6	INVESTIGATION OF SNP ASSOCIATION WITH PHENOTYPE SEVERITY IN THE <i>MYBPC3</i> C.3330+2T>G VARIANT COHORT	175
5.6.1	SPECIFIC METHODS	175
5.6.2	RESULTS	177
5.6.3	DISCUSSION	180

6 CONCLUDING REMARKS 185

6.1	CLINICAL BENEFITS TO THE AMISH COMMUNITY	185
6.2	BENEFITS TO THE WIDER SOCIETY AND SCIENTIFIC BENEFITS	189

7	SUPPLEMENTAL DATA.....	192
7.1	ADDITIONAL INFORMATION FOR <i>HYAL2</i> PROJECT	192
7.2	ADDITIONAL INFORMATION FOR HYPERTROPHIC CARDIOMYOPATHY PROJECT	197
7.2.1	THE ANCESTORS OF THE CARRIERS OF THE C.3330+2T>G VARIANT	197
7.2.2	PEDIGREES OF <i>MYBPC3</i> C. 3330+2T>G VARIANT CARRIERS	199
7.2.3	INVESTIGATION OF THE PHENOTYPE OF CARRIERS OF THE <i>MYBPC3</i> C.3330+2T>G VARIANT	208
7.2.4	DIASTOLIC FUNCTION IN <i>MYBPC3</i> C.3330+2T>G VARIANT CARRIERS WITHOUT HYPERTROPHY .	223
7.2.5	EXOME SEQUENCE ANALYSIS OF <i>MYBPC3</i> C.3330+2T>G VARIANT CARRIERS WITH SEVERE PHENOTYPE	228
7.2.6	INFORMATION LEAFLET ON HYPERTROPHIC CARDIOMYOPATHY IN THE ANABAPTIST COMMUNITIES 229	
7.2.7	INFORMATION DAY ON HYPERTROPHIC CARDIOMYOPATHY	231
7.3	CONTRIBUTIONS FROM INDIVIDUALS AND OTHER INSTITUTIONS TO THE PRACTICAL WORK DESCRIBED IN THIS THESIS	232
7.3.1	MUTATIONS IN <i>HYAL2</i> , ENCODING HYALURONIDASE 2, CAUSE A SYNDROME OF COR TRIATRIATUM SINISTER AND OROFACIAL CLEFTING IN HUMANS AND MICE	232
7.3.2	INVESTIGATION OF THE PHENOTYPE AND GENETIC MODIFIERS IN HYPERTROPHIC CARDIOMYOPATHY CAUSED BY THE C.3330+2T>G VARIANT IN MYOSIN BINDING PROTEIN C.....	233
	REFERENCES	235

List of tables

TABLE 1-1: STAGES OF HUMAN DEVELOPMENT WITH CORRESPONDING STAGES OF CARDIAC DEVELOPMENT IN THE HUMAN AND MOUSE.....	24
TABLE 1-2: CHARACTERISTICS OF PLAIN POPULATIONS CONDUCIVE TO GENETIC STUDIES	51
TABLE 3-1: ROMHILT-ESTES POINT SCORE SYSTEM.....	61
TABLE 3-2: TOUCHDOWN PCR PROGRAM	68
TABLE 3-3: REAGENTS FOR A 20 μ L PCR REACTION	69
TABLE 3-4: COMPONENTS OF A 10 μ L SEQUENCING REACTION.....	71
TABLE 3-5: PROGRAM FOR THE SEQUENCING REACTION.....	71
TABLE 3-3-6: COMPONENTS OF A 50ML PCR REACTION	75
TABLE 3-7: RT-PCR PROGRAM	75
TABLE 4-1: CLINICAL CHARACTERISTICS OF INDIVIDUALS HOMOZYGOUS FOR PATHOGENIC <i>HYAL2</i> VARIANTS.....	90
TABLE 5-1: DEMOGRAPHICS AND CLINICAL CHARACTERISTICS OF ALL <i>MYBPC3</i> 3330+2T>G CARRIERS	117
TABLE 5-2: CHARACTERISTICS OF <i>MYBPC3</i> VARIANT CARRIERS WHO DIED SUDDENLY	118
TABLE 5-3: SYMPTOMS OF CARRIERS OF <i>MYBPC3</i> 3330+2T>G VARIANT WITH AND WITHOUT MYECTOMY	121
TABLE 5-4: ECG CHARACTERISTICS (PART 1) OF INDIVIDUALS WITH AND WITHOUT MYECTOMY	122
TABLE 5-5: HOLTER CHARACTERISTICS OF INDIVIDUALS WITH AND WITHOUT MYECTOMY	123
TABLE 5-6: RELATION BETWEEN ECG NORMAL/ ABNORMAL AND MAXIMUM WALL THICKNESS IN ALL CARRIERS OF THE <i>MYBPC3</i> 3330+2T>G VARIANT	125
TABLE 5-7: AGE AT ONSET OF SYMPTOMS OF <i>MYBPC3</i> c.3330+2T>G VARIANT CARRIERS.....	128
TABLE 5-8: COMPARISON OF ECHOCARDIOGRAPHIC CHARACTERISTICS (PART 1) OF MALE AND FEMALE CARRIERS OF <i>MYBPC3</i> 3330+2T.G VARIANT	130
TABLE 5-9: COMPARISON OF DEMOGRAPHIC AND CLINICAL CHARACTERISTICS OF YOUNG (18-40 YEARS) AND OLDER ($\geq$ 41 YEARS) CARRIERS OF <i>MYBPC3</i> c.3330+2T>G VARIANT.....	134
TABLE 5-10: GENOMICS ENGLAND PANELAPPS USED FOR FILTERING OF VARIANTS	157
TABLE 5-11: CLINICAL CHARACTERISTICS OF SEVEN <i>MYBPC3</i> c.3330+2T>G VARIANT CARRIERS WITH SEVERE PHENOTYPE FROM OHIO INCLUDED IN THE EXOME STUDY.....	160
TABLE 5-12: SUMMARY OF POTENTIAL GENETIC MODIFIERS IDENTIFIED IN 10 CARRIERS OF <i>MYBPC3</i> c.3330+2T>G WITH A SEVERE PHENOTYPE.....	165
TABLE 5-13: THE SEVEN ADJACENT SNPs IN GENOME-WIDE SIGNIFICANT GENOMIC REGION ON CHROMOSOME 2q36.1	178
TABLE 7-1: SURVIVAL OF PROGENY OF <i>HYAL2</i> ^{+/+} INTERCROSSES	192
TABLE 7-2: DEMOGRAPHIC AND CLINICAL CHARACTERISTICS OF CARRIERS OF <i>MYBPC3</i> c.3330+2T>G VARIANT WITH AND WITHOUT SURGICAL MYECTOMY.....	208

TABLE 7-3: ECG CHARACTERISTICS (PART 2) OF INDIVIDUALS WITH AND WITHOUT MYECTOMY	209
TABLE 7-4: ECHOCARDIOGRAPHIC PARAMETERS (PART 1) OF INDIVIDUALS WITH AND WITHOUT MYECTOMY (STRUCTURE AND FUNCTION).....	210
TABLE 7-5: ECHOCARDIOGRAPHIC PARAMETERS (PART 2) OF INDIVIDUALS WITH AND WITHOUT MYECTOMY (DIASTOLIC FUNCTION)	211
TABLE 7-6: ECHOCARDIOGRAPHIC PARAMETERS (PART 3) OF INDIVIDUALS WITH AND WITHOUT MYECTOMY	212
TABLE 7-7: COMPARISON OF DEMOGRAPHIC AND CLINICAL CHARACTERISTICS OF MALE AND FEMALE CARRIERS OF <i>MYBPC3</i> C.3330+2T>G VARIANT.....	213
TABLE 7-8: COMPARISON OF SYMPTOMS OF MALE AND FEMALE CARRIERS OF <i>MYBPC3</i> C.3330+2T>G VARIANT.....	214
TABLE 7-9: COMPARISON OF ECG CHARACTERISTICS OF MALE AND FEMALE CARRIERS OF <i>MYBPC3</i> C.3330+2 T>G VARIANT.....	215
TABLE 7-10: COMPARISON OF HOLTER CHARACTERISTICS BETWEEN MALE AND FEMALE CARRIERS OF <i>MYBPC3</i> C.3330+2 T>G VARIANT	216
TABLE 7-11: COMPARISON OF ECHOCARDIOGRAPHIC CHARACTERISTICS (PART 2) OF MALE AND FEMALE CARRIERS OF <i>MYBPC3</i> C.3330+2T.G VARIANT	217
TABLE 7-12: COMPARISON OF ECHOCARDIOGRAPHIC CHARACTERISTICS (PART 3) OF MALE AND FEMALE CARRIERS OF <i>MYBPC3</i> C.3330+2T.G VARIANT	218
TABLE 7-13: COMPARISON OF SYMPTOMS OF YOUNG (18-40 YEARS) AND OLDER (≥ 41 YEARS) CARRIERS OF <i>MYBPC3</i> C.3330+2T>G VARIANT.....	219
TABLE 7-14: COMPARISON OF ECG CHARACTERISTICS OF YOUNG (18-40 YEARS) AND OLDER (≥ 41 YEARS) CARRIERS OF THE <i>MYBPC3</i> C.3330+2T>G VARIANT.....	220
TABLE 7-15: COMPARISON OF HOLTER CHARACTERISTICS OF YOUNG (AGE 18-40) AND OLDER (≥ 41 YEARS) CARRIERS OF THE <i>MYBPC3</i> C.3330+2T>G VARIANT.....	221
TABLE 7-16: COMPARISON OF ECHOCARDIOGRAPHIC PARAMETERS (PART 2) OF YOUNG (AGE 18-40) AND OLDER (≥ 41 YEARS) CARRIERS OF THE <i>MYBPC3</i> C.3330+2T>G VARIANT	222
TABLE 7-17: SUMMARY OF GENES INCLUDED IN VIRTUAL PANEL FOR EXOME SEQUENCE ANALYSIS (DERIVED FROM PANELAPP).....	228

List of figures

FIGURE 1-1: FORMATION OF THE GERM DISC WITH THREE GERM CELL LAYERS	16
FIGURE 1-2: FOLDING OF THE EMBRYO AND FORMATION OF THE HEART TUBE	18
FIGURE 1-3: FORMATION OF THE CARDIAC CHAMBERS.....	20
FIGURE 1-4: FORMATION OF THE VENOUS POLE	22
FIGURE 1-5: GENETIC CAUSES OF CONGENITAL HEART DISEASE	27
FIGURE 1-6: A SCHEMATIC OF GENES IMPLICATED IN HYPERTROPHIC CARDIOMYOPATHY.....	33
FIGURE 1-1-7: PROPORTION OF HCM PATIENTS WITH PATHOGENIC OR LIKELY PATHOGENIC VARIANTS.....	35
FIGURE 1-8: DETERMINANTS OF PHENOTYPE IN HYPERTROPHIC CARDIOMYOPATHY (HCM)	38
FIGURE 1-9: THE 'CHARMING NANCY'	44
FIGURE 1-10: AMISH FARM AND HORSE AND BUGGY.....	45
FIGURE 1-11: AMISH FARMER	46
FIGURE 1-12: TWO AMISH GIRLS IN TRADITIONAL CLOTHING	47
FIGURE 3-1: INFINIUM ASSAY PROTOCOL	73
FIGURE 4-1: TRANSTHORACIC ECHOCARDIOGRAPHY OF TWO AFFECTED AMISH INDIVIDUALS.....	87
FIGURE 4-2: PEDIGREES AND GRAPHICAL REPRESENTATION OF GENOTYPES IN REGION OF SHARED AUTOZYGOSITY ACROSS93	
FIGURE 4-3: PEDIGREES, CLINICAL FEATURES OF INDIVIDUALS HOMOZYGOUS FOR HYAL2 MUTATION AND IDENTIFIED HYAL2 MUTATIONS	94
FIGURE 4-4: EXPRESSION, HA BINDING, AND LOCALIZATION OF WILD TYPE (WT) AND K148R-HYAL2	96
FIGURE 4-5: COMPARATIVE MOLECULAR MODEL OF HUMAN HYAL2 BOUND TO A HEXAMER OF HYALURONIC ACID.....	98
FIGURE 4-6: CARDIAC ABNORMALITIES IN HYAL2 ^{-/-} MICE	100
FIGURE 5-1: HISTOGRAM SHOWING THE AGE DISTRIBUTION OF 313 CARRIERS OF THE <i>MYBPC3</i> 3330+2T>G VARIANT	115
FIGURE 5-2: MAXIMUM WALL THICKNESS IN ALL CARRIERS OF <i>MYBPC3</i> 3330+2T>G VARIANT (INDIVIDUALS WITH PREVIOUS MYECTOMY ARE EXCLUDED)	124
FIGURE 5-3: RELATION BETWEEN MAXIMUM WALL THICKNESS AND ECG NORMAL/ ABNORMAL IN ALL CARRIERS OF THE <i>MYBPC3</i> 3330+2T>G VARIANT	126
FIGURE 5-4: MAXIMAL WALL THICKNESS IN FEMALE AND MALE CARRIERS OF <i>MYBPC3</i> c.3330+2T>G VARIANT	131
FIGURE 5-5: RELATIONSHIP BETWEEN ECG ABNORMALITIES AND MAXIMUM WALL THICKNESS IN MALE AND FEMALE CARRIERS OF <i>MYBPC3</i> c.3330+2T>G VARIANT	132
FIGURE 5-6: AGE RELATED PENETRANCE OF CLINICAL HCM IN c.3330+2T>G <i>MYBPC3</i> CARRIERS BASED ON WALL THICKNESS MEASUREMENT.....	146
FIGURE 5-7: AGE RELATED PENETRANCE OF CLINICAL HCM IN c.3330+2T>G <i>MYBPC3</i> CARRIERS BASED ON ECG CRITERIA	146

FIGURE 5-8: LEVELS OF <i>MYBPC3</i> WILDTYPE TRANSCRIPTS IN BLOOD AND HEART OF CONTROLS	150
FIGURE 5-9: THE DISTRIBUTION OF <i>MYBPC3</i> ISOFORMS IN PERIPHERAL BLOOD RNA FROM CONTROLS AND MUTATION CARRIERS.....	151
FIGURE 5-10: MANHATTAN PLOT FOR THE “DEEP T-WAVE INVERSION” PHENOTYPE	179
FIGURE 5-11: SCHEMATIC SHOWING LOCATION OF THE ~140KB GENOMIC REGION IDENTIFIED BY THE 7 SNPs IN RELATION TO <i>EPHA4</i> AND <i>PAX3</i> GENES	184
FIGURE 7-1 MICRO-CT IMAGES OF THE SKULLS OF <i>HYAL2</i> ^{-/-} AND <i>HYAL2</i> ^{+/-} MICE	193
FIGURE 7-2 SKELETAL STAINING AND MICRO-CT IMAGING OF SKULLS OF <i>HYAL2</i> ^{-/-} AND <i>HYAL2</i> ^{+/-} MICE	194
FIGURE 7-3 ANALYSIS OF THE PALATES OF <i>HYAL2</i> ^{-/-} , <i>HYAL2</i> ^{+/-} , AND <i>HYAL2</i> ^{+/+} MICE	195
FIGURE 7-4 EVALUATION OF HEARING IN <i>HYAL2</i> ^{-/-} AND CONTROL (<i>HYAL2</i> ^{+/+}) MICE.....	196
FIGURE 7-5: MAIN PEDIGREE GEauga COUNTY, OHIO	199
FIGURE 7-6: PEDIGREE 1 HOLMES COUNTY, OHIO	200
FIGURE 7-7: PEDIGREE 2 HOLMES COUNTY, OHIO	201
FIGURE 7-8: PEDIGREE 3 HOLMES COUNTY, OHIO	202
FIGURE 7-9: PEDIGREE 4 HOLMES COUNTY, OHIO	203
FIGURE 7-10: PEDIGREE 5 HOLMES COUNTY, OHIO	204
FIGURE 7-11: PEDIGREE 6 HOLMES COUNTY, OHIO	205
FIGURE 7-12: PEDIGREE 7 HOLMES COUNTY, OHIO	206
FIGURE 7-13: PEDIGREE 8 HOLMES COUNTY, OHIO	207
FIGURE 7-14: MV INFLOW E/A RATIO IN CARRIERS OF THE <i>MYBPC3</i> c.3330+2T>G VARIANT WITHOUT HYPERTROPHY	223
FIGURE 7-15: PULMONARY VENOUS DOPPLER S/D RATIO IN CARRIERS OF THE <i>MYBPC3</i> c.3330+2T>G VARIANT WITHOUT HYPERTROPHY.....	224
FIGURE 7-16: PULMONARY VENOUS DOPPLER A IN CARRIERS OF THE <i>MYBPC3</i> c.3330+2T>G VARIANT WITHOUT HYPERTROPHY	225
FIGURE 7-17: LATERAL MITRAL ANNULUS TISSUE DOPPLER E’/A’ IN CARRIERS OF THE <i>MYBPC3</i> c.3330+2T>G VARIANT WITHOUT HYPERTROPHY.....	226
FIGURE 7-18: E/E’ LATERAL MITRAL ANNULUS IN CARRIERS OF THE <i>MYBPC3</i> c.3330+2T>G VARIANT WITHOUT HYPERTROPHY	227

1 Introduction

1.1 Cardiac embryological development

1.1.1 Formation of the linear heart tube

The heart is vital in supporting the circulatory system from invertebrates such as worms to larger vertebrates such as man. It is therefore not surprising that it is the first functioning organ during human development and starts beating around day 21 after fertilization. The heart is mainly a derivative of the mesoderm like all other cardiovascular structures, with some contributions from the ectoderm-derived cardiac neural crest.

Following blastocyte implantation in the uterus, the three germ cell layers (ectoderm, mesoderm and endoderm) develop forming the flat tri-laminar germ disc (Fig 1-1).

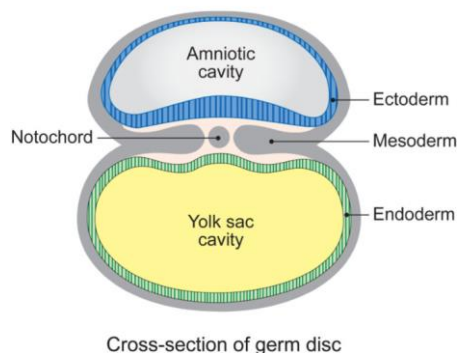


Figure 1-1: Formation of the germ disc with three germ cell layers

Adapted from (Kloesel, DiNardo et al. 2016)

The cardiac precursor cells lie in two symmetrical portions of the mesoderm, lateral to the stomato-pharyngeal membrane (SM), the future mouth. The bilateral groups of precursors converge in the midline, cranial to the SM, establishing the cardiac

crescent (figure 1-2 A, B) (Moorman, Christoffels et al. 2007, Sizarov, Ya et al. 2011). The differentiation of the cardiac crescent into cardiomyocytes is regulated by a complex system of signals from the adjacent endoderm (Harvey 2002). Between days 19-21 of human development, the flat embryo folds. The heart ends up ventrally to the foregut, which runs from the SM to the cloacal membrane (CM), caudally to the head and cranially to the umbilical cord and transverse septum, which contributes to the formation of the diaphragm.

The cardiac crescent develops into a tube with two inlets on the caudal side (venous pole), and one outlet on the cranial side (arterial pole). The peripheral part of the crescent forms the venous pole and ends up adjacent to the transverse septum. The central portion of the crescent forming the outflow tract is contiguous with the pharyngeal mesenchyme (Lescroart, Kelly et al. 2010). The close proximity of the cardiac and facial region during embryogenesis may explain the high incidence of combined cardiac and facial malformations.

The cells that contribute to the formation of the initial linear heart tube are referred to as the first heart field, whereas cells that are added to the heart tube at a later stage during development are denoted the second heart field (Buckingham, Meilhac et al. 2005, Dyer and Kirby 2009, Rochais, Mesbah et al. 2009). In the cardiac crescent, the second heart field is situated medially to the first heart field (figure 1-2 B). Following the folding of the embryo and the creation of the heart tube, the second heart field is positioned in the splanchnic mesoderm of the dorsal pericardial wall. All subsequent growth of the heart tube is mediated by addition of newly differentiated cardiomyocytes from the second heart field. Significantly, the identity of cardiomyocytes changes dependent on their position in the developing heart.

Simultaneously with the myocardium, the endocardium develops from a different set of precursor cells in the splanchnic mesoderm. An extracellular matrix, called cardiac jelly, lies between the two layers.

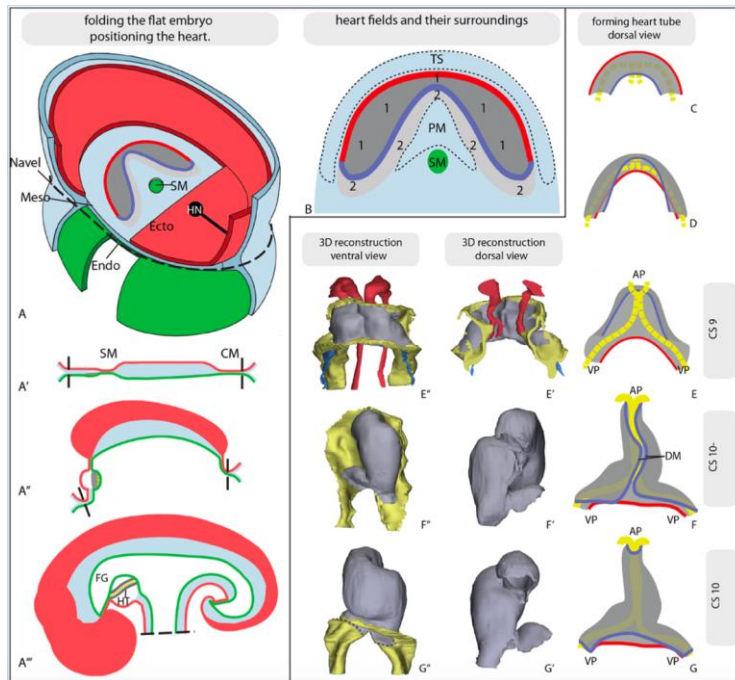


Figure 1-2: Folding of the embryo and formation of the heart tube

Adapted from (Sylva, van den Hoff et al. 2014)

Panel A: The embryo starts as a flat disc with the three germ layers ectoderm (Ecto), mesoderm (Meso) and endoderm (Endo); the mesoderm adjacent to the endoderm, between the future transverse septum (TS) and the future pharyngeal mesoderm (PM) will give rise to the heart

Panels A-A''': the embryonic gut running from the stomato-pharyngeal membrane (SM) to the cloacal membrane (CM) is formed during folding of the embryo; the heart (HT) becomes positioned ventrally to the foregut (FG), caudally to the head and cranially to the umbilical cord and transverse septum

Panel B: The first heart field (1) and the second heart field (2)

Panels C-G: formation of the heart tube from a flat horseshoe-shaped crescent to a tube by merging at the backside of the heart; the connection of the heart to the dorsal body wall is called the dorsal mesocardium (DM); the heart tube (grey) develops by addition of cells from the bordering splanchnic mesoderm (yellow); red line: peripheral border; blue line: central border

1.1.2 Development of the four heart chambers

Dependent on their location within the heart, cardiomyocytes exhibit different profiles of automaticity, conduction velocity and contractility. T-box transcription factors play a vital role in determining cardiomyocyte identity and the individual members of this family of transcription factors are expressed in different parts of the developing heart.

A process called looping starts between day 22-23 and results in an S-shaped heart with inner and outer curvatures. Re-initiation of proliferation and differentiation in localized areas of the primary heart tube follow, facilitating the ballooning of the future heart chambers. At the outer curvature, the future ventricles balloon and at the dorso-lateral venous pole the atrial appendices start to develop. They will ultimately extend cranially like in the formed heart (figure 1-3 K). The future chambers of the heart, ballooning from the primary heart tube (primary myocardium), are also referred to as secondary or chamber myocardium. The chamber myocardial cells start to express different genes including atrial natriuretic factor (*ANF/NPPA*) (figure 1-3 A-D) and the gap-junction protein connexin40 (*Cx40/GJA5*) facilitating higher conduction velocities.

The atrial chamber myocardium is bordered by the primary myocardium of the sinus venosus (SV) and atrioventricular canal (AVC). The SV myocardium separates the systemic veins from the atria, whereas the AVC myocardium separates the atria from the ventricles. Of note is, that the atrial chamber myocardium developed at this stage will merely constitute the trabeculated atrial appendices in the formed heart. The remaining smooth-wall myocardium in the atria is added later in the embryological development.

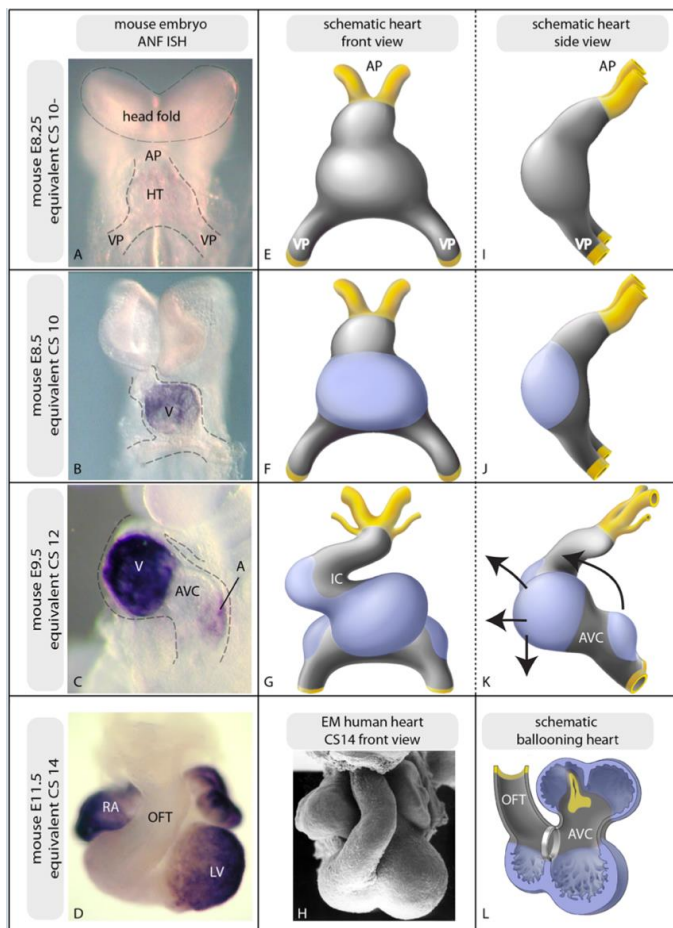


Figure 1-3: Formation of the cardiac chambers

Adapted from (Sylva, van den Hoff et al. 2014)

Panels A-D: developmental series of mouse embryos; whole mount RNA in situ hybridization for the embryonic chamber marker atrial natriuretic factor (ANF) is used as a marker for differentiation into chamber myocardium

Panels A, E, I: The heart tube (HT) consists from venous to arterial pole (VP, AP) of primary myocardium (grey)

Panels B, F, J: The first chamber to start ballooning is the ventricle (V) at the outer curvature of the heart tube

Panels C, G, K: The heart tube has started to loop and acquired an S-shape; the atria (A) start to balloon; the outflow tract (OFT), inner curvature (IC), and atrioventricular canal remain primary myocardium

Panel H: Electron microscopic photograph of a human heart (day 31-35) demonstrating the similarity with the mouse heart (E11.5) and the schematic shown in panel L

The systemic veins are connected to the venous pole of the heart. The left and right superior and inferior cardinal veins combine into the common cardinal veins, which drain into the sinus venosus or left and right sinus horns. Both sinus horns initially connect to the right atrium. The right sinus horn eventually becomes incorporated into the dorsal part of the formed right atrium and the left sinus horn becomes the coronary sinus. Figure 1-4 illustrates further details about the development of the venous pole.

Development of the pulmonary veins differs in that the initially single pulmonary vein forms and connects to the heart after the formation of the heart tube and the start of chamber formation. The primary atrial septum initially develops at the right atrial side resulting in the pulmonary vein being incorporated into the left atrium.

1.1.3 Septation and valve formation

Septation occurs in four different parts of the developing heart- the atria, the ventricles, the primary myocardium of the atrioventricular canal and primary foramen and the outflow tract.

Atrial septation occurs in stages during embryonic development allowing right-left atrial shunting. Initially the primary septum forms, starting out from the cranio-dorsal wall of the atrium at the right side of the midline towards the atrioventricular canal. The primary atrial foramen (ostium primum) permits right-left shunting at this stage. Whilst the ostium primum at the caudal end of the septum primum is closing, small holes develop in the cranial septum primum mediated by apoptosis and form the secondary foramen (ostium secundum). Subsequently the secondary septum

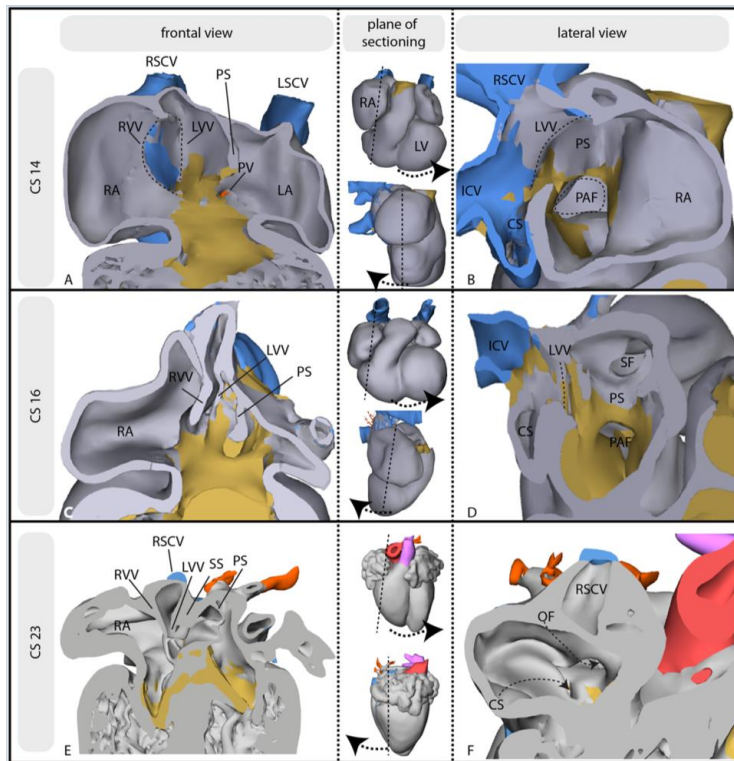


Figure 1-4: Formation of the venous pole

Adapted from (Sylva, van den Hoff et al. 2014)

Panels A-F: Cross-section of 3D reconstruction of human heart; myocardium is in grey, systemic veins in orange and cardiac mesenchyme in yellow

Panels A, B: Developing atria; the entire systemic venous pole consisting of the right superior caval vein, left superior caval vein and the inferior caval vein (RSCV, LSCV, ICV) is connected to the right atrium via the coronary sinus (CS) at this stage; the primary septum (PS) is growing to the right of the pulmonary vein (PV) orifice

Panels C, D: the RSCV and CS are still connected through a single orifice to the atrium; the secondary foramen (SF) appears in the PS

Panels E, F: all veins are connected separately to the atrium now; the primary atrial foramen (PAF) is closed and the secondary septum (SS) is growing between the left venous valve (LVV) and the primary septum (PS); the right venous valves (RVV) are now separated in a part flanking the right superior caval vein (Eustachian valve) and coronary sinus (Thebesian valve); OF: oval foramen

grows at the right side of the primary septum and covers the secondary foramen. Following birth when the blood pressure in the left atrium rises, the primary and secondary septum are pressed together, and the secondary foramen closes completely. The oval fossa is the part of the primary septum that is not covered by the secondary septum.

Ventricular septation starts from the cells in between the areas that balloon from the primary heart tube and eventually form the two ventricles. The septum grows by cells being added to the future interventricular septum mainly from the left ventricular free wall.

During embryonic development, the primary foramen provides a connection between the right atrium and ventricle in diastole and the left ventricle and aorta in systole. A separation into a left and right part by the membranous septum occurs between days 44 and 48 of human development. The atrioventricular canal and outflow tract are divided by endocardial cushions which are acellular structures filled with cardiac jelly produced by the myocardium.

Unidirectional blood flow in the primary heart tube is achieved by complete obliteration of the cardiac lumen during peristaltic contractions. In the formed heart, unidirectional blood flow is ensured by the atrioventricular and semilunar valves. Endocardial cushions and ridges made up of cardiac jelly and populated by endocardial, epicardial and neural crest cells form the basis for valve formation which is completed by a maturation process forming different layers of different types of extracellular matrix.

The heart is fully formed by day 60 of human and day 17 of mouse development post conception (table 1-1) (Sylva, van den Hoff et al. 2014, Kloesel, DiNardo et al. 2016).

Carnegie stage	Human DPC	Mouse DPC	
CS8	17–19	7	The cardiac crescent forms
CS9	19–21	7.5	The embryo folds, the pericardiac cavity is placed in its final position, gully of myocardium forms, the endocardial plexus forms, cardiac jelly forms
CS10	22–23	8	The heart beats, the endocardial tubes fuse, the mesocardium perforates, looping starts, the ventricle starts ballooning
CS11	23–26	8.5	The atria balloon, the pro-epicardium forms
CS12	26–30	9.5	The septum primum appears, the right venous valve appears, the muscular part of the ventricular septum forms, cells appear in the cardiac jelly, the epicardial growth starts
CS13	28–32	10.5	The atrioventricular-cushions form, the pulmonary vein attaches to the atrium, the left venous valve appears, epicardial mesenchyme appears first in the atrioventricular sulcus
CS14	31–35	11.5	The atrioventricular-cushions approach one another, the outflow ridges become apparent, capillaries form in the epicardial mesenchyme
CS15	35–38	12	The atrioventricular cushions oppose one another, the secondary foramen forms, the distal outflow tract septates the outflow tract ridges reach the primary foramen
CS16	37–42	12.5	The primary atrial septum closes, the outflow tract ridges approach the interventricular septum. The entire heart is covered in epicardium
CS17	42–44	13.5	Secondary atrial septum appears, the sinus node becomes discernable, the left and right atrioventricular connection becomes separate, the proximal outflow tract becomes septated, the semilunar valves develop
CS18	44–48	14.5	Papillary muscles appear, the atrioventricular valves start to form
CS19	48–51	15	The left venous valve fuses with the secondary septum, the mural leaflets of the mitral and tricuspid valve are released
CS21	53–54	16	The main branches of the coronary artery become apparent
CS22	54–56	16.5	The chorda tendinae form
CS23	56–60	17.5	The septal leaflet of the tricuspid valve delaminates

Table 1-1: Stages of human development with corresponding stages of cardiac development in the human and mouse

Adapted from (Sylva, van den Hoff et al. 2014)

1.2 Congenital heart disease

1.2.1 Epidemiology

Congenital heart disease (CHD) represents a structural abnormality of the heart and great vessels that is present at birth (Triedman and Newburger 2016). Being the most common birth defect, it affects ca. 1% of all live-born children (van der Linde, Konings et al. 2011). About a third of affected infants require surgery within the first year of life (Leirgul, Fomina et al. 2014). Enormous advancements of medical and surgical treatments have been made since the first systemic to pulmonary artery shunt procedure was performed by Alfred Blalock, with the assistance of Helen Taussig and Vivien Thomas, in 1944 (Blalock 1945). Survival of patients undergoing CHD surgery has continuously improved and the number of adults living with CHD now exceeds the number of children (Marelli, Ionescu-Iltu et al. 2014). However, CHD still remains the leading cause of mortality from birth defects in the developed world (Gelb 2015).

The categories of CHD are based on a combination of anatomical and physiological phenotypes resulting from a perturbation of the normal embryological cardiac development. The main types are conotruncal defects affecting the ventricular septum and outflow tract, defects leading to obstruction of the left ventricular outflow, disorders resulting from abnormal left-right relationships within the heart (heterotaxy), defects affecting the inflow like mitral and tricuspid valve abnormalities in atrioventricular canal defects and a number of other lesions including isolated atrial and ventricular septal defects. About 14% of affected individuals have extra-cardiac co-morbidities including neurodevelopmental disabilities that may additionally affect their quality of life (Zaidi and Brueckner 2017).

1.2.2 Genetics of congenital heart disease

The causes of CHD remain poorly understood and are thought to be a combination of genetic and environmental factors. A number of observations imply a major contribution of genetics including the relatively stable incidence of CHD of ca. 1%

over time and in different populations. There are variations in the types of CHD observed in different populations, for example an increased incidence of left ventricular outflow tract lesions in white and increased right ventricular outflow obstruction in Chinese children, indicative of population-specific genetic contributions (Jacobs, Leung et al. 2000, van der Linde, Konings et al. 2011). In addition, there is evidence for greater concordance of CHD in monozygotic compared to dizygotic twins (Nora, Dodd et al. 1969, Wang, Li et al. 2014), and the recurrence risk ratio of CHD among siblings is elevated compared to the general population prevalence, ranging from 3.4 for atrial septal defects to 79.1 for heterotaxy in the Danish national cohort study (Oyen, Poulsen et al. 2009). Rare Mendelian forms of CHD have been described including families with atrial septal defects (Schott, Benson et al. 1998), heterotaxy (Gebbia, Ferrero et al. 1997), mitral valve prolapse (Durst, Sauls et al. 2015) and bicuspid aortic valves (Garg, Muth et al. 2005) and there is an increased incidence of CHD in consanguineous populations (Shieh, Bittles et al. 2012). However, the majority of cases of CHD are sporadic and occur in families with no prior history of CHD, indicating that de-novo mutations, compound heterozygosity, digenic or polygenic inheritance or the interaction of genetic and environmental triggers are important factors. As severe CHD often impairs reproductive fitness, there are rarely large pedigrees with several affected members described. Recent studies using next-generation sequencing suggest that de-novo mutations contribute to 10% of CHD and at least 20% of syndromic cases of CHD with associated extra-cardiac and neurodevelopmental abnormalities (Zaidi, Choi et al. 2013, Homsy, Zaidi et al. 2015).

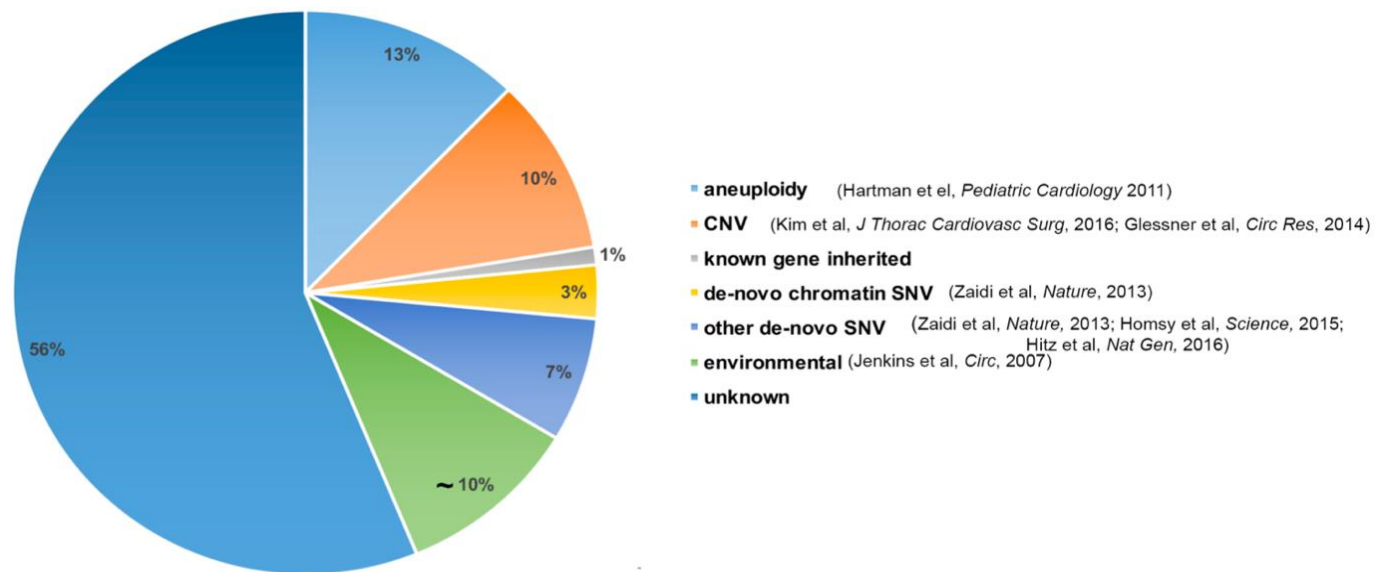


Figure 1-5: Genetic causes of congenital heart disease

adapted from (Zaidi and Brueckner 2017)

CNV: Copy number variations; SNV: Single nucleotide variant

The earliest established causes of CHD were aneuploidies. About 9-18% of cases of CHD are associated with cytogenetic abnormalities (Hartman, Rasmussen et al. 2011). Given the large number of genes affected, defects are often severe and 98% of children with CHD and cytogenetic abnormalities will have extra-cardiac manifestations. 35-50% of individuals with Trisomy 21, 60-80% of individuals with trisomy 13 and trisomy 18 and 33% with monosomy X have CHD (Wimalasundera and Gardiner 2004). A diverse spectrum of CHD is associated with the different aneuploidies. Given the large number of genes affected by cytogenetic abnormalities, it has only been possible to identify the likely embryological mechanism in some of the aneuploidies, for example it has been suggested that *DSCAM* and *COL6A* contribute to Trisomy 21-associated CHD (Korbel, Tirosh-Wagner et al. 2009).

Another well recognized cause for a number of clinical syndromes that include CHD, are copy number variations. The most common human micro-deletion, Del22q11, results in DiGeorge or velocardio-facial syndrome. Clinical signs include CHD, palate abnormalities, characteristic facial features and neurodevelopmental abnormalities. Del22q11 includes the T-Box transcription factor *TBX1*, and there is evidence that haploinsufficiency of *TBX1* (Jerome and Papaioannou 2001) causes the cardio-pharyngeal phenotype. In addition to specific syndromes, studies have also suggested that CNVs contribute to at least 10-15% of non-syndromic forms of CHD (Kim, Kim et al. 2016).

A number of rare Mendelian inherited disorders have been identified by linkage analysis, positional cloning and gene (dideoxy and exome/genome) sequencing of CHD candidate genes. The main genes implicated in CHD encode cardiac transcription factors, signalling molecules and cellular structural components. Pathogenic variants in the cardiac transcription factor *NKX2.5* were amongst the first to be shown to cause human CHD in large pedigrees with isolated atrial septal defects (ASDs) and ASDs and conduction abnormalities (Benson, Silberbach et al. 1999). Variants in *NKX2.5* have subsequently been associated with a wide spectrum

of CHD phenotypes including heterotaxy and Tetralogy of Fallot (McElhinney, Geiger et al. 2003). The causes for the significant phenotypic variation are unknown and speculated to be due to differences in genetic background, interactions between genotype and environmental factors that may include minor variations in the haemodynamics in utero during critical times of cardiac development. Other transcription factors implicated in CHD are *GATA4*, a zinc-finger transcription factor essential for cardiogenesis that associates with *NKX2.5* (Garg, Kathiriya et al. 2003), and *TBX5*, a T-box protein. Variants in *TBX5* were found in families with Holt-Oram syndrome, which is characterized by cardiac defects and upper limb malformations (Basson, Bachinsky et al. 1997). *TBX5* is expressed in the developing heart and forelimb buds and, like other T-box proteins, regulates cell fate and developmental processes.

A number of genes in the Notch signalling pathway have been implicated in various types of CHD. Notch signalling mediates intercellular communication during crucial processes in cardiac embryogenesis including formation of the left-right organizer, blood vessel development and ventricular chamber development (Siekman and Lawson 2007, Boskovski, Yuan et al. 2013, Han, Bloomekatz et al. 2016). The first gene from the Notch pathway described in families with heterogeneous phenotypes ranging from bi-leaflet aortic valves to hypoplastic left heart syndrome, was *NOTCH1* (Garg, Muth et al. 2005).

There are fewer studies on recessive inheritance in CHD. There is evidence for an increased incidence of CHD in consanguineous populations, especially heterotaxy and the associated complex cardiac malformations (Gatrad, Read et al. 1984). Genes implicated in recessively inherited heterotaxy include the cytoskeletal protein *SHROOM3*, the cilia-associated *WDR16*, matrix metalloproteinase 21 (*MMP21*) and nephronophthisis 4 (*NPHP4*) (Tariq, Belmont et al. 2011, French, van de Laar et al. 2012, Perles, Moon et al. 2015, Ta-Shma, Perles et al. 2015).

Despite enormous advances, the genetic basis of over 50% of CHD is unknown (figure 1-5). The significant genetic heterogeneity and limited genotype-phenotype

correlations pose challenges to the comprehensive analysis of the genetics of CHD. Some of the unexplained cases of CHD may be due to mutations in non-coding DNA, somatic mutations, complex inheritance including polygenic disease and modifier genes, and gene-environment interactions. Investigation of the genetic basis of CHD to date has identified critical physiological pathways implicated in CHD including chromatin remodelling, Notch signalling, cilia function, sarcomere structure and function, and RAS signalling and significantly contributed to our understanding of the embryological development of the heart. An increased appreciation of the genetics underlying CHD promises a more tailored approach to treatment and risk stratification for patients with CHD in the future (Zaidi and Brueckner 2017).

1.3 Hypertrophic cardiomyopathy

1.3.1 Epidemiology, morphology and clinical presentation

Hypertrophic cardiomyopathy (HCM) is the most common inherited cardiovascular disease worldwide with cases reported in more than 50 countries on all continents, affecting people of different ethnical and racial backgrounds and both sexes (Maron 2004). The estimated prevalence is 1:500 young adults, which translates to about 120,000 affected people in the UK and 600,000 in the USA. However, the true prevalence is likely to be higher as this estimate is based on detection of a wall thickness of >15 mm in a relatively young population and penetrance is known to be age dependent (Maron, Epstein et al. 1986, Maron, Gardin et al. 1995). Additionally many patients may only show mild hypertrophy, or develop hypertrophy in later life (Niimura, Patton et al. 2002). For example, HCM was detected in 0.29% (1:333) of 60-year olds being screened with echocardiography for cardiovascular evaluation in the United States (Maron, Mathenge et al. 1999). This estimated high occurrence in the general population contrasts with a far lower rate of diagnosis in clinical practice (Maron, McKenna et al. 2003) implying that many individuals remain undiagnosed throughout life.

Morphologically, HCM is characterised by hypertrophy of the left (and sometimes the right) ventricle which is often asymmetric without evidence of any other cardiac or systemic disease (e.g. systemic hypertension) that could produce the extent of hypertrophy present (Maron, McKenna et al. 2003). The left ventricle is generally non-dilated with preserved or enhanced systolic function except in some individuals where the disease progresses to an end-stage with left ventricular dilatation and systolic dysfunction associated with symptoms and signs of heart failure. This end-stage or “burnt-out” phase of HCM occurs in about 3% of patients (Melacini, Basso et al. 2010).

Dynamic subaortic gradients are a frequent feature of HCM which led to the historic nomenclature of idiopathic hypertrophic subaortic stenosis or hypertrophic obstructive cardiomyopathy before the recognition that left-ventricular outflow tract obstruction is not invariably present in individuals with HCM. One study reported that 70% of a hospital-based cohort had outflow tract obstruction at rest or with exercise, but the remaining 30% had a non-obstructive form (Maron, Olivotto et al. 2006). The proportion of individuals with non-obstructive HCM has no doubt increased over time, as earlier diagnosis of affected individuals is facilitated by genetic cascade screening and improved imaging technology especially cardiac MRI scanning.

Presentation of HCM can occur at any time between infancy to old age which is rather unusual among cardiovascular diseases. Many individuals with HCM achieve normal life expectancy with little or no disability. In others, severe complications due to heart failure, arrhythmias, embolic stroke or sudden cardiac death ensue. Common presentations in symptomatic individuals are chest pain and dyspnoea secondary to myocardial hypoperfusion, increased oxygen demand and diastolic dysfunction particularly during exercise. Palpitations are common and may be associated with light-headedness or dizziness and occasionally with syncope. Even though syncope is an infrequent symptom, it may herald sudden cardiac death (SCD) (Kofflard, Ten Cate et al. 2003). SCD secondary to ventricular tachyarrhythmia may

also be the first manifestation of HCM particularly in young and competitive athletes amongst whom it presents the most frequent cause of SCD (Maron, Doerer et al. 2009). Supraventricular arrhythmias including atrial fibrillation, which develops in about 20% of patients with HCM are also common, may lead to embolic strokes (prevalence 6%, incidence 0.8% per year) (Maron, Olivotto et al. 2002) and are often associated with adverse clinical outcome (Robinson, Frenneaux et al. 1990). Overall mortality rates are reported to be ca. 1% per year in unselected cohorts of patients with HCM (Maron, McKenna et al. 2003).

Pathological features of HCM comprise myocyte hypertrophy and disarray, interstitial fibrosis, hypertrophy and hyperplasia of the media of intramural coronary arteries and mitral valve leaflet anomalies. The degree of cardiac hypertrophy and interstitial fibrosis are major determinants of mortality and morbidity in HCM (Spirito, Bellone et al. 2000).

1.3.2 Genetic basis of hypertrophic cardiomyopathy

HCM is familial in 50-70% of cases (Marian 2010). The mode of inheritance is usually autosomal dominant with variable expression and age-related penetrance. Children of affected individuals therefore typically have a 50% probability (dependent on the specific genetic cause) of inheriting the causative mutation. Additionally, sporadic cases may arise due to *de novo* mutations in the proband, which are absent in the parents.

For the first 25 years after the initial report of HCM by Prof RD Teare in 1958, diagnosis of HCM could only be made by the integration of clinical examination, electrocardiogram and invasive angiographic/haemodynamic studies. With the increasing availability of echocardiography in the early 1970s, non-invasive visualization of the left ventricle became possible simplifying the investigation of individuals at risk and diagnosing HCM. In 1989 the advancement of molecular methodology allowed the mapping of the first causative locus for HCM to chromosome 14q1 (Jarcho, McKenna et al. 1989). Subsequently, sequence analysis

of a candidate gene identified the first missense mutation in the sarcomeric gene beta-myosin heavy chain (*MYH7*; p.Arg403Gln) responsible for HCM in 1990 (Geisterfer-Lowrance, Kass et al. 1990). Five years later the first mutation in another sarcomeric gene, *MYBPC3*, encoding myosin-binding protein C, was identified. Pathogenic variants in six other sarcomeric genes were discovered (*TNNT2*, *TPM1*, *MYL2*, *MYL3*, *TNNI3*, *ACTC1*) through linkage studies in multiple, large, affected pedigrees. The segregation of genetic variant with disease phenotype provided definitive evidence of disease association. Since then many 1000s of pathogenic variants are recognised in these eight core-sarcomeric genes, and the importance of sarcomeric gene mutation in the pathogenesis of HCM is indisputable.

More recent advances in DNA sequencing technologies have facilitated intense research in the genetic aetiology of HCM and over the last 15 years almost 250 variants in over 40 additional, mainly non-sarcomeric genes have been implicated in HCM. These have mainly been discovered with a candidate gene approach investigating genes with a hypothetical role in HCM, for example genes encoding the Z-disc and calcium handling proteins (figure 1-6).

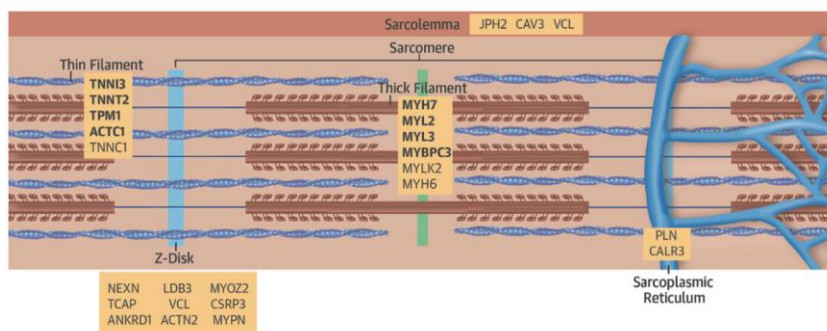


Figure 1-6: A schematic of genes implicated in hypertrophic cardiomyopathy

Adapted from (Burke, Cook et al. 2016)

Definitive (bold) and posited HCM genes with the subcellular localization of the encoded protein

The validity and relevance of some of these additional putative HCM genes has recently been questioned, as variant pathogenicity was often inferred based only on small family cohorts, and rarely supported by functional studies. Large gene sequencing projects (eg 1000 Genomes, the Exome Sequencing Project and the Exome Aggregation Consortium [ExAC]) have highlighted, that rare protein-altering variants are more common in the general population than previously assumed. Some variants were considered 'too common' at the population level to be causative of rare Mendelian disorders and, for particular genes, individually rare variants are collectively quite common. A recent study compared the level of rare variation between ExAC reference samples and a large cardiomyopathy cohort of over 6000 cases. The study found a significant excess of rare variants in the eight main sarcomere genes in HCM cases, but very minor or no excess burden of rare variants in the limited number of non-sarcomeric genes assessed in the clinical laboratories. (Walsh, Buchan et al. 2017). In fact, 99% of variants classified as pathogenic or likely pathogenic discovered by clinical genetic testing were found in the eight core-sarcomeric genes (figure 1-7).

To further explore the role of genes other than the major sarcomeric genes, a comprehensive panel of genes including 31 non-core sarcomeric and non-syndromic putative HCM genes in over 800 unrelated individuals with HCM were investigated and the prevalence of rare variants to reported frequencies in the ExAC database was compared. A significant excess of rare variants in the cases was only detected in three genes that had additional supporting human genetic evidence, *CSRP3* (Z-disc), *FHL1* (desmosome), *PLN* (calcium signalling), though the excess was less than 1% for each gene. The authors conceded, that for some of the other putative HCM genes, like the Z-disc genes *ACTN2* and *MYOZ2*, which previously co-segregated in relatively large HCM pedigrees, certain variants may be disease-causing. However, these findings cast doubt on the role of a large number of putative HCM genes previously classified as causative.

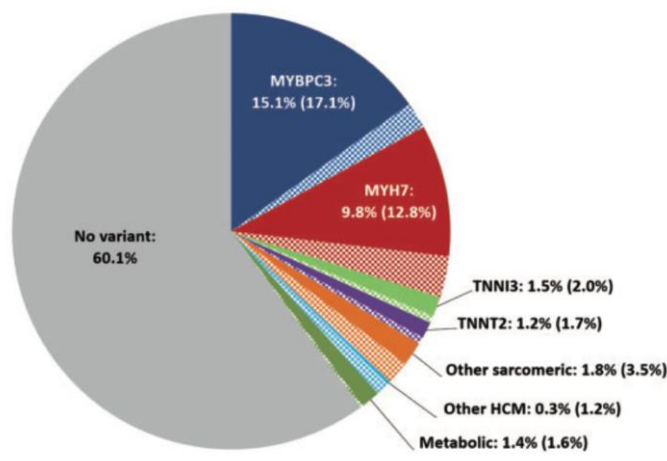


Figure 1-1-7: Proportion of HCM patients with pathogenic or likely pathogenic variants

Oxford Molecular Genetics Laboratory and Partners Laboratory for Molecular Medicine

Adapted from (Walsh, Buchan et al. 2017)

Percentages in parentheses refer to the case excess over the frequency of rare variation in ExAC (variants with unknown significance in check-patterned slices)

Other sarcomeric: ACTC1, MYL2, MYL3, TPM1

Other HCM: CSRP3, FHL1, PLN

Metabolic cardiomyopathies: GLA, LAMP2, PRKAG2

The vast majority of pathogenic mutations in HCM (90%) are missense variants with the exception of the mutations in *MYBPC3*, which mainly consist of insertions/deletions and premature truncation variants due to frameshift and splice-site mutations. Missense variants change the amino acid composition of the encoding protein and may alter protein structure and function, whereas truncated transcripts are often targeted by nonsense-mediated decay and truncated proteins degraded by the ubiquitin-proteasome system, leading to haploinsufficiency.

Only a small number of the causal mutations recur in different families with about 65% of probands carrying so-called “private mutations” (Maron, Maron et al. 2012) (Marian 2010).

1.3.3 Variable phenotypic expression and genetic modifiers

The phenotypic expression of HCM shows remarkable variability regarding the severity of cardiac hypertrophy and the risk of sudden cardiac death. This is not only true for carriers of different pathogenic variants, but also for members of a single family who share the same mutation. Initial phenotype-genotype studies suggested certain genotype specific features in disease expression, however subsequent studies contradicted or failed to replicate these findings (Van Driest, Vasile et al. 2004, Landstrom and Ackerman 2010). Often *MYH7* mutations seemed to be associated with an early onset and significant cardiac hypertrophy (Van Driest, Jaeger et al. 2004). In contrast, *MYBPC3* mutations have often been described as causing less severe hypertrophy which may develop later in life (Niimura, Bachinski et al. 1998). However, a recent investigation of a large cohort of carriers of *MYBPC3* mutations revealed significant phenotypic variability rendering previous suggestions that *MYBPC3* HCM is typically a late-onset and more benign form of HCM apparently premature (Page, Kounas et al. 2012).

One reliable conclusion that can be drawn from a positive HCM genetic test result in regards to clinical expression, seems to be the fact that individuals with a positive genetic test per se, regardless of the type of gene or mutation, tend to have a more severe phenotype than individuals with a negative genetic test (Van Driest, Ommen et al. 2005). In addition, gene dosage seems to influence the severity of disease. 3% to 5% of affected individuals are compound heterozygotes for two mutations in the same gene present on different alleles, or double heterozygotes for mutations in two distinct genes both implicated in HCM. Further, rare cases of triple mutations involving three genes and also homozygosity for the same mutation in a single gene have been described. These individuals tend to have more severe disease (Maron, Maron et al. 2012).

At present the exact mechanisms that lead to the diverse phenotypes of HCM remain incompletely understood, a fact that was highlighted in a report of the National Heart, Lung and Blood Institute examining research priorities in HCM (Force, Bonow et al. 2010). Genetic variants affecting sarcomeric protein structure or expression level lead to molecular and cellular dysfunction with activation of various signalling pathways. The resulting phenotype is the consequence of interactions between different intermediary molecular pathways and environmental influences. Cardiac hypertrophy and other morphological features of HCM are thought to be compensatory and likely to be modulated by a large number of factors (Marian 2002). In addition, current disease-causing variants do not completely explain certain morphological features such as mitral valve leaflet enlargement, microvascular abnormalities and segmental left-ventricular hypertrophy. Also, there is significant variability in phenotypic expression of hypertrophy and occurrence of sudden death between unrelated individuals or even among family members sharing the same mutation. Furthermore, significant variability in age-dependent penetrance exists and acceleration of development of hypertrophy during puberty is a typical feature. Altogether, these observations suggest the involvement of a number of factors distinct from the causal mutation in the evolution of the phenotype including environment, growth factors, hormones, epigenetic factors, post-translational modification of proteins and modifier genes (figure 1-8).

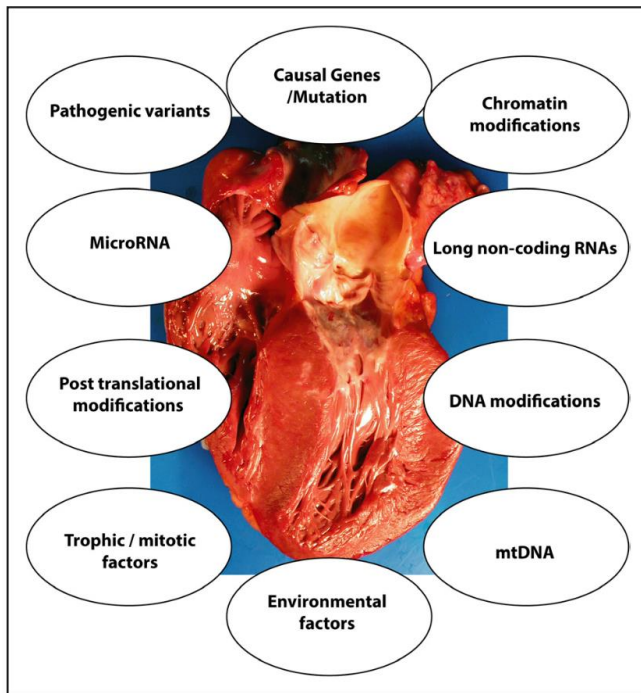


Figure 1-8: Determinants of phenotype in hypertrophic cardiomyopathy (HCM)

The causal mutation exerts the main effect and several others, like genetic variants (modifiers), genomics (such as noncoding RNAs), proteomics (such as post-translational modifications), and environmental factors (such as isometric exercises) contribute to the expression of the phenotype

Adapted from (Marian and Braunwald 2017)

Genetic modifiers may be defined as gene variants other than the causative mutation that affect the phenotypic expression of a genetic disorder. They are neither necessary nor sufficient to cause the disease, but for example in HCM, may influence the severity of cardiac hypertrophy and the risk of SCD. The observation of phenotypic variability between family members carrying the same HCM mutation, and also between individuals from different families with identical causal mutations, highlights the likely importance of modifier genes. This inter-individual variability of

cardiac hypertrophy is certainly partly due to the considerable differences in the genomes of individuals with the most commonly recognized being single nucleotide polymorphisms (SNPs). As elucidated by recent advances in whole exome and genome sequencing, an enormous degree of variation exists between genomes of different individuals, which very likely effects the expression of the phenotype of HCM as many other phenotype and disease traits. Each human nuclear genome contains on average 11 000 non-synonymous variants, 160 premature truncation variants and 500 000 variants in the known regulatory regions (Genomes Project, Auton et al. 2015). Given the diversity of the human genome, the composition of modifier variants likely also differs between individuals partly explaining the variability in phenotypic expression.

Modifier genes have not been systematically investigated in HCM with one of the most limiting factors being the genetic heterogeneity and the relatively small number of affected individuals carrying the same causative mutation. The only genome-wide study so far included 100 members of an HCM family including 36 with the InsG791 mutation in *MYBPC3* (Daw, Chen et al. 2007). Four loci were identified that affected expression of cardiac hypertrophy. The effect size varied from an 8g shift in left ventricular mass for 10p13 locus heterozygosity for the common allele to a 90g shift for 3q26.2 locus homozygosity for the uncommon allele. Refining of the 10p13 locus identified the potential candidate modifier genes *ITGA8* and *CARP* which are implicated in cardiac fibrosis and apoptosis.

Polymorphisms in regulatory regions of genes might also have a considerable effect on the transcription and translation of the corresponding gene and therefore influence gene expression and disease severity. Recently variants in *MEF2C*, which is a transcriptional regulator of several cardiac genes, were identified with a candidate gene approach. The identified variants were associated with a slightly increased transcriptional activity of *MEF2C* and more severe disease expression (Alonso-Montes, Naves-Diaz et al. 2012).

The most commonly studied putative modifiers in HCM are the genes encoding proteins of the renin-angiotensin-aldosterone system (RAAS). Even though some studies produced conflicting results, polymorphisms in the RAAS system have been associated with increased left ventricular mass, progressive septal hypertrophy and increased severity of left ventricular outflow obstruction, which are all risk factors for poor outcome including sudden cardiac death (Orenes-Pinero, Hernandez-Romero et al. 2011). A recent study in the Netherlands included 368 carriers of one of three mutations in *MYBPC3*. This represents the largest study cohort so far with functionally equivalent truncating mutations in one gene causative of HCM. Five candidate RAAS polymorphisms (*ACE*, rs4646994; *AGTR1*, rs5186; *CMA*, rs1800875; *AGT*, rs699; *CYP11B2*, rs1799998) were investigated. In contrast to previous studies, in this large study population of HCM patients with functionally equivalent mutations in the *MYBPC3* gene no major association between the investigated RAAS polymorphisms and the phenotypic expression of HCM were found (Kolder, Michels et al. 2012).

