Targeted Re-sequencing of a Large South African Cardiomyopathy Cohort

By

Polycarp Ndibangwi

NDBPOL001

University of Cape Town

Thesis Presented for the Degree of

DOCTOR OF PHILOSOPHY

in the Department of Medicine

UNIVERSITY OF CAPE TOWN

February 2024

Supervisors:

A/Prof Gasnat Shaboodien

Prof Ntobeko Ntusi

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

University of Cape Town

Dedication

I dedicate this work to the IMHOTEP study participants and my late parents, Menaih J. and Yengeh E. Bebanghai.

Acknowledgements

This research was successful thanks to the following people and institutions. I am glad to express my sincere appreciation and gratitude to: -

- My supervisors, A/Prof. Gasnat Shaboodien and Prof. Ntobeko Ntusi, thank you for your professional and critical remarks, especially your meticulous guidance and morals.
- The Cardiovascular Genetics Group staff, Mrs Lameez Peace and Mrs Kim Brooks thank you for always being willing to assist with the logistics and other laboratory aspects.
- All participants and parents of paediatric probands for their willingness to allow their children to participate in this study.
- The Cape Heart Institute staff and director, Prof. Karen Sliwa, for being a great mentor.
- Our University of Oxford collaborators, Prof Hugh Watkins, Michael Bowman, and Dr Kate Thompsom, thank you for enabling me to investigate the genetic variants reported in this thesis.
- Our University of Manchester collaborators, Prof Bernard Keavney and Dr Simon Williams, thank you for all the bioinformatics support and pipelines designed for this study.
- UCT Directorate of Postgraduate Studies for its institutional research seminars, workshops, and training sessions to improve my scientific writing skills.
- My family, thank you for your patience and, most importantly, for encouraging me and believing in my abilities throughout my entire study. Especially Mrs Sibongiseni Ndibangwi for the unequal support she showed me throughout this study.
- The IMHOTEP team, all the PIs and staff.

Declaration

I, Polycarp Ndibangwi, student number NDBPOL001, solemnly declare that this thesis entitled "Targeted Re-sequencing of a Large South African Cardiomyopathy Cohort" is my original work. All sources used or quoted in the study have been indicated and acknowledged through complete references.

Signature: Signed by candidate Date: February 09, 2024

Updated: June 23, 2024

Table of contents

Contents

Dedication ... i
Acknowledgements ... ii
Declaration ... iii
Table of contents ... iv
Abbreviation ... ix
List of Figures ... xiii
List of Tables ... xv
Abstract ... xvi
Chapter 1: Cardiovascular disease ... 1
1.1 INTRODUCTION ... 1
1.1.1 Cardiomyopathies (CMOs) ... 2
1.1.2 Next-generation sequencing (NGS) ... 21
1.1.3 *In silico* prediction tools for cardiomyopathy ... 27
1.1.4 Treatment and management of cardiomyopathy ... 28
1.2 RATIONALE ... 29
1.3 AIM AND OBJECTIVES ... 30
1.3.1 Aim ... 30
1.3.2 Objectives ... 30
Chapter 2: Methods ... 31
2.1 STUDY DESIGN ... 31
2.1.1 Overview ... 31
2.1.2 Inclusion and exclusion criteria ... 31
2.1.3 Sample size ... 32
2.1.4 Enrolment of participants ... 32
2.2 ETHICAL CONSIDERATION ... 33
2.2.1 HREC approval ... 33
2.2.2 Informed consent ... 33
2.2.3 Data management plan (DMP) ... 33
2.2.4 Other ethical considerations ... 34
2.3 MOLECULAR METHODS, DATA GENERATION AND PROCESSING ... 35
2.3.1 Baseline data ... 35

2.3.2 Genomic deoxyribonucleic acids (gDNA) extraction 36
2.3.3 Raw gDNA quality and integrity checks 36
2.3.4 Next-generation sequencing (NGS) 37
2.3.5 Processing NGS data using *H*aloplex *A*nalysis *P*ipeline in *Py*thon (HAPPy) 39
2.3.6 *V*ariant *e*ffect *p*redictor (VEP) and analysis pipeline 40
2.3.7 Variant filtering and interpretation 40
2.3.8 Deleterious variants validation, relative screening, and segregation analysis 42
2.3.9 Genotype-phenotype correlation 44
2.4 STATISTICAL ANALYSIS 44
2.5 STUDY MEASUREMENT AND ENDPOINT 45
Chapter 3: Dilated Cardiomyopathy (DCM) 46
3.1 INTRODUCTION 46
3.1.1 Clinical background 46
3.1.2 Signs and symptoms 47
3.1.3 DCM disease burden 48
3.1.4 Aetiology of DCM 48
3.1.5 Types of DCM 49
3.1.6 DCM genotype-phenotype correlation 54
3.1.7 Therapeutics 57
3.2 METHODS 59
3.2.1 Ethical compliance 59
3.2.2 Study population 59
3.2.3 DNA extraction and quality control 59
3.2.4 Targeted sequencing and variant calling 60
3.2.5 Biostatistical analysis 61
3.3 RESULTS 62
3.3.1 Adult-onset DCM probands 62
3.3.2 Peripartum cardiomyopathy (PPCM) 101
3.3.3 Paediatric DCM probands 110
3.3.4 Possible founder class 4 and 5 variants in the DCM cohort 117
3.4 DISCUSSION 118
3.4.1 Adult-onset DCM findings 119
3.4.2 Peripartum cardiomyopathy (PPCM) findings 148
3.4.3 Paediatric DCM findings 151
3.4.4 Possible founder variants in the DCM cohort 153

3.4.5 Digenic and compound heterozygous pathogenic variants....153
3.4.6 Polygenic risk scores....155
3.5 CONCLUSION....157
Chapter 4: Hypertrophic Cardiomyopathy (HCM)....159
4.1 INTRODUCTION....159
4.1.1 Clinical background....159
4.1.2 Signs and symptoms....159
4.1.3 HCM disease burden....159
4.1.4 Aetiology of HCM....160
4.1.5 Types of HCM....162
4.1.6 Genetics of HCM....163
4.1.7 HCM genotype-phenotype correlation....163
4.1.8 Therapeutics....164
4.1.9 HCM Genetic Modifiers....165
4.2 METHODS....166
4.2.1 Ethical compliance....166
4.2.2 Study population....166
4.2.3 DNA Extraction and quality control....166
4.2.4 Targeted Sequencing and Variant Calling....167
4.2.5 Biostatistical analysis:....167
4.3 RESULTS....168
4.3.1 Baseline characteristics for the adult-onset HCM cohort....168
4.3.2 Baseline characteristics for the adult-onset HCM population groups....169
4.3.3 Targeted sequencing for adult-onset HCM probands....171
4.3.4 Paediatric HCM probands....192
4.3.5 Possible founder variants in the HCM cohort....193
4.4 DISCUSSION....194
4.4.1 Adult-onset HCM probands' baseline findings....194
4.4.2 Adult-onset HCM probands' targeted sequencing findings....195
4.4.3 Paediatric HCM....208
4.5 CONCLUSION....209
Chapter 5: Restrictive Cardiomyopathy (RCM)....210
5.1 INTRODUCTION....210
5.1.1 Clinical background....210
5.1.2 Signs and symptoms....210

5.1.3 RCM disease burden 210
5.1.4 Aetiology of RCM 211
5.1.5 Types of RCM 211
5.1.6 Genetics of RCM 213
5.1.7 RCM genotype-phenotype correlation 214
5.1.8 Therapeutics 216
5.1.9 RCM Genetic Modifiers 216
5.2 METHODS 218
5.2.1 Ethical Compliance 218
5.2.2 Study Population 218
5.2.3 DNA Extraction and Quality Control 218
5.2.4 Targeted Sequencing and Variant Calling 219
5.2.5 Biostatistical Analysis: 219
5.3 RESULTS 220
5.3.1 Baseline characteristics for the adult RCM cohort 220
5.3.2 Baseline characteristics for the adult RCM according to probands' ancestry 221
5.3.3 Targeted sequencing result for the adult RCM cohort 224
5.4 DISCUSSION 235
5.4.1 RCM probands' baseline findings 236
5.4.2 RCM probands' targeted sequencing findings 237
5.5 CONCLUSION 244
Chapter 6: Arrhythmogenic Cardiomyopathy (ACM) 245
6.1 INTRODUCTION 245
6.1.1 Clinical background 245
6.1.2 Signs and symptoms 246
6.1.3 ACM disease burden 247
6.1.4 Aetiology of ACM 247
6.1.5 Types of Arrhythmogenic Cardiomyopathy (ACM) 247
6.1.6 Genetics of ACM 249
6.1.7 ACM genotype-phenotype correlation 250
6.1.8 Therapeutics 252
6.1.9 ACM genetic modifiers 255
6.2 METHODS 257
6.2.1 Ethical Compliance: 257
6.2.2 Study Population: 257

6.2.3 DNA Extraction and Quality Control 257
6.2.4 Targeted Sequencing and Variant Calling 258
6.2.5 Biostatistical Analysis: 258
6.3 RESULTS 259
6.3.1 Baseline characteristics for the adult-onset ACM cohort 259
6.3.2 Baseline characteristics for the adult-onset ACM population groups 260
6.3.3 Targeted sequencing for adult-onset ACM probands 263
6.4 DISCUSSION 278
6.4.1 ACM probands' baseline findings 278
6.4.2 ACM probands' targeted sequencing findings 278
6.5 CONCLUSION 288
Chapter 7: A summary of the IMHOTEP study 289
7.1 Baseline characteristics for the probands in the IMHOTEP study 289
7.1.1 Adult-onset Cardiomyopathies 289
7.1.2 Paediatric Cardiomyopathies 294
7.2 Genetic findings for the IMHOTEP study 295
7.2.1 Adult-onset Cardiomyopathies 295
7.2.2 Paediatric Cardiomyopathies 300
7.3 CONCLUSION 301
Reference 302
Appendix 327
APPENDIX A- INCLUSION AND EXCLUSION CRITERIA 327
APPENDIX B 1-HREC 009/2020 328
APPENDIX B 2-HREC 009/2020 Renewal 329
APPENDIX C - Puregene Blood Core Kit C (Qiagen) 330
APPENDIX D - PAXgene® Blood DNA Kit protocol 331
APPENDIX E – Buffers and reagents: Agarose gel electrophoresis 332
APPENDIX F –The targeted panel genes list 334
APPENDIX G – VCF file annotation (GRCh37) and analysis pipeline on UCT HPC cluster 336
APPENDIX H – CVG SOP: Primer design 339
APPENDIX I- Reagents master mix 1x reaction 342
APPENDIX J- Recruitment sites 344
APPENDIX K – List of amino acids with abbreviations and classifications 345
APPENDIX L – IMHOTEP Phase I Variant Filtering SOP 348
Supplementary Tables 350

Abbreviation

1000G	One Thousand Genomes Project
ACM	Arrhythmogenic cardiomyopathy
ACMG	American College of Medical Genetics and Genomics
AD	Autosomal dominant
AFR	African (1000 Genomes Project superpopulation group)
AMR	Admixed American (1000 Genomes Project superpopulation group)
ANOVA	Analysis of variance
AR	Autosomal recessive
ARVC	Arrhythmogenic right ventricular cardiomyopathy
B	Benign
BA	Black-African ancestry
BP	Benign supporting
bp	Base pairs
CADD	Combined Annotation Dependent Depletion
CAF	Central Analytical Facilities
cDNA	Complementary DNA
CDS	coding sequences
CMO	Cardiomyopathy
CMR	Cardiac magnetic resonance imaging
COVID19	Coronavirus disease 2019
CVD	Cardiovascular disease
CVG	Cardiovascular Genetics group
dbSNP	The Single Nucleotide Polymorphism Database
DCM	Dilated cardiomyopathy
DMP	Data Management Plan
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate

EAS	East Asian (1000 Genomes Project superpopulation group)
ECG	Electrocardiography
EDTA	Ethylenediaminetetraacetic acid
ESC	European Society of Cardiology
EUR	European (1000 Genomes Project superpopulation group)
ExAC	Exome Aggregation Consortium
ExoI	Exonuclease I
ES	Exome sequencing
FH	Family history
FHS	Faculty of Health Sciences
FWA	Federalwide Assurance
GATK	Genome analysis toolkit
GS	Genome sequencing
GL-3	Globotriaosylceramide
gnomAD	Genome Aggregation Database
GRCh37	Genome Reference Consortium Human Build 37
GSH	Groote Schuur Hospital
GTex	Genotype-Tissue Expression
HAPPy	**H**aloplex **A**nalysis **P**ipeline in **Py**thon
HCM	Hypertrophic cardiomyopathy
HF	Heart failure
HGNC	Human Genome Organisation Gene Nomenclature Committee
HGVS	Human Genome Variation Society
HREC	Human Research Ethics Committee
HRM	High-resolution melt
ICTS	Information and Communication Technology Services
ID	Identifier
IA	Indian ancestry
IMHOTEP	African Cardiomyopathy and Myocarditis Registry Programme
iPSC	Induced pluripotent stem cell
IQR	Interquartile range

LB	Likely benign
LBBB	Left bundle branch block
LVEF	Left ventricular ejection fraction
LoF	loss-of-function
LMICs	Low- and middle-income countries
LP	Likely pathogenic
LVNC	Left ventricular non-compaction cardiomyopathy
MAF	Minor allele frequency
M-CAP	Mendelian Clinically Applicable Pathogenicity
$MgCl_2$	Magnesium chloride
MRI	Magnetic resonance imaging
MA	Mixed ancestry
mRNA	Messenger RNA
n	Number
NCBI	National Centre for Biotechnology Information
NCD	Non-communicable disease
NGS	Next-generation sequencing
No.	Number
NYHA	New York Heart Association
OMGL	Oxford Medical Genetics Laboratories
P	Pathogenic
PCR	Polymerase chain reaction
PM	Pathogenic moderate
PP	Pathogenic supporting
PolyPhen	Polymorphism Phenotyping
PPCM	Peripartum cardiomyopathy
PS	Pathogenic strong
PVS	Pathogenic very strong
RBBB	Right bundle branch block
RCH	Red Cross War Memorial Children’s Hospital
RCM	Restrictive cardiomyopathy

RefSeq	Reference Sequence
RNA	Ribonucleic acid
rsID	reference SNP cluster identifier
RV	Right ventricular
SAS	South-East Asian (1000 Genomes Project superpopulation group)
SCD	Sudden cardiac death
SIFT	Sorting Intolerant from Tolerant
SS	Supplementary
SSA	Sub-Saharan Africa
T_a	Annealing temperature
TM	Trademark
U	Enzyme unit
UCT	University of Cape Town
UK	The United Kingdom
USA	United States of America
VEP	Variant Effect Predictor
VT	Ventricular tachycardia
VUS	Variant of uncertain significance
WA	White ancestry
WTCHG	Wellcome Trust Centre for Human Genetics

List of Figures

Figure 1.1: Global distribution of the cardiovascular diseases. 1
Figure 1.2: The principal causes of heath failure. 3
Figure 1.3: Classification of cardiomyopathies. 5
Figure 1.4: Cardiomyopathy phenotypes. 6
Figure 1.5: An overview of the clinical investigation. 7
Figure 1.6: Cardiomyopathy classification timeline. 8
Figure 1.7: South African historical map. 12
Figure 1.8: Population substructures in South Africa. 14
Figure 1.9: Clinical summary of heritable cardiomyopathies. 15
Figure 1.10: Biomechanisms of a failing heart. 18
Figure 1.11: Common genes involved in cardiomyopathies. 19
Figure 1.12: Next-generation sequencing technologies. 22
Figure 1.13: Next-generation sequencing approaches. 23
Figure 1.14: Targeted re-sequencing pipeline. 24
Figure 2.1: An overview of the clinical application of this study. 35
Figure 2.2: Next-generation sequencing workflow. 38
Figure 2.3: Target enrichment amplicon. 39
Figure 2.4: Variant investigation strategy. 41
Figure 3.1: Global burden of cardiomyopathy. 46
Figure 3.2: The cardiac sarcomere and its components. 56
Figure 3.3: Summary of predominant phenotypic features in DCM patients. 57
Figure 3.4: Adult DCM probands family history. 63
Figure 3.5: Adult DCM probands' age, sex assigned at birth and aetiology distribution. 65
Figure 3.6: Adult DCM probands' diagnostic yield. 67
Figure 3.7: Age and sex distribution for adult-DCM class 4 and 5 variant carriers. 68
Figure 3.8: Genes found in adult-onset HCM cohort with class 4 and/or class 5 variants. 74
Figure 3.9: DCM probands' genotype result grouped according to ancestry. 77
Figure 3.10: Adult DCM probands of Black-African ancestry. 78
Figure 3.11: Adult DCM probands of Mixed ancestry. 83
Figure 3.12: Adult DCM probands of White ancestry. 88
Figure 3.13: Family segregation of *BAG3* c.925C>T variant. 89
Figure 3.14: Family segregation of *TNNT2* c.421C>T variant. 90
Figure 3.15: Family segregation of *FLNC* c.4021C>T variant. 91
Figure 3.16: Family segregation of *SCN5A* c.589G>T variant. 92
Figure 3.17: Adult DCM class 3 variant reported evidence. 93
Figure 3.18: Adult DCM class 3 variant distribution according to sex. 94
Figure 3.19: Adult DCM class 3 variant carriers of Black-African ancestry. 97
Figure 3.20: Adult DCM class 3 variant carriers of Mixed ancestry. 98
Figure 3.21: Summary of adult DCM classes 1 to 5 variants carriers. 99
Figure 3.22: PPCM probands family history. 101
Figure 3.23: PPCM probands' age, sex assigned at birth and aetiology distribution. 102
Figure 3.24: PPCM class 3 variant carriers according to ancestry. 106
Figure 3.25: PPCM class 3 variant carriers of Black-African ancestry. 107
Figure 3.26: PPCM class 3 variant carriers of Mixed ancestry. 108
Figure 3.27: Summary of PPCM classes 1 to 5 variants carriers. 109
Figure 3.28: Paediatric DCM probands' age, sex assigned at birth and aetiology distribution. 110
Figure 3.29: Paediatric DCM class 3 variant distribution according to sex. 113
Figure 3.30: Paediatric DCM class 3 variant carriers of Black-African ancestry. 115
Figure 3.31: Paediatric DCM class 3 variant carriers of Mixed ancestry. 116

Figure 3.32: Adult DCM possible founder variants. 117
Figure 3.33: Titin's class 4 and 5 variants identified in the adult DCM cohort. 127
Figure 3.34: Schematic structure of *MYH7* with *DCM-related variants' location*. 131
Figure 3.35: Schematic structure of *LMNA* with DCM-related variants' location. 133
Figure 3.36: Schematic protein structure of human RBM20 with corresponding exons. 134
Figure 3.37: Schematic structure of *BAG3* with DCM-related variants' location. 136
Figure 3.38: Schematic representation of the FLNC gene with its protein-coding domains. 138
Figure 4.1: Hypertrophic cardiomyopathy proteins: cellular locations and functions. 164
Figure 4.2: Adult HCM probands family history. 169
Figure 4.3: Adult HCM probands' age, sex assigned at birth and aetiology distribution. 170
Figure 4.4: Adult HCM probands' diagnostic yield. 172
Figure 4.5: Age and sex distribution for adult-HCM class 4 and 5 variant carriers. 174
Figure 4.6: Genes found in adult-onset HCM cohort with class 4 and/or class 5 variants. 176
Figure 4.7: Adult HCM probands of Mixed ancestry. 180
Figure 4.8: Adult HCM probands of White ancestry. 182
Figure 4.9: Adult HCM probands of Black-African ancestry. 184
Figure 4.10: Adult HCM class 3 variant reported evidence. 186
Figure 4.11: Adult HCM class 3 variant distribution according to sex. 187
Figure 4.12: Adult HCM class 3 variant carriers of Black-African ancestry. 188
Figure 4.13: Summary of adult HCM classes 1 to 5 variants carriers. 190
Figure 4.14: Schematic structure of *MYH7* with HCM-related variants' location. 199
Figure 4.15: Schematic structure of *MYBPC3* with HCM-related variants' location. 201
Figure 4.16: Schematic structure of *GLA* with HCM-related variant's location. 204
Figure 5.1: Overview of RCM proposed classification. 212
Figure 5.2: Overview of RCM-related genes. 215
Figure 5.3: RCM probands family history. 221
Figure 5.4: RCM probands' age, sex assigned at birth and aetiology distribution. 223
Figure 5.5: RCM probands' diagnostic yield. 225
Figure 5.6: Age and sex distribution for RCM class 4 and 5 variant carriers. 226
Figure 5.7: Genes found in RCM cohort with class 4 and/or class 5 variants. 227
Figure 5.8: Summary of class 4 and 5 diagnostic yields for the RCM population groups. 230
Figure 5.9: Family segregation of *DES* c.1360C>T variant. 231
Figure 5.10: Family segregation of *FLNC* c.6031G>A variant. 232
Figure 5.11: Summary of adult RCM classes 1 to 5 variants carriers. 234
Figure 5.12: Schematic representation of the *DES* gene. 240
Figure 6.1: A heart sample from an ACM patient. 246
Figure 6.2: Protein components with reported ACM-causing variants. 250
Figure 6.3: ACM genes presentation and overlap. 251
Figure 6.4: A flow chart showing the clinical treatment for patients affected by ACM. 252
Figure 6.5: Implantable cardioverter-defibrillator therapy risk pyramid for ACM. 254
Figure 6.6: ACM probands family history. 260
Figure 6.7: ACM probands' age, sex assigned at birth and aetiology distribution. 262
Figure 6.8: ACM probands' diagnostic yield. 264
Figure 6.9: Genes found in ACM cohort with class 4 and/or class 5 variants. 267
Figure 6.10: ACM probands of White ancestry. 270
Figure 6.11: ACM probands of Mixed ancestry. 272
Figure 6.12: Summary of adult ACM classes 1 to 5 variants carriers. 276
Figure 7.1: Adult-onset IMHOTEP probands' age and sex compared to a European cohort. 290
Figure 7.2: The various ancestry of adult IMHOTEP probands in this study. 291
Figure 7.3: Genetics results for 594 adult IMHOTEP probands. 295
Figure 7.4: Genetics yield for adult IMHOTEP probands per phenotype. 298
Figure 7.5: The possible founder variants in the IMHOTEP study. 299
Figure 7.6: Map showing that our yields are much lower than in most studies. 300

List of Tables

Table 1.1: Common genes associated with heritable cardiomyopathies. 20
Table 1.2: Genetic variants *in silico* prediction tools 27
Table 3.1: The most reported heritable DCM-causing genes 50
Table 3.2: Baseline results for the DCM cohort, including adults and paediatric probands 62
Table 3.3: Baseline results for the adult-onset DCM cohort according to ACMG classification 66
Table 3.4: Class 4 and 5 variants (n=51) in the adult-onset DCM probands (n=76) spanning three population groups 70
Table 3.5: The cellular location of the 14 genes with class 4 and 5 variants in the adult DCM cohort 75
Table 3.6: Class 3 variants of interest in the adult-onset DCM probands 95
Table 3.7: Baseline results for the PPCM cohort according to ACMG classification 103
Table 3.8: Class 4 and 5 variants (n=12) in the PPCM probands (n=13) spanning two population groups 105
Table 3.9: Baseline results for the paediatric DCM probands cohort 111
Table 3.10: Class 4 and 5 variants found in the paediatric DCM cohort 111
Table 3.11: Class 3 variants of interest in the paediatric DCM probands 114
Table 4.1: Prevalence of published adult HCM-related prevalent genes 163
Table 4.2: Baseline results for the HCM cohort, including adults and paediatric probands 168
Table 4.3: Baseline results for the adult-onset HCM cohort according to ACMG classification 171
Table 4.4: Class 4 and 5 variants (n=13) in the adult-onset HCM probands (n=14) spanning three population groups 173
Table 4.5: Class 3 variants of interest in the HCM probands 185
Table 4.6: Summary of combined class 4 and 5 variants in our adult HCM cohort. 191
Table 4.7: Baseline results for the paediatric HCM probands cohort 192
Table 5.1: Prevalence of published adult RCM-related genes 214
Table 5.2: Baseline results for the RCM cohort 220
Table 5.3: Baseline results for the RCM cohort according to ACMG classification 224
Table 6.1: Baseline results for the ACM cohort 259
Table 6.2: Baseline results for the ACM cohort according to ACMG classification 263
Table 6.3: Class 4 and 5 variants (n=11) in the adult-onset ACM probands (n=12) spanning three population groups 265
Table 6.4: Class 3 variants of interest in the ACM probands 273

Abstract

Introduction: Cardiomyopathy is a major cause of heart failure and transplantation globally. In sub-Saharan Africa, cardiomyopathies are ranked as the third greatest contributing cause of cardiovascular diseases and account for about 30% of adults hospitalised with heart failure. The prevalence of heart failure due to cardiomyopathies is not well established in the South African population. To address this knowledge gap, we aimed to study the genetic cause of the disease in a large cohort of South African patients with cardiomyopathy.

Design: We recruited participants from multiple centres in South Africa and Mozambique between 2015 and 2022. Cases were classified according to phenotypes by a team of experts. Using the Illumina platform, we used targeted sequencing on a panel of 38 known genes that cause primary cardiomyopathies. We used the ACMG classification to investigate class 3, 4 and 5 variants. Variants were validated using Sanger sequencing.

Result: We recruited 690 cardiomyopathy probands (594 adults and 96 paediatrics). The 594 adults include dilated cardiomyopathy (n=450), hypertrophic cardiomyopathy (n=60), restrictive cardiomyopathy (n=43) and arrhythmogenic cardiomyopathy (n=41) probands. The adult DCMs constituted 75.8% (450/594) of the IMHOTEP study. The DCM probands had a mean age of 35.6 years at diagnosis, and a 56% preponderance of females were seen; the dominant populations recruited were Black-African (58%) and Mixed ancestry (33%). We reported a diagnostic yield of 16.9% (76/450) for the DCM probands, where 68.4% of the probands had pathogenic *TTN* truncating variants. The adult HCM cohort constituted 10.1% (60/594) of the IMHOTEP study probands, with a mean age of 41.3 years at diagnosis and a male preponderance of 65%. The dominant populations recruited were 57% Mixed and 25% Black-African. The diagnostic yield for the adult HCM cohort was 23.3% (14/60), with *MYH7* (40%) and *MYBPC3* (27%) found to be the predominant genes. The adult RCM cohort constituted 7.2% (43/594) of the IMHOTEP study probands, with a mean age of 33.0 years at diagnosis and a 65% female preponderance. The dominant populations recruited were 81% Black-African and 12% Mixed ancestry. We reported a diagnostic yield of 9.3% (4/43) for the RCM probands. The adult ACM cohort constituted 6.9% (41/594) of the IMHOTEP study probands, with a mean age of 40.6 years at diagnosis and a 65.9% male preponderance. The dominant populations recruited were 66% White and 20% Mixed. We reported a diagnostic yield of 29.3% (12/41), with *PKP2* accounting for 62% of the variants. Meanwhile, paediatric

probands constituted 13.9% (96/690) of the IMHOTEP study, and 70% were Black-Africans. The diagnostic yield for the paediatric cohort was 2.1% (2/96).

Conclusion: This study summarises the findings of the largest cardiomyopathy cohort in Africa. We show a very low genetic yield across all cohorts and that most probands are younger than reported internationally. We also report that DCM has emerged as the dominant phenotype in South Africa. EMF and RCM were categorised as restrictive cardiomyopathies; however, this study identified no single genetic cause in the Mozambique cohort. This study highlights that there is still much work to be done as we have only identified the genetic cause of disease in a small proportion of cardiomyopathy patients.

Chapter 1: Cardiovascular disease

1.1 **INTRODUCTION**

Cardiovascular diseases (CVDs) have reached epidemic proportions in Sub-Saharan Africa (SSA). While several studies in high-income countries (HIC) have found a decline in cardiovascular deaths (1), Africa has experienced a nearly 50% increase in CVD burden over the past three decades (2). In 2013, an estimated one million deaths were attributable to CVDs in SSA, which constituted 5.5% of the global fatalities related to CVDs and 11.3% of all deaths in Africa (3). Between 1990 and 2013, SSA remained the only geographical region where CVD-related deaths increased, as shown in Figure 1.1 (1, 3).

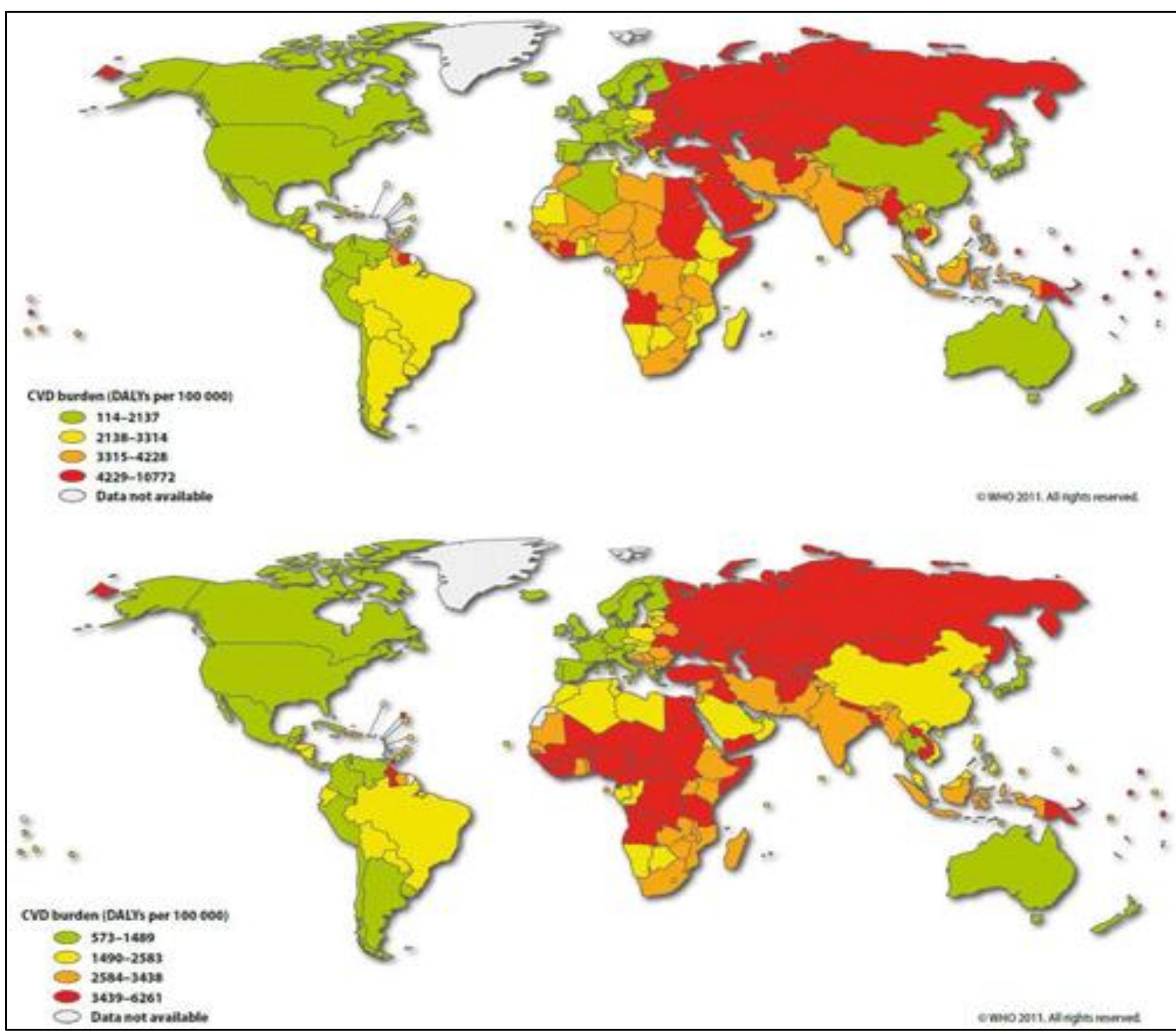


Figure 1.1: Global distribution of the cardiovascular diseases. *Global distribution of the CVDs-adjusted life years with several SSA countries depicted as recording some of the highest rates in the world in males (top) and females (bottom) (3).*

CVD has moved from being the sixth to the second leading cause of death between 1990 and 2019, with the CVD burden in SSA projected to double by 2030 (3). Population-based studies are showing that there is currently an epidemiological transition on the continent, where non-communicable diseases (NCDs) are projected to outpace communicable diseases within the current decade, with CVDs in particular, amongst the leading causes of morbidity and mortality worldwide, contributing to about a third of all deaths (4).

This dramatic shift in the CVD profile appears to be driven by increasing urbanisation, modernisation, westernisation and lack of socioeconomic development among some of the most vulnerable communities in the world (3). These uncontrolled increases in CVD risk factors have contributed to the growing burden of three major CVDs—hypertension (5), atherosclerotic and cardiomyopathy (CMO), as well as diseases that cause high rates of stroke and heart failure (HF) (5). The cardiomyopathies are also the primary cause of heart transplants worldwide as a result of HF (6).

Throughout this thesis, we mainly focussed on cardiomyopathies to fill in the significant gaps in knowledge in South Africa.

1.1.1 Cardiomyopathies (CMOs)

Cardiomyopathies (CMOs) are a complex and often silent heart condition that affects millions of individuals worldwide. They are a group of diseases that affect the heart muscle, impairing its ability to pump blood effectively. Patients with CMOs often have overt mechanical and electrical abnormalities, which can lead to inappropriate ventricular dilatation or hypertrophy and serious complications such as arrhythmias, HF, and, in extreme cases, sudden cardiac death (SCD) (5).

1.1.1.1 Disease burden

In the general population, the occurrence of heritable CMO in adults has ranged from 1:500 to 1:200 for the two common types and 1:5000 for the rarer phenotypes (7). Cardiomyopathy phenotypes in children (less than ten years old) have had an incidence of 1.3:100,000 (8) and 8.34:100,000 in infants (9).

Globally, the prevalence of CMO has increased by about 12% in the last decade (10) and in Africa, the few studies that have been conducted have reported a CMO prevalence rate of 12%

in Sudan, 17% in Kenya and 22% in Uganda (11). What is clear, though, is that many more detailed studies need to be conducted to determine the true prevalence of the CMOs in Africa.

Although it is unclear how much CMO has contributed to HF in Africa (12), the few hospital-based studies that have been done, mostly in South Africa, have shown that 40% of HF is caused by CMO (13, 14). We have summarised the prevalence data from a recently reviewed article on the genetics of inherited cardiomyopathies in Africa, as shown in Figure 1.2 (14).

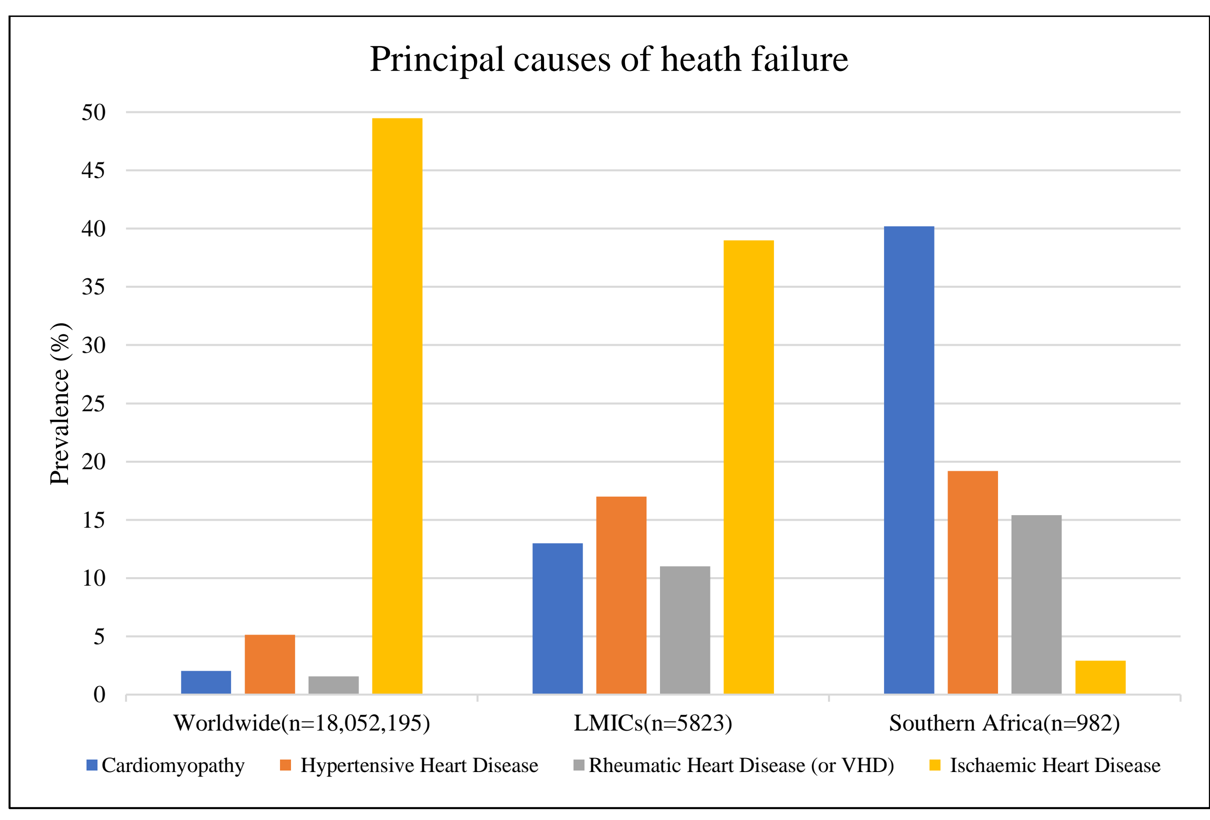


Figure 1.2: The principal causes of heath failure. *The global estimated prevalence of heart failure compared to the low-and-middle-income and Southern African countries. Drawn by P. Ndibangwi.*

Some hospital-based studies have even shown that twice as many South African males as females have died of CMO (15). In some instances, more than 70% of the cases were male (16-18). The reason for the male preponderance in DCM could be due to hormonal influences, gender, differences in environmental exposures or comorbid conditions, or biological factors such as intrinsic differences in the male and female heart (15, 19-21). Studies have also shown that patients of Black-African ancestry (BA) have a higher risk of developing HF at younger ages; deaths related to CMO can occur at any age, and the risk of death increases with age (22). The morbidity and mortality rates of CMO-induced HF in South Africa are less established than in high-income countries like the Netherlands (23). Thus, the information from these high-income countries has outpaced that from South Africa (14). However, the recent advancements

in Eurocentric medical knowledge have resulted in a better understanding of various aspects, such as the pathophysiology, genetic predisposition, diagnostic tools, management, and outcomes of conditions like peripartum cardiomyopathy (PPCM), which is highly prevalent in Africa (24).