Treatment with Losartan, an angiotensin II type 1 receptor antagonist, reversed interstitial fibrosis in a genetic mouse model of HCM (Lim, Lutucuta et al. 2001). This result raises the possibility of new therapeutic options for human HCM and supports the hypothesis that interstitial fibrosis, a predictor of sudden cardiac death in HCM, may be reversible. A subsequent small clinical study investigating the effect of Candesartan in human HCM also showed promising results (Penicka, Gregor et al. 2009) and further larger clinical trials are underway to investigate the effect of angiotensin receptor antagonists on disease progression.

It has recently been shown that *MYBPH* functions jointly with *cMYBPC* in regulating cardiomyocyte contraction in response to β -adrenergic stress (Mouton, Loos et al. 2015). Knock-down of either *MYBPH* or *cMYBPC* had little effect on cardiomyocyte contraction, while simultaneous knockdown significantly reduced contraction. Investigation of seven single nucleotide polymorphisms and haplotypes in *MYBPH* in three unique South African HCM-causing founder mutations (p.R403W and p.A797T

in β -myosin heavy chain gene (*MYH7*) and p.R92W in the cardiac troponin T gene (*TNNI2*) revealed a significant association between SNP rs2250509 and hypertrophy traits in the p.A797T *MYH7* mutation group suggesting that variation in this gene may modulate the severity hypertrophy in HCM.

Similarly, the role of variants in muscle ring finger (MuRF) proteins, a class of muscle-specific ubiquitin E3-ligases that appear to modulate cardiac mass and function by regulating the ubiquitin-proteasome system, has recently been investigated in HCM patients. Sequencing of the three members of the MuRF family, MuRF1, MuRF2 and MuRF3, in 594 unrelated HCM patients and 307 healthy controls revealed, that the prevalence of rare variants in both MuRF1 and MuRF2 in HCM patients was higher than that in control subjects. Patients with rare variants in MuRF1 or MuRF2 were younger and had greater maximum left ventricular wall thickness than those without such variants. These findings suggest that dysregulation of the ubiquitin-proteasome system contributes to the pathogenesis of HCM (Su, Wang et al. 2014).

More recently, certain polymorphisms in *ACE2*, *CALM3* and *TNNI3K* were proposed as potential disease modifiers in hypertrophic and dilated cardiomyopathies. Prevalence of *ACE2* (7160726 C>G) variant genotypes (CG and GG) was significantly higher in DCM subjects as compared to controls. Prevalence of *TNNI3K* (3784 C>T) and *CALM3* (-34T>A) variant homozygous genotype were significantly higher in HCM and DCM subjects as compared to controls. DCM patients with CT genotype showed significant decrease in left ventricular ejection fraction as compared to CC genotype ($p < 0.03$). Presence of *ACE2* (7160726 C>G) and *CALM3* (-34T>A) variant genotypes in HCM patients with underlying pathogenic variants (sarcomeric or non sarcomeric genes) was associated with increased mean septal thickness, further suggesting a role of these gene variants in modifying disease phenotype (Kumar, Rani et al. 2018).

Another group suggested that hypertrophic cardiomyopathy was the result of a vicious cycle triggered by sarcomere variants and secondary disease hits. It has been shown, that HCM patients and HCM animal models display excessive cellular reactive oxygen species (ROS). Oxidative stress markers including oxidized proteins, DNA, and

lipids are increased in the heart and serum of HCM patients. Furthermore, increased mitochondrial ROS production and changes in endogenous antioxidants are reported in HCM. Mutant sarcomeric protein may drive excessive levels of cardiac ROS via changes in cardiac efficiency and metabolism, mitochondrial activation and/or dysfunction, impaired protein quality control, and microvascular dysfunction (Wijnker, Sequeira et al. 2018).

Overall, data on genetic modifiers in HCM remains preliminary. Most putative modifiers have not been confirmed in control populations or results from different studies are contradictory. Identification of novel genetic modifiers of HCM may elucidate new pathways that are involved in the complex and largely unknown pathogenesis of myocardial hypertrophy, disarray and fibrosis. New insights in the disease process could inform the development of pharmacological interventions targeting modifier genes and their effects, which could prevent or induce regression of cardiac hypertrophy and fibrosis and as a result improve morbidity and the risk of sudden cardiac death. In addition, as cardiac hypertrophy is a common phenotypic response of the heart to a number of stimulants and an important determinant of morbidity, mortality and sudden cardiac death in all forms of cardiovascular disease (Haider, Larson et al. 1998), these findings could have significant impact on prevention and treatment of various other cardiovascular disorders.

1.4 The Amish population and the Windows of Hope Project

1.4.1 Amish history and culture

The Amish are descendants of the 16th century European Protestant Reformation. They form a genetically isolated founder population and are the offspring of Swiss Anabaptists that immigrated to the United States in the 1700s and 1800s. The Anabaptist movement was founded on the principle that only adults, and not children, can make the decision to commit their lives to God, which is symbolized by baptism. Anabaptists also advocate a separation of church and state and believe that

the church has the right to discipline disobedient members and that believers should live a life according to the teachings of the New Testament including non-resistance (Hostetler, J. 1993). In the 1500s these beliefs were considered a capital offense against the Catholic Church and led to the persecution of many Anabaptists. The Amish were a group of Anabaptists that followed a leader called Jacob Ammann. This is also where the name “Amish” stems from. Jacob Ammann promoted strict church discipline through the use of the *Meidung*, or shunning, and *Bann*, or excommunication. Due to the severe persecution that they experienced the Amish emigrated to America. Other contemporary groups with early Anabaptist roots include the Mennonites, Dunkards, Landmark Baptists, Hutterites, Beachy and Brethren.

Immigration into America occurred in two main waves in the 1700s and 1800s. In the first wave of immigration the Amish settled in Pennsylvania and subsequently in the early 1800s some of these immigrants moved to Holmes County, Ohio (Beachy, L. 2011, *Unser Leit: the story of the Amish*. Millersburg, OH: Goody Heritage Books). The second wave of immigration from Europe established additional communities in Ohio and Indiana (Hostetler, J. 1993, *Amish Society*, 4th ed. Baltimore, MD: Johns Hopkins University Press). Following that, more communities were formed in Indiana by settlers moving from Pennsylvania and Ohio, who were seeking new farmland (Amish Heritage Committee, 2009, *Amish and Mennonites in eastern Elkhart & LaGrange Counties, Indiana 1841-1991*. 2nd printing, Goshen, Indiana).

Since that time the Amish have settled in 31 American states and three Canadian provinces. They are the fastest-growing rural group in the United States and their population doubles every 18-20 years. In 2017 the Amish population was estimated to be ca. 320,000, which represents a 16% increase since 2012. This is attributed to the high birth rate (most families have five or more children) and a retention rate of over 85%. The largest Amish settlements remain in Ohio, Pennsylvania and Indiana. The settlement in Holmes County, Ohio counts 274 church districts and an estimated 35,130 people (“Amish Population Trends 2008-2013, 5-Year Highlights.” Young

Center for Anabaptist and Pietist Studies, Elizabethtown College.
www2.etown.edu/amishstudies).



Figure 1-9: The ‘Charming Nancy’

In 1737 the ship ‘Charming Nancy’ brought 21 Amish families to Pennsylvania

The Amish have a very distinct life-style that is deeply rooted in their culture. They live in rural communities and traditionally their main employment and livelihood has been farming. More recently, as a result of their increasing numbers and the limited amount of affordable farmland, Amish men also work in factories or smaller industries such as carpentry and construction. The Amish keep technology at a distance and only use it selectively, which helps to maintain their culture and heritage. Their main mode of transportation is the horse and buggy, but the majority will use local “English” drivers for longer journeys in cars when necessary.

Amish children attend local purely Amish schools up to eighth grade. Science is not taught in Amish schools and textbooks are often outdated. A recent study investigating the knowledge about cardiovascular disease amongst the Amish found, that there is a significant lack of basic understanding of cardiovascular function which may be a hindrance to health promotion and medical care. The chapter on the circulatory system in a current health book for seventh and eighth graders states:

“Many ailments are blamed on the heart which really have nothing to do with the heart. Fatigue and dizzy spells often come with ear infections. Many chest pains are merely gas pains. Pain from the heart very often comes after undue exertion. It may induce a feeling of great pressure in the chest, or a crushing pain which may extend into the arms. Cardiac, or heart, pain is usually under the breastbone (p. 62)” (Gillum, Staffileno et al. 2011)

The Amish use three languages in specific cultural contexts. The first language children will learn is Pennsylvania Dutch, originally a Swiss-German dialect which has evolved into a distinct language incorporating English elements, which is spoken at home and within the community. The formal language used in church, reading the bible and singing out of the *Ausbund*, the Anabaptist hymn book, is High German. In school, children are taught to speak English in order to be able to communicate with the “outside” world.



Figure 1-10: Amish farm and horse and buggy

The Amish lead an active life which is mainly due to their physical work in farming and manual labour jobs. Women usually stay at home, look after the children and do house work, which also entails a high degree of physical activity as for example the washing for the many household members is done by hand, and a significant

proportion of the food they consume is home-grown. A study investigating the physical activity of Amish from Ontario found that males averaged 18,425 steps/day and had 10 hours of vigorous activity and 42.8 hours of moderate activity per week. Females averaged 14,196 steps/day and 3.4 hours of vigorous and activity and 39.2 hours of moderate activity per week, indicating activity levels that are much higher than in the non-Amish (Bassett, Schneider et al. 2004). Smoking is often prohibited by the *Ordnung* of the church district which makes assessment of smoking rates difficult as questions about smoking habits may not be answered truthfully in fear of being reprimanded.



Figure 1-11: Amish farmer

Most Amish do not accept traditional health care insurance or participate in Medicare, the national social insurance program, administered by the U.S. Federal government, that guarantees access to health insurance for Americans aged 65 and older and younger people with disabilities or end stage renal disease, or Medicaid, a government insurance program for persons of all ages whose income and resources are insufficient to pay for health care. In some areas, Amish communities have negotiated discounted rates from local healthcare providers in exchange for payment in cash. If a single family cannot pay the bill, the church district will help out

to pay the bill or organize fundraisers to support the family (Gillum and Staffileno 2011).

The Amish observe strict endogamy which makes them a well-defined and isolated or even closed founder population, as the original founders were related themselves. The constraints on transport limit the geographic distribution of local communities. Given their relatively uniform life-style including employment in mainly manual jobs and similar diets, the limitations imposed on the Amish by their church which prohibit smoking and alcohol consumption, the variation caused by environmental factors on disease expression are minimized. Any variation in disease phenotype is therefore more likely to be caused by genetic factors. In addition, the Amish keep meticulous and extensive genealogical records (Anabaptist Genealogy Database) which allows the construction of expansive family trees linking large numbers of affected individuals.



Figure 1-12: Two Amish girls in traditional clothing

1.4.2 The Windows of Hope Project

The Amish Windows of Hope project (WOH; www.WOHproject.org) was established by Professor Andrew Crosby and Professor Harold Cross (University of Arizona) in 2000 to identify the causes of inherited conditions amongst the Amish and translate the findings of the research into benefits for the patient, family and community. The program comprises a large collaborative group of scientists, clinicians and other healthcare professionals in Ohio (USA) and the UK. The project aims to translate beneficial findings of genetic research to aid diagnosis and counselling, patient management and ultimately drive the development of new treatments for families with conditions present in the community. Professor Cross was raised amongst the Indiana Amish community and speaks the native language, Pennsylvania Dutch. His interest in inherited disease amongst the Amish stems back to his time as a PhD student in the 1960s, when he investigated inherited diseases under the supervision of Professor Victor McKusick. The WOH project has a base of operation in Ohio, as part of a co-staffed clinical facility (the Newleaf Center; <http://newleafclinic.org/>), and works closely with clinical centres in all of the other major Amish settlements in the USA (Holmes County and Geauga County, Ohio, Pennsylvania and Indiana). Over the last 19 years the WOH team has identified the genes for more than 20 inherited diseases present amongst the Amish. This includes a description of the molecular basis and new pathogenic genes responsible for four forms of motor neuron degenerative disease, five forms of neurodevelopmental delay, and a syndrome comprising myopia and deafness in the Amish community and several other disorders (Rogozin and Milanesi 1997, Patel, Cross et al. 2002, Proukakis, Hart et al. 2002, Simpson, Cross et al. 2009, Tekin, Chioza et al. 2013). The fact that the causative gene for many of these conditions was delineated within a 36-month period from project outset emphasises the power of such founder population studies.

1.5 Cardiovascular disorders in the Anabaptist communities

1.5.1 Hypertrophic cardiomyopathy in the Amish

The first description of asymmetric septal hypertrophy in a large Amish kindred that segregated in an autosomal dominant manner goes back to 1975 (Bingle, Dillon et al. 1975). The kindred were ascertained through a 3-month old infant and her father who were both affected. In 2008 Professor Crosby's team, in collaboration with researchers at Cleveland Clinic (USA), described 10 children with a severe infantile form of hypertrophic cardiomyopathy amongst Holmes and Geauga County (Ohio) Amish, associated with a homozygous essential splice site mutation in intron 30 of *MYBPC3* (3330+2T>G) (Zahka, Kalidas et al. 2008). Affected children die of progressive heart failure within the first year of life unless they undergo cardiac transplantation, which is rarely performed in the Amish community due to cultural and financial constraints. Post-mortem examination showed the hallmark of HCM, myocyte disarray throughout the right and left ventricles. The mutation was not present in 100 control chromosomes of European ancestry and computational splice site prediction analysis using the Spliceview program (Rogozin and Milanesi 1997) highlights the detrimental effect of the 3330+2T>G mutation on the probability of splicing occurring at the exon30/intron30 splice donor site. The computational prediction suggests either the continued reading of exon 30 into intron 30 which contains an in-frame stop codon 111bp into the intron, or potential skipping of the 140bp exon 30, which would disrupt the reading frame and lead to premature termination. At the same time another group investigating genetic disease in the Amish described the same disease in the Amish in Geauga County Ohio (Xin, Puffenberger et al. 2007). Transcript studies of peripheral leucocytes of heterozygous carriers revealed that the mutation resulted in skipping of the 140-bp exon 30, which led to a frame shift and premature stop codon in exon 31 resulting in an aberrant, shorter transcript in addition to the wild type transcript (p.Asp1064GlyfsX38) (Xin, Puffenberger et al. 2007).

Subsequently, the Geauga group undertook a small study of heterozygous carriers of the 3330+2T>G variant, investigating the phenotypic expression of individuals without obvious hypertrophy (De, Borowski et al. 2011). The investigators reported minimal differences in conventional 2-dimensional, Doppler, and speckle-tracking-derived parameters of systolic and diastolic function in a cohort of phenotype-negative carriers compared with healthy subjects. Regional alterations in strain indicative of the presence of underlying subclinical disease were observed and deemed to require further validation.

1.5.2 Other cardiovascular disorders in Anabaptist communities

The Amish, Mennonite and Hutterite communities (Anabaptist or Plain people) of North America have contributed much to the understanding of human genetics over the last 50 years. Seminal work undertaken by Victor McKusick in the 1960s described many characteristics of the Plain people that are conducive to genetic studies (see table 1-2). Since then, numerous genetic disorders have been defined in the Anabaptist communities (reviewed in (Strauss and Puffenberger 2009)) that have greatly helped to understand the molecular basis of diseases both outside and inside these populations.

Cardiovascular phenotypes have been investigated in large-scale genome-wide and also smaller family-based studies describing rare Mendelian disorders. Amish cohorts have been used to describe associations with complex disease traits including blood pressure, hypercholesterolaemia and coronary artery calcification, blood pressure and type 2 diabetes (Hsueh, Mitchell et al. 2000, Rampersaud, Damcott et al. 2007, McArdle, Rutherford et al. 2008, Shen, Damcott et al. 2010). In one study, an Amish cohort of 851 individuals was used to investigate cardiac hypertrophy-associated polymorphisms (Parsa, Chang et al. 2011). The study found that cardiac hypertrophy was heritable, and a genome wide association study revealed 21 candidate hypertrophy-associated SNPs. In a case- control study of unrelated Caucasians, one of the SNPs was associated with hypertrophy ($p=8 \times 10^{-6}$) and heart failure ($p=0.0019$) and in heart failure cases the same SNP was associated with increased

mortality (HR+1.51, p=0.0004). The SNP is in a gene desert of chromosome 20p12 and the study authors concluded that this region has biologic importance in heart failure and the SNP may be useful for predicting a more progressive form of heart failure.

Characteristics of Plain populations conducive to genetic studies

- A self-defined, closed population (little gene inflow)
- Western European origin
- Small number of founding members leading to a restricted gene pool
- Concentrated population settlements (relative immobility)
- Extensive genealogical records
- High standards of living (i.e., nutrition, hygiene) and physical health
- Strong interest in inherited illness and the healing arts
- Relatively uniform social, economic, and occupational circumstances
- Large sibships (typically 5–8 per family)
- Low illegitimacy rate
- Children with birth defects or genetic disorders cared for at home
- Multiple discrete subpopulations (demes) allow comparative analysis
- Identity-by-descent genomic regions flanking many disease gene loci
- Mutation homogeneity (founder effect) commonly responsible for inherited disease
- Population-specific allele frequencies

Table 1-2: Characteristics of Plain populations conducive to genetic studies
adapted from (Strauss and Puffenberger 2009)

Amish cohorts have also been used in modifier studies in cardiovascular disease. One study replicated the effect of previously described genetic variants in *NOS1AP* on QT length (Post, Shen et al. 2007) providing further evidence for the biological function of this gene. Another study investigated the effects of different mitochondrial DNA (mtDNA) haplogroups on the phenotypic expression of a form of mitochondrial cardiomyopathy (Strauss, DuBiner et al. 2013). The severity of the cardiac phenotype of an autosomal recessive form of myopathy and cardiomyopathy caused by a frameshift mutation in nuclear encoded mitochondrial DNA in a large Mennonite pedigree segregated with maternal lineage. Sequencing of mtDNA revealed two different European haplogroups, one of which was associated with more severe and rapidly progressive cardiomyopathy.

Due to their genetic isolation and cultural confinement, there is an increased prevalence of autosomal recessive disorders amongst the Anabaptist communities, in particular amongst smaller groups which previously 'split off' from the more mainstream Amish and remained isolated. A number of rare Mendelian diseases have been described which contributed to the understanding of molecular pathways and disease pathology, and importantly have often subsequently been described outside the Anabaptist communities highlighting the global relevance of such studies. Many of the described syndromes are associated with congenital heart disease including Ellis-van-Creveld syndrome, primary ciliary dyskinesia, Alström syndrome, DCMA syndrome (dilated cardiomyopathy with ataxia), Renpenning syndrome, PMSE syndrome (Polyhydramnios, Megalencephaly, and Symptomatic Epilepsy syndrome), Sitosterolemia, beta-glucuronidase deficiency mucopolysaccharidosis, McKusick-Kaufman syndrome and limb girdle muscular dystrophy type 21 (Wu, Tomatsu et al. 1994, Stone, Agarwala et al. 1998, Frosk, Greenberg et al. 2005, Davey, Parboosingh et al. 2006, Marshall, Hinman et al. 2007, Puffenberger, Strauss et al. 2007, O'Connor, Rider et al. 2011, Ferkol, Puffenberger et al. 2013) and www.ncbi.nlm.nih.gov/pmc/articles/PMC3077314/).

In addition to disorders and genes that have first or only been described in plain people, cardiac diseases and their underlying gene defects discovered outside the Amish communities have also been described in Anabaptist communities. These may arise due to novel variants and may be autosomal recessive in nature rather than dominant. Examples are a form of dilated cardiomyopathy with AV block caused by a mutation in *LMNA*, dilated cardiomyopathy with arrhythmia caused by a mutation in desmoplakin, and dilated cardiomyopathy associated with propionic acidaemia (*PCCB*) (clinicforspecialchildren.org/Research). Arrhythmia syndromes have also been described, including a recessive form of Long QT syndrome associated with deafness (Jervell and Lange-Nielsen syndrome) associated with a mutation in *KCNQ1* (Chen, Zhang et al. 1999) and Timothy syndrome caused by a mutation in *CACNA1C* (clinicforspecialchildren.org/Research).

2 Original hypotheses and aims

2.1 Syndrome of orofacial clefting and cardiac anomalies

This aspect of the project involved the delineation of the genetic and molecular basis and clinical outcomes of a novel syndrome characterised by cleft lip and palate, pectus excavatum, myopia and cardiac anomalies present in the Amish community. The overarching premise underpinning this study is that the condition arises due to the presence of an autosomal recessively acting founder mutation present at increased frequency in the community, located in a gene not previously associated with human disease.

The specific aims of the project entailed:

- 1) detailed phenotyping of affected individuals,
- 2) defining the underlying molecular basis of this new syndrome using autozygosity mapping, next generation and Sanger sequencing technologies,
- 3) functionally characterising the putative disease gene and pathogenic variant identified,
- 4) investigating findings in a knock-out mouse model of disease,
- 5) translating beneficial research findings into direct benefits to the affected individuals including improved diagnosis and prognosis, genetic counselling and potentially therapies.

2.2 Hypertrophic cardiomyopathy

The overarching hypotheses underpinning this project relate to the investigation of an *MYBPC3* variant (c.3330+2T>G), believed to be the principle cause of HCM in this large founder population. The condition presents with highly variable phenotype as is typically seen in other forms of HCM. Clinical outcomes may be influenced either by alterations to wild type or alternate mutant mRNA transcript outcomes arising from the mutant *MYBPC3* allele, or to genetic modifiers; these modifiers themselves are hypothesised to (primarily) entail founder sequence variants, and so be enriched

within the community aiding their definition. To investigate these hypotheses, a three-pronged approach was employed entailing a genome-wide SNP study, exome sequencing analysis and transcript studies.

The aims of this project entailed:

- 1) phenotypic description of the largest family cohort sharing the same HCM mutation described to date, permitting unique insights into the phenotypic variability and clinical expression associated with a single *MYBPC3* mutation,
- 2) identification of genetic modifiers of HCM in this community, to enable
 - an improved understanding of the disease mechanisms and pathways involved in the development of the variable phenotype typical of HCM,
 - the elucidation of pathways of cardiac hypertrophy, fibrosis, myocyte disarray and arrhythmogenesis,
 - the translation of genetic findings into direct patient benefit including improved risk prediction and prognosis,
 - the potential to develop novel therapeutic strategies for prevention and treatment of HCM,
 - an improved understanding of the mechanisms leading to left ventricular hypertrophy in general in other cardiovascular disorders.

3 Methods

3.1 Ethical approval

The Windows of Hope Project studies were reviewed and approved by the Institutional Review Board of the Office for Responsible Conduct of Research, University of Arizona, Tucson, USA (10-0050-10) and St George's University of London. The HCM study was also approved by the Institutional Review Board of University Hospitals Case Medical Center in Cleveland. All studies were conducted in accordance with the principles of the Declaration of Helsinki. Participants were required to sign a consent form and provide clinical details and a blood or buccal sample.

3.2 Recruitment and phenotypic assessment

3.2.1 Orofacial clefting and cardiac anomalies project

A full medical and developmental history were taken, and a clinical examination was performed for all individuals recruited. Photographs and relevant investigation reports were obtained with informed consent. Phenotypic assessment of dysmorphic features was confirmed by Dr Emma Baple, Consultant Clinical Geneticist, Royal Devon and Exeter NHS Trust. Audiograms were reviewed by Dr Wendy Albuquerque at St George's Healthcare NHS Trust. Height and weight were measured barefoot in light clothing. Occipito-frontal circumference (OFC) was measured according to standard procedures.

Height, weight and OFC standard deviations and Z-scores were calculated using a Microsoft Excel add-in to access growth references based on the LMS method (Pan H, Cole TJ. LMS growth, a Microsoft Excel add-in to access growth references based on the LMS method. Version 2.77, <http://www.healthforallchildren.co.uk/>;2012) using a reference European population (Cole, Freeman et al. 1998). 12 lead digital

Electrocardiograms using laptop-based ECG software (Cardiosoft™ GE Healthcare, UK) and transthoracic Echocardiograms with acquisition of 2D, colour and tissue Doppler (TDI) images using a portable Vivid I Echo machine, GE Healthcare, were performed on all individuals with orofacial clefting.

3.2.2 Hypertrophic cardiomyopathy project

3.2.2.1 Study participants

First degree relatives of any infant homozygous for the *MYBPC3* 3330+2T>G variant, relatives of known gene carriers or anyone known to be affected by HCM and their first-degree relatives living in Holmes or Geauga County, Ohio, USA, were offered a genetic test for the known pathogenic *MYBPC3* variant through regional clinical services and to participate in our research study. Only adults above the age of 18 were enrolled in the research study. Informed consent was obtained, and two tubes of blood were drawn into potassium EDTA tubes. One sample was sent to the accredited and diagnostically approved genetic testing centre at University Hospitals Case Medical Center, Cleveland (UH Cleveland) and one sample was sent to St George's University of London. Sample collection and recruitment were performed by nurses working for Windows of Hope (WOH). After test reports from UH Cleveland were obtained, participants who tested positive for the *MYBPC3* variant were invited to a consultation, electrocardiogram, transthoracic echocardiogram and 24-hour holter monitor. These assessments took place at the Windows of Hope Genetics Information Center in Walnut Creek, Holmes County and at the Birthing Center in Middlefield, Geauga County in Ohio, USA. A copy of each positive genetic test report was also sent to the participant's General Practitioner or Cardiologist.

Commented [CA1]: Lets just leave this as not relevant to the study parameters

3.2.2.2 Clinical assessment

The following information was recorded from each participant at the time of their assessment:

- Age
- Gender

- Number of children with infantile HCM and healthy children
- Occupation
- Symptoms and age at onset: exertional chest pain, exertional shortness of breath, palpitations, dizziness, syncope, fatigue
- Past medical history including direct questioning regarding history of hypertension, hypercholesterolaemia, coronary artery disease, HCM and atrial fibrillation
- Previous assessments by Cardiologist
- Drug history
- Family history, in particular history of HCM or sudden death

The following measurements were taken

- Height
- Weight
- Up to three measurements of blood pressure on the left arm after at least 10 minutes of sitting or lying down, separated by at least 10 minutes

The following investigations were performed:

- 10-second 12-lead digital Electrocardiogram using laptop-based ECG software (Cardiosoft™ GE Healthcare, UK) in the supine position
- Transthoracic Echocardiogram with acquisition of 2D, colour and tissue Doppler (TDI) images using a portable Vivid I Echo machine (GE Healthcare)
- 24-hour ambulatory monitoring using digital 10-electrode 12-channel recording devices with a sampling frequency of 1024Hz (CardioMemR CM 3000-12, Getemed, Germany), applied in the Mason-Likar configuration

Commented [CA2]: I don't think this adds much....

3.2.2.3 Electrocardiography

12 -lead digital electrocardiograms were analysed using a custom-developed software program written in Matlab (MathWorks, Massachusetts, USA) by Dr

Velislav Batchvarov, St George's University of London. Electronic callipers were used for all measurements.

The following measurements and observations were obtained:

at 10 mm/mV and 25mm/s:

- Rhythm
- QRS axis
- Largest R or S in limb leads
- Rhythm strip length
- Number of RR intervals

At 20 mm/mV and 50 mm/s:

- Right atrial enlargement (P wave ≥ 2.5 mm in II, III, or aVF)
- P wave duration (in lead II or I/III)
- PR interval (lead II)
- QT interval (lead II or V5), QTc was calculated using the Bazett's correction and the mean RR interval of the rhythm strip
- Depth and length of negative portion of P wave in V1
- Left atrial enlargement (prolonged P wave duration of > 120 ms in leads I or II with negative portion of the P wave ≥ 1 mm in depth and ≥ 40 ms in duration in lead V1, (Sharma, Drezner et al. 2018))
- Largest S wave in V1/V2
- Largest R wave in V1/V2
- Largest S in V5/V6
- Largest R in V5/V6
- Presence of RVH ((R wave height in V1 + deepest S wave in V5-V6 > 1.05 mV)
- Presence of LVH (S wave depth in V1 + R wave height in V5/V6 ≥ 35 mm)

At 20mm/mV and 100 mm/s:

- Length of intrinsicoid deflection in V5/V6

- QRS duration (lead V5)
- Presence of LBBB
- Presence of RBBB
- QRS fragmentation (in patients with QRS duration <120 milliseconds: the presence of an additional R wave (R' prime), notching in nadir of the S wave, notching of R wave, or the presence of more than one R prime in two contiguous leads corresponding to the left ventricular (LV) segment; in patients with right or left bundle branch block (QRS duration ≥120 milliseconds): various RsR' pattern with >2 R', >2 notches in the R wave, or >2 notches in the downstroke or upstroke of the S wave, in 2 contiguous leads corresponding to the LV segment; (Konno, Hayashi et al. 2015))
- Pathological Q waves (Q wave >1/4 of the ensuing R wave in depth and/or >40 milliseconds in duration in at least two leads except aVR)
- T wave inversion (T wave deflection of ≥ -0.1 mV in at least two contiguous leads excluding AVR, V1 or II)
- Deep T wave inversion (T-wave deflection of ≥ -0.2 mV in at least two contiguous leads excluding AVR, V1 or II)
- ST depression (ST shift of ≥-0.1 mV in two or more contiguous leads)
- ST elevation (ST shift of ≥ 0.1 mV in two or more contiguous leads)

The Romhilt-Estes score, a point score system for the diagnosis of left ventricular hypertrophy (LVH) from the ECG, which has previously been shown to have a specificity of 96.8% and a sensitivity of 60% in an autopsy series (Romhilt and Estes 1968) was calculated for each individual (table 5-1).

“Abnormal ECG” was defined as positive if any of the following was present:

- Romhilt-Estes score of ≥ 4
- Pathological Q waves (as defined above)
- T wave inversion (as defined above)
- ST depression (as defined above)

Criteria	Points
Amplitude Positive if (1) Largest R or S wave in the limb leads ≥ 20 mm or (2) S wave in V1 or V2 ≥ 30 mm or (3) R wave in V5 or V6 ≥ 30 mm	3
ST-T segment Without digitalis With digitalis Positive if typical ST-T pattern of left ventricular strain is present (ST-T segment vector shifted in direction opposite to mean QRS vector)	3 1
Left atrial involvement Positive if the terminal negativity of the P wave in V1 is 1mm or more in depth with a duration of 0.04 seconds or more	3
Left axis deviation Positive if left axis deviation of -30° or more is present in frontal plane	2
QRS duration Positive if QRS duration is ≥ 0.09 seconds	1
Intrinsicoid deflection Positive if intrinsicoid deflection in V5 or V6 ≥ 0.05 seconds	1
Maximum total	13

Table 3-1: Romhilt-Estes point score system

Five points is read as LVH, four points is read as probable LVH

3.2.2.4 Transthoracic Echocardiography

Echocardiogram studies were recorded and stored on DVDs in CD/DVD archive mode. Offline analysis was performed with the Echopac software (GE Healthcare).

The following measurements and observations were obtained:

- Assessment of degree/type of left ventricular hypertrophy, wall thickness and relative wall thickness:
 - Parasternal long-axis view: IVSd, LVDd, PWd, relative WT
 - Parasternal short-axis at view mitral valve level: WT anterior/lateral/inferior/septal, anterior-inferior diameter and septal-lateral diameter
 - Parasternal short axis view at papillary muscle level: WT anterior/lateral/inferior/septal, anterior-inferior diameter and septal-lateral diameter
 - Parasternal short axis view at apical level: WT anterior/lateral/inferior/septal
 - Type of LVH: sigmoid septum, reverse septal curvature, concentric, apical, segmental anterior/lateral, dilated phase
- Assessment of LV size and systolic function
 - Parasternal long-axis view: LVDd, LVDs
 - Apical four and two chamber view: Simpson's biplane ejection fraction
 - Apical four chamber view: MAPSE lateral and septal
- Assessment of LV diastolic function
 - Four chamber view: tissue Doppler septal and lateral mitral annulus
 - Two chamber view: tissue Doppler inferior and anterior mitral annulus
 - Four chamber view: mitral valve inflow peak E and A velocity and E deceleration time

- Four chamber view: peak pulmonary vein flow velocities S, D and A, atrial reversal duration
- Four/five chamber view: IVRT
- Assessment of left atrial volume
 - Four chamber view: LA area and volume (area length method)
 - Two chamber view: LA area and volume (area length method)
- Assessment of left ventricular outflow tract velocity in five chamber view
- Length of the anterior mitral valve leaflet in PLAX, four chamber and three chamber view
- Systolic anterior motion of the mitral valve/cordae
- Papillary muscle morphology (attachment, hypertrophy, bifid, apical displacement, hypermobile)

3.2.2.5 Holter monitoring

Holter monitoring was performed for a 24-hour period in a randomly selected subset of study participants (n=65). Digital 10-electrode 12 channel recording devices with a sampling frequency of 1024Hz (CardioMem^R CM 3000-12, Getemed, Germany) were applied in the Mason-Likar configuration. The low pass filter setting for the 12 lead Holters was 70Hz. Analysis was performed on a workstation using commercial Holter analysis software (Cardioday^R, Getemed, Germany).

Manual over-reading was performed to eliminate artefact and correct the automated identification of ventricular premature beats (VPB). Maximum, minimum and mean heart rate, and total number of VPBs, number of episodes of bigemini and number of episodes and duration of non-sustained VT were recorded.

3.3 Data management

Following recruitment of an individual, each family was given a pedigree ID, each individual an individual ID, and each sample a sample ID. All clinical and molecular

information is recorded in password protected databases. Family pedigrees were constructed with the help of the online Swiss Anabaptist Geneological Association database (www.saga-omii.org) and stored using Progeny (800 Palm Trail, Suite 200, Delray Beach, FL 33483 USA (<http://www.progenygenetics.com/>)). All ECG and ECHO data is stored in anonymised form on a dedicated laptop and DVDs respectively.

3.4 Molecular methods

3.4.1 DNA extraction

Lymphocyte DNA was extracted from whole blood using the Promega ReliaPrep™ Blood gDNA Miniprep System.

The blood samples were stored at -20°C until DNA extraction was performed. Prior to DNA extraction, the samples were thawed, and adequate mixture of the different blood components was ensured. Filter tip pipettes were used at all stages to prevent contamination of samples and equipment.

20µl of Proteinase K were dispensed into a 1.5ml microcentrifuge tube. 200µl of blood were mixed with the Proteinase K by repeat pipetting. Then 200µl of cell lysis buffer were added to the microcentrifuge tube. The tube contents were vortexed for 10-20 seconds to ensure thorough mixing and incubated at 56°C for 10 minutes.

Following the incubation period, 250µl of Binding Buffer A were added and the microcentrifuge tube contents were vortexed for 10-20 seconds. The lysate appeared dark green in colour at this stage. In order to remove any residue on the tube lid the tubes were briefly centrifuged. The lysate was transferred to a ReliaPrep binding column which was placed into a collection tube. The collection tube was then placed in a microcentrifuge and spun at maximum speed (13,000rpm) for one minute. If there was any lysate visible on top of the column membrane afterwards, the tube was spun for a further minute. The column was placed in a clean collection tube and the flow through was discarded. Then 500µl of Column Wash Solution

were added to the column and the column was spun in the microcentrifuge for three minutes at maximum speed. The flow through was discarded as hazardous waste. This washing step was repeated two more times.

The column was then placed in a clean 1.5ml microcentrifuge tube and 200µl of nuclease-free water were added. Following that, the column was spun for one minute at maximum speed to elute the DNA.

After elution, the DNA concentration and purity were measured using the Nanodrop spectrophotometer. Sterile distilled water was used for dilution to achieve the desired DNA concentration of 10-30ng/µl.

3.4.2 RNA extraction

Lymphocyte RNA was extracted from whole blood collected in PAX gene RNA blood tubes. Tube racks and pipettes were cleaned with RNase Zap (Ambion). Filter tip pipette tips were used in order to prevent contamination of samples and equipment. The protocol for the Qiagen PAX gene blood RNA kit handbook for extraction and purification of RNA was followed as summarised below.

The PAX gene RNA blood tubes were centrifuged at 3000x g for 10 minutes in a swing-out centrifuge. The supernatant was removed with a pipette without disturbing the pellet. After adding 4ml of RNase free water to the pellet, the blood tube was closed with a new BD Hemogard closure. The tube was vortexed until the pellet was dissolved and then centrifuged at 3000x g for 10 minutes. The supernatant was removed with a pipette. 350µl of buffer BR1 were added to the pellet, the sample was vortexed until the pellet was re-suspended again, and the sample was transferred to a 1.5ml microcentrifuge tube. 300µl of buffer BR2 and 40µl of proteinase K were added and the sample was vortexed for 5 seconds. The sample was incubated at 55°C in a shaker-incubator set at 1400 rpm for 10 minutes and then transferred to a PAX gene shredder spin column in a 2ml processing tube. The sample was again centrifuged at 14,000xg for 3 minutes. The flow-through was transferred to a fresh 1.5ml microcentrifuge tube. After adding 350µl of 100% purity

grade ethanol, the sample was mixed by vortexing for 10 seconds. To remove any fluid from the lid, the tube was briefly centrifuged (1-2 seconds).

700µl of sample were pipetted in a PAX gene RNA spin column placed in a 2ml processing tube and centrifuged for 1min at 14,000x g. The spin column was transferred to a fresh processing tube. The flow-through was discarded as hazardous waste. After repeating this step with the remaining sample, 350µl of wash buffer 1 (BR3) were added to the PAX gene spin column. The spin column was centrifuged for 1 minute at 14,000x g and then transferred to a new 2ml processing tube. The old tube and flow-through were discarded.

70µl of DNA digestion buffer and 10µl of DNase I stock solution (supplied with the kit) were added to a 1.5ml microcentrifuge tube. The tube was gently flicked to mix the solutions whilst preventing denaturation of the enzyme. The tube was centrifuged for 1-2 seconds to remove any liquid from the lid and the side of the tube. The solution was transferred onto the spin column membrane and incubated at room temperature for 15mins. 350µl of wash buffer 1 (BR3) were added to the spin column and the column centrifuged at 14,000x g for 1 minute and then transferred to a new processing tube. After adding 500µl of wash buffer 2 (BR4) to the spin column, the tube was at 14,000x g for 3 minutes. The spin column was placed in a fresh tube and centrifuged for a further 1 minute at 14,000x g. The tube was discarded, and the spin column placed in a 1.5ml microcentrifuge tube. 40µl of elution buffer (BR5) were added onto the spin column membrane and the tube was centrifuged at 14,000x g for 1minute to elute the RNA. This step was repeated using a further 40µl of elution buffer in the same microcentrifuge tube. In order to denature the RNA, the tube and contents were incubated at 65°C for 5 minutes using the desk-top shaker-incubator. The eluted RNA was placed on ice and stored in the -80°C freezer.

3.4.3 Primer design

The relevant gene sequences were obtained from the Ensembl Genome Browser (February 2009 assembly, <http://www.ensembl.org/index.html>). Primer design was performed with the freely available Primer3 software version 0.4.0 (<http://frodo.wi.mit.edu/primer3/>).

The following criteria were used for designing primers:

- Primer sizes were 18-27 bases in length (specificity and annealing temperature are dependent on the length of the primer).
- Melting temperatures were kept as similar as possible for both the left and right primers to maintain PCR efficiency (ideally between 57 - 63°C).
- Where possible the guanine - cytosine (GC) base content was kept between 40-60%
- Primers with inter or intra-primer efficiency extending for more than 3 bases were avoided, to prevent the formation of secondary structures and primer dimers.
- In silico PCR analysis and BLAST searches were performed using the UCSC Genome Bioinformatics website (<http://genome.ucsc.edu/>) to ensure the primer sequences were specific and a perfect match to the region to be amplified.

3.4.4 Polymerase chain reaction

Polymerase chain reaction (PCR) is used for the cloning of DNA. It allows the exponential amplification of a specific part of the DNA with only very small amounts of DNA as a starting point. A touchdown PCR protocol was used to improve the specificity of the PCR by avoiding amplification of non-specific sequences. The annealing temperature is lowered in decrements of 2°C every 2 cycles from 2°C above the expected melting temperature (T_m) of the primers to 2°C below the T_m . The primers anneal initially at the highest temperature and the resulting template is

used in subsequent cycles. The T_m of the primer defines the upper limit of the annealing temperature. At temperatures just above this point, only very specific binding of the primer to the relevant DNA sequence will occur. At lower temperatures, the primers bind less specifically. If non-specific primer binding occurs in the early steps of amplification, these non-specific sequences will get amplified exponentially and obscure the PCR results.

During a touch down PCR protocol, the primer will anneal at the highest temperature that it is able to tolerate which will be specific to the DNA sequence of interest. Thus, the first sequence amplified is the one between the regions of greatest primer specificity. These DNA fragments will then be further amplified during subsequent rounds at lower temperatures and will “out-compete” the nonspecific sequences to which the primers may bind at those lower temperatures. Touchdown increases specificity of the reaction at higher temperatures and increases the efficiency in subsequent cycles by lowering the annealing temperature.

Number of cycles	Temperature [°C]	Time [seconds]
1	98	120
2	95	30
	Anneal Temperature + 4	45
	72	45
2	95	30
	Anneal Temperature + 2	45
	72	45
35	95	30
	Anneal Temperature	45
	72	45
1	72 (elongation)	300

Table 3-2: Touchdown PCR program

The following reagents were added to a 0.2ml 96-well microtiter plate or 0.2ml tubes for a PCR reaction:

Reagent	Volume [μ l]
Left primer (10 μ M - Sigma Aldrich)	0.6
Right primer (10 μ M - Sigma Aldrich)	0.6
dNTPs (10mM)	0.6
10xBuffer (20mM MgCl ₂)	2.0
Sterile distilled water	15.1
Red hot <i>Taq</i> polymerase (5U/ μ l –Fisher Thermoscientific)	0.1
DNA (10ng/ μ l)	1
Total	20

Table 3-3: Reagents for a 20 μ l PCR reaction

Amplification conditions for a primer pair were optimised according to the protocol below (5.4.5 Optimisation of amplification conditions) and initially used with control DNA samples that were known to amplify well in other PCR reactions. A water control was used in each reaction to exclude contamination of the PCR products. An Eppendorf 96 well mastercycler, thermal cycler machine was used for the PCR reactions.

3.4.5 Optimisation of amplification conditions

A Temperature gradient of 55-67°C was applied across the PCR block in order to establish the optimal annealing temperature for the primer pair both with and without the addition of 10% Dimethylsulfoxide (DMSO) and higher concentration of magnesium chloride(3mM).

3.4.6 Agarose gel electrophoresis

Adequate amplification of the DNA was checked by performing agarose gel electrophoresis with the products of the PCR reactions. For a 2% agarose gel, 2g of agarose were added to 100ml of 1x tri-acetate EDTA (TAE) buffer. The solution was heated in a microwave for 4 min until the agarose was dissolved. This was followed by the addition of 2 μ l of a 10mg/ml solution of ethidium bromide after the gel solution had cooled for two minutes. Ethidium bromide is a DNA intercalating agent that fluoresces brightly when exposed to ultraviolet light. A casting tray with combs to create wells was prepared and the gel poured into it. The gel was left to set for ca. 30 minutes and then placed into an electrophoresis tank filled with TAE buffer to cover the surface of the gel. To assess the size of the DNA products, 2 μ l of a DNA ladder (1KBplus from Invitrogen) made up with agarose gel loading buffer (final concentration 50ng/ μ l) was loaded into the first well. 5 μ l of each PCR product were mixed with 3 μ l of loading dye and loaded into each subsequent well. Following that, electrophoresis through the gel was performed for 40 minutes at 100 volts. The PCR products were then visualized by placing the gel in a UV transilluminator and photographs were taken with a filtered camera.

3.4.7 Purification of PCR products

Unincorporated primers and dNTPs were removed prior to the sequencing reaction with the use of ExoSAP-IT, a product that contains Exonuclease-1 (EXO) and Shrimp Alkaline Phosphatase (SAP). Exonuclease-1 degrades single stranded DNA in a 3 to 5 direction, releasing deoxyribonucleoside 5-monophosphates in a stepwise manner. This process releases dNTPs from unincorporated primers. Shrimp Alkaline Phosphatase (SAP) catalyses the release of 5'- and 3'-phosphate groups from extra nucleotides. 2 μ l of ExoSAP-IT were added to 5 μ l of each PCR product. The samples were then incubated at the optimal temperature for the enzyme activity (at 37°C) for 15 minutes and subsequently at 85°C for 15 minutes to inactivate the enzyme.

3.4.8 Sequencing reaction

Following purification, the PCR products were sequenced using the BigDye Terminator Cycle Sequencing Kit v3.1 (ABI, Applied Biosystems). The kit contains Taq DNA Polymerase, dNTPs, ddNTPs-Dye terminators and MgCl₂ buffer. Separate reactions were performed with forward and reverse primers. Each of the four dideoxynucleotide chain terminators is labelled with a different fluorescent dye. Each fluorescent dye emits light at a different wavelength. The components of the sequencing reaction are shown in table 5.4.

Reagent	Volume [μl]
Big Dye	0.5
Big Dye buffer	1.75
Sterile distilled water	4.25
Left or right primer (10mM)	0.5
Purified PCR product	3.0

Table 3-4: Components of a 10μl sequencing reaction

The program for the sequencing reaction is outlined in table 5.5.

Number of cycles	Temperature [°C]	Time [seconds]	Process
25	96	30	Denaturation
	50	15	Annealing
	60	4	Elongation

Table 3-5: Program for the sequencing reaction

3.4.9 Sequencing reaction purification

The products from the sequencing reaction were purified before being processed on an automated DNA sequencer. The purification was performed with the BigDye

XTerminator purification kit (ABI, Applied Biosystems). 45µl of SAM solution and 10µl of XTerminator solution were added to each sample in a 96 well plate. The contents of the plate were mixed thoroughly using a desk top shaker set at full speed for 30 minutes and then centrifuged for 2 minutes at 1000g. The samples were run on a sixteen capillary 3130XLA Applied Biosystems Automated Sequencer. A 36cm array POP7 polymer program setting was used on the sequencer. The genomic sequence was checked using Chromas light (Technelysium Pty Ltd), Gene tool (Biotoool Incorporated) and Mutation Surveyor (Soft Genetics) software.

3.4.10 Single nucleotide polymorphism (SNP) genotyping

For the *HYAL2* project, a genome wide microarray scan using Illumina Human CytoSNP-12v2.1 arrays was undertaken using DNA from affected family members following the manufacturer's protocol. Each chip genotypes 12 samples. The assay requires 200ng of DNA per sample at a concentration of 50ng/µl and is conducted over a three-day period (Figure 3-1).

For the HCM project, samples of 202 gene carriers were used for genotyping on the Affymetrix Genome-Wide Human SNP Array 6.0 featuring 1.8 million genetic markers, including more than 906,600 single nucleotide polymorphisms (SNPs) and more than 946,000 probes for the detection of copy number variation. This work was outsourced to the company EpiGenDx (www.epigendx.com).

3.4.11 Exome sequencing

Whole-exome sequencing of individuals was performed using two different sequencing platforms, Otogenetics and the Baylor College of Medicine Human Genome Sequencing Center (BCM-HGSC). At Otogenetics, genomic DNA was captured using the SureSelect Human All Exon V4 (Agilent Technologies) exome enrichment kit and sequenced on an Illumina HiSeq2000. At BCM-HGSC, genomic DNA was captured with the VCRome 2.1 enrichment kit (52 Mb, NimbleGen, Roche Sequencing) according to the manufacturer's protocol and sequenced on an Illumina HiSeq2000.

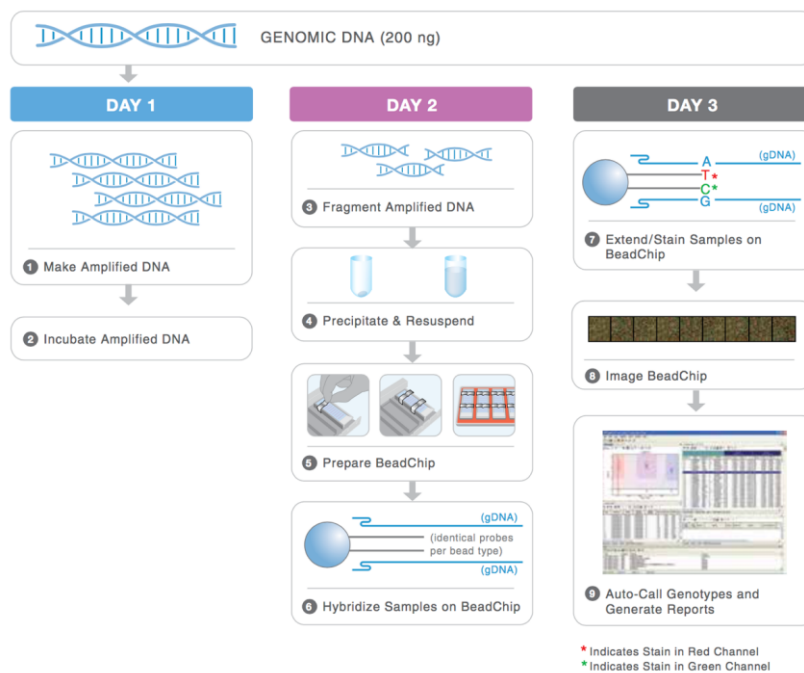


Figure 3-1: Infinium assay protocol

The FASTQ files obtained from Otogenetics and BGI were mapped to the reference genome using the Burrows-Wheeler Aligner BWA-MEM (v0.7.12) algorithm (Li, Handsaker et al. 2009, Li and Durbin 2010). This algorithm was used due to its improved performance, compared to other BWA algorithms, being faster and more accurate than previous versions as well as providing higher quality queries. The sequence alignment map (SAM), was converted to binary SAM format (BAM) to produce a smaller file and to increase the processing speed.

Duplicate reads were marked by Picard (version 1.129). The BAM file was then realigned, to account for indels, with variants called using GATK-HaplotypeCaller and subsequently quality filtered based on mapping quality, strand bias, the average position of a variant in a read and SNP quality. The variant call file (VCF) file was

annotated using the Alamut Software (v1.4.4) Suite. Variants were quality control (QC) checked and filtered for rare, non-synonymous exonic or splice variants, with a population frequency of <0.005 in control databases (including the Genome Aggregation Database; gnomAD, the Exome Aggregation Consortium; ExAC, and the 1000 Genomes Project) (Interactive Biosoftware).

Additional information about identified variants and associated phenotypes was obtained through interrogation of ClinVar (Landrum, Lee et al. 2016), a public archive of reports of the relationships among human variations and phenotypes, with supporting evidence and pubmed.

3.4.12 Reverse transcription PCR

Reverse transcriptase PCR (RT-PCR) is a methodology to indirectly amplify RNA by initially making a complementary DNA (cDNA) copy. The amplified cDNA is subsequently sequenced, and the RNA sequence inferred.

Prior to starting, the bench top, pipettes and racks were cleaned with 70% ethanol and subsequently RNaseZap (Ambion) in order to remove any traces of RNase. Filter tip pipette tips were used throughout the process. RNA was extracted from whole blood as previously described and stored on ice until required. The Titanium one-step RT-PCR kit and protocol (Clontech) were used. RT-PCR primers were designed targeting exonic sequences separated by at least one intron. Primer design was performed using Primer 3. The programme settings were changed so that the primers met the following specifications:

1. Primer melting temperature of approximately 70°C.
2. Primer length at least 22 nucleotides (25-30 preferable).
3. G-C content of 45-60%.

The primers were diluted, and a primer-mix created with a final concentration of 45µM for each primer.

A master mix was prepared: 1ng - 1µg of patient-RNA was added to each subsequent reaction and the total volume was made up to 50µl with RNase free water. In addition to a negative control in which water replaced RNA to control for contamination, an additional negative control was included. One reason for concern performing RT-PCR is, that genomic DNA may be amplified instead of RNA. In order to control for this possibility, the titanium taq was replaced with standard taq used for PCR, and RNA was replaced with water. A positive control was also included. Primers specific for mouse β actin were used to amplify β actin from mouse liver total RNA (supplied with the one- step kit). Please see table 3-6 for details of the reaction set-up.

Component	Volume [µl]
Primer mix 45 mM	1.0
50x dNTP mix	1.0
Recombinant RNase inhibitor (400u/ µl)	0.5
Thermostabilizing reagent	25.0
GC melt	10.0
Oligo (dT) primer	1.0
50x Titanium Taq RT enzyme mix	1.0

Table 3-3-6: Components of a 50µl PCR reaction

The reactions were put in a thermal cycler. The programme in table 3-7 was used.

Number of cycles	Temperature	Time
1	50	1 hr
1	94	5 min
	94	30 s
40	65	30 s
	68	1 min
1	68	2 min

Table 3-7: RT-PCR program

The products were electrophoresed on a 2.5% agarose gel for 90 mins with the appropriate DNA ladder to confirm successful cDNA amplification and to establish the size of the products.

3.4.13 Other methods used in the *HYAL2* project

3.4.13.1 Experimental animals

As previously described (Jadin, Wu et al. 2008), *Hyal2*^{-/-} (null) mice and littermate controls (*Hyal2*^{+/-} or *Hyal2*^{+/+}) were obtained from heterozygous intercrosses of mice maintained on an outbred (C129; CD1; C57BL/6) background. The morning a vaginal plug was discovered, embryos were considered E0.5. Genotyping was performed according to previously established methods (Chowdhury, Hemming et al. 2013). Animals were euthanized by isoflurane overdose and embryos collected for morphological and histological studies. All procedures were in compliance with the Canadian Council on Animal Care and followed a protocol approved by the University of Manitoba Animal Care Committee.

3.4.13.2 Animal imaging, hearing analysis and histological studies

Ultrasound imaging was performed on mice anesthetized with 2% isoflurane, using a Visual Sonics Vevo 2100 ultrasound equipped with a 40 MHz transducer. Micro-CT imaging was performed at 9µm resolution with a Skyscan 1176 micro-CT and images were reconstructed and analyzed using Bruker's NRecon, Data Viewer, CT Analyzer, or CT Vox software. Auditory brainstem response (ABR) testing was performed as described previously (Hochman, Blakley et al. 2006). Skeletal staining of neonatal mice was performed as described in (McLeod 1980)[54]. For histological studies, tissues were fixed in 1% cetylpyridinium chloride in 10% buffered formalin and decalcified if appropriate. After tissues were imbedded in paraffin, 5µm sections were obtained. Sections were then stained with H & E for analysis of morphology. HA was detected using the HA binding protein (HABP) following previously established protocols (Chowdhury, Hemming et al. 2013).

3.4.13.3 HA and hyaluronidase assays

HA levels were determined using a DuoSet enzyme linked immunosorbent assay (ELISA) development kit (R & D Systems).

3.4.13.4 Comparative molecular model of HYAL2 bound to HA

Human HYAL1 (NCBI Ref: NP_695013.1) and HYAL2 (NCBI Ref: NP_149348.2) protein sequences aligned using ClustalW2.1 (Larkin, Blackshields et al. 2007) provided the basis for a comparative molecular model of HYAL2 built using Modeller v9.13 (Marti-Renom, Stuart et al. 2000)[56] and the crystal structure of HYAL1 (PDB ID: 2PE4) (Chao, Muthukumar et al. 2007) as a 3D template. A Michaelis complex of HYAL2 bound to a hexamer of HA was generated by first superposing the crystal structure of bee venom hyaluronidase bound to a tetrameric HA product (PDB ID: 1FCV) (Markovic-Housley, Miglierini et al. 2000) onto the HYAL2 model, which placed the HA tetramer from the bee venom structure into the sugar binding subsites -1 through -4 (see nomenclature in (Davies, Wilson et al. 1997)) of the HYAL2 active site groove. With the HA tetramer in place, the substrate distortion that occurs in the -1 subsite of retaining glycosidases (Davies, Planas et al. 2012) was modelled by placing the non-reducing GlcNAc of a correctly distorted chitobiose disaccharide (from a crystal structure of a chitobiase from *Serratia marcescens* (PDB ID: 1QBB) (Tews, Perrakis et al. 1996) into the -1 subsite of HYAL2, replacing the terminal GlcNAc of the HA tetramer. This reducing GlcNAc was converted into a GlcUA sugar. An additional GlcNAc was added to the GlcUA in the +1 subsite via a (1->3) linkage and fit into the putative +2 subsite of HYAL2, which was found to be proximal to Lys148.

3.4.13.5 Transient transfection and western blot analysis

For transfection, pCMV6-HYAL2, pCMV6-HYAL2K148R, pIRES_HYAL1, pIRES_HYAL1K144R, or empty vector were co-transfected with pRcCMV β -gal using GenJet reagent (SignaGen) or Turbofect (Thermo Fisher) following the manufacturer's instructions. Transfections of HYAL2-expressing vectors were performed into mouse embryonic fibroblasts (MEFs) derived from a Hyal2^{-/-}

embryos and HYAL1-expressing vectors into MEFs derived from Hyal1^{-/-} embryos (Hemming, Martin et al. 2008). Cells and culture medium were collected 48 hours post-transfection and the β -galactosidase activity was determined using O-nitrophenol- β -D-galactopyranoside as a substrate (Rosenthal 1987). Protein concentrations in the lysates were determined using a BioRad protein assay kit based on the Bradford method, and β -galactosidase levels were used to normalize for transfection efficiency in all assays. For the assessment of HA binding, 50 μ g of rooster comb HA were added to 10-30 μ g of protein lysate prepared in RIPA buffer (2% Igepal, 1% sodium deoxycholate, 0.2% SDS in pBS). Complexes were precipitated with an equal volume of 2% hexadecylpyridinium chloride (CPC), collected by centrifugation, and washed with 1% CPC before preparation of the sample and separation on an SDS-PAGE for western blot analysis. Western blot analysis was performed as described previously, but using polyclonal anti-HYAL2 (Proteintech) or monoclonal anti-HYAL1 antibody, 1D10 (Hemming, Martin et al. 2008). For immunofluorescence, the cells were re-plated at a lower concentration 24 h post-transfection, and HYAL2 was detected using a 1/250 dilution of anti-HYAL2 and an alexafluor 568 conjugated secondary antibody, following previously described procedures (Hemming, Martin et al. 2008).

4 Mutations in *HYAL2*, encoding hyaluronidase 2, cause a syndrome of cor triatriatum sinister and orofacial clefting in humans and mice

4.1 Summary

Orofacial clefting is amongst the most common of birth defects, with both genetic and environmental components. Although numerous studies have been undertaken, the genetic etiology of this heterogeneous condition remains incompletely understood. Congenital heart disease (CHD) is commonly associated with syndromic orofacial clefting. Cor triatriatum sinister is a rare congenital cardiac anomaly characterized by division of the left atrium. To date, no molecular cause has been identified for cor triatriatum in humans.

Five Amish children were identified with a syndrome including orofacial clefting and cor triatriatum sinister. Using a combined approach of autozygosity mapping and exome sequence analysis, a likely pathogenic variant was identified in *HYAL2* encoding hyaluronidase 2 (K148R), a gene not previously associated with human disease. *HYAL2*, encoding hyaluronidase 2, degrades extracellular hyaluronan, a critical component of the developing heart and palatal shelf matrix. Transfection assays demonstrated that the gene variation destabilizes the molecule, dramatically reducing *HYAL2* protein levels. Consistent with the clinical presentation in affected individuals, investigations of *Hyal2*^{-/-} mice revealed craniofacial abnormalities, including submucosal cleft palate and cor triatriatum sinister in a proportion of *Hyal2*^{-/-} mice. Taken together these findings identify a new genetic cause of orofacial clefting in humans and mice and define the first molecular cause of human cor triatriatum sinister, illustrating the fundamental importance of *HYAL2* and hyaluronan turnover for normal human and mouse development.

4.2 Introduction

4.2.1 Cor triatriatum sinister

Cor triatriatum (atrial heart) is a rare congenital anomaly which has been reported in 0.1-0.4 % of individuals with congenital heart disease (Jegier, Gibbons et al. 1963, Talner 1998). It is characterized by division of either the left (cor triatriatum sinister) or less commonly the right atrium (cor triatriatum dexter). Cor triatriatum sinister was first described in 1868 by Church (Church WS. Congenital malformation of the heart: Abnormal septum in left auricle. 1868; 19:188-90) and since then a number of variants have been reported. At the core of the classic lesion is an oblique fibromuscular membrane dividing the left atrium into a superior and inferior compartment. The pulmonary veins usually drain into the superior part and the vestibule and atrial appendage are part of the inferior compartment. The body of the atrium may be divided in various ways between the two parts and the oval fossa opens either into the superior or inferior compartment (Bharucha, Spicer et al. 2015).

The majority of cases of cor triatriatum sinister are associated with other cardiac lesions including atrial septal defects, patent ductus arteriosus and anomalous pulmonary venous drainage (Humpl, Reineker et al. 2010). Clinical manifestations and age of presentation are dependent on the degree of obstruction between the proximal and distal compartments (determined by the number and size of openings in the membrane) and the concomitant occurrence of other cardiac anomalies. Isolated cor triatriatum sinister may present with symptoms of pulmonary congestion and pulmonary arterial hypertension, and in younger patients failure to thrive or, if associated with other cardiac anomalies, cyanosis (Humpl, Reineker et al. 2010). Without significant obstruction, cor triatriatum sinister may, however, also be an incidental finding in adults (Horowitz, Zager et al. 1993) . The lesion can readily be diagnosed with transthoracic echocardiography and if there is obstruction to blood flow surgical resection of the membrane is the treatment of choice.

A number of theories have been proposed to explain the morphogenesis of cor triatriatum sinister, but so far no unifying explanation for the variations of the divided left atrium has been found (Bharucha, Spicer et al. 2015) and no molecular cause for cor triatriatum has been described. One of the suggested theories depends on the persistence of the left superior caval vein which may impinge on the left atrium and cause the growth of the abnormal membrane (Gharagozloo, Bulkley et al. 1977). However, although there are cases associated with persistence of the left superior caval vein, which runs obliquely across the dorsal wall of the left atrium parallel to the dividing shelf and drains into the right atrium via an enlarged coronary sinus, the majority of patients with cor triatriatum sinister do not have a persistent superior left caval vein. In addition, there is no sign of encroachment of the left atrium during embryogenesis in mice, who retain the left superior caval vein as part of their normal development (Bharucha, Spicer et al. 2015). Other proposed theories depend on malseptation, malincorporation or entrapment of the pulmonary vein (Van Praagh and Corsini 1969). It has been shown that the pulmonary vein initially develops as a solitary channel adjacent to the left atrio-ventricular junction. Subsequently the pulmonary venous component of the left atrium enlarges. Abnormal incorporation of the pulmonary vein could explain cases where the oval fossa drains into the inferior chamber, but it would be more difficult to explain cases with drainage of the oval fossa into the superior chamber (Bharucha, Spicer et al. 2015). In addition, there are cases of cor triatriatum sinister associated with partially or totally anomalous pulmonary venous drainage (Thilenius, Bharati et al. 1976). A unifying explanation for the morphogenesis of cor triatriatum sinister has therefore not been established yet and the discovery of the first molecular cause may contribute to elucidating the pathways that are involved.