The high prevalence of CMO in South Africa results from various factors: firstly, there is a rise in poverty levels and limited access to healthcare and education (25). Secondly, infectious diseases such as HIV/AIDS have led to further damage to the heart muscle (26). Thirdly, high rates of obesity and malnutrition have contributed to cardiovascular system disorders. Lastly, genetics have also played a significant role (27), as the history of human migration into South Africa has resulted in various CMO phenotypes. For example, hypertrophic CMO (HCM) is the most common among White (and Mixed) patients of European descent in South Africa (28).

1.1.1.2 Signs and symptoms

The signs and symptoms of CMO may include an irregular heartbeat (arrhythmia), shortness of breath (dyspnoea), extreme tiredness (fatigue), fainting episodes (syncope), pounding in the chest (palpitations), light-headedness, dizziness and swelling of the legs and feet (29-31). In some individuals, the first sign is (SCD (32).

1.1.1.3 Clinical characteristics

Clinical characterisation protocols for assessing and categorising all forms of CMO have included a genetic component (33). For example, the inclusion of genetics investigation has revealed that about half of patients clinically diagnosed with CMO have had familial aetiology (34). Although the inclusion of genetics investigation for assessing CMO has been accepted globally (35), most of the data used in developing these guidelines (such as the human genome project) have failed to represent the entire human population equally (36, 37).

The onset and development of CMO abnormalities have involved complex pathophysiological and biological mechanisms (38, 39). In all aspects of CMO pathomechanisms, an early age of onset and failing chambers in the heart have been associated with severe phenotypes (40). Few studies on CMO emerging from Africa have investigated the pathophysiology more than the pathomechanisms.

Figure 1.3 shows the ESC classification of primary types of CMO based on the gross morpho-functional appearance of the myocardium are arrhythmogenic cardiomyopathy (ACM), dilated cardiomyopathy (DCM), hypertrophic cardiomyopathy (HCM), and restrictive cardiomyopathy (RCM) (41).

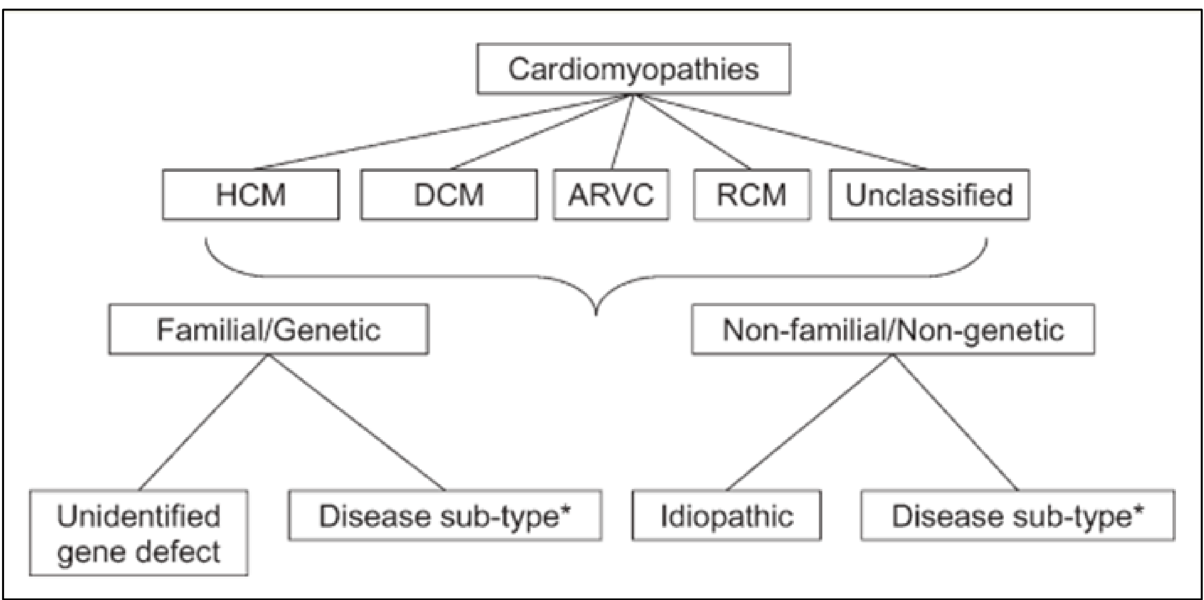


Figure 1.3: Classification of cardiomyopathies. *The ESC established classification system of primary types of cardiomyopathies (41).*

The four forms of CMOs shown in Figure 1.4 are reviewed further within the context of the specific phenotype in chapter 4 (DCM), chapter 5 (HCM), chapter 6 (RCM) and chapter 7 (ACM).

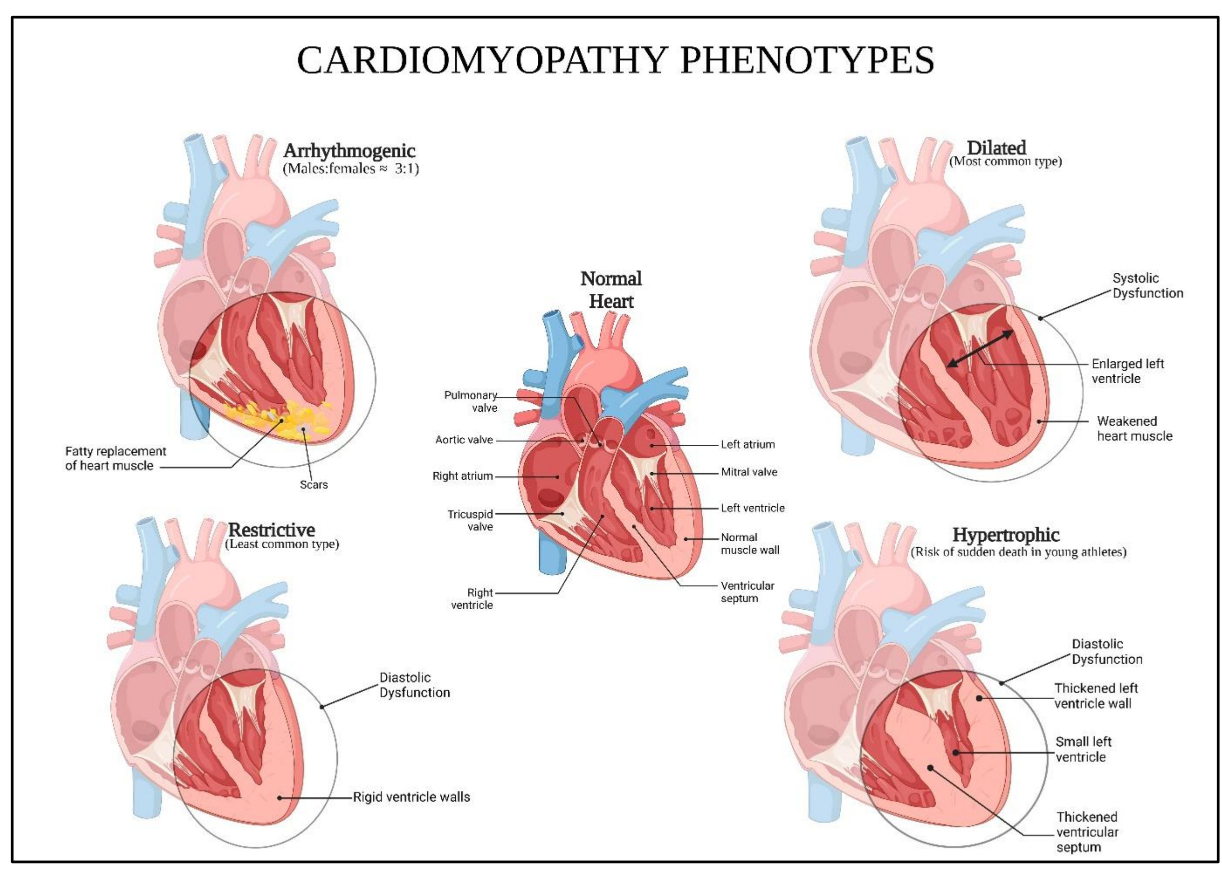


Figure 1.4: Cardiomyopathy phenotypes. *The various forms of heritable cardiomyopathy have distinct myocardial morpho-functional structures. The illustration was created by P. Ndibangwi using BioRender.*

Clinical phenotyping of the probands requires the attending physician to be able to differentiate between the various phenotypes (or subtypes) and determine the severity for management (42). Treatment is crucial for patients who may require heart transplantation at their initial presentation (43). Early identification of individuals "at risk" has required thorough clinical assessment and accurate classification (44). Figure 1.5 concisely summarises clinical investigations needed for adolescent and adult patients with or suspected of having CMO (45-48).

CLINICAL DIAGNOSIS AND EVALUATION FOR PATIENTS WITH OR SUSPECTED OF HAVING CARDIOMYOPATHY

TEST	INVESTIGATION	DILATED	HYPERTROPHIC	RESTRICTIVE	AHRRYTHMOGENIC
Biochemical	Enzymes: • Creatine kinase • Troponin T, Troponin I • N-terminal prohormone of brain natriuretic peptide (pro-BNP)	• Higher NT-proBNP due to systolic dysfunction • Elevated CK-MB and cardiac troponin-I	Elevated: • CK and CK-MB • cardiac troponin-T	• Elevated NT-proBNP reflecting amyloid deposite • Cardiac troponin higher	Elevated: • cardiac troponins
	Inflammation Biomarkers: • c-reactive protein (CRP) • Erythrocyte sedimentation rate • White Blodd Cells Count (WBC)	Elevated: • CRP (due to endothelial dysfunction) • ratio of neutrophile to lymphocyte	Elevated: • Cytokines (IL-6, TNF-*α*) • *CRP (with LVOT obstruction)*	• Eosinophilic infiltration • T lymphocytes present	Elevated: • Cytokines • Neutrophiles
	Electrolytes: • Sodium (Na^{+}) • Calcium (Ca^{2+}) • Creatinine	• Hypocalcaemia known rare cause of DCM	Hypophosphatemia and hypomagnesaemia are uncommon electrolytes imbalance capable of causing cardiomyopathy		
Clinical	Electrocardiogram	• *Late potential on signal-average ECG* • *(in)complete RBBB* • *Epsilon wave*	• *q & t waves frequent in sarcomeric disease* • *QTc prologation in advanced HCM*	• *AV block* • *Atrisl & ventricular arrhythmia* • *Low voltage QRS*	• Inverted T-waves, • QRS ≥55ms, • Epsilon wave, • low QRS voltages (<0.5 mV peak to peak)
	Echocardiography	• LV end-diastolic dimension >112% • LV enlargement • LV wall thinning	• *max LV thickeness* ≥15mm (or 13-14mm) consisten with HCM • Left atrial dilatation • Diastolic dysfunction	• *LV diastolic dysfunction* • *Biatrial dilatation* • *Near-normal systolic function*	• Global systolic dysfunction • Stria pattern
	Cardiac magnetic resonance imaging	• LVEF more predictive of DCM mortality • LVEF <50% indicates systolic dysfunction	• abnormal asymmetrical and variable wall thickening	• *Bilateral enlargement* • *Specific LGE patterns e.g. spotty intramural in sarcoidosis* • *Low T2 in siderosis*	• Fibrous replacement of the myocardium • akinesis/diskinesis
Genetic	Genetic counseling	Counseling is provided prior to family history. Pedigree charts are drawn to hold phenotyping summary, medical flagged notes or reports or special clinical diagnoses/procedures to show the relationship between proband and individuals across several generations			
	Next-Generation Sequencing *(targeted or whole exome or genome sequencing)*	Next-Generation sequencing is use for routine and discovery investigations in Mendelian diseses such as heritable cardiomyopathies			
	Sanger Sequencing *(cascade testing of fisrt degree relatives)*	Sanger sequencing is used as a gold standard to confirm NGS results for diagnosis and surveilance for LVH development in negetive relatives			

Figure 1.5: An overview of the clinical investigation. *The diagnostic clinical investigation and test performed on patients suspected of cardiomyopathy was drawn by P. Ndibangwi using BioRender.*

1.1.1.4 Classification

Cardiomyopathies have undergone various classifications over the past few decades, as shown in Figure 1.6, with the guidelines integrating morphological, functional, and genetic descriptors (49, 50). For example, the MOGE (S) nosology system embodies all these characteristics. It describes the morphofunctional phenotype (M), organ (s) involvement (O), genetic inheritance pattern (G), etiological annotation (E), including genetic defect or underlying disease/substrate, and the functional status (S) of the disease using both the American College of Cardiology/American Heart Association stage and New York Heart Association (NYHA) functional class (50). However, the guidelines are constantly updated, targeting specific forms of CMOs and, more significantly, focusing on specific geographical regions (51).

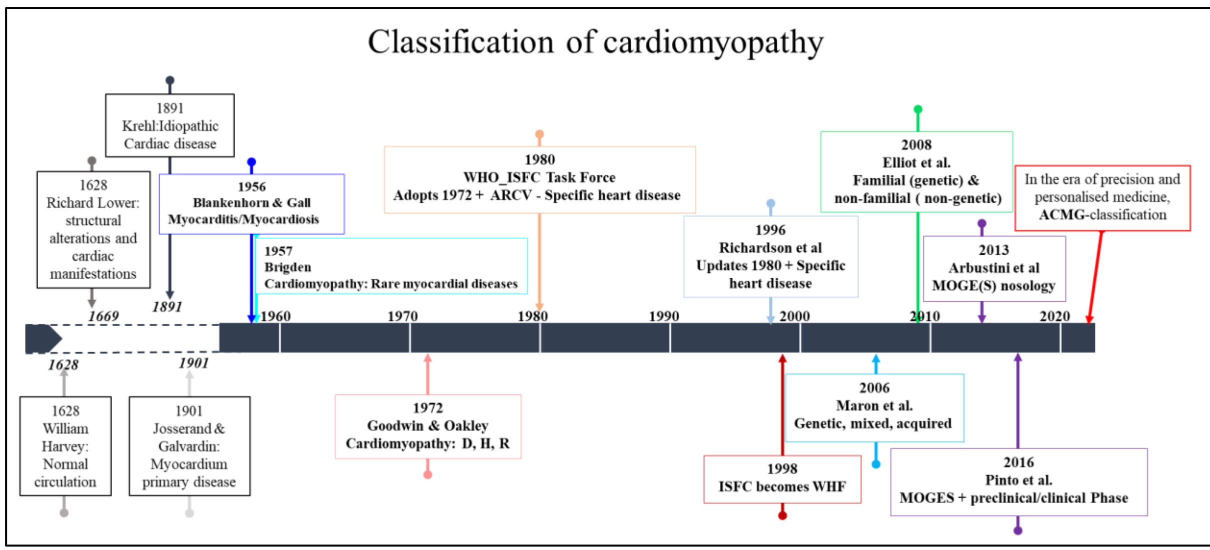


Figure 1.6: Cardiomyopathy classification timeline. *Historical classification of the morpho-functional and genetic aspects of the cardiomyopathies (52).*

For an accurate clinical genetic diagnosis, a different classification system that requires a multiparametric approach was proposed for identifying the genetic aetiology of CMOs (35). The American College of Medical Genetics and Genomics/Association for Molecular Pathology (ACMG/AMP) CMO working group task teams developed standardised five-tier system guidelines for curating and investigating genetic variants (35). The five-tier system considered several critical factors, including the frequency of the variant in the population, functional data, *in silico* predictions, segregation data, *de novo* variants, disease-specific databases, allelic heterogeneity, functional domains, and clinical phenotype.

Rare genetic variants are more likely to be deleterious, while functional studies showing the variant's impact on the gene or protein's function increase the likelihood of pathogenicity (35). Computational tools can predict the potential effects of the variant, and if it segregates with the disease within families, it supports a pathogenic (P) role. *De novo* variants and information from disease-specific databases or literature documenting the association of the variant with the specific disease are also considered. Multiple P variants in the same gene associated with similar phenotypes support the pathogenicity of a novel variant in that gene. The variant's location within the functional domains of the gene is essential and should be considered. It supports pathogenicity if the variant is frequently associated with the observed clinical features. The combination of these factors contributes to an overall variant classification according to the ACMG/AMP guidelines. Variants may be classified as P or likely pathogenic (LP) if they meet specific criteria across multiple categories. If the evidence is inconclusive or conflicting, the variant may be classified as uncertain significance (VUS) or (likely benign (LB) or benign

(B). Regular updates and re-evaluation of variants are necessary as new evidence becomes available.

Despite the importance of having new data representing the whole human race, Africa, for example, has had no population-based data for CMO and related diseases (53). Therefore, there is an urgent need for more diverse data to be collected and included in developing guidelines for CMO assessment, management, and treatment (51). Towards this end, the Afr*i*can Cardio*m*yopat*h*y and My*o*cardi*t*is R*e*gistry *P*rogram (IMHOTEP) study designed in 2015 aimed to investigate the genetic characteristics of heritable forms of CMO in a larger African cohort.

1.1.1.5 Aetiology

CMOs are distinct disorders with a wide range of known and unknown aetiologies. They could be hereditary or caused by other conditions such as infections or stress (or autoimmune), lifestyle (toxins, alcohol, or heavy metals), amyloidosis or pregnancy (23, 54). The causes of CMO have remained largely unknown, and the clinical aetiology has ranged from idiopathic to sporadic cases, including those with familial history (55). Although CMO could occur as a primary or secondary (due to viral infections or therapeutic procedures) disorder with similar aetiology, we have focused on the primary phenotypes in this study.

We explore the aetiologies under the various types of cardiomyopathies in chapters 3 (DCM); section 3.1.4, chapter 4 (HCM); section 4.1.4, chapter 5 (RCM); section 5.1.4 and chapter 6 (ACM); section 6.1.4.

1.1.1.6 Environmental risk factors

Environmental factors modulate the penetrance and expressivity of CMOs (54). You have the environmental risk factor (e.g. alcohol, virus, drugs) that can lead to secondary cardiomyopathy, such as CMO secondary to alcohol abuse, CMO secondary to chemotherapeutic agents, CMO secondary to severe thyroid disease, CMO secondary to sarcoidosis, CMO due to infectious processes (HIV, intravenous drug abuse, beriberi, severe chronic anaemia, hemochromatosis, pheochromocytoma, acromegaly, genetic storage or neuromuscular disease, connective tissue disease, amyloidosis, neoplastic heart disease, cardiac surgery, or penetrating trauma to the heart) (21, 54).

Certain infectious diseases prevalent in Africa, including Chagas disease and endomyocardial fibrosis, have been associated with the development of CMO (25, 56). Malnutrition and specific nutrient deficiencies may also contribute to this disease. For example, deficiencies in selenium and thiamine have been linked to CMO (57, 58). Exposure to certain environmental toxins and pollutants, such as heavy metals or specific chemicals, can also contribute to this disease. Also, exposure to toxins from contaminated water sources may be a concern in some regions. Parasitic infections may be relevant in certain African regions (59), including those caused by parasites such as *Trypanosoma cruzi* and parasites that can cause eosinophilic myocarditis. Environmental factors that contribute to hypertension, such as lifestyle and diet, may indirectly influence the risk of developing CMO (53). High blood pressure is a known risk factor for various types of CMO. Limited access to healthcare facilities and medical resources may impact CMO prevention, diagnosis, and management in some regions (11). It is also important to note that these factors can vary across Africa (11, 53, 60).

Environmental risk factors can also impact disease development if a patient already has a genetic predisposition. For example, *TTN*tv represents a prevalent genetic predisposition for alcoholic CMOs and is also associated with a worse LVEF in DCM patients who consume alcohol above recommended levels (54). They identified a direct interaction between *TTN*tv and alcohol consumption in the context of typical DCM: cases with a *TTN*tv and excess alcohol consumption had a markedly reduced LVEF compared with those with low alcohol intake. They postulate that certain drugs, nutritional deficiencies, recreational drug use, and viral myocarditis are additional environmental factors that might work in tandem with *TTN*tv (61).

Another example is the impact of sex-related differences on genetic cardiomyopathies (62, 63). Argiro and others described how phenotypic expression of CMOs may differ profoundly between sexes (63). For example, in patients with familial DCM that is acquired in an autosomal dominant (AD) manner and carries *TTN*tv, the mutation has less impact on women and is associated with better systolic function and a lower rate of atrial fibrillation. On the other hand, women who carry *LMNA* variants have about 45% lower risk of life-threatening ventricular arrhythmia (63). In contrast, progression to end-stage HF in *LMNA* DCM does not appear to vary by sex (64, 65). The study also found that women with HCM were typically diagnosed at an older age than men and were more likely to have a sarcomere mutation. Meanwhile, men with ACM had increased prevalence and worse outcomes compared with women; the impact of sex hormones and physical activity on this observation has been suggested.

1.1.1.7 Heredity risk factors

Heredity risk factors contributing to CMOs include genetic mutations, family history (FH), genetic diversity, consanguinity, and genetic testing. Specific gene mutations often cause inherited CMOs. For example, mutations in genes like *MYH7*, *MYBPC3*, and *TNNT2* have been linked to hypertrophic cardiomyopathy (66, 67). A significant risk factor for developing CMO or SCD is having a FH of the disease (68). If someone has a close relative who has been diagnosed with any CMO, other members of the family may also be at an increased risk. Africa is known for its high genetic diversity due to various ethnic groups and populations (37). Different regions may have specific genetic variants linked to CMO. Also, in some areas, marriages between close relatives (consanguinity) are more common, which can increase the risk of inherited genetic conditions, including CMOs, if there is a FH (69). Further, advances in genetic testing technologies can help identify specific genetic mutations associated with CMOs. However, the availability and accessibility of genetic testing can differ across various regions of Africa (57).

1.1.1.8 South Africa's genetic populations

Before we can delve into the genetics of cardiomyopathy in South Africa, it is important that we first address the turbulent history of a country considered to be one of the most genetically diverse populations in the world and the numerous colonisation events throughout its history that shaped this "rainbow nation". The history of this country highlights the interplay between different genetic backgrounds and environmental factors and how it shaped the unique genetic makeup of individuals in South Africa (70-72). Unravelling the intricacies of the South African population's genetic makeup has become increasingly important as it can unlock insights into various genetic traits, diseases, and responses to medications specific to this population (70, 73, 74). Understanding this genetic diversity is crucial for medical research and personalised healthcare.

Embedded in the DNA of all South Africans is a history of migration events (Figure 1.7) that dates back to the beginnings of our species, about 300,000 years ago. Over time, these population movements have significantly influenced and altered the genetic landscape of Southern African groups, resulting in a mix of genetics, languages, cultures, and traditions found nowhere else in the world (27, 75)—the most significant of these being the colonising events of South Africa, mainly attributed to the Europeans.

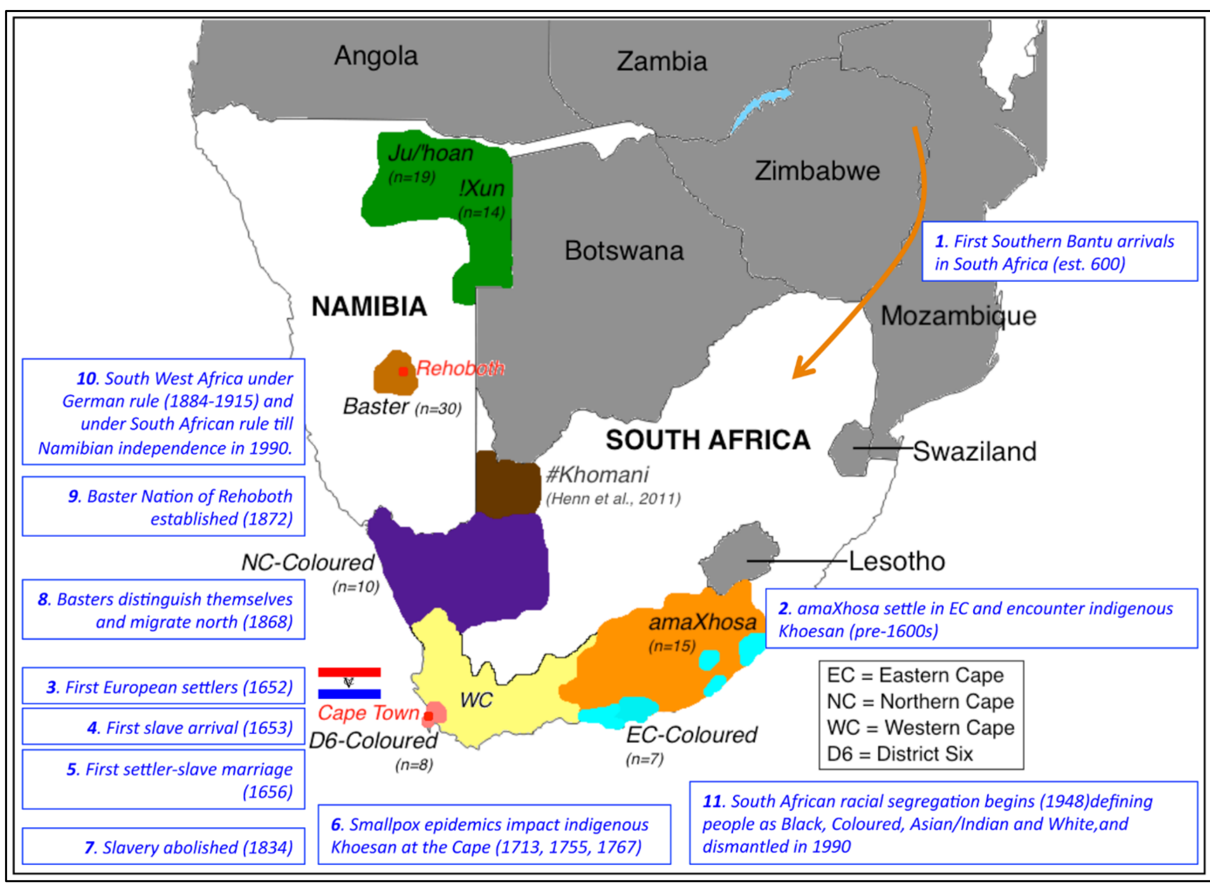


Figure 1.7: South African historical map. *A map of Southern Africa showing significant historical events that have shaped regional population admixture is indicated chronologically.*

Jan van Riebeeck sailed from the Netherlands on December 24th, 1651, to establish a refreshment station at the Cape of Good Hope. The station was to provide supplies to Dutch ships travelling to the East under a contract with the Dutch East India Company (VOC). Sailing on the Dromedaris with two other ships, the Rejiger and De Goede Hoop, Van Riebeeck was accompanied by 82 men and eight women. On April 5th, 1652, they sailed into Cape of Good Hope, where they immediately started construction of the refreshment station. The arrival of Van Riebeeck marked the beginning of the permanent European settlement in the region.

The colonisation of South Africa is mostly attributed to the Europeans. A closer investigation shows that the colonisation events can be divided into three categories: an 'unofficial colonisation', two 'official colonisations' and an 'internal colonisation'. The first was about 2000 years ago, an unofficial colonisation by the Bantu people from the North (Figure 1.7). This colonisation ended in 1880 when the country was divided into four polities, two governed by the British and two by the Afrikaners. The two official colonising events occurred when the Dutch VOC formally colonised South Africa in 1652, starting from the south. Great Britain again colonised South Africa after the British acquired the country from the Dutch in 1806.

When the nation became a Republic in 1961, this era ended. White Afrikaners began their internal colonisation of the nation in 1961 under the "Apartheid" regime, which ended in 1994 when democracy was established (76).

According to research by Petersen et al., genetic diversity within populations is highest in Africa, whereas genetic diversity between populations is directly correlated with geographic distance (Figure 1.8). Here, the Khoesan, or click-speaking forager peoples of southern Africa, are amongst Southern Africa's most divergent contemporary human populations. Their research showed that the complex admixture patterns between historically geographically isolated Khoesan and, more recently, diverged populations have resulted from intra- (Bantu expansion) and inter-continental migration (European-driven colonisation). They also demonstrated how the aboriginal Khoesan contributions to the Cape Coloured (or Mixed people as they are known today) and Baster populations were mainly a result of the Khoesan's assimilation with European settlement at the southernmost tip of Africa (27).

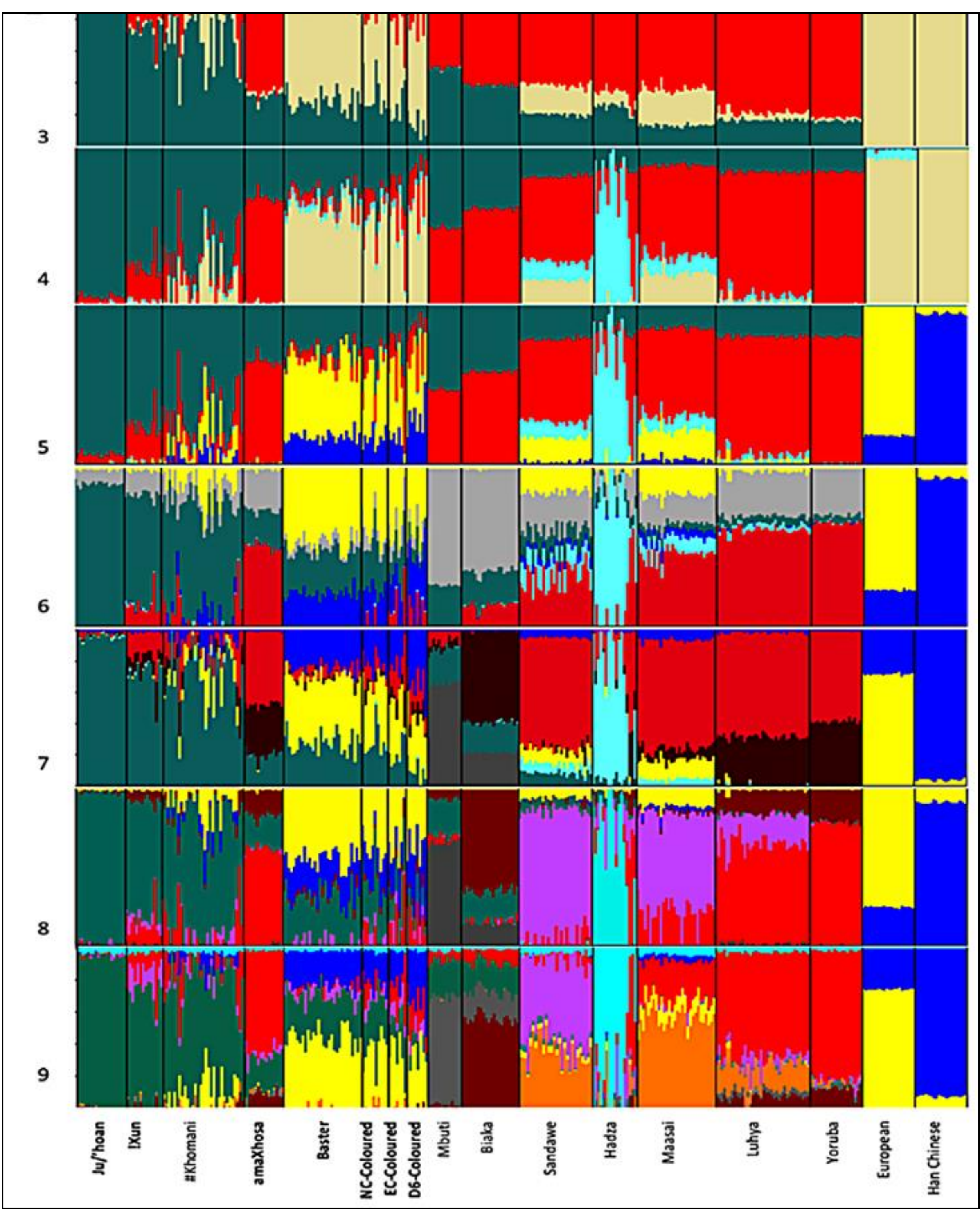


Figure 1.8: Population substructures in South Africa. *Shows the population substructure within the five southern African populations defined as Khoesan (specifically Ju/' hoan and !Xun), Southern Bantu (specifically amaXhosa), and European-initiated admixed populations (specifically Mixed) (27).*

When the groups from the north and the south arrived in this region of Africa, the Cape was not home to the Cape Coloured/Mixed (27). The Dutch arrival in South Africa is officially credited with giving rise to the Mixed people. First, when Dutch and Malay slaves arrived at the Cape in 1657, there were Mixed children. The Khoikhoi, the San, and eventually the Xhosa

people were among the Mixed children of the soldiers and settlers who arrived at the Cape (76). Additionally, the 170 years of East Indian slave trade and intracontinental migrations are responsible for the complex genetic pattern and highly specific admixture produced in the Cape Coloured/Mixed populations (27).

1.1.1.9 Genetics of cardiomyopathy in South Africa

This brings us to the topic of the genetics of cardiomyopathy in South Africa and the impact of this diversity on clinical research.

1.1.1.10 Heritable cardiomyopathy

1.1.1.11 Types of heritable cardiomyopathy

Cardiomyopathy is a rare heart muscle disease with genetic determinants (77), and over 50% of CMO patients have a variant that causes the condition (34). The four forms of heritable CMOs with genetic aetiologies are reviewed further within the context of the specific phenotype in chapter 4 (DCM), chapter 5 (HCM), chapter 6 (RCM) and chapter 7 (ACM); however, Figure 1.9 summarises the distinct clinical features reported in the literature.

CLINICAL SUMMARY OF HERITABLE CARDIOMYOPATHIES

Type of Cardiomyopathy	Arrhythmogenic	Dilated	Hypertrophic	Restrictive	Unclassified
Relevant Clinical Notes	Progressive replacement of the ventricular myocardium with fibrofatty tissue, often associated with arrhythmia, syncope, palpitations or sudden cardiac death (SCD)	Left or biventricular dilatation and impaired systolic function in the absence of CAD, hypertension, valvular disease or CHD	Unexplained LVH (increased LV wall thickness > 15mm) in the absence of hypertension or valvular disease	Restrictive ventricular physiology (increased myocardial stiffness causing reduced ventricular filling), sometimes with reduced systolic and diastolic volumes	**Non-compaction:** Prominent LV and/or RV trabeculae which may be associated with LV dilatation or impaired systolic function
International Prevalence	1/5,000 to 1/2,000	1/250	1/500	Rare	Rare
Common Signs or Symptoms	• Palpitations • Syncope • SCD	• EF <50% • Heart Failure • Arrhythmia	• Arrhythmia • Chest pain • Syncope • HF, SCD or • Thromboembolism	• Arrhythmia, HF • Severe diastolic dysfunction • Pulmonary hypertension • conduction defects	• Arrhythmia • Syncope • HF • SCD • Thromboembolism
Age of Onset (years)	~20 to 40	~20 to 50	Any	Any	Any (median age 40–50years)
Genetics: Known Genes Involved	Desmosomal genes (e.g. *PKP2, DES, DSP, DSC2, DSG2*) Other genes include: TMEM43, MYBPC3, etc 20-35% Penetrance	Genes involved in • force generation e.g. *TTN* & *ACTC1* • transduction e.g. *TTN* & *MYBPC3* • nuclear proteins e.g. *LMNA* & *RBM20* • others *e.g. BAG3, etc*	Sarcomeric genes (e.g. *MYH7, MYBPC3, TNNC1, TNNI3, TNNT2, TPM1, MYL2, MYL3*) other genes include: *TTN, GLA, PLN*, etc	DCM and HCM genes (e.g. *MYH7, FLNC, TTN*)	Genes for the cardiomyopathy identified in association with the non-compaction phenotype
Genetics: Diagnostic Yield	Detectable in 10-50%	Detectable in 10-40%	Detectable in 30-60%	Detectable in 10-60%	Unknown
Patients Management and outcomes	• Avoidance of strenuous exercise • Risk of SCD • Treatment with ICD in high-risk patients to prevent SCD	• Treatment of HF symptoms • Heart transplantation • High risk of mortality from HF or SCD	• Risk of HF or SCD • Treatment with ICD in high-risk patients to prevent SCD	High mortality within the 1st few years often due to • HF • SCD • Heart block	• Treatment of HF symptoms • ICD implantation • Risk assessment

Figure 1.9: Clinical summary of heritable cardiomyopathies. *Clinical and genetics summary of the various forms of heritable cardiomyopathy, drawn by P. Ndibangwi using BioRender.*

The identification of the pathogenic variants results in the proband and immediate family being earmarked for a clinical management program (consisting of genetic counsellors and clinicians) for inherited cardiovascular disorders (78, 79), including conditions that can cause SCD, such as CMOs, aortopathies, inherited arrhythmia syndromes, and lipid disorders (80). However, genetic testing results are difficult to interpret due to incomplete penetrance and variable expressivity of genes, which are hallmarks of these conditions (81). Many genes commonly used in genetic testing panels for heritable CMOs and SCD lack strong gene-disease associations or accurate gene penetrance estimates (82). As a result, there is a risk of misclassifying variants detected in these genes (83, 84), especially when identified without clinical symptoms or a FH of related disease.