In two of the affected children in this study, a dilated coronary sinus consistent with a persistent left superior vena cava was detected without any obvious signs of cor triatriatum on echocardiography. Whereas it is normal for mice to have a left sided superior vena cava, there is regression of the left superior cardinal vein caudal to the

innominate vein during normal human development. Only the failure of this regression leads to the persistence of a left sided superior vena cava (PLSVC). A PLSVC represents the most common congenital thoracic venous anomaly with a prevalence of 0.3–0.5% in the general population (Wood PH. Diseases of the Heart and Circulation. 2 ed. Philadelphia: J. B. Lippincott; 1956 1956). In 80–90% of individuals, the PLSVC drains into the right atrium via the coronary sinus, which leads to a dilated coronary sinus, but usually has no hemodynamic consequence (Goyal, Punnam et al. 2008).

4.2.2 Orofacial clefting

Orofacial clefts, which include cleft lip with or without cleft palate (CL, CL/P) and cleft palate alone (CP), are amongst the most common birth defects in all populations worldwide (Mossey, Little et al. 2009). The average incidence of CL/P is 1 in 700 new-borns with the highest occurrence rates in Asia and the lowest in Africa (Wyszynski, Lewanda et al. 1996, Rahimov, Jugessur et al. 2012). Affected individuals have difficulties feeding after birth, will often develop conductive hearing loss and experience speech and dental problems. Patients commonly undergo multiple craniofacial and dental surgeries, speech and hearing therapies and may experience life-long social and psychological effects from the malformation.

Orofacial clefts result from a failure of the complex processes involved in craniofacial development including cell growth, migration, differentiation and apoptosis. All human embryos start off with clefts of the lips and palate which fuse to result in the normal structures. The development of the face begins very early in embryogenesis (week 4) with the migration of neural crest cells to form the five facial primordia: the frontonasal prominence, the paired mandibular processes and the paired maxillary processes. By week 5 the paired medial and lateral nasal processes are formed and by week 6 the medial nasal processes have merged with the maxillary processes to form the upper lip and primary palate. The secondary palate begins to develop in week 7 when bilateral palatal shelves emerge from the maxillary processes, which initially grow vertically downwards, but later elevate into

a horizontal plane above the tongue and fuse in the middle to form the secondary palate. The secondary palate is complete by week 12 and separates the oral and nasal cavities (Jiang, Bush et al. 2006, Gritli-Linde 2007).

CL/P cases are categorized according to the presence or absence of additional congenital anomalies in syndromic (SCL/P, 30%) and non-syndromic CL/P (NSCL/P, 70%) (Beriaghi, Myers et al. 2009, Luijsterburg, Rozendaal et al. 2014). Almost 300 different syndromes in which CL/P is a prominent feature have been described (www.ncbi.nlm.nih.gov/OMIM) and in about 75% of these the genetic cause has been established (Leslie and Marazita 2013). Congenital heart diseases (CHD) are amongst the most frequent anomalies associated with CL/P with a reported incidence of 1.3 to 27% (Pannbacker 1968, Shah, Pruzansky et al. 1970, Mackeprang and Hay 1971, Geis, Seto et al. 1981, Voisin, Montoya et al. 1981). In a study of 78 patients with CL/P and CHD, conotruncal defects were the most common with Tetralogy of Fallot present in 24% of cases (Wyse, Mars et al. 1990). Underlying causes for the CL/P associated with CHD include mutations in single genes, chromosomal anomalies and teratogens with neural crest cells likely being the common factor as they are crucial in the development of both tissues. It has also been speculated, that the close proximity of the cardiac and facial region during embryogenesis may explain the high incidence of combined cardiac and facial malformations. For example, individuals with DiGeorge syndrome, also known as 22q11.2 deletion or velocardiofacial syndrome, often exhibit CL/P and CHD. Deletion of the gene *TBX1* within the 3Mb deleted region is thought to be responsible for the major phenotypes seen in this syndrome (Yagi, Furutani et al. 2003).

4.2.3 Hyaluronan and hyaluronidase 2

Hyaluronan (HA) is a high molecular mass, linear glycosaminoglycan which is ubiquitously present as a major component of the extracellular matrix (ECM) of many vertebrate tissues including skin, heart valves, vitreous of the eye and synovial fluid. The viscoelastic properties of HA influence the constitution of the ECM and are important for cell proliferation, differentiation and migration (Chowdhury, Hemming

et al. 2013). HA is central during embryogenesis both for palatal and cardiac development. HA expansion is critical during palatal shelf reorientation (Richman and Lee 2003) and high levels of HA have been detected in the developing heart (Camenisch, Spicer et al. 2000). In addition, a deficiency of HA synthesis in hyaluronan synthase 2 knock-out mice has been shown to lead to loss of endocardial cushion swelling and a lack of epithelial-to-mesenchymal transition with subsequent death of the mice at embryonic day 9.5 (Camenisch, Biesterfeldt et al. 2001, Camenisch, Schroeder et al. 2002). There are six human hyaluronidase-related genes (*HYAL1*, *HYAL2*, *HYAL3*, *HYALp1*, *HYAL4*, and *SPAM1*) (Csoka, Scherer et al. 1999), but only *HYAL1* and *HYAL2* encoding hyaluronidase 1 (*HYAL1*) and hyaluronidase 2 (*HYAL2*) are considered to have a significant role in HA degradation (Stern 2004). *HYAL2* is a glycosylphosphatidylinositol (GPI) anchored, lipid raft associated cell surface protein (Andre, Duterme et al. 2011), with weak activity towards high molecular mass HA (Lepperdinger, Mullegger et al. 2001). High molecular mass HA is bound to cell surfaces by HARE (hyaluronan receptor for endocytosis) (Falkowski, Schledzewski et al. 2003) and CD44 receptors, from which *HYAL2* is thought to commence HA degradation into intermediate size fragments (Bourguignon, Singleton et al. 2004). The intermediate size fragments are then endocytosed and further degraded to oligosaccharides in lysosomes by *HYAL1* and exoglycosidases (Gushulak, Hemming et al. 2012). In addition, *HYAL2* interacts with tumour-growth factor- β to mediate some of its intracellular actions (Hsu, Schultz et al. 2009), acts as receptor for ovine oncogenic retroviruses (Miller 2008) and, in platelets cleaves the extracellular HA cables that bind leucocytes (de la Motte, Nigro et al. 2009).

Despite the proposed role of *HYAL2* in HA degradation, only a deficiency in *HYAL1* has previously been associated with a human disorder; mucopolysaccharidosis IX. Only four affected individuals have been described with this condition, in which phenotypic features include periarticular soft tissue masses, joint effusions, synovial proliferation, and a bifid uvula. In affected individuals plasma HA levels were markedly elevated, there was complete deficiency of plasma hyaluronidase activity

and pathological storage of mucopolysaccharide in lysosomes of histiocytes and fibroblasts (Natowicz, Short et al. 1996) (Imundo, Leduc et al. 2011) *HYAL1* mutations, inherited in an autosomal recessive fashion, were identified in these patients (Triggs-Raine, Salo et al. 1999) (Imundo, Leduc et al. 2011). Previous studies of *Hyal2*^{-/-} mice revealed craniofacial abnormalities, chronic anaemia, heart valve thickening, unexplained pre-weaning lethality and in a proportion of mice, atrial dilatation and pulmonary fibrosis with accumulation of HA in the extracellular matrix (Jadin, Wu et al. 2008, Chowdhury, Hemming et al. 2013).

4.3 Study-specific methods

4.3.1 Exome sequencing

Whole exome sequencing was performed by the Otogenetics Corporation using the SureSelect Human All Exon V4 (Agilent Technologies) exome enrichment kit on an Illumina HiSeq2000. Support for the analysis of the exome sequencing data was provided by Dr Barry Chioza and Prof Michael Wheedon, based at University of Exeter.

4.3.2 Functional studies

Functional studies were performed to assess the impact of p.K148R and p.P250L on the HYAL2 protein, including the introduction of the variants into a human HYAL2 construct and transfecting it into mouse embryonic fibroblasts, western blot analysis, immunofluorescence microscopy and activity assays in collaboration with Professor B Triggs-Raine's group, University of Manitoba, Winnipeg, Canada.

4.3.3 Mouse studies

Mouse studies were designed by Prof Triggs-Raine and Dr B Chowdhury based at University of Manitoba, Winnipeg, Canada. Following discussion of the human phenotype with myself, Dr Chowdhury performed detailed analysis of the mouse phenotype.

4.4 Results

4.4.1 Summary of phenotype data

4.4.1.1 Amish family

During recruitment of individuals for the HCM study, we became aware of a family with two siblings with CLP. After further questioning it transpired that three distant cousins who lived quite closely, were also affected. Given the occurrence of multiple cases in one generation, we suspected that a recessive founder mutation was responsible. CLP is often associated with congenital heart disorders, which prompted us to evaluate the children from a cardiological point of view even though none displayed any cardiac symptoms. This led to the identification of the associated cardiac features in these children with CLP.

Altogether five individuals aged between 4 and 16 years affected by a syndromic form of CLP were identified within a single extended Amish pedigree (family 1). Four individuals exhibited surgically corrected bilateral CLP and one individual unilateral cleft lip on the right. Craniofacial features included frontal bossing, hypertelorism, widened and flattened nasal bridge and tip, thickened helices and micrognathia. Three out of five children had pectus excavatum. Single palmar creases were noted and 2–3 soft tissue syndactyly of the toes was found in the affected individuals. However, syndactyly was also seen in unaffected members of the family.

On echocardiography, one individual showed non-obstructive cor triatriatum dividing the pulmonary venous confluence from the body of the left atrium with an anterior fenestration in the membrane. Two individuals had dilated coronary sinuses consistent with persistent left superior venae cavae. In one child both atria were dilated, and one other individual had a mildly dilated left atrium. The oldest individual (aged 16) showed some minor aortic valve thickening with slightly restricted motion of the right cusp.

Four individuals had varying degrees of myopia, which was corrected with glasses. Two out of these four individuals have significant anisometropia (myopia in one eye, hyperopia in the other eye), one of whom has been diagnosed with a staphyloma (protrusion of uveal tissue through a weak point in the eyeball) in the severely myopic eye with myopic degeneration of the retina and a posterior subcapsular polar cataract. Four of the children had varying degrees of conductive hearing loss presumably secondary to recurrent otitis media in childhood and one individual had severe bilateral sensorineural hearing loss with a cochlear implant in one ear. All affected individuals had normal height and occipital frontal circumference (OFC) for their age and mostly reached their developmental milestones at the appropriate time apart from the individual with the cochlear implant who had slightly delayed speech and academic performance (table 4-1 for full clinical characteristics).

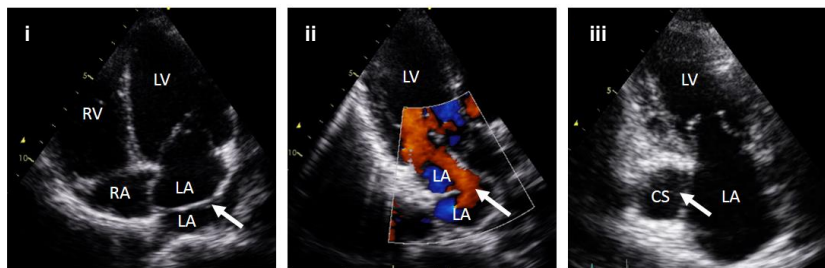


Figure 4-1: Transthoracic echocardiography of two affected Amish individuals

i: Apical four chamber view of Amish individual XII:3; arrow indicates the membrane across the left atrium dividing the pulmonary venous confluence from the body of the left atrium (cor triatriatum sinister).

ii: Apical three chamber view with colour Doppler of individual XII:3; arrow indicates the anterior fenestration in the membrane dividing the left atrium; the laminar flow through the fenestration implies there is no obvious obstruction to flow across the membrane.

iii: Apical two chamber view of individual XII:9; arrow indicates the dilated coronary sinus indicative of a persistent left superior vena cava.

RV: right ventricle, RA: right atrium, LV: left ventricle, LA: left atrium, CS: coronary sinus

4.4.1.2 Saudi Arabian family

During our investigation, we established a collaboration with Professor Fowzan S. Alkuraya's group at the Department of Genetics, King Faisal Specialist Hospital and Research Center, Riyadh, Saudi Arabia, who had previously described two children aged 7 and 12 years from an extended consanguineous Arab family from Northern Saudi Arabia with a very similar phenotype to the children in our study (family 2) (Shaheen, Patel et al. 2016). Most notably, the presence of bilateral cleft lip and palate in one affected girl and a similar facial appearance with frontal bossing, hypertelorism, micrognathia, and a broad flat nasal bridge. Both girls also had bilateral single palmar creases, lens opacities, staphyloma and were myopic. Echocardiography in the younger girl revealed an abnormal mitral valve with accessory tissue and a small restrictive peri-membranous ventricular-septal defect covered with tricuspid aneurysmal tissue. The older affected girl had pectus excavatum and a normal echocardiogram (table 4-1 for full clinical characteristics).

	FAMILY 1					FAMILY 2	
INDIVIDUAL	XII:3	XII:5	XII:7	XII:9	XII:12	V:3	VI:2
GENDER	Male	Male	Female	Female	Male	Female	Female
HYAL2 MUTATION	c.443A>G	c.443A>G	c.443A>G	c.443A>G	c.443A>G	c.749C>T	c.749C>T
AGE (YRS)	12.3	15.5	4.3	9.9	12.7	7.1	12.3
GROWTH Height (SDS)	-0.54	0.45	-1.07	0.29	-0.40	-2.87	-2.87
Head circumference (SDS)	-0.67	-0.22	-0.28	-0.49	-1.60	-3.13	-2.59
SKELETON Pectus excavatum	-	✓	✓	-	✓	-	✓
CRANIOFACIAL FEATURES	Bilateral CLP Not known	Rt unilateral CLP Not known	Bilateral CLP ✓ ✓ ✓ ✓ -	Bilateral CLP - ✓ ✓ ✓ Over folded superior helices	Bilateral CLP ✓ ✓ ✓ ✓ -	None - ✓ ✓ ✓ Cupped ears	Bilateral CLP - ✓ ✓ ✓ Cupped ears
HEART Congenital cardiac malformations	Lt cor triatriatum	Dilated coronary sinus consistent with persistent LSVC	Normal Echocardiogram	Dilated coronary sinus consistent with persistent LSVC	Normal Echocardiogram	Abnormal MV with accessory tissue, small restrictive perimembranous VSD covered with tricuspid aneurysmal tissue	Normal Echocardiogram
Other cardiac abnormalities		Thickened AV, restricted Rt aortic leaflet, mild AR		Dilated left and right atrium		Mild TR, Lt-Rt shunt and mildly dilated LV	

EYES <i>Visual acuity</i>	Glasses since age 2; Rt myopia -8.00 Lt hyperopia +3.25	Glasses since age 6; Rt myopia -1.50 Lt hyperopia +1.75	Normal	Glasses since age 9	Glasses since age 9 Rt myopia -2.50 Lt myopia -2.50	Glasses since age 3 Rt myopia -16.75 Lt myopia -15.00	Glasses since age 4 Rt myopia -9.75 Lt myopia -16.00
<i>Other ocular abnormalities</i>	Rt staphyloma, myopic macular degeneration, posterior subcapsular polar cataract/ Mittendorf dot	-	-	-	-	Cortical lens opacities and staphyloma with small disc	Corneal opacity flat retina staphyloma and left Retinal scar
HEARING <i>Conductive hearing loss</i>	Rt moderate; postlingual	-	Bilateral mild	-	Lt moderate; prelingual	Rt Mild	-
<i>Sensorineural hearing loss</i>	-	Bilateral severe to profound; prelingual; Lt cochlear implant	-	-	-	-	-
NORMAL PSYCHOMOTOR DEVELOPMENT	✓	✓	✓	✓	✓	✓	✓
HANDS AND FEET	Not known	Not known	Single palmar crease Lt hand Bilateral 2-3 and 3- 4 skin syndactyly of toes	Bilateral 2-3 and 3-4 skin syndactyly of toes	Bilateral single palmar creases Bilateral 2-3 skin syndactyly of toes	Bilateral single palmar creases Broad thumbs; Acromelia,	Bilateral single palmar creases

Table 4-1: Clinical characteristics of individuals homozygous for pathogenic HYAL2 variants

4.4.2 Identification of putative pathogenic variants in *HYAL2*

4.4.2.1 Amish family

Assuming that a recessive founder mutation was responsible, a genome-wide SNP study in affected members of family 1 was undertaken in order to map the chromosomal location of the disease gene. This identified a single notable (>1Mb) shared autozygous region of 10.18Mb on chromosome 3p21.31 as likely to correspond to the disease locus. The region is delimited by SNPs rs6441961 and rs2201057 (chr3:46310893-56499374[hg38]) and is predicted to contain 235 RefSeq alignment-supported transcripts (Fig 4-1). Autozygosity of this genomic region specifically within affected family members was confirmed by microsatellite marker analysis (data not shown). In support of this, multipoint linkage analysis with Simwalk2 (Sobel, Sengul et al. 2001) assuming autosomal recessive inheritance, full penetrance and a disease allele frequency of 0.0001, achieved a LODMAX = 10.37 corresponding to the autozygous interval.

To identify the causative mutation, whole exome sequence analysis of a single affected individual (XII:3) was undertaken (Otogenetics Corporation). After filtering for call quality, potential pathogenicity, population frequency and localization within the candidate interval, we identified only a single likely deleterious variant located in the *HYAL2* gene (*HYAL2* chr3:g.50320047T>C [hg38]; NM_003773.4:c.443A>G; p.K148R; PolyPhen = probably damaging (1); SIFT = deleterious (0.03) MutationTaster = disease causing (probability 0.999997). The variant was validated by Sanger dideoxy sequencing, which was also used to confirm co-segregation of the variant within the extended Amish pedigree (figure 4-2 A). Seven heterozygous carriers of the c.443A>G variant were identified in 532 Amish control chromosomes examined, indicating a carrier frequency of approximately 1:76 in the population (allele frequency 1:152). Using the Hardy-Weinberg principle, this translates into 1:23 000 homozygous and therefore affected individuals. However, this number may be different in the Amish population given the innate characteristics of a closed

founder population, the limited mobility of individuals and tendency to settle locally, and large number of children, giving rise to potentially higher chances of homozygosity and clustering of larger number of cases. The variant was not listed in the Exome Variant Server (NHLBI GO Exome Sequencing Project (ESP), Seattle, WA; <http://evs.gs.washington.edu/EVS/>), 1000 Genomes browser (<http://browser.1000genomes.org/index.html>), the Exome Aggregation Consortium (ExAC) database (<http://exac.broadinstitute.org/>) or gnomAD database (<http://gnomad.broadinstitute.org/>).

4.4.2.2 Saudi Arabian family

Interestingly, *HYAL2* had recently been identified as a candidate gene as a cause of short stature and facial dysmorphism in a study of multiplex consanguineous families (Shaheen, Patel et al. 2016). This study undertook genetic studies in 33 heterogeneous families with dysmorphic and other systemic clinical manifestations and identified a sequence alteration in *HYAL2* as a possible cause of the clinical features in a single family comprising two affected individuals (figure 6-2). Although inconclusive, when considered together with the genetic, clinical and functional data from our Amish study, this permits the robust genetic definition and clinical description of a new form of syndromic CLP. Genetic studies in Family 2 comprised whole genome SNP mapping which identified a single autozygous interval of 13.49Mb on chromosome 3p21 (rs7650433-rs9811393; chr3:40857673–54345775 [hg38]) shared by two affected family members, in which a missense *HYAL2* variant (chr3:g.50319741G>A [hg38]; NM_003773.4: c.749C>T; p.P250L in Family 2; (figure 6-2 A, B) absent in 817 ethnically matched individuals, was identified as the candidate genetic cause. This variant (rs781999115) was not listed in the Exome Variant Server or the 1000 Genomes browser but two carriers were reported on the ExAC database and three carriers in gnomAD database (<http://gnomad.broadinstitute.org/gene/ENSG00000068001>). The genomic autozygous region (chr3:46310893–54379802 [hg38]) common to both family

cohorts is 8.0Mb and predicted to contain 270 RefSeq alignment-supported transcripts.

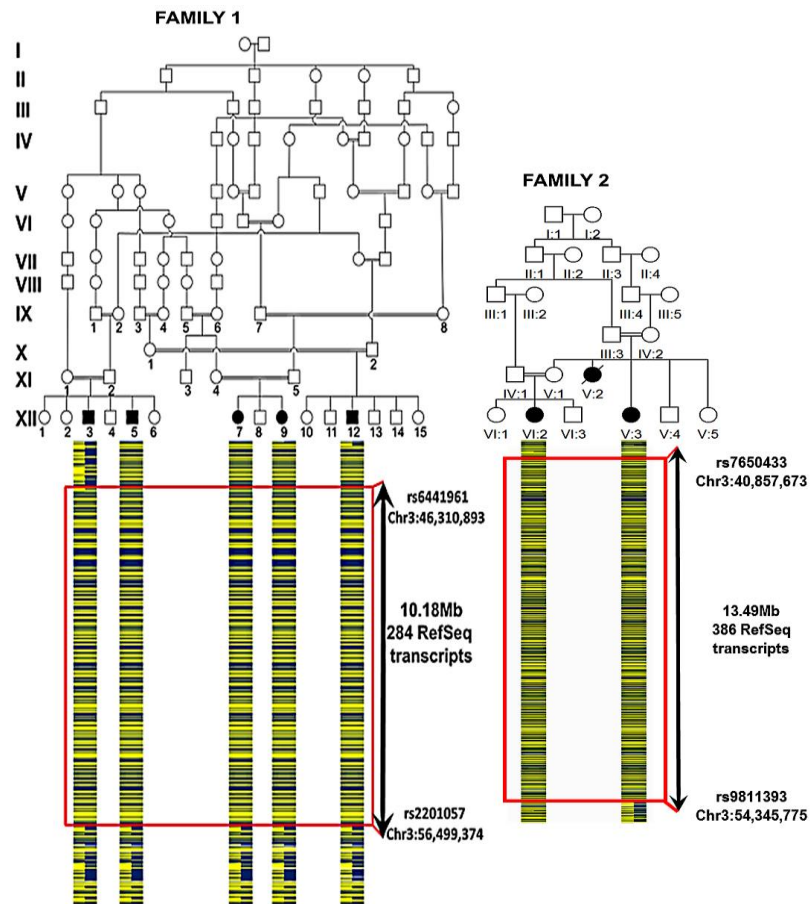


Figure 4-2: Pedigrees and graphical representation of genotypes in region of shared autozygosity across

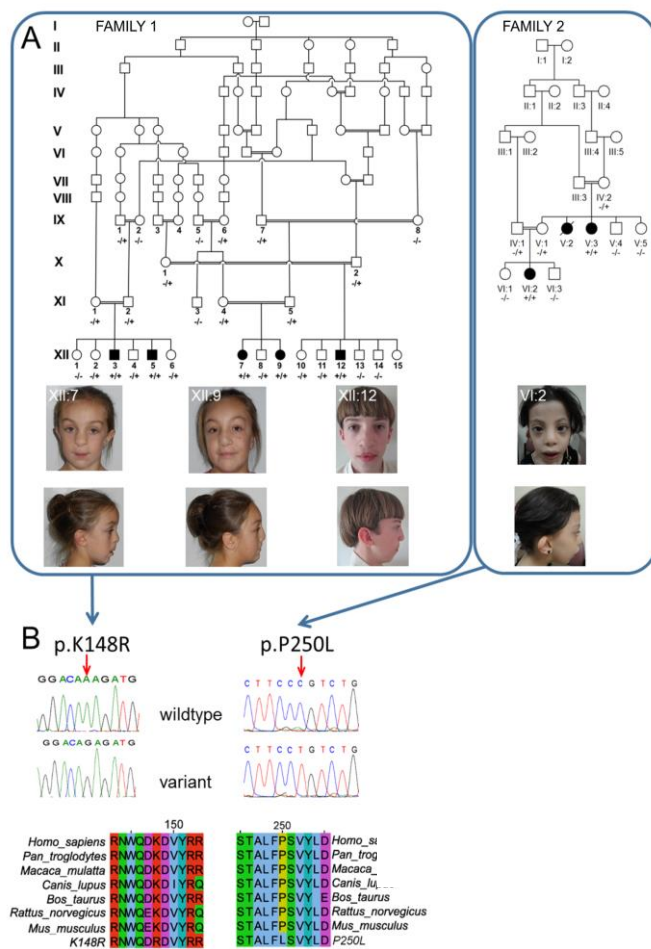


Figure 4-3: Pedigrees, clinical features of individuals homozygous for HYAL2 mutation and identified HYAL2 mutations

(A) Pedigree diagrams and facial phenotype of individuals (Amish Family 1: XII:7; XII:9; XII:12 and Saudi Family 2: VI:2) with HYAL2 deficiency. Note the craniofacial similarities including frontal bossing, hypertelorism, widened nasal bridge, flattened broad nasal tip and cupped ears/over-folding of the superior helices. Consent for publication of these photographs was obtained

(B) Electropherograms showing the identified c.443A>G & c.749C>T mutations and conservation of protein sequence across species.

4.4.3 *HYAL2* p.K148R and p.P250L significantly reduce *HYAL2* protein levels

The functional studies described below were undertaken in collaboration with Dr Triggs-Raine (Winnipeg, Canada).

HYAL2 encodes hyaluronidase 2, a membrane localized protein (Andre, Duterme et al. 2011) with weak activity towards hyaluronic acid (HA) (Lepperdinger, Mullegger et al. 2001), which is an extracellular matrix glycosaminoglycan that is abundant during development. To assess the impact of p.K148R and p.P250L on the *HYAL2* protein, the mutations were each introduced into a human *HYAL2* expression construct and transiently transfected into mouse embryonic fibroblasts (MEFs) deficient in *HYAL2*. Western blot analysis revealed a profound effect of the p.K148R and p.P250L mutations on protein levels, resulting in 11- and 20-fold reductions, respectively, compared to wild type (Fig 4-4 A, B, F).

The residual K148R-containing *HYAL2* was readily precipitated with added hyaluronan (Fig 4-4 C, D), and immunofluorescence microscopy showed its subcellular localization was broadly similar to that of wild type *HYAL2*, which localized to both the membrane and cytoplasm (Fig 4-4 E). Although breakdown products do appear in extracts of the p. K148R-containing *HYAL2* protein over time, this data indicates that the K148R mutation likely results in partial *HYAL2* deficiency.

4.4.4 *HYAL2* p.K148R significantly reduces enzyme activity in *HYAL1*

In parallel with these studies we also investigated the effect of the p.K148R substitution on catalytic function. Unfortunately, due to the extremely low levels of activity known to be associated specifically with *HYAL2*, reliable activity assays are not available. We therefore took advantage of the almost complete functional identities between the active sites of *HYAL2*, and its more significantly active paralogue *HYAL1* (Fig. 4-5, A-B), for which such activity assays may be readily

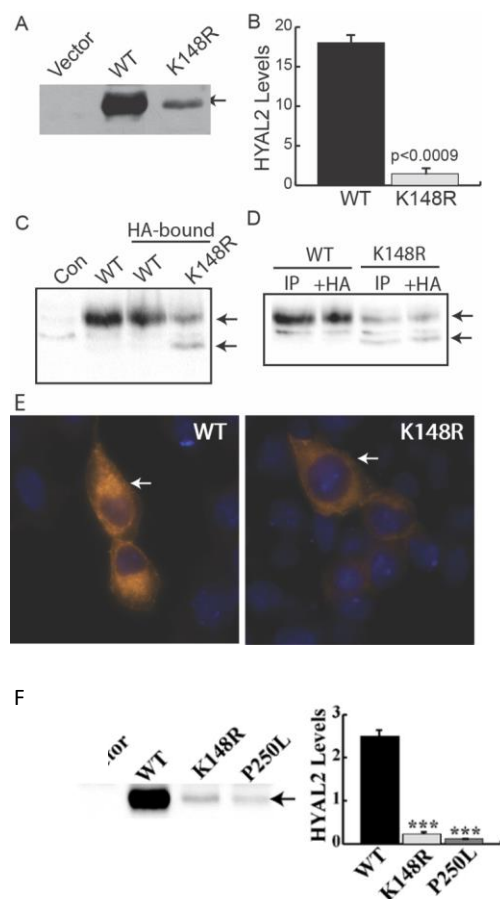


Figure 4-4: Expression, HA binding, and localization of wild type (WT) and K148R-HYAL2

MEFs deficient in HYAL2 were transfected with empty vector, pCMV6-HYAL2, pCMV6-HYAL2K148R, or pCMV6-HYAL2P250L

(A) Western blot analysis of expressed HYAL2 protein normalized to co-transfected β -galactosidase activity. An arrow indicates HYAL2.

(B) HYAL2 levels from WT or K148R-transfected cells were quantified by imaging the chemiluminescent signal using a BioRad ChemiDoc. The columns represent the average level of HYAL2 $\pm$ SEM (n=4). Significance was determined using the student's T test.

(C) Western blot analysis of HA bound HYAL2 from WT and K148R-HYAL2 transfection lysates compared to the HYAL2 signal detected in lysates from non-transfected (Con) and

WT HYAL2-transfected cells. Arrows indicate HYAL2; note the K148R bound HYAL2 degradation product.

(D) Western blot analysis comparing input (IP) levels of WT or K148R-HYAL2 compared to HA bound HYAL2 (+HA). Arrows indicate the position of HYAL2, again showing degradation products in the K148R HYAL2 samples.

(E) Localization of WT and K148R-HYAL2 (orange) in transfected MEF cells by immunofluorescence. The arrows indicate the membrane localization.

(F) Expression of wild type (WT), K148R and P250L-HYAL2. Western blots were performed on lysates prepared from MEFs deficient in HYAL2 that were transfected with empty vector, pCMV6-HYAL2, pCMV6-HYAL2K148R, or pCMV6-HYAL2P250L. An arrow indicates HYAL2. HYAL2 levels shown in the graph were quantified by imaging the chemiluminescent signal using a BioRad ChemiDoc. The columns represent the average level ($\times 10^6$ light units) of $\text{HYAL2} \pm \text{SEM}$ ($n = 4$). Significance was determined using the student's T test. *** indicates $p < 0.0001$.

Data courtesy of Professor B Triggs-Raine, University of Manitoba, Winnipeg, Canada

performed by introducing the equivalent missense mutation (p.K144R) into this functional paralogue. Wild type and $\text{HYAL1}^{\text{K144R}}$ mutant constructs were transiently expressed in MEFs deficient in HYAL1, and western blotting was used to confirm similar levels of expression and secretion (Fig 4-5, C). Neither the wild type nor $\text{HYAL1}^{\text{K144R}}$ exhibited activity at pH 6.0 or pH 8.0, indicating that the mutation had not altered the pH optimum of the enzyme (not shown). While significant residual activity remained, zymography and activity assays revealed diminished acid-active activity associated with the K144R substitution (Fig 4-5, D) at ~50% of the level of the wild type enzyme (Fig 4-5, E). As transient overexpression systems typically over-amplify residual activities (Mahuran 1999) the 50% activity detected in this assay is likely to significantly over-represent the actual in-vivo activity of the mutant molecule. Therefore, while such activity assays undertaken in the paralogous molecule should be interpreted with caution, this data together with the significant reduction in HYAL2 levels associated with the p.K148R HYAL2 substitution, indicates that *HYAL2* p.K148R is likely to result in a significant deficiency of HYAL2 activity.

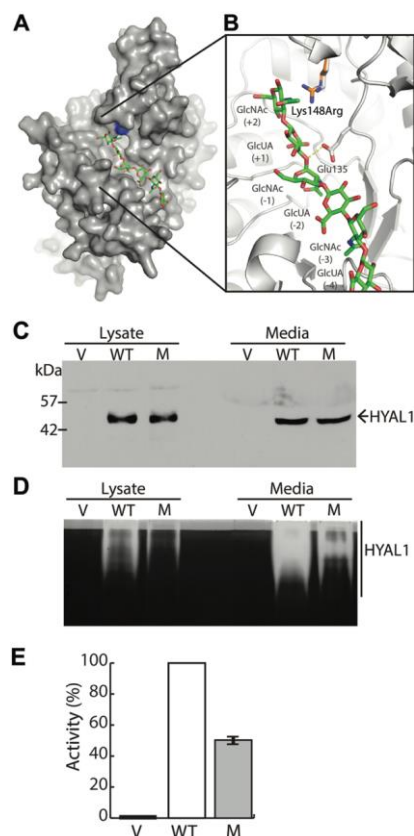


Figure 4-5: Comparative molecular model of human HYAL2 bound to a hexamer of hyaluronic acid

The model predicts the enzyme-substrate complex that is expected to form between HYAL2 and HA prior to glycosidic bond cleavage (which occurs between the +1 and -1 sugar binding subsites (nomenclature in (Davies, Wilson et al. 1997))).

(A) Overall surface representation of the HYAL2 model (shown in grey) with the HA hexamer bound within the active site groove (shown in stick format, green carbons). The location of Lys148, near the +2 sugar binding subsite is shown in blue.

(B) The model accounts for the distortion of the GlcNAc sugar in the -1 subsite towards a 1S3 conformation that is predicted to occur in GH56 glycosidases and has been experimentally determined to occur in many configuration retaining glycosidases. Accounting for the substrate distortion in the -1 subsite predicts the locations of the +1 and +2 subsites of Hyal2 and, interestingly, places the GlcNAc sugar bound to the +2 subsite to within $\sim 4\text{\AA}$ of Lys148 (shown in stick format, grey carbons), suggesting that this lysine residue comprises part of the +2 subsite. In light of this prediction, the Lys148Arg mutation (Arg shown in stick format,

orange carbons) would likely distort the shape of the +2 subsite and thereby reduce substrate binding affinity and catalytic efficiency of the enzyme. Sugars comprising the HA hexamer are labeled with their respective sugar-binding subsites indicated underneath in brackets. The predicted general acid/base residue of Hyal2, Glu135 (Glu131 in Hyal1), is shown in stick format (grey carbons), which forms a 2.7Å hydrogen bond (dashed yellow line) to the glycosidic oxygen of the scissile bond that spans the +1 and -1 subsites.

(C) Western blot analysis of K144R and wild type HYAL1. Lysates and media from transiently transfected HYAL1-deficient fibroblasts were analysed for the presence of HYAL1 (arrow).

(D) Zymography of lysates and media from transfected cells. The activity is indicated by the clearing of the gel and migrates as a smear in this non-denaturing gel likely because of interactions with other proteins.

(E) Hyaluronidase activity in the media of transfected cells.

Data courtesy of Professor B Triggs-Raine, University of Manitoba, Winnipeg, Canada

4.4.5 *HYAL2*^{-/-} mice exhibit a similar phenotype to humans with *HYAL2* mutations

4.4.5.1 Evaluation of the heart in *Hyal2*^{-/-} mice

Prof Triggs-Raine's group had previously investigated the phenotype in *Hyal2*^{-/-} mice. Skeletal abnormalities and severe pre-weaning lethality with only 1/3 of pups surviving to the stage of weaning had been described. Of the surviving *Hyal2*^{-/-} mice, 54% showed grossly enlarged left or right atria, and developed apparent rapidly progressive heart failure that required euthanasia at an average age of three months. The other 46% of *Hyal2*^{-/-} mice developed a more chronic form of heart failure, requiring euthanasia at an average age of 5.6 months. Histological and biochemical examination of the hearts in *Hyal2*^{-/-} mice revealed the accumulation of HA in heart valves and myocardial interstitium (Chowdhury, Hemming et al. 2013).

Following our findings of cardiac anomalies including cor triatriatum sinister in children with putative pathogenic *HYAL2* variants, we initiated a thorough assessment of cardiac structures using echocardiography in *Hyal2*^{-/-} mice. Ultrasound imaging as described previously (Corrigan, Brazil et al. 2010), revealed a centrally fenestrated membrane dividing the left atrium into a proximal and distal chamber in *Hyal2*^{-/-} mice. Half (7 of 14) of the *Hyal2*^{-/-} mice, including those with or

without significant atrial enlargement, had cor triatriatum sinister (Fig. 4-6 A, B). Pulse wave imaging revealed turbulent flow across the fenestration, suggesting a degree of obstruction. Examination of control mice revealed no evidence of cor triatriatum. In order to characterise the structural abnormalities in the hearts further, histological analyses were performed. Consistent with the ultrasound findings, membranous tissue was identified, that divided the left atrium of *Hyal2*^{-/-} mice in a proximal and distal compartment (Fig. 4-6 C–F) that stained strongly for hyaluronic acid (Fig. 4-6 G).

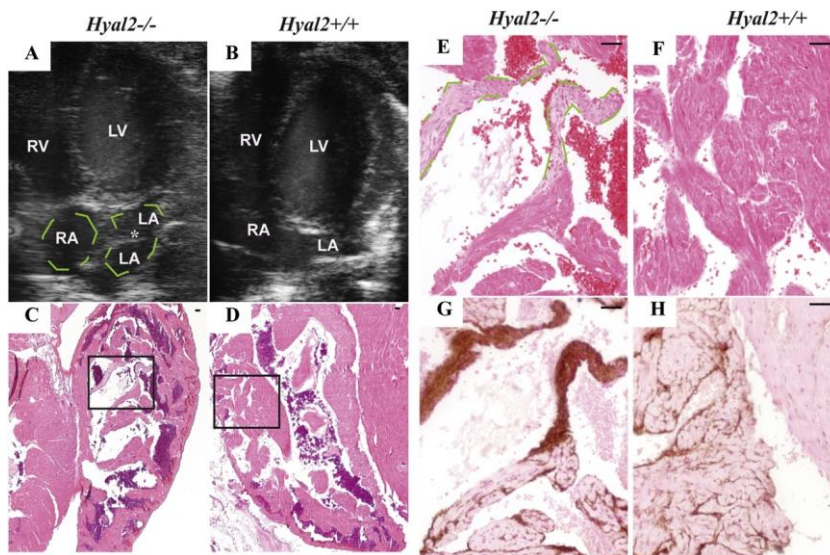


Figure 4-6: Cardiac abnormalities in *Hyal2*^{-/-} mice

A, B: High resolution (40 MHz) ultrasound images of *Hyal2*^{-/-} and control hearts at 4 weeks of age, apical 4-chamber view, (*) marks a septum in the left atrium of the *Hyal2*^{-/-} mouse resulting in cor triatriatum sinister (green dashes);

C, D: transverse sections of the left atrium (H&E stained) show valve-like tissue (black box) in the *Hyal2*^{-/-} mice (C) but not in control mice (D);

E, F: Enlarged view of the area in the box in C and D. A green dashed line outlines the valve-like structure in the atrium of a *Hyal2*^{-/-} mouse;

G, H: strong hyaluronic acid (HA) staining with the HA binding protein was shown in the valve-like structure in the *Hyal2*^{-/-} atrium (G) that is absent in the control atrium (H). Images are representative of 3 pairs of mice. Scale bar = 50 μ m

Data courtesy of Professor B Triggs-Raine, University of Manitoba, Winnipeg, Canada

4.4.5.2 Evaluation of craniofacial features and palate in *HYAL2*^{-/-} mice

HYAL2^{-/-} mice have previously been described to exhibit a short broad nose, increased interorbital space, and Wormian bones in the inter-frontal suture (Jadin, Wu et al. 2008), consistent with the facial dysmorphism in the human affected individuals. Wormian bones or intra sutural bones, are extra bone pieces that occur within a suture in the cranium. Examination of skull X-rays of affected children showed no evidence of Wormian bones, but frontal bossing was noted in all the affected individuals.

In order to investigate whether orofacial clefting present in the human affected individuals is also a feature of *Hyal2*^{-/-} mice, we examined the time of pre-weaning lethality previously described to occur in 2/3 of knockouts (Jadin, Wu et al. 2008). Orofacial clefting would impair sucking and therefore be a feasible cause of demise in the pre-weaning stage. Out of 201 progeny of *Hyal2*^{+/-} intercrosses followed from birth to weaning, 16 died at postnatal day (P)1, 16 died between P2-P9, and 16 *Hyal2*^{-/-} mice survived. Assuming all pups that died were *Hyal2*^{-/-}, 48 of 201 progeny (24%) would be *Hyal2*^{-/-}, approximating the Mendelian frequency for an autosomal recessive condition. We also examined ninety E18.5 embryos from *Hyal2*^{+/-} intercrosses and found two *Hyal2*^{-/-} and one *Hyal2*^{+/-} embryo(s) that had died at approximately E15.5. Considering these intra-uterine deaths, the proportion of *Hyal2*^{+/+} mice still exceeded that of *Hyal2*^{-/-} mice (30% vs 24%). Although this difference was not statistically different, it may suggest reduced survival of *Hyal2*^{-/-} and *Hyal2*^{+/-} embryos during early gestation (Supplemental data, table 7-1).

CL/P is a likely cause of death in *Hyal2*^{-/-} mice at P1 and possibly between P2-P9 due to impaired sucking, whereas heart defects are more likely to be causative at E15.5.

Although most deceased pups were cannibalized before their genotype could be assessed, four of the pups that were found dead at P1, without milk in their stomachs, were identified as *Hyal2*^{-/-}. Micro-CT studies of these pups revealed an under-developed and under-ossified viscerocranium compared to littermate controls (supplemental data, figure 7-1 i-ii). Several central palate bones were under-developed, and the vomer did not fuse centrally or form a head that articulated with the maxilla (supplemental data, fig 7-1 iii-vi). The ethmoid bone was nearly absent (supplemental data, fig 7-1 i- ii), or severely reduced in size (supplemental data, figure 7-2).

This underdevelopment of the viscerocranium indicated a defect in intramembranous ossification that was more prominent in the anterior midline and varied in severity among affected skulls (supplemental data, figure 7-2).

Characterization of the palates of *Hyal2*^{-/-} E18.5–19.5 embryos under a dissecting microscope revealed partial clefts and/or shortening of the secondary palate as well as abnormally formed rugae in most *Hyal2*^{-/-} mice and a small number of *Hyal2*^{+/-} mice (supplemental data, figure 7-3 A-D). Micro-CT confirmed reduced ossification and under-development of the viscerocranial bones, particularly the vomer, consistent with submucosal cleft palate (SMCP) in 15 of 18 (83%; 95% CI 58–96%) *Hyal2*^{-/-}, 3 of 54 (5.6%; 95% CI 1–16%) *Hyal2*^{+/-}, and 0 of 23 (0%; 95% CI 0–18%) *Hyal2*^{+/+} mice. Palatal malformation and clefting are likely to contribute significantly to the pre-weaning lethality in *Hyal2*^{-/-} mice, indicating further phenotypical overlap with the human condition.

Histological studies of P1 mice confirmed that the viscerocranial bones of all *Hyal2*^{-/-} mice were underdeveloped (supplemental data, figure 7-3 E and F, and arrows in N and O). Coronal sections showed the failed fusion between the epithelial surface of the vomero-nasal organ and the dorsal side of the palate shelf (figure 7-3 L and N), leaving an average gap of $472 \pm 50\mu\text{m}$ ($n = 3$) in palates of *Hyal2*^{-/-} mice that was absent in control palates (figure 7-3 M and O). HA levels were clearly increased in the *Hyal2*^{-/-} tissues (figure 7-3 J and N vs K and O). Excess mesenchymal cells and their

surrounding matrix were also evident, particularly in the anterior head (figure 7-3 H and I).

4.4.5.3 *HYAL2*^{-/-} mice have significant hearing loss

Most of the affected children in this study had conductive hearing loss which may be secondary to the orofacial clefting and recurrent otitis media. However, one affected individual was found to have profound sensorineural hearing loss. In order to assess hearing in *HYAL2*^{-/-} mice (8–12 weeks of age), the auditory brainstem response (ABR) was evaluated (supplemental data fig 7-4). *HYAL2*^{-/-} mice exhibited a significantly higher threshold at all frequencies tested, (mean $\pm$ 95% confidence interval for *HYAL2*^{-/-} and *HYAL*^{+/+} mice were 17.6 $\pm$ 7.16 and 5.00 $\pm$ 3.88dB respectively; Mann-Whitney U-test U = 1433, $p < 0.001$), indicating loss of hearing in all *HYAL2*^{-/-} mice. The ABR test parameters are unable to differentiate conductive and sensorineural hearing loss. However, these tests were performed in adult mice and severe craniofacial defects were excluded. The presence of hearing loss in mice without severe palate defects suggests further investigation is required to ascertain whether *HYAL2* is involved in the development and/or maintenance of normal hearing.

4.5 Discussion

Up to now, no molecular cause for cor triatriatum sinister has been described, and even though a number of theories have been proposed, no unifying explanation for the morphogenesis of cor triatriatum sinister has been established. The molecular basis of human cleft lip and palate (CLP) is incompletely understood, despite its frequent occurrence and associated morbidity.

In this study the investigation of multiple Amish families linking into an extended pedigree, led to the robust mapping of the likely disease locus by identification of a single autozygous region peculiar to affected family members only on chromosome 3p21.3. Exome sequence analysis identified a single candidate variant genome-wide

located in the critical interval in *HYAL2*, co-segregating as appropriate for an autosomal recessive disease in the Amish families.

Subsequently, a second (non-Amish) family was identified through our collaborative links in Saudi Arabia with a distinct mutation in *HYAL2* and clear similarities in the clinical phenotype. This study showed that the likely outcome of both mutations is instability of the mutant polypeptide resulting in notably reduced levels and loss of function of *HYAL2*.

These findings provide the first molecular explanation for cor triatriatum sinister in humans and mice and demonstrate, that a deficiency of the hyaluronic acid (HA)-degrading enzyme *HYAL2* is a novel cause of syndromic CLP in humans and submucosal cleft lip and palate (SMCP) in mice. All but one affected individual exhibited orofacial clefting with craniofacial dysmorphism being a consistent finding. Myopia and other ocular abnormalities, pectus excavatum, single palmar creases, cor triatriatum sinister, a persistent left superior vena cava and other cardiac features were variably present. These studies show that *Hyal2*^{-/-} mice universally exhibit facial dysmorphism, an under- development of the viscerocranium and cardiac valve thickening as well as variably penetrant hearing loss, partial or SMCP (83%), cor triatriatum sinister (50%) and atrial dilatation (50%). The significant similarities between the human and mouse phenotypes provide strong evidence that *HYAL2*-deficiency is the cause of these developmental abnormalities.

In the heart, fibroelastic tissues derived from endothelial cells form the heart valves and septum primum. The thickened valves and increased extracellular matrix in *Hyal2*^{-/-} mice suggest, that the presence of fibromuscular tissues in the *Hyal2*^{-/-} atrium originated during development and persists into adulthood.

The requirement for HA during cardiac development has been demonstrated in *HA synthase 2* deficient mice which die at E9.5 due to a disruption of normal cardiac morphogenesis (Lagendijk, Szabó et al. 2013). High molecular mass HA was confirmed to be critical for the formation of heart chambers and epithelial to

mesenchymal transition (EMT) of the cardiac cushion. Low molecular mass HA was shown to attenuate EMT and promote partitioning of the heart chambers through a vascular endothelial growth factor mediated pathway (Lagendijk, Szabó et al. 2013). These findings suggest that hyaluronidase is involved in the break-down of high molecular mass HA and/or the production of low molecular mass HA to regulate EMT and heart development. HYAL2 is usually highly expressed in the developing heart, and excess HA is found in the tissues of Hyal2^{-/-} mice (Chowdhury, Hemming et al. 2013). The sustained and increased presence of HA in the developing heart of Hyal2^{-/-} mice may result in excess EMT to produce the fibromuscular tissue that promotes the formation of triatrial heart.

This study demonstrates that the likely loss of HYAL2 activity, resulting in higher levels of HA, increases the risk of formation of triatrial heart. This is the first description of a molecular cause of cor triatriatum sinister in both mice and humans. The Hyal2^{-/-} mice provide an excellent model to investigate approaches to manage triatrial heart and to identify factors that may modify the risk for this malformation. Levels of HA may be increased in response to a number of growth factors and other influences, like high glucose levels, which increase the availability of the sugar precursors to produce HA in some tissues. By examining the impact of factors that increase or decrease HA in the context of HYAL2 deficiency it may be possible to identify risk factors for triatrial heart and possibly other congenital heart abnormalities in the developing human embryo.

These data imply that the differences between the human and mouse phenotypes may be largely explained by the complete absence of HYAL2 in Hyal2^{-/-} mice, whereas the destabilizing and deactivating p.K148R and p.P250L HYAL2 substitutions identified in affected humans lead primarily to profound reductions in HYAL2 levels, which may permit limited residual functionality. Progressive accumulation of HA in the circulation, pulmonary fibrosis, and cardiac dysfunction lead to premature heart failure in surviving Hyal2^{-/-} mice (Chowdhury, Hemming et al. 2013). The absence of elevated circulating HA in the affected individuals may partially or completely

protect them from these long-term cardiopulmonary complications. The residual activity associated with the amino acid substitutions in human subjects likely supports ongoing constitutive turnover of HA, however, is insufficient for the rapid and regulated turnover required during development. Further support for this is provided by a study of *Hyal2*^{+/-} mice where a small number showed developmental abnormalities in the absence of elevated circulating HA (Chowdhury, Hemming et al. 2013). This may, however, be an overly simplistic interpretation as circulating HA levels are approximately 10–20 fold higher in the mouse, despite similar rates of turnover in mice and humans (Jadin, Bookbinder et al. 2012). The identification of additional *HYAL2*-deficient individuals may provide a more complete understanding of the phenotypic spectrum associated with *HYAL2*-deficiency.

SMCP was identified as a likely cause of early pre-weaning lethality in *Hyal2*^{-/-} mice. The under-developed viscerocranium likely weakens the palate, and in combination with the facial dysmorphism, results in significant feeding difficulties for *Hyal2*^{-/-} suckling pups. As the cribriform plate of the ethmoid is important in supporting the olfactory bulb and in preventing the leakage of cerebrospinal fluid into the nasal cavity, neurological abnormalities may also be a contributing factor to the demise of the *Hyal2*^{-/-} mice. The bones of the viscerocranium of human carriers of the *HYAL2* variants have not yet been examined, however defects of the vomer are common in individuals with CLP (Olofsson, Porsch et al. 2014). The identification of developmental abnormalities in a small number of heterozygous (*Hyal2*^{+/-}) mice suggests that haploinsufficiency of *HYAL2* may also pose the risk of a similar outcome in humans. Consistent with the concept of altered *HYAL2* levels affecting palatal development, decreased HA has previously been associated with increased risk for CP in the *Tbx1*^{-/-} mouse (Goudy, Law et al. 2010), whereas increased HA has been associated with CP in *Sim2*^{-/-} mice (Shamblott, Bugg et al. 2002). Hence *HYAL2* variants may be considered as candidates, even in the heterozygous form, to confer risk for the development of CLP. However, screening a cohort of non-syndromic CLP individuals, provided no evidence for a role of *HYAL2* in this cohort (personal

communication- P. Stanier). The cohort consisted of 380 human individuals (176 with unilateral or bilateral cleft lip and palate, 38 with cleft lip only, 121 with cleft palate only and 46 with submucous cleft palate) of a mixture of ethnicities with the majority being of white European origin (72%) from the North East region of London, UK. Approximately 22% of cases were familial with affected first or second-degree relatives. Whilst this cohort is too small to draw definitive genotype-phenotype conclusions, this finding is consistent with *HYAL2* variants being associated with a rare, autosomal recessive syndrome.

As mentioned above, previous studies in mice deficient in HA synthesis (*Has2*^{-/-}) demonstrated that high molecular mass HA was crucial for epithelial to mesenchymal transition (EMT) in the developing heart, whereas HA oligosaccharides inhibited EMT and promoted angiogenesis (Camenisch, Spicer et al. 2000), (Rodgers, Lalani et al. 2006) (Lagendijk, Szabó et al. 2013). *HYAL2* may modulate the size of HA to inhibit EMT and promote differentiation. *HYAL2* deficiency may therefore result in increased levels of high molecular mass HA, an overproduction of mesenchymal cells, and a decrease in differentiation. Indeed, the thickened cardiac valves and excess mesenchymal cells in some parts of the head of *Hyal2*^{-/-} mice, suggest that EMT is upregulated in *Hyal2*^{-/-} mice, and that *HYAL2* plays a role in inhibiting this process. In histological sections of the nasopharynx and secondary palate, there is an abundance of cartilage but a deficiency in bone. The deficient ossification in the skull of *Hyal2*^{-/-} mice, and underdevelopment of the viscerocranium, suggests a failure in the development of osteoblasts, as has also been seen in *Tbx22*^{-/-} mice (Pauws, Hoshino et al. 2009).

In addition to the direct effects of HA on signalling pathways, it is possible that an abundance of HA in *Hyal2*^{-/-} tissues leads to variations in the gradients of morphogens or signal transduction factors, impacting on normal development. Clearly further studies are needed to investigate this possibility. Treatment strategies for this developmental disorder also deserve further research, as HA synthesis can be inhibited using 4-methylumbelliferone, an FDA-approved drug (Nagy, Kaber et al.

2015). Furthermore, commercial preparations of human hyaluronidase that function at a neutral pH could be used to remove extracellular HA. These enzymes are available for human use (Wasserman 2014), and formulations that are stable in the circulation are under development (Jiang, Li et al. 2012). Taken together, our findings highlight the vital role hyaluronidase enzymes play in both human and mouse development, and in particular have revealed a previously unrecognized pathway involved in the pathogenesis of cor triatriatum sinister, orofacial clefting and myopia, which may ultimately be amenable to treatment strategies.

Future work relating to this study is ongoing. Both our group and other research groups around the world have identified additional individuals with disease-causing variants in *HYAL2*. After publication of our findings, we were contacted by other research groups who had been studying individuals with similar phenotypes, and after sequencing of *HYAL2* identified causative variants in this gene. Up to now, an additional six individuals have been identified who are either homozygous or compound heterozygous for disease-causing *HYAL2* mutations. Of the six individuals, two are Amish, one non-Amish North American, two European and one is of Turkish ancestry. Assessment of additional individuals with *HYAL2* mutations has provided further insight into the phenotypic spectrum of disease expression, which seems to include a more variable expression of CLP and other congenital cardiac anomalies (ventricular septal defect). Seven new mutations have been identified, including predicted loss-of-function variants in compound heterozygous configuration with missense mutations. Future studies need to further explore the role of *HYAL2* both during development and later on in life, to gain a better understanding of the disease process and the effects of the different mutations on protein function, with the ultimate goal to identify potential treatment targets.

5 Investigation of the phenotype and genetic modifiers in hypertrophic cardiomyopathy caused by the c.3330+2T>G variant in *MYBPC3*

5.1 Summary

Hypertrophic cardiomyopathy (HCM) is the most common inherited cardiovascular disorder worldwide (1:500) and characterised by asymmetric hypertrophy of the left ventricle. The most commonly implicated gene is myosin binding protein C (*MYBPC3*). Significant phenotypic variation exists between affected individuals and even between carriers of the same HCM causing mutation. The exact underlying mechanisms for this variable expression are yet to be defined, however, genetic modifiers likely play an important role.

The splice site mutation c.3330+2T>G in *MYBPC3* has previously been identified as a founder mutation causing HCM in the Anabaptist communities. Homozygous carriers of the mutation universally die of a severe infantile type of HCM. The aims of this study were to describe the phenotypic expression of HCM in adult heterozygous carriers of the c.3330+2T>G variant in *MYBPC3*, and to investigate potential genetic modifiers of disease expression using a three-pronged approach of *MYBPC3* transcript analysis, exome sequence analysis and a genome wide SNP association study.

The study population comprises 313 carriers of the *MYBPC3* c.3330+2T>G variant and constitutes the largest cohort of HCM founder mutation carriers investigated to date. The phenotypic description of the carriers illustrates the broad spectrum of clinical disease associated with the *MYBPC3* c.3330+2T>G variant. Disease penetrance is 50% and age related, all morphological variants of cardiac hypertrophy are observed, and sudden death seems to occur at any age. The fact that the phenotype can vary to such a degree among a large group of carriers of the same

HCM mutation confirms that disease modifying factors are very powerful and may exert as much of an effect on disease expression as the mutation itself.

The aim of the *MYBPC3* transcript analysis was to investigate the transcriptional outcomes deriving from the pathogenic c.3330+2T>G variant in *MYBPC3*. The study confirmed a distinct transcription profile of *MYBPC3* in heart and blood of controls with notably increased expression levels of the consensus transcript in heart compared with blood. In blood of gene carriers, the key transcripts present are the consensus transcript, and an additional transcript deleted for exon 30. There was no significant association of transcript levels in the blood of *MYBPC3* c.3330+2T>G variant carriers and disease severity.

Exome sequence analysis was performed in ten carriers of the c.3330+2T>G variant in *MYBPC3* with a severe HCM phenotype with the aim to identify rare disease modifying variants in coding regions or at splice sites. A number of putative disease-modifying variants in genes previously associated with cardiomyopathies or cardiac arrhythmias were identified. These variants deserve further investigation with segregation and association studies with phenotype severity within the larger community cohort of *MYBPC3* gene carriers, and additional functional studies, however, some promise to be important modifiers of disease expression in c.3330+2T>G *MYBPC3* associated and potentially other types of HCM.

The aim of the SNP association study was to identify common variants associated with HCM phenotype severity in the *MYBPC3* c.3330+2T>G variant cohort. A potentially important finding was a genome-wide significant association between a region on chromosome 2q36.1, involving seven adjacent SNPs and an ECG marker of clinical HCM severity. The finding needs to be validated in larger Amish and non-Amish HCM cohorts to establish whether this constitutes a true association. The nearest genes to the association signal are *EPHA4* and *PAX3*, and future functional studies need to define the role of these genes in the myocardium to assess whether they may be biologically plausible modifiers of the phenotype.

5.2 Introduction

Hypertrophic cardiomyopathy (HCM) is the most common inherited cardiovascular disease worldwide with an estimated prevalence of 1:500 individuals. Morphologically, HCM is characterised by hypertrophy of the left (and sometimes the right) ventricle which is often asymmetric without evidence of any underlying cardiac or systemic disease (Maron, McKenna et al. 2003). Symptoms can be minimal or progressive with resultant heart failure or sudden death. Pathological features of HCM comprise myocyte hypertrophy and disarray, interstitial fibrosis and mitral valve leaflet anomalies. The degree of cardiac hypertrophy and interstitial fibrosis are major determinants of mortality and morbidity in HCM (Spirito, Bellone et al. 2000).

HCM is familial in 50-70% of cases (Marian 2010). The mode of inheritance is usually autosomal dominant with variable expression. In familial cases a pathogenic variant in one of the sarcomeric genes (*MYH7*, *MYBPC3*, *TNNT2*, *TPM1*, *MYL2*, *MYL3*, *TNNI3*, *ACTC1*) can be identified in up to 80% of families (Walsh, Thomson et al. 2017). Most pathogenic variants causing HCM are so-called “private mutations”, and only identified in one family. By investigating larger cohorts of individuals with HCM, a number of founder mutations have been identified over the last 20 years. These are mutations that seemingly occur in different families, however, reside on the same haplotype and therefore must have arisen in a common ancestor and passed down through the generations. This distinguishes them from recurring mutations in the same genetic (hot-)spot, which can also be present in different families, however they do not reside in the same genetic haplotypes. Most HCM founder mutations occur in *MYBPC3* (Moolman-Smook, De Lange et al. 1999, Kubo, Kitaoka et al. 2005, Dhandapany, Sadayappan et al. 2009, Oliva-Sandoval, Ruiz-Espejo et al. 2010, Teirlinck, Senni et al. 2012, Jaaskelainen, Helio et al. 2013, Adalsteinsdottir, Teekakirikul et al. 2014, Sabater-Molina, Saura et al. 2017, van Velzen, Schinkel et al. 2017) (Teirlinck, Senni et al. 2012). Typically, founder disease mutations are recessive or if dominant, affect adults after prime reproductive age, and therefore

are passed on (often unknowingly) from the original founder to many of his or her descendants. The number of carriers will be dependent on the growth-rate of the population the variant has arisen in.

The splice site mutation c.3330+2T>G in *MYBPC3* has previously been identified as a founder mutation causing hypertrophic cardiomyopathy in the Anabaptist communities of North America (Xin, Puffenberger et al. 2007, Zahka, Kalidas et al. 2008). Homozygosity for this variant results in a lethal form of infantile hypertrophic cardiomyopathy. Affected children uniformly show progressive deterioration of ventricular systolic function which results in death in the first year of life due to inadequate cardiac output, unless orthotopic heart transplantation is performed. Microscopy of affected hearts shows the hallmark feature of HCM- widespread myocyte disarray affecting the right and left ventricles. The majority of the parents of the identified homozygous children were unaware of being (obligate) carriers for the c.3330+2T>G variant in *MYBPC3* and had not been diagnosed with hypertrophic cardiomyopathy. Cases of HCM associated with heterozygous carrier status in the Anabaptist communities however had been observed, often occurring in other close relatives of the affected children (personal communication Dr K Zahka and Dr H Lever, Cleveland Clinic, USA).

The phenotypic expression of autosomal dominant HCM shows remarkable variability regarding the severity of cardiac hypertrophy and the risk of sudden cardiac death. This is not only true for carriers of different pathogenic variants, but also for members of a single family who share the same mutation. At present the exact mechanisms that lead to the diverse phenotypes of HCM remain incompletely understood. Genetic variants affecting sarcomeric protein structure or expression level lead to molecular and cellular dysfunction with activation of various signalling pathways. The resulting phenotype is the consequence of interactions between different intermediary molecular pathways and environmental influences. Altogether, these observations suggest the involvement of a number of factors distinct from the causal mutation in the evolution of the phenotype including

environment, growth factors, hormones, epigenetic factors, post-translational modification of proteins and modifier genes.

This study investigates the phenotypic spectrum of HCM in heterozygous carriers of the *MYBPC3* founder mutation c.3330+2T>G in the Anabaptist communities of Ohio, and the presence of potential genetic modifiers with a three-pronged approach entailing *MYBPC3* transcript studies, exome sequence analysis and a genome-wide SNP study.

5.3 Investigation of the phenotype of carriers of the *MYBPC3* c.3330+2T>G variant

5.3.1 Specific methods

In total 614 individuals were genetically tested for the 3330+2T>G variant in *MYBPC3* (428 individuals in Holmes County, 186 in Geauga County). 331 carriers of the *MYBPC3* variant were identified. After test reports from UH Cleveland were obtained, participants who tested positive for the *MYBPC3* variant were invited to receive their test results in person and have a consultation, electrocardiogram, transthoracic echocardiogram and 24-hour holter monitor performed. 18 individuals were excluded from further analysis in the research study for the following reasons:

- Seven women were pregnant
- Three individuals had previous unrelated cardiac surgery (mitral valve replacement, aortic valve replacement and mitral valve repair, mitral valve replacement for rheumatic fever as a child)
- One obligate carrier (he had a homozygous baby) refused genetic testing
- One carrier died of a neurodegenerative disorder after recruitment
- Three individuals were dwarfs

313 individuals were included in the clinical data analysis.

Commented [CA3]: Stated earlier....

Data were analysed using IBM SPSS software, version 24 (Armonk, NY) and expressed as mean +/- standard deviation, range and percentages as appropriate. Categorical dependent and independent variables were compared using the Pearson Chi-square test, as long as less than 20% of cells in the test table had an expected count of less than 5. Column proportions were compared using the z-test and p-values were adjusted with the Bonferroni method. Continuous dependent and categorical independent variables were compared with the independent-samples T test. Levene's test for equality of variances was used to assess the assumption of homogeneity of variance.

The relationship between continuous dependent and independent variables was assessed with the bivariate procedure, as long as scatterplots indicated a linear relationship between the variables. The Pearson correlation was used to quantify the strength and direction of the relationship between the two continuous variables. Probability values of ≤ 0.05 were considered statistically significant.

5.3.2 Results

5.3.2.1 Demographic and clinical details of 313 carriers of the *MYBPC3* c.3330+2T>G variant

In total 614 individuals were tested for the *MYBPC3* c.3330+2T>G variant (428 individuals in Holmes County, 186 in Geauga County, Ohio). 331 carriers of the *MYBPC3* variant were identified. All tested individuals could be traced back to being descendants of an Anabaptist couple who emigrated from Europe to Somerset County in Pennsylvania in the 18th century. The husband, MT, was born in 1730 and died in 1807. The wife, MM, was born in 1736 and died in 1807. Eight out of their fourteen children have descendants in Holmes and Geauga County (further details and pedigrees in appendix 7.2.1 and 7.2.2).

Figure 5-1 shows the age distribution of the 313 carriers of the *MYBPC3* 3330+2T>G variant included in the data analysis. Ten individuals declined attending for their clinical assessment. Demographic and clinical details were incomplete in nine

individuals due to some questions not being answered or some assessments not completed. The demographic and clinical characteristics of all *MYBPC3* variant carriers are summarised in table 5-1.

159 of 313 (50.8%) gene carriers were male. The mean age at assessment was 42.7 +/- 16.4 years. Every gene carrier had on average 4.3 +/- 3.4 healthy children. 32 gene carriers had between one and four homozygous babies with severe infantile HCM. In total, 31 homozygous babies were reported by the participants. They had been born to 19 different couples. Apart from three homozygous children who received heart transplants, all homozygous babies died within the first year of life due to heart failure. Out of the three children who had received heart transplants, one died at the age of 16 after his second transplant.

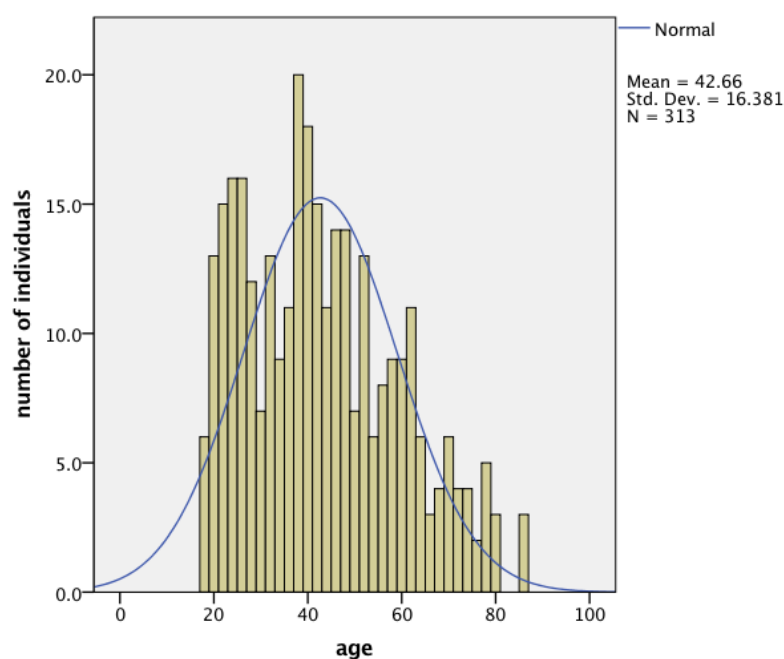


Figure 5-1: Histogram showing the age distribution of 313 carriers of the *MYBPC3* 3330+2T>G variant

48 out of 294 individuals (16.3%) had previously been diagnosed with HCM. The diagnosis was based on a clinical assessment in 23 (47.9%) and a genetic test in 17 individuals (35.4%). In eight participants (16.7%) it was not possible to establish how a diagnosis of HCM had been made. The average age at diagnosis was 42.2 +/- 12.6 years in the individuals with known HCM or a known positive genetic result for the *MYBPC3* variant. 22.1% of participants (65/294) had previously seen a Cardiologist on at least one occasion.

27 out of 294 gene carriers (8.6%) reported a first degree relative who died suddenly after being diagnosed with HCM, or under circumstances suspicious of sudden cardiac death secondary to undiagnosed HCM. In individuals without a known diagnosis of HCM, only sudden deaths which occurred in seemingly healthy individuals without obvious cause were considered likely secondary to HCM. A post-mortem examination was only performed in one of the sudden death victims (20-year-old female, post-mortem examination showed HCM with septal hypertrophy and myocyte disarray). In total 16 sudden deaths were registered. Of note is that in three nuclear families two sudden deaths in two generations occurred respectively. 11/16 (68.8%) sudden death victims were male. The mean age was 42.5 +/- 20.7 years with sudden deaths occurring in each decade between the ages of 14 to 69 years old (table 5-2).

16 out of 294 gene carriers (5.4%) had a history of atrial fibrillation or were in atrial fibrillation at the time of assessment. The mean age at onset of atrial fibrillation was 57.9 +/- 15.4 years. Five individuals had either a permanent pacemaker or internal defibrillator (1.7%, 5/294). The mean age at implantation was 50.8 +/- 8.8 years. Six individuals (1.9%, 6/313), all male, had undergone surgical myectomies for outflow tract obstruction at a mean age of 53.0 +/- 3.8 years.

male	50.8% (159/313)
age [y]	42.7 +/- 16.4
number of healthy children	4.3 +/- 3.4
weight [kg]	80.5 +/-16.8
height [cm]	169.5 +/- 9.1
BMI [kg/m ²]	28.0 +/- 5.8
systolic BP [mmHg]	122.0 +/- 16.5
diastolic BP [mmHg]	77.1 +/- 10.2
hypertension	11.8% (34/289)
diabetes mellitus	3.7% (11/294)
hypercholesterolemia	4.8% (14/294)
coronary disease	1.7% (5/294)
known diagnosis of HCM	16.3% (48/294)
age at diagnosis of HCM [y]	42.2 +/- 12.6
previously seen Cardiologist	22.1% (65/294)
first degree relative SCD	8.6% (27/294)
atrial fibrillation	5.4% (16/294)
age at onset AF [y]	57.9 +/-15.4
ICD/pacemaker	1.7% (5/294)
age at ICD/pacemaker [y]	50.8 +/- 8.8
surgical myectomy	1.9% (6/313)
age at myectomy [y]	53.0 +/- 3.8
on beta-blocker	12.5% (39/294)
on NDH channel blocker	1.7% (5/294)

Table 5-1: Demographics and clinical characteristics of all *MYBPC3* 3330+2T>G carriers

Hypertension defined as known hypertension or average BP > 140/90 mmHg at time of assessment

individual	age	gender	circumstances of death	number of first-degree relatives identified as carriers of <i>MYBPC3</i> variant in this study
1	14	female	whilst running after complained of burning in chest, previously had chest pain on exertion, had been avoiding running and playing at school	1
2	15	male	whilst playing basketball, previously fit and well	5
3	16	male	whilst swimming, previously fit and well	1
4	20	female	Sudden death, post-mortem examination showed HCM with septal hypertrophy and myocyte disarray	1
5	21	male	collapsed and fell into machine, previously well	1
6*	27	male	whilst out running, previously well	2
7	37	male	sudden death, previously well	1
8 [#]	47	female	sudden death, known heart condition	1
9	49	female	Sudden death, known HCM	8
10	55	male	sudden death, known HCM, previous myectomy	3
11	57	male	sudden death, known heart problem	1
12 [#]	58	male	sudden death, known HCM, previous myectomy and pacemaker	1
13 ⁺	63	male	sudden death, previously fit and well	1
14 ⁺	65	female	sudden death, known congestive heart failure for 13 years	1
15	67	male	sudden death, previously fit and well	1
16*	69	male	sudden death, previously well	2

Table 5-2: Characteristics of *MYBPC3* variant carriers who died suddenly

*: individual 6 and 16 are son and father
#: individual 8 and 12 are mother and son
+: individual 13 and 14 are son and mother

5.3.2.2 Carriers of the *MYBPC3* c.3330+2T>G variant with and without previous myectomy

Demographic and clinical characteristics

Six out of 313 carriers of the *MYBPC3* c.3330+2T>G variant had undergone surgical myectomies. All six individuals were male and the mean age at myectomy was 53 +/- 3.8 years. One of the carriers with myectomy declined to attend for a clinical assessment. The mean age at the time of assessment of the individuals with previous myectomy was 55.2 +/- 8.0 years. HCM had been diagnosed at a mean age of 47.4 +/- 8.3 years. A history of atrial fibrillation (AF) was more common in carriers with myectomy (66.7%, 4/6) compared to carriers without previous myectomy (4.2%, 12/288). The mean age at onset of AF was 50.5 +/- 6.1 years compared to 60.4 +/- 17.0 in carriers without myectomy. One of the individuals who had undergone myectomy had an ICD (16.7%) compared to 1.4% (4/288) of the carriers without myectomy (appendix table 7-2).

The most common symptom prior to myectomy in 2/6 individuals (33.3%) was sudden syncope at rest. In carriers without myectomy, 5/285 (1.8%) reported sudden syncope at rest. One individual with myectomy who experienced syncope, subsequently had ventricular tachycardia on holter monitoring and an episode of aborted sudden cardiac death after admission to hospital. None of the carriers with myectomy reported exertional chest pain as presenting symptom. One individual each reported shortness of breath, palpitations and fatigue (table 5-3).

ECG characteristics

P wave duration, PR interval, QTc and QRS duration all seemed longer in individuals with myectomy. The average Romhilt Estes score in carriers with myectomy was 8.0 +/- 2.1 (n=5) compared to 2.0 +/- 1.4 (n=239) in individuals without myectomy (table 5-4). 80% (4/5) of individuals with myectomy were in sinus rhythm at the time of assessment compared to 95.8% (229/239) of carriers without myectomy. Only 2/5 (40%) had a normal QRS axis compared to 84.5% (202/239) of carriers without

myectomy. 80% (4/5) fulfilled the Sokolow-Lyon criteria for LVH compared to 9.3% (22/237) of individuals without myectomy.

Only 20% (1/5) of the individuals who had undergone myectomy showed fragmentation of the QRS complex. This seemed more common in carriers without myectomy (67.4%, 155/230). None of the carriers with myectomy had pathological Q waves compared to 7.8% (18/232) of individuals without myectomy. All individuals with previous myectomy had T wave inversion/ deep T wave inversion compared to 7.6% (18/237) and 3.4% (8/237) of carriers without myectomy respectively. The most common site for TWI in carriers with myectomy was V4-V6 (80%, 4/5). All five individuals with myectomy who underwent a clinical assessment had an abnormal ECG and a Romhilt Estes score of ≥ 5 , compared to 24.0% (56/233) and 7.1% (17/239) of carriers without myectomy respectively (appendix table 7-3).

	myectomy (n=6)	no myectomy
exertional chest pain	0%	14.6% (42/287)
exertional shortness of breath	16.7% (1/6)	45.6% (131/287)
palpitations	16.7% (1/6)	39.9% (114/286)
syncope	33.3% (2/6)	24.5% (70/286)
neurocardiogenic syncope	0%	19.2% (55/286)
sudden syncope at rest	33.3% (2/6)	1.8% (5/285)
aborted SCD	16.7% (1/6)	0%
fatigue	16.7% (1/6)	6.6% (16/241)

Table 5-3: Symptoms of carriers of MYBPC3 3330+2T>G variant with and without myectomy

Aborted SCD: aborted sudden cardiac death

	myectomy	no myectomy
P wave duration [ms]	127.6 +/- 11.5 (n=5)	110.0 +/- 13.5 (n=232)
PR interval [ms]	229.8 +/- 30.8 (n=4)	171.0 +/- 25.4 (n=232)
QT duration [ms]	456.0 +/- 22.3 (n=4)	404.3 +/- 32.6 (n=235)
QTc duration [ms]	441.4 +/- 18.6 (n=4)	422.6 +/- 26.3 (n=235)
QRS duration [ms]	150.6 +/- 13.1 (n=5)	82.9 +/- 13.6 (n=237)
intrinsicoid deflection [ms]	66.4 +/- 15.0 (n=5)	37.5 +/- 7.7 (n=234)
Romhilt Estes score	8.0 +/- 2.1 (n=5)	2.0 +/- 1.4 (n=239)

Table 5-4: ECG characteristics (part 1) of individuals with and without myectomy

Holter monitoring

24-hour-holter monitoring was performed in two individuals with previous myectomy and 63 gene carriers without previous myectomy. The recordings were only analysed for ventricular arrhythmias.

The two individuals with previous myectomy had a small number of premature ventricular contractions only (PVCs, 4.5 +/- 3.5) and neither of them had any non-sustained ventricular tachycardia (NSVT). The individuals without myectomy had significant variation in the number of PVCs observed with a mean number of 585.9 +/- 1541.5, ranging from 0 to 6506 in a 24-hour period. Nine out of 63 individuals (14.3%) had at least one episode of NSVT (table 5-5).

Transthoracic echocardiography

All individuals who had previously undergone a surgical myectomy, had a residual maximal wall thickness (WT) of 13 mm or more. 50.0% of individuals without myectomy (133/266) had a max WT of 13 mm or more, fulfilling echocardiographic diagnostic criteria for HCM in known carriers of a pathogenic variant causing HCM. Figure 5-2 shows the distribution of maximum wall thickness in all carriers of the *MYBPC3* c.3330+2T>G variant without myectomy.

Mean septal thickness in diastole (IVSd) in the parasternal long axis view (PLAX) in individuals with previous myectomy was 13.8 +/- 2.2 mm compared to 11.8 +/- 3.6 (n=266) in gene carriers without myectomy. Septal thickness was on average larger than posterior wall thickness (PWd) in both groups (12.0 +/- 2.7 mm and 9.4 +/- 1.8 mm respectively). None of the gene carriers with previous myectomy and only 1.5% (3/200) of individuals without myectomy had a resting left ventricular outflow tract (LVOT) gradient of more than 30 mmHg (appendix table 7-4).

	myectomy (n=2)	no myectomy (n=63)
number of PVCs	4.5 +/- 3.5	585.9 +/- 1541.5 (0-6506)
Number of episodes of bigeminy	0	18.7 +/- 86.6 (0-663)
Number of couplets	0	11.1 +/- 73.2 (0-577)
Number of triplets	0	0.45 +/- 2.4 (0-19)
Individuals with NSVT ≥ 4 beats [%]	0	14.3 (9/63)
Min heart rate [bpm]	56.0 +/- 0	54.8 +/- 6.8 (37-71)
Max heart rate [bpm]	123.0 +/- 9.9	129.4 +/- 15.6 (102-174)
Mean heart rate [bpm]	68.5 +/- 13.4	77.7 +/- 8.2 (50-95)

Table 5-5: Holter characteristics of individuals with and without myectomy

Mean values for mitral valve inflow E/A ratio (1.36 +/- 0.55), mitral valve inflow deceleration time (172.9 +/- 42.8 ms), intraventricular relaxation time (IVRT, 85.0 +/- 21.4 ms), E'/A' lateral (1.54 +/- 0.94) and septal (1.31 +/- 0.73) and average E/ E' (lateral and mitral, 7.64 +/- 3.80) in carriers of the *MYBPC3* 3330+2T>G variant without myectomy were within normal limits indicating on average normal diastolic function. Mean values for indices of diastolic function in individuals with previous myectomy tended to be higher, indicating on average mild diastolic dysfunction (appendix table 7-5).

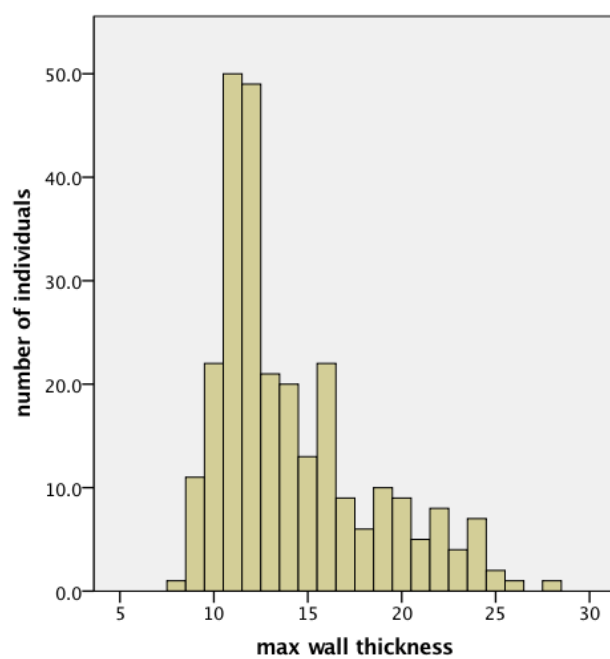


Figure 5-2: Maximum wall thickness in all carriers of *MYBPC3* 3330+2T>G variant (individuals with previous myectomy are excluded)

As mentioned above, all individuals with myectomy had an abnormal ECG and a residual maximal WT ≥ 13 mm.

In the individuals without myectomy, the ECG had a sensitivity of 38.1% to correctly identify individuals with a positive phenotype defined by a maximum wall thickness of ≥ 13 mm on echocardiography. The ECG had a specificity of 89.9% to identify individuals with a maximum wall thickness of < 13 mm. The positive predictive value of an abnormal ECG was 78.4% and the negative predictive value of a normal ECG 60.1% (table 5-6 and figure 5-3).

	max WT < 13 mm	max WT ≥ 13 mm	total
normal ECG	98 (60.1%)	65 (39.9%)	163 (100.0%)
abnormal ECG	11 (21.6%)	40 (78.4%)	51 (100.0%)
total	109 (50.9%)	105 (49.1%)	218

Table 5-6: Relation between ECG normal/ abnormal and maximum wall thickness in all carriers of the *MYBPC3* 3330+2T>G variant

(excluding individuals with myectomy)

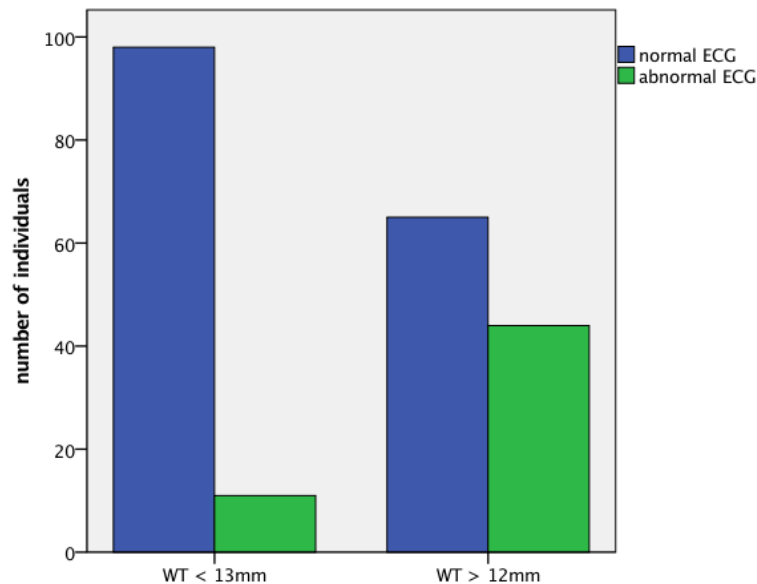
Most individuals without previous myectomy and a max WT ≥13mm either had a sigmoid septum (25.2%, 63/250) or reverse septal curvature (20.0%, 50/250). 2.4% (6/250) had apical hypertrophy, 1.6% (4/250) concentric hypertrophy, 2.0% (5/250) segmental hypertrophy in the anterior or lateral wall and 0.4% (1/250) had progressed to the dilated stage. Three out of four (75%) individuals with previous myectomy showed the reverse septal hypertrophy pattern and one (25%) had concentric hypertrophy.

12.2% (24/196) of carriers of the *MYBPC3* c.3330+2T>G variant without myectomy had systolic anterior motion of the mitral valve (SAM). 75.0% (3/4) of individuals with previous myectomy showed SAM. All of the individuals with previous myectomy had hypertrophy of the papillary muscles compared to 54.3% (100/184) of individuals without previous myectomy (appendix table 7-6).

5.3.2.3 Male and female carriers of the *MYBPC3* c.3330+2T>G variant

Demographic and clinical characteristics

The following data includes 307 carriers (six carriers with previous myectomy were excluded). 49.8 % of the carriers were male (153/307). The average age of male and female carriers was not significantly different (41.3 +/- 16.9 vs. 43.5 +/- 15.8 years). A higher proportion of female carriers had hypertension (16.0% vs. 7.9%, p=0.037).



2

Figure 5-3: Relation between maximum wall thickness and ECG normal/abnormal in all carriers of the *MYBPC3* 3330+2T>G variant

Individuals who have undergone septal myectomy are excluded

Male carriers were more likely to have been diagnosed with HCM previously (18.8% vs. 10.4%, $p=0.045$) and to have concomitant diagnosed coronary artery disease (3.5% vs. 0.0%, $p=0.024$). There was no significant difference between male and female carriers in BMI, systolic and diastolic blood pressure, diabetes mellitus and hypercholesterolaemia. The age at previous diagnosis in individuals with known HCM was comparable (41.0 $\pm$ 13.8 vs. 43.3 $\pm$ 11.9 years). The rate of first-degree relatives who died suddenly, prevalence of atrial fibrillation, implantable pacemakers and defibrillators and treatment with beta-blockers and NDH calcium-channel blockers was not significantly different (appendix table 7-7).

Overall, female carriers were more likely to report palpitations (52.8% vs. 26.8%, $p<0.001$). There was no significant difference between the proportion of male and female carriers reporting palpitations consistent with ectopic beats (10.6% vs. 17.4%, $p=n.s.$) or supraventricular arrhythmias/ atrial fibrillation (2.8% vs. 4.2%) either documented or based on the description. There was a significant difference between male and female carriers with palpitations that were undetermined based on the description or available monitoring (14.1% vs. 30.6%, $P=0.001$). A similar proportion of male and female carriers reported exertional chest pain, exertional shortness of breath, dizziness and fatigue. A larger proportion of female carriers reported syncope consistent with neurocardiogenic syncope based on the history (23.6% vs. 14.8%, $p=0.058$), and a larger proportion of male carriers reported exertional syncope (4.3% vs. 0.0%, $p=0.012$). The rate of sudden syncope at rest was not significantly different between male and female carriers (0.7% vs. 2.8%, $p=n.s.$). The proportion of carriers reporting significant symptoms, which were defined as chest pain, shortness of breath, dizziness or syncope on exertion or sudden dizziness or syncope at rest, was comparable between male and female carriers (29.5% vs. 28.2%, $p=n.s.$) (appendix table 7-8). There was no significant difference in the age at onset of symptoms between male and female carriers apart from the age at onset of palpitations (32.8 +/- 12.8 vs. 39.8 +/- 12.0 years, $p=0.02$) (table 5-7).

	male	female	p value
age at first chest pain	42.0 +/- 19.2 (n=11)	43.1 +/- 16.3 (n=17)	n.s.
age at first SOB	40.0 +/- 15.7 (n=47)	40.9 +/- 12.9 (n=60)	n.s.
age at first palpitations	32.8 +/- 12.8 (n=27)	39.8 +/- 12.0 (n=48)	0.02
age at first dizziness	35.6 +/- 14.3 (n=14)	37.5 +/- 14.3 (n=15)	n.s.
age at first syncope	26.9 +/- 20.8 (n=26)	24.0 +/- 20.2 (n=29)	n.s.
age at first fatigue	56.3 +/- 10.2 (n=4)	44.8 +/- 17.9 (n=9)	n.s.
age at first symptom	33.2 +/- 17.5 (n=71)	32.7 +/- 15.2 (n=92)	n.s.

Table 5-7: Age at onset of symptoms of MYBPC3 c.3330+2T>G variant carriers
ECG characteristics

Female carriers had a slightly higher resting heart rate on 12-lead ECG recordings (68.1 +/- 10.9 vs. 65.0 +/- 10.5 bpm, $p = 0.026$). Male carriers' ECGs were more likely to fulfil Sokolow-Lyon criteria for RVH and LVH than female carriers' ECGs (33.9% vs 7.2%, $p < 0.001$ and 17.0% vs 2.4%, $p < 0.001$) and had longer QRS duration (85.8 +/- 13.3 ms vs. 80.3 +/- 13.4 ms, $p = 0.002$). A larger proportion of male carriers had pathological Q waves on ECG than female carriers (11.9% vs. 4.1%, $p = 0.025$). QTc duration in female carriers was longer than in male carriers (430.4 +/- 27.0 ms vs. 413.9 +/- 22.5 ms, $p < 0.001$). There was no statistically significant difference in any other ECG parameters between male and female carriers, however there seemed to be a trend towards a higher proportion of male carriers with abnormal ECGs than female carriers (29.1% vs. 19.5%, $p = n.s.$) (appendix table 7-9).

Holter monitoring

24-hour holter monitoring was performed in 32 male and 33 female randomly selected carriers. There were no significant differences in the number of recorded premature ventricular complexes, episodes of bigeminy, number of couplets and triplets, minimum and mean heart rate between male and female carriers. There seemed to be a higher proportion of female than male carriers with non-sustained

ventricular tachycardia of four or more beats (21.2% vs. 6.3%), however the investigated numbers were too small to perform valid statistical analysis. The average maximum recorded heart rate was slightly higher in men than in women (133.3 +/- 16.8 vs 125.4 +/- 13.1 beats per minute, $p=0.039$) (appendix table 7-10).

Transthoracic Echocardiography

Male carriers had on average larger wall thickness measurements than female carriers. The intraventricular septum in diastole in the long-axis parasternal view in male carriers measured on average 12.4 +/- 3.5 mm compared to 11.2 +/- 3.6 in female carriers ($p=0.007$). The posterior wall measurements were 9.8 +/- 1.8 mm in male vs. 9.1 +/- 1.7 mm in female carriers ($p=0.001$). Wall measurements in the short axis view at mitral, papillary and apical level were equally larger in male than female carriers (table 5-8). The average maximum wall thickness in males measured 15.0 +/- 4.1 mm compared to 13.3 +/- 3.8 mm in females. 61.7% of male carriers had a maximum wall thickness of ≥ 13 mm in contrast to only 39.1% of female carriers.

Left ventricular cavity dimensions and volumes were also larger in males. Left ventricular ejection fraction and relative wall thickness were not significantly different between males and females. Left atrial volumes were larger in male than female carriers (65.1 +/- 28.3 ml vs. 54.9 +/- 18.3 ml, $p=0.004$).

As mentioned above, a larger proportion of female carriers of the *MYBPC3* variant were phenotype negative with a maximum wall thickness of less than 13 mm compared to male carriers (60.8% vs 38.3%, $p<0.001$). Out of the individuals with a wall thickness of 13mm or more, a larger proportion of males had a pattern of reverse septal curvature (26.7% vs. 13.8%, $p=0.011$) and apical hypertrophy (5.0% vs 0.0%, $p=0.010$) compared to females.

Males seemed to be more likely to have systolic anterior motion of the mitral valve (16.0% vs 8.8%), but the difference was not statistically significant (appendix table 7-11).

	male	female	p value
IVSd (PLAX) [mm]	12.4 +/- 3.5 (n=128)	11.2 +/- 3.6 (n=138)	0.007
PWd (PLAX) [mm]	9.8 +/- 1.8 (n=127)	9.1 +/- 1.7 (n=138)	0.001
LVdD (PLAX) [mm]	44.8 +/- 5.9 (n=127)	40.3 +/- 5.5 (n=138)	<0.001
LVSd (PLAX) [mm]	33.5 +/- 6.1 (n=127)	30.6 +/- 5.5 (n=127)	<0.001
LVEDv BP [ml]	86.5 +/- 23.2 (n=118)	69.5 +/- 16.0 (n=116)	<0.001
LVESv BP [ml]	34.3 +/- 12.0 (n=118)	26.7 +/- 8.1 (n=116)	<0.001
EF BP [%]	60.7 +/- 7.6 (n=118)	61.7 +/- 7.2 (n=116)	n.s.
relative WT (PLAX)	0.51 +/- 0.16 (n=127)	0.52 +/- 0.17 (n=138)	n.s.
max mitral level WT (PSAX) [mm]	13.8 +/- 3.9 (n=125)	12.5 +/- 4.0 (n=132)	0.009
max papillary level WT (PSAX) [mm]	13.4 +/- 3.8 (n=125)	11.7 +/- 3.2 (n=131)	<0.001
max apical WT [mm]	12.8 +/- 4.0 (n=66)	10.3 +/- 1.7 (n=67)	<0.001
max WT [mm]	15.0 +/- 4.1 (n=128)	13.3 +/- 3.8 (n=138)	0.001
max WT ≥ 13 mm	61.7% (n=128)	39.1% (n=138)	<0.001
LAv (a-l) mean 4Ch/2Ch [ml]	65.1 +/- 28.3 (n=87)	54.9 +/- 18.3 (n=97)	0.004
LVOT gradient [mmHg]	5.3 +/- 2.2 (n=97)	7.8 +/- 12.6 (n=103)	n.s.
LVOT gradient ≥ 30 mmHg	0.0 % (n=97)	2.9% (n=103)	n.s.
MAPSE lateral [mm]	14.0 +/- 2.9 (n=97)	14.4 +/- 2.8 (n=108)	n.s.
MAPSE septal [mm]	12.5 +/- 2.9 (n=97)	12.4 +/- 2.3 (n=108)	n.s.

Table 5-8: Comparison of echocardiographic characteristics (part 1) of male and female carriers of MYBPC3 3330+2T.G variant

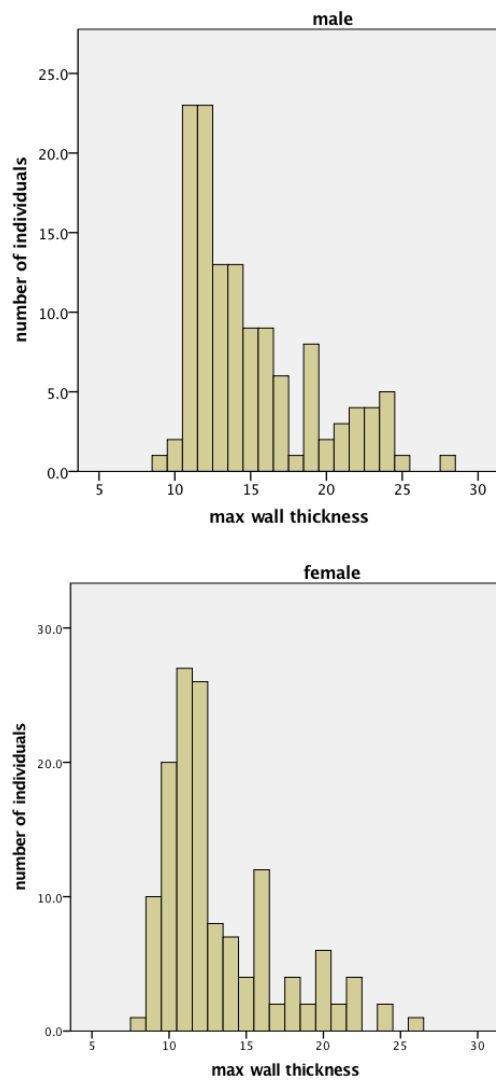


Figure 5-4: Maximal wall thickness in female and male carriers of *MYBPC3* c.3330+2T>G variant

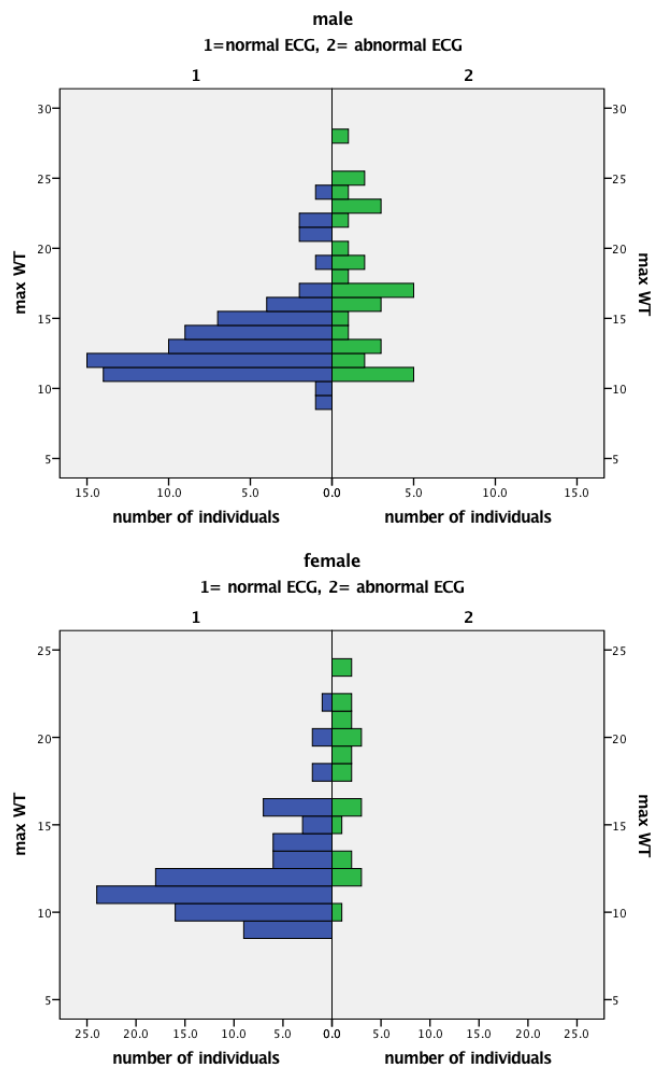


Figure 5-5: Relationship between ECG abnormalities and maximum wall thickness in male and female carriers of *MYBPC3* c.3330+2T>G variant

5.3.2.4 Age related penetrance in carriers of *MYBPC3* c.3330+2T>G variant

Demographic and clinical characteristics

The median age of all carriers of the *MYBPC3* c.3330+2T>G variant was 40 years. The carriers were divided in a group of younger (age 18- 40 years, n=156) and older carriers (41 years and older, n=151). There was no significant difference between the proportion of male individuals in the group of younger carriers and older carriers (54.5% vs. 45.0%, n.s.). Older carriers had a larger BMI than younger carriers (29.3 +/- 5.6 kg/m² vs. 26.8 +/- 5.8 kg/m², p<0.001). The average systolic blood pressure in younger carriers was lower than in older carriers (116.5 +/- 14.2 mmHg vs. 127.6 +/- 17.0 mmHg, p<0.001) and a larger proportion of older carriers had hypertension (19.6%, 27/138 vs. 4.8%, 7/145). A larger proportion of the older carriers had known co-morbidities including diabetes mellitus (7.8%, 11/141 vs. 0%, p=0.001), hypercholesterolaemia (7.8%, 11/141 vs. 1.4%, 2/147, p=0.008) and coronary disease (3.5%, 5/141 vs. 0%, p=0.021) and a larger proportion of older carriers had previously seen a Cardiologist (29.8%, 42/141 vs. 11.6%, 17/147, p<0.001). Older carriers had a higher prevalence of atrial fibrillation (7.8%, 11/141 vs. 0.7%, 1/147, p=0.003) and a larger proportion had an internal cardio-defibrillator or pacemaker than younger carriers (2.8%, 4/141 vs. 0%, p=0.040). More older carriers were taking betablockers compared to younger carriers (19.9%, 28/141 vs. 4.1%, 6/147), p<0.001 (table 5-9).

Regarding symptoms, a larger proportion of older carriers reported exertional shortness of breath (54.6%, 77/141 vs. 37.0%, 54/146, p=0.003) and palpitations (47.9%, 67/140 vs. 32.2%, 47/146, p=0.007). Five older carriers reported sudden syncope at rest compared to none of the younger carriers (3.6%, 5/140 vs. 0%, p=0.022). Younger individuals were less likely to report fatigue compared to older individuals (3.2%, 4/126 vs. 10.4%, 12/115, p=0.024) (appendix table 7-13).

	age 18-40	age ≥ 41	p value
Number of carriers	156	151	n.s.
male	54.5% (85/156)	45.0% (68/151)	n.s.
number of alive children	2.8 +/- 2.8 (n=122)	5.8 +/- 3.3 (n=118)	<0.001
weight [kg]	78.2 +/- 17.5 (n=140)	82.4 +/- 16.0 (n=135)	0.043
height [cm]	170.9 +/- 9.2 (n=141)	167.9 +/- 8.9 (n=137)	0.007
BMI [kg/m ²]	26.8 +/- 5.8 (n=140)	29.3 +/- 5.6 (n=135)	<0.001
systolic BP [mmHg]	116.5 +/- 14.2 (n=144)	127.6 +/- 17.0 (n=137)	<0.001
diastolic BP [mmHg]	75.4 +/- 10.4 (n=144)	78.7 +/- 9.8 (n=137)	0.006
hypertension	4.8% (7/145)	19.6% (27/138)	<0.001
diabetes mellitus	0.0%	7.8% (11/141)	0.001
hypercholesterolemia	1.4% (2/147)	7.8% (11/141)	0.008
coronary disease	0.0%	3.5% (5/141)	0.021
known diagnosis of HCM	10.9% (16/147)	18.4% (26/141)	n.s.
age at diagnosis of HCM [y]	29.4 +/- 5.5 (n=15)	49.2 +/- 10.2 (n=25)	<0.001
previously seen Cardiologist	11.6% (17/147)	29.8% (42/141)	<0.001
first degree relative SCD	7.5% (11/147)	11.3% (16/141)	n.s.
atrial fibrillation	0.7% (1/147)	7.8% (11/141)	0.003
age at onset AF [y]	32 (n=1)	63.0 +/- 15.2 (n=11)	n.s.
ICD/pacemaker	0.0%	2.8% (4/141)	0.040
age at ICD/pacemaker [y]	-	49.8 +/- 9.8 (n=4)	-
on beta-blocker	4.1% (6/147)	19.9% (28/141)	<0.001
on NDH channel blocker	0.0%	3.5% (5/141)	0.021

Table 5-9: Comparison of demographic and clinical characteristics of young (18-40 years) and older (≥ 41 years) carriers of *MYBPC3* c.3330+2T>G variant

Electrocardiogram

Comparing the ECG characteristics, there was no significant difference in resting heart rate between younger and older carriers (66.0 +/- 10.3 bpm vs. 67.3 +/- 11.2 bpm, n.s.) and a similar proportion seemed to be in sinus rhythm (96.7%, 118/122 vs. 94.9%, 111/117). A larger proportion of younger carriers had a rightward QRS axis (9.8%, 12/122 vs. 2.6%, 3/117, $P=0.020$). P wave duration was longer in older carriers (113.8 +/- 14.6 ms vs. 106.3 +/- 11.3 ms, $p<0.001$) and a larger proportion of older carriers demonstrated left atrial enlargement on ECG (9.9%, 11/111 vs. 1.7%, 2/121, $p=0.006$). PR duration was also significantly longer in older carriers (176.7 +/- 28.9 ms vs. 165.7 +/- 20.5, $p=0.001$).

Younger individuals were more likely to fulfil criteria for RVH and LVH according to the Sokolow-Lyon criteria (25.6% vs. 13.8%, $p=0.022$ and 15.7% vs. 2.6%, $p=0.001$). Older individuals had significantly longer QTc intervals than younger carriers (431.4 +/- 27.4 ms vs. 414.3 +/- 22.3, $p<0.001$). A significantly larger proportion of older carriers had T wave inversion on ECG compared to younger carriers (13.8%, 16/116 vs. 1.7%, 2/121, $p<0.001$) and a larger proportion of older individuals had a Romhilt Estes score of 4 or more (17.9%, 21/117 vs. 8.2%, 10/122, $p=0.025$). Overall a significantly larger proportion of older individuals had an abnormal ECG compared to younger individuals (34.8%, 40/115 vs. 13.6%, 16/118, $p<0.001$) (appendix table 7-14).

Holter monitoring

On holter monitoring, older carriers had a significantly larger number of premature ventricular ectopics compared to younger individuals (1108.2 +/- 2002.9 vs. 11.6 +/- 47.5, $p=0.004$) and a larger proportion of older carriers had episodes of non-sustained ventricular tachycardia of four beats or more (24.2% (8/33) vs. 3.3% (1/30), $p=0.018$). The maximum heart rate in younger individuals was significantly higher than in older individuals (135.6 +/- 16.4 bpm vs. 123.9 +/- 12.5 bpm, $p=0.002$) (appendix table 7-15).

Echocardiography

On echocardiography, older carriers had highly significantly larger wall thickness measurements compared to younger carriers. Septal thickness in diastole in the parasternal long axis view in older carriers was on average 13.4 +/- 3.6 mm compared to 10.3 +/- 2.7 mm in younger carriers ($p<0.001$) and posterior wall thickness 10.0 +/- 1.9 mm compared to 8.9 +/- 1.5 mm ($p<0.001$). Short axis measurements of wall thickness at mitral, papillary and apical level were also all significantly larger in older compared to younger carriers. Maximal wall thickness in older carriers on average was 15.6 +/- 4.3 mm compared to 12.7 +/- 3.2 mm in younger carriers ($p<0.001$) and a significantly larger proportion of older individuals had a wall thickness of 13 mm or more compared to younger carriers (65.9%, 85/129 vs. 35.0%, 48/137, $p<0.001$).

Left ventricular diastolic dimensions were significantly smaller in older individuals compared to younger carriers (41.0 +/- 6.8 mm vs. 43.8 +/- 5.1 mm, $p<0.001$) whereas systolic dimensions were not significantly different (32.0 +/- 7.0 vs. 32.0 +/- 4.7 mm, n.s.). Both diastolic and systolic volumes were significantly smaller in older carriers (27.7 +/- 10.9 ml vs. 33.1 +/- 10.2 ml, $p<0.001$ and 70.9 +/- 19.7 ml vs. 84.7 +/- 21.4 ml, $p<0.001$) with similar preserved ejection fractions in both age groups (61.1 +/- 9.1% vs. 61.2 +/- 5.4%, n.s.). Relative wall thickness was significantly larger in older carriers (0.59 +/- 0.18 vs. 0.44 +/- 0.10, $p<0.001$).

Older carriers had larger left atrial volumes than younger carriers (63.4 +/- 30.8 ml vs. 56.3 +/- 14.6 ml, $p=0.048$). Long axis function both in the mitral and lateral wall was greater in younger individuals compared to older individuals, whereas right ventricular long axis function showed no significant difference (table 5-10).

	18-40 years	≥ 41 years	p value
IVSd (PLAX) [mm]	10.3 +/- 2.7 (n=137)	13.4 +/- 3.6 (n=129)	<0.001
PWd (PLAX) [mm]	8.9 +/- 1.5 (n=137)	10.0 +/- 1.9 (n=128)	<0.001
LVdD (PLAX) [mm]	43.8 +/- 5.1 (n=137)	41.0 +/- 6.8 (n=128)	<0.001
LVSd (PLAX) [mm]	32.0 +/- 4.7 (n=137)	32.0 +/- 7.0 (n=128)	n.s.
LVEDv BP [ml]	84.7 +/- 21.4 (n=121)	70.9 +/- 19.7 (n=113)	<0.001
LVESv BP [ml]	33.1 +/- 10.2 (n=121)	27.7 +/- 10.9 (n=113)	<0.001
EF BP [%]	61.2 +/- 5.4 (n=121)	61.1 +/- 9.1 (n=113)	n.s.
relative WT (PLAX)	0.44 +/- 0.10 (n=137)	0.59 +/- 0.18 (n=128)	<0.001
max mitral level WT (PSAX) [mm]	11.8 +/- 3.1 (n=134)	14.6 +/- 4.5 (n=123)	<0.001
max papillary level WT (PSAX) [mm]	11.6 +/- 3.3 (n=134)	13.6 +/- 3.7 (n=122)	<0.001
max apical WT [mm]	11.0 +/- 3.1 (n=79)	12.3 +/- 3.4 (n=54)	0.024
max WT [mm]	12.7 +/- 3.2 (n=137)	15.6 +/- 4.3 (n=129)	<0.001
max WT ≥ 13 mm	35.0% (48/137)	65.9% (85/129)	<0.001
LAV (a-l) mean 4Ch/2Ch [ml]	56.3 +/- 14.6 (n=95)	63.4 +/- 30.8 (n=89)	0.048
LVOT gradient [mmHg]	5.9 +/- 8.6 (n=102)	7.4 +/- 9.9 (n=97)	n.s.
LVOT gradient ≥ 30 mmHg [%]	1.0% (1/102)	2.0% (2/98)	n.s.
MAPSE lateral [mm]	15.2 +/- 2.5 (n=108)	13.1 +/- 2.7 (n=97)	<0.001
MAPSE septal [mm]	13.3 +/- 2.2 (n=108)	11.5 +/- 2.7 (n=97)	<0.001
TAPSE [mm]	25.1 +/- 4.2 (n=108)	24.6 +/- 4.6 (n=93)	n.s.

Table 5-10: Comparison of echocardiographic parameters of young (age 18-40) and older (≥ 41 years) carriers of the MYBPC3 c.3330+2T>G variant (structure and function)

5.3.3 Discussion

This study presents data on the phenotypic spectrum of hypertrophic cardiomyopathy associated with heterozygous carrier status for the c.3330+2T>G founder variant in *MYBPC3* in the Anabaptist communities, and constitutes the far largest cohort of individuals with the same HCM causing pathogenic variant described. The major findings of the study were: (1) clinical disease expression is variable, (2) disease expression is age-related (3) sudden death appears to occur at any age (4) disease penetrance is higher in men than in women (5) symptoms were reported by a third of gene carriers (6) The average age at onset of symptoms was 35-45 years (6) severe left ventricular hypertrophy is rare, and (7) the ECG has a low sensitivity to identify phenotype positive individuals.

The extreme variation of clinical expression of the disease in individuals and families with HCM is well described (Maron 2002, Elliott and McKenna 2004). Early studies tried to explain phenotypic variation in terms of gene-specific or mutation-specific correlates, concluding that pathogenic variants in *MYBPC3* cause a more benign phenotype with late-onset of disease (Niimura, Bachinski et al. 1998) (Charron, Dubourg et al. 1998). More recently larger cohorts of *MYBPC3* mutation carriers have been investigated in which it was shown that the extremely variable phenotype in heterogeneous HCM cohorts is also observed with pathogenic variants in *MYBPC3* (Page, Kounas et al. 2012).

This variation in phenotypic expression has also been described in moderately sized cohorts sharing the same founder HCM mutations. The majority of HCM founder mutations have been identified in *MYBPC3* in families from Finland (Jaaskelainen, Helio et al. 2013), the Netherlands (van Velzen, Schinkel et al. 2017), India (Dhandapany, Sadayappan et al. 2009), Japan (Kubo, Kitaoka et al. 2005), South Africa (Moolman-Smook, De Lange et al. 1999), Spain (Oliva-Sandoval, Ruiz-Espejo et al. 2010) (Sabater-Molina, Saura et al. 2017) France (Teirlinck, Senni et al. 2012) and Iceland (Adalsteinsdottir, Teekakirikul et al. 2014). Disease penetrance in founder mutations has been reported to be between 61% - 82% with the average maximum

left ventricular wall thickness between 20 and 25 mm. Despite comparative rates of sudden deaths in founder and non-founder mutations, founder mutations largely escape negative selection pressure and cause HCM across many generations. One of the reasons may be that disease expression in the majority of individuals is delayed until after the reproductive years when the founder variant has already been passed on to the next generation. Unfortunately, in some individuals, sudden death and severe phenotypic manifestation can occur at a much younger age consistent with the phenotypic variability described in heterogenous HCM cohorts.

This was evident in our study as 8.6% of mutation carriers reported a sudden death in a first degree relative. Sudden deaths occurred in each decade between the ages of 14 to 69 years old. Only one female individual aged 20 had a post-mortem examination which confirmed hypertrophic cardiomyopathy. Three teenagers aged 14 to 16 died suddenly during exercise and in each case one of their parents subsequently tested positive for the *MYBPC3* variant, which makes it highly likely that the children died of sudden arrhythmic death secondary to HCM, whereas the parent was asymptomatic beyond their reproductive years. Some of the other sudden death victims had previously been diagnosed with HCM. As no post-mortem examinations were performed, coronary artery disease as a contributing factor cannot be excluded in some of the older individuals. One individual who had previously undergone a myectomy had ventricular tachycardia on holter monitoring and an episode of aborted sudden cardiac death after admission to hospital before undergoing implantation of an internal cardio-defibrillator.

About a third of carriers reported significant symptoms that were deemed to likely or possibly be secondary to HCM (chest pain or shortness of breath on exertion, palpitations, and syncope on exertion or sudden syncope at rest suggestive of an arrhythmic event), and the average age of onset of symptoms was from the mid-thirties to the mid-forties both in men and women. Women and older individuals were more likely to report palpitations, and a larger proportion of older individuals reported shortness of breath (54.6%).

The six individuals who reported syncope on exertion were all male. Three of them were phenotype negative on echocardiography, and the other three had asymmetric septal hypertrophy with an average maximal wall thickness of 20.3 mm. None of the study participants had significant left ventricular outflow tract (LVOT) gradients at rest, but outflow tract gradients were not routinely evaluated after Valsalva manoeuvre or exercise. Syncope on exertion together with asymmetric septal hypertrophy may suggest latent LVOT obstruction, and together with the fact that the five individuals who had previously undergone septal myectomy were all male, this suggests that male mutation carriers are more likely to develop significant LVOT obstruction. Overall the majority of individuals reported no symptoms including two 86-year old males, one still a practicing bishop in the Amish community.

All patterns of hypertrophy were observed including sigmoid septum, reverse septal curvature, apical hypertrophy, concentric hypertrophy, and segmental hypertrophy in the anterior or lateral wall. Male individuals were more likely to have reverse septal curvature and apical hypertrophy compared to females. One male individual had progressed to “end stage” or “burnt-out” HCM associated with LV cavity dilatation, systolic dysfunction, and heart failure. This has previously been described in *MYBPC3* mutation carriers (Kubo, Kitaoka et al. 2005) (Page, Kounas et al. 2012) and is a rare complication of hypertrophic cardiomyopathy.

Overall disease penetrance (defined as a maximal wall thickness of 13 mm or more) was 50.0%. This is slightly lower than the previously reported disease penetrance of 56.9% in a cohort of individuals with different *MYBPC3* mutations (Page, Kounas et al. 2012) and may be partly due to the composition of our study cohort which contains a large proportion of “relatives” compared to “probands”. We offered screening to all first-degree relatives cascading from homozygous babies (parents typically were asymptomatic and unaware of their carrier status), any newly identified carriers and only in a small proportion of cases starting from individuals with known HCM (who would traditionally be called “probands”).

This is also reflected in the fact that only 48 out of 294 study participants (16.3%) had previously either been diagnosed with HCM or found to be carriers of the *MYBPC3* variant. In addition, at least a third of these were diagnosed as carriers of the *MYBPC3* variant as part of previous research studies rather than because they presented with symptoms suggestive of HCM, reducing the number of “probands” even more. On the other hand, a large proportion of individuals who participated in the study and were symptomatic, had not obtained medical advice. This is mainly due to the fact that Anabaptists, and in particular the more conservative Amish, mostly have no health insurance and medical care is often unaffordable or has to be shouldered by the community, which a lot of individuals try to avoid unless absolutely necessary. In different cultural circumstances, the proportion of previously clinically diagnosed individuals would therefore likely be higher.

Only a small percentage of individuals (5.4%) had a history of atrial fibrillation (AF) or were in AF at the time of the assessment. This again may be due to the structure of this cohort but is probably also confounded by the fact that very few individuals would be under regular cardiological follow-up and paroxysmal atrial fibrillation is likely underdiagnosed. Even though about 20% of the participants have previously seen a cardiologist, few would be under 12-18 monthly follow-up with repeat holter monitoring as is standard care in the UK and USA for individuals affected by HCM. The same applies for regular assessments of risk of sudden death, which is reflected in the very low percentage of individuals having internal cardio-defibrillators.

Disease penetrance was higher in men than women, which is consistent with previous reports. Penetrance in male participants was 61.7% and in female participants 39.1%, which again is slightly lower compared to a cohort of different *MYBPC3* mutation carriers, in which penetrance in males was 65.1% and 48.1% in females (Page, Kounas et al. 2012). The average maximum wall thickness in males measured 15.0 +/- 4.1 mm (range 9-28 mm) compared to 13.3 +/- 3.8 mm in females (range 8-26 mm). It has previously been speculated that biological gender specific factors, such as chromosomal or hormonal factors may influence the pathogenesis of

hypertrophy as seen in experimental animal models (Malhotra, Buttrick et al. 1990). Gender specific genetic or epigenetic factors and environmental or behavioural factors, like the amount of physical activity, may also be relevant. Another factor may be the use of the cut-off value of 13 mm to clinically diagnose HCM. A previous study reported that the gender differences in wall thickness disappeared when values were indexed for body surface area (BSA), which may mean that women tend to be clinically under-diagnosed or men over-diagnosed (van Velzen, Schinkel et al. 2017). The wall thickness measurements in this study were not adjusted for BSA, however, there was no linear correlation between maximum wall thickness and BSA. American recommendations for cardiac chamber quantification include different normal ranges for septal and posterior wall thickness in females (6-9 mm) and males (6-10), which may be another reason to consider different cut-off values for the diagnosis of HCM in females and males.

Average maximum wall thickness (MWT) in clinically affected individuals (defined as maximal wall thickness ≥ 13 mm) was 17.0 ± 3.5 mm. This was very similar in females (16.9 ± 3.4 mm) and males (17.0 ± 3.7 mm), which was a little surprising, given disease penetrance is significantly higher in males. It seems that once hypertrophy is present, the degree of hypertrophy may be similar in males and females. Overall maximum wall thickness again seems lower in our cohort compared to previous studies in a cohort of different *MYBPC3* mutation carriers and other founder mutation carriers, which again may be due to the large percentage of asymptomatic relatives included as outlined above, and possibly also due to a degree of natural selection, with severely affected individuals deceased given the limited access to medical care and awareness of the condition in the community.

Of note is that none of the mutation carriers had a wall thickness above 28 mm and only 35/266 (13.2%) of gene carriers had a wall thickness of 20 mm or above. The average age of gene carriers with a wall thickness of 20 mm or more was 53.5 ± 17.3 years (range 18 to 85 years). The fact that severe hypertrophy was rarely observed may be due to a number of factors. It is conceivable that limited

hypertrophy is a characteristic of the c.3330+2T>G *MYBPC3* variant due to its innate properties and consequences on protein levels, irrespective of the influence of other modifying factors. It may also be possible, that the genetic background of the Amish community is enriched for “protective” modifiers, that influence the phenotypic development in a positive way with amelioration of degree of hypertrophy that develops in gene carriers. At the same time, given the founder effect in the Amish population which results in a lack of genetic diversity, certain genetic modifiers that are present in the “outside” general population and exacerbate the phenotype in HCM, may be absent in the Anabaptist communities. In addition, the distinct Amish lifestyle with rural living, home-grown food and avoidance of many modern commodities of life may also have an influence on the phenotypic expression. Another contributing factor may be a degree of survivor bias, as severely affected individuals may have passed away without seeking medical attention and being assessed for risk of sudden cardiac death, and therefore not receiving appropriate treatment including primary prevention automated defibrillator implants.

Even though there was no linear correlation between age and MWT, older individuals on average had larger MWT compared to younger individuals (15.6 +/- 4.3 mm vs. to 12.7 +/- 3.2 mm) and a significantly larger proportion of older individuals had a wall thickness of 13 mm or more compared to younger carriers (65.9% vs. 35.0%). The proportion of gene carriers with clinical HCM (defined as wall thickness $\geq$ 13 mm) steadily increased with each decade reaching 100% in gene carriers above the age of 77 (figure 5-6). Older individuals had larger BSA, higher systolic blood pressures and a larger proportion had hypertension, however again there was no association between these and MWT.

The ECG was abnormal (defined as a Romhilt-Estes score of $\geq$ 4, or the presence of pathological Q waves, or T wave inversion, or ST depression) in a quarter of all gene carriers. Only 7.1% of gene carriers had a Romhilt Estes score $\geq$ 5 indicating left ventricular hypertrophy. Again, ECG abnormalities were significantly more common in older gene carriers compared to younger gene carriers (34.8%vs. 13.6%). Overall

the positive predictive value of an abnormal ECG was 78.4% and the negative predictive value of a normal ECG only 60.1%. This means a relatively large proportion (39.9%) of individuals with a normal ECG had a max WT $\geq$ 13 mm, indicating that an ECG is not a good screening test for the presence of hypertrophy in this cohort. Disease penetrance based on echocardiographic measurement of a WT $\geq$ 13 mm compared to being based on an abnormal ECG was higher in gene carriers at any age (figure 5-6 and 5-7). This is in contrast to previously reported higher disease penetrance based on ECG abnormalities compared to echocardiographic criteria, indicating that ECG abnormalities develop prior to significant hypertrophy on echocardiography and are indicative of early disease manifestation (Page, Kounas et al. 2012). This does not seem to be the case in our cohort. Despite similar criteria in assessing ECG abnormalities, this difference may partly be explained by differences in methodology and inter-observer variability, however, could also represent a characteristic feature of c.3330+2T>G *MYBPC3* carriers. Screening for HCM in the Amish communities with an ECG alone would therefore not be a reliable strategy and underdiagnose a significant proportion of affected individuals. Even though conventionally as definitely abnormal accepted ECG criteria were relatively rare, fragmentation of the QRS complex (defined in 3.2.2.3) was quite common in gene carriers (67.4%) and may represent an early manifestation of the disease process in c.3330+2T>G *MYBPC3* carriers.

On holter monitoring, older carriers had a significantly larger number of premature ventricular ectopics compared to younger individuals (1108.2 $\pm$ 2002.9 vs. 11.6 $\pm$ 47.5) and a larger proportion of older carriers had episodes of non-sustained ventricular tachycardia of four beats or more (24.2% vs. 3.3%). This finding is consistent with the age-related penetrance of HCM and the greater degree of hypertrophy and therefore arrhythmogenic substrate observed in older individuals. A significantly larger proportion of female carriers had non-sustained ventricular tachycardia on holter monitoring than male carriers (21.2% vs. 6.3%). This may imply that female carriers are at increased risk of ventricular arrhythmias, however these

individuals were not matched for age, degree of hypertrophy or other risk factors for life-threatening arrhythmias, and therefore this may be a spurious observation. A larger number of females also reported sudden syncope at rest which may be suggestive of an arrhythmia, however this difference was not statistically significant.

In summary, this Anabaptist population with the HCM causing founder mutation c.3330+2T>G in *MYBPC3* provides a unique opportunity to observe phenotypic disease expression having controlled for genetic heterogeneity. Living in a geographically defined area with a relatively homogenous lifestyle, also diminishes the environmental influences on phenotypic expression. Despite this, the cohort displays a broad spectrum of clinical disease with no specific phenotype apparent. Disease penetrance was 50% and age related, all morphological variants of cardiac hypertrophy were observed, and sudden death seems to occur at any age. The fact that disease expression can vary to such a degree among carriers of the same mutation, confirms that disease modifying factors are very powerful and may exert as much of an effect on disease expression as the mutation itself.

Longitudinal follow-up data on the carriers of the c.3330+2T>G variant in *MYBPC3* in this study would give crucial insight into the natural disease progression. It may be possible to identify high risk features at an early stage in the disease which could prompt recommendations for closer follow-up and early treatment. On the other hand, certain low risk features may predict a benign prognosis and individuals could be reassured and only followed- up very infrequently. This would have the added benefit of reduced health care costs for the Amish given a minority have health insurance. It would also be interesting to look for potential environmental or behavioural modifiers of disease expression, such as diet, personality types, perceived level of stress, or activity levels.

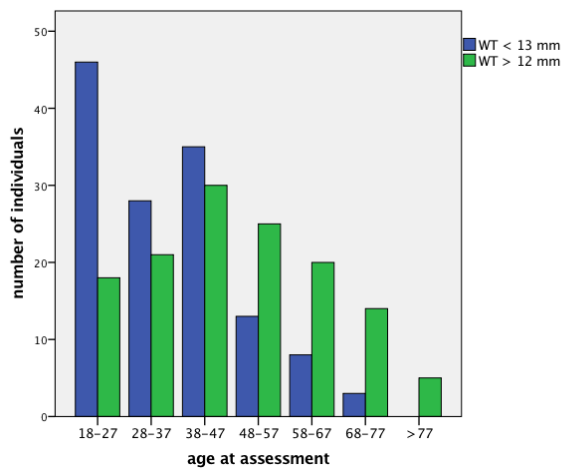


Figure 5-6: Age related penetrance of clinical HCM in c.3330+2T>G MYBPC3 carriers based on wall thickness measurement

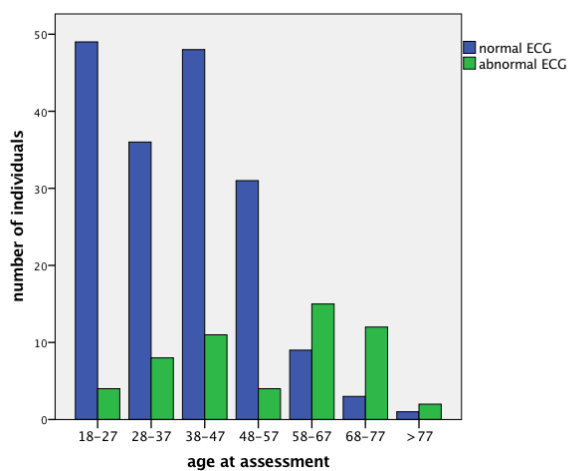


Figure 5-7: Age related penetrance of clinical HCM in c.3330+2T>G MYBPC3 carriers based on ECG criteria

5.4 Investigation of the transcriptional outcomes deriving from the pathogenic c.3330+2T>G variant in *MYBPC3*

5.4.1 Specific methods

48 carriers of the *MYBPC3* variant were selected at random (24 with mild phenotype, 24 with significant phenotypic expression) to be included in the transcript study. The predicted splicing outcomes of the founder *MYBPC3* variant were investigated by RTPCR and isoform-specific qRTPCR. Peripheral blood samples were taken from 10 UK controls, 10 Amish controls and the 48 Amish cases into PAX gene tubes (Qiagen, Crawley, UK). RNA was extracted using Qiagen PAX gene blood RNA kit handbook for extraction and purification of RNA. 200 ng total RNA from each sample was reversed transcribed in 20 µL reactions using the Superscript III VILO kit (Thermo Fisher, Foster City, USA). Initially, M13-tagged primers were designed to amplify the entire *MYBPC3* transcript in overlapping fragments. Individual PCR products from each primer set were run on 0.8% agarose gels to visualise bands corresponding to different splice products. Individual splice products were isolated by removing small samples from the gels using a pipette tip and subsequent reamplification of DNA from the agarose gel plugs (bandstab PCR). The abundance of the consensus *MYBPC3* transcript and each novel splice variant was determined by isoform-specific qRTPCR, using probes targeted across unique exon boundaries. Assays were validated for accuracy and efficiency by standard curve analysis using seven serial 1:2 dilutions of peripheral blood cDNA.