1.1.1.12 The genetic landscape of heritable cardiomyopathy

Researchers have identified thousands of variants in over 200 CMO genes that could explain the observed phenotypes in most patients (9, 85, 86). As genetic data collected from patients and their relatives grows, the American College of Medical Genetics and Genomics/Association for Molecular Pathology (ACMG/AMP) CMO working group task teams have developed standardised guidelines for curating and investigating genetic variants (35). These guidelines are necessary due to the vast amount of data available, including detectable genes and phenotype-genotype variability (35, 87). However, implementing the ACMG/AMP five-tier evidence and rules has gained polarity when classifying disease-causing variants in CMOs (22). For example, missense variants occurring in the β-myosin heavy chain (*MYH7*) gene are evaluated differently in DCM and HCM. Besides, most genes have been cautiously analysed for possible CMO-causing variants, leaving class 3 variants largely untouched. Some class 3 variants have disease-modifying effects, which can result in incomplete penetrance, phenotypic variation, and overlap (88, 89). As a result, the ACMG guidelines are reviewed continuously by specific forms of disease teams of experts. Therefore, this study classified class 3 variants with unknown genotype-phenotype correlation as possible genetic modifiers. These variants are plausibly pathogenic but currently lack the evidence to be classified as likely pathogenic (class 4).

1.1.1.13 Prevalent genes

The genes responsible for causing CMO have varying underlying biological mechanisms, resulting in structural or functional damage to the heart muscle. The genes' pathomechanisms can be complex and involve multiple factors, including genetic variations, gene-environment interactions, and lifestyle choices. For example, genetic variation in a failing heart could result in various phenotypes, as summarised in Figure 1.10 (90). HF is primarily initiated at the myocardial and cardiomyocyte levels. The pathology within the myocardium can include reactive or replacement myocardial interstitial fibrosis, microvascular rarefaction characterised by decreased microvascular density, and cardiomyocyte hypertrophy and apoptosis. Important mechanisms at the cardiomyocyte level include increased oxidative stress and damaging X-reactive oxygen species (X-ROS) signalling, altered energetics with decreased fatty acid β-oxidation and increased glycolysis, and impaired contractility due to pathogenic variants in genes encoding sarcomeric proteins. These pathological alterations can lead to structural remodelling of the heart, resulting in either dilatation or hypertrophy of the left ventricle. This can lead to RCM, which has lower stroke volume and higher filling pressures, or DCM, which is characterised by decreased myocardial contractility (91).

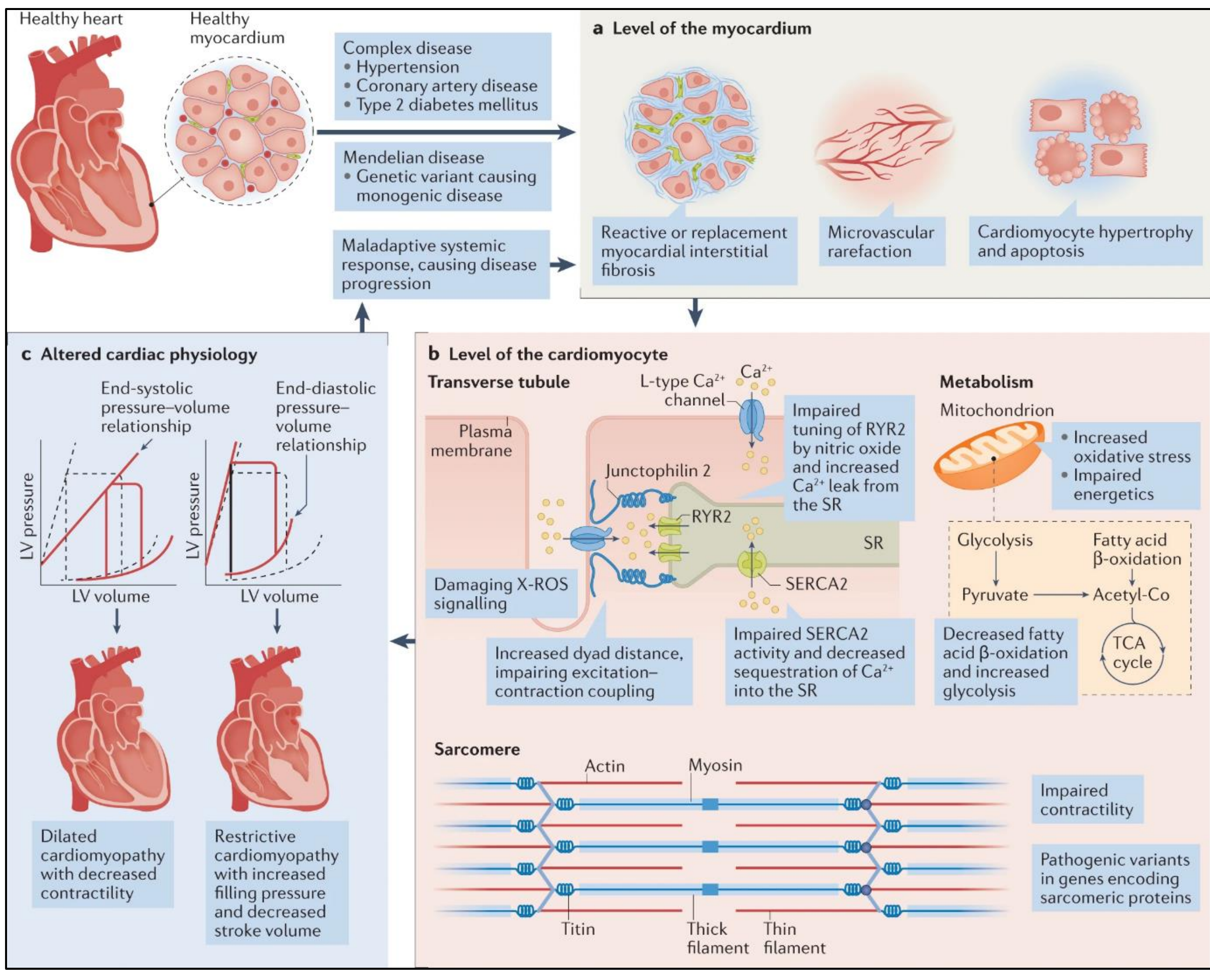


Figure 1.10: Biomechanisms of a failing heart. *The onset of heart failure occurs at the myocardial and cardiomyocyte levels. a) Pathology within the myocardium can include reactive or replacement myocardial interstitial fibrosis. b) At the level of the heart muscle cell, important mechanisms include a decrease in the coupling of excitation to contraction, coupled with impaired handling of Ca^{2+}. c) The pathological changes cause structural remodelling of the heart, resulting in the enlargement or thickening of the left ventricle. LV, left ventricular; RYR2, ryanodine receptor 2; SERCA2, sarcoplasmic– endoplasmic reticulum Ca^{2+} ATPase 2; SR, sarcoplasmic reticulum; TCA, tricarboxylic acid (91).*

Finding the genes and understanding the various pathomechanisms is important for our population, as new findings are crucial for developing effective treatments and management plans. Notably, the current understanding of these genes' biological mechanisms has helped develop routine genetic diagnostic strategies in countries like the United Kingdom (UK).

Genes that cause cardiomyopathy have been found to cluster within various compartments of cardiomyocytes (Figure 1.11), including plasmamembrane–associated genes, sarcoglycans, the cytoplasm of the cardiomyocyte, the nuclear membrane genes and genes involved in Ca^{2+} homeostasis regulation. These compartments have been associated with particular phenotypes such as DCM, HCM, ACM and RCM (92).

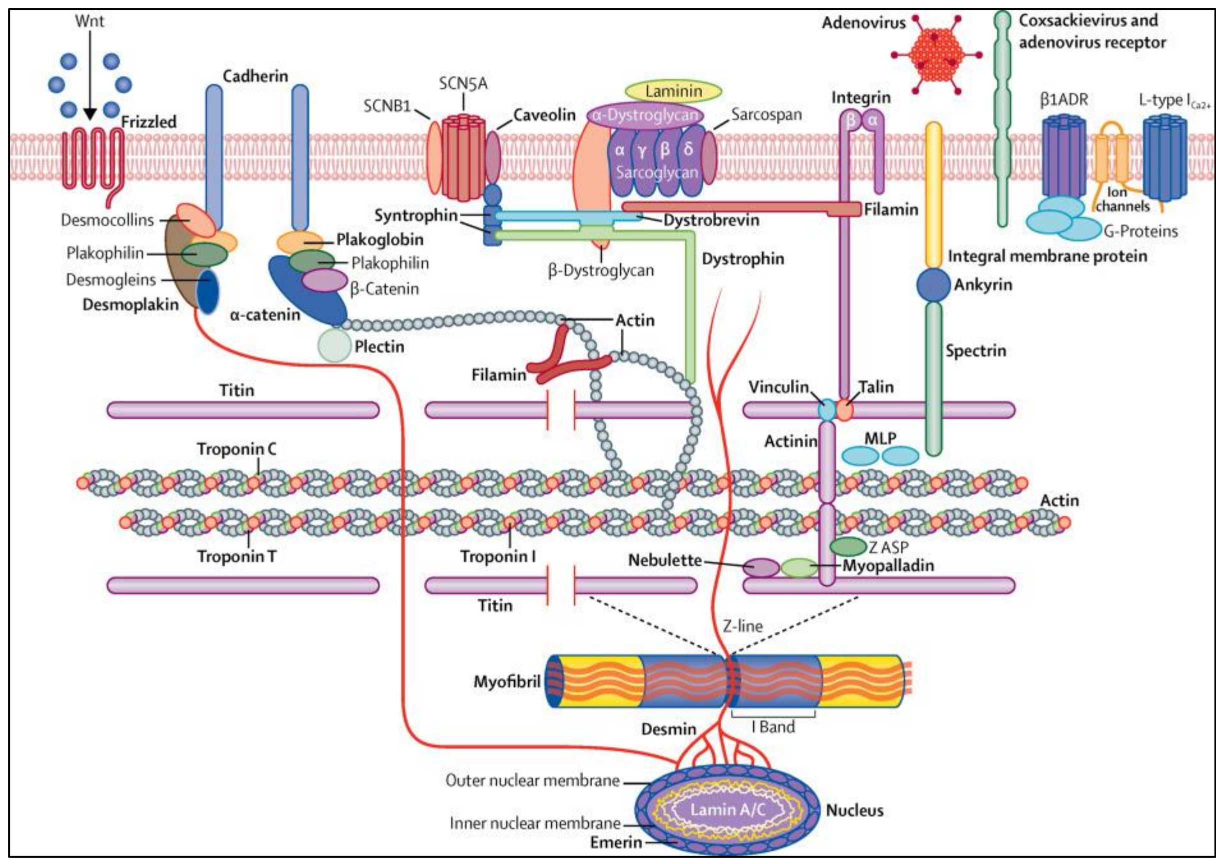


Figure 1.11: Common genes involved in cardiomyopathies. *Mutations in plasma membrane-associated proteins such as dystrophin and associated proteins can lead to ACM. Mutations in sarcomere proteins result in HCM, DCM, or a mixed phenotype. Nuclear membrane gene mutations, such as lamins A and C, also cause inherited DCM and cardiac conduction system disease (92).*

For example, the *TTN* gene in the sarcomere, whose protein product (titin) attaches to the z-disc and stretches across the thin and thick filaments to the M-Line, is associated with most titinopathy. Then, there is the filamin C (*FLNC*) gene that has been linked to myocardial damage in various types of CMO (93, 94). Table 1.1 shows the prevalent cardiomyopathy-causing genes in South Africa (95-98). In contrast, the *PKP2* gene product (plakophilin 2), which makes up the desmosomal plaque, is more strongly linked with ACM (95).

Table 1.1: Common genes associated with heritable cardiomyopathies.

Common Genes Associated with Heritable Cardiomyopathies								
Gene			Genotype-Phenotype(s) Prevalence (%)				Inheritance	
Name	Symbol	Function	*ACM*	*DCM*	*HCM*	*RCM*	Mode	Features
α-actin	*ACTC1*	Sarcomere	*Rare*	<1	*Rare - 1*	*Rare*	*AD*	Few mutations, linked to overlapping phenotypes
α-actinin 2	*ACTN2*	Z-disc	*Rare*	Rare	*Rare - 1*	*Rare*	*AR*	Left ventricular hypertrophy
BCL2-associated athanogene 3	*BAG3*	Z-disc	*Rare*	0-1	*Rare*	*Rare*	*AD*	Myofibrilar myopathy
Cadherin-2	*CDH2*	Intercalated disc	*Rare - 2*	Rare	*Rare*	*Rare*	*AD*	No specific genotype–phenotype relationship identified
Desmin	*DES*	Intermediate filament	*0- 2*	*Rare*	*Rare*	*Rare*	*AD/AR*	Fully penetrant; associated with LV-dominant ACM
Filamin C	*FLNC*	Cytoskeletal structure	*Rare - 3*	<1	*Rare - 1*	*Rare*	*AD*	High penetrant, associated with LV-dominant ACM
Lamin A/C	*LMNA*	Nuclear envelope	*Rare - 4*	4 - 8	*Rare - 1*	*Rare*	*AD/AR*	More common in severe forms of ACM with a dilated
Myosin-binding protein C	*MYBPC3*	Sarcomere	*Rare*	*2*	20 - 40	*Rare*	*AD*	Men are more likely to have LVH at a younger age
β-myosin heavy chain	*MYH7*	Sarcomere	*Rare*	*4 - 10*	15 - 25	*Rare*	*AD*	More cardiac conduction disease, ventricular arrhythmia
Plakophilin 2	*PKP2*	Desmosome	*20 - 45*	Rare	*Rare*	*Rare*	*AD*	Most commonly mutated
Sodium channel protein type 5 subunit alpha	*SCN5A*	Cardiac sodium channel	*2*	2 - 3	*Rare*	*Rare*	*AD*	Prolonged QRS interval
Transmembrane protein 43	*TMEM43*	Cytoskeletal structure	*Rare-2*	*Rare*	*Rare*	*Rare*	*Rare*	Fully penetrant; males more severely affected than females (frequent in Newfoundland, Canada)
Tropomyosin 1	*TPM1*	Sarcomere	0.5–1.0	Rare - 5	*Rare*	*Rare*	X-linked	X-linked Ebstein anomaly
Titin	*TTN*	Sarcomere	*Rare-10*	12 - 25	*Rare - 5*	*Rare*	*AD*	Higher risk of supraventricular tachycardia and progression to heart failure
Desmoplakin	*DSP*	Desmosome	1–3	3–4	*Rare*	*Rare*	*AR*	AR mutation associated with Carvajal syndrome

In the last decade, there has been an unprecedented increase in the number of genetic variants and genes discovered, thanks to next-generation sequencing (NGS) technology. We will dive into a more detailed explanation of NGS in section 1.4 of this chapter.

1.1.2 Next-generation sequencing (NGS)

The discovery of NGS, a high-throughput method that can sequence millions of DNA copies at once, revolutionised biomedical research and dramatically increased sequencing data output. Before NGS, *in vitro* DNA replication was achieved using the Sanger sequencing method. Sanger sequencing technology is not a high-throughput technique, but using it with the polymerase chain reaction (PCR) process (99) has enabled exponential amplification of very minute amounts of double-stranded (ds)DNA into millions of copies. The PCR amplification methods have revolutionised genetics in tandem with Sanger sequencing (100). PCR during sequencing has enabled the development of robust non-Sanger-based high-throughput sequencing platforms. For example, Illumina's and ThermoFishers' NGS platforms use bridging PCR and emulsion PCR technologies, respectively.

1.1.2.1 Next-generation sequencing technologies

Non-Sanger-based high-throughput sequencing technologies fall into two groups based on the output read length (101): the short-read (or second-generation sequencing) and long-read (or third-generation sequencing) technologies. The short-read technologies generate reads <600bps in length and *vice versa*. Figure 1.12 illustrates an overview of these NGS technologies.

Next Generation Sequencing Technologies: An Overview

	Short-read Technologies sequencing platforms		Long-read Technologies sequencing platforms	
Company	Illumina	ThermoFisher	Pacific Biosciences	Oxford Nanopore
System Platforms	iSeq MiniSeq MiSeq NextSeq NovaSeq	GeneStudio Genexus Ion-PGM-Dx	Sequel Sequel II Sequel IIe	Flongle MinION GridION PromethION
Sequencing Principle	Sequencing by Synthesis		Single Molecule Sequencing	
Detection	Fluorescent	Ion	Fluorescent	Electrical conductivity
Application	Small WGS, Small RNA, WGS, WES, TS, transcriptome and epigenome sequencing	Exome, TS, epigenome and transcriptome sequencing	Whole genome *de novo* assembly, Indels, SVs detection, full length transcript, targeted, and metagenomics sequencing, 5mC in CpG contexts	cDNA, DNA, amplicons, Direct RNA, Whole genome, metagenomics and epigenetics sequencing
Accuracy/Quality Score	≥99.66/Q30≥80%	≥99	HiFi Reads >99%: Up to 5000000	R9.4.1 Q50 at ~100X, >98.3%, R10.4 Q47 at ~ 60X with ~80 Mb N50 >99.3% modal
Advantages	High accuracy with good depth of coverage		Very long reads; no DNA amplification required; comparatively fast turnaround time	Fast-sequencing; small instrument footprint; portability; real-time data analysis
Disadvantages	Long run time with phasing difficulties		Expensive equipment; large instrument footprint; Historically higher error rate (continue to improve)	Historically higher error rate (continue to improve)

Figure 1.12: Next-generation sequencing technologies. *A schematic overview of NGS technologies, drawn by P. Ndibangwi using BioRender.*

1.1.2.2 Next-generation sequencing principles

NGS is a sequencing technology that uses two principles: sequencing by synthesis (SBS) for short reads and single-molecule real-time (SMRT) sequencing for long reads. Illumina's SBS technologies use fluorescently labelled nucleotides to sequence millions of DNA clusters in parallel, while ThermoFisher's SBS technologies detect, and measure hydrogen ions released during *in vitro* DNA polymerisation. On the other hand, PacBio's SMRT sequencing technology uses four fluorescently labelled nucleotides, and Nanopore's SMRT channels technology detects and measures electrical conductivity during *in vitro* DNA polymerisation. Although these companies use different principles, the NGS approach remains the same across all technologies.

1.1.2.3 Next-generation sequencing approaches

There are different approaches to NGS, including genome sequencing (GS), exome sequencing (ES), and targeted sequencing of specific regions (also known as panel sequencing). Figure 1.13 illustrates an overview of these NGS approaches.

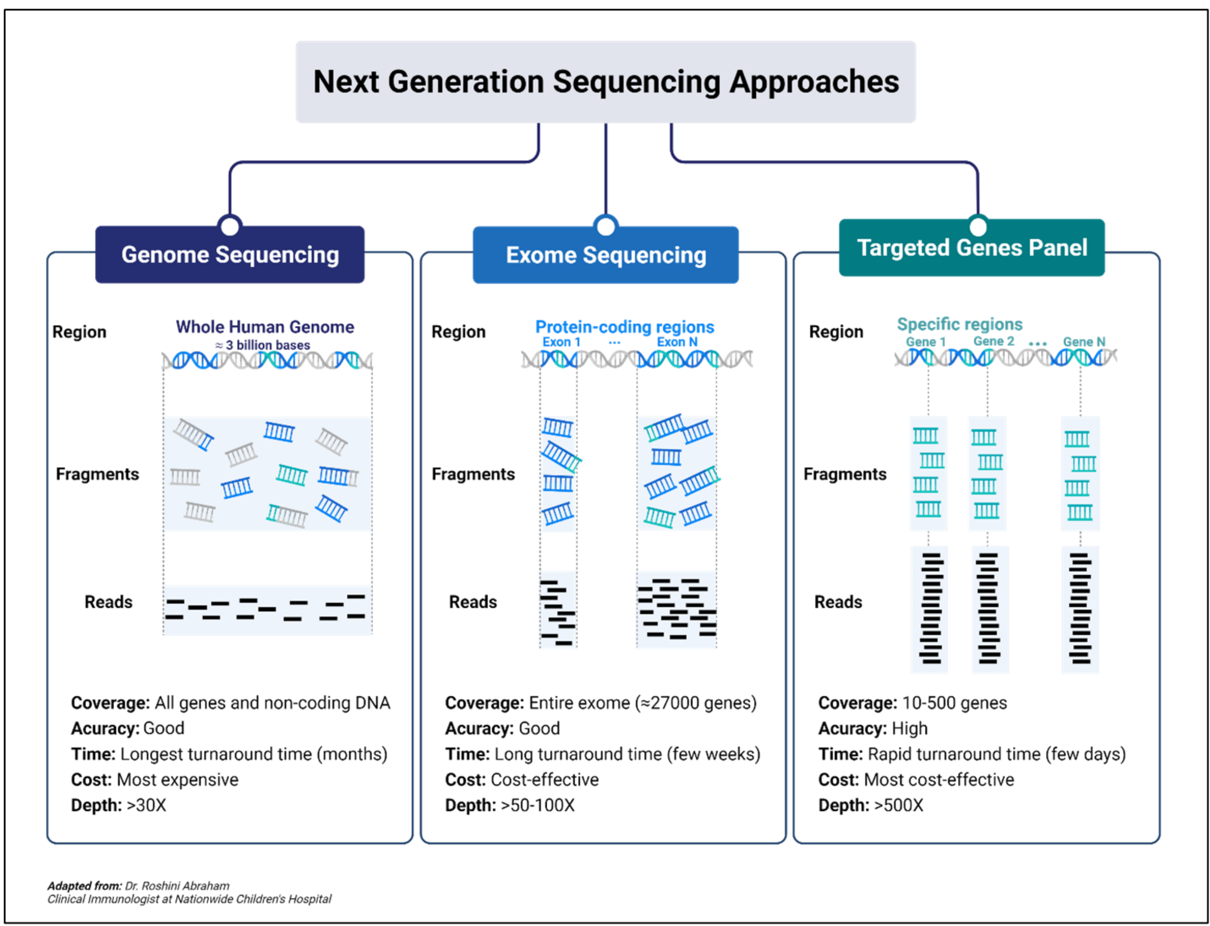


Figure 1.13: Next-generation sequencing approaches. *A schematic comparison of various NGS approaches, illustrated by P. Ndibangwi using a BioRender template.*

These approaches involve 1) fragmenting nucleic acid into smaller pieces to prepare samples, 2) sequencing and 3) data analysis (filtering and bioinformatics analysis) and storage of the data. Figure 1.14 depicts the pipeline strategy used in DNA re-sequencing. Briefly, sequencing is a process that involves using a sequencer to analyse an individual's DNA. The sequencer generates raw read data, which is then aligned, sorted, and mapped onto the human reference genome sequence. Bioinformatic tools are then used to filter nucleotides in the targeted regions that vary from those in the human reference genome sequence. These variations are further analysed to determine their possible involvement in the observed phenotype progression. This process is known as targeted re-sequencing.

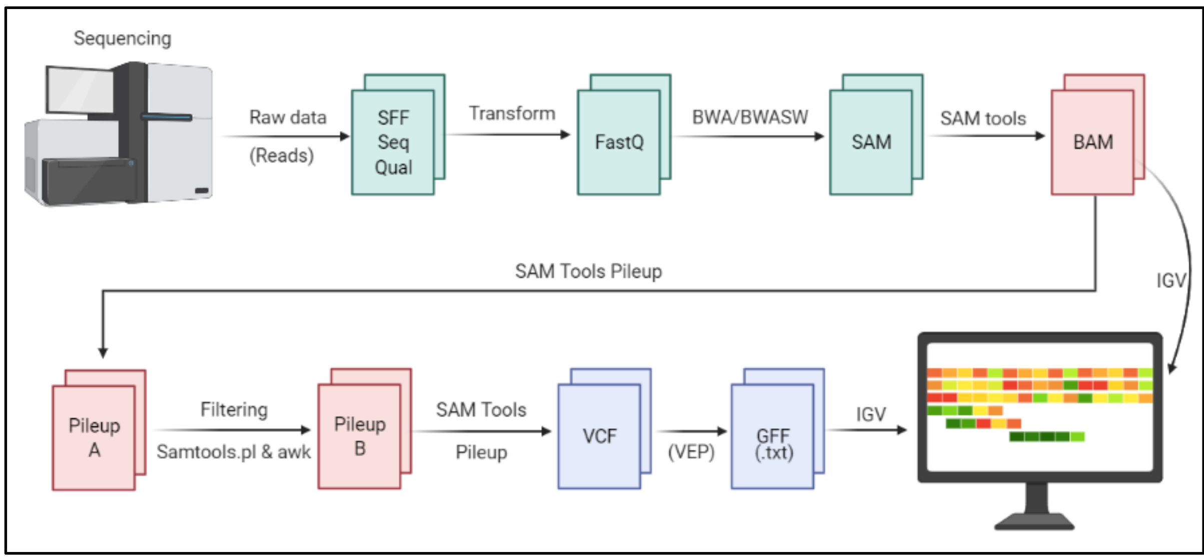


Figure 1.14: Targeted re-sequencing pipeline. *A schematic representation of a targeted re-sequencing pipeline, drawn by P. Ndibangwi using BioRender.*

1.1.2.4 Genome sequencing (GS)

The human reference genome contains about three billion nucleotides, and less than 2% (≈27000 genes) code for functional proteins (102, 103). GS is a technique used to sequence the genome (all genes and regulatory regions) without bias. It can detect variants in regulatory regions, but interpreting noncoding variants is still challenging. Compared to other NGS approaches, GS has more uniform exon coverage and is more sensitive in detecting rare exonic variants, although neither NGS approach fully covers their genomic targets (104). With the development of NGS, hundreds of genes linked to the pathogenesis of heritable CMOs have been identified; however, many causal genes remain undiscovered (105). As NGS technology advances, the cost of GS and its bioinformatic burden, particularly data storage and analysis, could become more affordable. However, the GS and its bioinformatics data storage and analysis load burden are still expensive.

Africa's diverse genetic backgrounds could contribute to unique health challenges, and GS can improve precision medicine efforts tailored to the region's specific genetic diversity (72). However, implementing GS on a large-scale is faced with challenges such as high costs, infrastructure limitations, and the need for specialised expertise. Moreover, addressing ethical concerns regarding data privacy, consent, and potential stigmatisation is essential when using genomic information in healthcare for diverse African populations (106).

1.1.2.5 Exome sequencing (ES)

The coding portions (less than 2%) of the genome are called the exome and constitute the genome's exons. ES is an alternative sequencing technique to GS that investigates the protein-coding regions, and it is relatively cost-effective compared to GS. ES has been used to identify about 26 putative new disease genes for genetic CMOs (105) and has yielded a genetic diagnostic rate of <60% for the various forms of CMO (87). For example, ES (targeted analysis of GS) has recently identified the *FLNC* gene as a significant cause of genetic CMO (39). Literature has shown that *FLNC* gene loss-of-function (LoF) variants have accounted for DCM in 1% to 4.5% of patients diagnosed clinically with the disease, in about 3.2% to 3.3% of patients with ACM, and *FLNC* missense variants have been observed in 1.3% to 9.7% of HCM cases (107). Another milestone achieved with ES was the complete gene sequence of titin in a shorted time compared to using the Sanger sequencing approach (108, 109). Titin (*TTN*) is a massive protein in vertebrate striated muscle that maintains myofibril elasticity and structural integrity. Titin's A-band region is known for causing adult-onset CMO, and one pathomechanism of A-band mutation-related COM is associated with dominant negative (110).

Despite being effective for gene discovery, ES coverage appeared inadequate for some genes, such as *TNNI3* and *PLN*, compared to the TS approach (111). Also, ES variant interpretation has remained an enormous challenge in large-scale studies (112). The increased number of genes tested using ES can increase the diagnostic yield. However, this comes at the cost of a longer turnaround time from sequencing and analytical processing, leading to higher identification of (VUS, which can increase clinical uncertainty (83).

An increase in ES clinical uncertainty compounded by the cost makes it more challenging to be implemented as a clinical routine diagnostic approach in Africa. In addition to Africa's diverse genome, there is also a lack of specialised skilled experts, few infrastructures, a high level of poverty, and mental health challenges (25, 83). Thus, the DGEMBE (Developing GEnomic Medicine BEtween Africa and the UK) program was established with a focus on CVDs to address some of the skills and infrastructural shortages in Africa.

1.1.2.6 Targeted sequencing (TS)

The targeted sequencing approach focuses on specific genes associated with phenotypes (in our context, DCM, HCM, RCM or ACM). The approach significantly reduces the sequencing

cost as well as the bioinformatics burden of data storage and analysis load compared to sequencing the exome or genome. For example, a study conducted in the UK examined the value of screening 51 non-sarcomeric candidate genes for HCM in 240 patients who were negative for P or LP sarcomeric variants using NGS. The study found only one additional variant, indicating that more extensive testing would have a negligible diagnostic yield in this setting. Furthermore, the increased cost and complexity associated with such testing would outweigh any potential benefits. Additionally, the study indicates the low risk of reporting uninterpretable findings to patients using TS compared to ES or GS, which could further complicate clinical certainty (113).

Similarly, extended gene panels in a Dutch DCM study resulted in more detection of (VUS without a corresponding increase in clinically relevant variants (114). Also, a strong preference for targeted panel tests over ES or GS has been published (115). Participants in the study indicated that they were more likely to choose panel sequencing over other types of NGS due to its potential for higher yield, lower cost, and fewer VUS. According to their preference, the participants perceive panel sequencing to offer better value for money in the clinical setting (105). However, when given the option of no testing, participants still preferred NGS testing, regardless of the type. One of the reasons could be the inflexibility of TS; it is restricted to a specific set of genes targeting a particular disease (116).

TS could be a cost-effective NGS strategy for less resource regions such as Africa. However, with the diversity of the African genome, TS might not effectively detect insertions, deletions, or other rearrangements (104). Although in the ClinVar database, about 95% of rare CMOs are detected through targeted re-sequencing (database searched, September 2023), most of the data were not for Africans. African populations have exhibited a very high genetic diversity, and rare pathogenic or novel variants could be missed by TS (117).

1.1.3 *In silico* prediction tools for cardiomyopathy

1.1.3.1 *In silico* prediction tools for cardiomyopathy

Notably, platforms such as ClinVar and other databases have been pivotal *in silico* tools in the prediction and characterisation of molecular genetics of Mendelian disorders. Online databases are utilised to predict the pathogenicity of variants reported following re-sequencing. Common databases and tools used to categorise variants in CMOs are listed in Table 1.2. The list of online tools can become exhaustive and time-consuming as the number of tools increases. If the tools are privatised, it can also become expensive. The categories of variants, database and prediction threshold are indicated in the first, third and last columns, respectively.

Table 1.2: Genetic variants *in silico* prediction tools

Category	Basis	Prediction database	Prediction Threshold
SNVs	Minor allelic frequency	gnomAD	<0.01 rare variant
	Percentage	CADD	>20 rare variant
Missense SNVs	Protein structure	Mutation taster	>0.5 disease-causing
	Function	PolyPhen2	≥0.85 damaging
	Evolutionary	SIFT	<0.05 deleterious
	Conservation	FATHMM	≥0.7 pathogenic
Nonsense	Function	DANN	>0.96 pathogenic
INDELS	Evolutionary	SIFT-INDEL	<0.05 deleterious
Splice-sites	Gene expression	GTEx, SpliceAI	≥0.7 pathogenic

1.1.3.2 Functional models for cardiomyopathy

1.1.3.2.1 Cardiomyopathy *in vitro* cell line models

For many years, cells and tissue culture models have proven useful in evaluating genetic variants' molecular consequences and protein function. Cardiac proteins involved in cell adhesion (e.g., desmosomes), ion transport, and communication have been candidates for these models. The available cell and tissue models include cardiac cell lines (such as H9C2), non-cardiac cell lines (such as HEK293), and induced pluripotent stem cells (iPSCs).

1.1.3.2.2 Cardiomyopathy *in vivo* animal models

Advances in CRISPR/Cas9, mutagenesis, bioimaging and genetic material purification have continued to improve our understanding and usage of animals for research. Using purified or synthetic genetic molecules in animals has yielded tremendous findings in medical research. Current animal models in biological studies, such as rats, mice, and zebrafish, have had high *in*

vivo experimental throughput. For example, our research group has used the zebrafish model to demonstrate the functional effect of the *POLG* c.2492A>G variant (118).

1.1.4 Treatment and management of cardiomyopathy

Without treatment, more than one-third of adults with heritable CMO do not survive more than five years after clinical diagnosis (55). Some patients have developed features of other heart defects and have poor prognoses. Others have developed potentially fatal HF, which has required heart transplantation at first presentation (43). The severity of the condition has varied among affected individuals, even among members of the same family. The disease is managed and treated as HF using medications such as beta-blockers or ACE inhibitors and lifestyle changes (12). Some patients have successfully prevented their phenotypes from becoming severe by eating healthy, avoiding the use of alcohol, exercising regularly, avoiding stress, and getting enough sleep. Early diagnosis of probands has been instrumental in helping at-risk relatives make informed decisions regarding their lifestyle (119). The success of CMO management in high-income countries like the UK relies on identifying and characterising diseases for personalised treatment and prevention.

We explore the various treatment options and management plans under the various types of cardiomyopathies in chapters 3 (DCM); section 3.1.7, chapter 4 (HCM); section 4.1.8, chapter 5 (RCM); section 5.1.8 and chapter 6 (ACM); section 6.1.8.

1.2 **RATIONALE**

The extraordinary advancements in molecular genetics research in the past 25 years have identified over a thousand mutations in several genes with different ontologies (120-122). These mutations reveal the different molecules and signalling pathways that lead to arrhythmogenic, dilated, hypertrophic, and restrictive cardiomyopathies (122).

However, these triumphs are overshadowed by a glaring failure. For the past two decades, the vast majority of cardiovascular genomics research has been Eurocentric, and this narrow focus on a small fraction of the world's population is leading to racial disparities in clinical genomics that will be further magnified as precision medicine becomes broadly adopted in the future (123).

Studies indicate that genetic data originating from various ancestries is valuable for therapeutic interventions and advances research (84, 124). According to a recent study using self-identified ancestry data, Black patients were found to have a higher estimated prevalence of familial DCM than White patients, and Black patients' first-degree relatives were found to carry a higher risk of developing DCM than White patients' relatives (84, 124). Additionally, Black African patients exhibited the disease earlier, were more prevalent, and had higher rates of morbidity and mortality (124, 125). This implies that the genetic cause of Black patients is either the same or greater than that of White patients (126).

Despite all the evidence, there is a concerning lack of genetic data for people of African ancestry, which restricts the ability of researchers in Africa to correctly interpret the pathogenicity of variants and counsel appropriately (125, 127).

Through our study called the Afr***i***can Cardio***m***yopat***h***y and My***o***carditis R***e***gistry ***P***rogram (IMHOTEP); HREC Ref. №. 766/2014), we intend to fill these gaps in knowledge and show the impact of the genetic landscape across the South Africa cardiomyopathy cohorts. We hypothesised that cardiomyopathies in the South African population have unknown genetic determinants. This study is part of the IMHOTEP project, a significant investigation into the incidence and genetics of cardiomyopathy in Africa. One of the main aims of IMHOTEP has been to characterise cardiomyopathies using cutting-edge genetics technologies such as next-generation sequencing (NGS), where entire genomes can now be sequenced by small laboratories, and researchers are able to study biological systems at a level never before possible.

1.3 **AIM AND OBJECTIVES**

1.3.1 Aim

This study aims to determine the genetic cause of cardiomyopathy in South Africa.

1.3.2 Objectives

The objectives set in this study have been to:

1. Identify the disease-causing variants and determine their frequency in probands with DCM, HCM, RCM, and ACM using a targeted re-sequencing strategy.
2. Use various bioinformatic pipelines to create a list of novel pathogenic variants in the South African population.
3. Determine the frequency and distribution of possible founder disease-causing variants in the DCM, HCM, RCM, and ACM cohorts.
4. Screen affected family members already in the IMHOTEP registry for variants identified in the probands in point 1 above.

Chapter 2: Methods

2.1 STUDY DESIGN

2.1.1 Overview

This molecular genetic research formed part of the Afr*i*can Cardio*m*yopat*h*y and My*o*cardi*t*is R*e*gistry ***P***rogramme (IMHOTEP; HREC Ref. №. 766/2014) study. The IMHOTEP study is a hospital-based, longitudinal observational study across multiple centres. It aims to identify the baseline characteristics and clinical course of various morphological types of heart muscle diseases, particularly emphasising primary (genetic) CMO. All patients who expressed a willingness to take part in the IMHOTEP study were required to sign a consent or assent form. The patients were then screened for eligibility at seven cardiac clinics in six centres across South Africa and one in Mozambique. At each clinic, the participants received special clinical investigations provided by at least one cardiologist following the ESC criteria. The clinical findings defined our participants' CMO phenotype (i.e., ACM, DCM, HCM, and RCM). The participants eligible for inclusion provided peripheral whole blood samples (5-25mL or 1-5 teaspoons) that we collected into either ethylenediaminetetraacetic acid (EDTA) or PAXgene blood DNA collection tubes. We stored the EDTA samples' buffy and the PAXgene whole blood samples at -80°C in the biobank repository at the Cardiovascular Genetics Laboratory, UCT. Genomic DNAs were extracted from the samples and used in high-throughput sequencing experiments to locate and classify variants in 38 genes associated with heritable cardiomyopathies. The phenotype and genotype results were correlated, and the findings were shared with the participants via the clinicians or genetic counsellors and the public via the UCT OpenAccess publishing platforms.

2.1.2 Inclusion and exclusion criteria

In addition to the IMHOTEP inclusion and exclusion criteria (Appendix A), participants were enrolled on this sub-study given that:

i. They were clinically diagnosed with unexplained cardiomyopathy.
ii. They were informed and consented to participate. The IMHOTEP team recruited consenting relatives to identify reliable genetic causes of the proband's disease.
iii. Once off, enough peripheral blood samples were obtainable.

However, we excluded participants when:

i. Lost during follow-up.
ii. Unable to collect a rebleed sample or rebleed samples were unobtainable.
iii. They decided to unenroll from the IMHOTEP project.

2.1.3 Sample size

We designed this study primarily to ascertain the genetic determinants of familial or sporadic cardiomyopathies in South Africa. Specifically, to identify the genetic variants that could explain the observed CMO phenotypes. Our ascertainment strategy ensured that a minimum number of 600 probands (i.e., with ACM (n=100), DCM (n=300), HCM (n=100), or RCM (n=100)) were recruited. These numbers were based primarily on a feasibility study in the vanguard phase of the IMHOTEP project over three years with finite resources. This thesis had a sufficient sample size required to address the question of the frequency of familial CMOs in South Africa.

2.1.4 Enrolment of participants

We enrolled 719 consecutive multiracial participants from seven sites. At the commencement of this study, the participants' recruitment register, informed consent or assent forms obtained before enrolment into the IMHOTEP registry were checked and verified (Appendix B). We scanned the quick response (QR) coded stickers with study identification details fixed onto the forms (and samples) into an Excel spreadsheet. We then crossmatched the scanned information to the stored samples' identification and pedigree charts. Finally, accepted samples were selected and thawed gradually, and we extracted their genomic deoxyribonucleic acids (gDNA) at room temperature.

2.2 ETHICAL CONSIDERATION

We performed this study in compliance with the Declaration of Helsinki and good clinical practice principles. Also, we obtained Ethical clearance and accounted for other considerations, as stated in section 2.2.

2.2.1 HREC approval

Ethics approval, **HREC REF:009/2020**, was obtained from the UCT Faculty of Health Sciences Human Research Ethics Committee (FHS HREC) before the commencement of this study. The progress of this sub-study was reported to the HREC annually for the duration of this research. Besides, there were no risks to the participants already recruited into the IMHOTEP main study (HREC Ref. №. 766/2014). However, the leading recruitment site, Groote Schuur Hospital, was responsible for indemnity under the UCT Federalwide Assurance (FWA) to protect human participants.

2.2.2 Informed consent

The IMHOTEP team collected informed consent (and assent) forms meticulously. All forms were reviewed prior to the molecular experiments.

2.2.3 Data management plan (DMP)

Our data management, sharing and preservation strategy was created using the UCT online generic template of DMP (https://dmp.lib.uct.ac.za/plans/1537). The full DMP contained ten sections with 21 questions that dealt with:

i. Data collection and the estimated size of the dataset
ii. Repositories where the final data were stored.
iii. Storage requirements for all stages of the lifecycle of the data
iv. Budget considerations for data management and curation
v. Metadata standards used.
vi. Roles and responsibilities
vii. Relevant policy documents

2.2.4 Other ethical considerations

The investigators of this study considered the incredible impact of genetic testing on both the proband and their family members. Hence, we pledged to communicate the positive genetic result with the participants' clinician and genetic counsellor immediately after we agreed with the finding. Our clinical team was mandated and could inform the affected individual and the family. Furthermore, scientific evidence has shown that family screening provides the opportunity to offer diagnosis, risk stratification and preventive intervention early in the disease process.

During the COVID-19 pandemic, we adhered to all protocols and national recommendations to minimise the spread of the virus.

Finally, we made this thesis available to the scientific community through the UCT publishing platforms.

2.3 MOLECULAR METHODS, DATA GENERATION AND PROCESSING

We have schematically presented the methods used to generate and collect vital clinical and genetic data for applying the molecular genetic findings in managing and treating our patients in Figure 2.1.

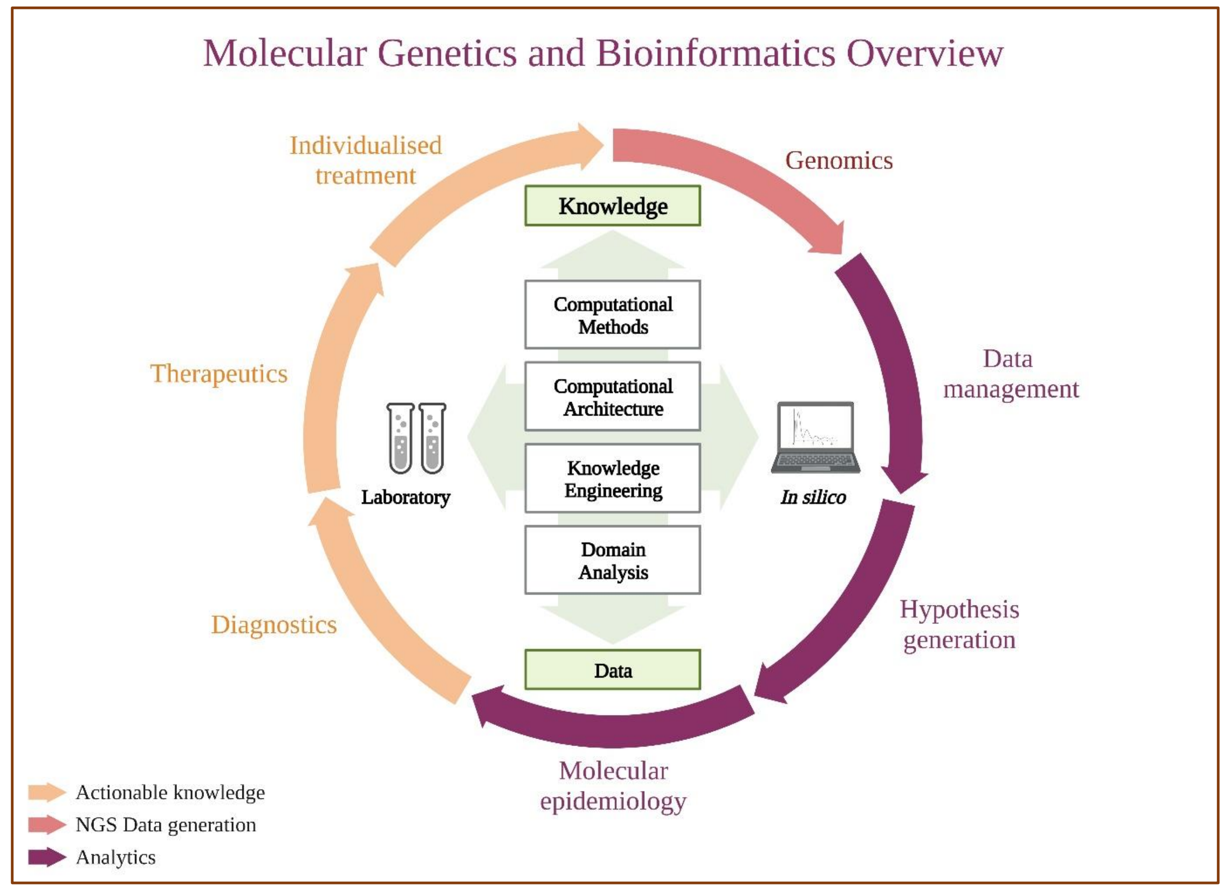


Figure 2.1: An overview of the clinical application of this study. *Drawn by P. Ndibangwi using BioRender.*

2.3.1 Baseline data

Probands coded consenting documents, pedigree charts and requests for molecular test forms from February 2015 to October 2021 were accessed. The probands' CMO phenotype, sex assigned at birth, age at diagnosis, self-reported ancestry, and genetic history were retrieved. All information was captured and used for descriptive analysis in Excel or R and RStudio. We also used pedigree charts to select families for segregation analysis and screening. The clinical evaluation notes retrieved were the family medical assessment notes, provisional diagnoses, and any other special enquires.

2.3.2 Genomic deoxyribonucleic acids (gDNA) extraction

We used the samples' storage catalogues to locate tubes with frozen (-80ºC) buffy or whole blood for the probands. The samples were allowed to thaw gradually from -80ºC to -20 ºC to 4 ºC and at room temperature. The samples' gDNA were extracted from the buffy or whole blood leukocytes following the published manufacturer's methods for Puregene Blood Core Kit C (Appendix C) or PAXgene Blood DNA kit (Appendix D), respectively. The raw gDNA extracts were suspended in a hydration solution overnight at room temperature to dissolve and homogenise. We checked all raw gDNA sample quality and computed the results in an Excel spreadsheet stored in the UCT research data database. Where initial samples were insufficient or failed to give good quality gDNA, we contacted the participants for consent, and once off-rebled, whole blood tissues were collected.

2.3.3 Raw gDNA quality and integrity checks

a) gDNA quantification

After extracting the DNA, we used one micro-litre (µL) of every homogeneous raw gDNA to determine the sample's concentration. We followed the manufacturer's method using a calibrated NanoDrop^{TM} 2000 spectrophotometer (Thermo SCIENTIFIC) with software v2.6 on a Windows 10 computer for the quantification. The NanoDrop software measured the samples' optical densities at 230nm, 260nm, and 280nm wavelengths and used the readings to determine the samples' purity. An A260/A280 (1.76-1.83) or A260/A230 (1.9-2.1) ratio in the ranges indicated confirms pure raw gDNA with an insignificant quantity of proteins or organic compounds (such as residual phenol or carbohydrates), respectively. We confirmed the intactness of good quality uncontaminated samples using agarose gel.

b) gDNA intactness

When we were satisfied with the purity and quality of the raw gDNA, we then used a 1% (w/v) agarose gel electrophoresis image to visualise if the samples were damaged. We prepared a 1% gel using 1g agarose powder mixed with 100 mL of 1X TBE buffer in a 250 mL conical flask. We heated the mixture for the agarose to dissolve fully, cooled it to about 50°C and poured it into a casting chamber. After 25 minutes, we loaded the set gel into the chamber. We added a

mixture of 3µL gDNA sample or molecular marker (1kbp DNA ladder, Fermentas) and 5µL loading dye (Appendix E) into the wells. The gel electrophoresis was programmed to run at 110V for 75 minutes. The gel was visualised in a Uvitec Xplorer D55 gel doc using an Xplorer 1D software version 15.08 (Uvitec Cambridge, Cambridge, UK) on a computer operating with a 2007 Windows XP. All gel images were saved in BMP format, and damaged or degraded samples should form a smear in the image.

c) Raw gDNA dilution (sample prep)

We used a portion of the sample for every pure and intact raw gDNA to make a working dilution. We diluted the sample with nuclease-free water (omega, BIO-TEK) to a volume (V_f) of 200µL and a final concentration (c_f) of ≥120ng/µL. We used the formula $c_iV_i=c_fV_f$ and the NanoDrop readings [initial concentration (c_i)] to determine the initial volume (V_i).

The quality and integrity of the dilutions were checked (same as in sections 2.3.3 **a** and **b**). Eventually, we made a 75µL aliquot of each checked diluted sample. The aliquots were packaged in 96 well combo plates (QR coded), put on dry ice and sent to the Oxford Medical Genetics Laboratories (OMGL) or Wellcome Trust Centre for Human Genetics (WTCHG) via TNT (FedEx) courier company.

The remaining portion of each raw gDNA and left-over quality-controlled dilutions were catalogued and stored at -80°C at the CVG Molecular Genetics Laboratory, UCT, for any subsequent experiment such as Sanger sequencing.

2.3.4 Next-generation sequencing (NGS)

At the OMGL, the diluted gDNA samples were prepped for NGS using custom-designed kits targeting the coding regions and exon-intron boundaries (± 10bp) of 38 genes (Section 1.3) for sequencing on an Illumina NGS platform. Figure 2.2 illustrates an overview of the NGS workflow used.

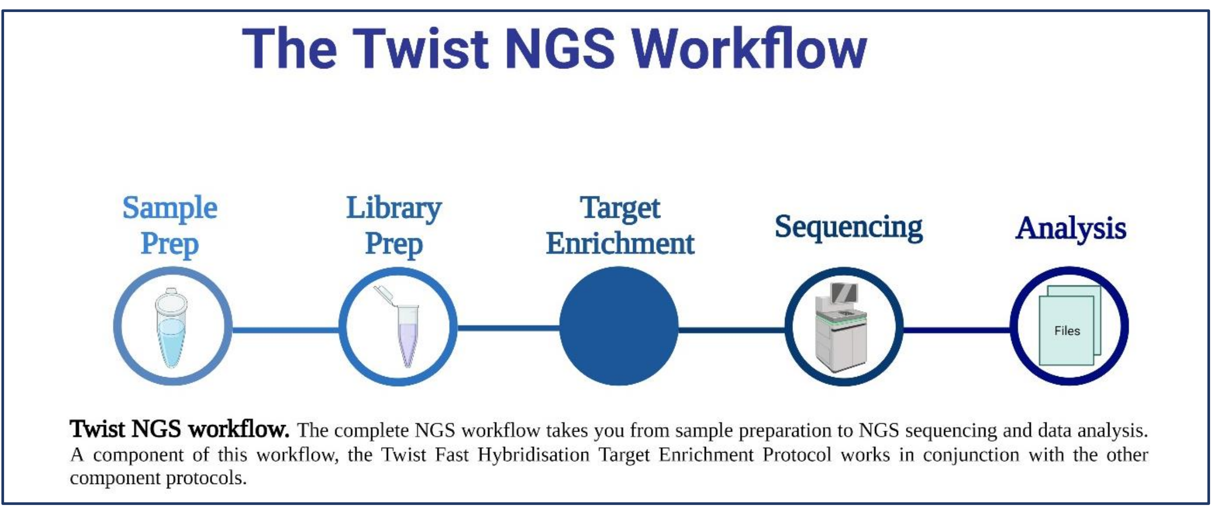


Figure 2.2: Next-generation sequencing workflow. *The next-generation sequencing workflow used in this study was drawn by P. Ndibangwi using BioRender.*

The Illumina NGS components' protocols presented in Figure 2.2 have been described briefly:

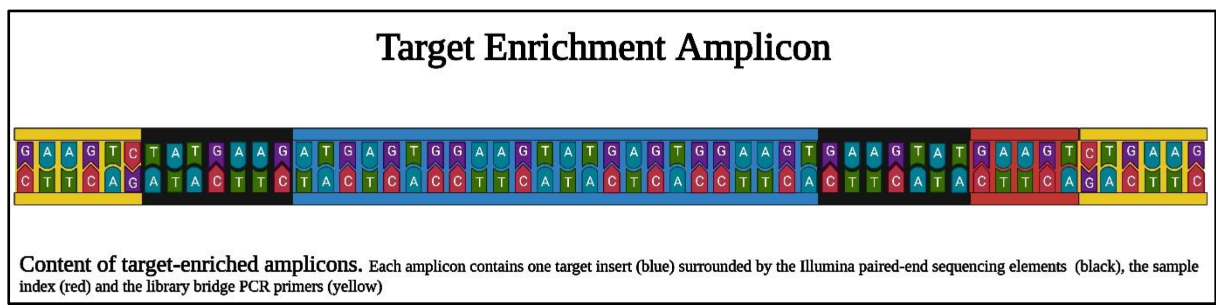
a) Library prep and target enrichment

To convert gDNA into libraries for Illumina sequencing platforms, the OMGL used a human core Exome EF multiplex complete kit and an Agilent HaloPlex Target Enrichment System kit manufactured by TWIST Bioscience Corporations and Agilent Technologies, respectively. The manufacturers designed the kits to create gDNA libraries specific to 38 CMO disease-causing genes, listed in Appendix F. The targeted panel kits utilised enzymatic fragmentation and synthetic probes to capture these 38 specific gene sequences.

b) To create an enriched target human gDNA library

The steps followed according to the kit manufacturers' method to create an enriched target are (128): i) digest the gDNA with 16 different restriction enzymes to create a library; ii) hybridise the gDNA restriction fragments collection to a synthetic probe capture library for target enrichment and sample indexing; iii) capture the circularised target gDNA-probe hybrids, containing biotin on streptavidin beads; iv) add DNA ligase to the capture reaction to close nicks in the circularised probe-target gDNA hybrids; v) elute the captured DNA libraries after ligation with freshly prepared NaOH; vi) amplified the captured target libraries using a PCR master mix and program; vii) purify, validate and quantify samples for sequencing. Enriched library samples were qualitatively analysed using gel electrophoresis. Figure 2.3 shows the

content of a target-enriched amplicon and viii) pool purified target-enriched sample with different indexes for direct sequencing.



Content of target-enriched amplicons. Each amplicon contains one target insert (blue) surrounded by the Illumina paired-end sequencing elements (black), the sample index (red) and the library bridge PCR primers (yellow)

Figure 2.3: Target enrichment amplicon. *The schematic representation of a target enrichment amplicon used for next-generation sequencing was drawn by P. Ndibangwi using BioRender.*

c) Re-sequencing

The target-enriched libraries containing exons of the 38 genes and their exon-intron boundaries (± 10bp) were batched and sequenced. Standard Illumina paired-end primers and chemistry (an Agilent's HaloPlex Target Enrichment on an Illumina® MiSeq system at OMGL or a TWIST molecular autopsy panel on NovaSeq system at the WCTHG) were used for the sequencing by synthesis. The chemistry enabled a sequencing coverage of 99.5% at 1x, 99.3% at 20x or more, and a high concordance of single nucleotide variants to the dbSNP database. The sequencing machine output raw data were processed using an in-house automated pipeline that fetched the data from the sequencer, aligned and mapped the reads to the human genome reference sequence GRCh37/hg19.

2.3.5 Processing NGS data using ***H***aloplex ***A***nalysis ***P***ipeline in ***Py***thon (HAPPy)

The sequencing reads generated by the Illumina® MiSeq or NovaSeq systems were processed using an in-house automated ***H***aloplex ***A***nalysis ***P***ipeline in ***Py***thon (HAPPy). The HAPPy automatically fetched raw FASTQ reads from the OMGL MiSeq or WCTHG FTP site and copied them to a Linux server before execution. We enlisted an independent checker to ensure data were submitted to HAPPy correctly. The check occurred in two parts: to match the worklist file contents to the MiSeq/NovaSeq Run Worksheet and the backup directory to the worklist filename and the MiSeq/NovaSeq flow cell identifier.

PuTTY and FileZilla were used to run the generated data in the in-house automated pipeline, HAPPy. The first few steps of the sequencer's analysis pipeline were to (*i*) identify the indexes, Figure 2.3; (*ii*) trim indexes and paired-end sequencing elements, and (*iii*) remove low-quality

reads. Good quality paired-end reads were aligned to the reference human genome (GRCh37/hg19) using the SAM toolkit software to create BAM files. The tools calculated the coverage of all targeted regions; bases with less than 15 reads were considered "low coverage". VCF toolkits (GATK v3.8-1-0 and Platypus v 0.8.1) were used for variant calling to convert the BAM files into VCF files. The BAM, VCF, mutation reports, and other files generated were saved on the server. The files were loaded into the respective software platforms as and when required for storage, transfer, and variants investigation. To investigate the variants, we started by annotating the VCF files generated in machine-readable formats using the Ensembl variant effect prediction (VEP) tool to make them human-readable. The VEP pipeline identified single-base substitutions and short insertions/deletions.

2.3.6 *V*ariant *e*ffect *p*redictor (VEP) and analysis pipeline

To annotate the hg19 VCF files using VEP, we uploaded the samples into the UCT HPC cluster and converted files into human-readable text files (Appendix G). To achieve this, we used (*i*) a Linux *sbatch* command to run an Ensembl VEP shell script to annotate the VCF files, (*ii*) a *Perl 5* script to convert the output VCF format to a text file while retaining the columns of interest, (*iii*) another *Perl 5* script to annotated the output text files containing all variants for each individual, (*iv*) used another *Perl 5* script to combine all the samples, (*v*) used another *Perl 5* script to create a summary of variants per gene, and (*vi*) downloaded the VEP output text files on to the UCT research data database for further filtering using a specific set of criteria.

2.3.7 Variant filtering and interpretation

We selected variants for further analysis using the American College of Medical Genetics and Genomics/Association for Molecular Pathology (ACMG/AMP) set of filtering criteria, Figure 2.4. We selected (*i*) variants based on protein consequences (e.g. start lost, stop gained, to name a few), (*ii*) variants with a gnomAD minor allele frequency (MAF) of <1%, (*iii*) variants predicted by bioinformatics tools such as SIFT and PolyPhen2 to probably have damaging or unknown *protein* function and (*iv*) variants classified by Alamut™ Visual Plus version 1.6.1 and ClinVar (2022-12-17) as damaging. Finally, we classified the variants according to the

refined ACMG guidelines criteria as class 1 (B), class 2 (LB), class 3 (VUS), class 4 (LP) and class 5 (P).

We created a list of class 3, 4, or 5 variants with LoF (null known, known variant, or in-frame non-repetitive), *de novo* inheritance, or in protein domain hotspots. The ACMG/AMP rule abbreviations are provided in supplementary (SS) Tables 1 and 2.

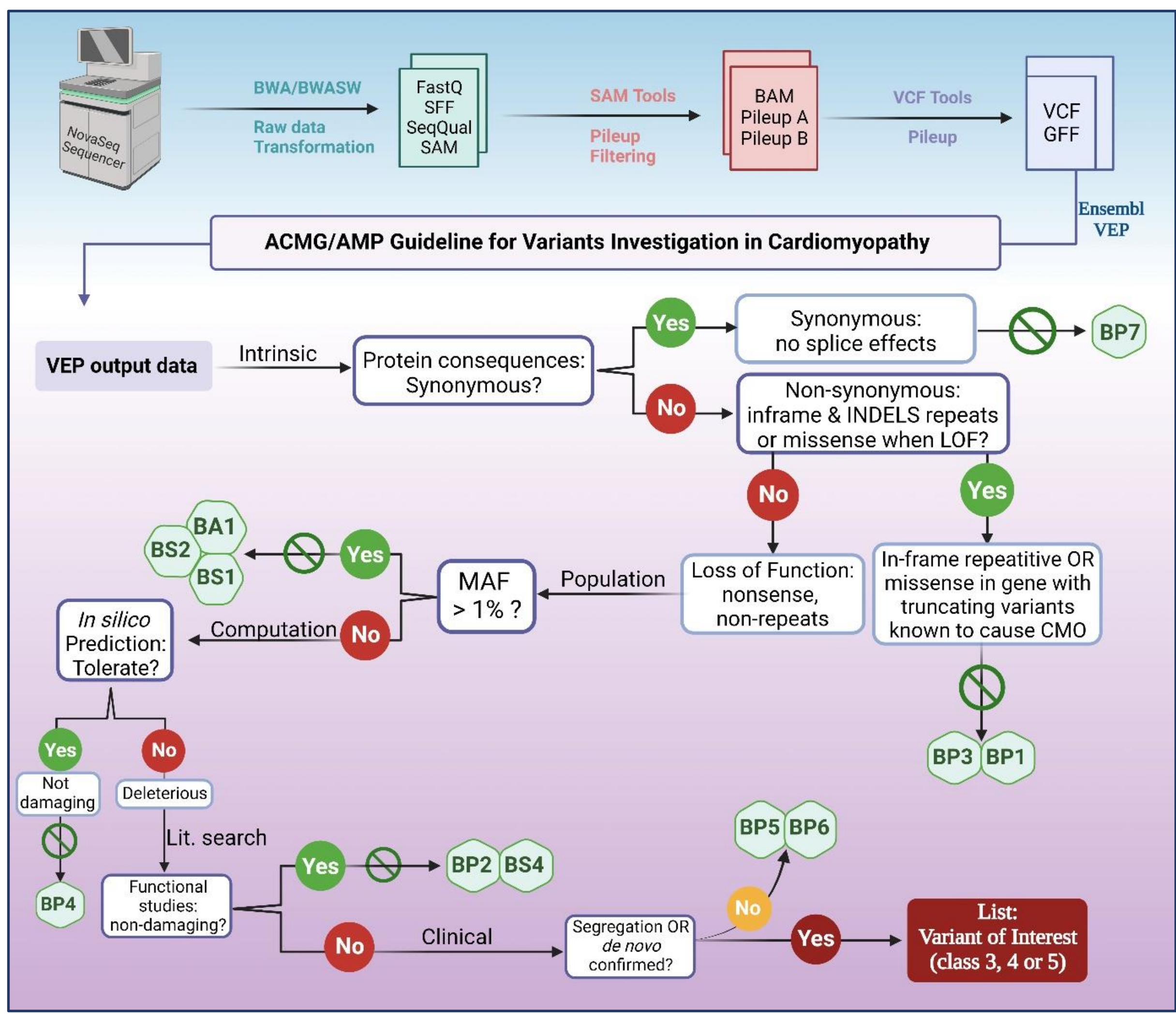


Figure 2.4: Variant investigation strategy. *Schematic representation of variants filtering and classification strategies, drawn by P. Ndibangwi using BioRender.*

We also looked at the class 4 and 5 variant caller scores for Haplocaller and Platypus for threshold values above 400 and 700, respectively. Further, we searched current published data intensively to support variants' pathogenicity, visualised the variants' BAM files in the Alamut^{TM} Visual Plus to eliminate false positives, summarised variants per gene and individual, regularly conducted in-house discussions to correlate genotype to the observed

phenotype, and recommended variants for segregation analysis, family screening and founder variants investigation.

We recommended further molecular genetic experiments for the probands' samples, in whom a class 4 or 5 variant was unidentified. We recommended some selected families (i.e., trios and large) for ES to maximise the identification of reproducible genetic causes of cardiomyopathy. Finally, the variants categorised as class 4 or 5 were validated using Sanger sequencing.

2.3.8 Deleterious variants validation, relative screening, and segregation analysis

We processed the samples of probands carrying class 4 and 5 variants via Sanger sequencing. We included the relatives' samples (with consent for molecular genetics study) enrolled on the IMHOTEP study for screening purposes. New primer sets were designed and synthesised for the Sanger sequencing during screening when required.

a) Primer's design

We downloaded the genes' regions with variants of interest from the Ensembl GenBank database. Primer pairs were carefully chosen at least 60bps before and after each variant location. We used the IDT oligoAnalyzer tool to pick primers (amplicon size between 150 and 380 bps). For details of this process, refer to CVG SOP: '*Primer Design*' (Appendix H). The Inqaba Biotechnical Industries PTY Ltd and Whitehead Scientific synthesised our primers. Refer to '*Primers List*' (supplementary sheet 3) for detailed primer pair sequences and information.

Polymerase chain reaction (PCR) optimisation for optimum magnesium and temperature conditions was not required.

b) High-Resolution Melt (HRM)

We used the PCR to create amplicons for post-PCR HRM reactions on Rotor-Gene™ 6000 (Corbette Research, Celtic Molecular Diagnostics) real-time PCR machine. In a 25µL PCR reaction mix, we amplified 120ng/µL of gDNA using 0.5µL of each primer (20µM) and 23µL 1X master mix (For the PCR master mix reagents' concentration, volume, and reaction

conditions details, refer to Appendix I, '*HRM 1X master mix*'). First, the PCR reaction mix was incubated at 95ºC for 10 minutes for the gDNA to denature and followed by 50 cycles of 95ºC for 5s, 50ºC for 10s and 72ºC for 10s.

After the last circle of the PCR, we generated melt curves by ramping the PCR reaction mix from 72ºC (90s wait pre-melt) to 95ºC, at +0.1ºC increment (2s wait after each step). In the reaction, double-stranded (ds) DNA intercalated with a fluorescent dye (EvaGreen™) heats up, and as it melts into single-stranded DNA, the dye falls off. The machine measured the decreasing fluorescence and plotted the readings on graphs. We analysed the melt curves to determine the genotype of the probands (and their relatives when applicable) according to the Rotor-Gene manufacturer's manual. For routine diagnostic reporting, the samples were purified and used for Sanger sequencing.

c) Sanger sequencing

A clean-up experiment was performed to purify the samples for Sanger sequencing. A 20 µL clean-up reaction was set up in a microtube (Whitehead Scientific) containing 5 µL PCR products and 15 µL master mix (13.9µL NFW plus 1 unit of Shrimp Alkaline Phosphatase plus 0.1unit of exonuclease). Each reaction tube was incubated in an Eppendorf thermal cycler at 37ºC for 60 minutes and 75ºC for 15 minutes to digest the left-over products and inactivate the enzymes. During incubation, the exonuclease and phosphatase removed left-over primers and dNTPs.

We used the clean-up products to perform a Big-Dye v3.1 re-sequencing reaction. A 20 µL Big-Dye sequencing reaction was set up in a microtube (Whitehead Scientific) containing 15µL master mix (9 µL NFW plus 4 µL buffer (5X) plus 2 µL Big-Dye), 3 µL clean-up product and 2 µL primer (sense or antisense). Each reaction tube was incubated in an Eppendorf thermal cycler to run at 1) at 96ºC for 5 minutes, 2) at 25 cycles of 96ºC for 30 seconds, 50ºC for 15 seconds, 60ºC for 4 minutes, and 3) at 4ºC for 30 seconds. We programmed the machine to idle at 4ºC until the reaction tubes were removed.

The Stellenbosch University central analytical facilities (CAF) or Inqaba Biotech used the Big-Dye reaction products created to generate electropherograms in sequencing electrophoresis-only reactions. We visualised the electropherograms to confirm and validate the variants in Alamut Visual Plus version 1.6.1 or BioEdit sequence alignment editor version 7.2.6.1

(https://bioedit.software.informer.com/7.2/). Finally, we held in-house discussions to determine if the genotype-positive results could explain the observed phenotype in each participant.

2.3.9 Genotype-phenotype correlation

We implemented the ACMG/AMP guidelines to generate secondary data on an Excel spreadsheet indicating the variants' status (class 1 to 5). We appended the proband's clinical information, the variants coding sequences (CDS), reference SNP cluster identifier (rsID), gene symbol and protein changes to the Excel list. The variants were categorised using the ACMG status. This classification separated the probands into those with disease-causing variants (class 4 or 5), those with variants of uncertain significance (class 3), and those with unknown genetic causes (class 1 or 2). Published evidence, such as in ClinVar and gnomAD v2.1, and the ACMG rules were recorded for the variants. In addition, we explored the proband's demographics, pedigree charts, clinical notes, and variant status to find explainable links between the genotype and phenotype. Notably, novel class 4 and 5 variants with unclear phenotype correlation were listed for detailed segregation and other molecular studies, e.g., using either cell line or zebrafish model. Also, we recommended variants found in multiple carriers for founder variants analysis.

2.4 STATISTICAL ANALYSIS

We performed a descriptive analysis of cardiomyopathy's baseline characteristics and genetic outcomes (n = 690 probands). The data presented in this thesis were analysed using the UCT HPC cluster, Microsoft Excel 365, and R and RStudio (version 4.1.2). In RStudio, base R, *tidyverse*, *finalfit*, *lessR* and *knitr* R packages were used to analyse and fit logistic regression models. We presented the final results as mean plus and minus one standard deviation ($\boldsymbol{\bar{x} = \pm 1\delta}$) [or median with the interquartile range (IQR)] for numerical variables and total counts with percentages (or proportions) for categorical variables.

2.5 STUDY MEASUREMENT AND ENDPOINT

Identifying a definite disease-causing variant that could explain the observed phenotype marked the endpoint. Thus, all novel and known class 4 and 5 variants were validated using the Sanger sequencing method to confirm their presence; known polymorphisms were not confirmed by Sanger sequencing. Larger families were used to assess deleterious variants' pathogenicity, and we recommended novel class 4 and 5 variants in singletons for further molecular functional studies. The findings of the known class 4 and 5 variants were reported and discussed with the clinical team for diagnostic reporting purposes.

Chapter 3: Dilated Cardiomyopathy (DCM)

3.1 **INTRODUCTION**

According to the Global Burden of Disease survey, there are approximately 2.5 million cases of cardiomyopathy (CMO) worldwide (129), and dilated CMO is the second most frequent cause of cardiac arrest and cardiac transplantation, preceded only by coronary artery disease. The global burden of CMO (mainly DCM) varies according to geographical regions; for example, Figure 3.1 illustrates the 2019 prevalence. Studies from high-income nations like the UK, the United States of America (USA), and the Netherlands have reported about 20% to 50% of occurrences of CMO and approximately one-third of cases of (HF attributed to DCM (57, 58, 95).

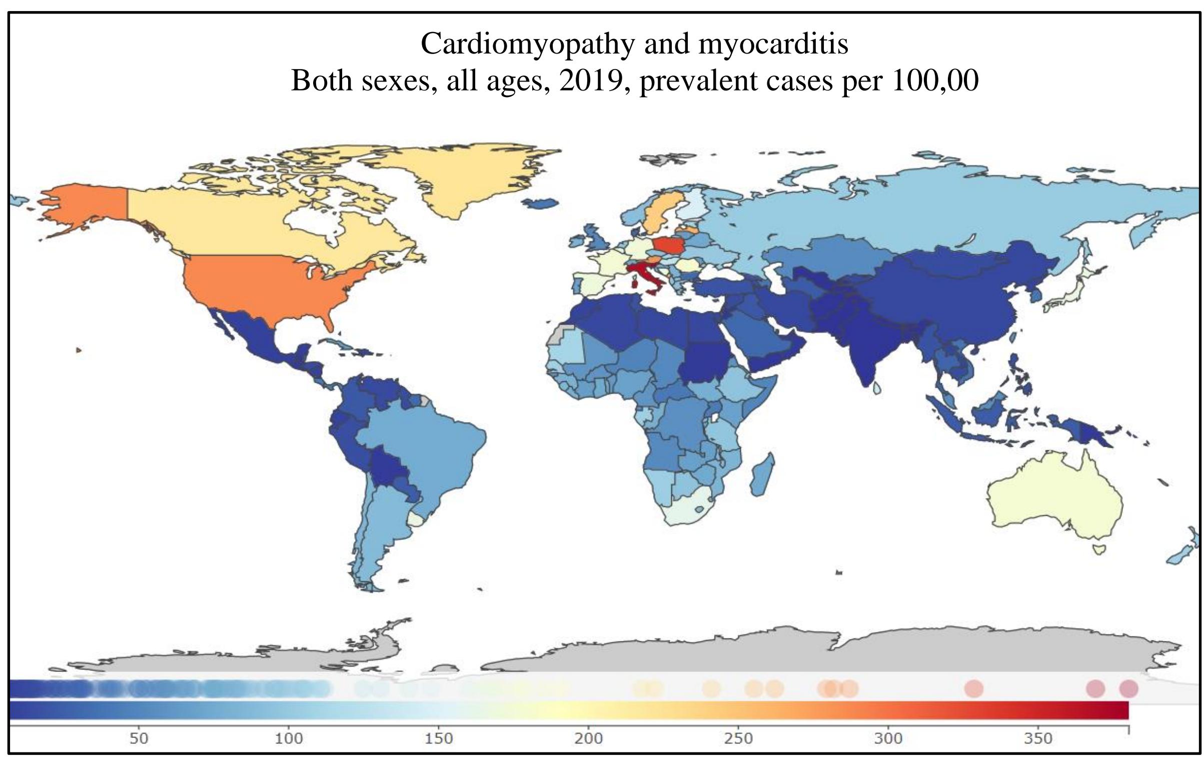


Figure 3.1: Global burden of cardiomyopathy. *The 2019 Global Burden of Disease study estimated the world's prevalence of cardiomyopathy and myocarditis at 2.5 million, but there is no data for Africa. Drawn by P. Ndibangwi.*

3.1.1 Clinical background

DCM is a disease where the heart cannot maintain normal blood flow because the left ventricle is enlarged and weakened. The left ventricle is the heart's main pumping chamber, and its enlargement impairs the systolic function without coronary artery disease (CAD),

hypertension, valvular disease, or congenital heart disease (CHD). DCM is a non-ischaemic (decreased heart function not caused by heart attack or blockages in arteries) disorder that causes mechanical or electrical dysfunction in the heart muscle. DCM occurs due to the heart's left or biventricular enlargement that impairs systolic function, and the major complications include ventricular arrhythmia, HF, or (SCD. DCM present differently in adult and paediatric cases, and we first reviewed the adult-onset disease and followed this up with the paediatric onset.