The expression level of the consensus *MYBPC3* isoform and all novel transcripts identified were then quantified in a series of 10 UK controls, 10 Amish controls and 48 Amish cases using the comparative Ct approach Pfaffl. Expression levels were calculated relative to the mean crossing point of three endogenous control genes β -glucuronidase (GUSB), ubiquitin C (UBC) and β -2-microglobulin (B2M), which had been empirically determined to be stable between groups and were expressed relative to the median level of the consensus *MYBPC3* isoform in Amish controls and

UK controls combined. Levels of consensus and novel *MYBPC3* transcripts in a pooled sample of heart RNA (Clontech, UK) were also assessed using the same techniques for comparison. We then assessed whether levels of *MYBPC3* transcripts were associated with patient phenotype as defined by maximum wall thickness and diastolic septal thickness in parasternal long axis view on echocardiography by One way ANOVA, with adjustment for age, sex, BMI, systolic blood pressure. Statistics were carried out using SSPSv21.

Transcript studies were undertaken in collaboration with Prof Lorna Harries at University of Exeter. The studies were carried out working closely with Dr Ben Lee and Dr Gaurav Harlalka based at the University of Exeter.

5.4.2 Results

To gain insight into transcriptional outcomes associated with the *MYBPC3* c.3330+2T>G gene variant, isoform-specific qRT-PCR was used to investigate the presence and abundance of distinct alternatively processed mRNA transcripts in gene carriers and controls. Blood samples from 48 Amish *MYBPC3* c.3330+2T>G carriers, 10 Amish control blood samples and 10 UK control blood samples, as well as commercially available RNA samples extracted from heart tissue from pooled unaffected individuals (Clontech, UK), were assessed. Two transcriptional outcomes were predicted to be most likely: (i) a continued reading of exon 30 into intron 30, which would result in the inclusion of an in-frame stop codon 111bp into the intron causing premature termination of protein translation, and (ii) the skipping of the 140bp comprising exon 30, predicted to disrupt the reading frame and result in protein truncation.

In order to aid this evaluation, a detailed assessment of the *MYBPC3* transcriptional profile in peripheral blood RNA from control samples using conventional RTPCR was initially undertaken. This identified six alternate *MYBPC3* transcripts: the consensus full length transcript, a transcript deleted for exons 6-9 inclusive, a transcript deleted for exons 32-34 along with part of the 3' untranslated region, a transcript deleted for

exons 17 and 18, a transcript deleted for exon 30 alone, and one deleted for exons 30 and 32. qRT-PCR analysis revealed that, in normal heart, by far the most predominant transcript corresponded to the consensus *MYBPC3* transcript, although transcripts demonstrating deletion of exons 6-9, 32-34 +’UTR or exon 30 alone were also detected at low levels. In peripheral blood only the consensus transcript, and the exon 30 deleted transcript, were detected at notable levels. The transcript deleted for exons 17 and 18 and the transcript deleted for exons 30 and 32 were not detected, and as such may be most likely to represent errors of splicing or stochastic events (figure 5-8).

In the 48 blood samples from Amish individuals heterozygous for the *MYBPC3* founder mutation, the levels of *MYBPC3* wild type transcript remained largely unaltered, and as such displayed no correlation to clinical phenotype severity (assessed as maximal wall thickness and septal thickness in diastole in parasternal long axis view on echocardiography). However, the alternative exon 30-deleted transcript is significantly up-regulated to approximately 2.5-fold that of wild type. While the precise levels of this transcript varied to a minimal degree, the exon 30-deleted transcript consistently represented the major transcript in gene carriers (figure 5-9).

The clinical parameters of patients with particularly high levels of the exon 30 del transcript were then compared to those with lower than average levels, to assess for correlation with clinical phenotypes. However, no correlation between phenotype and isoform expression levels was apparent.

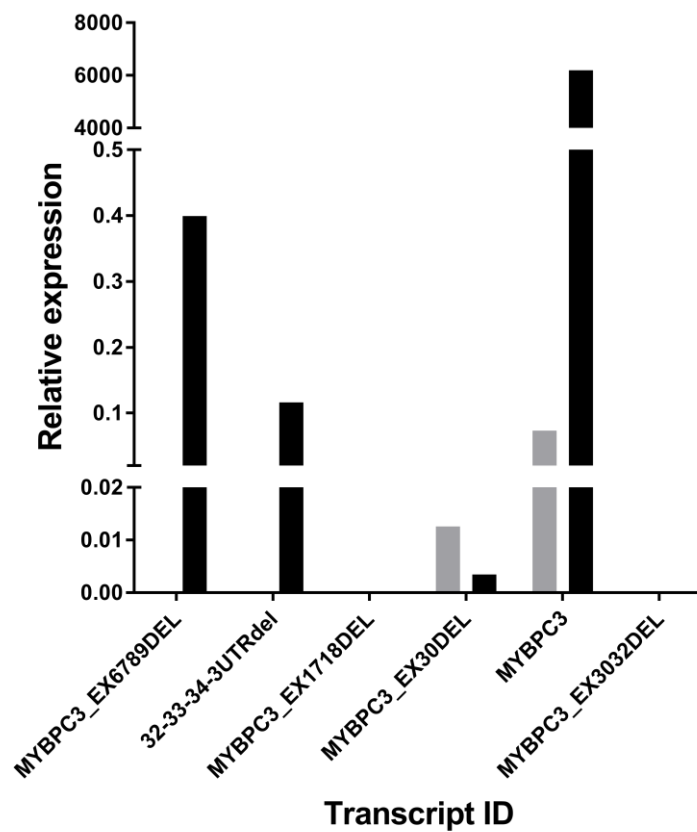


Figure 5-8: Levels of *MYBPC3* wildtype transcripts in blood and heart of controls

Displaying the abundance of each detected MYBCP3 transcript in peripheral blood (grey) and heart tissue (black). Transcript abundance is given on the Y axis and refers to relative quantification (RQ) values calculated from the abundance of each transcript relative to the Mean of the beta-2-microglobulin (B2M), beta glucuronidase (GUSB) and ubiquitin C (UBC) endogenous control genes and normalised to the median level of the wild type *MYBCP3* transcript in control samples. The identity of each transcript tested is given on the X axis.

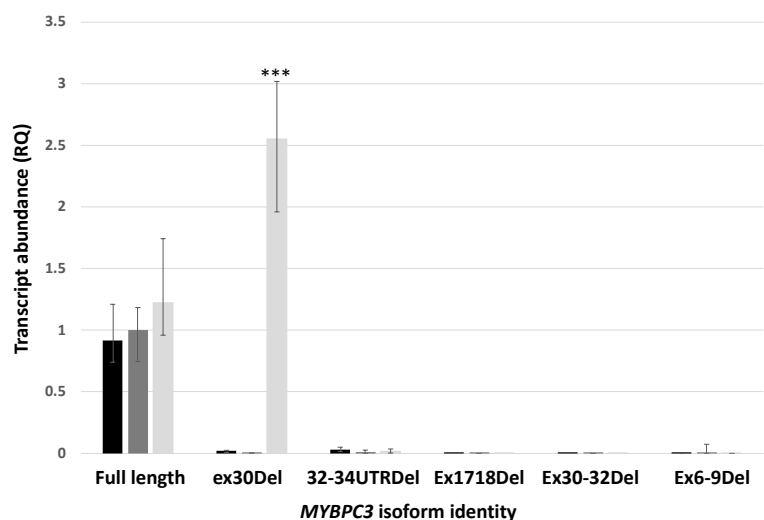


Figure 5-9: The distribution of *MYBPC3* isoforms in peripheral blood RNA from controls and mutation carriers

Displaying the abundance of each detected *MYBPC3* transcript. Transcript abundance is given on the Y axis and refers to relative quantification (RQ) values calculated from the abundance of each transcript relative to the mean of the beta-2-microglobulin (B2M), beta glucuronidase (GUSB) and Ubiquitin C (UBC) endogenous control genes and normalised to the median level of the wild type *MYBPC3* transcript in control samples. The identity of each transcript tested is given on the X axis. Transcript levels in UK controls, Amish control and Amish mutation carriers are given in black, mid grey and light grey respectively. Error bars refer to the standard deviation of expression between subjects. *** = $p < 0.0001$.

5.4.3 Discussion

The aim of this aspect of the study was to investigate the transcriptional outcomes deriving from the pathogenic c.3330+2T>G variant in *MYBPC3*. As heart specimens of *MYBPC3* c.3330+2T>G mutation carriers were unavailable, we assessed the transcriptional outcomes in blood of heterozygous carriers of the *MYBPC3* c.3330+2T>G mutation carriers to see if there may be a correlation with disease severity. If so, apart from furthering our understanding of the pathological processes involved in truncating *MYBPC3* mutations and HCM, this information may form a biomarker for the degree of hypertrophy and phenotype severity in HCM that would be relatively easily obtainable and measurable from peripheral blood of gene carriers. It may also be useful as a marker of disease progression, or to assess the response to disease modifying therapies.

Previous studies have shown that *MYBPC3* is expressed at high levels in the heart, with only low levels of transcript being detected in peripheral blood (<https://www.gtexportal.org/home/gene/ENSG00000134571.6>). To corroborate this, we firstly studied transcript levels of *MYBPC3* in blood and heart tissue of control individuals. This confirmed a distinct transcription profile in each tissue, revealing notably increased expression levels of the consensus transcript in heart compared with blood (figure 5-8).

Our studies of *MYBPC3* transcript outcomes in blood also determined that in gene carriers, the key transcripts present are the consensus transcript, and an additional transcript deleted for exon 30. Interestingly the levels of WT transcript did not differ notably between controls and carriers of the c.3330+2T>G variant, indicating possible up-regulation of WT transcript in gene carriers. Also of note was that the exon 30 deleted transcript was identified as the most dominant transcript being present at ~2.5x level of that of WT transcript (figure 5-9). Finally, our studies to investigate a putative correlation between levels of WT transcript and/or exon 30 deleted transcript in the blood of c.3330+2T>G *MYBPC3* gene carriers with clinical outcomes (ventricular wall thickness), failed to identify any significant associations.

These results confirm that there are likely very distinct transcript regulatory mechanisms in blood compared to the heart, and that it is not possible to conclusively correlate the transcript outcome in the blood with the clinical outcome in the heart in carriers of the c.3330+2T>G variant in *MYBPC3*.

Multiple independent studies have found that in cardiac tissue of patients with truncating mutations in *MYBPC3* on average 70% of wild-type MYBPC3 protein present in control individuals is expressed (Jacques, Hoskins et al. 2008, Marston, Copeland et al. 2009, van Dijk, Dooijes et al. 2009, van Dijk, Paalberends et al. 2012, Helms, Davis et al. 2014, van Dijk, Boontje et al. 2014, McNamara, Li et al. 2017). In addition, truncated protein has not been detected in cardiac tissue (Jacques, Hoskins et al. 2008, Marston, Copeland et al. 2009, van Dijk, Dooijes et al. 2009). Based on these observations, the disease mechanism originally assumed to underlie cardiac hypertrophy arising from *MYBPC3* gene mutation of truncated MYBPC3 acting as a poison peptide within the sarcomere (Yang, Sanbe et al. 1998) has been rejected and haploinsufficiency is considered the most likely mechanism driving disease in these cases. It then also follows that, even if our finding of pre-dominant levels of exon 30 deleted transcript in blood of *MYBPC3* c.3330+2T>G mutation carriers was replicated in heart tissue, truncated protein would unlikely be synthesised, or undergo rapid degradation.

In addition, a reduction in the abundance of premature-termination codon (PTC)-containing mutant transcript relative to wild-type has been observed in human heart in truncating *MYBPC3* mutation carriers. Mutant transcripts have been reported to constitute 0-44% of the total mRNA pool (Marston, Copeland et al. 2009, van Dijk, Dooijes et al. 2009, Helms, Davis et al. 2014) indicating allelic imbalance favouring an increased ratio of wild-type to mutant transcript. This indicates that it is unlikely that the exon 30 deleted transcript would be the pre-dominant transcript in heart tissue of *MYBPC3* c.3330+2T>G mutation carriers as we observed in the blood. Allelic imbalance has been observed in the transcription of 25% of human genes (Vidal, de Souza et al. 2011). Mechanisms resulting in allelic imbalance can originate at the

DNA, mRNA, and protein levels (Gaur, Li et al. 2013). Epigenetic modifications including gene imprinting or differential methylation/acetylation associated with cis-acting mutations on one allele may influence accessibility of a gene to transcript machinery. Further, mutations in trans-acting regulatory machinery such as microRNA, transcription factors, and proteins can interact with wild-type and mutant alleles unequally. For example, pathogenic variants may lead to decreased transcript stability by affecting the 3'-untranslated region of microRNA binding on the mRNA level.

Furthermore, mutant mRNA transcripts with PTC are susceptible to nonsense-mediated mRNA decay (NMD) (Vignier, Schlossarek et al. 2009, Rivas, Pirinen et al. 2015). It is conceivable that there are significant differences in the degree of allelic imbalance and NMD in *MYBPC3* truncating mutations dependent on tissue type. In addition it has previously been shown that 5-25% of nonsense transcripts escape degradation (Mort, Ivanov et al. 2008), which may be contributing to the abundance of exon 30 deleted transcript in blood of *MYBPC3* c.3330+2T>G mutation carriers. Further, the overall low expression of *MYBPC3* in blood (in addition to potential allelic imbalance favouring the wild-type allele in mutation carriers) may be one of the reasons why we did not detect a difference in the expression levels between controls and mutation carriers.

The previously observed levels of WT *MYBPC3* in truncating mutation carriers in the heart (on average 70%) are higher than the expected level of 50% in heterozygous carriers of a pathogenic variant. In addition, allelic imbalance at the mRNA level has been described with reduction in the abundance of mutant transcript relative to wild-type transcript. These observations suggest compensatory mechanisms that incompletely overcome the loss of the mutant allele. The underlying transcription factors or feedback mechanisms responsible for this compensatory mechanism are to be defined and would constitute attractive targets for disease modifying therapies, as it has been shown that an increase in wild-type *MYBPC3* protein ameliorates the phenotype of hypertrophic cardiomyopathy both in stem cell based

systems (Monteiro da Rocha, Guerrero-Serna et al. 2016, Prondzynski, Kramer et al. 2017) and animal models (Gedicke-Hornung, Behrens-Gawlik et al. 2013, Mearini, Stimpel et al. 2014).

As mentioned above, previous studies in heart tissue of truncating *MYBPC3* mutation carriers indicate that the reduction in wild-type MYBPC3 is partially compensated for by up-regulation of the wild-type allele, with on average 70% wildtype MYBPC3 compared to controls. This observation raises the interesting question of whether phenotype negative individuals with truncating mutations in *MYBPC3* have higher levels of wild-type MYBPC3, and therefore better developed compensatory mechanisms than phenotype positive individuals. It is also conceivable that a certain threshold of wild-type MYBPC3 has to be reached before the phenotype of HCM develops (Glazier, Thompson et al. 2018). In addition to genetic and epigenetic factors that influence allelic imbalance, there may be environmental or hormonal factors that trigger a reduction in wild-type transcript and protein levels. In support of this theory is the observation that stress may trigger reduction in MYBPC3 levels and development of the HCM phenotype in asymptomatic individuals, which may be one of the factors contributing to the variable age of onset of clinical HCM in MYBPC3 mutation carriers (Strande 2015). Adrenergic challenge with isoprenaline and phenylephrine infusions, and both sham operations and transverse aortic restriction surgery in *MYBPC3* animal models, induced reduction in MYBPC3 protein levels and worsening of HCM phenotype (Schlossarek, Schuermann et al. 2012, Barefield, Kumar et al. 2015), also suggesting that stress may influence phenotype severity in truncating *MYBPC3* mutations.

As part of future work on this project, it would be extremely interesting to assess *MYBPC3* expression and protein levels in heart tissue of *MYBPC3* c.3330+2T>G mutation carriers with varying degree of phenotypic expression. However, this work would be challenging, as whilst some affected individuals had surgical myectomies performed potentially providing access to cardiac tissue, the number of available samples would be limited and only represent a small spectrum of the phenotype

associated with this variant. In particular, heart tissue samples of phenotype negative individuals would likely not be available. However, one possible avenue of future research to investigate this is the creation of cardiomyocytes derived from human-induced pluripotent stem cells (hiPS-CM) from a spectrum of *MYBPC3* c.3330+2T>G mutation carriers, to assess the degree of allelic imbalance and haploinsufficiency. Previous studies suggesting amelioration of phenotype with increased levels of *MYBPC3* indicate that phenotypic severity is associated with levels of wild type *MYBPC3* and our cohort would be the ideal platform to investigate this theory given the large number of individuals carrying the same truncating *MYBPC3* mutation and the wide phenotypic spectrum observed. These hiPS-CM models could also be used to assess the influence of various putative modifiers, either genetic, epigenetic or environmental on transcript and protein levels and phenotypic expression, and the effect of potential therapies.

Another fascinating project would be to take advantage of emerging genome-wide allele specific approaches that allow genome-wide evaluation of allelic imbalance to further evaluate the influence of cis-regulatory variants and the association between allelic imbalance and disease severity in our large cohort of truncating *MYBPC3* mutation carriers with variable expression of phenotype. Overall a better understanding of the mechanisms that regulate *MYBPC3* mRNA and protein homeostasis may lead to disease-modifying treatment strategies for HCM, where natural regulatory and compensatory processes could be leveraged to restore sufficient *MYBPC3* protein levels to overcome haploinsufficiency and ameliorate or ultimately avoid the development of the phenotype of hypertrophic cardiomyopathy.

5.5 Exome sequence analysis of carriers of the *MYBPC3* c.3330+2T>G variant with severe phenotype

5.5.1 Specific methods

Whole exome sequencing on DNA of ten carriers of the *MYBPC3* 3330+2T>G variant with severe HCM phenotype was performed. Following stringent bioinformatic filtering as described above (3.4.11), the resulting variant call files were screened with a virtual panel created by combining the green genes (genes with a high level of evidence of association with the disorder) for different types of cardiomyopathies and cardiac arrhythmias from the Genomics England PanelApp [<https://panelapp.genomicsengland.co.uk>] as listed in table 5-11. After exclusion of duplicates, this virtual panel included 76 genes (appendix table 7-17 for full list of genes). Variants were then prioritised by (1) severity of effect (loss-of-function > missense > synonymous) and (2) longest transcript present. Dr J Fasham and Dr B Chioza, based at the University of Exeter, provided support for the analysis of the exome sequencing data.

Genomics England PanelApp	disorder
#47	Dilated cardiomyopathy and conduction defects (37 genes)
#49	Hypertrophic cardiomyopathy (17 genes)
#134	Arrhythmogenic cardiomyopathy (9 genes)
#238	Left ventricular non-compaction cardiomyopathy (6 genes)
#468	Cardiac arrhythmias (26 genes)

Table 5-10: Genomics England PanelApps used for filtering of variants

5.5.2 Results

In this part of the study whole exome sequencing of DNA of ten carriers of the *MYBPC3* c.3330+2T>G variant with severe HCM phenotype was performed. The aim of the study was to identify variants in the coding regions or splice-sites, that may act as potential genetic modifiers and worsen the clinical phenotype of HCM caused by the c.3330+2T>G variant in *MYBPC3*. Seven individuals from the Anabaptist communities in Ohio were included, and three carriers of the *MYBPC3* c.3330+2T>G variant were recruited by our collaborators in Wisconsin and Pittsburgh.

The individuals were selected based on the age of diagnosis and the severity of symptoms or degree of phenotypic expression on echocardiography. For example, a 14- year old girl from Ohio was included in this part of the study (ID 4005), who was not part of the adult HCM cohort described above. She was included as she had an aborted sudden cardiac arrest at the age of 10 with subsequent internal defibrillator implant. Her echocardiogram showed severe apical hypertrophy with a maximal wall thickness of 23 mm. Sadly she passed away following recruitment into the study aged 16. Also included are three carriers of the *MYBPC3* c.3330+2T>G variant who expressed a dilated form of cardiomyopathy, which is generally referred to as “end-stage” or “burnt-out” phase of HCM (ID 2481, A1750, A1837). One of these three individuals had previously also undergone a septal myectomy (A1837). Two other individuals have previously undergone septal myectomies (3725, 2487) and the remaining individuals have severe hypertrophy, some of them at a relatively young age (3187, 4186, 4260, A1840). The clinical details and investigation results of the individuals recruited from Ohio are summarised in table 5.12.

ID	age	sex	clinical characteristics	family history	symptoms	age at diagnosis	drugs	Co-morbidities
2481	70	m	dilated phase of HCM	grandson died of HCM as infant	shortness of breath, palpitations	48	Carvedilol, Digoxin, Spironolactone, Omega 3	nil
3187	18	m	significant hypertrophy in adolescence, max WT 23 mm	father known HCM; paternal uncle known HCM and SCD	palpitations since age 14	n/a	nil	nil
3725	45	m	septal myectomy aged 39	brother aged 41 HCM, significant hypertrophy (22 mm, septum), symptoms consistent with LVOT obstruction (pre-syncope on exertion)	nil	38	nil	nil
4005	14	f	aSCD aged 10, AICD, died aged 16	monozygous twin sister HCM	fatigue since age 6	10	vitamins	nil
4186	35	m	apical HCM, max WT 24 mm	father known HCM, myectomy, SCD aged 55; maternal uncle HCM	palpitations since age 30, one episode of pre-syncope associated with chest pain and nausea after exertion aged 27	34	Metoprolol 25 mg BD, Atorvastatin	Hypercholesterolaemia
4260	41	m	significant symptoms and hypertrophy at young age, max WT 22 mm, AICD	brother HCM and maternal uncle HCM and AICD	chest pain, shortness of breath and dizziness on exertion since age 25, palpitations	32	Atenolol 25 mg OD, Fluoxetine, Quetiapine, Clonazepam, 3 more antidepressants	depression
2487	68	m	Septal myectomy aged 58	Mother carrier, alive and well aged 80, declined Echo	nil	57	Metoprolol, simvastatin	Hypercholesterolaemia

ID	height [cm]	weight [kg]	average BP [mmHg]	ECG	holter	ECHO
2481	183	85.3	118/80	SR, extreme axis, 1st degree AV block (PR 268 ms), LBBB, QRS fragmentation, Q waves (I,II,aVL; UDMI), TWI (V3-V6), deep TWI (V4-V6), RE score 2	n/a	IVSd 10 mm, LVEDd 57 mm, max WT 11 mm, LVEDV BP 122 ml, LV EF 41%,
3187	181	64.5	119/61	n/a	n/a	max WT 23 mm (septum), LVEDd 45 mm, LVEF 68%, reverse septal curvature
3725	176	99.5	130/80	SR, normal axis, LBBB, deep S in V1 (34 mm), Q waves (aVR, I), TWI (I, II, aVR, V5, V6), deep TWI (aVR, V5, V6), ST depression I, II, aVF, V5, V6, ST elevation (aVR), RE score 8	2 PVCs, min HR 56 bpm, max HR 130 bpm, mean HR 78 bpm	LVEDd 48 mm, max WT 17 mm (septum), LVEF 56%, LVOT gradient at rest 5 mmHg, reverse septal curvature
4005	157.5	52.6	n/a	SR, normal axis, LA enlargement, QRS fragmentation (V6, aVf, aVL, aVR, II, I), Q wave lead III, TWI (I, II, aVR, aVL, V5, V6), deep TWI (aVR), ST depression (V1, V2), RE score 6	n/a	LVEDd 54 mm, LVESd 40 mm, LVEF 53%, max WT 23 mm (apex), apical hypertrophy
4186	183	103.5	133/83	SR, normal axis, LA enlargement, QRS fragmentation (V3-V6, II, aVF, aVL, III), TWI (V3-V6, I, aVL), deep TWI (V4-V6, I, aVL), RE score 4	261 PVCs, 3 episodes bigeminy, 1 couplet, min HR 43 bpm, max HR 111 bpm, mean HR 58 bpm	LVEDd 51 mm, LVEF 68%, max WT 24 mm, apical hypertrophy
4260	187	109.6	141/90	SR, normal axis, LA enlargement, QRS fragmentation (V5, aVF, aVL, I, II, III), RE score 4	n/a	IVSd 19 mm, LVEDd 35 mm, LVESd 27 mm, max WT 22 mm (septum), sigmoid septum
2487	165	94.7	133/72	SR, normal axis, 1 st degree AV block (PR duration 268 ms), LBBB, Q waves (aVR and V1), TWI (V4-V6, aVf, I, II, III), deep TWI (V5-V6, aVF, I, II, III), ST depression (II, III, aVF, V5, V6), ST elevation (aVL, aVR), RE score 5	7 PVCs, min HR 56 bpm, max HR 116 bpm, mean HR 59 bpm,	IVSd 16 mm, LVEDd 44 mm, LVESd 31 mm, max WT 25 mm

Table 5-11: Clinical characteristics of seven *MYBPC3* c.3330+2T>G variant carriers with severe phenotype from Ohio included in the exome study

Max WT: maximal wall thickness on echocardiography; (a)SCD: (aborted) sudden cardiac death; LVOT: left ventricular outflow tract; AICD: automated internal cardiac defibrillator; SR: sinus rhythm; AV block: atrio-ventricular block; LBBB: left bundle branch block; IVSd: intra-ventricular septal thickness in diastole in parasternal long-axis (PLAX) view on echocardiography; LVEDd: left ventricular end-diastolic dimension in in PLAX view; LVESd: left ventricular end-systolic dimension in PLAX view; LVEDv: left ventricular end-diastolic volume; LVEF: left ventricular ejection fraction; TWI: T wave inversion; PVC: premature ventricular contraction; HR: heart rate; RE score: Romhilt Estes score

The clinical details regarding the three individuals recruited by our collaborators are as follows:

ID A1750: This individual was recruited in Pittsburgh, and stems from a family with a malignant history of multiple sudden deaths and several relatives with dilated cardiomyopathy (DCM). He was diagnosed with DCM at the age of 16 years and a cardiac transplant was performed a few months later. Notably, an aunt who is negative for the *MYBPC3* c.3330+2T>G variant developed peri-partum dilated cardiomyopathy.

ID A1837: This 55-year old gentleman was recruited in Wisconsin. He was diagnosed with HCM at the age of 35 and had a septal myectomy aged 42. He has recurrent persistent atrial fibrillation and previously had an episode of decompensated heart failure which improved with medical management. His symptoms include shortness of breath and chest pain on exertion. He has a family history of sudden death in a cousin who collapsed whilst running and playing ball. His blood pressure was 110/77 mmHg and ECG showed LBBB (QRS duration 166 ms). Echo showed biventricular systolic dysfunction, left ventricular ejection fraction 37%, dilated left ventricle, left ventricular end-diastolic diameter 63 mm, left ventricular end-systolic diameter 50 mm. His medication comprised: Lisinopril, metoprolol, furosemide, atorvastatin, aspirin, warfarin.

ID A1840: This 37-year old man was also recruited in Wisconsin. He was diagnosed with HCM at the age of 34. He suffers from shortness of breath and chest pain on exertion. His BP was 118/60 mmHg and ECG showed sinus rhythm. Transthoracic echocardiography showed asymmetric septal hypertrophy with a maximal septal thickness of 19 mm and systolic anterior motion of the mitral valve, with no significant outflow tract gradient at rest. His current medication is metoprolol.

D	pNomen	gene	gene description	chrom	Hets (HCM)	Hom (HCM)	Hets (Amish)	Hom (Amish)	gnomad AltFreq_nfe
1	p.Glu332Ala	NEXN	Nexilin (F actin binding protein)	1	1	0	3	0	0.00319
2	p.Thr363Ile	NEXN	Nexilin (F actin binding protein)	1	1	0	0	0	0
3	p.Asp398del	CASQ2	Calsequestrin 2 (cardiac muscle)	1	1	0	3	0	0.007575
4	p.Ala28Val	TNNT2	Troponin T2, cardiac type	1	1	0	0	0	0.000663
5	p.Val34965Leu	TTN	Titin	2	2	0	3	0	0.000024
6	p.Arg33432His	TTN	Titin	2	0	1	11	0	0.000024
7	p.Asp31617Glu	TTN	Titin	2	1	0	2	0	0.003013
8	p.Thr26505Ile	TTN	Titin	2	2	0	3	0	0.000016
9	p.Pro21891Leu	TTN	Titin	2	2	0	5	0	0.000048
10	p.Trp19898*	TTN	Titin	2	2	0	0	0	0
11	p.Gly18301=	TTN	Titin	2	0	1	11	0	0.000902
12	p.Pro12977Ser	TTN	Titin	2	1	0	1	0	0.005716
13	p.Pro11593Leu	TTN	Titin	2	1	0	3	0	0.000019
14	p.Pro11492Thr	TTN	Titin	2	1	0	0	0	0.00502
15	p.Thr4957Ser	TTN	Titin	2	1	0	2	0	0.003285
16	p.Tyr3824His	TTN	Titin	2	2	0	0	0	0
17	p.Asn3721Tyr	TTN	Titin	2	1	0	0	0	0
18	p.Thr137Met	SCN10A	Sodium channel, voltage gated, type X alpha subunit	3	1	0	2	0	0.000027
19	p.Val1486Leu	ANK2	Ankyrin 2	4	2	0	6	0	0.000396
20	p.Gly603Ala	TRDN	Triadin	6	1	0	0	0	0.000065
21	p.Val675=	VCL	Vinculin	10	1	0	6	0	0.004609
22	p.Val118Met	LDB3	LIM domain binding 3	10	2	0	10	1	0.006825
23	p.Val298Met	BAG3	BCL2-associated athanogene 3	10	1	0	0	0	0.000159
24	p.Trp4Arg	CSRP3	Cysteine and glycine rich protein 3	11	1	0	5	0	0.004162
25	p.Glu1179Lys	MYBPC3	Myosin binding protein C, cardiac	11	1	0	4	0	0.000294
26	p.Asp1064GlyfsX38	MYBPC3	Myosin binding protein C, cardiac	11	10	0	0	0	0.000067
27	p.Ala210Val	SCN3B	Sodium channel, voltage-gated, type III, beta	11	1	0	1	0	0.000024
28	p.Asp208Asn	MYH6	Myosin, heavy chain 6, cardiac muscle, alpha	14	1	0	1	0	0.007759
29	p.Trp244Arg	IDH2	Isocitrate dehydrogenase (NADP(+)) 2, mitochondrial	15	1	0	0	0	0.000036
30	p.Gly880Ser	TRPM4	Transient receptor potential cation channel subfamily M member 4	19	1	0	0	0	0.000089
31	p.*124Ileext*14	KCNE2	Potassium voltage-gated channel, Isk-related family, member 2	21	1	0	2	0	0.000087

ID	pNomen	gene	Input pos	Input ref	Input alt	Coding effect	Var location	HGMD	ClinVar Clin Sign	SIFT	Revel	#id
1	p.Glu332Ala	NEXN	78395131	A	C	missense	exon	DM?	VUS, Benign	Deleterious	0.157	CM159729
2	p.Thr363Ile	NEXN	78399001	C	T	missense	exon			Deleterious	0.157	
3	p.Asp398del	CASQ2	116243874	ATCG	A	in-frame	exon					
4	p.Ala28Val	TNNT2	201337340	G	A	missense	exon	DM	Likely benign, Pathogenic, VUS	Tolerated	0.61	CM063210
5	p.Val34965Leu	TTN	179396449	C	G	missense	exon		VUS		0.43	
6	p.Arg33432His	TTN	179401179	C	T	missense	exon		VUS		0.437	
7	p.Asp31617Glu	TTN	179411207	A	T	missense	exon		VUS, Likely benign, Benign		0.362	
8	p.Thr26505Ile	TTN	179431345	G	A	missense	exon		VUS		0.25	
9	p.Pro21891Leu	TTN	179447858	G	A	missense	exon		VUS		0.653	
10	p.Trp19898*	TTN	179456938	C	T	stop gain	exon		Likely pathogenic			
11	p.Gly18301=	TTN	179467226	G	C	syn - splice	exon		Likely benign, Benign			
12	p.Pro12977Ser	TTN	179517605	G	A	missense	exon		Benign			
13	p.Pro11593Leu	TTN	179538368	G	A	missense	exon		VUS		0.421	
14	p.Pro11492Thr	TTN	179540461	G	T	missense	exon		Benign			
15	p.Thr4957Ser	TTN	179600303	G	C	missense	exon		Conflicting interpretations		0.101	
16	p.Tyr3824His	TTN	179615657	A	G	missense	exon				0.067	
17	p.Asn3721Tyr	TTN	179621042	T	A	missense	exon				0.069	
18	p.Thr137Met	SCN10A	38830507	G	A	missense	exon	DM?	VUS	Tolerated	0.316	CM147565
19	p.Val1486Leu	ANK2	114274230	G	C	missense	exon		VUS	Tolerated	0.05	
20	p.Gly603Ala	TRDN	123581780	C	G	missense	exon		VUS	Tolerated	0.002	
21	p.Val675=	VCL	75863580	G	A	syn - splice	exon		Likely benign, Benign			
22	p.Val118Met	LDB3	88441223	G	A	missense	exon	DM?	Likely benign, Benign	Tolerated	0.022	CM152475
23	p.Val298Met	BAG3	121432151	G	A	missense	exon	DM?	VUS	Tolerated	0.07	CM1711645
24	p.Trp4Arg	CSRP3	19213986	A	G	missense	exon	DM?	Conflicting interpretations	Deleterious	0.474	CM023060
25	p.Glu1179Lys	MYBPC3	47354209	C	T	missense	exon	DM?	Likely benign, VUS	Deleterious	0.464	CM132320
26	p.Asp1064GlyfsX38	MYBPC3	47354743	A	C		splice site	DM	Pathogenic, Likely pathogenic			CS034593
27	p.Ala210Val	SCN3B	123504870	G	A	missense	exon		Likely benign	Tolerated	0.29	
28	p.Asp208Asn	MYH6	23873940	C	T	missense	exon	DM?	Benign, Likely benign	Tolerated	0.076	CM165004
29	p.Trp244Arg	IDH2	90630756	A	G	missense	exon			Deleterious	0.506	
30	p.Gly880Ser	TRPM4	49700124	G	A	missense	exon		VUS	Tolerated	0.294	
31	p.*124Ileext*14	KCNE2	35743145	CCT	C	stop loss	exon	DM?	VUS			CD097716

Table 5-12: Summary of potential genetic modifiers identified in 10 carriers of *MYBPC3* c.3330+2T>G with a severe phenotype

Allele frequencies are given within the cohort of ten *MYBPC3* 3330+2T>G variant carriers with severe phenotype in heterozygous configuration (Hets HCM) and homozygous configuration (Hom HCM), within our database of 107 exome sequence files of Anabaptist individuals (non-carriers of *MYBPC3* c.3330+2T>G variant) in heterozygous (Het Amish) and homozygous configuration (Hom Amish), and within the gnomad European (non-finnish) population (gnomad AltFreq_nfe).

pNomen: protein-level nomenclature; chrom: chromosome; HGMD: Human Gene Mutation Database; DM: deleterious mutation; ClinVar Clin Sign: reported clinical significance in ClinVar submissions; SIFT (Sorting Intolerant From Tolerant): in-silico tool to predict effect of SNP on protein function, score ranges from 0.0 (deleterious) to 1.0 (tolerated); Revel: in-silico tool to predict the pathogenicity of rare missense variants, score ranges from 0 (tolerated) to 1 (deleterious); #id: HGMD variant ID;

Following stringent bioinformatic filtering of the ten exome-sequence variant files, a total of 31 candidate variants were identified in the 76 genes on the virtual cardiomyopathy/ cardiac arrhythmia panel (table 5-13, table 7-17). In addition to identifying the HCM associated variant (c.3330+2T>G in *MYBPC3*) in all 10 individuals, a number of additional variants were identified that potentially contribute to the severe disease expression in these carriers.

In one gene carrier (A1840), a second variant in *MYBPC3* was identified (p.Glu1179Lys). This variant is present at relatively high frequency within the Amish community in non-carriers of the c.3330+2T>G variant in *MYBPC3* (4/107, 3.7%) and has a gnomAD allele frequency in the European Non-Finish population (gnomAD AltFreq_nfe) of 0.000024%. The variant has previously been reported on ClinVar as likely benign and variant of unknown significance (VUS), however is predicted to be deleterious by SIFT and has a moderate Revel score of 0.464.

A total of 13 variants were identified across the cohort in *Titin*. Two of these (p.Arg33432His and p.Gly18301=) were found in a homozygous configuration in the same individual (A1837), indicating that they are located in cis on the same (likely ancestral Amish) chromosome. The missense variant p.Arg33432His is present at a high carrier frequency in the Amish (11/107, 10.3%), although at a very low frequency in gnomAD (0.000024%). It has previously been reported as a VUS on ClinVar and has a Revel score of 0.437. The p.Gly18301= variant is a synonymous splice-site variant not resulting in an amino-acid change, and also present at high carrier frequency in the Amish (11/107, 10.3%). Its frequency in gnomAD is reported as 0.000902% and it has previously been reported on ClinVar as likely benign and benign.

A further 10 missense variants in *Titin* were present in heterozygous configuration in one to two individuals each. Their carrier frequency is variable in the Amish population ranging between 0% and 4.7% (5/107), although all are rare or ultra-rare in gnomAD 90% - 0.005716%). Most notably one truncating *Titin* variant

(p.Trp19898*) was identified in two individuals (A1750 and 2481). This variant has previously been reported on ClinVar as likely pathogenic.

A missense variant (p.Ala28Val) in the sarcomeric protein Troponin T2 (*TNNI2*) was identified in one individual (3725). This variant was not identified in the Amish control cohort, and the gnomAD frequency is low at 0.000663%. This variant has previously been reported as likely benign, VUS and pathogenic on ClinVar. SIFT predicts it to be tolerated and its Revel score is 0.61.

A missense variant (p.Val298Met) in BLC2-associated athanogene 3 (*BAG3*) was identified in one individual (A1750). This variant was not been present in our Amish control cohort, and its' gnomAD frequency is reported as 0.000159%. ClinVar contains a previous report as VUS, SIFT predicts it to be tolerated and the Revel score is 0.07.

Another missense variant (p.Val1486Leu) was identified in Ankyrin 2 (*ANK2*) in two individuals (3187 and 4260). The frequency in the Amish control group is relatively high at 5.6% (6/107), and the frequency in gnomAD is 0.000396%. This variant has previously been reported as VUS on ClinVar and SIFT predicts it to be tolerated with a Revel score of only 0.05.

Three variants in sub-units of ion channels were also identified, in one individual each. The missense variant p.Thr137Met in *SCN10A* (in individual 4260), encoding the alpha subunit of the sodium channel NaV1.8, has an Amish frequency of 1.9% (2/107) and a gnomAD frequency of 0.000027%, has previously been reported as a VUS on ClinVar and is predicted tolerated on SIFT with a Revel score of 0.316. The missense variant p.Ala210Val in *SCN3B* (in individual 2487), encoding the beta subunit 3 of the sodium voltage-gated channel, has an Amish control group frequency of 0.9% (1/107) and a gnomAD frequency of 0.000024%, has previously been reported as likely benign on ClinVar and is predicted to be tolerated on SIFT with a Revel score of 0.29. The variant p.*124Ileext*14 in *KCNE2* (in individual 4005), leading to the loss of a stop-codon, has a frequency in the Amish cohort of

1.9% (2/107) and in gnomAD of 0.000087%. It had previously been reported to ClinVar as VUS.

5.5.3 Discussion

This part of the study investigated the presence of rare variants, located in coding regions or at splice sites, with a potential modifying effect on clinical outcomes in ten carriers of the c.3330+2T>G variant in *MYBPC3* with a severe clinical phenotype of HCM. In order to identify variants most likely to have a clinical impact, the analysis was focused on genes previously implicated in inherited cardiac disease. A total of 31 candidate variants in genes previously associated with cardiomyopathy or cardiac arrhythmia were identified.

In a 37-year old male gene carrier from Wisconsin (A1840) a second variant in *MYBPC3* was identified (p.Glu1179Lys, c.3535G>A, rs199669878). Individual A1840 was diagnosed with HCM at the age of 34 and echocardiography showed him to have significant asymmetric septal hypertrophy (max wall thickness 19 mm). He suffers from symptoms of shortness of breath and chest pain on exertion. The p.Glu1179Lys variant is present at relatively high frequency within the Amish community (3.7%), but has a low frequency in the European (non-Finnish) population on gnomAD (0.000024%). Notably, p.Glu1179Lys has previously been reported in association with HCM (Rodriguez-Garcia, Monserrat et al. 2010, Bick, Flannick et al. 2012, Maron, Maron et al. 2012, Kassem, Azer et al. 2013) and sudden infant death syndrome (Brion, Allegue et al. 2012). However, Rodriguez-Garcia et al. (2010) for example, only sequenced *MYBPC3* in their study, so it is unknown whether a mutation in another disease-causing gene may have been present in the affected individual investigated in the study. Similarly, Kassem, Azer et al (2013) sequenced only *MYBPC3*, *MYH7*, and *TNNT2*, and therefore the same might apply for their case. P.Glu1179Lys is also reported on ClinVar in a number of patients with different types of cardiomyopathy, but often in conjunction with a variant in another cardiomyopathy causing gene with more convincing evidence for pathogenicity (Landrum, Lee et al. 2016). Nevertheless, the p.Glu1179Lys variant

represents a non-conservative amino acid substitution, and may certainly impact on secondary protein structure as these residues differ in polarity, charge, size and potentially other properties. This substitution occurs at a position that is conserved across species and two out of three *in silico* prediction programs predict this variant to be damaging/ probably damaging to the protein structure and/or function. Conversely, there are currently 115 individuals listed in the gnomAD database (of 141,456 individuals), and in the Exome Aggregation Consortium (ExAC) and the NHLBI Exome Sequencing Project report, p.Glu1179Lys was observed in 22/15,208 alleles from individuals of South Asian ancestry and 6/8,254 alleles from individuals of European ancestry, respectively. It is not possible to conclude from these low frequencies, and the data above, whether this may represent a pathogenic or rare benign variant in these populations.

Overall p.Glu1179Lys has been classified in ClinVar as a variant of uncertain significance or as a likely benign variant (Landrum, Lee et al. 2016) based on existing information. It may be conceivable though, that, even if not on its own disease causing, the variant may in conjunction with a second (pathogenic) variant in *MYBPC3*, such as the c.3330+2T>G variant, exacerbate haploinsufficiency of *MYBPC3* or worsen the phenotypic expression of HCM via a different mechanism. It certainly is known that individuals who are compound heterozygous for two (pathogenic) variants in the same HCM causing gene present on different alleles, express more severe phenotypes, and this occurs in about 3% to 5% of individuals (Maron, Maron et al. 2012).

For future studies, it would be interesting to see whether the presence of the p.Glu1179Lys variant in carriers of the c.3330+2T>G variant in the nuclear family of individual A1840, and also in the wider community, is associated with a more severe HCM phenotype. Additionally, given the relatively high frequency of the p.Glu1179Lys variant in the Amish, clinical assessment of individuals carrying this variant alone (without the c.3330+2T>G variant) would be interesting to see whether a HCM phenotype or family history of HCM may be associated with it.

Another notable finding relates to gene carriers A1750 and 2481, who expressed a phenotype of dilated cardiomyopathy (DCM), often referred to as “end-stage” or “burnt-out” HCM. Both are carriers of the same truncating variant in titin (*TTN*; p.Trp19898*, c.59693G>A). Individual A1750 was recruited in Pittsburgh, and stems from a family with a malignant history of multiple sudden deaths and several relatives with DCM. He was diagnosed with DCM at the age of 16 years and a cardiac transplant was performed a few months later. Interestingly, one of his aunts who is negative for the *MYBPC3* c.3330+2T>G variant, developed peri-partum dilated cardiomyopathy, and made a good recovery. Individual 2481 is a 70-year old Amish gentleman who was diagnosed with HCM at the age of 48. He suffers from shortness of breath on exertion and palpitations. More recently his echocardiogram has shown a dilated left ventricle with reduced function and thin walls (maximal wall thickness 11 mm). Of note is that his grandson, who was a heterozygous carrier for the *MYBPC3* c.3330+2T>G variant, died of HCM as an infant.

The p.Trp19898* variant in *TTN* has not been previously reported in individuals with cardiomyopathy or in large population studies (ClinVar). This nonsense variant leads to a premature termination codon at position 19898, which is predicted to lead to a truncated or absent protein. Truncating variants in the giant sarcomeric protein titin (*TTN*tv) have been found to be the most frequent causes of DCM. Clinically relevant *TTN*tv have been identified in about 20% of unselected and familial DCM cases (Herman, Lam et al. 2012, Akinrinade, Ollila et al. 2015, Akinrinade, Alastalo et al. 2016). DCM-causing *TTN*tv seem to be clustered in the A-band region of the gene, whereas *TTN*tv in reference and healthy individuals are clustered in the I-band region. Due to the usage of an alternative promoter, *TTN*tv in the I-band region may be spliced out from mature long *TTN* isoforms resulting in minimal or no functional consequences (Deo 2016). If located in an exon that is highly expressed in the heart (Roberts, Ware et al. 2015), A-band *TTN*tv have been estimated to have a nearly 98% chance of being disease causing when identified in unselected DCM patients (Akinrinade, Alastalo et al. 2016). The p.Trp19898* variant is located in the

A-band of TTN in the highly expressed exon 251, and although additional studies are required to fully establish its clinical significance, the variant may therefore be likely pathogenic (ClinVar), and a likely disease causing or at least modifying factor in these two individuals who are carriers of the *MYBPC3* c.3330+2T>G variant, but express a DCM phenotype. It is conceivable that individuals with so called “end-stage” or “burnt-out” HCM in general actually harbour an unidentified DCM causing variant, which alters their clinical phenotype from HCM to DCM.

Interestingly, one of the aunts of individual A1750, who was not a carrier for the *MYBPC3* c.3330+2T>G variant, developed peri-partum cardiomyopathy. It is unknown whether she is a carrier for the p.Trp19898* variant, however if she was, this may be a significant factor as truncating variants in TTN have previously been also reported as the most prevalent genetic predisposition in peri-partum cardiomyopathy (Ware, Seidman et al. 2016). Also of note is that individual 2481 had a grandson who died of heart failure secondary to a cardiomyopathy as an infant. Unlike the majority of cases of lethal infantile cardiomyopathy in the Amish communities, the boy was not homozygous for the *MYBPC3* c.3330+2T>G variant. He was a heterozygous carrier for the *MYBPC3* variant, and it may be conceivable that in conjunction with another insult, either environmental or in the form of additive genetic variation, the truncating TTN variant contributed to the severe cardiomyopathy leading to his death. None of the Amish individuals in the control database was a carrier for the TTNtv p.Trp19898* variant. It would be interesting to assess other families with DCM in the community, who are not carriers of the *MYBPC3* c.3330+2T>G variant, to see whether the TTNtv is disease causing on its own.

We also identified 12 additional distinct missense variants in TTN (two in homozygous and 10 in heterozygous configuration) in our *MYBPC3* variant carriers. A recent European multi-centre DCM study assessed the possible role of TTN missense and NFS-INDEL (non-frameshift insertion/ deletion) variants in DCM. It showed that there is a statistically insignificant enrichment of rare TTN missense

and NFS-INDEL variants in DCM patients compared to control individuals, suggesting that TTN missense and NFS-INDEL variants do not cause DCM and should be classified as “likely benign”. The authors however also concluded, that some rare TTN missense variants might have modifier effects for the phenotype (Akinrinade, Helio et al. 2019), which again may be relevant in our cohort and deserves further exploration.

In addition to the truncating variant p.Trp19898* in Titin, individual A1750, who was diagnosed with DCM at the age of 16 and subsequently underwent cardiac transplantation, was also found to be a carrier for the missense variant p.Val298Met in BLC2-associated athanogene 3 (*BAG3*). This variant has not previously been identified in our Amish control cohort and its gnomAD frequency is reported to be low (0.000159%). ClinVar contains a previous report as VUS, SIFT predicts it to be tolerated, and its Revel score is 0.07. However, mutations in *BAG3* have recently been described in DCM, and seem to be characterised by high penetrance and a high risk of progressive heart failure (Dominguez, Cuenca et al. 2018). *BAG3* protein has been shown to be essential for the stability of the cardiac sarcomere (Ulbricht and Hohfeld 2013) where it interacts with the actin capping protein CapZ, stabilizing the myofibril structure in response to mechanical stress (Hishiya, Kitazawa et al. 2010). Further characterisation of this variant is necessary, in addition to segregation studies in this large family cohort with multiple members with DCM and a malignant history of sudden deaths, but it is very conceivable that both the TTNtv p.Trp19898* and the *BAG3* p.Val298Met variant play an important role in the pathogenesis of the different cardiomyopathy phenotypes and the high incidence of sudden death in this family.

In two individuals (3187 and 4260) a missense variant in Ankyrin 2 (*ANK2*) was identified (p.Val1486Leu, c.4456G>C). Individual 3187 is only 18 years of age and already expresses severe asymmetric septal hypertrophy with a reverse septal curvature phenotype, and a maximal wall thickness of 23 mm. He has a family history of HCM in his father, and of HCM and sudden death in a paternal uncle. He

was unaware of his diagnosis until taking part in this study, however, has been experiencing palpitations since the age of 14. Individual 4260 was diagnosed with HCM at the age of 32. He had been experiencing chest pain, shortness of breath and dizziness on exertion since the age of 25. He is now 41 and has an automated internal cardio-defibrillator (AICD). His echo showed a sigmoid septum with a maximal wall thickness of 22 mm. His brother also has known HCM, and so has a maternal uncle who also has an AICD.

The frequency of the *ANK2* variant p.Val1486Leu in our Amish controls is relatively high at 5.6%, with a gnomAD frequency of only 0.000396%. This variant has previously been reported as a VUS on ClinVar, and algorithms developed to predict the effect of missense changes on protein structure and function including SIFT, PolyPhen-2, Revel and Align-GVGD suggest that this variant is likely to be tolerated, but these predictions have not been confirmed by published functional studies. This sequence change replaces valine with leucine at codon 1486 of the *ANK2* protein. The valine residue is weakly conserved and there is only a small physico-chemical difference between valine and leucine. The variant is also present in other population databases (rs149678604, ExAC 0.03%) but has not been reported in the literature in individuals with an *ANK2*-related disease. In summary, this variant is a rare missense change that is not predicted to affect protein function and is found in the population at an appreciable frequency and at relatively high frequency in our Amish cohort. Overall the variant is not anticipated to cause disease, however, the available evidence is currently insufficient to confirm this conclusively. Therefore, it has been classified as a variant of uncertain significance on ClinVar.

A previous study investigating the hypothesis that variation in non-sarcomere genes may influence the phenotype in HCM, reported that patients with HCM with rare variants in *ANK2* (half of which also carried variants in sarcomeric proteins), had greater maximum left ventricular wall thickness (MLVWT) ($p=0.0005$). The proportion of patients with a MLVWT ≥ 30 mm was also greater in carriers of an *ANK2* rare variant ($p<0.0018$). This was also true when considering individuals who

carried only a sarcomeric gene variant (Lopes, Syrris et al. 2015). This observation would fit with the finding in this study that the two individuals with an *ANK2* variant had above average wall thickness at a young age. *ANK2*, or ankyrin B, stabilises membrane ion-channels in cardiomyocytes and pathogenic variants in the gene have been associated with long QT syndrome 4, ventricular arrhythmias and sinus node disease (Webster and Berul 2013). Ankyrins interact with proteins influencing calcium homeostasis and β -adrenergic signalling, and it is therefore possible that they affect the cellular phenotype resulting from a sarcomeric gene variant.

In conclusion, a number of putative disease-modifying variants in genes previously associated with cardiomyopathies or cardiac arrhythmia have been identified in the ten *MYBPC3* c.3330+2T>G variant carriers with a severe type of hypertrophic cardiomyopathy. These variants deserve further investigation with segregation and association studies with phenotype severity within the nuclear family cohorts and the larger community cohort of *MYBPC3* gene carriers to determine their potential to influence the phenotypic expression in *MYBPC3* c.3330+2T>G HCM. The relatively high frequency of most variants in the Amish, where each likely represents a founder mutation, empowers genetic studies and allows the relevance of each variant to disease to be more conclusively determined. If association with disease severity in the Amish is confirmed, association with phenotypic expression in other HCM cohorts should be investigated, to establish a potential modifying role in HCM caused by different variants and other genes. Functional studies with cardiomyocytes derived from human-induced pluripotent stem cells (hiPS-CM) harbouring the c.3330+2T>G *MYBPC3* variant with and without the identified variants would allow the assessment of the effect of the putative modifiers on haploinsufficiency of *MYBPC3*, disease severity and the mechanisms involved. An alternative way of assessing the effect of putative genetic modifiers would be the development of a mouse model carrying the c.3330+2T>G *MYBPC3* variant. An additional line of future work would also include exome sequence analysis of additional individuals with severe phenotypic expression. This would likely reveal

further variants potentially involved in disease expression. It would also be worthwhile to investigate individuals with very mild phenotype at an old age, to determine whether “protective” variants may be present that may contribute to alleviate haploinsufficiency of MYBPC3 or disease severity via other mechanisms. The number of potential disease-modifying variants identified in this small cohort of ten individuals alludes to the complexity likely underlying the mechanisms involved in the biology of the disease and the variable expression of the phenotype.

5.6 Investigation of SNP association with phenotype severity in the *MYBPC3* c.3330+2T>G variant cohort

5.6.1 Specific methods

202 carriers of the *MYBPC3* c.3330+2T>G variant were selected at random for genotyping and inclusion in the genome wide SNP study. The study was performed with support of Dr Michael Weedon and Dr Sam Jones, based at the University of Exeter.

The samples were genotyped on Affymetrix Genome-Wide Human SNP Array 6.0 including 909,626 SNPs by EpiGenDx (www.epigendx.com). Genotypes were manually converted to PLINK (PMID: 17701901) format from the birdseed files. PLINK was used to remove variants with 5% or more missing genotypes and Hardy-Weinberg equilibrium $P < 1E-6$, after which 661,864 variants remained for imputation. Individuals with evidence of a mismatch between genetically determined sex and reported sex were removed using X-chromosome SNPs. The genotype data was phased with SHAPEIT 2 with the use of a genetic map from 1000 Genomes phase 3. The genotypes were imputed using IMPUTE2 using the 1000 Genomes reference set. A total of 20,608,282 autosomal variants were imputed. A genome wide association study was performed using GCTA (v 1.26.0) to account for relatedness in a maximum of 170 samples. variants with quality <0.9 and minor

allele frequency <0.5% were excluded, leaving a maximum of 6,170,969 variants for any one GWAS. All continuous phenotypes were rank-normalised to ensure a normal distribution when running the GWAS. We adjusted for the co-variables age, gender, systolic blood pressure and Body surface area at GWAS runtime. Because of the small number of samples for binary traits, we performed Fisher's exact test in Stata v13 as a sensitivity check for any variants that were genome-wide significant in the GCTA analysis.

The following phenotypes were used for the association study:

- (1) Presence or absence of deep T wave inversion on ECG
- (2) Romhilt Estes score
- (3) maximal wall thickness on echocardiography
- (4) mitral wall thickness score

The mitral wall thickness score was the sum of the wall thickness measurements of the septal, anterior, lateral and inferior wall at the mitral level in the parasternal short axis view.

- (5) maximal wall thickness score

The maximal wall thickness score was the maximum value out of the three wall thickness scores (mitral, papillary and apical score). The individual scores at different levels of the left ventricle were derived by adding up the wall thickness in the four segments (septal, anterior, lateral, inferior) as measured in the parasternal short axis view.

- (6) relative wall thickness mitral and papillary level

This measurement was the sum of the relative wall thickness at mitral and papillary level. The relative wall thickness at mitral/papillary level was obtained by dividing the mitral/ papillary wall thickness score by the sum of the septal-lateral and anterior-inferior diameter at mitral/ papillary level.

5.6.2 Results

Statistical analyses of genome-wide SNP data against two dichotomous and four continuous cardiomyopathy-related phenotypes were performed as defined in 5.3.4 (deep T wave inversion, Romhilt Estes score, maximal wall thickness, maximal wall thickness score, mitral wall thickness score, relative wall thickness mitral and papillary) on a maximum of 6,170,969 imputed variants to identify potential common variant genetic modifiers. GCTA (PMID: 21167468) was used to attempt to account for relatedness in the association analyses. This identified only a single significant genomic region genome-wide on chromosome 2q36.1, involving seven adjacent SNPs (table 5-14) for the dichotomous “deep T-wave inversion ≥ 2 mm” phenotype (Figure 5-10). The SNP, rs186189876 was associated with an odds ratio of 0.39 ($P=3.1 \times 10^{-9}$) for the uncommon A allele. The nearest genes to the association signal are *EPHA4* and *PAX3*. It may be that the SNP influences a regulatory region for one of these genes or the signal represents a founder haplotype which includes a coding variant in one of these genes. However, the association signal needs to be viewed with some caution because of the low allele frequency entailing only 5 heterozygotes in 145 samples included in this study. When using a Fisher’s exact test as a sensitivity test (which does not account for relatedness but is a more appropriate statistical test for small sample sizes), the association falls below the genome-wide significant threshold ($P=3.1 \times 10^{-5}$). Replication in an independent or larger dataset should therefore be sought for this association signal.

Chr	SNP	bp	A1	A2	Freq	b	se	p	N
2	rs143308503:222547780:C:T	222547780	C	T	0.975666	-0.758712	0.133341	1.27E-08	145
2	rs146673543:222607101:A:G	222607101	A	G	0.975886	-0.760567	0.133313	1.16E-08	145
2	2:222618929:C:A	222618929	C	A	0.984555	-0.996631	0.172359	7.37E-09	145
2	2:222619686:T:C	222619686	T	C	0.984555	-0.996631	0.172359	7.37E-09	145
2	rs78732070:222634757:A:G	222634757	A	G	0.975886	-0.760567	0.133313	1.16E-08	145
2	rs186189876:222642998:G:A	222642998	G	A	0.982803	-0.937092	0.158056	3.05E-09	145
2	rs187941303:222689859:A:G	222689859	A	G	0.984941	-1.0252	0.177017	6.97E-09	145

Table 5-13: The seven adjacent SNPs in genome-wide significant genomic region on chromosome 2q36.1

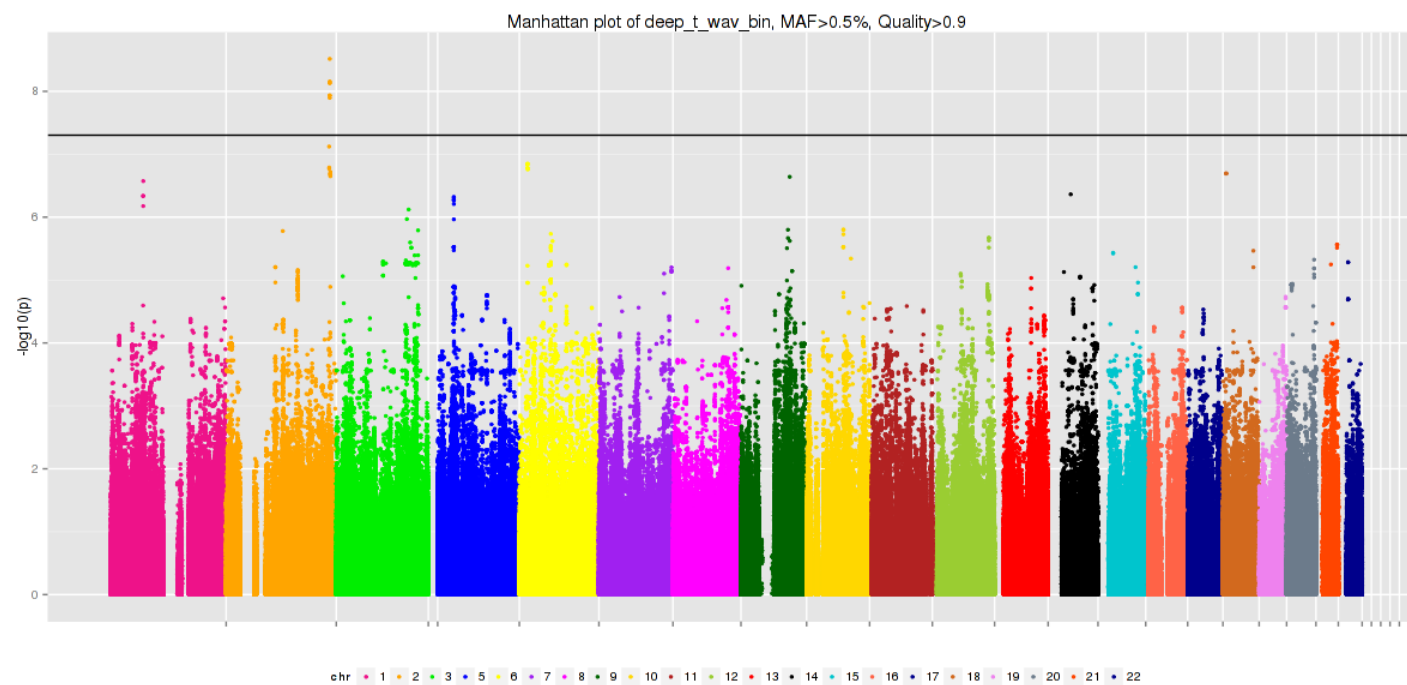


Figure 5-10: Manhattan plot for the “deep T-wave inversion” phenotype

5.6.3 Discussion

Association studies measure the correlation between a genotype at a marker locus and disease phenotype, with the aim to determine whether a particular allele correlates with disease in the population studied. This correlation could indicate a direct causal relationship between the genotyped marker allele and the phenotype which allows the investigation of mechanisms and pathways in disease progression. Alternatively, a positive association may indicate an indirect relationship, due to linkage disequilibrium (LD) between the test variant and alleles at a nearby causal variant. In this case the association study may aid in localising the causal variant.

In the past association studies were limited to candidate genes, but with the advances in high-through-put genotyping the age of whole genome association studies (GWAS) was entered, and association studies commonly now involve 500,000 to 4 million markers across the genome. Although this has increased the power of association studies considerably, it does give rise to a number of bioinformatic and multiple-testing issues given the large datasets. Microarrays used for genotyping were developed incorporating so called tagging SNPs which were chosen based on information derived from the HapMap project (<http://hapmap.ncbi.nlm.nih.gov/>) to improve efficiency of the platforms at the expense of reduced redundancy.

GWAS have been hugely successful in identifying common variants associated with complex diseases. A number of high profile GWAS identified susceptibility loci for breast cancer, rheumatoid arthritis, type 1 and 2 diabetes, atrial fibrillation, obesity and many other common diseases (Easton, Pooley et al. 2007, Plenge, Seielstad et al. 2007, Todd, Walker et al. 2007, Zeggini, Scott et al. 2008, Nielsen, Fritsche et al. 2018, Turcot, Lu et al. 2018). The success of these studies has largely been due to large sample sizes (>2000 cases), stringent quality control procedures for discarding low-quality SNPs and/or samples and the use of stringent significance thresholds ($p = 5 \times 10^{-8}$) to account for multiple testing and low prior probability of true effect. Another important consideration in GWAS studies is the replication in an independent data set, ideally from the same population studied.