3.1.2 Signs and symptoms

3.1.2.1 Adult-onset DCM

DCM's most common signs are reduced ability to perform physical activities and oedema of the abdomen, ankles, legs, and feet. The symptoms of DCM include fatigue, dyspnoea, chest pains, and palpitations. Other symptoms might consist of syncope, symptoms of arrhythmias, HF or sudden death. The DCM clinical diagnosis shown in Table 1.1 section1.2.5 has involved the use of 1) 12-lead Electrocardiography ((ECG) to detect late potential on signal-average ECG; Epsilon wave; left or bi-atrial enlargement; atrial fibrillation; Abnormal Q waves; Left bundle branch block (LBBB); ventricular ectopy, 2) echocardiography to reveal LV end-diastolic dimension >112%; LV enlargement and wall thinning; global LV hypokinesia and systolic impairment; functional mitral regurgitation, and 3) cardiac magnetic resonance imaging (MRI) to indicate LVEF of <50% diastolic dysfunction, and LVEF is more predictive of DCM mortality (49).

3.1.2.2 Paediatric DCM

Paediatric DCM is a rare form of heart muscle disease that results from intrinsic or extrinsic insults to cardiomyocyte mechanics, calcium signalling, and downstream intracellular signalling pathways (130). Paediatric DCM has a spectrum of signs and symptoms, such as inability to perform physical activities, chest pain or oedema (131). Other patients are asymptomatic cases diagnosed incidentally and could require a heart transplant at first presentation (132) or risk SCD.

3.1.3 DCM disease burden

3.1.3.1 Adult-onset DCM

DCM has an estimated global incidence of 1:500 to 1:250 in the adult general population, and patients with established genetic aetiologies are known for developing HF or requiring heart transplants (133). SSA studies have attributed over 76% of adult-onset populations to CMO, mainly DCM (14). However, extensive studies have yet to be conducted on DCM in Africa, leaving significant gaps in our knowledge. Most African DCM studies have been hospital-based, showing that 17% to 48% of patients hospitalised with HF in the SSA have had the disease (134). In South Africa, hospital-based studies found that 26.6% to 28% of hospitalised patients had idiopathic DCM (15). Additionally, in the South African adult-onset population, CMO incidence, mainly DCM, is estimated to have reached 35.3% (15); however, the incidence of some DCM phenotypes, such as the peripartum, is unknown.

3.1.3.2 Paediatric DCM

Paediatric DCM results in significant morbidity and mortality (42, 132). The disease prevalence data are inconsistent, such as in Australia (132), but prevalences of 1 in 2700 have been reported in the USA and, 1 in 7000 in Japan (86), 2.6 cases per 100 000 children in Finland (131). Paediatric DCM has also been estimated to have an international incidence of 0.57-1.5 per 100,000 children (135), and over 50% of the patients have had familial DCM (42). In the SSA region, studies have attributed over 20.8% of certain paediatric populations to CMO, mainly DCM (14). In one paediatric study from Egypt, 25 of the 68 paediatric patients analysed had positive genetic findings for DCM (136). However, more research on paediatric DCM in Africa is needed as no extensive studies have been done. The incidence of paediatric DCM phenotypes in South Africa is unknown (14).

3.1.4 Aetiology of DCM

3.1.4.1 Adult-onset DCM

The natural progression of DCM is significantly influenced by genetic predisposition as well as environmental factors (137). DCM has diverse aetiologies; for example, *i*) infectious pathogens such as viruses, bacteria, protozoa, helminths, and fungi may cause inflammatory DCM; *ii*) exposure to toxins such as alcohol, amphetamines; anthracyclines, cannabis, catecholamines, cocaine, 5-fluorouracil, lithium, heavy metals (cobalt, lead, and mercury) and

carbon may lead to secondary DCM, and *iii*) metabolic or endocrine dysfunction or neuromuscular diseases or pregnancy such as PPCM (57).

The correlation between these factors, such as alcohol consumption or low socioeconomic status and risk alleles, remains understudied in Africa (25, 138). Studies show that alcohol exposure can trigger the onset of DCM in patients with rare genetic mutations (e.g., titin gene variants) associated with the condition (54, 139). Epigenetic interactions with the environment can also impact gene expression. Thus, a better understanding of the environmental and epigenetic factors can help improve the assessment of familial DCM (140).

Also, researchers have shown that up to half of the DCM cases globally have rare variants (126, 141), and the idiopathic forms in adults are the second most common cause of HF in Africa (57). DCM with genetic causes manifests as severe morpho-functional abnormalities that lead to substantial morbidity and mortality or heart transplantation (139).

3.1.4.2 Paediatric DCM

The spectrum of DCM with genetic causes in paediatrics has a wide range of determinants and has varied from birth to adolescence (42). About 30% to 50% of paediatric DCM cases have an identifiable aetiology (130) encompassing syndromic, metabolic, and neuromuscular causes (42, 96). The aetiology of paediatric DCM, particularly genetic causes, is more heterogeneous.

3.1.5 Types of DCM

3.1.5.1 Genetics of classified DCM

3.1.5.1.1 Adult-onset DCM

Most DCM patients may transmit the disease in all modes of inheritance pattern, but the most common inheritance pattern is autosomal dominant (AD) (60). Researchers have discovered DCM-causing variants in over 250 genes spanning >10 gene ontologies (120, 121). Most reported DCM-causing genes encode sarcomere (e.g., titin (*TTN*)), cytoskeleton (e.g., filamin C (*FLNC*)) and nuclear envelope lamin A/C (*LMNA*) proteins. These protein structures are essential in maintaining cardiomyocytes' structural integrity, communication, force generation and transmission. Plausible (likely) disease-causing variants in any of these protein-coding sequences may affect the protein's expression level or change their typical structural conformation, impairing cardiomyocyte proteins' function. Poor-quality protein (e.g.,

shortened or misfolded) will cause the protein products' haploinsufficiency or disrupt various signalling pathways or myocyte disarray and collagen deposition. These disruptions impair myocardial force generation, increase susceptibility to arrhythmias and abnormal remodelling, or disrupt cardiomyocyte alignment, promoting abnormal electrical conduction. Hence leading to arrhythmias or diastolic dysfunction and subsequently causing HF.

The genetics of heritable DCM in Africa have been reviewed recently (14, 57), indicating the paucity of data, for example, the lack of routine diagnostic yield data. However, research data mainly from high-income countries have shown that more than 50 high-impact heritable DCM-causing genes have been discovered, and these genes are being used in these countries for primary DCM routine genetic testing. In these countries, routine genetic testing for DCM has reached a diagnostic yield of about 20% to 50% (142). Table 3.1 contains the genes most commonly responsible for heritable DCM in these high-income countries.

Table 3.1: The most reported heritable DCM-causing genes

Gene Symbol	Function	Prevalence (%)	Ref
TTN	Sarcomere	12 - 44.2	(95)
MYH7	Sarcomere	4 - 20	(98)
LMNA	Nuclear envelope	4 - 19.4	(95)
RBM20	Nucleus	2 - 10.1	(95)
PLN	Calcium regulatory protein	Rare - 9.7	(95)
SCN5A	Cardiac sodium channel	2 - 9.5	(143)
FLNC	Actin crosslinking	Rare - 8.3	(143)
TNNT2	Sarcomere	2 - 7.4	(95)
DES	Intermediate filament	Rare - 6	(98)
DSP	Desmosome	3 - 6	(98)
BAG3	Z-disc (Molecular chaperone)	0 - 4.6	(95)
TNNI3	Sarcomere	4 - 5	(98)
MYBPC3	Sarcomere	2 - 3.1	(143)
TNNC1	Sarcomere	Rare - 1	(98)
TPM1	Sarcomere	Rare - 1	(142)
DMD	Sarcolemma	Rare - 0.5	(95)
TTR	Golgi apparatus	Rare - 0.5	(144)
ACTC1	Sarcomere	<1	(142)
VCL	Z-disc	<1	(142)

Large global datasets from high-income countries have shown inconsistent genetic causes of DCM in the adult population. Currently, the ascertainment of DCM genetic causes depends on identifying patients' disease-causing variant (s) using the NGS and Sanger sequencing

approach. The genetic causes yield ranges from 15% to 36% and 44% to 64% in sporadic and familial adult-onset DCM patients, respectively (17, 145, 146). In most patients, one class 4 or 5 variant has caused the DCM; however, at least 3% to 20% of DCM patients might have carried multiple DCM-causing variants (121, 139). Notably, the type and number of DCM-causing variant (s) would determine the severity and manifestation of the disease. For example, patients carrying multiple (compound heterozygous or digenic or homozygous) Class 4 and 5 variants have had a younger onset age or severe DCM with a poor prognosis compared to those carrying one DCM-causing variant (96, 147).

The type of variants commonly associated with DCM are those that introduce a premature termination codon in the protein-coding sequence; for example, truncating variants (tv) in the *TTN* gene have triggered about 20% to 30% of familial DCM (148). Although truncating variants have caused most end-stage DCM, especially in younger patients, other cardiac insults, DCM-causing genes, and genetic-environment interactions have put most patients at risk of HF or SCD (148). Significantly, missense variants in the *LMNA* and *MYH7* genes have been reported in 5% of DCM patients with an early onset age with a high risk of HF or SCD and phenotypic expression (149).

3.1.5.1.2 Paediatric DCM

Most of the DCM paediatric patients have been reported as carriers of heterozygous DCM-causing variants primarily inherited in an AD pattern or *de novo* (43). The DCM-causing variants have been identified at different paediatric time points, although unparalleled genetic testing in paediatrics has continued compared with the NGS testing approach used in adults. For example, DCM-causing variants in the *TPM1* gene have occurred in infancy (150), *TNNI3* in toddlers, *TTN*tv and *BAG3* in adolescents (142), and variants in the *MYH7* and *TNNT2* genes have presented at different time points (96, 130). Some of the paediatric probands carrying DCM-causing variant (s) in *TPM1* or *VCL* have been reported with severe early-onset disease (150, 151). Also, paediatric probands carrying multiple DCM-causing variants in the *MYH7* and *LAMA4* genes have been reported with severe infantile DCM (152). However, in many paediatric cohorts, monogenic variants have caused DCM (142). These known wide-ranging genetic determinants have made paediatric DCM categorisation more difficult.

Paediatrics diagnosed with heritable DCM have had severely damaged myocardium, especially at autopsy (153). Besides, unparalleled genetic testing in paediatrics has continued even though

similar genetic variants have caused DCM in about 25% to 50% of paediatric and adult patients (132, 154). Further, the few genetic test results published for paediatric cases have associated DCM-causing variants in many familial paediatric patients with low heart transplant-free survival (136). Also, a similar rate of testing positive in paediatric patients with (27%) or without (33%) FH has demonstrated insignificant diagnostic yield differences within some cohorts (130). These authors have generally searched for *de novo* variants; however, the South African paediatrics cohort has an unestablished genetic characteristic.

Few genetics studies have been conducted in South Africa for familial or non-familial DCM, and the clinical findings showed that DCM patients in the country have been documented with a spectrum of non-ischemic cardiac phenotypes (14). The genotype-phenotype correlations have been less established in South Africa, especially in the paediatric DCM. For example, no data is published for South African paediatric patients relating to the variants affecting several cytoskeletal protein-protein interactions that have caused phenotypes such as LV non-compaction (14, 155, 156).

3.1.5.2 Genetics of unclassified DCM

3.1.5.2.1 Peripartum cardiomyopathy (PPCM)

PPCM manifest as a systolic HF with reduced left ventricular ejection fraction (LVEF) in women of childbearing age, specifically during pregnancy or in the early post-partum period. The aetiology is non-specific (157), and PPCM global incidence has varied significantly by geographical region (158). Amongst other epidemiological findings, nationwide studies in the (USA have reported PPCM in more than 40% of women with African ancestry. Notable geographical hotspots have occurred in predominantly African and Afro-American populations, with incidences ranging from 1 in 100 deliveries in Nigeria to 1 in 300 deliveries in Haiti (159). The incidence of PPCM in South Africa is close to 1 in 1000 deliveries, predominantly in young women of African origin (160). Noticeably, these young South African PPCM patients have demonstrated less favourable six-month recovery rate outcomes compared to those in Germany (161).

The findings in literature for PPCM genetics studies illustrate that the disease-causing genes are similar to that of adult-onset DCM, with the global prevalence for the genetic causes estimated to be about 22%, and the most common of these DCM genes found in about 15% of

PPCM patients are the *BAG3* and *TTN* genes (159). In some cohorts, about 1.4% to 10% of the participants have had PPCM caused by *TTN*tv (162, 163). Most genetics studies have identified *TTN*tv in the sarcomere gene's A-band region in PPCM patients. Titin A-band region CMO-causing variants are known for evoking hypercontractility, inconsistent relaxation, and increased energy consumption by the heart, thus increasing patients' risk of arrhythmias and HF (160).

3.1.5.2.2 Left ventricular non-compaction cardiomyopathy (LVNC)

LVNC is a rarer form of unclassified DCM encountered in patients presenting with heart muscle diseases (152). In most LVNC patients, the heart muscle is compact and spongy, with several trabeculae visible on echocardiography (ECHO) or cardiac magnetic resonance (CMR) images. The primary complications include life-threatening arrhythmias, HF, or SCD in about 40% of the cases analysed (164). Life-threatening LVNC is commonly underestimated; however, it has been reported in 47% to 52% of individuals clinically diagnosed with familial HF (33).

In SSA, the paediatric populations have had an LVNC incidence rate of about 6.9% in a reviewed study (14). The reviewed research further showed that about 6.6% of the LVNCs in South Africa were paediatric patients carrying known genetic variants. Notably, South African individuals with young adult onset of LVNC are expected to have different genetic determinants compared to paediatrics (60). That is, about 4.7% of healthy young adult South Africans have been reported with isolated LVNC trabeculation, and the affected individuals were predominantly young males (31).

LVNC genetic determinants have been found in genes encoding for mainly the sarcomere proteins (e.g., *MYH7* and *TTN* genes) (147). Sarcomere variations in *TTN (*19%), *MYH7* (10%) and *MYH6* (6%)) genes have accounted for about 27% to 56% of LVNC (154, 165). Sarcomeric variants commonly associated with adult-onset DCM phenotype have caused about 40.8% of heritable LVNC, and the *MYH7* gene has harboured almost 25.2% of the reported Class 4 and 5 variants in some cohorts (166). Remarkably, the genes that have mutated most are the *MYH7* and *TTN* in paediatric (19%) and adult-onsets (7%), respectively (154). Besides, other gene variations, such as mitochondrial dysfunction in the TAZ gene, have caused about 15.2% to 20.4% of LVNC in some cohorts (165). The chance of finding a disease-causing variant in

the *MYH7*, *TAZ* and *TPM1* is high in paediatrics, and disease-causing variants in these genes have occurred more frequently in children (44%) compared to adult-onsets (30%) (165).

The genetics have revealed complex characteristics, specifically in patients with severe heritable non-compaction, and the complex phenotype is commonly associated with compound heterozygous truncating variants (154). Also, very young age and multiple variants have conferred a higher risk of HF in about 16% of LVNC patients (166). The few reports on multiple variants (e.g., *TTN*tv) have shown that the carriers have a higher risk of hospitalisation and HF events (33). Notably, LVNC patients' risk of developing major adverse cardiac events, especially in paediatric patients carrying multiple variants in the *MYH7*, *MYBPC3* and *TTN genes,* have been least established in this study's population.

3.1.6 DCM genotype-phenotype correlation

The genetic aetiology of heritable DCM is commonly caused by the genes encoding for the sarcomere (e.g., *TTN* gene), nuclear envelope (e.g., *LMNA*), cell-junction (e.g., *SCN5A*), Z-disc (e.g., *BAG3*) and other cytoskeletal structure (e.g., *FLNC* and *DES*) proteins. For example, truncating variants (tv) in *TTN* (11.35%) and *DSP* (1.44%) and nontruncating variants in *MYH7* (2.88%) and *TNNT2* (1.25%) genes have accounted for about 17% of significant rare variants identified in patients with severe DCM (142).

Disease-causing variants in the protein-coding sequence of sarcomere structural proteins are the leading cause of heritable DCM (167, 168). Notably, amongst the sarcomere protein variants (e.g., *TTN*, *MYH7*, *TNNT2* and *TNNI3*), the *TTN*tv has occurred in about 25% of familial DCM patients (148). The reported frequency of DCM-related *TTN* variants is 12% to <45% (14, 95). Specifically, A-band *TTN* Class 4 and 5 variants have had higher penetrance and commonly have occurred even in patients with most underdiagnosed subtypes of DCM, such as the left-ventricular non-compaction (12, 169). The various subtypes of DCM with unconnected features in many patients include the DCM and two unclassified subtypes– peripartum (PPCM) and left-ventricular non-compaction (LVNC) (60).

Class 4 and 5 variants that alter the nuclear protein's structure are associated with several phenotypes and carried by about 30% to 40% of adult-onset DCM patients (95). Most patients presenting with ventricular arrhythmias resulting from these nuclear envelope variants have been reported as having a higher risk of SCD. For example, patients carrying Class 4 and 5

variants in the lamin A/C (*LMNA*) gene have resulted in the most severe phenotype, with variant-positive patients experiencing SCD or requiring urgent heart transplants. Notably, the frequency of Class 4 and 5 variants in *LMNA* ranges from 4% to <20% of DCM patients (95).

Although rare disease-causing variants in desmosome genes that encode the cell-membrane and junctions, such as the desmoplakin (DSP) protein, have been reported with a prevalence range from 3% to 6% in DCM patients (98), these proteins have been known to interfere mainly with the function of wild-type cytoplasmic proteins (170). For example, Class 4 and 5 variants in *DSP* with a frequency of 3% to 6% in DCM patients have disrupted the function of wild-type cytoplasmic proteins such as sarcomere, leading to either ventricular tachycardia (VT) or CMO. Class 4 and 5 variants in the desmosome genes are more specific with ACM, and DCM-causing variants in genes such as the *DSP* indicate the possibility of phenotypes overlap.

As summarised in Figure 3.2, phenotype overlap in some patients has been caused by variants affecting several cytoskeletal protein interactions, especially with the nuclear, sarcomere, and desmosome proteins. The cytoskeletal structural protein network is involved in several cardiac functions. Ultimately, class 4 and 5 variants altering the cytoskeletal protein structures will be associated with cardiovascular disorders such as cardiac conduction, ventricular arrhythmia, and CMOs. For example, patients carrying multiple variants in the *MYH7* gene have been diagnosed with DCM-related HCM disorders (147, 149, 171). In most studies, genotype-phenotype overlapping has been known to increase the complexity of CMOs (mainly DCM). What was noted was that most of these studies were conducted using European patients, while very few studies have been conducted in Africa that have shown the various CMO phenotypes. However, the sarcomere genes are the dominant genetic causes of DCM. For example, genetic variation in sarcomere may result in multiple phenotypes (90).

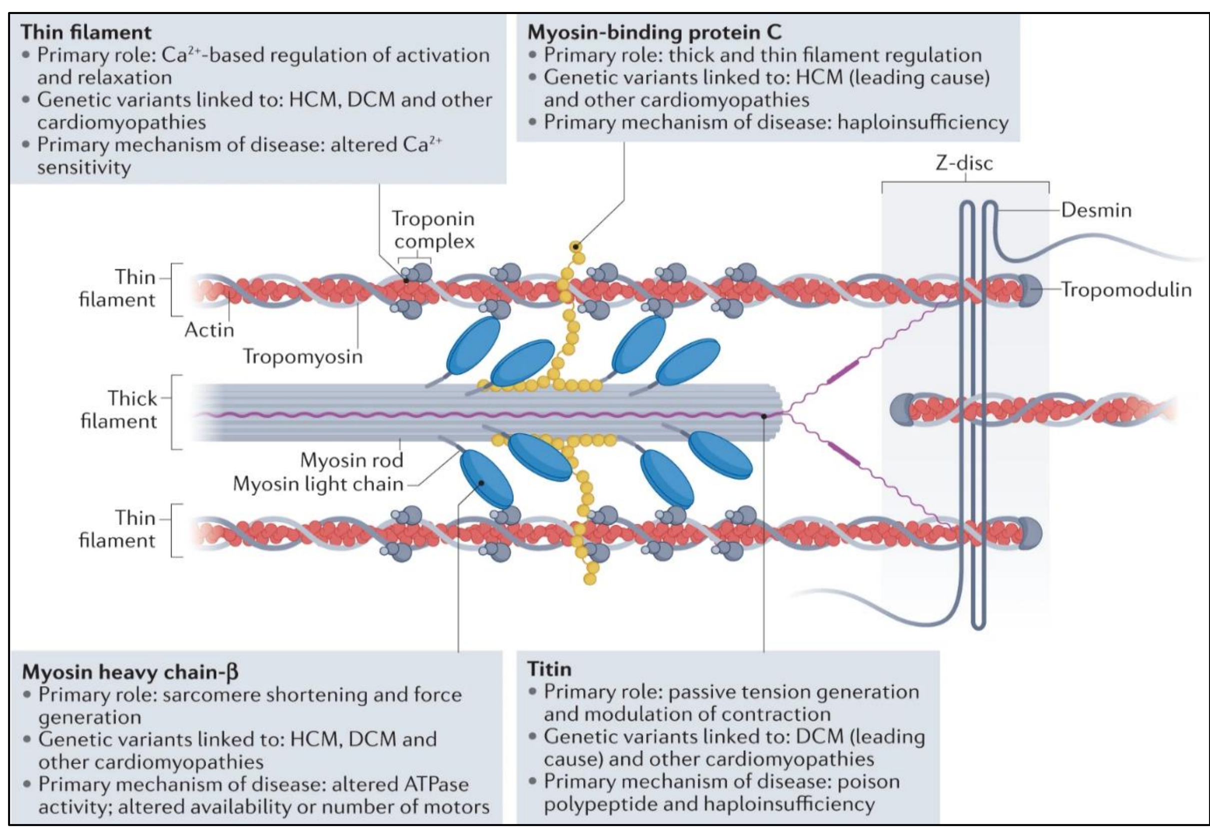


Figure 3.2: The cardiac sarcomere and its components. *The cardiac sarcomere comprises regulatory thin filaments, force-generating thick filaments (myosin), regulatory myosin-binding protein C and tension-sensing titin. These genes' mutations are linked to various genetic cardiomyopathies, such as hypertrophic cardiomyopathy (HCM) and dilated cardiomyopathy (DCM), typically characterised by a unique molecular disease mechanism (90).*

In most studies, DCM patients with complex genotype-phenotype correlations have been reported more often in younger males and women of childbearing age. In older patients, cardiac insults such as multiple variants have influenced the disease manifestation and penetrance. Most DCM-modifying insults have put patients, especially paediatrics and women of childbearing age, more at risk of HF. That is, severe DCM has also been reported in patients carrying oligogenic or compound heterozygote variants. Examples are the multiple variants of the carriers of *MYH7* and *LAMA4* genes (152). Some of these genes encode proteins involved in numerous pathways and disrupting them would cause a spectrum of cardiovascular system disorders. For example, the vinculin protein anchors the actin cytoskeletal complex to the integrin- and cadherin-containing adhesion molecules on the cell membrane, and any potential deleterious variant in this gene may destabilise cardiomyocyte functional and structural networks (172). Therefore, DCM patients with genetic aetiologies (Figure 3.3) may have a wide range of phenotypes with distinct determinants in adults, paediatrics and possibly women of childbearing age (173).

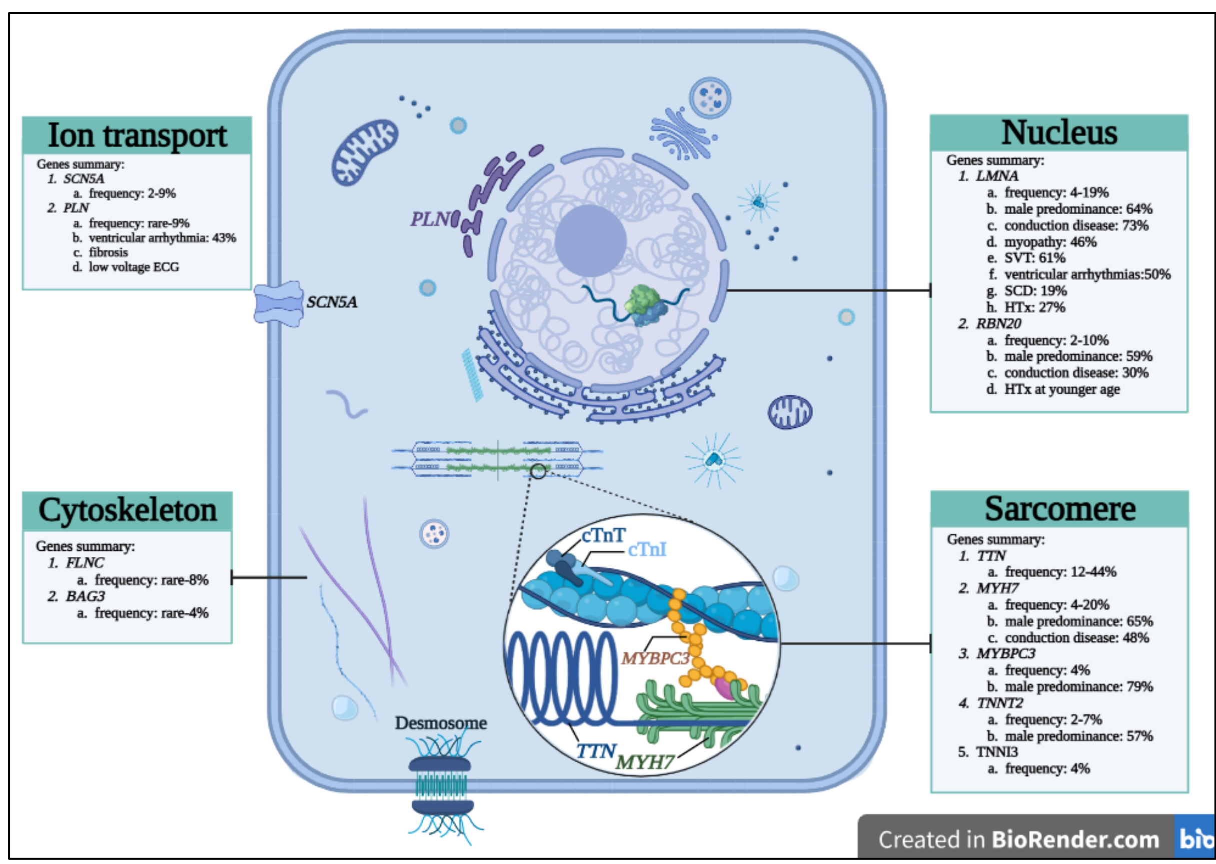


Figure 3.3: Summary of predominant phenotypic features in DCM patients. *Mutations in PLN, RBM20, TTN, MYH7, MYBPC3, TNNT2, TNNI3, FLNC, BAG3, LMNA, and SCN5A genes. SVT-supraventricular tachycardia, SCD-sudden cardiac death, HTx-heart transplantation, cTnT- cardiac troponin T, cTnI- cardiac troponin I. Drawn by P. Ndibangwi using BioRender.*

3.1.7 Therapeutics

The treatments available for DCM mainly aim to manage the symptoms, enhance cardiac function, and address the condition's underlying causes. The treatment may involve medications, lifestyle changes, gene therapy, or surgery (51, 174).

Medications such as ACE inhibitors and ARBs, beta-blockers, diuretics, aldosterone antagonists and digoxin play a vital role in managing heart-related health conditions (174, 175). ACE inhibitors and ARBs help dilate blood vessels, decrease afterload, and improve the heart's pumping ability. Meanwhile, beta-blockers work by reducing the heart rate, decreasing the workload on the heart, and improving its overall efficiency. Diuretics can help alleviate fluid retention and reduce oedema by increasing urine output.

In severe cases, implantable Cardioverter-Defibrillator (ICD) can be implanted in patients at risk of life-threatening arrhythmias. ICD is designed to monitor and regulate heart rhythms, thus preventing sudden cardiac arrest (175). Also, Cardiac Resynchronization Therapy (CRT)

is a specialised treatment that involves using a special pacemaker to coordinate the contractions of the heart's ventricles. CRT helps improve overall cardiac function, benefiting individuals suffering from HF or other cardiac conditions (176).

Recently, CRISPR-Cas9 has been used to correct *TTN*tv-related DCM (168). The researchers reversed depressed contractility caused by *TTN*tv using CRISPR-Cas9 to correct the mutation, which eliminated truncated titin proteins and raised wild-type titin content in engineered heart muscle generated from hiPSC-CMs. Gene therapy is a form of treatment for DCM that aims to address the underlying genetic factors contributing to the disease (177). Gene therapy can be done by introducing, modifying, or replacing genes to correct or compensate for genetic defects (141, 178). Gene therapy trials are studying the safety and feasibility of treating DCM. Researchers aim to improve cardiac function by targeting calcium handling, contractility, and myocardial structure genes. However, efficient gene delivery to cardiac cells is a significant challenge. Viral vectors commonly used to deliver genes for therapy include adeno-associated viruses (AAVs), but face challenges such as immune responses and off-target effects; however, ongoing research aims to optimize gene-based interventions (179). Clinical trials are underway to evaluate gene therapy for various cardiovascular conditions, including DCM.

Lifestyle modifications that may improve cardiovascular health include making dietary changes, exercising regularly, and quitting smoking (138). A heart-healthy diet low in salt and saturated fats is often recommended, along with individualized exercise recommendations. Quitting smoking is essential for overall heart health. Identifying and addressing the underlying causes of DCM, including inherited mutations, viral infections, autoimmune disorders, or metabolic conditions, is crucial to managing the condition effectively (48). In severe cases of DCM, where other treatments have been unsuccessful, heart transplantation may be considered as a final option (180, 181). Genetic counselling could be recommended for cases with a genetic component to assess risk to family members and provide guidance on preventive measures.

3.2 **METHODS**

The methods discussed in Chapter 2 are again briefly described in Section 3.2 for the DCM cohort.

3.2.1 Ethical compliance

Ethical clearance with reference number **HREC 009/2020** was obtained for this study from the Human Research Ethics Committees (HREC) at the Faculty of Health Sciences (FHS), University of Cape Town (UCT). All participants were informed before they provided a signed consent form in the main study IMHOTEP registry (HREC 766/2014).

3.2.2 Study population

We recruited 539 unrelated consecutive probands clinically diagnosed with DCM at six tertiary cardiac clinics in South Africa (Appendix J) from February 2018 to October 2021. At least one cardiologist examined the probands at the participating site. The records were again re-evaluated in Cape Town by a team of at least one cardiologist, a research fellow, and a clinic doctor. The clinical examination included baseline assessment, complete medical history, pedigree, ECG and transthoracic echocardiography. The clinical diagnostic criteria for DCM were LVEF of < 45%, associated with left ventricular end-diastolic dimensions > 117% of the predicted value corrected for age and body surface area (182). The inclusion and exclusion criteria are described in section 2.1.2 of chapter two.

3.2.3 DNA extraction and quality control

We extracted the genomic DNAs (gDNA) from probands' peripheral whole-blood leukocytes. Each gDNA was extracted from the buffy or whole blood following the published manufacturer's methods for Puregene Blood Core Kit C (Appendix C) or PAXgene Blood DNA kit (Appendix D), respectively. We then assessed the gDNA samples' integrity, quality, and quantity. We computed the gDNA purity and concentrations with a calibrated NanoDrop^{TM} 2000 spectrophotometer (Thermo SCIENTIFIC) software v2.6 on a Windows 10 computer. At the same time, we checked the samples' intactness with a 1% agarose gel. We programmed the

gel electrophoresis to run at 110V for 75 minutes. The gel was visualised in a Uvitec Xplorer D55 gel doc running on an Xplorer 1D software version 15.08 (Uvitec Cambridge, Cambridge, UK).

3.2.4 Targeted sequencing and variant calling

A custom-designed list of 38 CMO disease-causing genes was created by the (OMGL in the UK. The custom-designed chip contained 38 specific genes, including *BAG3*, *DES*, *DSP*, *DMD*, *FLNC*, *LMNA*, *PLN*, *SCN5A*, *RBM20*, *TTN*, *VCL*, and other DCM-related genes. The complete genes list is attached as Appendix F. The chip's chemistry enabled a targeted sequencing read coverage of ≥ 99.3% at a depth of >500x. All gDNA samples underwent targeted resequencing, initially performed using the Haloplex kit but subsequently replaced with the Twist NGS kit in September 2021.

After DNA sequencing, we used the FastQC standard tool to check the quality of reads generated on an Illumina platform to remove low-quality base calls. We also used the Illumina QC tools to identify and remove poor-quality nucleotides and adaptor-contaminated reads. We mapped QC-passed reads to the human reference genome sequence (GRCh37), calculated the percentage of aligned reads, and removed artefacts.

We transformed the data generated into variant call format (VCF) using the Genome Analysis Toolkit (GATK) v3.8-1-0 and Platypus v 0.8.1. Using the GATK software, we applied filters against poor-quality sequences for the variants calling. We then annotated the VCF files using the Ensembl Variants Effect Predictor (VEP) for further analysis. We analysed the annotated VCF file for each proband and curated all possible DCM-causing variants harboured by the 38 genes sequenced.

Through a specific set of filtering criteria, we selected variants for further analysis. The variants were classified according to the refined ACMG guidelines criteria as class 1 (B), class 2 (LB), class 3 (VUS), class 4 (LP) and class 5 (P). We selected variants: i) with a gnomAD minor allele frequency (MAF) of <1%; ii) variants that were predicted by bioinformatics tools, SIFT and PolyPhen2 to have damaging deleterious or unknown protein function probably; iii) we analysed the variants further using Alamut™ Visual Plus version 1.6.1 and ClinVar (2022-12-17); iv) we also looked at the variant caller scores for Haplocaller and Platypus for threshold values above 400 and 700, respectively, as variants below these thresholds are artefacts or

polymorphisms; and v) we visualised the BAM files for probands carrying class 4 and 5 variants in the Alamut as this may identify possible false positive variants. Finally, we validated all identified class 4 and 5 variants using Sanger sequencing.

3.2.5 Biostatistical analysis

In the *Results* section, we analysed the data using Microsoft Excel, R, and RStudio (version 4.1.2). For numerical variables, we presented the results as the $\bar{x} = \pm 1\delta$, while for categorical variables, we reported total counts with percentages.

3.3 **RESULTS**

This study enrolled 539 probands diagnosed with DCM, comprising 83.5% adult-onset and 16.5% paediatrics DCM (Table 3.2). Of the 450 adult participants enrolled, 30.2% (136/450) were diagnosed with peripartum CMO (PPCM), while 24.7% (22/89) of the paediatrics probands were diagnosed with non-compaction CMO (LVNC). The results for the adult (≥ 13 years) and paediatric (< 13 years) are presented in sections 3.3.1 and 3.3.3, respectively.

Table 3.2: Baseline results for the DCM cohort, including adults and paediatric probands

Label	Levels	Total IMHOTEP cohort	DCM Adult-onset Probands	DCM Paediatrics Probands
Number of probands with genetics data analysed, n (%)		690	450 (65.2%)	89 (12.9%)
Age at diagnosis (years)	Mean (SD)	31.7 (16.4)	35.6 (11.8)	3.1 (3.5)
Left ventricular ejection fraction (%)	Mean (SD)	36.1 (19.6)	28.0 (11.9)	-
New York Heart Association classification	Class I	83 (15.7)	28 (6.9)	-
	Class II	162 (30.6)	113 (27.8)	-
	Class III	223 (42.1)	207 (50.9)	-
	Class IV	52 (9.8)	50 (12.3)	-
	ND	10 (1.9)	9 (2.2)	-
Familial	Yes	171 (24.8)	128 (28.4)	-
	No	359 (52.0)	279 (62)	-
	No data	160 (23.2)	43 (9.5)	-
Sex assigned at birth	Female	364 (52.8)	252 (56.0)	49 (55.1)
	Male	326 (47.2)	198 (44.0)	40 (44.9)
Ancestry	Black-African	381 (55.2)	260 (57.8)	62 (69.7)
	Mixed	221 (32.0)	149 (33.1)	25 (28.1)
	White	83 (12.0)	40 (8.9)	2 (2.2)
	Indian	5 (0.7)	1 (0.2)	0 (0.0)
Values are mean ± 1SD or n (%) based on non-missing values of the total analysed.				

3.3.1 Adult-onset DCM probands

3.3.1.1 Baseline characteristics for the adult DCM cohort

In Table 3.2, we showed that the adult cohort had a mean age of 35.6 years (SD 11.8, range 14-75) at diagnosis, with 56.0% recorded as female, and that >50% of the probands were classified as NYHA class III. The 450 adult probands' data also revealed that the self-reported ancestry found 57.8% Black-African, 33.1% Mixed, 8.9% White and 0.2% Indian ancestry (IA) probands. We presented the FH data of their disease in Figure 3.3, which shows that 62.0% (279/450) of the DCM phenotype was caused by non-familial disease, 28.4% (101/450) by familial DCM, and 9.6% (43/450) of the probands had no FH data. In addition, some probands' aetiology was attributed to peripartum CMO in 30.2% (136/450) and other causes (e.g., toxin-associated DCM or tachycardiomyopathy) in 13.1% (59/450).

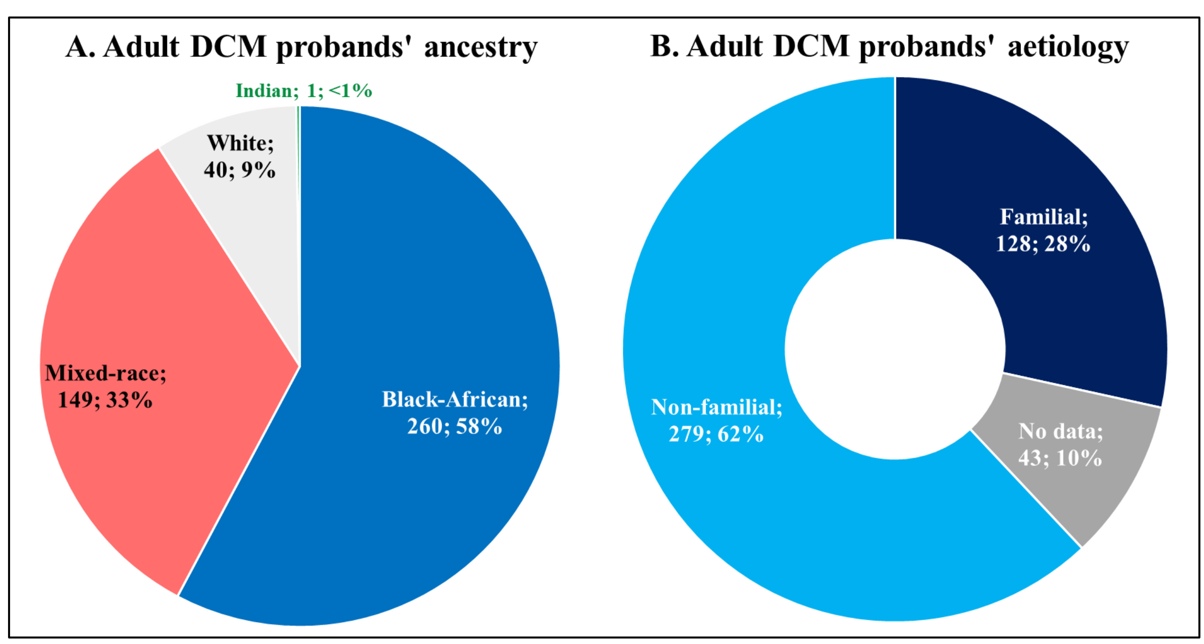


Figure 3.4: Adult DCM probands family history. ***A.*** *The self-reported ancestry for the adult DCM probands was recorded at enrolment.* ***B.*** *The clinical aetiologies assigned at diagnosis; n=450 probands.*

3.3.1.2 Baseline characteristics for the adult DCM according to probands' ancestry

As South Africa has a uniquely mixed population, and the cohort's results had shown differences in age and sex assigned at birth, we took the opportunity to investigate any potential trends. We investigated the diagnostic classification across the four dominant population groups of South Africa: Black Africa ancestry (BA), Indian ancestry (IA), Mixed ancestry (MA), and White ancestry (WA), as shown in Figure 3.5. However, we used caution in the interpretation because some population groups were in small numbers.

The 260 BA probands were the youngest group in this cohort (p-value = 0.11), with a mean age of 34 years (SD 11.7 years, range 14-72) (Figure 3.5 A), with 63.5% (165/450) being female (Figure 3.5 B). The aetiology showed that 68.5% (178/260) were non-familial, only 22.7% (59/260) had a familial DCM, and 8.8% (23/260) had no aetiology data for their HF history (Figure 3.5 D).

In comparison, the MA probands were, on average, about two years older than the BA probands with a mean age of 36 (SD 10.8, range 14-68 years). Also, 50.3% (75/149) of MA probands recruited were females. The aetiology showed that 57.7% (86/149) were non-familial, 32.9% (49/149) had familial aetiology, and 9.4% (14/149) had no data for the cause or HF history.

We also recruited 40 WA probands, and the data showed that they had a mean age of 43.1 years (SD 13.8, range 16-75), which was older than the other population groups (p-value = 0.11).

Unlike in the BA and MA groups, the majority of the WA probands, 70% (28/40) were males. The aetiology showed that 47.5% (19/40) were familial, 37.5% (15/40) had non-familial aetiology, and 15% (6/40) had no data for the cause or HF history (Figure 3.5 D).

Finally, we recruited one male IA proband, who was diagnosed at the age of 36 with familial DCM.

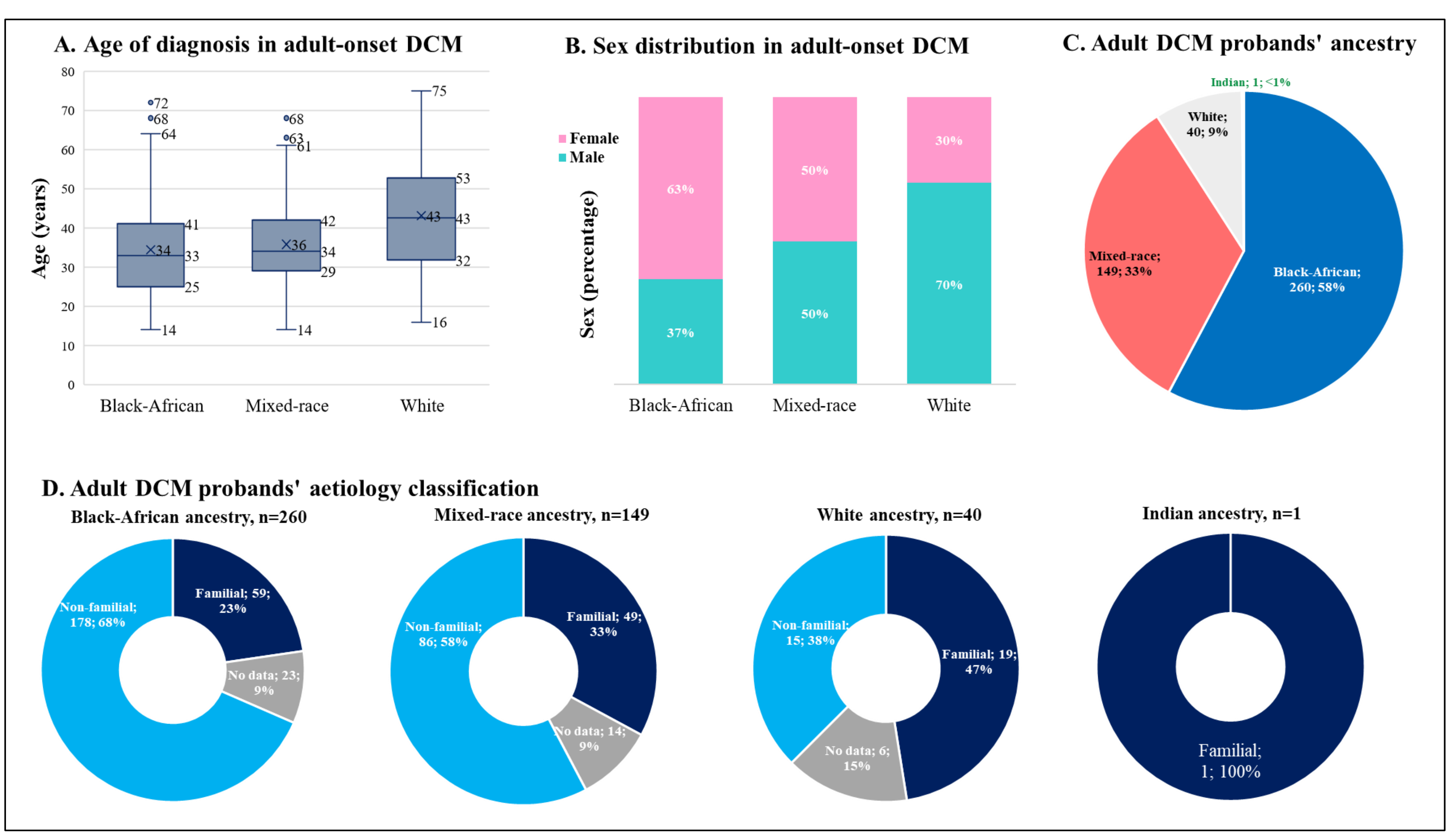


Figure 3.5: Adult DCM probands' age, sex assigned at birth and aetiology distribution. *Graphs showing comparisons between adult-onset DCM population groups regarding age, sex assigned at birth and diagnostic classification. A. Population groups vs. age of diagnosis in adult-onset DCM; B. Population groups vs. sex distribution; C. The self-reported ancestry recorded at enrolment, D. Population groups vs. diagnostic classification.*

3.3.1.3 Targeted sequencing for adult-onset DCM

All variants reported in this DCM cohort were curated using the refined ACMG classification system 2018. Table 3.3 reports the baseline characteristics of ACMG classes 1-5 variant carriers for the adult-onset DCM cohort. Whereas SS Table 3 shows the fitted logistic regression model for adult DCM probands. This chapter also delves into these variants under the various ACMG class headings. According to the ACMG guidelines, class 5 is P, class 4 is LP, class 3 is a VUS, class 2 is LB, and class 1 is B.

Table 3.3: Baseline results for the adult-onset DCM cohort according to ACMG classification

LABEL	LEVELS	TOTAL	POSITIVE	UNCERTAIN	NEGATIVE	P-VALUE
DCM diagnostic yield, n (%)		450	76 (16.9)	223 (49.6)	151 (33.6)	
AGE	Mean (SD)	35.6 (11.8)	36.4 (11.8)	34.9 (11.5)	36.3 (12.4)	0.401
LVEF	Mean (SD)	28.0 (11.9)	25.7 (12.1)	28.8 (11.8)	28.0 (11.9)	0.179
NYHA						0.045
	Class I	28 (6.9)	3 (4.4)	11 (5.5)	14 (10.1)	1.00
	Class II	113 (27.8)	25 (36.8)	45 (22.4)	43 (31.2)	0.251
	Class III	207 (50.9)	29 (42.6)	108 (53.7)	70 (50.7)	0.098
	Class IV	50 (12.3)	9 (13.2)	31 (15.4)	10 (7.2)	7.39e-3
	ND	9 (2.2)	2 (2.9)	6 (3.0)	1 (0.7)	0.065
FAMILIAL*						0.010
	Yes	128 (28.4)	34 (44.7)	55 (24.7)	39 (25.8)	1.74e-5
	No	279 (62.0)	34 (44.7)	146 (65.5)	99 (65.6)	0.321
	No data	43 (9.6)	8 (10.5)	22 (9.9)	13 (8.6)	0.977
SEX*						0.020
	Female	252 (56.0)	32 (42.1)	135 (60.5)	85 (56.3)	4.01e-7
	Male	198 (44.0)	44 (57.9)	88 (39.5)	66 (43.7)	0.929
ANCESTRY*						0.021
	Black-African	260 (57.8)	39 (51.3)	137 (61.4)	84 (55.6)	2.44e-8
	Mixed	149 (33.1)	24 (31.6)	76 (34.1)	49 (32.5)	0.904
	White	40 (8.9)	13 (17.1)	10 (4.5)	17 (11.3)	0.207
	Indian	1 (0.2)	0 (0.0)	0 (0.0)	1 (0.7)	0.979
DCM SUBTYPE						0.001
	Dilated	314 (69.8)	63 (82.9)	138 (61.9)	113 (74.8)	9.67e-7
	Peripartum	136 (30.2)	13 (17.1)	85 (38.1)	38 (25.2)	0.098

VALUES ARE MEAN ± 1SD OR N (%) BASED ON NON-MISSING VALUES OF THE TOTAL ANALYSED:450. POSITIVE-CLASS 4 & 5, UNCERTAIN-CLASS 3, AND NEGATIVE-CLASS 1&2: LVEF-LEFT VENTRICULAR EJECTION FRACTION, NYHA-NEW YORK HEART ASSOCIATION. ASTERISKS (*)-SELF-REPORTED.

Overall, targeted sequencing identified 76 (16.9%) DCM probands carrying 51 class 4 and 5 variants; 51 probands had 30 class 5 variants, and 25 probands carried 21 class 4. We also identified 223 probands that carried variants of uncertain significance (class 3), and 151 probands had genotype-negative (class 1 or 2) variants (Figure 3.6 A and B). Figure 3.6 B also shows an overview of the distribution of novel and reported variants across class 3, 4 and 5 variants for this DCM cohort. This graph highlights that in South Africa, a significant

proportion of variants, especially class 3, require further evaluation to determine if they are P, B or possible modifiers.

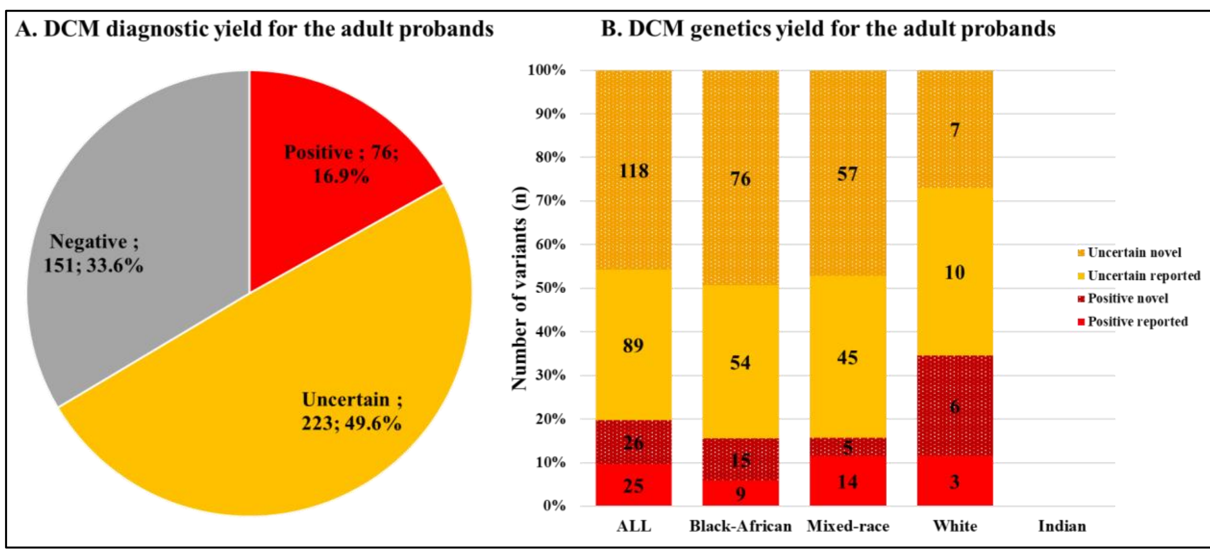


Figure 3.6: Adult DCM probands' diagnostic yield. *A. Pie chart indicating the diagnostic yield for the adult DCM cohort, B. Bar graph showing the count of probands with variants grouped as novel vs reported within the DCM cohort.*

Figure 3.6 B further illustrates the low class 4 and 5 variant yields across the BA, MA, and WA probands and the large yield of class 3 variants, especially in both the BA and MA population groups. The variants in the various ACMG classes will be further explored in the subsequent sections. We first described class 4 and 5 variants and followed up with a report on class 3 variants and their possible modifying effects on DCM. Note that classes 1 and 2 (Negative) variants are beyond the scope of this thesis as we attempted to focus on the variants that have a higher likelihood of causing DCM.

To understand the genetic architecture of South Africa's diverse populations, we also looked at the DCM ACMG classification yields across the three ethnic groups of BA, MA and WA.

3.3.1.3.1 Probands with pathogenic (class 4 and 5) variants

3.3.1.3.1.1 Baseline characteristics

We analysed the probands' genetics data and found that 16.9% (76/450) carried DCM-causing variants: three of the probands carried two class 4 or 5 variants. The positive genetic findings in Table 3.4 show that the 76 DCM genotype-positive participants consist of 39 (51.3%) BA probands, 24 (31.6%) MA and 13 (17.1%) WA probands. Table 3.4 also shows 14 genes

harbouring 51 class 4 and 5 variants we identified. Because the number of variants (n=51) is fewer than that of probands (n=76), we hypothesise they are recurring variants in our DCM cohort. We therefore analysed the positive genotype findings as a whole and according to the various probands' self-reported ancestry.

When we proceeded to summarise the combined data for the probands positive group (class 4 and 5 variants carriers), we noted a significant age difference (p=0.00076) between the MA probands [34.0 years (SD 8.6, range 22-60)] as shown in Figure 3.7. Figure 3.7A shows that the probands with class 4 variants were much younger for the MA [mean age of 26.3 years (SD 2.9, range 22-30)] compared to their BA [mean age of 37.3 years (SD 11.9, range 25-64)] and WA counterparts [mean age of 42.2 years (SD 11.7, range 26-58)].

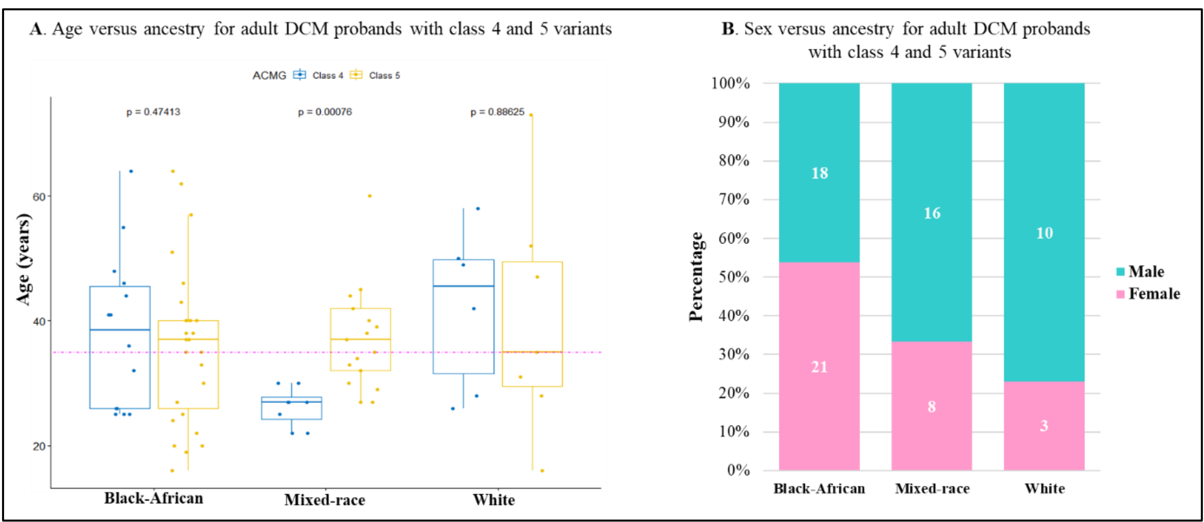


Figure 3.7: **Age and sex distribution for adult-DCM class 4 and 5 variant carriers.** ***A.*** *The age-to-ancestry summary for adult DCM probands with class 4 or 5 variants. The magenta line indicates the median age of 35 for the 76 positive probands.* ***B.*** *The proportion of adult DCM-positive probands with class 4 or 5 variants according to sex assigned at birth.*

When comparing adult DCM probands with class 4 and 5 variants and sex assigned to them at birth, we noted 57.9% (44/76) males vs 42.1% (32/76) females (Table 3.3). Further analysis of the sex distribution showed that 70.5% (31/44) of the males had class 5 variants compared to 29.5% (13/44) with class 4 variants. Similarly, more female probands, 68.8% (22/32), had class 5 variants compared to 31.3% (10/32) with class 4 variants. However, Figure 3.7B shows that more female (53.8%) BA probands had class 4 and 5 variants than male BA probands. In contrast, over two-thirds of the MA and WA probands with class 4 and 5 variants were males.

3.3.1.3.1.2 Diagnostic findings

We analysed the probands' genetics data and found that 16.9% (76/450) carried DCM-causing class 4 and 5 variants, as shown in Figure 3.6A, while three probands each carried two class 4 or 5 variants. We also note that 83.1% of the DCM cohort had uncertain or negative results, Figure 3.6 A.

The 76 probands with the 51 class 4 and 5 variants (21 class 4 and 30 class 5 variants) are shown in Table 3.4.

Table 3.4: Class 4 and 5 variants (n=51) in the adult-onset DCM probands (n=76) spanning three population groups

PROBAND_ID	SEX	ANCESTRY	AGE	GENE	PROTEIN CHANGE	CDNA CHANGE	CONSEQUENCES	EVIDENCE	dbSNP_ID	ZYGOSITY	ACMG
10100007	Female	Black-African	19	*BAG3*	p.Trp36Ter	c.107G>A	stop_gained	Novel	None	Het	Class 5
10100024	Female	Mixed	45	*LMNA*	p.Arg190Trp	c.568C>T	missense	Reported	rs59026483	Het	Class 5
10100043	Male	White	16	*LMNA*	p.Arg541His	c.1622G>A	missense	Reported	rs61444459	Het	Class 5
10100045	Male	Mixed	30	*TTN*	Splice Site	c.106531+1G>A	splice_donor	Reported	rs760915007	Het	Class 5
10100047	Male	Black-African	44	*SCN5A*	p.Tyr1228Cys	c.3683A>G	missense	Novel	rs746509665	Het	Class 4
10100065	Female	Mixed	27	*TTN*	p.Arg20626Ter	c.61876C>T	stop_gained	Reported	rs72646846	Het	Class 5
10100082	Female	Mixed	27	*TTN*	p.Ile14101SerfsTer40	c.42300del	frameshift	Novel	None	Het	Class 4
10100305	Female	Black-African	38	*TTN*	p.Glu20759Ter	c.62275G>T	stop_gained	Novel	None	Het	Class 5
10100306	Male	White	47	*TTN*	p.Gly1316Ter	c.3946G>T	stop_gained	Novel	None	Het	Class 5
10100309	Female	Black-African	20	*TTN*	Splice Site	c.9704-2A>G	splice_acceptor	Novel	None	Het	Class 5
10100312	Male	Mixed	60	*BAG3*	p.Arg309Ter	c.925C>T	stop_gained	Reported	rs869248137	Het	Class 5
10100313	Male	Black-African	24	*TTN*	p.Arg31670Ter	c.95008C>T	stop_gained	Reported	rs1322596650	Het	Class 5
10100323	Male	Mixed	27	*TNNT2*	p.Arg141Trp	c.421C>T	missense	Reported	rs74315379	Het	Class 5
10100328	Male	White	35	*BAG3*	p.Glu72Ter	c.214G>T	stop_gained	Novel	None	Het	Class 5
10100329	Female	Mixed	39	*FLNC*	p.Gln2303Ter	c.6907C>T	stop_gained	Reported	None	Het	Class 5
10100338	Female	Black-African	41	*TTN*	p.Glu28397Ter	c.85189G>T	stop_gained	Novel	None	Het	Class 4
10100340	Male	White	28	*TTN*	p.Tyr29208Ter	c.87624C>A	stop_gained	Reported	rs772121356	Het	Class 4
10100344	Female	Mixed	32	*BAG3*	p.Arg309Ter	c.925C>T	stop_gained	Reported	rs869248137	Het	Class 5
10100345	Male	Black-African	40	*TTN*	p.Phe7410GlnfsTer19	c.22227_22228del	frameshift	Novel	None	Het	Class 5
10100351	Female	Black-African	22	*FLNC*	p.Arg1341Ter	c.4021C>T	stop_gained	Reported	rs1562998062	Het	Class 5
10100363	Male	White	52	*PKP2*	p.Ser837ValfsTer94	c.2509del	frameshift	Reported	rs727504432	Het	Class 5
10100365	Female	White	28	*TTN*	p.Trp26308Ter	c.78923G>A	stop_gained	Novel	None	Het	Class 5
10100375	Male	Mixed	33	*TTN*	p.Arg16724Ter	c.50170C>T	stop_gained	Reported	rs794729265	Het	Class 5
10100376	Male	Mixed	29	*TTN*	p.Arg16724Ter	c.50170C>T	stop_gained	Reported	rs794729265	Het	Class 5
10100378	Male	Mixed	27	*MYBPC3*	p.Arg726Cys	c.2176C>T	missense	Reported	rs752200396	Het	Class 4
10100378	Male	Mixed	27	*TTN*	p.Arg29947Ter	c.89839C>T	stop_gained	Reported	rs727505224	Het	Class 4
10100395	Male	Black-African	37	*TTN*	p.Arg31670Ter	c.95008C>T	stop_gained	Reported	rs1322596650	Het	Class 5
10100396	Female	Black-African	35	*TTN*	p.Arg31670Ter	c.95008C>T	stop_gained	Reported	rs1322596650	Het	Class 5
10100402	Female	Black-African	25	*TTN*	Splice Site	c.9704-2A>G	splice_acceptor	Novel	None	Het	Class 5

10100438	Male	Black-African	38	*TTN*	p.Arg31670Ter	c.95008C>T	stop_gained	Reported	rs1322596650	Het	Class 5
10100702	Male	Black-African	37	*TTN*	p.Asn16662ThrfsTer8	c.49985_49988del	frameshift	Novel	None	Het	Class 5
10100706	Female	Mixed	44	*FHL1*	p.Thr152Ile	c.455C>T	missense	Novel	rs1268621504	Het	Class 5
10100713	Female	Mixed	22	*TTN*	p.Arg28779Ter	c.86335C>T	stop_gained	Reported	rs1060500525	Het	Class 4
10100720	Male	Mixed	30	*TTN*	p.Tyr29208Ter	c.87624C>A	stop_gained	Reported	rs772121356	Het	Class 4
10100724	Male	White	31	*TTN*	p.Met26800AsnfsTer6	c.80398dup	frameshift	Novel	None	Het	Class 5
10100742	Male	Mixed	42	*TTN*	p.Pro21370GlnfsTer4	c.64109del	frameshift	Novel	None	Het	Class 5
10100747	Male	White	73	*PRKAG2*	Splice Site	c.1585-2A>G	splice_acceptor	Novel	None	Het	Class 5
10100750	Male	Mixed	37	*FLNC*	p.Gln2303Ter	c.6907C>T	stop_gained	Reported	None	Het	Class 5
101301023	Female	Black-African	26	*MYH7*	Splice Site	c.5791-2A>T	splice_acceptor	Reported	None	Het	Class 4
101301023	Female	Black-African	26	*MYH7*	p.Gly1931AlafsTer2	c.5792del	frameshiftt&splice_region	Novel	None	Het	Class 4
101301363	Male	Mixed	34	*TTN*	p.Gly27533ValfsTer47	c.82598del	frameshift	Reported	None	Het	Class 5
101301367	Female	Black-African	43	*TTN*	p.Arg31670Ter	c.95008C>T	stop_gained	Reported	rs1322596650	Het	Class 5
101301368	Female	Black-African	25	*TTN*	p.Lys13508AsnfsTer8	c.40521_40527del	frameshift	Novel	None	Het	Class 4
101301372	Female	White	50	*TTN*	p.Tyr29208Ter	c.87624C>A	stop_gained	Reported	rs772121356	Het	Class 4
101301389	Male	Mixed	40	*TTN*	p.Phe14847CysfsTer7	c.44540_44541del	frameshift	Novel	None	Het	Class 5
101301396	Female	Mixed	38	*LMNA*	p.Arg190Trp	c.568C>T	missense	Reported	rs59026483	Het	Class 5
101301409	Male	Mixed	25	*TTN*	p.Arg18140Ter	c.54418C>T	stop_gained	Reported	rs747236787	Het	Class 4
101301418	Male	Mixed	42	*LMNA*	p.Arg190Trp	c.568C>T	missense	Reported	rs59026483	Het	Class 5
101301438	Female	Black-African	35	*TTN*	p.Arg31670Ter	c.95008C>T	stop_gained	Reported	None	Het	Class 5
101301504	Male	Black-African	46	*TTN*	p.Arg34079SerfsTer9	c.102234_102237del	frameshift	Reported	rs754731686	Het	Class 4
101301508	Female	Black-African	36	*TTN*	p.Arg15898Ter	c.47692C>T	stop_gained	Reported	rs775186117	Het	Class 4
101301528	Male	Black-African	33	*TTN*	Splice Site	c.45350-1G>A	splice_acceptor	Novel	None	Het	Class 5
101301554	Male	Black-African	16	*RBM20*	p.Arg634Trp	c.1900C>T	missense	Reported	rs796734066	Het	Class 5
101301563	Male	Mixed	35	*LMNA*	p.Arg190Trp	c.568C>T	missense	Reported	rs59026483	Het	Class 5
101301569	Male	Mixed	30	*TTN*	p.Cys26564Ter	c.79692T>A	stop_gained	Novel	None	Het	Class 4
101301717	Female	Black-African	40	*TTN*	p.Phe29218LeufsTer7	c.87654del	frameshift	Novel	None	Het	Class 5
101304259	Female	White	49	*JUP*	p.Asn321Ter	c.960dup	frameshift	Novel	None	Het	Class 4
106304528	Female	Black-African	32	*TTN*	Splice Site	c.48638+2T>G	splice_donor	Reported	None	Het	Class 4
106304529	Male	White	58	*TTN*	p.Tyr29208Ter	c.87624C>A	stop_gained	Reported	rs772121356	Het	Class 4
106304532	Male	White	26	*TTN*	p.Tyr29208Ter	c.87624C>A	stop_gained	Reported	rs772121356	Het	Class 4

106304542	Female	Black-African	25	*MYH7*	p.Pro152Arg	c.455C>G	missense	Novel	None	Het	Class 4
106304543	Female	Black-African	55	*TTN*	Splice Site	c.48638+2T>G	splice_donor	Reported	None	Het	Class 4
106304545	Male	Black-African	51	*TTN*	p.Arg31670Ter	c.95008C>T	stop_gained	Reported	rs1322596650	Het	Class 5
106304552	Male	Black-African	40	*TTN*	p.Arg31670Ter	c.95008C>T	stop_gained	Reported	rs1322596650	Het	Class 5
106304556	Female	Black-African	46	*TTN*	p.Asp24091GlufsTer8	c.72269_72272dup	frameshift	Novel	None	Het	Class 5
106304557	Male	Black-African	40	*TTN*	p.Arg31670Ter	c.95008C>T	stop_gained	Reported	rs1322596650	Het	Class 5
106304560	Female	Black-African	64	*TTN*	p.Arg31670Ter	c.95008C>T	stop_gained	Reported	rs1322596650	Het	Class 5
106304560	Female	Black-African	64	*SCN5A*	p.Leu1338Ile	c.4012C>A	missense	Novel	None	Het	Class 4
106304569	Male	White	42	*TTN*	p.Tyr29208Ter	c.87624C>A	stop_gained	Reported	rs772121356	Het	Class 4
106304571	Female	Black-African	20	*TTN*	p.Arg31670Ter	c.95008C>T	stop_gained	Reported	rs1322596650	Het	Class 5
106304575	Male	Black-African	35	*DMD*	Splice Site	c.1603-2A>G	splice_acceptor	Reported	None	Hom	Class 5
106304577	Male	Black-African	27	*TTN*	p.Arg31670Ter	c.95008C>T	stop_gained	Reported	rs1322596650	Het	Class 5
106304582	Male	Black-African	57	*TTN*	p.Arg31670Ter	c.95008C>T	stop_gained	Reported	rs1322596650	Het	Class 5
106304597	Male	Mixed	22	*MYH7*	p.Arg1250Gly	c.3748C>G	missense	Reported	None	Het	Class 4
401303008	Female	Black-African	25	*TTN*	p.Arg15898Ter	c.47692C>T	stop_gained	Reported	rs775186117	Het	Class 4
401303009	Female	Black-African	62	*TTN*	p.Asn16662ThrfsTer8	c.49985_49988del	frameshift	Novel	None	Het	Class 5
401303019	Male	Black-African	41	*TTN*	p.Ser33240Ter	c.99719C>G	stop_gained	Reported	rs794727539	Het	Class 4
401303028	Male	Black-African	30	*TTN*	p.Arg31670Ter	c.95008C>T	stop_gained	Reported	rs1322596650	Het	Class 5
401303047	Male	Black-African	48	*FLNC*	p.Ile2003AspfsTer35	c.6006dup	splice_acceptor	Novel	None	Het	Class 4

The summarised data is shown in a sunburst chart (Figure 3.8), where we show the genetic distribution for the class 4 and class 5 variants in the DCM probands. The predominant gene in the DCM cohort was *TTN*, occurring in 68.4% (52/76) of the probands, which was followed by *LMNA*, *FLNC*, *BAG3* and MYH7 with 6.6% (5/76), 5.2% (4/76)), 5.2% (4/76) and 5.2% (4/76), respectively.

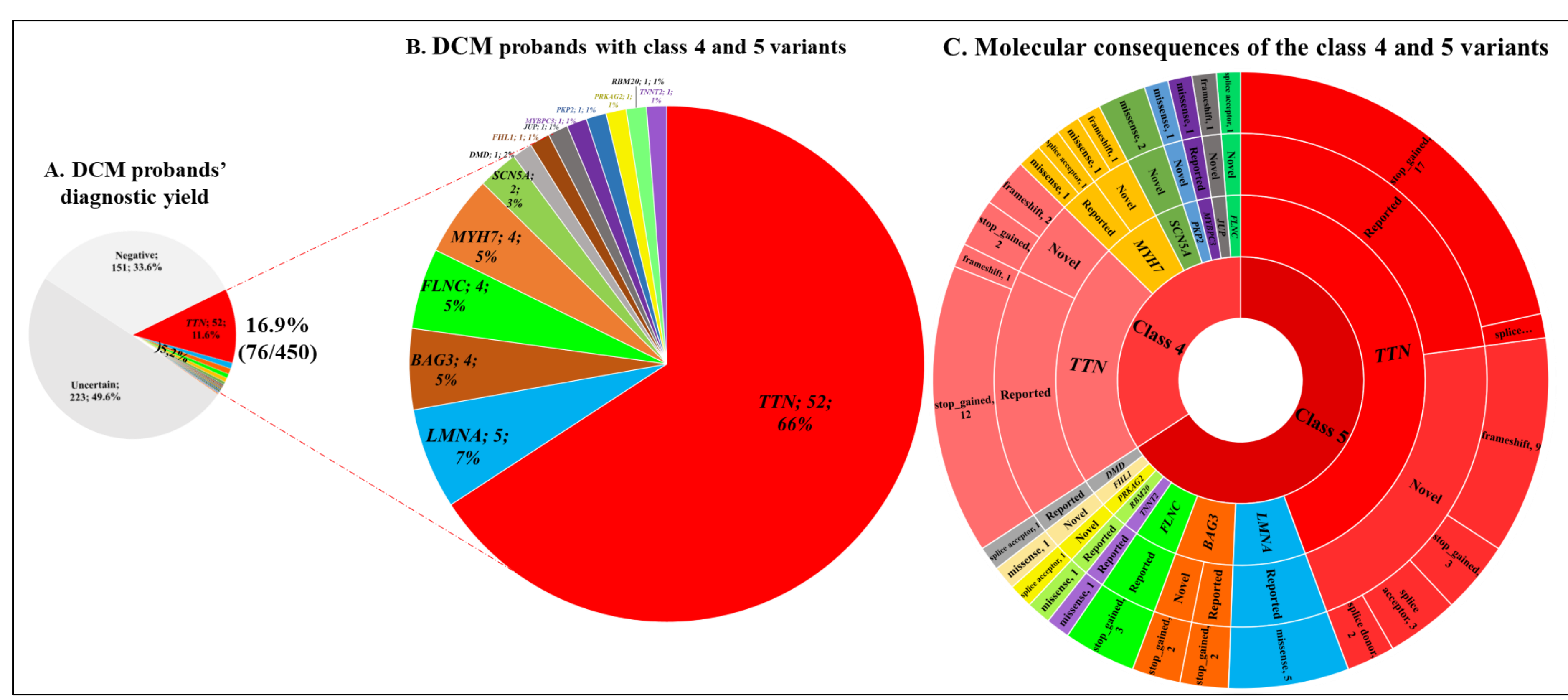


Figure 3.8: **Genes found in adult-onset HCM cohort with class 4 and/or class 5 variants.**

The class 4 and 5 variants span 14 genes that encode for the sarcomere (*TTN*, *MYH7*, *MYBPC3* and *TNNT2*), nuclear envelope (*LMNA*, *RBM20*, *PRKAG2*), Z-disc (*BAG3*), cytoskeletal (*FLNC*, *DMD*, *FHL1*) and cell-junction (*SCN5A*, *JUP*, *PKP2*) genes, Table 3.5. We also noted that *TTN* harboured seven (77.8%) of the recurring variants, followed by recurrent variants in *BAG3*, *FLNC* and *LMNA*. Further investigations are being conducted (not part of this study) to ascertain if the recurrent variants in Table 3.5 could be founder mutations.