GWAS have also been used in the search for genetic modifiers of disease severity and have been successful both in outbred as well as isolated founder populations (Wright, Strug et al. 2011, Zhang, Stein et al. 2012). A recent GWAS in 139 Ashkenazi Jewish patients with type 1 Gaucher disease, homozygous for the N370S mutation, demonstrates the power of association studies in isolated founder populations. In this study, several SNPs in linkage disequilibrium within the *CLN8* gene locus were found to be associated with Gaucher disease severity; the most significant SNP rs11986414 at a p-value of 1.26×10^{-6} . Even though this p-value is lower than the stringent genome wide significant p-value ($p = 5 \times 10^{-7}$ to $p = 5 \times 10^{-8}$), subsequent functional studies supported the modifier role of *CLN8* confirming it as a plausible candidate modifier gene in Gaucher disease (Zhang, Stein et al. 2012). Isolated or closed founder populations are genetically more homogenous than outbred populations, which minimizes population stratification. In addition, if the population is not only isolated but also closed (ie the founders themselves were related and there has been limited gene inflow into the community), the number of existing variants may be considerably reduced to start with. If the population is very prolific, such as is the case with the Amish, it is likely that rare variants become enriched and reside on preserved haplotype blocks. These otherwise rare variants are therefore likely to have larger effect sizes compared with common variants in outbred populations and be in linkage disequilibrium with tagged variants on the microarray chip. All these factors substantially increase the power of genome wide association studies in closed founder populations and facilitate the detection of genetic modifiers.

Given the increased power of association studies in closed founder populations, and the previous successful genome wide association studies with relatively small sample sizes in such populations, this part of the study aimed to investigate the potential association of common variants (or rare variants that may have become enriched as a result of the founder effect) with phenotype severity. Despite the main hindrance to the study of modest sample size of *MYBPC3* c.3330+2T>G variant carriers, a

potentially important finding was a genome-wide significant association between a region on chromosome 2q36.1, involving seven adjacent SNPs and the “deep T-wave inversion” phenotype. Deep T-wave inversion (≥ -0.2 mV) may be a marker of disease severity or reflect higher degree of myocardial remodelling in certain gene carriers, and conventionally is deemed a definitely abnormal finding on ECG if present in at least two contiguous leads excluding AVR, V1 or II. It is also commonly present in the subtype of apical HCM. When using Fisher’s exact test to investigate this putative association, which does not account for relatedness but is a more appropriate statistical test for small sample sizes, the association falls below the genome-wide significant threshold. Thus, while not considered statistically significant by both analytical methods, the putative association with this region should be considered further in case it may represent a genuine and biologically plausible association.

The nearest genes to the association signal are *EPHA4* and *PAX3* (figure 5-11). A recent study suggested that *EPHA4* may be a sex-specific hypertension risk gene in women (Zhang, Tremblay et al. 2019). *EPHA4* is a tyrosine kinase and ubiquitously expressed including in the vascular smooth muscle cells. Its’ stimulation reduces vascular smooth muscle cell contractility and this effect is enhanced by oestrogens, therefore explaining its’ potential association with hypertension risk in females. *EPHA4* is also expressed in the heart (<https://gtexportal.org/home/gene/EPHA4>) and it is conceivable that it may have an influence on myocardial contractility and disease biology in HCM. Future functional studies need to define its role in the myocardium to assess whether it may be a plausible modifier in HCM.

The *PAX3* gene encodes a transcription factor that is expressed during development of skeletal muscle, the central nervous system and neural crest derivatives. Germline mutations of the human *PAX3* genes cause deficiencies in these developmental lineages and result in Waardenburg syndrome (congenital hearing loss, changes in pigmentation). *PAX3* has also been implicated in the pathogenesis of rhabdomyosarcoma (Boudjadi, Chatterjee et al. 2018).

Although located at a greater distance from both genes than would be considered in a traditional GWAS, this signal may represent a genuine association in a founder haplotype setting such as in the Amish. In particular, the SNPs identified are located immediately upstream of *EPHA4*, close to the promoter region, and therefore may be located in a distant regulatory element for transcription from this gene. Additionally, both genes are located ca. 100Kb from the SNP cluster, and close enough to escape recombination in a founder community setting, in which they may be linked to an as yet unidentified sequence variant present in either gene with which they are in disequilibrium. Consequently, the association of this locus with the abnormal ECG marker in HCM might not have genome-wide significance, but may still represent a biologically important finding, as was the case in the above mentioned modifier study in Gaucher's disease in a different closed founder population (Zhang, Stein et al. 2012).

Future studies may confirm this by investigating larger Amish cohorts of individuals with the *MYBPC3* c.3330+2T>G variant, as well as in different HCM cohorts, to corroborate and validate the potential significant association of this locus with ECG abnormalities or disease severity. Additionally, exome sequencing of one or several of the Individuals with "deep T wave inversion" identified in the association may detect a variant in either *EPHA4* or *PAX3* that resides on the same haplotype and is in linkage disequilibrium with the associated SNPs. If so, this variant could then be directly tested for association with phenotype severity in future studies. Functional studies, such as studies in cardiomyocytes derived from human-induced pluripotent stem cells (hiPS-CM) harbouring the c.3330+2T>G *MYBPC3* variant, will need to define the potential role of variation of *EPHA4* and *PAX3* in the heart.

6 Concluding remarks

6.1 Clinical benefits to the Amish community

The Amish Heart Program is part of a long-running clinical-genetic project called “Windows of Hope” based in the Anabaptist communities of Ohio. The aims of the project are to determine the genetic basis of inherited disorders occurring in the Anabaptist communities and to translate these findings into direct medical benefit both within the community and for the general population. To date, the project has discovered the genes for over 20 novel conditions in the community. These findings have been instrumental for the diagnostic genetic testing and counselling of a large number of Amish individuals affected by these conditions. By using the insights into the molecular basis of these newly described syndromes, individuals affected by them often avoided lengthy and costly investigations which previously commonly did not culminate in a definitive diagnosis. Improved knowledge of the clinical features and molecular basis of these conditions amongst the local healthcare providers adds to the improved and quicker diagnosis of these conditions.

The comprehensive clinical, genetic and molecular studies in the individuals affected by the syndrome comprising congenital heart defects and orofacial clefting described in this thesis, equipped us with detailed knowledge about this syndrome, which is now being used to facilitate appropriate medical care and follow-up for affected individuals. Awareness of the full phenotype allows screening for all syndromic associations including congenital heart defects, which may require early intervention to prevent potentially serious long-term complications. All of the research results stemming from the Windows of Hope project are promptly fed back to the community during regular family support group meetings, educational events and by generating disease specific and community appropriate information material (please see appendix 7.2.6 and 7.2.7 for examples). We also regularly organise educational meetings for local health care professionals to communicate new findings, with the aim to improve health care outcomes through improved knowledge of the different

inherited conditions present in the communities. Our group has recently (partly based on the work performed for this thesis) been awarded an Alexander Fleming dissemination award from the MRC, Medical Research Foundation. The purpose of the award is to support the dissemination of MRC funded research results beyond the scientific peer reviewed press, especially to patients, participants of research studies and policy makers. This funding was awarded in recognition of the previous success of Windows of Hope engaging patients and healthcare workers to enable continuation and further development of the established community activities.

Next generation sequencing technologies have revolutionised diagnostic genetic testing allowing the screening of multiple genes in parallel in a much shorter time period. This technology has inspired a long-term vision of the Windows of Hope team to establish a newborn screening program consisting of a panel of genes involved in commonly occurring inherited conditions in the Amish communities. Genetic screening of newborns enables the early diagnosis and prognostication, appropriate counselling of parents and institution of prompt treatment strategies to ensure the best long-term health outcomes.

Undoubtedly the regionally most important pathogenic variant, which appears to be present at high frequency in all of the major Amish settlements, is the *MYBPC3* c.3330+2T>G variant. The high carrier frequency in the Anabaptist communities and the fact that, as is the case with other founder mutations, phenotypic expression is often delayed until after the reproductive years, have resulted in a number of couples (mostly unaware of their carrier status), having sometimes several homozygous children who passed away. This has understandably caused immense grief and sadness, and often guilt in these families.

In this study we offered any first degree relative of affected individuals free genetic testing in a CLIA approved clinical genetic laboratory. All gene carriers were offered to participate in the research study and underwent a clinical assessment. In addition, study participants received information on symptoms and signs of hypertrophic cardiomyopathy and its' genetic basis. Over 600 individuals took part in the study,

and over 300 gene carriers were identified. Most of the time individuals were not aware of their carrier status or that they were clinically affected by HCM. By reaching out to such a large number of individuals and their families, we significantly raised awareness for HCM and the implications of the high carrier frequency for the c.3330+2T>G variant in the community. Interest in the community was very high, as evidenced by the attendance of over 200 individuals at an educational event organised in conjunction with HCM experts from Cleveland Clinic (appendix 7.2.7). Appropriate cardiological follow-up was recommended for gene carriers and individuals with clinical HCM. In general Amish individuals rarely seek medical attention, partly due to the fact that they have no health insurance. Hopefully, better understanding and increased awareness of the disease process will improve the uptake of medical care and prevent complications from HCM in the long-term. Parents of homozygous infants who died of HCM, were very grateful for discovering the reasons and getting explanations for the tragedies that occurred in their families. A number of them reported feelings of guilt, which were alleviated to a degree by understanding the genetic nature of the disease. Even though family planning is not widely used within the conservative Amish communities, the knowledge about the recessive nature of infantile HCM and the awareness of carrier status, may influence some individuals' choices and help prevent future cases of lethal infantile HCM.

In addition, the description of the phenotypic spectrum in this large cohort of adult c.3330+2T>G variant carriers has provided valuable insight into the natural history and variable penetrance and expression of the disease. This information is extremely useful for local clinicians when counselling and following-up patients who are gene carriers or have HCM. Confirming that the c.3330+2T>G *MYBPC3* variant causes HCM with variable expression at any age and potential sudden death, reinforces the need for screening and continuing follow-up in these families. This study has shown that overall penetrance may be slightly lower compared to other HCM populations, but all types of HCM and its' complications may occur with the c.3330+2T>G *MYBPC3* variant, and therefore international guidelines for the management of HCM should

be followed as with other types of HCM. Longitudinal follow-up data would give additional valuable insight into disease progression and further inform prognostication and counselling of gene carriers.

The identification of putative genetic modifiers in HCM and additional potentially causative mutations for other types of cardiomyopathy may have wide reaching implications. Should future studies confirm the validity of the putative genetic modifiers identified in the Amish community, and prove their effect on phenotype expression, this information could be used to personalize follow-up, risk stratification and treatment of carriers of the c.3330+2T>G *MYBPC3* variant. This would not only improve individual care of patients with HCM, but also contribute to tackling the enormous financial burden that this disease poses to the communities by focusing follow-up and treatments, including extremely costly automated internal defibrillators, on the individuals who are most at risk. Furthermore, if pathogenicity is confirmed for distinct mutations causative of other types of cardiomyopathy, this information could be used in other families with cardiomyopathies not caused by c.3330+2T>G in *MYBPC3* and aid screening of at-risk individuals.

During recruitment for the HCM study we also encountered a number of other cardiac disorders and phenotypes amongst the Anabaptist communities which seem to have a genetic basis including a form of perinatal dilated cardiomyopathy, double outlet right ventricle, situs inversus, long QT syndrome, dilated cardiomyopathy associated with propionic acidemia (previously described), Brugada syndrome, and a form of sudden death suggestive of catecholaminergic polymorphic ventricular tachycardia (CPVT). Known causes have been excluded in some of these disorders and molecular studies are ongoing, which may give rise to establishing a genetic diagnosis in some of these families and new gene discoveries in the future.

6.2 Benefits to the wider society and scientific benefits

The benefits to the wider society and the contribution to scientific progress of community-based studies are numerous and wide ranging. Due to their genetic isolation and cultural confinement, there is an increased prevalence of autosomal recessive disorders amongst the Anabaptist communities, in particular amongst smaller groups which previously 'split off' from the more mainstream Amish and remained isolated. A number of rare Mendelian diseases have been described which contributed to the understanding of molecular pathways and disease pathology, and importantly have often subsequently been described outside the Anabaptist communities highlighting the global relevance of such studies. Many of the described syndromes are associated with congenital heart disease including Ellis-van-Creveld syndrome, primary ciliary dyskinesia, Alstroem syndrome, DCMA syndrome (dilated cardiomyopathy with ataxia), Renpenning syndrome, PMSE syndrome, Sitosterolemia, beta-glucuronidase deficiency mucopolysaccharidosis, McKusick-Kaufman syndrome and limb girdle muscular dystrophy type 21 (Wu, Tomatsu et al. 1994, Stone, Agarwala et al. 1998, Frosk, Greenberg et al. 2005, Davey, Parboosingh et al. 2006, Marshall, Hinman et al. 2007, Puffenberger, Strauss et al. 2007) (www.ncbi.nlm.nih.gov/pmc/articles/PMC3077314/).

Another example is the syndrome of congenital cardiac anomalies and orofacial clefting caused by mutation of *HYAL2* described in this thesis. This study highlights the vital role hyaluronidase enzymes play in both human and mouse development, and in particular has revealed a previously unrecognized pathway involved in the pathogenesis of cor triatriatum sinister, orofacial clefting and myopia, which may ultimately be amenable to treatment strategies. Since work started on this project, an additional six individuals from different ethnical and geographical backgrounds have been identified with similar phenotypes, who turned out to have pathogenic variants in *HYAL2*, highlighting the significance of these findings outside the Amish community. Since starting their work in 2000, the members of the Windows of Hope project alone have identified the genes for more than 20 inherited diseases present

amongst the Amish (Rogozin and Milanesi 1997, Patel, Cross et al. 2002, Proukakis, Hart et al. 2002, Simpson, Cross et al. 2009, Tekin, Chioza et al. 2013), demonstrating the scientific, clinical and financial advantages presented by community-based studies.

The presence of a founder mutation in the gene most commonly involved in HCM worldwide (*MYBPC3*) with likely hundreds and possibly thousands of gene carriers in the Anabaptist communities, provides unique opportunities to further our understanding of the disease processes involved in HCM. As previously discussed, in over 50% of cases of HCM mutations in sarcomeric proteins are identified. At present there is a lack of understanding of the mechanisms leading from mutations in sarcomeric proteins to the different pathologies (hypertrophy, fibrosis, myocyte disarray, arrhythmias). In addition, the pathogenic processes resulting in the extensive variation in expression even amongst carriers of the same mutation within one family are poorly understood. Furthermore, current therapeutic approaches focus on the alleviation of symptoms once disease is established, and the prevention of sudden cardiac death with implantable cardiac defibrillators, rather than altering the natural course of the disease process. In a special report commissioned by the National Heart, Lung, and Blood Institute, HCM has been identified as a prime example for the potential opportunities provided by harnessing modern genetic science in medicine to make gene-based diagnosis and prediction reality (Force, Bonow et al. 2010) (Force T et al, Circulation 2010). The report highlighted our lack of understanding what mechanisms are triggered by mutations in sarcomeric proteins that ultimately lead to the different phenotypes of HCM.

The phenotypic description and the identification of putative genetic modifiers in this unique Amish HCM population may have significant impact on our knowledge about disease mechanisms and future research in a number of areas. Firstly, the findings of this study contribute to our understanding and knowledge about the variable phenotypic spectrum of disease expression and its' mechanisms. The identification of novel genes implicated in the modification of the disease process

may reveal new pathways that have not previously been implicated in the molecular mechanisms that lead to the different phenotypes of HCM, which may lead to a better understanding of its pathophysiology. More specifically, pathways of disease expression that are responsible for the development of hypertrophy, adverse myocardial remodelling, atrial or ventricular arrhythmogenesis or even sudden cardiac death can be explored in this extensive cohort which may greatly extend current knowledge. This knowledge may also provide a more comprehensive understanding of molecular triggers and mechanisms for cardiac hypertrophy and remodelling in a wide range of other cardiovascular diseases (eg hypertension, heart failure).

Furthermore, the investigation of these putative modifiers in larger (heterogeneous) HCM cohorts, would allow the exploration of global relevance of our findings in the general (ie non- Amish) population. If confirmed, this information could be used to personalise management of HCM patients by improving prognostication and risk stratification, similarly as previously suggested for within the Amish community. Additionally, future studies using cardiomyocytes derived from human-induced pluripotent stem cells (hiPS-CM) or a mouse model carrying the c.3330+2T>G *MYBPC3* variant, may further evaluate the effect and contribution of the putative genetic modifiers identified in this study to the disease process. The identification of as yet unknown pathways involved in the disease process of HCM and their further investigation might eventually elucidate targets and therapeutic strategies to alter the natural history of the disease. Ultimately, the identification of suitable strategies for disease prevention or treatment may be explored in clinical trials to determine whether these novel therapies alter the natural history of patients with HCM in general. Appropriate treatment might be initiated at an earlier stage before the occurrence of complications and more accurate estimation of prognosis and disease progression might be possible.

7 Supplemental data

7.1 Additional information for *HYAL2* project

Stage	Total	<i>Hyal2</i> ^{-/-}	<i>Hyal2</i> ^{+/-}	<i>Hyal2</i> ^{+/+}
Dead at E18.5	6	5 ¹	1	
Live at E18.5	84	16	42	26
Total E18.5	90	21 ¹	43	26
Dead at P1	16	16 ²		
Dead at P2-P9	16	16 ²		
Live at P21	169	16	93	60
Total pups born	201	48 ²	93	60
Sum of embryos and pups				
Observed #	291	69	136	86
(%)	(100)	(24)	(46)	(30)
Expected #	291	72.75	145.5	72.75
(%)	(100)	(25)	(50)	(25)
Chi-square test (O-E) ² /E ³		0.19	0.62	2.41

Table 7-1: Survival of progeny of *Hyal2*^{+/-} intercrosses

¹ Includes 3 resorptions presumed to be *Hyal2* KO because no tissue remained

²Missing mice were presumed to be *Hyal2* KO; only 7 of 32 were confirmed

³3.23 p>0.5

Data courtesy of Professor B Triggs-Raine, University of Manitoba, Winnipeg, Canada

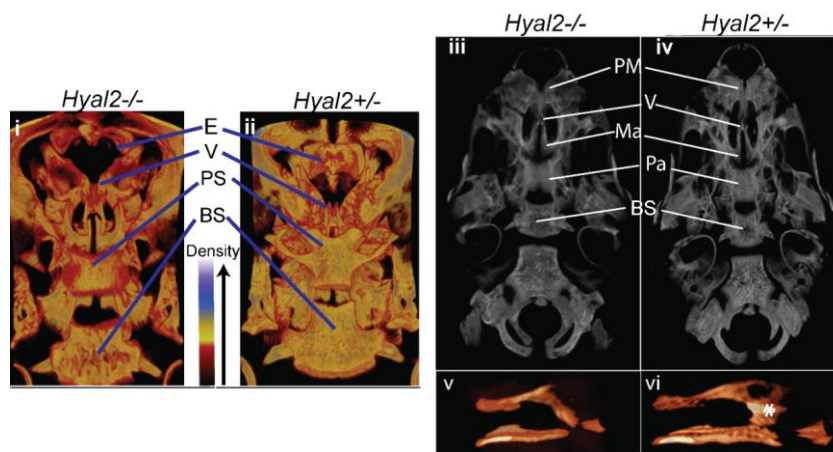


Figure 7-1 Micro-CT images of the skulls of *Hyal2*^{-/-} and *Hyal2*^{+/-} mice

3D Images of *Hyal2*^{-/-} mice that died at P1, and the control (*Hyal2*^{+/-}) which was sacrificed at P2, were created using CTvox software. All images were scanned and reconstructed using identical parameters and the dynamic range for the grey scale was held constant (i and ii). Dorsal views from the rostral end of the skull were prepared using a transfer function that colorized the grey scale such that areas of the bone with lower density are red (see density scale in image). (iii and iv). Ventral grey scale views of the skulls without the lower jaw. (v and vi). Space filled images of the vomer. * indicates the vomer head. Similar findings were evident in three additional *Hyal2*^{-/-} mice found dead at P1. These images are not to scale.

E, ethmoid; V, vomer; PS, presphenoid; BS, basisphenoid; PM, premaxilla; MA, maxilla; PA, palantine.

Data courtesy of Professor B Triggs-Raine, University of Manitoba, Winnipeg, Canada

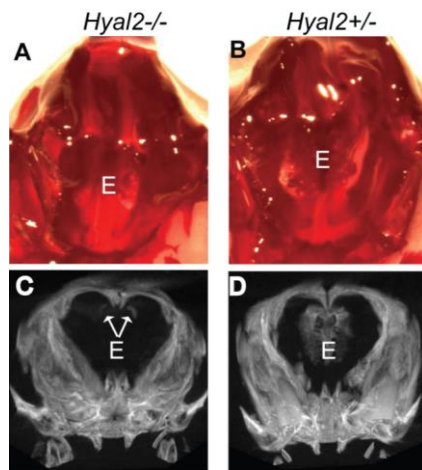


Figure 7-2 Skeletal staining and micro-CT imaging of skulls of *Hyal2*^{-/-} and *Hyal2*^{+/-} mice

Visualization of the skulls of *Hyal2*^{-/-} and *Hyal2*^{+/-} mice. The skulls of *Hyal2*^{-/-} mice that died at P1, and of a *Hyal2*^{+/-} mouse that was sacrificed at P2, were visualized by skeletal staining (A,B) or micro-CT imaging (C,D). (A and B) Photographs of the anterior region of the viscerocranium after removal of the skull. The ethmoid (E) bone is severely underdeveloped in the *Hyal2*^{-/-} mouse compared to the *Hyal2*^{+/-} mouse. For micro-CT imaging, the scanning and reconstruction parameters, including the dynamic range of the grey scale, were the same for all images. (C and D) 3D micro-CT grey scale images of the ethmoid bone that is almost absent in the *Hyal2*^{-/-} mouse, but well developed in the *Hyal2*^{+/-} control. E-ethmoid bone.

Data courtesy of Professor B Triggs-Raine, University of Manitoba, Winnipeg, Canada

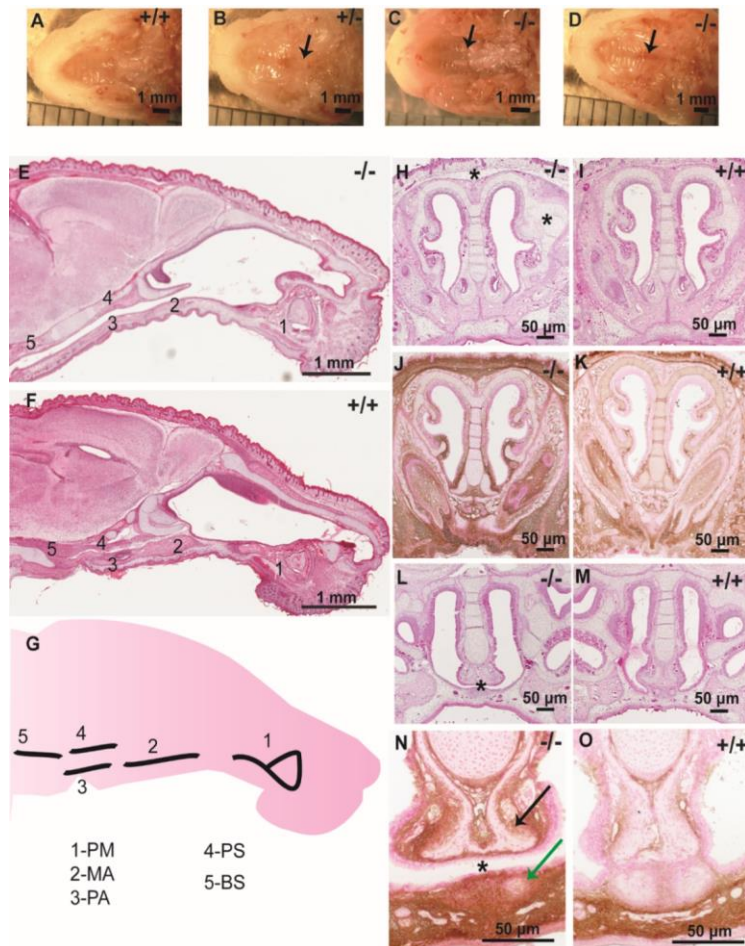


Figure 7-3 Analysis of the palates of *Hyal2*^{-/-}, *Hyal2*^{+/-}, and *Hyal2*^{+/+} mice

(A-D) Photographs were taken of the ventral surface of the head of E18.5 embryos after removing the bottom mandible. Arrows indicate partial cleft palates indicative of submucosal cleft palate. (E - G) Histological analysis of *Hyal2*^{-/-} and *Hyal2*^{+/+} P1 heads. Sagittal sections were stained with H & E. The darker pink staining with white spaces is bone; the schematic in (G) indicates the identity of the bones: PM, premaxilla; MA, maxilla; PA, palatine; PS, prephenoid; BS, basisphenoid (H-I) Coronal sections through the nasopharynx and secondary palate of E18.5 embryos were stained for morphological analysis with hematoxylin and eosin. Areas with excess mesenchymal cells in the *Hyal2*^{-/-} mouse are indicated with asterisks. (J and K) Hyaluronan (brown) detected in coronal sections of the nasopharynx. (L and M) H & E staining of the fusion between the vomeronasal organ and palate. An * indicates the failed fusion in the *Hyal2*^{-/-} mouse. (N and O) Hyaluronan (brown)

detection in the region of the vomeronasal/palate fusion. An * indicates the failed fusion between the vomeronasal organ and primary palate. A black arrow indicates the underdeveloped vomer and a green arrow indicates the underdeveloped maxilla. All images are representative of a minimum of 3 sets of Hyal2^{-/-} and control mice.

Data courtesy of Professor B Triggs-Raine, University of Manitoba, Winnipeg, Canada

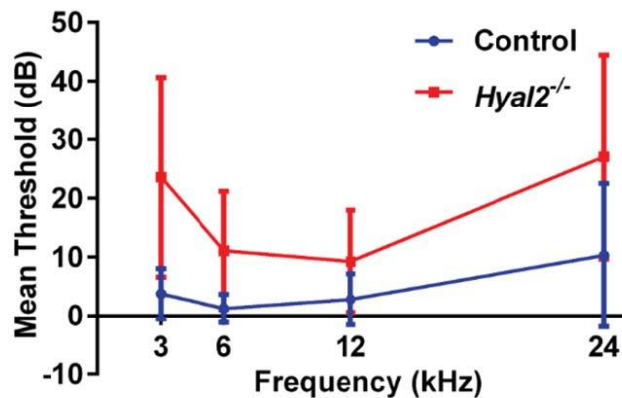


Figure 7-4 Evaluation of hearing in HYAL2^{-/-} and control (HYAL2^{+/+}) mice

Evaluation of hearing in HYAL2^{-/-} and control (HYAL2^{+/+}) mice was assessed by the ABR (auditory brain stem response) at varying frequencies and intensities. These values are based on the average of 16 ears (8 mice) for each genotype. The bars represent the mean threshold and the 95% confidence interval.

Data courtesy of Professor B Triggs-Raine, University of Manitoba, Winnipeg, Canada

7.2 Additional information for hypertrophic cardiomyopathy project

7.2.1 The ancestors of the carriers of the c.3330+2T>G variant

Michael Troyer (born in 1730 in Europe, died 1807 in Somerset County, Pennsylvania) and Magdalena Mast (born in 1736 Europe, died in 1807 in Somerset County, Pennsylvania) are the couple that all carriers of the *MYBPC3* variant can be traced back to. They both emigrated from Europe to Pennsylvania, USA in the 18th century. The variant may have been a de-novo variant in of them.

Michael and Magdalena had fourteen children. Three of their children were not married and had no descendants. Eight out of the eleven children who got married and had children, have living descendants that are carriers of the *MYBPC3* variant.

1. John Troyer (DOB 3 Feb 1753), not married, no descendants
2. Michael Troyer (DOB 4 Sep 1754), first marriage with Anna Rickenbach (deceased, five children), second marriage with Magdalena Rickenbach (four children), descendants live in Holmes County, Ohio (2 lines)
3. Christian Troyer (DOB 18 May 1756), First marriage with Barbara Yoder (deceased, four children), second marriage with Elizabeth Becker (two children), emigrated to York County, Ontario; son David C is the common ancestor for this branch and he immigrated to Holmes County (3 lines through David C only)
4. Barbara Troyer (DOB 3 Jan 1758), married to Peter Sipe, descendants in Somerset County, Pennsylvania; LaGrange County, Indiana and Holmes County, Ohio (3 lines)
5. Magdalena Troyer (DOB 20 Oct 1760), married to Christian Yoder, 14 children, no HCM pedigree lines trace through Magdalena
6. Anna Troyer (DOB 4 Sep 1962), not married, no descendants

7. Jacob Troyer (DOB 19 Sep 1764), first marriage with Catherine Mast (deceased, two children), second marriage with Mary Winger (eight children), no HCM pedigree lines trace through Jacob
8. Henry Troyer (DOB 13 Sep 1766), married to Veronica Kauffman (nine children), no HCM pedigree lines trace through Henry
9. Mary Troyer (DOB 6 Nov 1768), not married, no descendants
10. Joseph Troyer (DOB 3 Mar 1771), married to Mary Kauffman (10 children), descendants in Somerset County, Pennsylvania; Iowa; Illinois and Wayne/Holmes County, Ohio (2 lines)
11. Andrew Troyer (DOB 30 Nov 1773) married to Margaret Rickenbach (five children), descendants in Holmes County, Ohio (2 lines)
12. David Troyer (DOB 6 Jul 1777), married to Barbara Schrock (10 children), descendants in Holmes County, Ohio; most of Holmes County HCM families trace through David and Barbara (15 lines; Holmes families trace through 5 of the 10 children)
13. Veronica Troyer (DOB 7 May 1779), married to Jacob Miller (nine children), descendants in Elkhart County, Indiana; Middlebury, Ohio and Holmes County, Ohio (3 lines)
14. Elizabeth Troyer (DOB 21 Jan 1781), married to David Miller (14 children), descendants in Lagrange County, Indiana; Iowa; and Holmes County, Ohio (4 lines)

7.2.2 Pedigrees of *MYBPC3* c. 3330+2T>G variant carriers

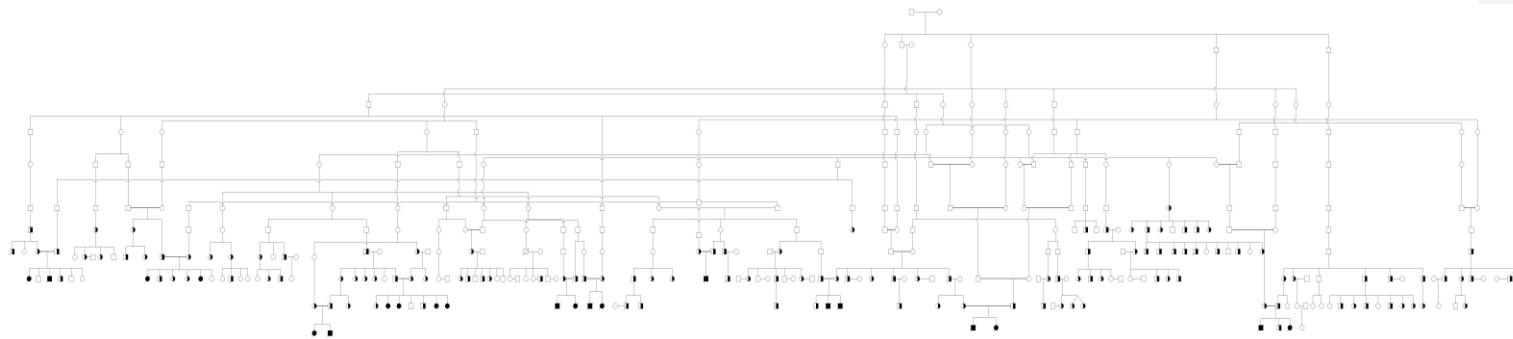


Figure 7-5: Main pedigree Geauga county, Ohio

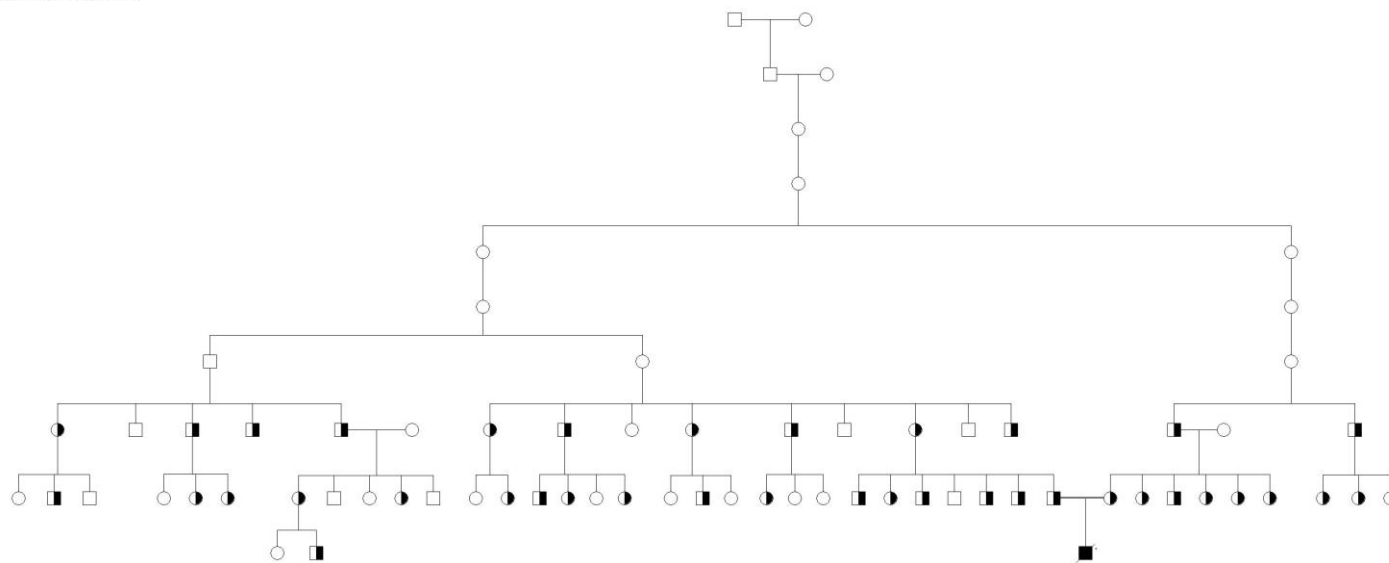


Figure 7-6: Pedigree 1 Holmes county, Ohio

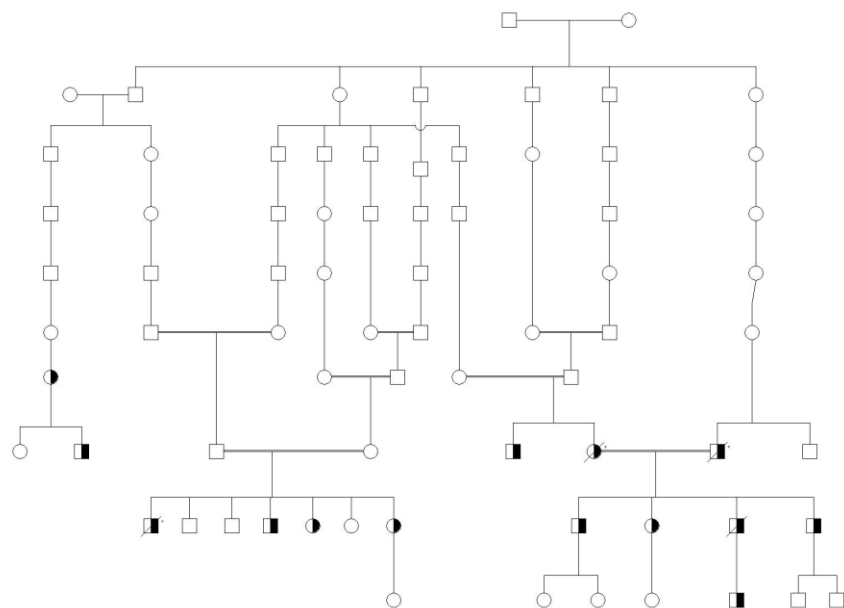


Figure 7-7: Pedigree 2 Holmes County, Ohio

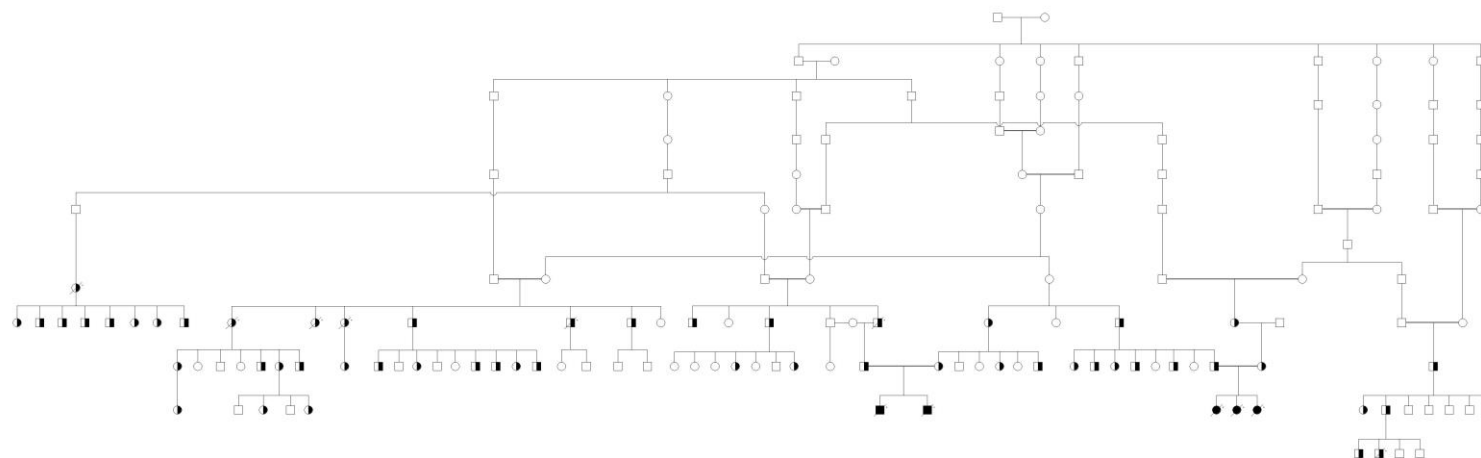


Figure 7-8: Pedigree 3 Holmes County, Ohio

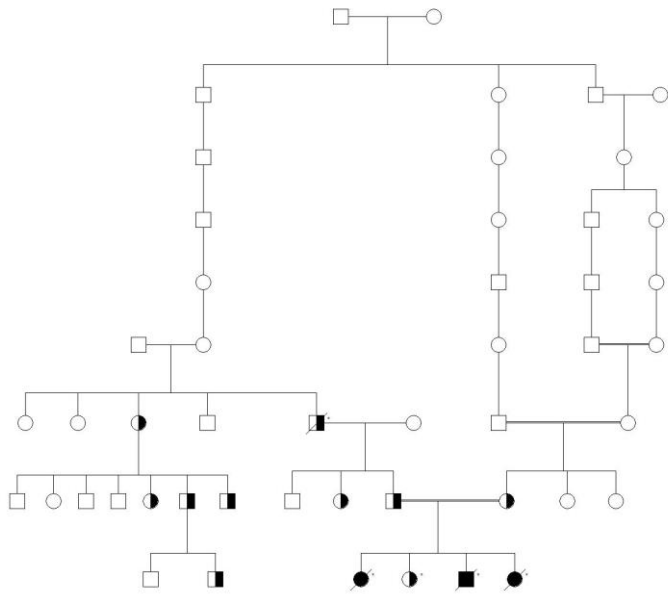


Figure 7-9: Pedigree 4 Holmes County, Ohio

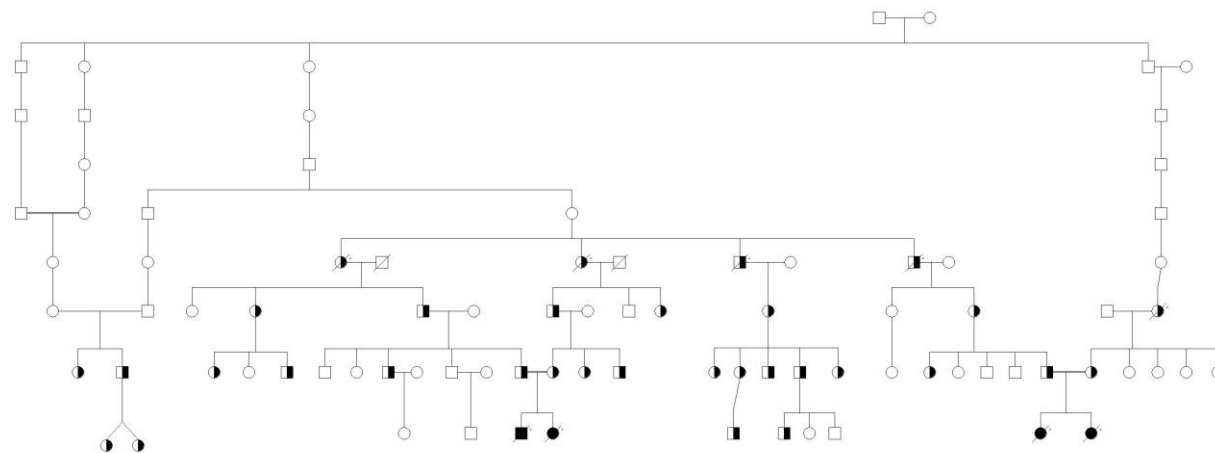


Figure 7-10: Pedigree 5 Holmes County, Ohio

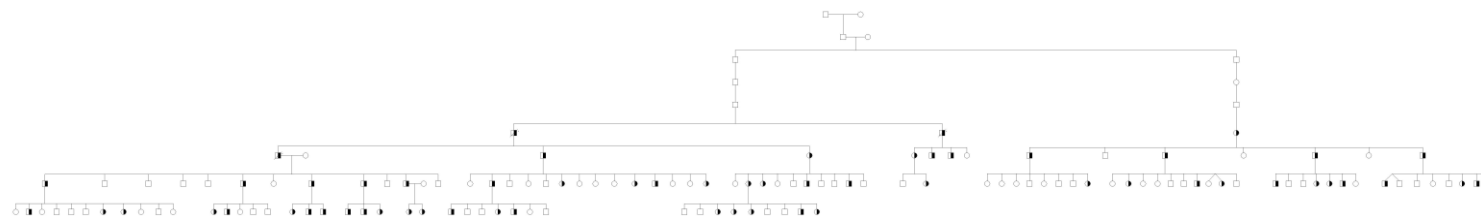


Figure 7-11: Pedigree 6 Holmes County, Ohio

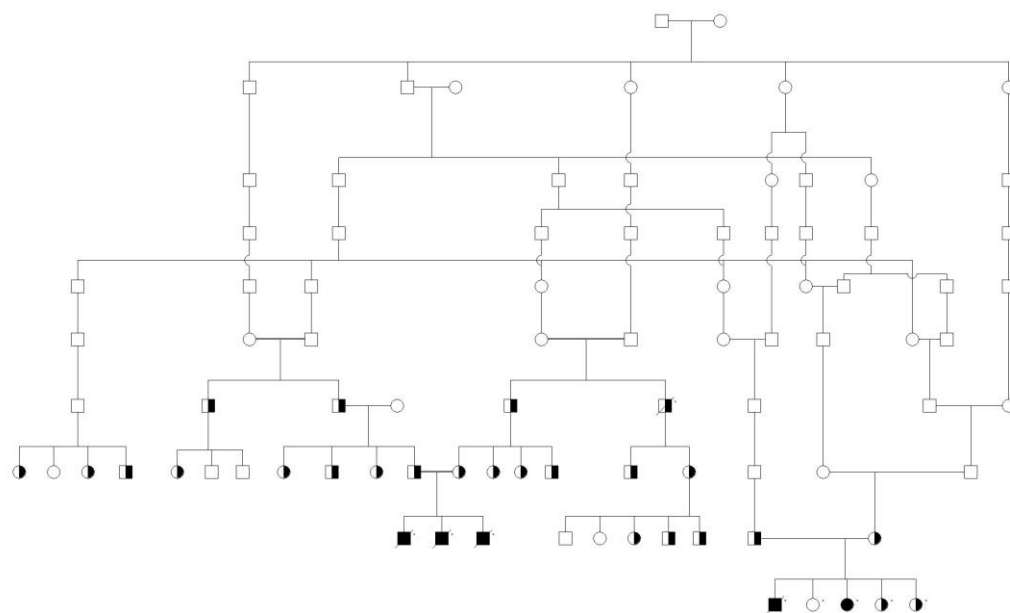


Figure 7-12: Pedigree 7 Holmes County, Ohio

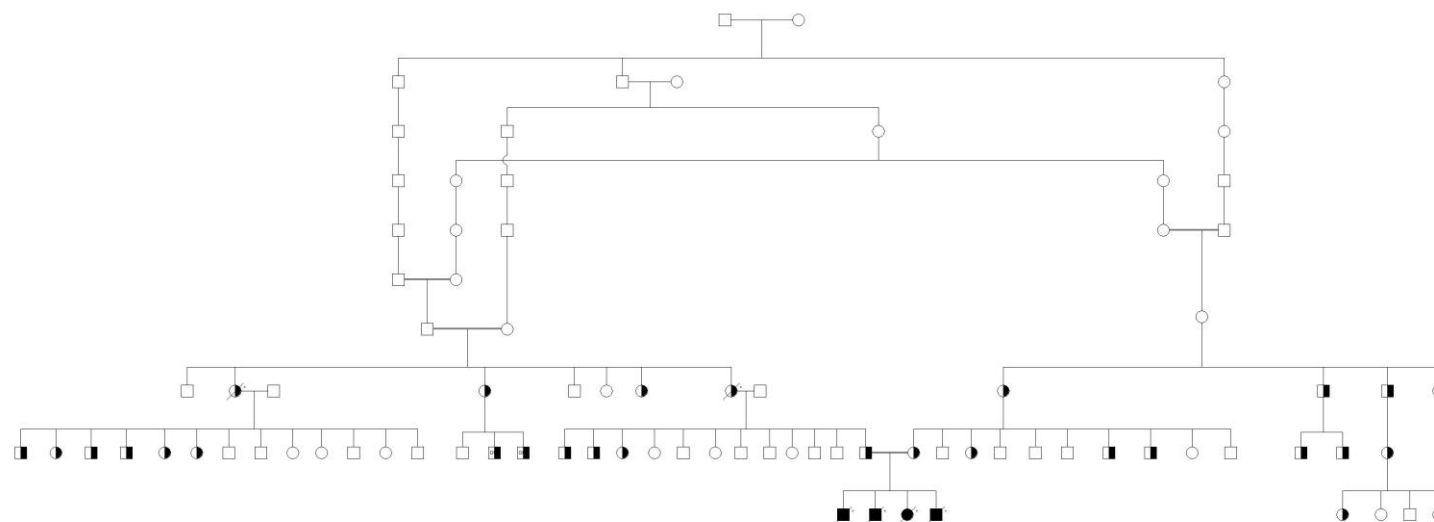


Figure 7-13: Pedigree 8 Holmes County, Ohio

7.2.3 Investigation of the phenotype of carriers of the *MYBPC3* c.3330+2T>G variant

	myectomy (n=6)	no myectomy
male	100%	49.8% (153/307)
age [y]	55.2 +/- 8.0	42.4 +/-16.4
number of healthy children	3.75 +/- 5.6 (0-12)	4.27 +/-3.4 (0-15)
weight [kg]	89.7 +/- 7.3	80.3 +/- 16.9
height [cm]	175.6 +/- 5.9	168.4 +/- 9.2
BMI [kg/m ²]	29.2 +/- 3.7	28.0 +/- 5.8
systolic BP [mmHg]	124.0 +/- 13.1	121.9 +/- 16.6
diastolic BP [mmHg]	77.8 +/- 6.6	77.0 +/-10.3
hypertension	0%	12.0% (34/283)
diabetes mellitus	0%	3.8% (11/288)
hypercholesterolemia	16.7 % (1/6)	4.5% (13/288)
coronary disease	0%	1.7% (5/288)
known diagnosis of HCM	100%	14.6% (42/288)
age at diagnosis of HCM [y]	47.4 +/- 8.3	41.8 +/- 13.0
first degree relative SCD	0%	9.4% (27/288)
atrial fibrillation	66.7% (4/6)	4.2% (12/288)
age at onset AF [y]	50.5 +/- 6.1	60.4 +/- 17.0
ICD/pacemaker	16.7% (1/6)	1.4% (4/288)
age at ICD/pacemaker [y]	55	49.8 +/- 9.8
age at myectomy [y]	53 +/- 3.8	n/a
on beta-blocker	83.3% (5/6)	11.8% (34/288)
on NDH channel blocker	0%	1.7% (5/288)

Table 7-2: Demographic and clinical characteristics of carriers of *MYBPC3* c.3330+2T>G variant with and without surgical myectomy

	myectomy (n=5)	no myectomy
sinus rhythm	80.0% (4/5)	95.8% (229/239)
normal QRS axis	40.0% (2/5)	84.5% (202/239)
right QRS axis	20.0% (1/5)	6.3% (15/239)
left QRS axis	20.0% (1/5)	7.1% (17/239)
extreme axis	0%	2.1% (5/239)
RA enlargement	0%	2.2% (5/230)
LA enlargement	20.0% (1/5)	5.6% (13/232)
RVH Sokolow-Lyon	0%	19.8 % (47/237)
LVH Sokolow Lyon	80.0% (4/5)	9.3% (22/237)
LBBB	0%	1.3% (3/236)
RBBB	0%	0.4% (1/236)
QRS fragmentation	20.0% (1/5)	67.4% (155/230)
Q waves	0%	7.8% (18/232)
T wave inversion (TWI)	100% (5/5)	7.6% (18/237)
TWI V1 - V4	20.0% (1/5)	5.1% (12/237)
TWI V4 - V6	80.0% (4/5)	4.6% (11/237)
TWI I + aVL	40.0% (2/5)	2.5% (6/237)
TWI II+III+aVF	20.0% (1/5)	0.8% (2/237)
Deep TWI	100% (5/5)	3.4% (8/237)
deep TWI V1 – V4	20.0% (1/5)	2.1% (5/237)
deep TWI V4 – V6	80.0% (4/5)	2.1% (5/237)
deep TWI I + aVL	40.0% (2/5)	0.4% (1/237)
Deep TWI II + III + aVF	20.0% (1/5)	0%
ST elevation	0%	0.5% (1/221)
RE score ≥ 4	100% (5/5)	13.0% (31/239)
RE score ≥ 5	100% (5/5)	7.1% (17/239)
Abnormal ECG	100% (5/5)	24.0% (56/233)

Table 7-3: ECG characteristics (part 2) of individuals with and without myectomy

	myectomy (n=5)	no myectomy
IVSd (PLAX) [mm]	13.8 +/- 2.2	11.8 +/- 3.6 (n=266)
PWd (PLAX) [mm]	12.0 +/- 2.7	9.4 +/- 1.8 (n=265)
LVDd (PLAX) [mm]	46.4 +/- 3.9	42.4 +/- 6.1 (n=265)
LVSd (PLAX) [mm]	32.4 +/- 5.1	32.0 +/- 6.0 (n=265)
LVEDv BP [ml]	107.0 +/- 30.7 (n=4)	78.1 +/- 21.7 (n=233)
LVESv BP [ml]	45.0 +/- 17.5 (n=4)	30.5 +/- 10.9 (n=233)
EF BP [%]	59.0 +/- 6.2 (n=4)	61.2 +/- 7.4 (n=234)
relative WT (PLAX)	0.57 +/- 0.15	0.51 +/- 0.16 (n=265)
max mitral level WT (PSAX) [mm]	19.2 +/- 3.6	13.1 +/- 4.0 (n=257)
max papillary level WT (PSAX) [mm]	17.2 +/- 2.3	12.5 +/- 3.6 (n=256)
max apical WT [mm]	14 (n=1)	11.52 +/- 3.3 (n=133)
max WT [mm]	19.2 +/- 3.6	14.1 +/- 4.1 (n=266)
max WT ≥ 13 mm	100%	50.0% (n=266)
LAv (a-l) 4Ch [ml]	125.5 +/- 34.1 (n=4)	60.1 +/- 24.1 (n=204)
LAv (a-l) 2Ch [ml]	93.0 +/- 18.4 (n=2)	62.5 +/- 26.7 (n=178)
LAv (a-l) mean 4Ch/2Ch [ml]	105.3 +/- 17.3 (n=2)	60.8 +/- 23.7 (n=178)
LVOT gradient [mmHg]	10.5 +/- 6.1 (n=4)	6.6 +/- 9.3 (n=199)
LVOT gradient ≥ 30 mmHg [%]	0% (n=4)	1.5% (n=200)
MAPSE lateral [mm]	13.0 +/- 4.7 (n=4)	14.2 +/- 2.8 (n=205)
MAPSE septal [mm]	9.5 +/- 4.1 (n=4)	12.5 +/- 2.6 (n=205)
TAPSE [mm]	15.3 +/- 4.9 (n=4)	24.9 +/- 4.3 (n=201)

Table 7-4: Echocardiographic parameters (part 1) of individuals with and without myectomy (structure and function)

	myectomy	no myectomy
MV inflow E [cm/s]	71.0 +/- 16.5 (n=3)	69.6 +/- 16.5 (n=148)
MV inflow A [cm/s]	133.7 +/- 108.5 (n=3)	56.1 +/- 16.9 (n=146)
MV inflow E/A	0.75 +/- 0.47 (n=3)	1.36 +/- 0.55 (n=145)
MV inflow DT (ms)	194.0 +/- 64.5 (n=3)	172.9 +/- 42.8 (n=146)
E' lateral [cm/s]	7.5 +/- 2.1 (n=2)	11.7 +/- 4.3 (n=202)
A' lateral [cm/s]	7.0 +/- 7.0 (n=2)	8.9 +/- 2.9 (n=201)
S' lateral [cm/s]	6.5 +/- 3.5 (n=2)	8.2 +/- 2.0 (n=202)
E' lat/ A' lat	2.5 +/- 2.8 (n=2)	1.54 +/- 0.94 (n=201)
E' septal [cm/s]	5.5 +/- 0.7 (n=2)	9.3 +/- 3.4 (n=199)
A' septal [cm/s]	6.5 +/- 6.4 (n=2)	8.0 +/- 2.3 (n=198)
S' septal [cm/s]	5.5 +/- 2.1 (n=2)	7.5 +/- 1.4 (n=199)
E' sep/ A' sep	1.52 +/- 1.38 (n=2)	1.31 +/- 0.73 (n=198)
E/ E' lateral	8.00 (n=1)	6.80 +/- 3.44 (n=145)
E/ E' septal	14.4 (n=1)	8.59 +/- 4.62 (n=143)
Average E/ E'	11.2 (n=1)	7.64 +/- 3.80 (n=143)
IVRT [ms]	120.0 +/- 36.4 (n=4)	85.0 +/- 21.4 (n=191)
PV _s [cm/s]	45.3 +/- 4.9 (n=3)	49.1 +/- 10.9 (n=163)
PV _D [cm/s]	38.7 +/- 8.5 (n=3)	45.2 +/- 10.8 (n=163)
PV _s / PV _D	1.20 +/- 0.18 (n=3)	1.15 +/- 0.36 (n=163)
PV _a [cm/s]	31.7 +/- 10.7 (n=3)	28.7 +/- 6.8 (n=113)
PV _{adur} [ms]	143.3 +/- 20.6 (n=3)	126.6 +/- 27.2 (n=111)

Table 7-5: Echocardiographic parameters (part 2) of individuals with and without myectomy (diastolic function)

	myectomy	no myectomy
No hypertrophy (max WT<13mm)	0% (0/4)	50.0 % (125/250)
Type of hypertrophy		
sigmoid septum	0% (0/4)	25.2% (63/250)
reverse septal curvature	75.0% (3/4)	20.0% (50/250)
apical hypertrophy	0% (0/4)	2.4% (6/250)
concentric hypertrophy	25.0% (1/4)	1.6% (4/250)
segmental hypertrophy	0% 0/4)	2.0% (5/250)
dilated phase	0% (0/4)	0.4% (1/250)
SAM	75.0% (3/4)	12.2% (24/196)
cordal/PM SAM	0% (0/4)	12.8% (25/196)
PM morphology		
PM normal	0%	41.8% (77/184)
PM hypertrophy	100% (4/4)	54.3% (100/184)
PM abnormal attachment	25.0% (1/4)	20.1% (37/184)
PM apical displacement	25.0% (1/4)	8.7% (16/184)
PM bifid	0%	1.6% (3/184)
length anterior MV leaflet (PLAX) [mm]	24.0 +/- 2.9 (n=4)	21.8 +/- 2.6 (n=192)
length anterior MV leaflet (4Ch) [mm]	25.0 +/- 4.8 (n=4)	21.2 +/- 2.4 (n=194)
length anterior MV leaflet (3Ch) [mm]	25.0 +/- 3.8 (n=4)	21.4 +/- 2.5 (n=188)

Table 7-6: Echocardiographic parameters (part 3) of individuals with and without myectomy

	male	female	p value
number of carriers	49.8% (153/307)	50.2 % (154/307)	n.s.
age [y]	41.3 +/- 16.9	43.5 +/- 15.8	n.s.
number of healthy children	4.0 +/- 3.5	4.5 +/- 3.3	n.s.
weight [kg]	85.2 +/- 13.3	75.7 +/- 18.5	<0.001
height [cm]	176.6 +/- 6.7	162.7 +/- 5.0	<0.001
BMI [kg/m ²]	27.4 +/- 4.3	28.6 +/- 7.0	n.s.
systolic BP [mmHg]	123.0 +/- 13.7	120.9 +/- 18.9	n.s.
diastolic BP [mmHg]	78.1 +/- 9.7	76.0 +/- 10.7	n.s.
hypertension	7.9% (11/139)	16.0% (23/144)	0.037
diabetes mellitus	2.1% (3/144)	5.6% (8/144)	n.s.
hypercholesterolemia	4.2% (6/144)	4.9% (7/144)	n.s.
coronary disease	3.5% (5/144)	0.0%	0.024
known diagnosis of HCM	18.8% (27/144)	10.4% (15/144)	0.045
age at diagnosis of HCM [y]	41.0 +/- 13.8	43.3 +/- 11.9	n.s.
previously seen Cardiologist	22.9% (33/144)	18.1% (26/144)	n.s.
first degree relative SCD	10.4% (15/144)	8.3% (12/144)	n.s.
atrial fibrillation	4.2% (6/144)	4.2% (6/144)	n.s.
age at onset AF [y]	61.8 +/- 21.5	59.0 +/- 13.1	n.s.
ICD/pacemaker	2.1% (3/144)	0.7% (1/144)	n.s.
age at ICD/pacemaker [y]	47.0 +/- 10.0	58	n.s.
on beta-blocker	12.5% (18/144)	11.1% (16/144)	n.s.
on NDH Ca ²⁺ -channel blocker	1.4% (2/144)	2.1% (3/144)	n.s.

Table 7-7: Comparison of demographic and clinical characteristics of male and female carriers of *MYBPC3* c.3330+2T>G variant

	male	female	p value
exertional chest pain	11.9% (17/143)	17.4% (25/144)	n.s.
exertional shortness of breath	43.4% (62/143)	47.9% (69/144)	n.s.
palpitations	26.8% (38/142)	52.8% (76/144)	<0.001
palpitations consistent with ectopics	10.6% (15/142)	17.4% (25/144)	n.s.
palpitations undetermined	14.1% (20/142)	30.6% (44/144)	0.001
palpitations other	2.8% (4/142)	4.2% (6/144)	n.s.
dizziness	35.5% (39/110)	36.8% (43/117)	n.s.
dizziness after standing	14.5% (16/110)	17.9% (21/117)	n.s.
dizziness after bending over	13.6% (15/110)	10.3% (12/117)	n.s.
exertional dizziness	13.6% (15/110)	9.4% (11/117)	n.s.
other dizziness	0.9% (1/110)	5.1% (6/117)	n.s.
syncope	21.8% (31/142)	27.1% (39/144)	n.s.
neurocardiogenic syncope	14.8% (21/142)	23.6% (34/144)	0.058
exertional syncope	4.3% (6/141)	0.0%	0.012
sudden syncope at rest	0.7% (1/141)	2.8% (4/144)	n.s.
indetermined syncope	2.1% (3/141)	2.1% (3/144)	n.s.
fatigue	6.0% (7/116)	7.2% (9/125)	n.s.
aborted SCD	0%	0%	
any significant symptom	29.5% (33/112)	28.2% (35/124)	n.s.

Table 7-8: Comparison of symptoms of male and female carriers of *MYBPC3* c.3330+2T>G variant

	male	female	p value
heart rate [bpm]	65.0 +/- 10.5 (n=113)	68.1 +/- 10.9 (n=125)	0.026
sinus rhythm	95.6% (109/114)	96.0% (120/125)	n.s.
normal QRS axis	79.8% (91/114)	88.8% (111/125)	n.s.
right QRS axis	7.9% (9/114)	4.8% (6/125)	n.s.
left QRS axis	8.8% (10/114)	5.6% (7/125)	n.s.
extreme axis	3.5% (4/114)	0.8% (1/125)	n.s.
P wave duration [ms]	110.9 +/- 12.8 (n=110)	109.9 +/- 14.0	n.s.
RA enlargement	0.9% (1/110)	3.3% (4/120)	n.s.
LA enlargement	7.3% (8/110)	4.1% (5/122)	n.s.
PR duration [ms]	173.8 +/- 28.4 (n=110)	168.4 +/- 22.2 (n=122)	n.s.
RVH Sokolow-Lyon	33.9% (38/112)	7.2% (9/125)	<0.001
LVH Sokolow Lyon	17.0% (19/112)	2.4% (3/125)	<0.001
Intrinsicoid deflection [ms]	38.2 +/- 8.5 (n=111)	36.8 +/- 6.9 (n=123)	n.s.
QRS duration	85.8 +/- 13.3 (n=112)	80.3 +/- 13.4 (n=125)	0.002
LBbB	0.9% (1/111)	1.6% (2/125)	n.s.
RBbB	0.9% (1/111)	0.0%	n.s.
QRS fragmentation	70.6% (77/109)	64.5% (78/121)	n.s.
Q waves	11.9% (13/109)	4.1% (5/123)	0.025
QT duration [ms]	401.0 +/- 30.8 (n=111)	407.3 +/- 34.1 (n=124)	n.s.
QTc duration [ms]	413.9 +/- 22.5 (n=111)	430.4 +/- 27.0 (n=124)	<0.001
T wave inversion (TWI)	6.3% (7/112)	8.8% (11/125)	n.s.
Deep TWI	3.6% (4/112)	3.2% (4/125)	n.s.
ST elevation	0.9% (1/106)	0.0%	n.s.
RE score ≥ 4	15.8% (18/114)	10.4% (13/125)	n.s.
RE score ≥ 5	7.0% (8/114)	7.2% (9/125)	n.s.
Abnormal ECG	29.1% (32/110)	19.5% (24/123)	n.s.

Table 7-9: Comparison of ECG characteristics of male and female carriers of MYBPC3 c.3330+2 T>G variant

	Male (n=32)	Female (n=33)	p value
number of PVCs	450.9 +/- 1366.9	681.8 +/- 1669.6	n.s.
Number of episodes of bigeminy	6.6 +/- 28.8	29.4 +/- 116.1	n.s.
Number of couplets	1.2 +/- 3.1	20.3 +/- 101.8	n.s.
Number of triplets	0.1 +/- 0.4	0.8 +/- 3.3	n.s.
Individuals with NSVT $\geq$ 4 beats [%]	6.3% (2/32)	21.2% (7/33)	-
Min heart rate [bpm]	54.3 +/- 6.5	55.3 +/- 6.9	n.s.
Max heart rate [bpm]	133.3 +/- 16.8	125.4 +/- 13.1	0.039
Mean heart rate [bpm]	78.0 +/- 8.0	76.9 +/- 8.9	n.s.