Table 3.5: The cellular location of the 14 genes with class 4 and 5 variants in the adult DCM cohort

Gene	Protein Change	cDNA Change	Consequences	Evidence	ACMG	Frequency
Sarcomeric genes						
TTN	p.Arg31670Ter	c.95008C>T	stop_gained	Reported	Class 5	14*
TTN	p.Tyr29208Ter	c.87624C>A	stop_gained	Reported	Class 4	6*
TTN	p.Arg15898Ter	c.47692C>T	stop_gained	Reported	Class 4	2*
TTN	Splice Site	c.48638+2T>G	splice_donor	Reported	Class 4	2*
TTN	p.Arg16724Ter	c.50170C>T	stop_gained	Reported	Class 5	2*
TTN	Splice Site	c.9704-2A>G	splice_acceptor	Novel	Class 5	2*
TTN	p.Arg34079SerfsTer9	c.102234_102237del	frameshift	Reported	Class 4	1
TTN	Splice Site	c.106531+1G>A	splice_donor	Reported	Class 5	1
TTN	p.Phe7410GlnfsTer19	c.22227_22228del	frameshift	Novel	Class 5	1
TTN	p.Gly1316Ter	c.3946G>T	stop_gained	Novel	Class 5	1
TTN	p.Lys13508AsnfsTer8	c.40521_40527del	frameshift	Novel	Class 4	1
TTN	p.Ile14101SerfsTer40	c.42300del	frameshift	Novel	Class 4	1
TTN	p.Phe14847CysfsTer7	c.44540_44541del	frameshift	Novel	Class 5	1
TTN	Splice Site	c.45350-1G>A	splice_acceptor	Novel	Class 5	1
TTN	p.Asn16662ThrfsTer8	c.49985_49988del	frameshift	Novel	Class 5	1
TTN	p.Arg18140Ter	c.54418C>T	stop_gained	Reported	Class 4	1
TTN	p.Arg20626Ter	c.61876C>T	stop_gained	Reported	Class 5	1
TTN	p.Glu20759Ter	c.62275G>T	stop_gained	Novel	Class 5	1
TTN	p.Pro21370GlnfsTer4	c.64109del	frameshift	Novel	Class 5	1
TTN	p.Asp24091GlufsTer8	c.72269_72272dup	frameshift	Novel	Class 5	1
TTN	p.Trp26308Ter	c.78923G>A	stop_gained	Novel	Class 5	1
TTN	p.Cys26564Ter	c.79692T>A	stop_gained	Novel	Class 4	1
TTN	p.Met26800AsnfsTer6	c.80398dup	frameshift	Novel	Class 5	1
TTN	p.Gly27533ValfsTer47	c.82598del	frameshift	Reported	Class 5	1
TTN	p.Glu28397Ter	c.85189G>T	stop_gained	Novel	Class 4	1
TTN	p.Arg28779Ter	c.86335C>T	stop_gained	Reported	Class 4	1
TTN	p.Phe29218LeufsTer7	c.87654del	frameshift	Novel	Class 5	1
TTN	p.Arg29947Ter	c.89839C>T	stop_gained	Reported	Class 4	1
TTN	p.Ser33240Ter	c.99719C>G	stop_gained	Reported	Class 4	1
MYH7	p.Arg1250Gly	c.3748C>G	missense	Reported	Class 4	1
MYH7	p.Pro152Arg	c.455C>G	missense	Novel	Class 4	1
MYH7	Splice Site	c.5791-2A>T	splice_acceptor	Reported	Class 4	1
MYH7	p.Gly1931AlafsTer2	c.5792del	frameshift	Novel	Class 4	1
MYBPC3	p.Arg726Cys	c.2176C>T	missense	Reported	Class 4	1
TNNT2	p.Arg141Trp	c.421C>T	missense	Reported	Class 5	1
Nuclear envelope genes						
LMNA	p.Arg190Trp	c.568C>T	missense	Reported	Class 5	4*
LMNA	p.Arg541His	c.1622G>A	missense	Reported	Class 5	1
RBM20	p.Arg634Trp	c.1900C>T	missense	Reported	Class 5	1
PRKAG2	Splice Site	c.1585-2A>G	splice_acceptor	Novel	Class 5	1
Cytoskeletal genes						
FLNC	p.Gln2303Ter	c.6907C>T	stop_gained	Reported	Class 5	2*
FLNC	p.Arg1341Ter	c.4021C>T	stop_gained	Reported	Class 5	1
FLNC	p.Ile2003AspfsTer35	c.6006dup	splice_acceptor	Novel	Class 4	1
FHL1	p.Thr152Ile	c.455C>T	missense	Novel	Class 5	1
DMD	Splice Site	c.1603-2A>G*	splice_acceptor*	Reported	Class 5	1
Cell junction genes						
PKP2	p.Ser837ValfsTer94	c.2509del	frameshift	Reported	Class 5	1

JUP	p.Asn321Ter	c.960dup	frameshift	Novel	Class 4	1
SCN5A	p.Tyr1228Cys	c.3683A>G	missense	Novel	Class 4	1
SCN5A	p.Leu1338Ile	c.4012C>A	missense	Novel	Class 4	1
Z-disc genes						
BAG3	p.Arg309Ter	c.925C>T	stop_gained	Reported	Class 5	2*
BAG3	p.Trp36Ter	c.107G>A	stop_gained	Novel	Class 5	1
BAG3	p.Glu72Ter	c.214G>T	stop_gained	Novel	Class 5	1

*** Recurring variants in probands**

The 76 probands that were class 4 or 5 positive occurred in 51.3% (39/76) BA, 31.6% (24/76) MA, and 17.1% (13/76) WA probands. The BA and MA probands had more class 5 than class 4 variant carriers, whereas the probands of WA had an equal distribution of class 4 and 5 variant carriers.

Analysis showed that the predominant genes for the 39 positive BA probands were *TTN*; 79.5% (31/39), *MYH7*; 7.7% (3/39), *FLNC*; 5.1% (2/39), *SCN5A*; 5.1% (2/39), *BAG3*; 2.6% (1/39), *DMD*; 2.6% (1/39) and *RBM20*; 2.6% (1/39). In comparison, the predominant genes in the 24 MA probands were *TTN*; 54.2% (13/24), *LMNA*; 16.7% (4/24), *BAG3*; 8.3% (2/24), *FLNC*; 8.3% (2/24), *FHL1*; 4.2% (1/24), *TNNT2*; 4.2% (1/24), *MYBPC3*; 4.2% (1/24) and *MYH7*; 4.2% (1/24). The predominant genes in the 13 WA probands were *TTN*; 61.5% (8/13), *BAG3;* 7.7% (1/13), *LMNA;* 7.7% (1/13), *PKP2;* 7.7% (1/13), *PRKAG2;* 7.7% (1/13), and *JUP;* 7.7% (1/13), as shown in Figure 3.9 shows. Thus, we further analyse and present the genetic result per the probands' self-reported ancestry.

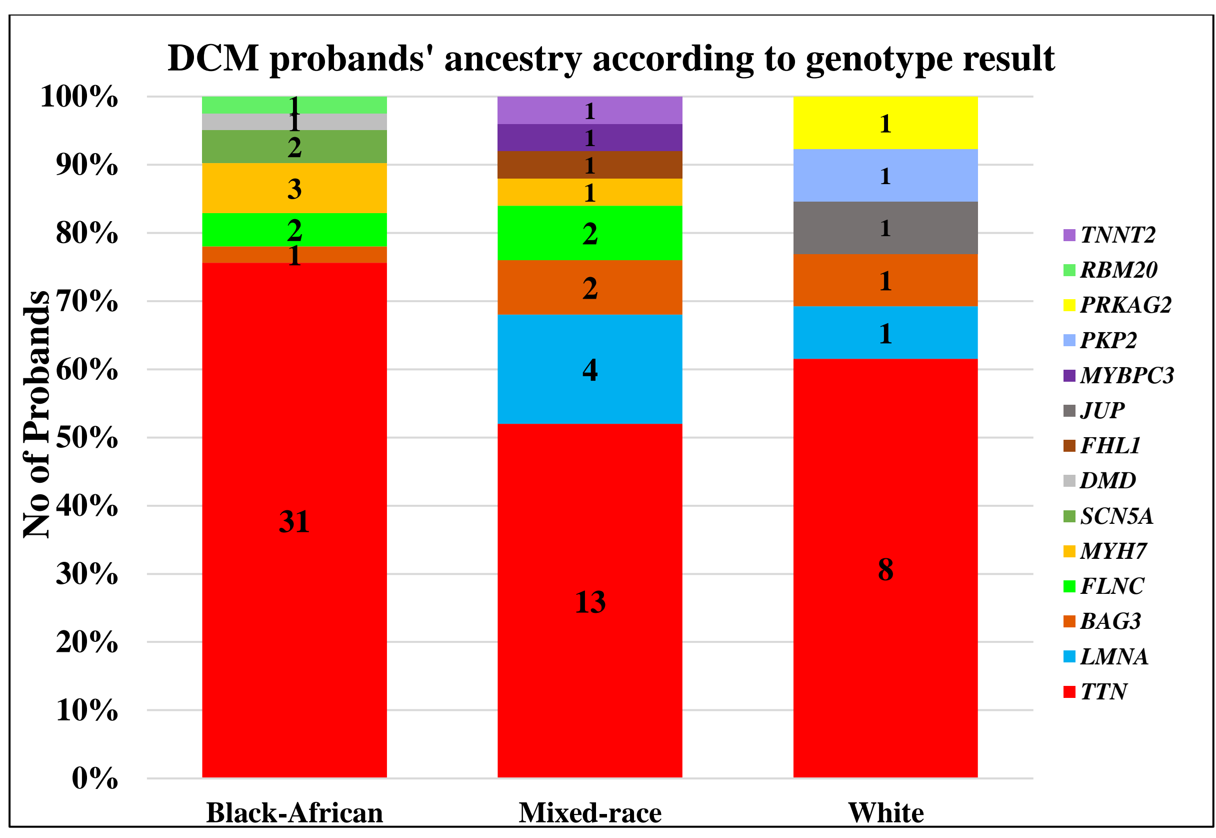
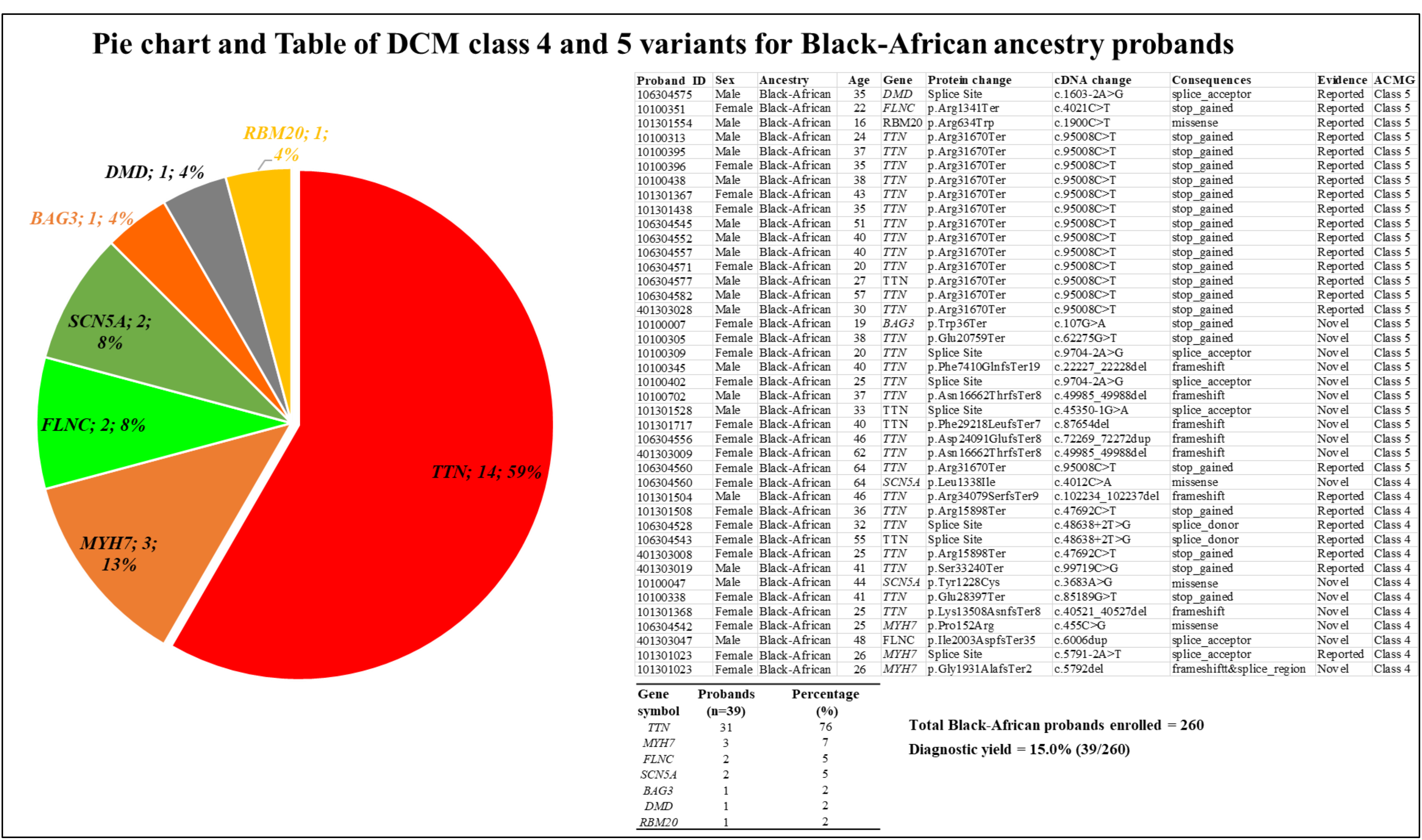


Figure 3.9: DCM probands' genotype result grouped according to ancestry. *Number of probands per gene harbouring class 4 or 5 variant according to self-reported ancestry.*

3.3.1.3.1.2.1 Class 4 and 5 variant carriers of Black-African ancestry (BA)

Of the 260 BA adult-onset DCM probands we recruited, we found 39 (63% females) carrying class 4 and 5 variants, as shown in Figure 3.10. The BA had 24 class 4 and 5 variants found in *TTN* [58.3% (14/24)], *MYH7* [12.5% (3/24)], *FLNC* [8.3% (2/24)], *BAG3* [4.2% (1/24)], *DMD* [4.2% (1/24)], *RBM20* [4.2% (1/24)] and *SCN5A* [4.2% (1/24)] genes.



Pie chart and Table of DCM class 4 and 5 variants for Black-African ancestry probands

Proband ID	Sex	Ancestry	Age	Gene	Protein change	cDNA change	Consequences	Evidence	ACMG
106304575	Male	Black-African	35	*DMD*	Splice Site	c.1603-2A>G	splice_acceptor	Reported	Class 5
10100351	Female	Black-African	22	*FLNC*	p.Arg1341Ter	c.4021C>T	stop_gained	Reported	Class 5
101301554	Male	Black-African	16	RBM20	p.Arg634Trp	c.1900C>T	missense	Reported	Class 5
10100313	Male	Black-African	24	*TTN*	p.Arg31670Ter	c.95008C>T	stop_gained	Reported	Class 5
10100395	Male	Black-African	37	*TTN*	p.Arg31670Ter	c.95008C>T	stop_gained	Reported	Class 5
10100396	Female	Black-African	35	*TTN*	p.Arg31670Ter	c.95008C>T	stop_gained	Reported	Class 5
10100438	Male	Black-African	38	*TTN*	p.Arg31670Ter	c.95008C>T	stop_gained	Reported	Class 5
101301367	Female	Black-African	43	*TTN*	p.Arg31670Ter	c.95008C>T	stop_gained	Reported	Class 5
101301438	Female	Black-African	35	*TTN*	p.Arg31670Ter	c.95008C>T	stop_gained	Reported	Class 5
106304545	Male	Black-African	51	*TTN*	p.Arg31670Ter	c.95008C>T	stop_gained	Reported	Class 5
106304552	Male	Black-African	40	*TTN*	p.Arg31670Ter	c.95008C>T	stop_gained	Reported	Class 5
106304557	Male	Black-African	40	*TTN*	p.Arg31670Ter	c.95008C>T	stop_gained	Reported	Class 5
106304571	Female	Black-African	20	*TTN*	p.Arg31670Ter	c.95008C>T	stop_gained	Reported	Class 5
106304577	Male	Black-African	27	TTN	p.Arg31670Ter	c.95008C>T	stop_gained	Reported	Class 5
106304582	Male	Black-African	57	*TTN*	p.Arg31670Ter	c.95008C>T	stop_gained	Reported	Class 5
401303028	Male	Black-African	30	*TTN*	p.Arg31670Ter	c.95008C>T	stop_gained	Reported	Class 5
10100007	Female	Black-African	19	*BAG3*	p.Trp36Ter	c.107G>A	stop_gained	Novel	Class 5
10100305	Female	Black-African	38	*TTN*	p.Glu20759Ter	c.62275G>T	stop_gained	Novel	Class 5
10100309	Female	Black-African	20	*TTN*	Splice Site	c.9704-2A>G	splice_acceptor	Novel	Class 5
10100345	Male	Black-African	40	*TTN*	p.Phe7410GlnfsTer19	c.22227_22228del	frameshift	Novel	Class 5
10100402	Female	Black-African	25	*TTN*	Splice Site	c.9704-2A>G	splice_acceptor	Novel	Class 5
10100702	Male	Black-African	37	*TTN*	p.Asn16662ThrfsTer8	c.49985_49988del	frameshift	Novel	Class 5
101301528	Male	Black-African	33	TTN	Splice Site	c.45350-1G>A	splice_acceptor	Novel	Class 5
101301717	Female	Black-African	40	TTN	p.Phe29218LeufsTer7	c.87654del	frameshift	Novel	Class 5
106304556	Female	Black-African	46	*TTN*	p.Asp24091GlufsTer8	c.72269_72272dup	frameshift	Novel	Class 5
401303009	Female	Black-African	62	*TTN*	p.Asn16662ThrfsTer8	c.49985_49988del	frameshift	Novel	Class 5
106304560	Female	Black-African	64	*TTN*	p.Arg31670Ter	c.95008C>T	stop_gained	Reported	Class 5
106304560	Female	Black-African	64	*SCN5A*	p.Leu1338Ile	c.4012C>A	missense	Novel	Class 4
101301504	Male	Black-African	46	*TTN*	p.Arg34079SerfsTer9	c.102234_102237del	frameshift	Reported	Class 4
101301508	Female	Black-African	36	*TTN*	p.Arg15898Ter	c.47692C>T	stop_gained	Reported	Class 4
106304528	Female	Black-African	32	*TTN*	Splice Site	c.48638+2T>G	splice_donor	Reported	Class 4
106304543	Female	Black-African	55	TTN	Splice Site	c.48638+2T>G	splice_donor	Reported	Class 4
401303008	Female	Black-African	25	*TTN*	p.Arg15898Ter	c.47692C>T	stop_gained	Reported	Class 4
401303019	Male	Black-African	41	*TTN*	p.Ser33240Ter	c.99719C>G	stop_gained	Reported	Class 4
10100047	Male	Black-African	44	*SCN5A*	p.Tyr1228Cys	c.3683A>G	missense	Novel	Class 4
10100338	Female	Black-African	41	*TTN*	p.Glu28397Ter	c.85189G>T	stop_gained	Novel	Class 4
101301368	Female	Black-African	25	*TTN*	p.Lys13508AsnfsTer8	c.40521_40527del	frameshift	Novel	Class 4
106304542	Female	Black-African	25	*MYH7*	p.Pro152Arg	c.455C>G	missense	Novel	Class 4
401303047	Male	Black-African	48	FLNC	p.Ile2003AspfsTer35	c.6006dup	splice_acceptor	Novel	Class 4
101301023	Female	Black-African	26	*MYH7*	Splice Site	c.5791-2A>T	splice_acceptor	Reported	Class 4
101301023	Female	Black-African	26	*MYH7*	p.Gly1931AlafsTer2	c.5792del	frameshiftt&splice_region	Novel	Class 4

Gene symbol	Probands (n=39)	Percentage (%)
TTN	31	76
MYH7	3	7
FLNC	2	5
SCN5A	2	5
BAG3	1	2
DMD	1	2
RBM20	1	2

Total Black-African probands enrolled = 260

Diagnostic yield = 15.0% (39/260)

Figure 3.10: Adult DCM probands of Black-African ancestry. *The pie chart and the corresponding tables show the molecular consequences of class 4 and 5 variants in the adult DCM probands of Black-African ancestry.*

Of the 24 variants, 37.5% (9/24) were reported, and 62.5% (15/24) were novel; 12 variants were class 5 (four reported and eight novel), and 12 were class 4 (seven reported and five novel). The majority, 83.3% (20/24), of the class 4 and 5 variants carried by the BA probands were truncating variants: seven stop-gained, seven frameshifts, five splice_acceptor and one splice_donor. The remaining four had missense molecular consequences, as shown in Figure 3.10. We also found that five of the *TTN* variants were recuring.

3.3.1.3.1.2.1.1 TTN c.95008C>T

The stop-gained class 5 variant, *TTN* c.95008C>T variant, recurred in 35.9% (14/39) of the BA probands. No proband in any other ethnic group had this variant. At the same time, another proband (ID 106304560) carried digenic variants: a *TTN*: c.95008C>T and *SCN5A*: c.4012C>A. The 14 BA probands had been classified as NYHA class II, III or IV, with an ejection fraction of ≤ 31% at diagnosis; there was no FH of disease in 78.6% (11/14) of the probands.

3.3.1.3.1.2.1.2 *TTN*tv c.47692C>T

The stop_gained *TTN*tv c.47692C>T (p.Arg15898Ter) class 4 variant recurred in 5.1% (2/39) of the BA probands with pathogenic variants. The two BA probands diagnosed at the ages of 36 and 25 were both classified as NYHA class III, with an ejection fraction of 30% and 20%, respectively. The variant creates a premature translational stop codon at residue 15898 in the *TTN* gene coding sequence.

3.3.1.3.1.2.1.3 TNN c.9704-2A>G

The novel splice-acceptor c.9704-2A>G recurred in two female BA probands diagnosed at ages 20 and 25 with an ejection fraction of 13% and 11%, respectively. The variant was not identified in any other ethnic group.

3.3.1.3.1.2.1.4 TTN c.49985_49988del

The novel frameshift c.49985_49988del variant recurred in two BA probands diagnosed at ages 37 and 62, classified as NYHA class II and III, with an ejection fraction of 30% and 35%, respectively.

3.3.1.3.1.2.1.5 RBM20 c.1900C>T

Proband (ID 101301554) is a 16-year-old BA male who had been classified as NYHA class IV, with an ejection fraction of 15%; there was no FH of heart disease. Targeted sequencing identified the *RBM20* c.1900C>T (p. Arg634Trp) missense class 5 variant as the pathogenic variant. The c.1900C>T variant replaces an Arginine with Tryptophan at residue 634 in the RS domain, an *RBM20*-related DCM hotspot. The change significantly alters the domain's properties, as a non-polar, hydrophobic amino acid with an aromatic side chain replaced the polar and hydrophilic basic amino acid. Also, computational prediction tools and conservation analysis are consistent with pathogenicity.

3.3.1.3.1.2.1.6 MYH7 c.5791-2A>T and c.5792del

Proband (ID 101301023) is a 26-year-old female proband of BA that had been classified as NYHA class II, with an ejection fraction of 41% at diagnosis; there was no FH of the disease. Targeted sequencing identified the compound heterozygous variants, the *MYH7* c.5791-2A>T reported splice_acceptor and c.5792del novel frameshift&splice-region class 4 variants as the likely cause of disease.

3.3.1.3.1.2.1.7 *MYH7* c.455C>G

Proband (ID 106304542) is a 25-year-old female proband of BA that had been classified as NYHA class II, with an ejection fraction of 23% at diagnosis; there was no FH of heart disease. Targeted sequencing identified the *MYH7* c.455C>G (p.Pro152Arg) class 4 variant as the pathogenic variant. The novel *MYH7* c.455C>G missense variant (Align GVGD: Class C65, CADD: 24.7, PolyPhen2: Probably damaging, SIFT: Deleterious) was predicted *in silico* to cause conformational changes leading to structural instability at the myosin motor binding site, thus causing a severe *MYH7*-related DCM in the proband. The c.455C>G variant replaces a Proline amino acid with Arginine. Proline is a nonpolar amino acid with a relatively rigid structure, while Arginine is a positively charged amino acid with a more flexible and large side chain. This property difference can change the protein's sequence size and charge at position 152.

3.3.1.3.1.2.1.8 *FLNC* c.4021C>T

Proband (ID 10100351) is a 22-year-old BA female who had been classified as NYHA class III, with an ejection fraction of 14%; there was no FH of the disease. Targeted sequencing identified the *FLNC* c.4021C>T (p.Arg1341Ter) class 5 variant as the likely cause of disease. The variant creates a premature translational stop codon at residue 1341 in the *FLNC* gene coding sequence.

3.3.1.3.1.2.1.9 FLNC c.6006dup

Proband (ID 401303047) is a 48-year-old BA male who had been classified as NYHA class IV, with an ejection fraction of 22%, and there was no FH of the disease. Targeted sequencing identified a novel *FLNC* c.6006dup (p.Ile2003AspfsTer35) class 4 variant as the likely cause of disease. The variant replaces Isoleucine with Aspartic acid, shifting the reading frame and terminating at Ter35.

3.3.1.3.1.2.1.10 BAG3 c.107G>A

Proband (ID 10100007) is a 19-year-old BA female that had been classified as NYHA class III, with an ejection fraction of 35%, and there was no FH of the disease. Targeted sequencing identified the *BAG3* c.107G>A (p.Trp36Ter) class 5 variant as the likely cause of disease. The novel *BAG3* c.107G>A in exon 1 is expected to introduce a premature translation termination codon at amino acid position 36 in a highly conserved residue within the WW domain of BAG3 protein.

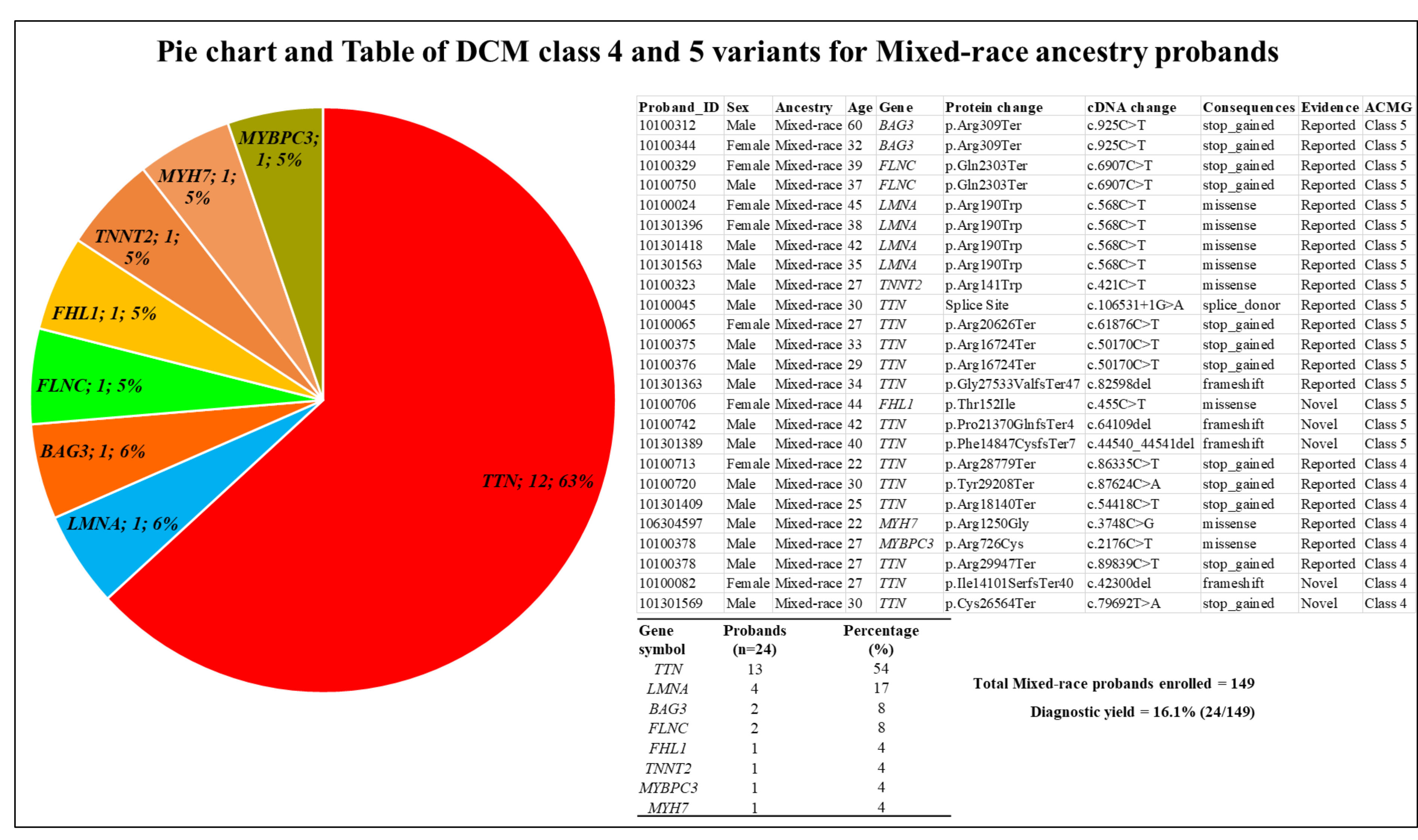
3.3.1.3.1.2.1.11 DMD c.1603-2A>G

Proband (ID 106304575) is a 35-year-old BA male classified as NYHA class II with an ejection fraction of 20%; there was no FH of the disease. Targeted sequencing identified the *DMD* c.1603-2A>G class 5 variant as the likely cause of the disease. The c.1603-2A>C variant is predicted to cause a natural splice acceptor site loss on exon 14 and cause exon 14-15 skipping.

3.3.1.3.1.2.2 Class 4 and 5 variant carriers of Mixed ancestry (MA)

Of the 149 BA adult-onset DCM probands we recruited, we found 24 (50% females) carrying class 4 and 5 variants. The MA had 19 class 4 and 5 variants found in *TTN* [63.2% (12/19)],

MYH7 [5.3% (1/19)], *FLNC* (5.3%), *BAG3* (5.3%), *LMNA* (5.3%), *TNNT2* (5.3%), *MYBPC3* (5.3%) and *FHL1* (5.3%) genes.



Proband_ID	Sex	Ancestry	Age	Gene	Protein change	cDNA change	Consequences	Evidence	ACMG
10100312	Male	Mixed-race	60	*BAG3*	p.Arg309Ter	c.925C>T	stop_gained	Reported	Class 5
10100344	Female	Mixed-race	32	*BAG3*	p.Arg309Ter	c.925C>T	stop_gained	Reported	Class 5
10100329	Female	Mixed-race	39	*FLNC*	p.Gln2303Ter	c.6907C>T	stop_gained	Reported	Class 5
10100750	Male	Mixed-race	37	*FLNC*	p.Gln2303Ter	c.6907C>T	stop_gained	Reported	Class 5
10100024	Female	Mixed-race	45	*LMNA*	p.Arg190Trp	c.568C>T	missense	Reported	Class 5
101301396	Female	Mixed-race	38	*LMNA*	p.Arg190Trp	c.568C>T	missense	Reported	Class 5
101301418	Male	Mixed-race	42	*LMNA*	p.Arg190Trp	c.568C>T	missense	Reported	Class 5
101301563	Male	Mixed-race	35	*LMNA*	p.Arg190Trp	c.568C>T	missense	Reported	Class 5
10100323	Male	Mixed-race	27	*TNNT2*	p.Arg141Trp	c.421C>T	missense	Reported	Class 5
10100045	Male	Mixed-race	30	*TTN*	Splice Site	c.106531+1G>A	splice_donor	Reported	Class 5
10100065	Female	Mixed-race	27	*TTN*	p.Arg20626Ter	c.61876C>T	stop_gained	Reported	Class 5
10100375	Male	Mixed-race	33	*TTN*	p.Arg16724Ter	c.50170C>T	stop_gained	Reported	Class 5
10100376	Male	Mixed-race	29	*TTN*	p.Arg16724Ter	c.50170C>T	stop_gained	Reported	Class 5
101301363	Male	Mixed-race	34	*TTN*	p.Gly27533ValfsTer47	c.82598del	frameshift	Reported	Class 5
10100706	Female	Mixed-race	44	*FHL1*	p.Thr152Ile	c.455C>T	missense	Novel	Class 5
10100742	Male	Mixed-race	42	*TTN*	p.Pro21370GlnfsTer4	c.64109del	frameshift	Novel	Class 5
101301389	Male	Mixed-race	40	*TTN*	p.Phe14847CysfsTer7	c.44540_44541del	frameshift	Novel	Class 5
10100713	Female	Mixed-race	22	*TTN*	p.Arg28779Ter	c.86335C>T	stop_gained	Reported	Class 4
10100720	Male	Mixed-race	30	*TTN*	p.Tyr29208Ter	c.87624C>A	stop_gained	Reported	Class 4
101301409	Male	Mixed-race	25	*TTN*	p.Arg18140Ter	c.54418C>T	stop_gained	Reported	Class 4
106304597	Male	Mixed-race	22	*MYH7*	p.Arg1250Gly	c.3748C>G	missense	Reported	Class 4
10100378	Male	Mixed-race	27	*MYBPC3*	p.Arg726Cys	c.2176C>T	missense	Reported	Class 4
10100378	Male	Mixed-race	27	*TTN*	p.Arg29947Ter	c.89839C>T	stop_gained	Reported	Class 4
10100082	Female	Mixed-race	27	*TTN*	p.Ile14101SerfsTer40	c.42300del	frameshift	Novel	Class 4
101301569	Male	Mixed-race	30	*TTN*	p.Cys26564Ter	c.79692T>A	stop_gained	Novel	Class 4

Gene symbol	Probands (n=24)	Percentage (%)
TTN	13	54
LMNA	4	17
BAG3	2	8
FLNC	2	8
FHL1	1	4
TNNT2	1	4
MYBPC3	1	4
MYH7	1	4

Figure 3.11: Adult DCM probands of Mixed ancestry. *The pie chart and the corresponding tables show the molecular consequences of class 4 and 5 variants in the adult DCM probands of Mixed ancestry.*

Of the 19 variants, 73.7% (14/19) were reported, and 26.3% (5/24) were novel; 11 variants were class 5 (eight reported and three novel), and eight were class 4 (six reported and two novel). Unlike in the BA probands, where 62.5% (15/24) of the variants were unknown, we found only five (26.3%) novel variants (Figure 3.11) in the MA probands. The majority, 73.7% (14/19), of the class 4 and 5 variants carried by the MA probands were truncating variants: nine stop-gained, four frameshifts and one splice_donor. The remaining five had missense molecular consequences, as shown in Figure 3.11.

Like in the BA probands, most (63.2%) of these MA probands carried *TTN*tv. However, we noted more genes with recuring variants in the MA probands cohort.

3.3.1.3.1.2.2.1 LMNA c.568C>T

The missense *LMNA c.568C>T* (p.Arg190Trp) variant recurred in 16.7% (4/24) of the genotype-positive probands of MA. Two of the four MA were females diagnosed with familial DCM, and the other two were males diagnosed with unknown aetiology for the disease. The four MA probands were diagnosed between the ages of 35 and 45. The missense *LMNA c.568C>T* is on the lamin A protein's central rod 1B region. The variant replaces a basic and polar Arginine amino acid with Tryptophan, which is a neutral and slightly polar residue.

3.3.1.3.1.2.2.2 BAG3 c.925C>T

The stop_gained *BAG3* c.925C>T (p.Arg309Ter) variant recurred in two MA probands. The variant was not identified in any other ethnic group. The two probands diagnosed at 32 and 60 with idiopathic DCM were classified as NYHA class IV, with an ejection fraction of 27% and 38%, respectively. The variant is in exon 4 of the *BAG3* and creates a premature stop codon at position 309 of the protein sequence.

3.3.1.3.1.2.2.3 FLNC c.6907C>T

The *FLNC* c.6907C>T (p.Gln2303Ter) variant recurred in two MA probands. The variant was not identified in any other ethnic group. The first MA proband is a 39-year-old female diagnosed with severe familial DCM, and the second is a 37-year-old male diagnosed with severe idiopathic DCM. Both probands are classified as NYHA class III, with an ejection

fraction of 21% and 12%, respectively. The variant is in exon 20 of the filamin C ROD2 domain and is expected to produce no or shortened filamin-C protein.

3.3.1.3.1.2.2.4 *TTN* c.50170C>T

The stop_gained *TTN* c.50170C>T (p.Arg16724Ter) variant recurred in two MA probands. The variant was not identified in any other ethnic group in this cohort. The two probands diagnosed at the ages of 29 and 33 as having familial DCM were classified as NYHA class II and III, with an ejection fraction of 36% and 20%, respectively.

3.3.1.3.1.2.2.5 *TNNT2* c.421C>T

We found a 27-year-old male proband of MA diagnosed with familial DCM carrying one *TNNT2* c.421C>T (p.Arg141Trp) missense class 5 variant. The 27-year-old had been classified as NYHA class II, with an ejection fraction of 30% and had a FH of HF.

3.3.1.3.1.2.2.6 TTN c.82598del

Proband (ID 101301363) is a 34-year-old MA male classified as NYHA class III, with an ejection fraction of 11% and no FH of the disease. Targeted sequencing identified the frameshift *TTN* c.82598del (p.Gly27533ValfsTer47) class 5 variant as the likely cause of disease. The *TTN* c.82598del results from a nucleotide deletion that creates a translational frameshift predicted to stop on the 47th codon downstream. The affected exon is in the A-band region of the titin protein.