Table 7-10: Comparison of holter characteristics between male and female carriers of *MYBPC3* c.3330+2 T>G variant

	male	female	p value
No hypertrophy (max WT < 13mm)	38.3% (46/120)	60.8% (79/130)	<0.001
Type of hypertrophy			
sigmoid septum	27.5% (33/120)	23.1% (30/130)	n.s.
reverse septal curvature	26.7% (32/120)	13.8% (18/130)	0.011
apical hypertrophy	5.0% (6/120)	0.0% (0/130)	0.010
concentric hypertrophy	2.5% (3/120)	0.8% (1/130)	n.s.
segmental hypertrophy	2.5% (3/120)	1.5% (2/130)	n.s.
dilated phase	0.8% (1/120)	0.0% (0/130)	n.s.
SAM	16.0% (15/94)	8.8% (9/102)	n.s.
cordal/PM SAM	11.7% (11/94)	13.7% (14/102)	n.s.
PM morphology			
PM normal	37.2% (32/86)	45.9% (45/98)	n.s.
PM hypertrophy	57.0% (49/86)	52.0% (51/98)	n.s.
PM abnormal attachment	22.1% (19/86)	18.4% (18/98)	n.s.
PM apical displacement	14.0% (12/86)	4.1% (4/98)	0.018
PM bifid	2.3% (2/86)	1.0% (1/98)	n.s.
length anterior MV leaflet (PLAX) [mm]	22.0 +/- 2.4 (n=94)	21.6 +/- 2.8 (n=98)	n.s.
length anterior MV leaflet (4Ch) [mm]	21.8 +/- 2.4 (n=94)	20.4 +/- 2.2 (n=100)	<0.001
length anterior MV leaflet (3Ch) [mm]	21.7 +/- 2.4 (n=94)	21.1 +/- 2.6 (n=94)	n.s.

Table 7-11: Comparison of echocardiographic characteristics (part 2) of male and female carriers of *MYBPC3* c.3330+2T.G variant

	male	female	p value
MV inflow E [cm/s]	64.7 +/- 15.4 (n=67)	73.2 +/- 15.6 (n=80)	0.001
MV inflow A [cm/s]	52.3 +/- 17.3 (n=67)	59.4 +/- 15.9 (n=79)	0.011
MV inflow E/A	1.39 +/- 0.60 (n=66)	1.33 +/- 0.51 (n=79)	n.s.
MV inflow DT (ms)	175.8 +/- 42.0 (n=66)	170.5 +/- 43.6 (n=80)	n.s.
E' lateral [cm/s]	11.7 +/- 4.2 (n=96)	11.7 +/- 4.4 (n=106)	n.s.
A' lateral [cm/s]	8.5 +/- 2.7 (n=96)	9.2 +/- 3.0 (n=105)	n.s.
S' lateral [cm/s]	8.0 +/- 1.9 (n=96)	8.4 +/- 2.1 (n=106)	n.s.
E' lat/ A' lat	1.60 +/- 0.94 (n=96)	1.49 +/- 0.95 (n=105)	n.s.
E' septal [cm/s]	9.1 +/- 3.3 (n=95)	9.6 +/- 3.5 (n=104)	n.s.
A' septal [cm/s]	7.9 +/- 2.3 (n=95)	8.1 +/- 2.3 (n=103)	n.s.
S' septal [cm/s]	7.5 +/- 1.4 (n=95)	7.6 +/- 1.3 (n=104)	n.s.
E' sep/ A' sep	1.29 +/- 0.74 (n=95)	1.33 +/- 0.73 (n=103)	n.s.
E/ E' lateral	6.37 +/- 2.96 (n=65)	6.87 +/- 2.81 (n=79)	n.s.
E/ E' septal	8.47 +/- 4.64 (n=64)	8.29 +/- 3.08 (n=78)	n.s.
Average E/ E'	7.36 +/- 3.51 (n=65)	7.53 +/- 2.70 (n=79)	n.s.
IVRT [ms]	87.3 +/- 21.7 (n=93)	82.8 +/- 21.0 (n=98)	n.s.
PV _s [cm/s]	46.5 +/- 11.4 (n=73)	51.3 +/- 10.1 (n=90)	0.005
PV _D [cm/s]	46.7 +/- 12.1 (n=73)	43.9 +/- 9.4 (n=90)	n.s.
PV _s / PV _D	1.06 +/- 0.36 (n=73)	1.22 +/- 0.35 (n=90)	0.004
PV _a [cm/s]	28.2 +/- 5.9 (n=54)	29.1 +/- 7.6 (n=59)	n.s.
PV a _{dur} [ms]	128.3 +/- 31.8 (n=53)	125.0 +/- 22.4 (n=58)	n.s.

Table 7-12: Comparison of echocardiographic characteristics (part 3) of male and female carriers of MYBPC3 c.3330+2T.G variant

	age 18-40	age ≥ 41	p value
exertional chest pain	11.6% (17/146)	17.7% (25/141)	n.s.
exertional shortness of breath	37.0% (54/146)	54.6% (77/141)	0.003
palpitations	32.2% (47/146)	47.9% (67/140)	0.007
palpitations consistent with ectopics	15.1% (22/146)	12.9% (18/140)	n.s.
palpitations indeterminated	15.8% (23/146)	29.3% (41/140)	0.006
palpitations other	2.1% (3/146)	5.0% (7/140)	n.s.
dizziness	41.9% (49/117)	30.0% (33/110)	n.s.
dizziness after standing	21.4% (25/117)	10.9% (12/110)	0.033
dizziness after bending over	16.2% (19/117)	7.3% (8/110)	0.037
exertional dizziness	9.4% (11/117)	13.6% (15/110)	n.s.
other dizziness	2.6% (3/117)	3.6% (4/110)	n.s.
syncope	24.0% (35/146)	25.0% (35/140)	n.s.
neurocardiogenic syncope	19.2% (28/146)	19.3% (27/140)	n.s.
exertional syncope	2.8% (4/145)	1.4% (2/140)	n.s.
sudden syncope at rest	0.0%	3.6% (5/140)	0.022
indetermined syncope	2.1% (3/145)	2.1% (3/140)	n.s.
fatigue	3.2% (4/126)	10.4% (12/115)	0.024
aborted SCD	0.0%	0.0%	n.s.
any significant symptom	24.0% (30/125)	34.2% (38/111)	n.s.

Table 7-13: Comparison of symptoms of young (18-40 years) and older (≥ 41 years) carriers of *MYBPC3* c.3330+2T>G variant

	age 18-40	age ≥ 41	p value
heart rate [bpm]	66.0 +/- 10.3 (n=121)	67.3 +/- 11.2 (n=117)	n.s.
sinus rhythm	96.7% (118/122)	94.9% (111/117)	-
normal QRS axis	85.2% (104/122)	83.8% (98/117)	n.s.
right QRS axis	9.8% (12/122)	2.6% (3/117)	0.020
left QRS axis	4.1% (5/122)	10.3% (12/117)	n.s.
extreme axis	0.8% (1/122)	3.4% (4/117)	-
P wave duration [ms]	106.3 +/- 11.3 (n=121)	113.8 +/- 14.6 (n=111)	<0.001
RA enlargement	1.7% (2/120)	2.7% (3/110)	-
LA enlargement	1.7% (2/121)	9.9% (11/111)	0.006
PR duration [ms]	165.7 +/- 20.5 (n=121)	176.7 +/- 28.9 (n=111)	0.001
RVH Sokolow-Lyon	25.6% (31/121)	13.8% (16/116)	0.022
LVH Sokolow Lyon	15.7% (19/121)	2.6% (3/116)	0.001
intrinsicoid deflection [ms]	37.6 +/- 6.8 (n=120)	37.5 +/- 8.6 (n=114)	n.s.
QRS duration [ms]	85.8 +/- 13.3 (n=112)	80.3 +/- 13.4 (n=125)	0.002
LBbB	0.8% (1/121)	1.7% (2/115)	-
RBbB	0.0%	0.9% (1/116)	-
QRS fragmentation	64.4% (76/118)	70.5% (79/112)	n.s.
Q waves	5.9% (7/118)	9.6% (11/114)	n.s.
QT duration [ms]	397.9 +/- 29.8 (n=121)	411.1 +/- 34.2 (n=114)	0.002
QTc duration [ms]	414.3 +/- 22.3 (n=121)	431.4 +/- 27.4 (n=114)	<0.001
T wave inversion (TWI)	1.7% (2/121)	13.8% (16/116)	<0.001
deep TWI	0.8% (1/121)	6.0% (7/116)	-
ST elevation	0.9% (1/114)	0.0%	-
RE score ≥ 4	8.2% (10/122)	17.9% (21/117)	0.025
RE score ≥ 5	4.1% (5/122)	10.3% (12/117)	n.s.
abnormal ECG	13.6% (16/118)	34.8% (40/115)	<0.001

Table 7-14: Comparison of ECG characteristics of young (18-40 years) and older (≥ 41 years) carriers of the MYBPC3 c.3330+2T>G variant

	18-40 years	≥ 41 years	p value
number of PVCs	11.6 +/- 47.5 (n=30)	1108.2 +/- 2002.9 (n=33)	0.004
number of episodes of bigeminy	0.1 +/- 0.5 (n=30)	35.6 +/- 117.9 (n=33)	n.s.
number of couplets	0.0 +/- 0.2 (n=29)	20.8 +/- 100.1 (n=33)	n.s.
number of triplets	0.0 +/- 0.2 (n=29)	0.8 +/- 3.3 (n=33)	n.s.
individuals with NSVT ≥ 4 beats [%]	3.3% (1/30)	24.2% (8/33)	0.018
min heart rate [bpm]	53.5 +/- 6.0 (n=30)	55.9 +/- 7.3 (n=33)	n.s.
max heart rate [bpm]	135.6 +/- 16.4 (n=30)	123.9 +/- 12.5 (n=33)	0.002
mean heart rate [bpm]	78.9 +/- 6.6 (n=30)	76.7 +/- 9.5 (n=33)	n.s.

Table 7-15: Comparison of holter characteristics of young (age 18-40) and older (≥ 41 years) carriers of the *MYBPC3* c.3330+2T>G variant

	18-40 years	≥ 41 years	p value
No hypertrophy (max WT<13mm)	66.4% (85/128)	32.8% (40/122)	<0.001
Type of hypertrophy			
sigmoid septum	18.0% (23/128)	32.8% (40/122)	0.007
reverse septal curvature	10.9% (14/128)	29.5% (36/122)	<0.001
apical hypertrophy	1.6% (2/128)	3.3% (4/122)	n.s.
concentric hypertrophy	1.6% (2/128)	1.6% (2/122)	-
segmental hypertrophy	1.6% (2/128)	2.5% (3/122)	-
dilated phase	0.0%	0.8% (1/122)	-
SAM	8.9% (9/101)	15.8% (15/95)	n.s.
cordal/PM SAM	10.9% (11/101)	14.7% (14/95)	n.s.
PM morphology			
PM normal	50.0% (48/96)	33.0% (29/88)	0.019
PM hypertrophy	49.0% (47/96)	60.2% (53/88)	n.s.
PM abnormal attachment	18.8% (18/96)	21.6% (19/88)	n.s.
PM apical displacement	7.3% (7/96)	10.2% (9/88)	n.s.
PM bifid	3.1% (3/96)	0.0% (0/88)	n.s.
length anterior MV leaflet (PLAX) [mm]	21.5 +/- 2.5 (n=102)	22.2 +/- 2.7 (n=90)	n.s.
length anterior MV leaflet (4Ch) [mm]	21.3 +/- 2.4 (n=101)	20.9 +/- 2.5 (n=93)	n.s.
length anterior MV leaflet (3Ch) [mm]	21.6 +/- 2.4 (n=99)	21.2 +/- 2.6 (n=89)	n.s.

Table 7-16: Comparison of echocardiographic parameters (part 2) of young (age 18-40) and older (≥ 41 years) carriers of the *MYBPC3* c.3330+2T>G variant

7.2.4 Diastolic function in *MYBPC3* c.3330+2T>G variant carriers without hypertrophy

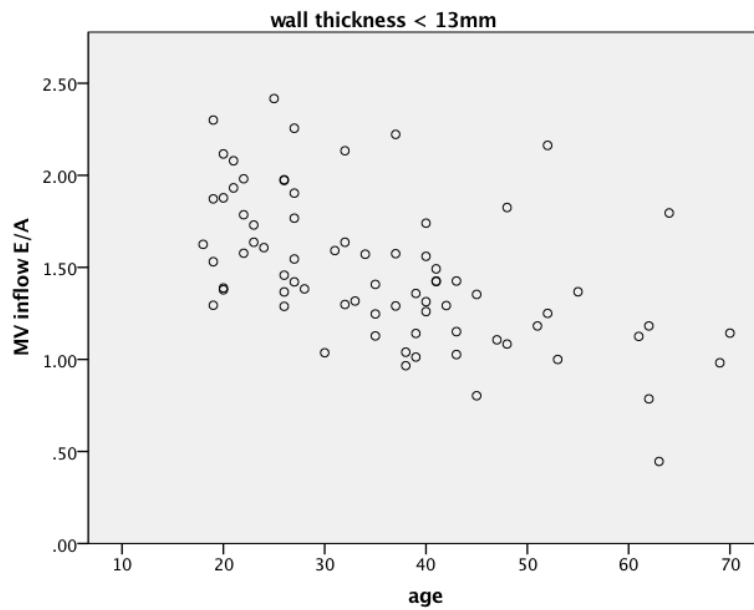


Figure 7-14: MV inflow E/A ratio in carriers of the *MYBPC3* c.3330+2T>G variant without hypertrophy

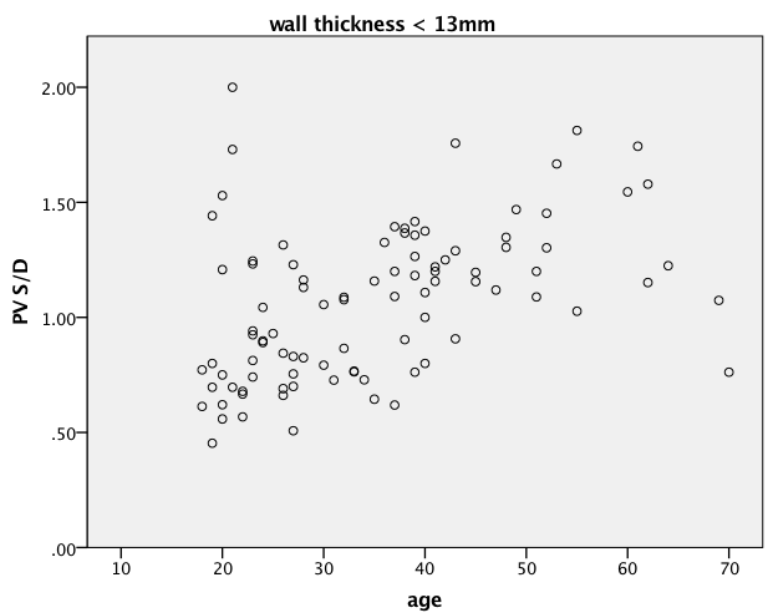


Figure 7-15: Pulmonary venous doppler S/D ratio in carriers of the *MYBPC3* c.3330+2T>G variant without hypertrophy

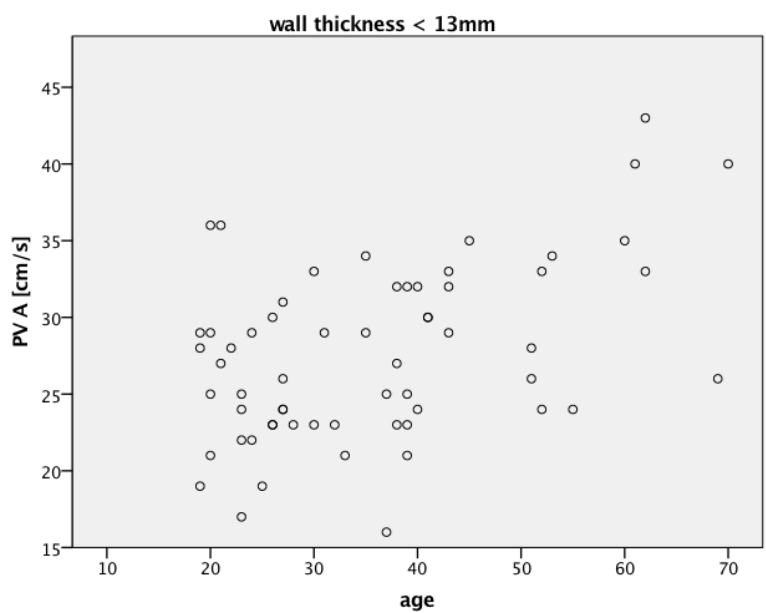


Figure 7-16: Pulmonary venous doppler A in carriers of the *MYBPC3* c.3330+2T>G variant without hypertrophy

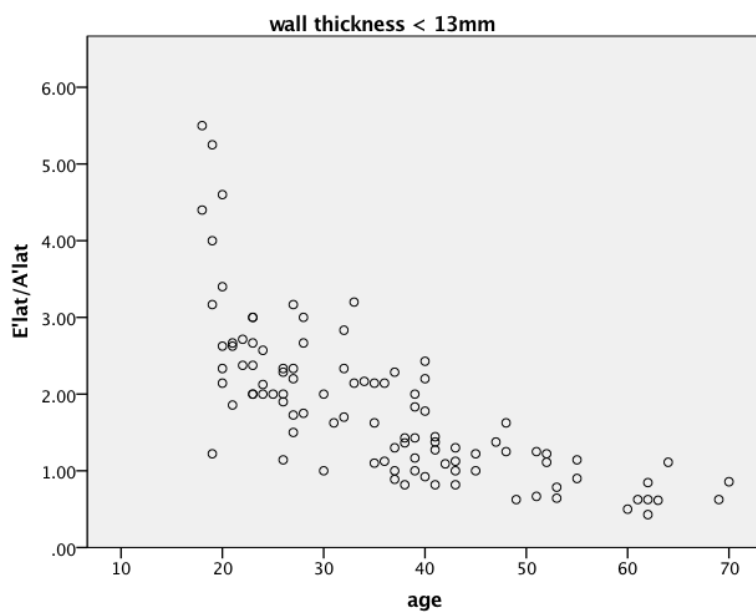


Figure 7-17: Lateral mitral annulus tissue Doppler E'/A' in carriers of the MYBPC3 c.3330+2T>G variant without hypertrophy

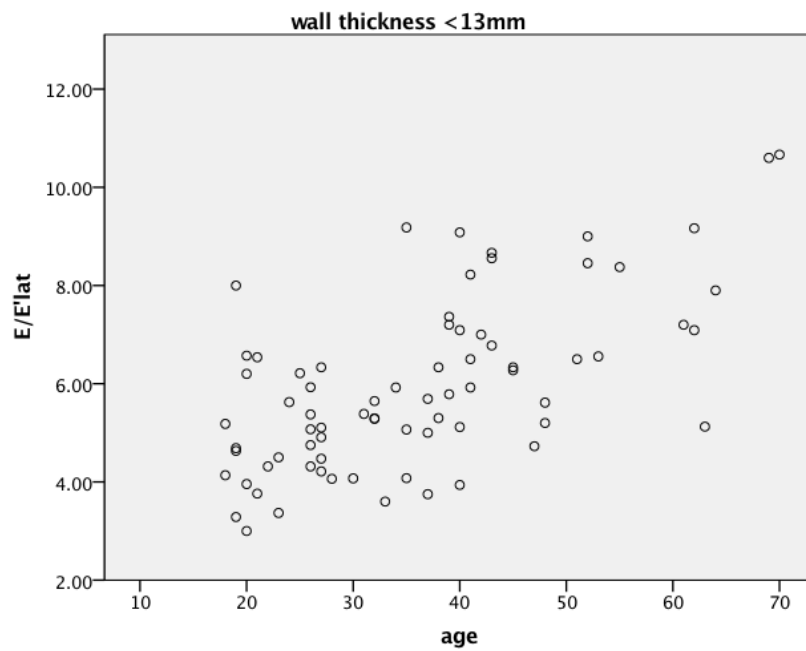


Figure 7-18: E/E' lateral mitral annulus in carriers of the *MYBPC3* c.3330+2T>G variant without hypertrophy

7.2.5 Exome sequence analysis of *MYBPC3* c.3330+2T>G variant carriers with severe phenotype

Arrhythmogenic cardiomyopathy	Cardiac arrhythmias	Dilated Cardiomyopathy & conduction defects	Hypertrophic cardiomyopathy - teen & adult	Left Ventricular Noncompaction Cardiomyopathy
DES	ANK2	ABCC9	ACTC1	ACTC1
DSC2	CACNA1C	ACTC1	CSRP3	MYBPC3
DSG2	CACNB2	ACTN2	FHL1	MYH7
DSP	CALM1	BAG3	GLA	TAZ
JUP	CALM2	CSRP3	LAMP2	TNNT2
PKP2	CASQ2	DES	LZTR1	TPM1
PLN	GPD1L	DMD	MYBPC3	
RYR2	HCN4	DSP	MYH7	
TMEM43	KCNE1	EPG5	MYL2	
	KCNE2	EYA4	MYL3	
	KCNE3	FKTN	PLN	
	KCNH2	HAMP	PRKAG2	
	KCNJ2	HFE	TNNC1	
	KCNJ5	HFE2	TNNI3	
	KCNQ1	IDH2	TNNT2	
	RYR2	LDB3	TPM1	
	SCN10A	LMNA	TTR	
	SCN1B	MYBPC3		
	SCN3B	MYH6		
	SCN5A	MYH7		
	SLC22A5	NEXN		
	SLC4A3	PLN		
	SNTA1	PPP1R13L		
	TNNI3K	RAB3GAP2		
	TRDN	RBM20		
	TRPM4	SCN5A		
		SGCD		
		SLC40A1		
		TAZ		
		TCAP		
		TFR2		
		TNNC1		
		TNNI3		
		TNNT2		
		TPM1		
		TTN		
		VCL		

Table 7-17: Summary of genes included in virtual panel for exome sequence analysis (derived from PanelApp)

7.2.6 Information leaflet on hypertrophic cardiomyopathy in the Anabaptist communities

Figure 1: An idealized family diagram illustrating the inheritance pattern of HCM when both parents possess one copy of the HCM gene.

The chance of passing on a gene with a mutation occurs randomly. The above chances apply to every pregnancy and are the same for boys and for girls.

Windows of Hope Genetics Information Center, Inc.
WHGIC helps individuals and families by providing genetic testing, information on disorders and counseling for coping with inherited health conditions. Information about known treatments or therapies is also provided.

Mission Statement
Windows of Hope Genetics Information Center shall provide information and resources for families and professionals to aid in the diagnosis, treatment and care of individuals with inherited disorders. It shall do studies of, provide education and training about, develop and co-ordinate treatment of genetic disease and risks.

WHGIC is a nonprofit 501(c)3 organization which operates by donations and support from individuals, local businesses, grants and health care funds. Your support is greatly appreciated.

To support and/or get involved with WHGIC, contact the WHGIC office or the Board of Directors at any time.

Anyone desiring help or information about hypertrophic cardiomyopathy or a genetic disorder please contact WHGIC at (330) 893-0220, or visit the following website for more information about genetic disorders.
www.WOHproject.org

Printed by Holmes Printing Solutions; Mt. Hope OH (888) 473-6870
Version 3; August 2013

**HYPERTROPHIC
CARDIOMYOPATHY**

**Serving Holmes County
and surrounding
Anabaptist Communities**

**Improving Care for Individuals
with Special Needs
and Addressing Health Care Costs
through Current Genetic Information**

**WHGIC, Inc.
P.O. Box 208
Walnut Creek, OH 44687
(330) 893-0220
info@wohgeneticinfo.org
www.wohgeneticinfo.org**

WHAT IS HYPERTROPHIC CARDIOMYOPATHY?

The heart is a specialized muscle that pumps the blood through the lungs and the body. In hypertrophic cardiomyopathy (HCM) the heart muscle cells become enlarged. The word "hypertrophy" means enlargement; in this case a thickening of the walls of the heart. The thickened walls make the heart stiff and prevent it from filling and pumping properly, especially with exertion. The thickening can also lead to abnormal and fast heartbeats, which can result in blackouts and in rare cases sudden death. The amount of hypertrophy that people with HCM develop varies greatly, even among members of the same family.

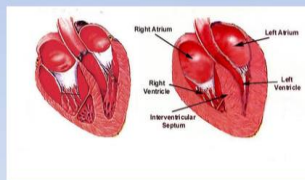


Figure showing the thickened heart muscle in patients with HCM

WHAT ARE THE SIGNS AND SYMPTOMS?

HCM varies widely in its initial symptoms and in how the disease progresses. A lot of people never experience any symptoms and lead normal lives. Others, however, have

chest pains, shortness of breath, irregular heartbeats, or blackouts. A few individuals develop worsening symptoms and lead severely restricted lives. Heart rhythm irregularities causing sudden death can be a rare manifestation of HCM, even in those who have no prior symptoms.

WHAT CAUSES HCM?

HCM affects about 1 in 500 people, although the disease may occur more frequently in Anabaptist communities. It can affect children, teenagers and adults of all ages. People generally develop HCM because they inherited an abnormal gene from a parent. As treatments are available, it is important that individuals with a family history of the condition are tested to see if they carry the abnormal gene. If they are found to carry the abnormal gene they can receive appropriate follow-ups, advice, and treatment.

HOW IS HCM DIAGNOSED AND TREATED?

Although there is no cure for HCM, it can be managed. Early detection is important in trying to ward off life-threatening problems. A gene test can determine if someone has the abnormal gene commonly causing HCM in the Anabaptist community, even before the onset of symptoms. Gene tests can be arranged by contacting WHGIC. Medications like beta blockers and calcium channel blockers can treat symptoms by improving

the function of the thickened heart muscle and preventing abnormal heart beats. In some patients, surgery may be used to remove portions of the thickened heart muscle. If necessary, a defibrillator (or pacemaker-like device) may be implanted under the skin to treat dangerous heart rhythms.

HCM IN ANABAPTISTS

HCM is usually caused by a single abnormal gene inherited from one parent. The abnormal gene is able to cause the disease despite the presence of a healthy copy of the gene inherited from the other parent (an autosomal dominant condition). We have discovered that a particular gene defect responsible for HCM is very common in the Anabaptist community. Therefore, on occasion, two affected individuals may marry who both possess a copy of the same abnormal gene responsible for HCM. Such couples have a 25% (1 in 4) chance that a child will inherit a copy of the abnormal gene from each parent (see Figure 1 on next page). Without having a normal copy of the gene, these children develop severe HCM with serious heart problems and usually do not live beyond a year of age. There is also a 50% (2 in 4) chance that any child will inherit just one copy of the abnormal gene and may develop HCM as a teenager or adult. Lastly, there is a 25% (1 in 4) chance that the child will inherit both normal copies of the gene. In this case the child will not have this type of HCM.

7.2.7 Information day on hypertrophic cardiomyopathy



Windows of Hope Genetics Information Center

Providing information and resources for families and professionals treating inherited disorders.

INFORMATION DAY ON HYPERTROPHIC CARDIOMYOPATHY

SATURDAY, NOVEMBER 2, 2013; 10:00AM – 3:00PM

LOCATED AT: THE PIER at Lighthouse of Hope
14070 Winesburg Road; Dundee, OH 44624

Everyone is welcome. We would particularly like to invite all the participants of the Heart Program and their families to this event. **Attendance of this meeting is free of charge** and we will provide a light lunch for everyone.

10:00am	Arrival and sign-in (coffee and refreshments)
10:30am	Welcome, Overview and Introductions (Windows of Hope team)
11:00am	“Hypertrophic Cardiomyopathy in Adults” (Dr. Harry Lever, Medical Director of the Hypertrophic Cardiomyopathy Clinic at Cleveland Clinic) <i>What is HCM, what are the complications and how is it treated?</i>
12:00pm	“Hypertrophic Cardiomyopathy in Children” (Dr. Ken Zahka, Pediatric Cardiologist at Cleveland Clinic) <i>Do children need to be tested for HCM? Are there any special considerations and how is it treated?</i>
1:00pm	Lunch (will be provided)
2:00pm	Q&A session with Dr. Lever and Dr. Zahka for local Physicians

Please register for this event by October 12th, 2013 by calling WHGIC at: (330) 893-0220 or returning the bottom portion to WHGIC at P.O. Box 208 Walnut Creek, OH 44687. Please indicate the number of attendees in your party.

For Geauga County, Ohio participants: please contact Kay Gurtz at (440) 667-3388. A bus is organized round-trip to the event for Geauga County families and will load from 7:45 – 8:15am at the Wal-Mart parking lot; Middlefield, OH (ticket \$10).

Number of person(s) attending Information Day on HCM: _____

Name: _____

Address: _____

Phone: _____

Please return to WHGIC; P.O. Box 208; Walnut Creek, OH 44687

7.3 Contributions from individuals and other institutions to the practical work described in this thesis

7.3.1 Mutations in *HYAL2*, encoding hyaluronidase 2, cause a syndrome of cor triatriatum sinister and orofacial clefting in humans and mice

Phenotypic investigation including performing physical examinations, 12-lead electrocardiograms and transthoracic echocardiograms of the five Amish children was performed by Dr M Muggenthaler.

Dr E Baple, Consultant Clinical Geneticist, Royal Devon and Exeter NHS Trust advised regarding the description of the dysmorphic features of the phenotype.

Audiograms were reviewed by Dr Wendy Albuquerque at St George's Healthcare NHS Trust.

Blood draws were performed by members of the Windows of Hope team.

DNA extraction was performed by Dr M Muggenthaler.

Whole exome sequencing was performed by the Otogenetics Corporation.

Genome wide microarray scan using Illumina Human CytoSNP-12v2.1 arrays was undertaken by Dr M Muggenthaler with help from Dr B Chioza, University of Exeter.

Support for the analysis of the SNP- genotype data and exome sequencing data was provided by Dr B Chioza and Prof M Wheedon, based at University of Exeter. Dr M Muggenthaler evaluated the identified potentially putative candidate variants.

Dr M Muggenthaler performed Sanger dideoxy sequencing to validate the c.443A>G *HYAL2* variant and confirm co-segregation of the variant within the extended Amish pedigree.

Dr G Harlalka performed Sanger sequencing to assess the carrier frequency in 532 Amish control chromosomes.

Functional studies were performed to assess the impact of p.K148R and p.P250L on the HYAL2 protein, including the introduction of the variants into a human HYAL2 construct and transfecting it into mouse embryonic fibroblasts, western blot analysis, immunofluorescence microscopy and activity assays by Professor B Triggs-Raine's group, University of Manitoba, Winnipeg, Canada.

Mouse studies were designed by Prof Triggs-Raine and Dr B Chowdhury based at University of Manitoba, Winnipeg, Canada. Following discussion of the human phenotype with Dr M Muggenthaler, Dr Chowdhury performed detailed analysis of the mouse phenotype.

Phenotypic assessment and genetic investigation of the two non-Amish study participants were performed by Professor Fowzan S. Alkuraya's group at the Department of Genetics, King Faisal Specialist Hospital and Research Center, Riyadh, Saudi Arabia.

7.3.2 Investigation of the phenotype and genetic modifiers in hypertrophic cardiomyopathy caused by the c.3330+2T>G variant in myosin binding protein C

Recruitment to the study was coordinated by Dr M Muggenthaler and supported by members of the Windows of Hope team (Roselyn Coblentz, Karin Wagner and Kay Gurtz), who also performed sample collections and all supporting administrative work.

331 carriers of the *MYBPC3* variant were identified and invited to a consultation, electrocardiogram, transthoracic echocardiogram and 24-hour holter monitor. All consultations and clinical assessments were performed by Dr M Muggenthaler.

Phenotype data were analysed using IBM SPSS software by Dr M Muggenthaler.

RNA was extracted using Qiagen PAX gene blood RNA kit by Dr M Muggenthaler and Dr G Harlalka.

DNA extraction was performed by Dr M Muggenthaler and Dr G Harlalaka.

Transcript studies were undertaken in collaboration with Prof Lorna Harries at University of Exeter. The studies were carried out by Dr M Muggenthaler working closely with Dr Ben Lee and Dr Gaurav Harlalaka based at the University of Exeter, who also performed a large part of the laboratory work. Prof Lorna Harries provided support with the data analysis.

Whole-exome sequencing of DNA of severely affected individuals was outsourced to Baylor College of Medicine Human Genome Sequencing Center (BCM-HGSC). Dr J Fasham and Dr B Chioza, based at the University of Exeter, provided support for the analysis of the exome sequencing data. Dr M Muggenthaler evaluated the identified putative modifiers as described.

Genotyping on the Affymetrix Genome-Wide Human SNP Array 6.0 of 202 HCM gene carriers was outsourced to the company EpiGenDx (www.epigenDx.com).

Dr M Muggenthaler chose the variables and parameters for the genome wide SNP association study. Dr Michael Weedon and Dr Sam Jones, based at the University of Exeter performed the genome wide SNP association study.

References

- Adalsteinsdottir, B., P. Teekakirikul, B. J. Maron, M. A. Burke, D. F. Gudbjartsson, H. Holm, K. Stefansson, S. R. DePalma, E. Mazaika, B. McDonough, R. Danielsen, J. G. Seidman, C. E. Seidman and G. T. Gunnarsson (2014). "Nationwide study on hypertrophic cardiomyopathy in Iceland: evidence of a MYBPC3 founder mutation." *Circulation* **130**(14): 1158-1167.
- Akinrinade, O., T. P. Alastalo and J. W. Koskenvuo (2016). "Relevance of truncating titin mutations in dilated cardiomyopathy." *Clin Genet* **90**(1): 49-54.
- Akinrinade, O., T. Helio, R. H. Lekanne Deprez, J. D. H. Jongbloed, L. G. Boven, M. P. van den Berg, Y. M. Pinto, T. P. Alastalo, S. Myllykangas, K. V. Spaendonck-Zwarts, J. P. V. Tintelen, P. A. van der Zwaag and J. Koskenvuo (2019). "Relevance of Titin Missense and Non-Frameshifting Insertions/Deletions Variants in Dilated Cardiomyopathy." *Sci Rep* **9**(1): 4093.
- Akinrinade, O., L. Ollila, S. Vattulainen, J. Tallila, M. Gentile, P. Salmenpera, H. Koillinen, M. Kaartinen, M. S. Nieminen, S. Myllykangas, T. P. Alastalo, J. W. Koskenvuo and T. Helio (2015). "Genetics and genotype-phenotype correlations in Finnish patients with dilated cardiomyopathy." *Eur Heart J* **36**(34): 2327-2337.
- Alonso-Montes, C., M. Naves-Diaz, J. L. Fernandez-Martin, J. Rodriguez-Reguero, C. Moris, E. Coto, J. B. Cannata-Andia and I. Rodriguez (2012). "New polymorphisms in human MEF2C gene as potential modifier of hypertrophic cardiomyopathy." *Mol Biol Rep* **39**(9): 8777-8785.
- Andre, B., C. Duterme, K. Van Moer, J. Mertens-Strijthagen, M. Jadot and B. Flamion (2011). "Hyal2 is a glycosylphosphatidylinositol-anchored, lipid raft-associated hyaluronidase." *Biochem Biophys Res Commun* **411**(1): 175-179.
- Barefield, D., M. Kumar, J. Gorham, J. G. Seidman, C. E. Seidman, P. P. de Tombe and S. Sadayappan (2015). "Haploinsufficiency of MYBPC3 exacerbates the development of hypertrophic cardiomyopathy in heterozygous mice." *J Mol Cell Cardiol* **79**: 234-243.
- Bassett, D. R., P. L. Schneider and G. E. Huntington (2004). "Physical activity in an Old Order Amish community." *Med Sci Sports Exerc* **36**(1): 79-85.
- Basson, C. T., D. R. Bachinsky, R. C. Lin, T. Levi, J. A. Elkins, J. Soultz, D. Grayzel, E. Kroumpouzou, T. A. Traill, J. Leblanc-Straceski, B. Renault, R. Kucherlapati, J. G. Seidman and C. E. Seidman (1997). "Mutations in human TBX5 [corrected] cause limb and cardiac malformation in Holt-Oram syndrome." *Nat Genet* **15**(1): 30-35.
- Benson, D. W., G. M. Silberbach, A. Kavanaugh-McHugh, C. Cottrill, Y. Zhang, S. Riggs, O. Smalls, M. C. Johnson, M. S. Watson, J. G. Seidman, C. E. Seidman, J. Plowden and J. D. Kugler (1999). "Mutations in the cardiac transcription factor NKX2.5 affect diverse cardiac developmental pathways." *J Clin Invest* **104**(11): 1567-1573.
- Beriaghi, S., S. L. Myers, S. A. Jensen, S. Kaimal, C. M. Chan and G. B. Schaefer (2009). "Cleft lip and palate: association with other congenital malformations." *J Clin Pediatr Dent* **33**(3): 207-210.

Bharucha, T., D. E. Spicer, T. J. Mohun, D. Black, G. W. Henry and R. H. Anderson (2015). "Cor triatriatum or divided atriums: which approach provides the better understanding?" Cardiol Young **25**(2): 193-207.

Bick, A. G., J. Flannick, K. Ito, S. Cheng, R. S. Vasan, M. G. Parfenov, D. S. Herman, S. R. DePalma, N. Gupta, S. B. Gabriel, B. H. Funke, H. L. Rehm, E. J. Benjamin, J. Aragam, H. A. Taylor, Jr., E. R. Fox, C. Newton-Cheh, S. Kathiresan, C. J. O'Donnell, J. G. Wilson, D. M. Altshuler, J. N. Hirschhorn, J. G. Seidman and C. Seidman (2012). "Burden of rare sarcomere gene variants in the Framingham and Jackson Heart Study cohorts." Am J Hum Genet **91**(3): 513-519.

Bingle, G. J., J. Dillon and R. Hurwitz (1975). "Asymmetric septal hypertrophy in a large amish kindred." Clin Genet **7**(3): 255-261.

Blalock, A. (1945). "Physiopathology and surgical treatment of congenital cardiovascular defects." Harvey Lect **41**: 90-116.

Boskovski, M. T., S. Yuan, N. B. Pedersen, C. K. Goth, S. Makova, H. Clausen, M. Brueckner and M. K. Khokha (2013). "The heterotaxy gene GALNT11 glycosylates Notch to orchestrate cilia type and laterality." Nature **504**(7480): 456-459.

Boudjadi, S., B. Chatterjee, W. Sun, P. Vemu and F. G. Barr (2018). "The expression and function of PAX3 in development and disease." Gene **666**: 145-157.

Bourguignon, L. Y., P. A. Singleton, F. Diedrich, R. Stern and E. Gilad (2004). "CD44 interaction with Na⁺-H⁺ exchanger (NHE1) creates acidic microenvironments leading to hyaluronidase-2 and cathepsin B activation and breast tumor cell invasion." J Biol Chem **279**(26): 26991-27007.

Brion, M., C. Allegue, M. Santori, R. Gil, A. Blanco-Verea, C. Haas, C. Bartsch, S. Poster, B. Madea, O. Campuzano, R. Brugada and A. Carracedo (2012). "Sarcomeric gene mutations in sudden infant death syndrome (SIDS)." Forensic Sci Int **219**(1-3): 278-281.

Buckingham, M., S. Meilhac and S. Zaffran (2005). "Building the mammalian heart from two sources of myocardial cells." Nat Rev Genet **6**(11): 826-835.

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

Camenisch, T. D., J. Biesterfeldt, T. Brehm-Gibson, J. Bradley and J. A. McDonald (2001). "Regulation of cardiac cushion development by hyaluronan." Exp Clin Cardiol **6**(1): 4-10.

Camenisch, T. D., J. A. Schroeder, J. Bradley, S. E. Klewer and J. A. McDonald (2002). "Heart-valve mesenchyme formation is dependent on hyaluronan-augmented activation of ErbB2-ErbB3 receptors." Nat Med **8**(8): 850-855.

Camenisch, T. D., A. P. Spicer, T. Brehm-Gibson, J. Biesterfeldt, M. L. Augustine, A. Calabro, Jr., S. Kubalak, S. E. Klewer and J. A. McDonald (2000). "Disruption of hyaluronan synthase-2 abrogates normal cardiac morphogenesis and hyaluronan-mediated transformation of epithelium to mesenchyme." J Clin Invest **106**(3): 349-360.

Chao, K. L., L. Muthukumar and O. Herzberg (2007). "Structure of human hyaluronidase-1, a hyaluronan hydrolyzing enzyme involved in tumor growth and angiogenesis." Biochemistry **46**(23): 6911-6920.

Charron, P., O. Dubourg, M. Desnos, M. Bennaceur, L. Carrier, A. C. Camproux, R. Isnard, A. Hagege, J. M. Langlard, G. Bonne, P. Richard, B. Hainque, J. B. Bouhour, K. Schwartz and M. Komajda (1998). "Clinical features and prognostic implications of familial hypertrophic cardiomyopathy related to the cardiac myosin-binding protein C gene." *Circulation* **97**(22): 2230-2236.

Chen, Q., D. Zhang, R. L. Gingell, A. J. Moss, C. Napolitano, S. G. Priori, P. J. Schwartz, E. Kehoe, J. L. Robinson, E. Schulze-Bahr, Q. Wang and J. A. Towbin (1999). "Homozygous deletion in KVLQT1 associated with Jervell and Lange-Nielsen syndrome." *Circulation* **99**(10): 1344-1347.

Chowdhury, B., R. Hemming, S. Hombach-Klonisch, B. Flamion and B. Triggs-Raine (2013). "Murine hyaluronidase 2 deficiency results in extracellular hyaluronan accumulation and severe cardiopulmonary dysfunction." *J Biol Chem* **288**(1): 520-528.

Cole, T. J., J. V. Freeman and M. A. Preece (1998). "British 1990 growth reference centiles for weight, height, body mass index and head circumference fitted by maximum penalized likelihood." *Stat Med* **17**(4): 407-429.

Corrigan, N., D. P. Brazil and F. M. Auliffe (2010). "High-frequency ultrasound assessment of the murine heart from embryo through to juvenile." *Reprod Sci* **17**(2): 147-157.

Csoka, A. B., S. W. Scherer and R. Stern (1999). "Expression analysis of six paralogous human hyaluronidase genes clustered on chromosomes 3p21 and 7q31." *Genomics* **60**(3): 356-361.

Davey, K. M., J. S. Parboosingh, D. R. McLeod, A. Chan, R. Casey, P. Ferreira, F. F. Snyder, P. J. Bridge and F. P. Bernier (2006). "Mutation of DNAJC19, a human homologue of yeast inner mitochondrial membrane co-chaperones, causes DCMA syndrome, a novel autosomal recessive Barth syndrome-like condition." *J Med Genet* **43**(5): 385-393.

Davies, G. J., A. Planas and C. Rovira (2012). "Conformational analyses of the reaction coordinate of glycosidases." *Acc Chem Res* **45**(2): 308-316.

Davies, G. J., K. S. Wilson and B. Henrissat (1997). "Nomenclature for sugar-binding subsites in glycosyl hydrolases." *Biochem J* **321** (Pt 2): 557-559.

Daw, E. W., S. N. Chen, G. Czernuszewicz, R. Lombardi, Y. Lu, J. Ma, R. Roberts, S. Shete and A. J. Marian (2007). "Genome-wide mapping of modifier chromosomal loci for human hypertrophic cardiomyopathy." *Hum Mol Genet* **16**(20): 2463-2471.

de la Motte, C., J. Nigro, A. Vasanji, H. Rho, S. Kessler, S. Bandyopadhyay, S. Danese, C. Fiocchi and R. Stern (2009). "Platelet-derived hyaluronidase 2 cleaves hyaluronan into fragments that trigger monocyte-mediated production of proinflammatory cytokines." *Am J Pathol* **174**(6): 2254-2264.

De, S., A. G. Borowski, H. Wang, L. Nye, B. Xin, J. D. Thomas and W. H. Tang (2011). "Subclinical echocardiographic abnormalities in phenotype-negative carriers of myosin-binding protein C3 gene mutation for hypertrophic cardiomyopathy." *Am Heart J* **162**(2): 262-267 e263.

Deo, R. C. (2016). "Alternative Splicing, Internal Promoter, Nonsense-Mediated Decay, or All Three: Explaining the Distribution of Truncation Variants in Titin." *Circ Cardiovasc Genet* **9**(5): 419-425.

Dhandapany, P. S., S. Sadayappan, Y. Xue, G. T. Powell, D. S. Rani, P. Nallari, T. S. Rai, M. Khullar, P. Soares, A. Bahl, J. M. Tharkan, P. Vaideeswar, A. Rathinavel, C. Narasimhan, D. R. Ayapati, Q. Ayub, S. Q. Mehdi, S. Oppenheimer, M. B. Richards, A. L. Price, N. Patterson, D. Reich, L. Singh, C. Tyler-Smith and K. Thangaraj (2009). "A common MYBPC3 (cardiac myosin binding protein C) variant associated with cardiomyopathies in South Asia." *Nat Genet* **41**(2): 187-191.

Dominguez, F., S. Cuenca, Z. Bilinska, R. Toro, E. Villard, R. Barriaes-Villa, J. P. Ochoa, F. Asselbergs, A. Sammani, M. Franaszczyk, M. Akhtar, M. J. Coronado-Albi, D. Rangel-Sousa, J. F. Rodriguez-Palomares, J. Jimenez-Jaimez, J. M. Garcia-Pinilla, T. Ripoll-Vera, M. V. Mogollon-Jimenez, A. Fontalba-Romero, D. Garcia-Medina, J. Palomino-Doza, D. de Gonzalo-Calvo, M. Cicerchia, J. Salazar-Mendiguchia, C. Salas, S. Pankuweit, T. M. Hey, J. Mogensen, P. J. Barton, P. Charron, P. Elliott, P. Garcia-Pavia and I. European Genetic Cardiomyopathies Initiative (2018). "Dilated Cardiomyopathy Due to BLC2-Associated Athanogene 3 (BAG3) Mutations." *J Am Coll Cardiol* **72**(20): 2471-2481.

Durst, R., K. Sauls, D. S. Peal, A. deVlaming, K. Toomer, M. Leyne, M. Salani, M. E. Talkowski, H. Brand, M. Perrocheau, C. Simpson, C. Jett, M. R. Stone, F. Charles, C. Chiang, S. N. Lynch, N. Bouatia-Naji, F. N. Delling, L. A. Freed, C. Tribouilloy, T. Le Tourneau, H. LeMarec, L. Fernandez-Friera, J. Solis, D. Trujillano, S. Ossowski, X. Estivill, C. Dina, P. Bruneval, A. Chester, J. J. Schott, K. D. Irvine, Y. Mao, A. Wessels, T. Motiwala, M. Puceat, Y. Tsukasaki, D. R. Menick, H. Kasiganesan, X. Nie, A. M. Broome, K. Williams, A. Johnson, R. R. Markwald, X. Jeunemaitre, A. Hagege, R. A. Levine, D. J. Milan, R. A. Norris and S. A. Slaugenhaupt (2015). "Mutations in DCHS1 cause mitral valve prolapse." *Nature* **525**(7567): 109-113.

Dyer, L. A. and M. L. Kirby (2009). "The role of secondary heart field in cardiac development." *Dev Biol* **336**(2): 137-144.

Easton, D. F., K. A. Pooley, A. M. Dunning, P. D. Pharoah, D. Thompson, D. G. Ballinger, J. P. Struwing, J. Morrison, H. Field, R. Luben, N. Wareham, S. Ahmed, C. S. Healey, R. Bowman, S. collaborators, K. B. Meyer, C. A. Haiman, L. K. Kolonel, B. E. Henderson, L. Le Marchand, P. Brennan, S. Sangrajrang, V. Gaborieau, F. Odefrey, C. Y. Shen, P. E. Wu, H. C. Wang, D. Eccles, D. G. Evans, J. Peto, O. Fletcher, N. Johnson, S. Seal, M. R. Stratton, N. Rahman, G. Chenevix-Trench, S. E. Bojesen, B. G. Nordestgaard, C. K. Axelsson, M. Garcia-Closas, L. Brinton, S. Chanock, J. Lissowska, B. Peplonska, H. Nevanlinna, R. Fagerholm, H. Eerola, D. Kang, K. Y. Yoo, D. Y. Noh, S. H. Ahn, D. J. Hunter, S. E. Hankinson, D. G. Cox, P. Hall, S. Wedren, J. Liu, Y. L. Low, N. Bogdanova, P. Schurmann, T. Dork, R. A. Tollenaar, C. E. Jacobi, P. Devilee, J. G. Klijn, A. J. Sigurdson, M. M. Doody, B. H. Alexander, J. Zhang, A. Cox, I. W. Brock, G. MacPherson, M. W. Reed, F. J. Couch, E. L. Goode, J. E. Olson, H. Meijers-Heijboer, A. van den Ouweland, A. Uitterlinden, F. Rivadeneira, R. L. Milne, G. Ribas, A. Gonzalez-Neira, J. Benitez, J. L. Hopper, M. McCredie, M. Southey, G. G. Giles, C. Schroen, C. Justenhoven, H. Brauch, U. Hamann, Y. D. Ko, A. B. Spurdle, J. Beesley, X. Chen, kConFab, A. M. Group, A. Mannermaa, V. M. Kosma, V. Kataja, J. Hartikainen, N. E. Day, D. R. Cox and B. A. Ponder (2007). "Genome-wide association study identifies novel breast cancer susceptibility loci." *Nature* **447**(7148): 1087-1093.

Elliott, P. and W. J. McKenna (2004). "Hypertrophic cardiomyopathy." *Lancet* **363**(9424): 1881-1891.

Falkowski, M., K. Schledzewski, B. Hansen and S. Goerdts (2003). "Expression of stabilin-2, a novel fasciclin-like hyaluronan receptor protein, in murine sinusoidal endothelia, avascular tissues, and at solid/liquid interfaces." *Histochem Cell Biol* **120**(5): 361-369.

Ferkol, T. W., E. G. Puffenberger, H. Lie, C. Helms, K. A. Strauss, A. Bowcock, J. L. Carson, M. Hazucha, D. H. Morton, A. C. Patel, M. W. Leigh, M. R. Knowles and M. A. Zariwala (2013). "Primary ciliary dyskinesia-causing mutations in Amish and Mennonite communities." *J Pediatr* **163**(2): 383-387.

Force, T., R. O. Bonow, S. R. Houser, R. J. Solaro, R. E. Hershberger, B. Adhikari, M. E. Anderson, R. Boineau, B. J. Byrne, T. P. Cappola, R. Kalluri, M. M. LeWinter, M. S. Maron, J. D. Molkentin, S. R. Ommen, M. Regnier, W. H. Tang, R. Tian, M. A. Konstam, B. J. Maron and C. E. Seidman (2010). "Research priorities in hypertrophic cardiomyopathy: report of a Working Group of the National Heart, Lung, and Blood Institute." *Circulation* **122**(11): 1130-1133.

French, V. M., I. M. van de Laar, M. W. Wessels, C. Rohe, J. W. Roos-Hesselink, G. Wang, I. M. Frohn-Mulder, L. A. Severijnen, B. M. de Graaf, R. Schot, G. Breedveld, E. Mientjes, M. van Tienhoven, E. Jadot, Z. Jiang, A. Verkerk, S. Swagemakers, H. Venselaar, Z. Rahimi, H. Najmabadi, H. Meijers-Heijboer, E. de Graaff, W. A. Helbing, R. Willemsen, K. Devriendt, J. W. Belmont, B. A. Oostra, J. D. Amack and A. M. Bertoli-Avella (2012). "NPHP4 variants are associated with pleiotropic heart malformations." *Circ Res* **110**(12): 1564-1574.

Frosk, P., C. R. Greenberg, A. A. Tennese, R. Lamont, E. Nylen, C. Hirst, D. Frappier, N. M. Roslin, M. Zaik, K. Bushby, V. Straub, M. Zatz, F. de Paula, K. Morgan, T. M. Fujiwara and K. Wrogemann (2005). "The most common mutation in FKRP causing limb girdle muscular dystrophy type 2I (LGMD2I) may have occurred only once and is present in Hutterites and other populations." *Hum Mutat* **25**(1): 38-44.

Garg, V., I. S. Kathiriyar, R. Barnes, M. K. Schluterman, I. N. King, C. A. Butler, C. R. Rothrock, R. S. Eapen, K. Hirayama-Yamada, K. Joo, R. Matsuoka, J. C. Cohen and D. Srivastava (2003). "GATA4 mutations cause human congenital heart defects and reveal an interaction with TBX5." *Nature* **424**(6947): 443-447.

Garg, V., A. N. Muth, J. F. Ransom, M. K. Schluterman, R. Barnes, I. N. King, P. D. Grossfeld and D. Srivastava (2005). "Mutations in NOTCH1 cause aortic valve disease." *Nature* **437**(7056): 270-274.

Gatrad, A. R., A. P. Read and G. H. Watson (1984). "Consanguinity and complex cardiac anomalies with situs ambiguus." *Arch Dis Child* **59**(3): 242-245.

Gaur, U., K. Li, S. Mei and G. Liu (2013). "Research progress in allele-specific expression and its regulatory mechanisms." *J Appl Genet* **54**(3): 271-283.

Gebbia, M., G. B. Ferrero, G. Pilia, M. T. Bassi, A. Aylsworth, M. Penman-Splitt, L. M. Bird, J. S. Bamforth, J. Burn, D. Schlessinger, D. L. Nelson and B. Casey (1997). "X-linked situs abnormalities result from mutations in ZIC3." *Nat Genet* **17**(3): 305-308.

Gedicke-Hornung, C., V. Behrens-Gawlik, S. Reischmann, B. Geertz, D. Stimpel, F. Weinberger, S. Schlossarek, G. Precigout, I. Braren, T. Eschenhagen, G. Mearini, S. Lorain, T. Voit, P. A. Dreyfus, L. Garcia and L. Carrier (2013). "Rescue of cardiomyopathy through U7snRNA-mediated exon skipping in Mybpc3-targeted knock-in mice." *EMBO Mol Med* **5**(7): 1128-1145.

Geis, N., B. Seto, L. Bartoshesky, M. B. Lewis and H. M. Pashayan (1981). "The prevalence of congenital heart disease among the population of a metropolitan cleft lip and palate clinic." Cleft Palate J **18**(1): 19-23.

Geisterfer-Lowrance, A. A., S. Kass, G. Tanigawa, H. P. Vosberg, W. McKenna, C. E. Seidman and J. G. Seidman (1990). "A molecular basis for familial hypertrophic cardiomyopathy: a beta cardiac myosin heavy chain gene missense mutation." Cell **62**(5): 999-1006.

Gelb, B. D. (2015). "History of Our Understanding of the Causes of Congenital Heart Disease." Circ Cardiovasc Genet **8**(3): 529-536.

Genomes Project, C., A. Auton, L. D. Brooks, R. M. Durbin, E. P. Garrison, H. M. Kang, J. O. Korbel, J. L. Marchini, S. McCarthy, G. A. McVean and G. R. Abecasis (2015). "A global reference for human genetic variation." Nature **526**(7571): 68-74.

Gharagozloo, F., B. H. Bulkley and G. M. Hutchins (1977). "A proposed pathogenesis of cor triatriatum: impingement of the left superior vena cava on the developing left atrium." Am Heart J **94**(5): 618-626.

Gillum, D. R. and B. A. Staffileno (2011). "An integrative review of the current knowledge of cardiovascular disease and associated risk factors in the Old Order Amish." J Transcult Nurs **22**(2): 182-190.

Gillum, D. R., B. A. Staffileno, K. S. Schwartz, L. Coke, L. Fogg and D. Reiling (2011). "Cardiovascular disease in the Amish: an exploratory study of knowledge, beliefs, and health care practices." Holist Nurs Pract **25**(6): 289-297.

Glazier, A. A., A. Thompson and S. M. Day (2018). "Allelic imbalance and haploinsufficiency in MYBPC3-linked hypertrophic cardiomyopathy." Pflugers Arch.

Goudy, S., A. Law, G. Sanchez, H. S. Baldwin and C. Brown (2010). "Tbx1 is necessary for palatal elongation and elevation." Mech Dev **127**(5-6): 292-300.

Goyal, S. K., S. R. Punnam, G. Verma and F. L. Ruberg (2008). "Persistent left superior vena cava: a case report and review of literature." Cardiovasc Ultrasound **6**: 50.

Gritli-Linde, A. (2007). "Molecular control of secondary palate development." Dev Biol **301**(2): 309-326.

Gushulak, L., R. Hemming, D. Martin, V. Seyrantepe, A. Pshezhetsky and B. Triggs-Raine (2012). "Hyaluronidase 1 and beta-hexosaminidase have redundant functions in hyaluronan and chondroitin sulfate degradation." J Biol Chem **287**(20): 16689-16697.

Haider, A. W., M. G. Larson, E. J. Benjamin and D. Levy (1998). "Increased left ventricular mass and hypertrophy are associated with increased risk for sudden death." J Am Coll Cardiol **32**(5): 1454-1459.

Han, P., J. Bloomekatz, J. Ren, R. Zhang, J. D. Grinstein, L. Zhao, C. G. Burns, C. E. Burns, R. M. Anderson and N. C. Chi (2016). "Coordinating cardiomyocyte interactions to direct ventricular chamber morphogenesis." Nature **534**(7609): 700-704.

Hartman, R. J., S. A. Rasmussen, L. D. Botto, T. Riehle-Colarusso, C. L. Martin, J. D. Cragan, M. Shin and A. Correa (2011). "The contribution of chromosomal abnormalities to congenital heart defects: a population-based study." Pediatr Cardiol **32**(8): 1147-1157.

Harvey, R. P. (2002). "Patterning the vertebrate heart." Nat Rev Genet **3**(7): 544-556.

Helms, A. S., F. M. Davis, D. Coleman, S. N. Bartolone, A. A. Glazier, F. Pagani, J. M. Yob, S. Sadayappan, E. Pedersen, R. Lyons, M. V. Westfall, R. Jones, M. W. Russell and S. M. Day (2014). "Sarcomere mutation-specific expression patterns in human hypertrophic cardiomyopathy." Circ Cardiovasc Genet **7**(4): 434-443.

Hemming, R., D. C. Martin, E. Slominski, J. I. Nagy, A. J. Halayko, S. Pind and B. Triggs-Raine (2008). "Mouse Hyal3 encodes a 45- to 56-kDa glycoprotein whose overexpression increases hyaluronidase 1 activity in cultured cells." Glycobiology **18**(4): 280-289.

Herman, D. S., L. Lam, M. R. Taylor, L. Wang, P. Teekakirikul, D. Christodoulou, L. Conner, S. R. DePalma, B. McDonough, E. Sparks, D. L. Teodorescu, A. L. Cirino, N. R. Banner, D. J. Pennell, S. Graw, M. Merlo, A. Di Lenarda, G. Sinagra, J. M. Bos, M. J. Ackerman, R. N. Mitchell, C. E. Murry, N. K. Lakdawala, C. Y. Ho, P. J. Barton, S. A. Cook, L. Mestroni, J. G. Seidman and C. E. Seidman (2012). "Truncations of titin causing dilated cardiomyopathy." N Engl J Med **366**(7): 619-628.

Hishiya, A., T. Kitazawa and S. Takayama (2010). "BAG3 and Hsc70 interact with actin capping protein CapZ to maintain myofibrillar integrity under mechanical stress." Circ Res **107**(10): 1220-1231.

Hochman, J., B. W. Blakley, M. Wellman and L. Blakley (2006). "Prevention of aminoglycoside-induced sensorineural hearing loss." J Otolaryngol **35**(3): 153-156.

Homsy, J., S. Zaidi, Y. Shen, J. S. Ware, K. E. Samocha, K. J. Karczewski, S. R. DePalma, D. McKean, H. Wakimoto, J. Gorham, S. C. Jin, J. Deanfield, A. Giardini, G. A. Porter, Jr., R. Kim, K. Bilguvar, F. Lopez-Giraldez, I. Tikhonova, S. Mane, A. Romano-Adesman, H. Qi, B. Vardarajan, L. Ma, M. Daly, A. E. Roberts, M. W. Russell, S. Mital, J. W. Newburger, J. W. Gaynor, R. E. Breitbart, I. Iossifov, M. Ronemus, S. J. Sanders, J. R. Kaltman, J. G. Seidman, M. Brueckner, B. D. Gelb, E. Goldmuntz, R. P. Lifton, C. E. Seidman and W. K. Chung (2015). "De novo mutations in congenital heart disease with neurodevelopmental and other congenital anomalies." Science **350**(6265): 1262-1266.

Horowitz, M. D., W. Zager, M. Bilsker, R. A. Perryman and M. H. Lowery (1993). "Cor triatriatum in adults." Am Heart J **126**(2): 472-474.

Hsu, L. J., L. Schultz, Q. Hong, K. Van Moer, J. Heath, M. Y. Li, F. J. Lai, S. R. Lin, M. H. Lee, C. P. Lo, Y. S. Lin, S. T. Chen and N. S. Chang (2009). "Transforming growth factor beta1 signaling via interaction with cell surface Hyal-2 and recruitment of WWOX/WOX1." J Biol Chem **284**(23): 16049-16059.

Hsueh, W. C., B. D. Mitchell, J. L. Schneider, M. J. Wagner, C. J. Bell, E. Nanthakumar and A. R. Shuldiner (2000). "QTL influencing blood pressure maps to the region of PPH1 on chromosome 2q31-34 in Old Order Amish." Circulation **101**(24): 2810-2816.

Humpl, T., K. Reineker, C. Manlhiot, A. I. Dipchand, J. G. Coles and B. W. McCrindle (2010). "Cor triatriatum sinistrum in childhood. A single institution's experience." Can J Cardiol **26**(7): 371-376.

Imundo, L., C. A. Leduc, S. Guha, M. Brown, G. Perino, L. Gushulak, B. Triggs-Raine and W. K. Chung (2011). "A complete deficiency of Hyaluronoglucosaminidase 1 (HYAL1) presenting as familial juvenile idiopathic arthritis." J Inherit Metab Dis **34**(5): 1013-1022.

Jaaskelainen, P., T. Helio, K. Aalto-Setälä, M. Kaartinen, E. Ilveskoski, L. Hamalainen, J. Melin, M. S. Nieminen, M. Laakso, J. Kuusisto, H. C. M. s. g. Fin, H. Kervinen, J. Mustonen, J.

Juvonen, M. Niemi, P. Uusimaa, M. Huttunen, M. Kotila and M. Pietila (2013). "Two founder mutations in the alpha-tropomyosin and the cardiac myosin-binding protein C genes are common causes of hypertrophic cardiomyopathy in the Finnish population." Ann Med **45**(1): 85-90.

Jacobs, E. G., M. P. Leung and J. Karlberg (2000). "Distribution of symptomatic congenital heart disease in Hong Kong." Pediatr Cardiol **21**(2): 148-157.

Jacques, A., A. C. Hoskins, J. C. Kentish and S. B. Marston (2008). "From genotype to phenotype: a longitudinal study of a patient with hypertrophic cardiomyopathy due to a mutation in the MYBPC3 gene." J Muscle Res Cell Motil **29**(6-8): 239-246.

Jadin, L., L. H. Bookbinder and G. I. Frost (2012). "A comprehensive model of hyaluronan turnover in the mouse." Matrix Biol **31**(2): 81-89.

Jadin, L., X. Wu, H. Ding, G. I. Frost, C. Onclinx, B. Triggs-Raine and B. Flamion (2008). "Skeletal and hematological anomalies in HYAL2-deficient mice: a second type of mucopolysaccharidosis IX?" FASEB J **22**(12): 4316-4326.

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

Jegier, W., J. E. Gibbons and F. W. Wigglesworth (1963). "Cortriatriatum: clinical, hemodynamic and pathological studies surgical correction in early life." Pediatrics **31**: 255-267.

Jerome, L. A. and V. E. Papaioannou (2001). "DiGeorge syndrome phenotype in mice mutant for the T-box gene, Tbx1." Nat Genet **27**(3): 286-291.

Jiang, P., X. Li, C. B. Thompson, Z. Huang, F. Araiza, R. Osgood, G. Wei, M. Feldmann, G. I. Frost and H. M. Shepard (2012). "Effective targeting of the tumor microenvironment for cancer therapy." Anticancer Res **32**(4): 1203-1212.

Jiang, R., J. O. Bush and A. C. Lidral (2006). "Development of the upper lip: morphogenetic and molecular mechanisms." Dev Dyn **235**(5): 1152-1166.

Kassem, H., R. S. Azer, M. Saber-Ayad, S. Moharem-Elgamal, G. Magdy, A. Elguindy, F. Cecchi, I. Olivotto and M. H. Yacoub (2013). "Early results of sarcomeric gene screening from the Egyptian National BA-HCM Program." J Cardiovasc Transl Res **6**(1): 65-80.

Kim, D. S., J. H. Kim, A. A. Burt, D. R. Crosslin, N. Burnham, C. E. Kim, D. M. McDonald-McGinn, E. H. Zackai, S. C. Nicolson, T. L. Spray, I. B. Stanaway, D. A. Nickerson, P. J. Heagerty, H. Hakonarson, J. W. Gaynor and G. P. Jarvik (2016). "Burden of potentially pathologic copy number variants is higher in children with isolated congenital heart disease and significantly impairs covariate-adjusted transplant-free survival." J Thorac Cardiovasc Surg **151**(4): 1147-1151 e1144.

Kloesel, B., J. A. DiNardo and S. C. Body (2016). "Cardiac Embryology and Molecular Mechanisms of Congenital Heart Disease: A Primer for Anesthesiologists." Anesth Analg **123**(3): 551-569.

Kofflard, M. J., F. J. Ten Cate, C. van der Lee and R. T. van Domburg (2003). "Hypertrophic cardiomyopathy in a large community-based population: clinical outcome and identification

of risk factors for sudden cardiac death and clinical deterioration." *J Am Coll Cardiol* **41**(6): 987-993.

Kolder, I. C., M. Michels, I. Christiaans, F. J. Ten Cate, D. Majoer-Krakauer, A. H. Danser, R. H. Lekanne Deprez, M. Tanck, A. A. Wilde, C. R. Bezzina and D. Dooijes (2012). "The role of renin-angiotensin-aldosterone system polymorphisms in phenotypic expression of MYBPC3-related hypertrophic cardiomyopathy." *Eur J Hum Genet* **20**(10): 1071-1077.

Konno, T., K. Hayashi, N. Fujino, R. Oka, A. Nomura, Y. Nagata, A. Hodatsu, K. Sakata, H. Furusho, M. Takamura, H. Nakamura, M. A. Kawashiri and M. Yamagishi (2015). "Electrocardiographic QRS Fragmentation as a Marker for Myocardial Fibrosis in Hypertrophic Cardiomyopathy." *J Cardiovasc Electrophysiol* **26**(10): 1081-1087.

Korbel, J. O., T. Tirosh-Wagner, A. E. Urban, X. N. Chen, M. Kasowski, L. Dai, F. Grubert, C. Erdman, M. C. Gao, K. Lange, E. M. Sobel, G. M. Barlow, A. S. Aylsworth, N. J. Carpenter, R. D. Clark, M. Y. Cohen, E. Doran, T. Falik-Zaccai, S. O. Lewin, I. T. Lott, B. C. McGillivray, J. B. Moeschler, M. J. Pettenati, S. M. Poeschel, K. W. Rao, L. G. Shaffer, M. Shohat, A. J. Van Riper, D. Warburton, S. Weissman, M. B. Gerstein, M. Snyder and J. R. Korenberg (2009). "The genetic architecture of Down syndrome phenotypes revealed by high-resolution analysis of human segmental trisomies." *Proc Natl Acad Sci U S A* **106**(29): 12031-12036.

Kubo, T., H. Kitaoka, M. Okawa, Y. Matsumura, N. Hitomi, N. Yamasaki, T. Furuno, J. Takata, M. Nishinaga, A. Kimura and Y. L. Doi (2005). "Lifelong left ventricular remodeling of hypertrophic cardiomyopathy caused by a founder frameshift deletion mutation in the cardiac Myosin-binding protein C gene among Japanese." *J Am Coll Cardiol* **46**(9): 1737-1743.

Kumar, A., B. Rani, R. Sharma, G. Kaur, R. Prasad, A. Bahl and M. Khullar (2018). "ACE2, CALM3 and TNNI3K polymorphisms as potential disease modifiers in hypertrophic and dilated cardiomyopathies." *Mol Cell Biochem* **438**(1-2): 167-174.

Lagendijk, A. K., A. Szabó, R. M. Merks and J. Bakkers (2013). "Hyaluronan: a critical regulator of endothelial-to-mesenchymal transition during cardiac valve formation." *Trends Cardiovasc Med* **23**(5): 135-142.

Landrum, M. J., J. M. Lee, M. Benson, G. Brown, C. Chao, S. Chitipiralla, B. Gu, J. Hart, D. Hoffman, J. Hoover, W. Jang, K. Katz, M. Ovetsky, G. Riley, A. Sethi, R. Tully, R. Villamarin-Salomon, W. Rubinstein and D. R. Maglott (2016). "ClinVar: public archive of interpretations of clinically relevant variants." *Nucleic Acids Res* **44**(D1): D862-868.

Landstrom, A. P. and M. J. Ackerman (2010). "Mutation type is not clinically useful in predicting prognosis in hypertrophic cardiomyopathy." *Circulation* **122**(23): 2441-2449; discussion 2450.

Larkin, M. A., G. Blackshields, N. P. Brown, R. Chenna, P. A. McGettigan, H. McWilliam, F. Valentin, I. M. Wallace, A. Wilm, R. Lopez, J. D. Thompson, T. J. Gibson and D. G. Higgins (2007). "Clustal W and Clustal X version 2.0." *Bioinformatics* **23**(21): 2947-2948.

Leirgul, E., T. Fomina, K. Brodwall, G. Greve, H. Holmstrom, S. E. Vollset, G. S. Tell and N. Oyen (2014). "Birth prevalence of congenital heart defects in Norway 1994-2009--a nationwide study." *Am Heart J* **168**(6): 956-964.

Lepperdinger, G., J. Mullegger and G. Kreil (2001). "Hyal2--less active, but more versatile?" *Matrix Biol* **20**(8): 509-514.

Lescroart, F., R. G. Kelly, J. F. Le Garrec, J. F. Nicolas, S. M. Meilhac and M. Buckingham (2010). "Clonal analysis reveals common lineage relationships between head muscles and second heart field derivatives in the mouse embryo." Development **137**(19): 3269-3279.

Leslie, E. J. and M. L. Marazita (2013). "Genetics of cleft lip and cleft palate." Am J Med Genet C Semin Med Genet **163C**(4): 246-258.

Li, H. and R. Durbin (2010). "Fast and accurate long-read alignment with Burrows-Wheeler transform." Bioinformatics **26**(5): 589-595.

Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin and S. Genome Project Data Processing (2009). "The Sequence Alignment/Map format and SAMtools." Bioinformatics **25**(16): 2078-2079.

Lim, D. S., S. Lutucuta, P. Bachireddy, K. Youker, A. Evans, M. Entman, R. Roberts and A. J. Marian (2001). "Angiotensin II blockade reverses myocardial fibrosis in a transgenic mouse model of human hypertrophic cardiomyopathy." Circulation **103**(6): 789-791.

Lopes, L. R., P. Syrris, O. P. Guttman, C. O'Mahony, H. C. Tang, C. Dalageorgou, S. Jenkins, M. Hubank, L. Monserrat, W. J. McKenna, V. Plagnol and P. M. Elliott (2015). "Novel genotype-phenotype associations demonstrated by high-throughput sequencing in patients with hypertrophic cardiomyopathy." Heart **101**(4): 294-301.

Luijsterburg, A. J., A. M. Rozendaal and C. Vermeij-Keers (2014). "Classifying common oral clefts: a new approach after descriptive registration." Cleft Palate Craniofac J **51**(4): 381-391.

Mackeprang, M. and S. Hay (1971). "Observations on congenital heart disease in a mortality study of children with cleft lip and palate." J Chronic Dis **24**(1): 39-43.

Mahuran, D. J. (1999). "Biochemical consequences of mutations causing the GM2 gangliosidosis." Biochim Biophys Acta **1455**(2-3): 105-138.

Malhotra, A., P. Buttrick and J. Scheuer (1990). "Effects of sex hormones on development of physiological and pathological cardiac hypertrophy in male and female rats." Am J Physiol **259**(3 Pt 2): H866-871.

Marelli, A. J., R. Ionescu-Ittu, A. S. Mackie, L. Guo, N. Dendukuri and M. Kaouache (2014). "Lifetime prevalence of congenital heart disease in the general population from 2000 to 2010." Circulation **130**(9): 749-756.

Marian, A. J. (2002). "Modifier genes for hypertrophic cardiomyopathy." Curr Opin Cardiol **17**(3): 242-252.

Marian, A. J. (2010). "Hypertrophic cardiomyopathy: from genetics to treatment." Eur J Clin Invest **40**(4): 360-369.

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

Markovic-Housley, Z., G. Miglierini, L. Soldatova, P. J. Rizkallah, U. Muller and T. Schirmer (2000). "Crystal structure of hyaluronidase, a major allergen of bee venom." Structure **8**(10): 1025-1035.

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

Maron, B. J. (2004). "Hypertrophic cardiomyopathy: an important global disease." Am J Med **116**(1): 63-65.

Maron, B. J., J. J. Doerer, T. S. Haas, D. M. Tierney and F. O. Mueller (2009). "Sudden deaths in young competitive athletes: analysis of 1866 deaths in the United States, 1980-2006." Circulation **119**(8): 1085-1092.

Maron, B. J., S. E. Epstein and W. C. Roberts (1986). "Causes of sudden death in competitive athletes." J Am Coll Cardiol **7**(1): 204-214.

Maron, B. J., J. M. Gardin, J. M. Flack, S. S. Gidding, T. T. Kurosaki and D. E. Bild (1995). "Prevalence of hypertrophic cardiomyopathy in a general population of young adults. Echocardiographic analysis of 4111 subjects in the CARDIA Study. Coronary Artery Risk Development in (Young) Adults." Circulation **92**(4): 785-789.

Maron, B. J., M. S. Maron and C. Semsarian (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.

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

Maron, B. J., R. Mathenge, S. A. Casey, L. C. Poliac and T. F. Longe (1999). "Clinical profile of hypertrophic cardiomyopathy identified de novo in rural communities." J Am Coll Cardiol **33**(6): 1590-1595.

Maron, B. J., W. J. McKenna, G. K. Danielson, L. J. Kappenberger, H. J. Kuhn, C. E. Seidman, P. M. Shah, W. H. Spencer, 3rd, P. Spirito, F. J. Ten Cate, E. D. Wigle, C. Task Force on Clinical Expert Consensus Documents. American College of and C. Committee for Practice Guidelines. European Society of (2003). "American College of Cardiology/European Society of Cardiology clinical expert consensus document on hypertrophic cardiomyopathy. A report of the American College of Cardiology Foundation Task Force on Clinical Expert Consensus Documents and the European Society of Cardiology Committee for Practice Guidelines." J Am Coll Cardiol **42**(9): 1687-1713.

Maron, B. J., I. Olivetto, P. Bellone, M. R. Conte, F. Cecchi, B. P. Flygenring, S. A. Casey, T. E. Gohman, S. Bongioanni and P. Spirito (2002). "Clinical profile of stroke in 900 patients with hypertrophic cardiomyopathy." J Am Coll Cardiol **39**(2): 301-307.

Maron, M. S., I. Olivetto, A. G. Zenovich, M. S. Link, N. G. Pandian, J. T. Kuvin, S. Nistri, F. Cecchi, J. E. Udelson and B. J. Maron (2006). "Hypertrophic cardiomyopathy is predominantly a disease of left ventricular outflow tract obstruction." Circulation **114**(21): 2232-2239.

Marshall, J. D., E. G. Hinman, G. B. Collin, S. Beck, R. Cerqueira, P. Maffei, G. Milan, W. Zhang, D. I. Wilson, T. Hearn, P. Tavares, R. Vettor, C. Veronese, M. Martin, W. V. So, P. M. Nishina and J. K. Naggert (2007). "Spectrum of ALMS1 variants and evaluation of genotype-phenotype correlations in Alstrom syndrome." Hum Mutat **28**(11): 1114-1123.

Marston, S., O. Copeland, A. Jacques, K. Livesey, V. Tsang, W. J. McKenna, S. Jalilzadeh, S. Carballo, C. Redwood and H. Watkins (2009). "Evidence from human myectomy samples that MYBPC3 mutations cause hypertrophic cardiomyopathy through haploinsufficiency." Circ Res **105**(3): 219-222.

Marti-Renom, M. A., A. C. Stuart, A. Fiser, R. Sanchez, F. Melo and A. Sali (2000). "Comparative protein structure modeling of genes and genomes." *Annu Rev Biophys Biomol Struct* **29**: 291-325.

McArdle, P. F., S. Rutherford, B. D. Mitchell, C. M. Damcott, Y. Wang, V. Ramachandran, S. Ott, Y. P. Chang, D. Levy and N. Steinle (2008). "Nicotinic acetylcholine receptor subunit variants are associated with blood pressure; findings in the Old Order Amish and replication in the Framingham Heart Study." *BMC Med Genet* **9**: 67.

McElhinney, D. B., E. Geiger, J. Blinder, D. W. Benson and E. Goldmuntz (2003). "NKX2.5 mutations in patients with congenital heart disease." *J Am Coll Cardiol* **42**(9): 1650-1655.

McLeod, M. J. (1980). "Differential staining of cartilage and bone in whole mouse fetuses by alcian blue and alizarin red S." *Teratology* **22**(3): 299-301.

McNamara, J. W., A. Li, S. Lal, J. M. Bos, S. P. Harris, J. van der Velden, M. J. Ackerman, R. Cooke and C. G. Dos Remedios (2017). "MYBPC3 mutations are associated with a reduced super-relaxed state in patients with hypertrophic cardiomyopathy." *PLoS One* **12**(6): e0180064.

Mearini, G., D. Stimpel, B. Geertz, F. Weinberger, E. Kramer, S. Schlossarek, J. Mourot-Filiatre, A. Stoehr, A. Dutsch, P. J. Wijnker, I. Braren, H. A. Katus, O. J. Muller, T. Voit, T. Eschenhagen and L. Carrier (2014). "Mybpc3 gene therapy for neonatal cardiomyopathy enables long-term disease prevention in mice." *Nat Commun* **5**: 5515.

Melacini, P., C. Basso, A. Angelini, C. Calore, F. Bobbo, B. Tokajuk, N. Bellini, G. Smaniotto, M. Zucchetto, S. Iliceto, G. Thiene and B. J. Maron (2010). "Clinicopathological profiles of progressive heart failure in hypertrophic cardiomyopathy." *Eur Heart J* **31**(17): 2111-2123.

Miller, A. D. (2008). "Hyaluronidase 2 and its intriguing role as a cell-entry receptor for oncogenic sheep retroviruses." *Semin Cancer Biol* **18**(4): 296-301.

Monteiro da Rocha, A., G. Guerrero-Serna, A. Helms, C. Luzod, S. Mironov, M. Russell, J. Jalife, S. M. Day, G. D. Smith and T. J. Herron (2016). "Deficient cMyBP-C protein expression during cardiomyocyte differentiation underlies human hypertrophic cardiomyopathy cellular phenotypes in disease specific human ES cell derived cardiomyocytes." *J Mol Cell Cardiol* **99**: 197-206.

Moolman-Smook, J. C., W. J. De Lange, E. C. Bruwer, P. A. Brink and V. A. Corfield (1999). "The origins of hypertrophic cardiomyopathy-causing mutations in two South African subpopulations: a unique profile of both independent and founder events." *Am J Hum Genet* **65**(5): 1308-1320.

Moorman, A. F., V. M. Christoffels, R. H. Anderson and M. J. van den Hoff (2007). "The heart-forming fields: one or multiple?" *Philos Trans R Soc Lond B Biol Sci* **362**(1484): 1257-1265.

Mort, M., D. Ivanov, D. N. Cooper and N. A. Chuzhanova (2008). "A meta-analysis of nonsense mutations causing human genetic disease." *Hum Mutat* **29**(8): 1037-1047.

Mossey, P. A., J. Little, R. G. Munger, M. J. Dixon and W. C. Shaw (2009). "Cleft lip and palate." *Lancet* **374**(9703): 1773-1785.

Mouton, J., B. Loos, J. C. Moolman-Smook and C. J. Kinnear (2015). "Ascribing novel functions to the sarcomeric protein, myosin binding protein H (MyBPH) in cardiac sarcomere contraction." *Exp Cell Res* **331**(2): 338-351.

Nagy, N., G. Kaber, P. Y. Johnson, J. A. Gebe, A. Preisinger, B. A. Falk, V. G. Sunkari, M. D. Gooden, R. B. Vernon, M. Bogdani, H. F. Kuipers, A. J. Day, D. J. Campbell, T. N. Wight and P. L. Bollyky (2015). "Inhibition of hyaluronan synthesis restores immune tolerance during autoimmune insulinitis." *J Clin Invest* **125**(10): 3928-3940.

Natowicz, M. R., M. P. Short, Y. Wang, G. R. Dickersin, M. C. Gebhardt, D. I. Rosenthal, K. B. Sims and A. E. Rosenberg (1996). "Clinical and biochemical manifestations of hyaluronidase deficiency." *N Engl J Med* **335**(14): 1029-1033.

Nielsen, J. B., L. G. Fritsche, W. Zhou, T. M. Teslovich, O. L. Holmen, S. Gustafsson, M. E. Gabrielsen, E. M. Schmidt, R. Beaumont, B. N. Wolford, M. Lin, C. M. Brummett, M. H. Preuss, L. Refsgaard, E. P. Bottinger, S. E. Graham, I. Surakka, Y. Chu, A. H. Skogholt, H. Dalen, A. P. Boyle, H. Oral, T. J. Herron, J. Kitzman, J. Jalife, J. H. Svendsen, M. S. Olesen, I. Njolstad, M. L. Lochen, A. Baras, O. Gottesman, A. Marcketta, C. O'Dushlaine, M. D. Ritchie, T. Wilsgaard, R. J. F. Loos, T. M. Frayling, M. Boehnke, E. Ingelsson, D. J. Carey, F. E. Dewey, H. M. Kang, G. R. Abecasis, K. Hveem and C. J. Willer (2018). "Genome-wide Study of Atrial Fibrillation Identifies Seven Risk Loci and Highlights Biological Pathways and Regulatory Elements Involved in Cardiac Development." *Am J Hum Genet* **102**(1): 103-115.

Niimura, H., L. L. Bachinski, S. Sangwatanaroj, H. Watkins, A. E. Chudley, W. McKenna, A. Kristinsson, R. Roberts, M. Sole, B. J. Maron, J. G. Seidman and C. E. Seidman (1998). "Mutations in the gene for cardiac myosin-binding protein C and late-onset familial hypertrophic cardiomyopathy." *N Engl J Med* **338**(18): 1248-1257.

Niimura, H., K. K. Patton, W. J. McKenna, J. Soultis, B. J. Maron, J. G. Seidman and C. E. Seidman (2002). "Sarcomere protein gene mutations in hypertrophic cardiomyopathy of the elderly." *Circulation* **105**(4): 446-451.

Nora, J. J., P. F. Dodd, D. G. McNamara, M. A. Hattwick, R. D. Leachman and D. A. Cooley (1969). "Risk to offspring of parents with congenital heart defects." *JAMA* **209**(13): 2052-2053.

O'Connor, M. J., N. L. Rider, R. Thomas Collins, B. D. Hanna, D. Holmes Morton and K. A. Strauss (2011). "Contemporary management of congenital malformations of the heart in infants with Ellis - van Creveld syndrome: a report of nine cases." *Cardiol Young* **21**(2): 145-152.

Oliva-Sandoval, M. J., F. Ruiz-Espejo, L. Monserrat, M. Hermida-Prieto, M. Sabater, E. Garcia-Molina, M. Ortiz, M. I. Rodriguez-Garcia, L. Nunez, J. R. Gimeno, A. Castro-Beiras and M. Valdes (2010). "Insights into genotype-phenotype correlation in hypertrophic cardiomyopathy. Findings from 18 Spanish families with a single mutation in MYBPC3." *Heart* **96**(24): 1980-1984.

Olofsson, B., H. Porsch and P. Heldin (2014). "Knock-down of CD44 regulates endothelial cell differentiation via NFkappaB-mediated chemokine production." *PLoS One* **9**(3): e90921.

Orenes-Pinero, E., D. Hernandez-Romero, E. Jover, M. Valdes, G. Y. Lip and F. Marin (2011). "Impact of polymorphisms in the renin-angiotensin-aldosterone system on hypertrophic cardiomyopathy." *J Renin Angiotensin Aldosterone Syst* **12**(4): 521-530.

Oyen, N., G. Poulsen, H. A. Boyd, J. Wohlfahrt, P. K. Jensen and M. Melbye (2009). "Recurrence of congenital heart defects in families." *Circulation* **120**(4): 295-301.

Page, S. P., S. Kounas, P. Syrris, M. Christiansen, R. Frank-Hansen, P. S. Andersen, P. M. Elliott and W. J. McKenna (2012). "Cardiac myosin binding protein-C mutations in families with hypertrophic cardiomyopathy: disease expression in relation to age, gender, and long term outcome." *Circ Cardiovasc Genet* **5**(2): 156-166.

Pannbacker, M. (1968). "Congenital malformations and cleft lip and palate." *Cleft Palate J* **5**: 334-339.

Parsa, A., Y. P. Chang, R. J. Kelly, M. C. Corretti, K. A. Ryan, S. W. Robinson, S. S. Gottlieb, S. L. Kardia, A. R. Shuldiner and S. B. Liggett (2011). "Hypertrophy-associated polymorphisms ascertained in a founder cohort applied to heart failure risk and mortality." *Clin Transl Sci* **4**(1): 17-23.

Patel, H., H. Cross, C. Proukakis, R. Hershberger, P. Bork, F. D. Ciccarelli, M. A. Patton, V. A. McKusick and A. H. Crosby (2002). "SPG20 is mutated in Troyer syndrome, an hereditary spastic paraplegia." *Nat Genet* **31**(4): 347-348.

Pauws, E., A. Hoshino, L. Bentley, S. Prajapati, C. Keller, P. Hammond, J. P. Martinez-Barbera, G. E. Moore and P. Stanier (2009). "Tbx22null mice have a submucous cleft palate due to reduced palatal bone formation and also display ankyloglossia and choanal atresia phenotypes." *Hum Mol Genet* **18**(21): 4171-4179.

Penicka, M., P. Gregor, R. Kerekes, D. Marek, K. Curila, J. Krupicka, H. Candesartan use in and I. Non-obstructive Cardiomyopathy Estate Study (2009). "The effects of candesartan on left ventricular hypertrophy and function in nonobstructive hypertrophic cardiomyopathy: a pilot, randomized study." *J Mol Diagn* **11**(1): 35-41.

Perles, Z., S. Moon, A. Ta-Shma, B. Yaacov, L. Francescatto, S. Edvardson, A. J. Rein, O. Elpeleg and N. Katsanis (2015). "A human laterality disorder caused by a homozygous deleterious mutation in MMP21." *J Med Genet* **52**(12): 840-847.

Plenge, R. M., M. Seielstad, L. Padyukov, A. T. Lee, E. F. Remmers, B. Ding, A. Liew, H. Khalili, A. Chandrasekaran, L. R. Davies, W. Li, A. K. Tan, C. Bonnard, R. T. Ong, A. Thalamuthu, S. Pettersson, C. Liu, C. Tian, W. V. Chen, J. P. Carulli, E. M. Beckman, D. Altshuler, L. Alfredsson, L. A. Criswell, C. I. Amos, M. F. Seldin, D. L. Kastner, L. Klareskog and P. K. Gregersen (2007). "TRAF1-C5 as a risk locus for rheumatoid arthritis--a genome-wide study." *N Engl J Med* **357**(12): 1199-1209.

Post, W., H. Shen, C. Damcott, D. E. Arking, W. H. Kao, P. A. Sack, K. A. Ryan, A. Chakravarti, B. D. Mitchell and A. R. Shuldiner (2007). "Associations between genetic variants in the NOS1AP (CAPON) gene and cardiac repolarization in the old order Amish." *Hum Hered* **64**(4): 214-219.

Prondzynski, M., E. Kramer, S. D. Laufer, A. Shibamiya, O. Pless, F. Flenner, O. J. Muller, J. Munch, C. Redwood, A. Hansen, M. Patten, T. Eschenhagen, G. Mearini and L. Carrier (2017). "Evaluation of MYBPC3 trans-Splicing and Gene Replacement as Therapeutic Options in Human iPSC-Derived Cardiomyocytes." *Mol Ther Nucleic Acids* **7**: 475-486.

Proukakis, C., P. E. Hart, A. Cornish, T. T. Warner and A. H. Crosby (2002). "Three novel spastin (SPG4) mutations in families with autosomal dominant hereditary spastic paraplegia." *J Neurol Sci* **201**(1-2): 65-69.

Puffenberger, E. G., K. A. Strauss, K. E. Ramsey, D. W. Craig, D. A. Stephan, D. L. Robinson, C. L. Hendrickson, S. Gottlieb, D. A. Ramsay, V. M. Siu, G. G. Heuer, P. B. Crino and D. H. Morton

(2007). "Polyhydramnios, megalencephaly and symptomatic epilepsy caused by a homozygous 7-kilobase deletion in LYK5." *Brain* **130**(Pt 7): 1929-1941.

Rahimov, F., A. Jugessur and J. C. Murray (2012). "Genetics of nonsyndromic orofacial clefts." *Cleft Palate Craniofac J* **49**(1): 73-91.

Rampersaud, E., C. M. Damcott, M. Fu, H. Shen, P. McArdle, X. Shi, J. Shelton, J. Yin, Y. P. Chang, S. H. Ott, L. Zhang, Y. Zhao, B. D. Mitchell, J. O'Connell and A. R. Shuldiner (2007). "Identification of novel candidate genes for type 2 diabetes from a genome-wide association scan in the Old Order Amish: evidence for replication from diabetes-related quantitative traits and from independent populations." *Diabetes* **56**(12): 3053-3062.

Richman, J. M. and S. H. Lee (2003). "About face: signals and genes controlling jaw patterning and identity in vertebrates." *Bioessays* **25**(6): 554-568.

Rivas, M. A., M. Pirinen, D. F. Conrad, M. Lek, E. K. Tsang, K. J. Karczewski, J. B. Maller, K. R. Kukurba, D. S. DeLuca, M. Fromer, P. G. Ferreira, K. S. Smith, R. Zhang, F. Zhao, E. Banks, R. Poplin, D. M. Ruderfer, S. M. Purcell, T. Tukiainen, E. V. Minikel, P. D. Stenson, D. N. Cooper, K. H. Huang, T. J. Sullivan, J. Nedzel, G. T. Consortium, C. Geuvadis, C. D. Bustamante, J. B. Li, M. J. Daly, R. Guigo, P. Donnelly, K. Ardlie, M. Sammeth, E. T. Dermitzakis, M. I. McCarthy, S. B. Montgomery, T. Lappalainen and D. G. MacArthur (2015). "Human genomics. Effect of predicted protein-truncating genetic variants on the human transcriptome." *Science* **348**(6235): 666-669.

Roberts, A. M., J. S. Ware, D. S. Herman, S. Schafer, J. Baksi, A. G. Bick, R. J. Buchan, R. Walsh, S. John, S. Wilkinson, F. Mazzarotto, L. E. Felkin, S. Gong, J. A. MacArthur, F. Cunningham, J. Flannick, S. B. Gabriel, D. M. Altshuler, P. S. Macdonald, M. Heinig, A. M. Keogh, C. S. Hayward, N. R. Banner, D. J. Pennell, D. P. O'Regan, T. R. San, A. de Marvao, T. J. Dawes, A. Gulati, E. J. Birks, M. H. Yacoub, M. Radke, M. Gotthardt, J. G. Wilson, C. J. O'Donnell, S. K. Prasad, P. J. Barton, D. Fatkin, N. Hubner, J. G. Seidman, C. E. Seidman and S. A. Cook (2015). "Integrated allelic, transcriptional, and phenomic dissection of the cardiac effects of titin truncations in health and disease." *Sci Transl Med* **7**(270): 270ra276.

Robinson, K., M. P. Frenneaux, B. Stockins, G. Karatasakis, J. D. Poloniecki and W. J. McKenna (1990). "Atrial fibrillation in hypertrophic cardiomyopathy: a longitudinal study." *J Am Coll Cardiol* **15**(6): 1279-1285.

Rochais, F., K. Mesbah and R. G. Kelly (2009). "Signaling pathways controlling second heart field development." *Circ Res* **104**(8): 933-942.

Rodgers, L. S., S. Lalani, K. M. Hardy, X. Xiang, D. Broka, P. B. Antin and T. D. Camenisch (2006). "Depolymerized hyaluronan induces vascular endothelial growth factor, a negative regulator of developmental epithelial-to-mesenchymal transformation." *Circ Res* **99**(6): 583-589.

Rodriguez-Garcia, M. I., L. Monserrat, M. Ortiz, X. Fernandez, L. Cazon, L. Nunez, R. Barriales-Villa, E. Maneiro, E. Veira, A. Castro-Beiras and M. Hermida-Prieto (2010). "Novel human pathological mutations. Gene symbol: MYBPC3. Disease: cardiomyopathy, hypertrophic." *Hum Genet* **127**(4): 484.

Rogozin, I. B. and L. Milanesi (1997). "Analysis of donor splice sites in different eukaryotic organisms." *J Mol Evol* **45**(1): 50-59.

Romhilt, D. W. and E. H. Estes, Jr. (1968). "A point-score system for the ECG diagnosis of left ventricular hypertrophy." *Am Heart J* **75**(6): 752-758.

Rosenthal, N. (1987). "Identification of regulatory elements of cloned genes with functional assays." *Methods Enzymol* **152**: 704-720.

Sabater-Molina, M., D. Saura, E. Garcia-Molina Saez, J. Gonzalez-Carrillo, L. Polo, I. Perez-Sanchez, M. D. C. Olmo, M. J. Oliva-Sandoval, R. Barriales-Villa, P. Carbonell, D. Pascual-Figal and J. R. Gimeno (2017). "A Novel Founder Mutation in MYBPC3: Phenotypic Comparison With the Most Prevalent MYBPC3 Mutation in Spain." *Rev Esp Cardiol (Engl Ed)* **70**(2): 105-114.

Schlossarek, S., F. Schuermann, B. Geertz, G. Mearini, T. Eschenhagen and L. Carrier (2012). "Adrenergic stress reveals septal hypertrophy and proteasome impairment in heterozygous Mybpc3-targeted knock-in mice." *J Muscle Res Cell Motil* **33**(1): 5-15.

Schott, J. J., D. W. Benson, C. T. Basson, W. Pease, G. M. Silberbach, J. P. Moak, B. J. Maron, C. E. Seidman and J. G. Seidman (1998). "Congenital heart disease caused by mutations in the transcription factor NKX2-5." *Science* **281**(5373): 108-111.

Shah, C. V., S. Pruzansky and W. S. Harris (1970). "Cardiac malformations with facial clefts; with observations on the Pierre Robin syndrome." *Am J Dis Child* **119**(3): 238-244.

Shaheen, R., N. Patel, H. Shamseldin, F. Alzahrani, R. Al-Yamany, A. L. A. N. Ewida, S. Anazi, M. Alnemer, M. Elsheikh, K. Alfaleh, M. Alshammari, A. Alhashem, A. A. Alangari, M. A. Salih, M. Kircher, R. M. Daza, N. Ibrahim, S. M. Wakil, A. Alaqeel, I. Altowaijri, J. Shendure, A. Al-Habib, E. Faqieh and F. S. Alkuraya (2016). "Accelerating matchmaking of novel dysmorphology syndromes through clinical and genomic characterization of a large cohort." *Genet Med* **18**(7): 686-695.

Shamblott, M. J., E. M. Bugg, A. M. Lawler and J. D. Gearhart (2002). "Craniofacial abnormalities resulting from targeted disruption of the murine Sim2 gene." *Dev Dyn* **224**(4): 373-380.

Sharma, S., J. A. Drezner, A. Baggish, M. Papadakis, M. G. Wilson, J. M. Prutkin, A. La Gerche, M. J. Ackerman, M. Borjesson, J. C. Salerno, I. M. Asif, D. S. Owens, E. H. Chung, M. S. Emery, V. F. Froelicher, H. Heidebuchel, C. Adamuz, C. A. Asplund, G. Cohen, K. G. Harmon, J. C. Marek, S. Molossi, J. Niebauer, H. F. Pelto, M. V. Perez, N. R. Riding, T. Saarel, C. M. Schmied, D. M. Shipon, R. Stein, V. L. Vetter, A. Pelliccia and D. Corrado (2018). "International recommendations for electrocardiographic interpretation in athletes." *Eur Heart J* **39**(16): 1466-1480.

Shen, H., C. M. Damcott, E. Rampersaud, T. I. Pollin, R. B. Horenstein, P. F. McArdle, P. A. Peyser, L. F. Bielak, W. S. Post, Y. P. Chang, K. A. Ryan, M. Miller, J. A. Rumberger, P. F. Sheedy, 2nd, J. Shelton, J. R. O'Connell, A. R. Shuldiner and B. D. Mitchell (2010). "Familial defective apolipoprotein B-100 and increased low-density lipoprotein cholesterol and coronary artery calcification in the old order amish." *Arch Intern Med* **170**(20): 1850-1855.

Shieh, J. T., A. H. Bittles and L. Hudgins (2012). "Consanguinity and the risk of congenital heart disease." *Am J Med Genet A* **158A**(5): 1236-1241.

Siekmann, A. F. and N. D. Lawson (2007). "Notch signalling limits angiogenic cell behaviour in developing zebrafish arteries." *Nature* **445**(7129): 781-784.

Simpson, M. A., H. E. Cross, L. Cross, M. Helmuth and A. H. Crosby (2009). "Lethal cystic kidney disease in Amish neonates associated with homozygous nonsense mutation of NPHP3." Am J Kidney Dis **53**(5): 790-795.

Sizarov, A., J. Ya, B. A. de Boer, W. H. Lamers, V. M. Christoffels and A. F. Moorman (2011). "Formation of the building plan of the human heart: morphogenesis, growth, and differentiation." Circulation **123**(10): 1125-1135.

Sobel, E., H. Sengul and D. E. Weeks (2001). "Multipoint estimation of identity-by-descent probabilities at arbitrary positions among marker loci on general pedigrees." Hum Hered **52**(3): 121-131.

Spirito, P., P. Bellone, K. M. Harris, P. Bernabo, P. Bruzzi and B. J. Maron (2000). "Magnitude of left ventricular hypertrophy and risk of sudden death in hypertrophic cardiomyopathy." N Engl J Med **342**(24): 1778-1785.

Stern, R. (2004). "Hyaluronan catabolism: a new metabolic pathway." Eur J Cell Biol **83**(7): 317-325.

Stone, D. L., R. Agarwala, A. A. Schaffer, J. L. Weber, D. Vaske, T. Oda, S. C. Chandrasekharappa, C. A. Francomano and L. G. Biesecker (1998). "Genetic and physical mapping of the McKusick-Kaufman syndrome." Hum Mol Genet **7**(3): 475-481.

Strande, J. L. (2015). "Haploinsufficiency MYBPC3 mutations: another stress induced cardiomyopathy? Let's take a look!" J Mol Cell Cardiol **79**: 284-286.

Strauss, K. A., L. DuBiner, M. Simon, M. Zaragoza, P. P. Sengupta, P. Li, N. Narula, S. Dreike, J. Platt, V. Procaccio, X. R. Ortiz-Gonzalez, E. G. Puffenberger, R. I. Kelley, D. H. Morton, J. Narula and D. C. Wallace (2013). "Severity of cardiomyopathy associated with adenine nucleotide translocator-1 deficiency correlates with mtDNA haplogroup." Proc Natl Acad Sci U S A **110**(9): 3453-3458.

Strauss, K. A. and E. G. Puffenberger (2009). "Genetics, medicine, and the Plain people." Annu Rev Genomics Hum Genet **10**: 513-536.

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

Sylva, M., M. J. van den Hoff and A. F. Moorman (2014). "Development of the human heart." Am J Med Genet A **164A**(6): 1347-1371.

Ta-Shma, A., Z. Perles, B. Yaacov, M. Werner, A. Frumkin, A. J. Rein and O. Elpeleg (2015). "A human laterality disorder associated with a homozygous WDR16 deletion." Eur J Hum Genet **23**(9): 1262-1265.

Talner, C. N. (1998). "Report of the New England Regional Infant Cardiac Program, by Donald C. Fyler, MD, Pediatrics, 1980;65(suppl):375-461." Pediatrics **102**(1 Pt 2): 258-259.

Tariq, M., J. W. Belmont, S. Lalani, T. Smolarek and S. M. Ware (2011). "SHROOM3 is a novel candidate for heterotaxy identified by whole exome sequencing." Genome Biol **12**(9): R91.

Teirlinck, C. H., F. Senni, R. E. Malti, D. Majoor-Krakauer, F. Fellmann, G. Millat, X. Andre-Fouet, F. Pernot, M. Stumpf, J. Boutarin and P. Bouvagnet (2012). "A human MYBPC3 mutation appearing about 10 centuries ago results in a hypertrophic cardiomyopathy with

delayed onset, moderate evolution but with a risk of sudden death." *BMC Med Genet* **13**: 105.

Tekin, M., B. A. Chioza, Y. Matsumoto, O. Diaz-Horta, H. E. Cross, D. Duman, H. Kokotas, H. L. Moore-Barton, K. Sakoori, M. Ota, Y. S. Odaka, J. Foster, 2nd, F. B. Cengiz, S. Tokgoz-Yilmaz, O. Tekeli, M. Grigoriadou, M. B. Petersen, A. Sreekantan-Nair, K. Gurtz, X. J. Xia, A. Pandya, M. A. Patton, J. I. Young, J. Aruga and A. H. Crosby (2013). "SLITRK6 mutations cause myopia and deafness in humans and mice." *J Clin Invest* **123**(5): 2094-2102.

Tews, I., A. Perrakis, A. Oppenheim, Z. Dauter, K. S. Wilson and C. E. Vorgias (1996). "Bacterial chitinase structure provides insight into catalytic mechanism and the basis of Tay-Sachs disease." *Nat Struct Biol* **3**(7): 638-648.

Thilenius, O. G., S. Bharati and M. Lev (1976). "Subdivided left atrium: an expanded concept of cor triatriatum sinistrum." *Am J Cardiol* **37**(5): 743-752.

Todd, J. A., N. M. Walker, J. D. Cooper, D. J. Smyth, K. Downes, V. Plagnol, R. Bailey, S. Nejentsev, S. F. Field, F. Payne, C. E. Lowe, J. S. Szeszkó, J. P. Hafler, L. Zeitels, J. H. Yang, A. Vella, S. Nutland, H. E. Stevens, H. Schuilenburg, G. Coleman, M. Maisuria, W. Meadows, L. J. Smink, B. Healy, O. S. Burren, A. A. Lam, N. R. Ovington, J. Allen, E. Adlem, H. T. Leung, C. Wallace, J. M. Howson, C. Guja, C. Ionescu-Tirgoviste, F. Genetics of Type 1 Diabetes in, M. J. Simmonds, J. M. Heward, S. C. Gough, C. Wellcome Trust Case Control, D. B. Dunger, L. S. Wicker and D. G. Clayton (2007). "Robust associations of four new chromosome regions from genome-wide analyses of type 1 diabetes." *Nat Genet* **39**(7): 857-864.

Triedman, J. K. and J. W. Newburger (2016). "Trends in Congenital Heart Disease: The Next Decade." *Circulation* **133**(25): 2716-2733.

Triggs-Raine, B., T. J. Salo, H. Zhang, B. A. Wicklow and M. R. Natowicz (1999). "Mutations in HYAL1, a member of a tandemly distributed multigene family encoding disparate hyaluronidase activities, cause a newly described lysosomal disorder, mucopolysaccharidosis IX." *Proc Natl Acad Sci U S A* **96**(11): 6296-6300.

Turcot, V., Y. Lu, H. M. Highland, C. Schurmann, A. E. Justice, R. S. Fine, J. P. Bradfield, T. Esko, A. Giri, M. Graff, X. Guo, A. E. Hendricks, T. Karaderi, A. Lempradl, A. E. Locke, A. Mahajan, E. Marouli, S. Sivapalaratnam, K. L. Young, T. Alfred, M. F. Feitosa, N. G. D. Masca, A. K. Manning, C. Medina-Gomez, P. Mudgal, M. C. Y. Ng, A. P. Reiner, S. Vedantam, S. M. Willems, T. W. Winkler, G. Abecasis, K. K. Aben, D. S. Alam, S. E. Alharthi, M. Allison, P. Amouyel, F. W. Asselbergs, P. L. Auer, B. Balkau, L. E. Bang, I. Barroso, L. Bastarache, M. Benn, S. Bergmann, L. F. Bielak, M. Bluher, M. Boehnke, H. Boeing, E. Boerwinkle, C. A. Boger, J. Bork-Jensen, M. L. Bots, E. P. Bottinger, D. W. Bowden, I. Brandslund, G. Breen, M. H. Brilliant, L. Broer, M. Brumat, A. A. Burt, A. S. Butterworth, P. T. Campbell, S. Cappellani, D. J. Carey, E. Catamo, M. J. Caulfield, J. C. Chambers, D. I. Chasman, Y. I. Chen, R. Chowdhury, C. Christensen, A. Y. Chu, M. Cocca, F. S. Collins, J. P. Cook, J. Corley, J. Corominas Galbany, A. J. Cox, D. S. Crosslin, G. Cuellar-Partida, A. D'Eustacchio, J. Danesh, G. Davies, P. I. W. Bakker, M. C. H. Groot, R. Mutsert, I. J. Deary, G. Dedoussis, E. W. Demerath, M. Heijer, A. I. Hollander, H. M. Ruijter, J. G. Dennis, J. C. Denny, E. Di Angelantonio, F. Drenos, M. Du, M. P. Dube, A. M. Dunning, D. F. Easton, T. L. Edwards, D. Ellinghaus, P. T. Ellinor, P. Elliott, E. Evangelou, A. E. Farmaki, I. S. Farooqi, J. D. Faul, S. Fauser, S. Feng, E. Ferrannini, J. Ferrieres, J. C. Florez, I. Ford, M. Fornage, O. H. Franco, A. Franke, P. W. Franks, N. Friedrich, R. Frikke-Schmidt, T. E. Galesloot, W. Gan, I. Gandin, P. Gasparini, J. Gibson, V. Giedraitis, A. P. Gjesing, P. Gordon-Larsen, M. Gorski, H. J. Grabe, S. F. A. Grant, N. Grarup, H.

L. Griffiths, M. L. Grove, V. Gudnason, S. Gustafsson, J. Haessler, H. Hakonarson, A. R. Hammerschlag, T. Hansen, K. M. Harris, T. B. Harris, A. T. Hattersley, C. T. Have, C. Hayward, L. He, N. L. Heard-Costa, A. C. Heath, I. M. Heid, O. Helgeland, J. Hernesniemi, A. W. Hewitt, O. L. Holmen, G. K. Hovingh, J. M. M. Howson, Y. Hu, P. L. Huang, J. E. Huffman, M. A. Ikram, E. Ingelsson, A. U. Jackson, J. H. Jansson, G. P. Jarvik, G. B. Jensen, Y. Jia, S. Johansson, M. E. Jorgensen, T. Jorgensen, J. W. Jukema, B. Kahali, R. S. Kahn, M. Kahonen, P. R. Kamstrup, S. Kanoni, J. Kaprio, M. Karaleftheri, S. L. R. Kardia, F. Karpe, S. Kathiresan, F. Kee, L. A. Kiemeny, E. Kim, H. Kitajima, P. Komulainen, J. S. Kooner, C. Kooperberg, T. Korhonen, P. Kovacs, H. Kuivaniemi, Z. Kutalik, K. Kuulasmaa, J. Kuusisto, M. Laakso, T. A. Lakka, D. Lamparter, E. M. Lange, L. A. Lange, C. Langenberg, E. B. Larson, N. R. Lee, T. Lehtimäki, C. E. Lewis, H. Li, J. Li, R. Li-Gao, H. Lin, K. H. Lin, L. A. Lin, X. Lin, L. Lind, J. Lindstrom, A. Linneberg, C. T. Liu, D. J. Liu, Y. Liu, K. S. Lo, A. Lophatananon, A. J. Lotery, A. Loukola, J. Luan, S. A. Lubitz, L. P. Lyytikäinen, S. Mannisto, G. Marenne, A. L. Mazul, M. I. McCarthy, R. McKean-Cowdin, S. E. Medland, K. Meidtner, L. Milani, V. Mistry, P. Mitchell, K. L. Mohlke, L. Moilanen, M. Moitry, G. W. Montgomery, D. O. Mook-Kanamori, C. Moore, T. A. Mori, A. D. Morris, A. P. Morris, M. Muller-Nurasyid, P. B. Munroe, M. A. Nalls, N. Narisu, C. P. Nelson, M. Neville, S. F. Nielsen, K. Nikus, P. R. Njolstad, B. G. Nordestgaard, D. R. Nyholt, J. R. O'Connell, M. L. O'Donoghue, L. M. Olde Loohuis, R. A. Ophoff, K. R. Owen, C. J. Packard, S. Padmanabhan, C. N. A. Palmer, N. D. Palmer, G. Pasterkamp, A. P. Patel, A. Pattie, O. Pedersen, P. L. Peissig, G. M. Peloso, C. E. Pennell, M. Perola, J. A. Perry, J. R. B. Perry, T. H. Pers, T. N. Person, A. Peters, E. R. B. Petersen, P. A. Peyser, A. Pirie, O. Polasek, T. J. Polderman, H. Puolijoki, O. T. Raitakari, A. Rasheed, R. Rauramaa, D. F. Reilly, F. Renstrom, M. Rheinberger, P. M. Ridker, J. D. Rioux, M. A. Rivas, D. J. Roberts, N. R. Robertson, A. Robino, O. Rolandsson, I. Rudan, K. S. Ruth, D. Saleheen, V. Salomaa, N. J. Samani, Y. Sapkota, N. Sattar, R. E. Schoen, P. J. Schreiner, M. B. Schulze, R. A. Scott, M. P. Segura-Lepe, S. H. Shah, W. H. Sheu, X. Sim, A. J. Slater, K. S. Small, A. V. Smith, L. Southam, T. D. Spector, E. K. Speliotes, J. M. Starr, K. Stefansson, V. Steinthorsdottir, K. E. Stirrups, K. Strauch, H. M. Stringham, M. Stumvoll, L. Sun, P. Surendran, A. J. Swift, H. Tada, K. E. Tansey, J. C. Tardif, K. D. Taylor, A. Teumer, D. J. Thompson, G. Thorleifsson, U. Thorsteinsdottir, B. H. Thuesen, A. Tonjes, G. Tromp, S. Trompet, E. Tsafantakis, J. Tuomilehto, A. Tybjaerg-Hansen, J. P. Tyrer, R. Uher, A. G. Uitterlinden, M. Uusitupa, S. W. Laan, C. M. Duijn, N. Leeuwen, J. van Setten, M. Vanhala, A. Varbo, T. V. Varga, R. Varma, D. R. Velez Edwards, S. H. Vermeulen, G. Veronesi, H. Vestergaard, V. Vitart, T. F. Vogt, U. Volker, D. Vuckovic, L. E. Wagenknecht, M. Walker, L. Wallentin, F. Wang, C. A. Wang, S. Wang, Y. Wang, E. B. Ware, N. J. Wareham, H. R. Warren, D. M. Waterworth, J. Wessel, H. D. White, C. J. Willer, J. G. Wilson, D. R. Witte, A. R. Wood, Y. Wu, H. Yaghootkar, J. Yao, P. Yao, L. M. Yerges-Armstrong, R. Young, E. Zeggini, X. Zhan, W. Zhang, J. H. Zhao, W. Zhao, W. Zhao, W. Zhou, K. T. Zondervan, J. I. Rotter, J. A. Pospisilik, F. Rivadeneira, I. B. Borecki, P. Deloukas, T. M. Frayling, G. Lettre, K. E. North, C. M. Lindgren, J. N. Hirschhorn, R. J. F. Loos, C. H. D. E. Consortium, E.-C. Consortium, B. P. C. Exome, C. Global Lipids Genetic, T. D. G. C. Go, E. I. Consortium, I. Study, C. ReproGen, T. D.-G. Consortium, M. Investigators and G. Understanding Society Scientific (2018). "Protein-altering variants associated with body mass index implicate pathways that control energy intake and expenditure in obesity." *Nat Genet* **50**(1): 26-41.

Ulbricht, A. and J. Hohfeld (2013). "Tension-induced autophagy: may the chaperone be with you." *Autophagy* **9**(6): 920-922.

van der Linde, D., E. E. Konings, M. A. Slager, M. Witsenburg, W. A. Helbing, J. J. Takkenberg and J. W. Roos-Hesselink (2011). "Birth prevalence of congenital heart disease worldwide: a systematic review and meta-analysis." *J Am Coll Cardiol* **58**(21): 2241-2247.

van Dijk, S. J., N. M. Boontje, M. W. Heymans, F. J. Ten Cate, M. Michels, C. Dos Remedios, D. Dooijes, M. A. van Slegtenhorst, J. van der Velden and G. J. Stienen (2014). "Preserved cross-bridge kinetics in human hypertrophic cardiomyopathy patients with MYBPC3 mutations." *Pflugers Arch* **466**(8): 1619-1633.

van Dijk, S. J., D. Dooijes, C. dos Remedios, M. Michels, J. M. Lamers, S. Winegrad, S. Schlossarek, L. Carrier, F. J. ten Cate, G. J. Stienen and J. van der Velden (2009). "Cardiac myosin-binding protein C mutations and hypertrophic cardiomyopathy: haploinsufficiency, deranged phosphorylation, and cardiomyocyte dysfunction." *Circulation* **119**(11): 1473-1483.

van Dijk, S. J., E. R. Paalberends, A. Najafi, M. Michels, S. Sadayappan, L. Carrier, N. M. Boontje, D. W. Kuster, M. van Slegtenhorst, D. Dooijes, C. dos Remedios, F. J. ten Cate, G. J. Stienen and J. van der Velden (2012). "Contractile dysfunction irrespective of the mutant protein in human hypertrophic cardiomyopathy with normal systolic function." *Circ Heart Fail* **5**(1): 36-46.

Van Driest, S. L., M. A. Jaeger, S. R. Ommen, M. L. Will, B. J. Gersh, A. J. Tajik and M. J. Ackerman (2004). "Comprehensive analysis of the beta-myosin heavy chain gene in 389 unrelated patients with hypertrophic cardiomyopathy." *J Am Coll Cardiol* **44**(3): 602-610.

Van Driest, S. L., S. R. Ommen, A. J. Tajik, B. J. Gersh and M. J. Ackerman (2005). "Yield of genetic testing in hypertrophic cardiomyopathy." *Mayo Clin Proc* **80**(6): 739-744.

Van Driest, S. L., V. C. Vasile, S. R. Ommen, M. L. Will, A. J. Tajik, B. J. Gersh and M. J. Ackerman (2004). "Myosin binding protein C mutations and compound heterozygosity in hypertrophic cardiomyopathy." *J Am Coll Cardiol* **44**(9): 1903-1910.

Van Praagh, R. and I. Corsini (1969). "Cor triatriatum: pathologic anatomy and a consideration of morphogenesis based on 13 postmortem cases and a study of normal development of the pulmonary vein and atrial septum in 83 human embryos." *Am Heart J* **78**(3): 379-405.

van Velzen, H. G., A. F. L. Schinkel, R. A. Oldenburg, M. A. van Slegtenhorst, I. M. E. Frohn-Mulder, J. van der Velden and M. Michels (2017). "Clinical Characteristics and Long-Term Outcome of Hypertrophic Cardiomyopathy in Individuals With a MYBPC3 (Myosin-Binding Protein C) Founder Mutation." *Circ Cardiovasc Genet* **10**(4).

Vidal, D. O., J. E. de Souza, L. C. Pires, C. Masotti, A. C. Salim, M. C. Costa, P. A. Galante, S. J. de Souza and A. A. Camargo (2011). "Analysis of allelic differential expression in the human genome using allele-specific serial analysis of gene expression tags." *Genome* **54**(2): 120-127.

Vignier, N., S. Schlossarek, B. Fraysse, G. Mearini, E. Kramer, H. Pointu, N. Mougenot, J. Guiard, R. Reimer, H. Hohenberg, K. Schwartz, M. Vernet, T. Eschenhagen and L. Carrier (2009). "Nonsense-mediated mRNA decay and ubiquitin-proteasome system regulate cardiac myosin-binding protein C mutant levels in cardiomyopathic mice." *Circ Res* **105**(3): 239-248.

Voisin, M., P. Montoya, R. Grolleau, F. Montoya, R. Dumas and R. Jean (1981). "[Labio-palatine clefts and congenital cardiopathies]." *Arch Mal Coeur Vaiss* **74**(12): 1437-1445.

Walsh, R., R. Buchan, A. Wilk, S. John, L. E. Felkin, K. L. Thomson, T. H. Chiaw, C. C. W. Loong, C. J. Pua, C. Raphael, S. Prasad, P. J. Barton, B. Funke, H. Watkins, J. S. Ware and S. A. Cook (2017). "Defining the genetic architecture of hypertrophic cardiomyopathy: re-evaluating the role of non-sarcomeric genes." *Eur Heart J* **38**(46): 3461-3468.

Walsh, R., K. L. Thomson, J. S. Ware, B. H. Funke, J. Woodley, K. J. McGuire, F. Mazzarotto, E. Blair, A. Seller, J. C. Taylor, E. V. Minikel, C. Exome Aggregation, D. G. MacArthur, M. Farrall, S. A. Cook and H. Watkins (2017). "Reassessment of Mendelian gene pathogenicity using 7,855 cardiomyopathy cases and 60,706 reference samples." *Genet Med* **19**(2): 192-203.

Wang, X., P. Li, S. Chen, L. Xi, Y. Guo, A. Guo and K. Sun (2014). "Influence of genes and the environment in familial congenital heart defects." *Mol Med Rep* **9**(2): 695-700.

Ware, J. S., J. G. Seidman and Z. Arany (2016). "Shared Genetic Predisposition in Peripartum and Dilated Cardiomyopathies." *N Engl J Med* **374**(26): 2601-2602.

Wasserman, R. L. (2014). "Overview of recombinant human hyaluronidase-facilitated subcutaneous infusion of IgG in primary immunodeficiencies." *Immunotherapy* **6**(5): 553-567.

Webster, G. and C. I. Berul (2013). "An update on channelopathies: from mechanisms to management." *Circulation* **127**(1): 126-140.

Wijnker, P. J. M., V. Sequeira, D. W. D. Kuster and J. V. Velden (2018). "Hypertrophic Cardiomyopathy: A Vicious Cycle Triggered by Sarcomere Mutations and Secondary Disease Hits." *Antioxid Redox Signal*.

Wimalasundera, R. C. and H. M. Gardiner (2004). "Congenital heart disease and aneuploidy." *Prenat Diagn* **24**(13): 1116-1122.

Wright, F. A., L. J. Strug, V. K. Doshi, C. W. Commander, S. M. Blackman, L. Sun, Y. Berthiaume, D. Cutler, A. Cojocaru, J. M. Collaco, M. Corey, R. Dorfman, K. Goddard, D. Green, J. W. Kent, Jr., E. M. Lange, S. Lee, W. Li, J. Luo, G. M. Mayhew, K. M. Naughton, R. G. Pace, P. Pare, J. M. Rommens, A. Sandford, J. R. Stonebraker, W. Sun, C. Taylor, L. L. Vanscoy, F. Zou, J. Blangero, J. Zielenski, W. K. O'Neal, M. L. Drumm, P. R. Durie, M. R. Knowles and G. R. Cutting (2011). "Genome-wide association and linkage identify modifier loci of lung disease severity in cystic fibrosis at 11p13 and 20q13.2." *Nat Genet* **43**(6): 539-546.

Wu, B. M., S. Tomatsu, S. Fukuda, K. Sukegawa, T. Orii and W. S. Sly (1994). "Overexpression rescues the mutant phenotype of L176F mutation causing beta-glucuronidase deficiency mucopolysaccharidosis in two Mennonite siblings." *J Biol Chem* **269**(38): 23681-23688.

Wyse, R. K., M. Mars, S. al-Mahdawi, I. M. Russell-Eggitt and K. D. Blake (1990). "Congenital heart anomalies in patients with clefts of the lip and/or palate." *Cleft Palate J* **27**(3): 258-264; discussion 264-255.

Wyszynski, D. F., A. F. Lewanda and T. H. Beaty (1996). "Phenotypic discordance in a family with monozygotic twins and non-syndromic cleft lip and palate." *Am J Med Genet* **66**(4): 468-470.

Xin, B., E. Puffenberger, J. Tumbush, J. R. Bockoven and H. Wang (2007). "Homozygosity for a novel splice site mutation in the cardiac myosin-binding protein C gene causes severe neonatal hypertrophic cardiomyopathy." *Am J Med Genet A* **143A**(22): 2662-2667.

Yagi, H., Y. Furutani, H. Hamada, T. Sasaki, S. Asakawa, S. Minoshima, F. Ichida, K. Joo, M. Kimura, S. Imamura, N. Kamatani, K. Momma, A. Takao, M. Nakazawa, N. Shimizu and R. Matsuoka (2003). "Role of TBX1 in human del22q11.2 syndrome." *Lancet* **362**(9393): 1366-1373.

Yang, Q., A. Sanbe, H. Osinska, T. E. Hewett, R. Klevitsky and J. Robbins (1998). "A mouse model of myosin binding protein C human familial hypertrophic cardiomyopathy." *J Clin Invest* **102**(7): 1292-1300.

Zahka, K., K. Kalidas, M. A. Simpson, H. Cross, B. B. Keller, C. Galambos, K. Gurtz, M. A. Patton and A. H. Crosby (2008). "Homozygous mutation of MYBPC3 associated with severe infantile hypertrophic cardiomyopathy at high frequency among the Amish." *Heart* **94**(10): 1326-1330.

Zaidi, S. and M. Brueckner (2017). "Genetics and Genomics of Congenital Heart Disease." *Circ Res* **120**(6): 923-940.

Zaidi, S., M. Choi, H. Wakimoto, L. Ma, J. Jiang, J. D. Overton, A. Romano-Adesman, R. D. Bjornson, R. E. Breitbart, K. K. Brown, N. J. Carriero, Y. H. Cheung, J. Deanfield, S. DePalma, K. A. Fakhro, J. Glessner, H. Hakonarson, M. J. Italia, J. R. Kaltman, J. Kaski, R. Kim, J. K. Kline, T. Lee, J. Leipzig, A. Lopez, S. M. Mane, L. E. Mitchell, J. W. Newburger, M. Parfenov, I. Pe'er, G. Porter, A. E. Roberts, R. Sachidanandam, S. J. Sanders, H. S. Seiden, M. W. State, S. Subramanian, I. R. Tikhonova, W. Wang, D. Warburton, P. S. White, I. A. Williams, H. Zhao, J. G. Seidman, M. Brueckner, W. K. Chung, B. D. Gelb, E. Goldmuntz, C. E. Seidman and R. P. Lifton (2013). "De novo mutations in histone-modifying genes in congenital heart disease." *Nature* **498**(7453): 220-223.

Zeggini, E., L. J. Scott, R. Saxena, B. F. Voight, J. L. Marchini, T. Hu, P. I. de Bakker, G. R. Abecasis, P. Almgren, G. Andersen, K. Ardlie, K. B. Bostrom, R. N. Bergman, L. L. Bonnycastle, K. Borch-Johnsen, N. P. Burtt, H. Chen, P. S. Chines, M. J. Daly, P. Deodhar, C. J. Ding, A. S. Doney, W. L. Duren, K. S. Elliott, M. R. Erdos, T. M. Frayling, R. M. Freathy, L. Gianniny, H. Grallert, N. Grarup, C. J. Groves, C. Guiducci, T. Hansen, C. Herder, G. A. Hitman, T. E. Hughes, B. Isomaa, A. U. Jackson, T. Jorgensen, A. Kong, K. Kubalanza, F. G. Kuruvilla, J. Kuusisto, C. Langenberg, H. Lango, T. Lauritzen, Y. Li, C. M. Lindgren, V. Lyssenko, A. F. Marvelle, C. Meisinger, K. Midtjell, K. L. Mohlke, M. A. Morken, A. D. Morris, N. Narisu, P. Nilsson, K. R. Owen, C. N. Palmer, F. Payne, J. R. Perry, E. Pettersen, C. Platou, I. Prokopenko, L. Qi, L. Qin, N. W. Rayner, M. Rees, J. J. Roix, A. Sandbaek, B. Shields, M. Sjogren, V. Steinthorsdottir, H. M. Stringham, A. J. Swift, G. Thorleifsson, U. Thorsteinsdottir, N. J. Timpson, T. Tuomi, J. Tuomilehto, M. Walker, R. M. Watanabe, M. N. Weedon, C. J. Willer, C. Wellcome Trust Case Control, T. Illig, K. Hveem, F. B. Hu, M. Laakso, K. Stefansson, O. Pedersen, N. J. Wareham, I. Barroso, A. T. Hattersley, F. S. Collins, L. Groop, M. I. McCarthy, M. Boehnke and D. Altshuler (2008). "Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes." *Nat Genet* **40**(5): 638-645.

Zhang, C. K., P. B. Stein, J. Liu, Z. Wang, R. Yang, J. H. Cho, P. K. Gregersen, J. M. Aerts, H. Zhao, G. M. Pastores and P. K. Mistry (2012). "Genome-wide association study of N370S homozygous Gaucher disease reveals the candidacy of CLN8 gene as a genetic modifier contributing to extreme phenotypic variation." *Am J Hematol* **87**(4): 377-383.

Zhang, Z., J. Tremblay, J. Raelson, T. Sofer, L. Du, Q. Fang, M. Argos, F. C. Marois-Blanchet, Y. Wang, L. Yan, J. Chalmers, M. Woodward, S. Harrap, P. Hamet, H. Luo and J. Wu (2019).

"EPHA4 regulates vascular smooth muscle cell contractility and is a sex-specific hypertension risk gene in individuals with type 2 diabetes." J Hypertens **37**(4): 775-789.