3.3.1.3.1.2.2.7 TTN c.89839C>T and MYBPC3 c.2176C>T

We also identified proband (ID 10100378), a 27-year-old MA male classified as NYHA class III with an ejection fraction of 12%; there was a FH of the disease. Targeted sequencing identified the *TTN* c.89839C>T stop_gained and *MYBPC3* c.2176C>T missense class 4 variants as the likely cause of disease. The *TTN* c.89839C>T is more likely to cause DCM; however, the already published c.2176C>T variant is in exon 23 of the coding DNA sequence and replaces an Arginine amino acid residue 726 with Cysteine in the immunoglobulin-like domain. Arginine has a positive charge and a relatively large side chain, while Cysteine is neutral with a small side chain. Cysteine's ability to form disulfide bonds can cause structural

changes. Thus, replacing Arginine with Cysteine in the immunoglobulin-like domain can alter its size, charge, function, and structure. This missense variant is also in a highly conserved cMyBP-C protein region associated with adult onset of CMO.

3.3.1.3.1.2.3 Class 4 and 5 variant carriers of White ancestry (WA)

We identified 13 (32.5%) of the 40 WA probands recruited with class 4 and 5 variants. The 13 probands carried nine variants we found in six genes: *TTN* [44.4% (4/9)], *BAG3* [11.1% (1/9)], *LMNA (*11.1%), *PKP2 (*11.1%), *PRKAG2 (*11.1%) and *JUP (*11.1%), as shown in Figure 3.12. The *TTN* variants were three stop-gained (c.3946G>T, c.78923G>A and c.87624C>A) and one frameshift (c.80398dup), the *BAG3* c.214G>T was also a stop_gained variant, the *PKP2* c.2509del and *JUP* c.960dup variants were frameshift, the *LMNA* c.1622G>A variant was a missense, and the *PRKAG2* harboured a c.1585-2A>G splice-acceptor variant. Six (66.7%) class 4 and 5 variants were novel; three were class 5 *TTN*tv.

3.3.1.3.1.2.3.1 TTN: c.87624C>A

We also identified stop_gained *TTN* c.87624C>A (p.Tyr29208Ter) class 4 variant that recurred in five WA probands and one MA proband. The six probands had been classified as NYHA class II, III or IV, with an ejection fraction of ≤ 51% at diagnosis; there was a FH of heart disease in 66.7% (4/6) of the probands.

3.3.1.3.1.2.3.2 LMNA c.1622G>A

Proband (ID 10100043) is a 16-year-old WA male who had been classified as NYHA class II, with an ejection fraction of 26%; there was a FH of the disease. Targeted sequencing identified the missense *LMNA* c.1622G>A (p.Arg541His) class 5 variant as the likely cause of the disease. The *LMNA* c.1622G>A variant alters the amino acid composition, specifically from Arginine to Histidine (with similar polarity and positive charge) at residue 541 on the tail domain of the lamin A protein.

3.3.1.3.1.2.3.3 BAG3 c.214G>T

Proband (ID 10100328) is a 35-year-old WA male who had been classified as NYHA class IV, with an ejection fraction of 17%; there was a FH of the disease. Targeted sequencing identified

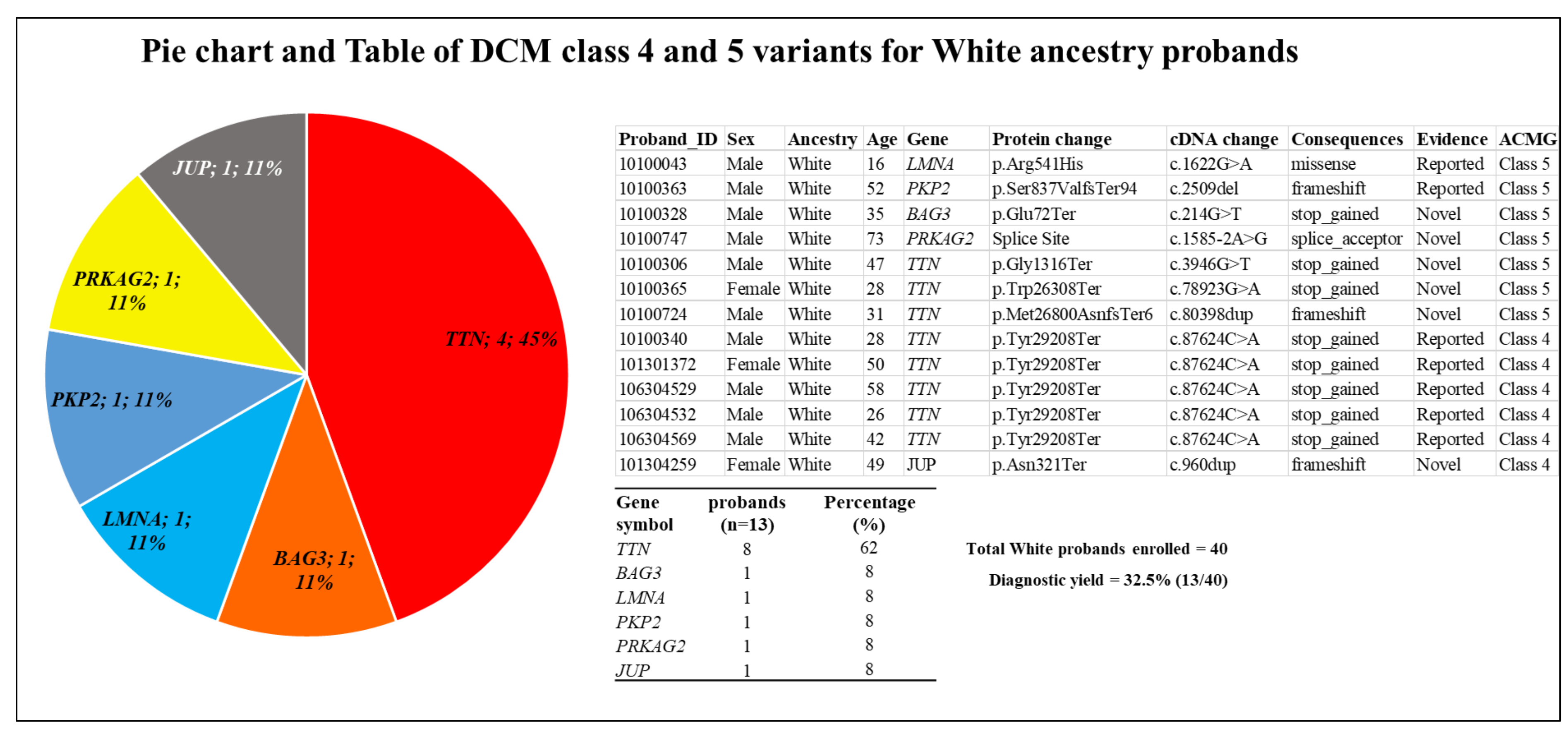
the missense *BAG3* c.214G>T (p.Glu72Ter) class 5 variant as the likely cause of the disease. The novel *BAG3* c.214G>T (p.Glu72Ter) stop-gained variant is in exon 2 of this molecular chaperon protein. The c.214G>T variant introduced a premature translation termination codon at amino acid position 76 in a moderately conserved amino acid residue region.

3.3.1.3.1.2.3.4 TTN c.80398dup

Proband (ID 10100724) is a 31-year-old WA male who had been classified as NYHA class III, with an ejection fraction of 13%; there was a FH of the disease. Targeted sequencing identified the novel frameshift *TTN* c.80398dup (p.Met26800AsnfsTer6) class 5 variant as the likely cause of disease. The variant is in the titin A-band region and is expected to create a truncated protein.

3.3.1.3.1.2.3.5 *PKP2* c.2509del

Proband (ID 10100363) is a 52-year-old WA male who had been classified as NYHA class IVI, with an ejection fraction of 49%; there was no FH of the disease. Targeted sequencing identified the frameshift *PKP2* c.2509del (p.Ser837ValfsTer94) class 5 variant as the likely cause of disease. The c.2509del variant is expected to create a longer protein product.



Proband_ID	Sex	Ancestry	Age	Gene	Protein change	cDNA change	Consequences	Evidence	ACMG
10100043	Male	White	16	*LMNA*	p.Arg541His	c.1622G>A	missense	Reported	Class 5
10100363	Male	White	52	*PKP2*	p.Ser837ValfsTer94	c.2509del	frameshift	Reported	Class 5
10100328	Male	White	35	*BAG3*	p.Glu72Ter	c.214G>T	stop_gained	Novel	Class 5
10100747	Male	White	73	*PRKAG2*	Splice Site	c.1585-2A>G	splice_acceptor	Novel	Class 5
10100306	Male	White	47	*TTN*	p.Gly1316Ter	c.3946G>T	stop_gained	Novel	Class 5
10100365	Female	White	28	*TTN*	p.Trp26308Ter	c.78923G>A	stop_gained	Novel	Class 5
10100724	Male	White	31	*TTN*	p.Met26800AsnfsTer6	c.80398dup	frameshift	Novel	Class 5
10100340	Male	White	28	*TTN*	p.Tyr29208Ter	c.87624C>A	stop_gained	Reported	Class 4
101301372	Female	White	50	*TTN*	p.Tyr29208Ter	c.87624C>A	stop_gained	Reported	Class 4
106304529	Male	White	58	*TTN*	p.Tyr29208Ter	c.87624C>A	stop_gained	Reported	Class 4
106304532	Male	White	26	*TTN*	p.Tyr29208Ter	c.87624C>A	stop_gained	Reported	Class 4
106304569	Male	White	42	*TTN*	p.Tyr29208Ter	c.87624C>A	stop_gained	Reported	Class 4
101304259	Female	White	49	JUP	p.Asn321Ter	c.960dup	frameshift	Novel	Class 4

Gene symbol	probands (n=13)	Percentage (%)
TTN	8	62
BAG3	1	8
LMNA	1	8
PKP2	1	8
PRKAG2	1	8
JUP	1	8

Total White probands enrolled = 40

Diagnostic yield = 32.5% (13/40)

Figure 3.12: Adult DCM probands of White ancestry. *The pie chart and the corresponding tables show the molecular consequences of class 4 and 5 variants in the adult DCM probands of White ancestry.*

3.3.1.3.1.3 Family segregation of class 4 and 5 variants

We screened 15 first-degree relatives of five probands with class 4 and 5 variants. Seven of the 15 relatives we screened were positive for the variants tested, as shown in Figures 3.13 to 3.16.

3.3.1.3.1.3.1 Family of 10100344 and 10100312 with BAG3 c.925C>T variant

We validated a known *BAG3* (p.Arg309Ter) c.925C>T stop_gained variant and screened the relatives' available gDNA samples in two families. In the first family, the proband (II:4) with study ID 10100312, the proband's second child (III:2) and his two grandchildren (IV:1 and IV:2) shown in the family pedigree in Figure 3.13 were positive (+/-) for the heterozygous *BAG3* c.925C>T variant. The III:2 participant was diagnosed with heart disease at the age of 55 years old with DCM, and her son (IV:2) was diagnosed with DCM at the age of 26 years old.

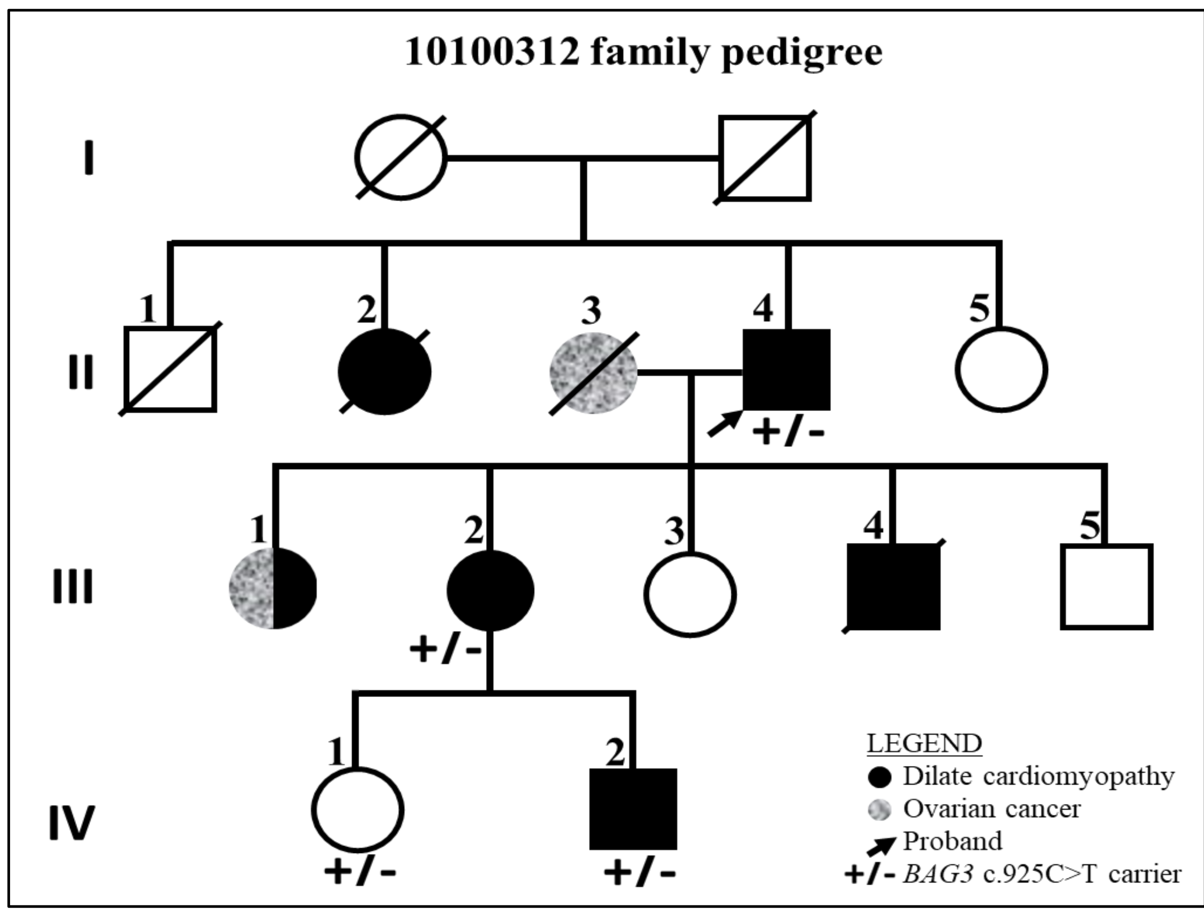


Figure 3.13: Family segregation of *BAG3* c.925C>T variant. *Pedigree chart showing proband and relatives clinically diagnosed with dilated cardiomyopathy and those tested positive for BAG3 c.925C>T variant.*

We also screened the relatives of a second family (pedigree not shown) with proband study ID 10100344, carrying the same *BAG3* c.925C>T variant. Of the three relatives screened, only the proband's clinically affected 22-year-old daughter was positive for the variant.

3.3.1.3.1.3.2 Family of 10100323 with TNNT2 c.421C>T variant

One sibling of proband (II:2) with ID 10100323 was screened for *TNNT2* (p.Arg141Trp) c.421C>T missense variant, as shown in the family pedigree in Figure 3.14. The proband's younger at-risk sister (II:4), suspected of having DCM at the age of 39 years, was found carrying the *TNNT2* c.421C>T variant. There was no sample for their deceased mother, who had a history of heart disease.

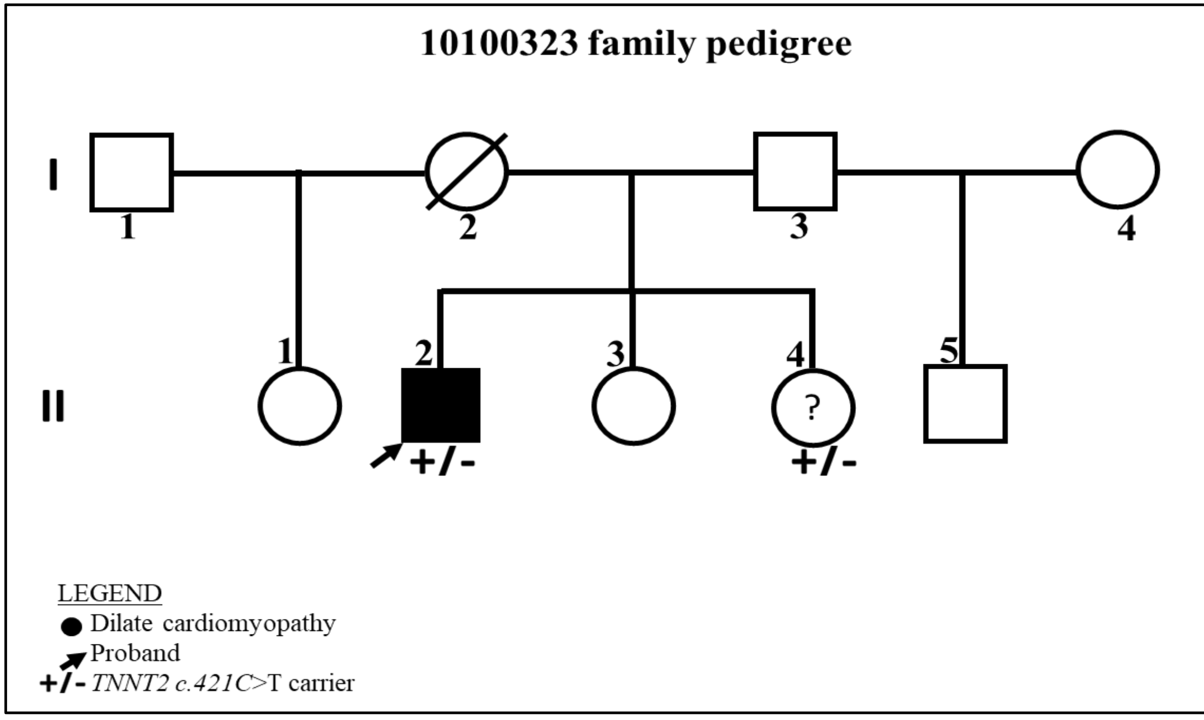


Figure 3.14: Family segregation of *TNNT2* c.421C>T variant. *Pedigree chart showing proband and relatives clinically diagnosed with dilated cardiomyopathy and those tested positive for TNNT2 c.421C>T variant.*

3.3.1.3.1.3.3 Family of 10100351 with FLNC c.4021C>T variant

The *FLNC* (p.Arg1341Ter) c.4021C>T stop_gained variant was carried by a 22-year-old female proband (II:4) with ID 10100351 that was validated, and relatives were screened. The affected older brother (II:1), clinically diagnosed with DCM at age 29 years, was positive for the variant, as shown in the family pedigree in Figure 3.15.

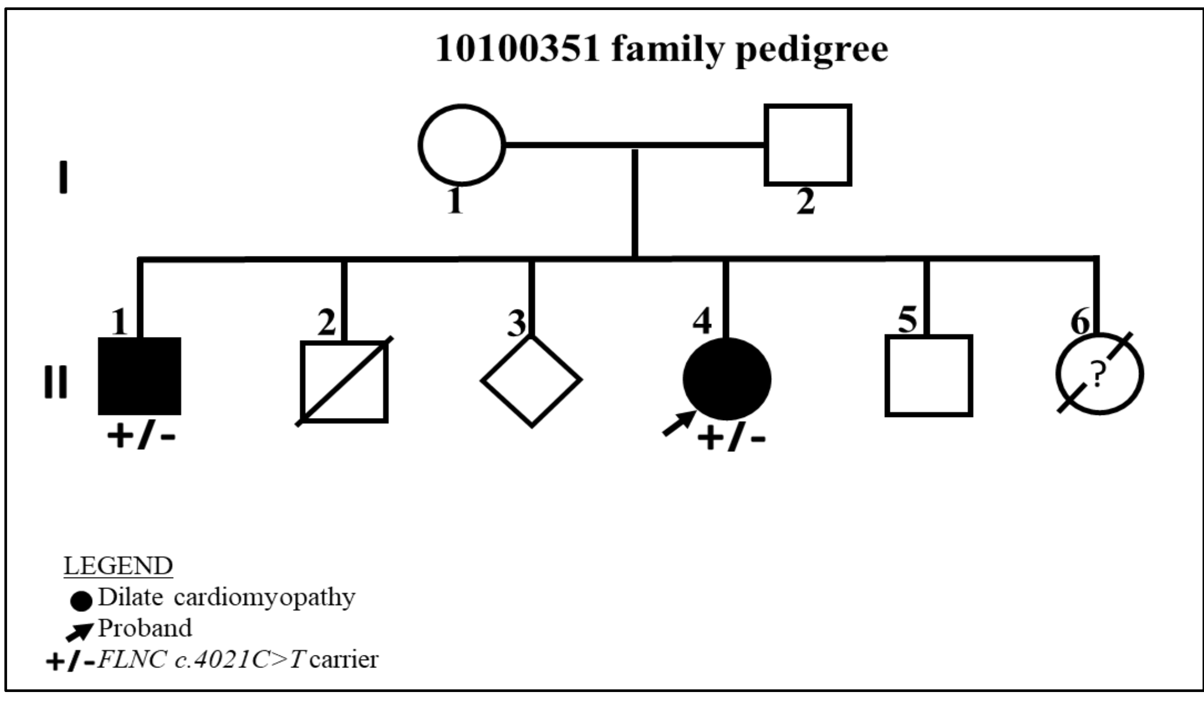


Figure 3.15: Family segregation of *FLNC* c.4021C>T variant. *Pedigree chart showing proband and relatives clinically diagnosed with dilated cardiomyopathy and those tested positive for FLNC c.4021C>T variant.*

3.3.1.3.1.3.4 Family of 101304303 with SCN5A c.589G>T variant

The proband (II:2) with ID 101304303, as shown in the family pedigree in Figure 3.16, carried the *SCN5A (*p.Asp197Tyr) c.589G>T missense variant. The proband was diagnosed at age 18 years and is a trio of family members whose at-risk parents were screened. The proband's father (I:2), who was 43 years old at the time of referral for familial DCM clinical assessment, tested positive for the variant. Neither affected relative carried any other suspected variant of interest that may be associated with heritable CMO.

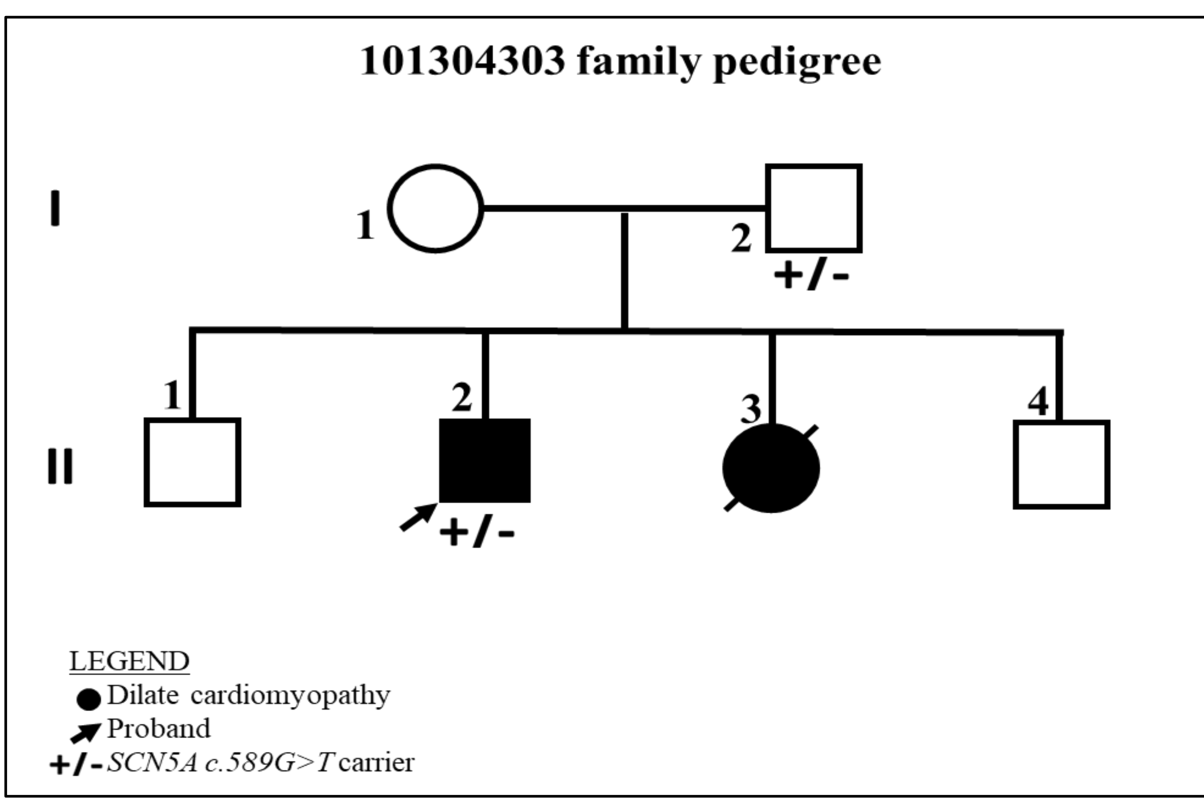


Figure 3.16: Family segregation of *SCN5A* c.589G>T variant. *Pedigree chart showing proband and relatives clinically diagnosed with dilated cardiomyopathy and those tested positive for SCN5A c.589G>T variant.*

3.3.1.3.2 Probands with variants of uncertain significance (class 3)

For probands with class 3 variants, we only focussed on the diagnostic gene list (n=18) provided by the OMGL, as these were the most likely genes to impact DCM. The genes were *ACTC1, ACTN2, BAG3, DES, DMD, DSP, FLNC, LMNA, MYH7, PLN, RBM20, SCN5A, TNNC1, TNNI3, TNNT2, TPM1, TTN* and *VCL.*

A total of 223 adult-onset DCM probands carried only class 3 variants: 137 BA, 76 MA, and 10 White probands, as shown in Table 3.3. The 223 probands had 29 genes harbouring 213 class 3 variants listed in SS Table 2. The class 3 variants were further analysed, and we found that most of the class 3 variants, 55% (123/213), were novel. Of these unknown class 3 variants, 13.8% (17/123) were truncating variants carried by 5.4% (12/223) of the adult-onset probands. Also, 38% (85/223) had both novel and reported class 3 variants. We summarised these probands' clinical and genotyped characteristics regarding their onset age compared to their self-reported sex and ancestry.

3.3.1.3.2.1 Baseline characteristics

Data analysis of the adult probands carrying only class 3 variants showed they were diagnosed at a mean age of 34.9 years (SD 11.5, range 14-73) in Table 3.3. There was thus no significant difference between the overall adult DCM age of diagnoses and probands positive for only class 3 variants. However, when we proceeded to separate the uncertain probands into groups of class 3 reported and novel variants, we noted a significant age difference (p=0.025) between the BA probands [mean age 33.7 years (SD 10.8, range 16-73)] in Figure 3.17. Figure 3.17 shows that the BA probands with novel class 3 variants were older than those carrying reported class 3 variants.

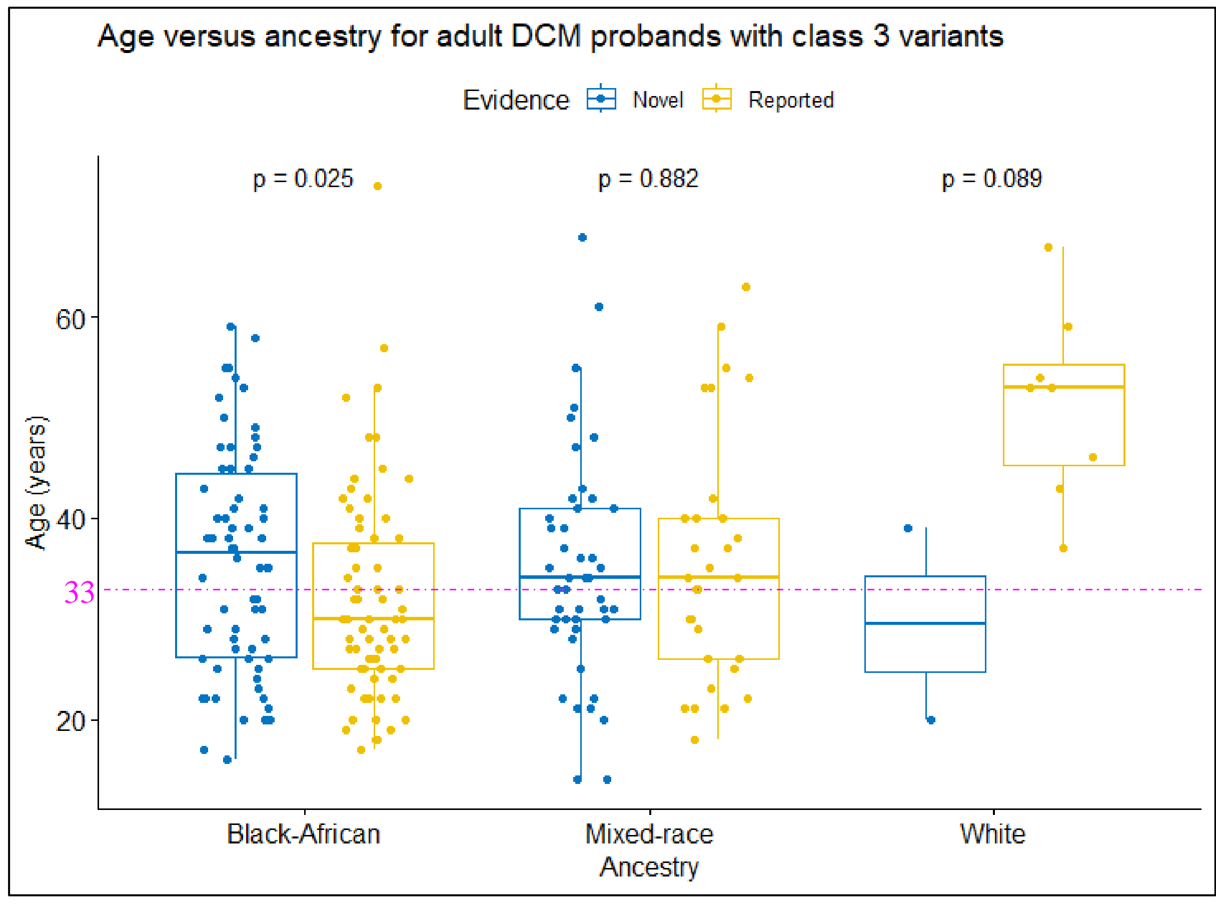


Figure 3.17: Adult DCM class 3 variant reported evidence. *The age distribution according to population ancestry for adult-onset DCM probands carrying only class 3 variants grouped by the variants' reported evidence.*

3.3.1.3.2.2 Diagnostic findings

We also noted a 1:1.5 ratio of male-to-female DCM probands carrying class 3 variants. When we separated the uncertain probands group into females and males, we noted a significant age difference (p=0.0163) between the BA and MA probands (p=0.0027), as shown in Figure 3.18. Figure 3.18 shows that the female adult DCM probands carrying class 3 variants were younger

than those with the same ancestry. We proceed further to investigate the genetic variants found per ancestry.

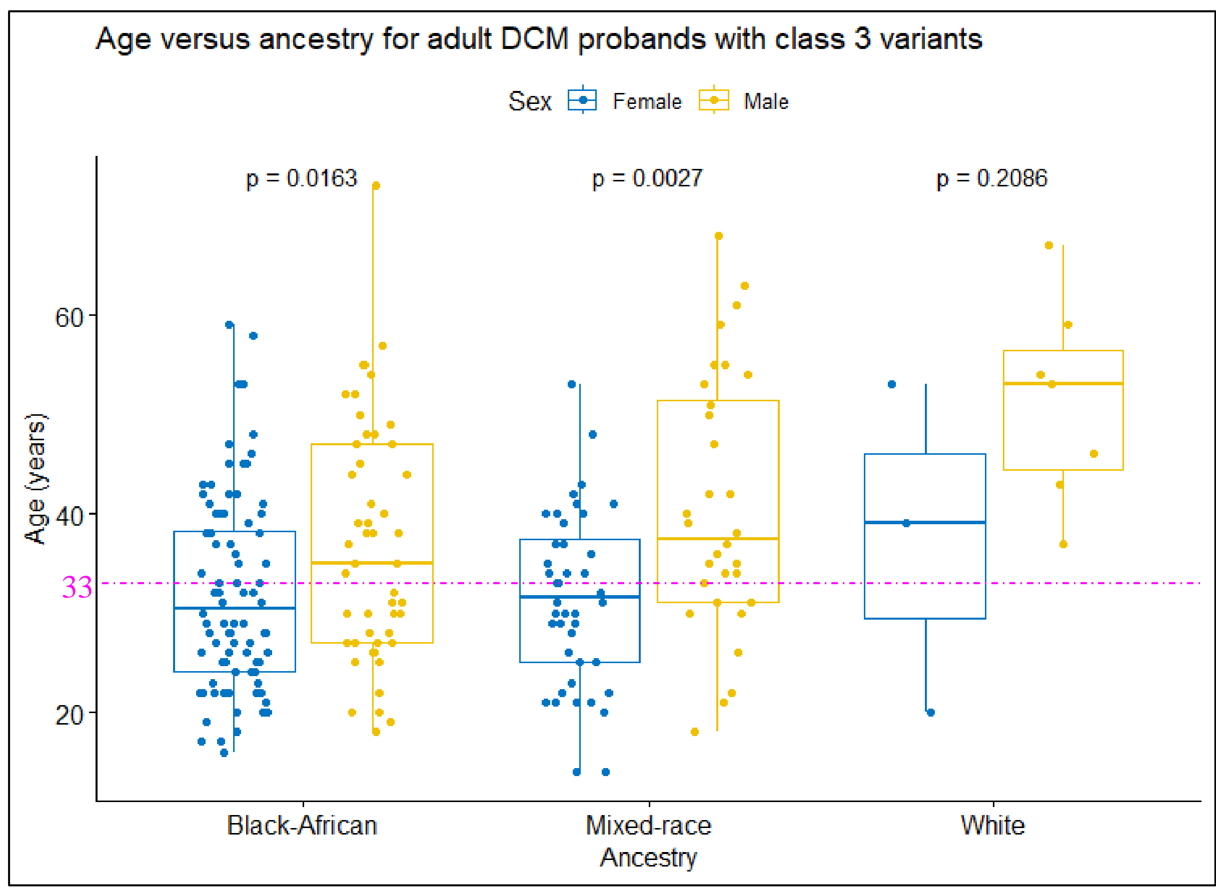


Figure 3.18: Adult DCM class 3 variant distribution according to sex. *The age distribution according to population ancestry for adult-onset DCM probands carrying only class 3 variants grouped by sex assigned at birth.*

The class 3 variants we found are listed in SS Table 2. In this section, we report on nine class 3 variants that barely met the class 4 variants requirement, being LP. We found these nine variants in 4.0% (9/223) of the class 3 variant carriers. The variants had missense molecular consequences and were either not scored or predicted *in silico* to have deleterious effects on the proteins involved by SIFT or PolyPhen2. The missense variants' biophysical characteristics prediction scores are shown in Table 3.6. We also summarised the findings listed in SS Table 2 for the variants of uncertain significance in this section per the probands' self-reported ancestry. In the SS Table 2, we have listed class 3 truncating variants in *DSP*, *TTN* and *VCL* to be re-evaluated as they lack sufficient evidence to meet class 4.

Table 3.6: Class 3 variants of interest in the adult-onset DCM probands

DEMOGRAPHICS				GENETICS				MISSENSE PREDICTION TOOLS		
Proband_ID	**Gender**	**Ancestry**	**Age**	**Gene**	**Protein change**	**cDNA change**	**Evidence**	**CADD**	**Align GVGD**	**MT**
10100394	Female	Black-African	26	*BAG3*	p.Pro234Ser	c.700C>T	rs1021293076	26.6	Class C0	B
101301508	Female	Black-African	36	*DMD*	p.Leu74Met	c.220C>A	None	25.3	Class C0	B
101301523	Male	Mixed	35	*DSP*	p.Asp468Asn	c.1402G>A	None	32.0	Class C0	D
102302041	Female	Mixed	14	*MYH7*	p.Asn977Ser	c.2930A>G	None	25.4	Class C45	D
106304547	Female	Black-African	45	*LMNA*	p.Asn142Lys	c.426C>A	None	22.5	Class C65	D
106304572	Female	Black-African	22	*FLNC*	p.Lys162Glu	c.484A>G	None	27.9	Class C55	D

CADD- COMBINED ANNOTATION DEPENDENT DEPLETION, ALIGN GVGD- GRANTHAM VARIATION GRANTHAM DIFFERENCE SCORE, MT-MUTATION TASTER, D-DELETERIOUS, B-BENIGN

3.3.1.3.2.2.1 Class 3 variant carriers of Black-African ancestry (BA)

We found 137 BA adult-onset DCM probands carrying only class 3 variants: 64.2% (88/137) females and 35.8% (49/137) males. The 137 BA adult-onset DCM probands carried 25 genes harbouring 130 class 3 variants.

Of the 130 class 3 variants carried by the BA probands with only class 3 variants, 58.5% (76/130) were novel, and 41.5% (54/130) were reported. The 76 novel and 54 reported class 3 variants were carried by 39 and 98 adult-onset BA probands, respectively. We also found that only 47.4% (65/137) of the BA probands with only class 3 variants carried multiple variants, and 52.6% (72/137) had only one class 3 variant each. The majority, 93.1% (121/130), of the class 3 variants in the BA probands had a missense molecular consequence. Many of these missense variants were in the *TTN* gene and are likely benign mutations. The remaining nine variants were six frameshift, three inframe_deletion and one splice-donor variant.

Figure 3.19 presents the nine genes with most of the class 3 variants found in the BA probands that are relevant to cause DCM. Figure 3.19 showed that most BA probands carrying only class 3 variants had a *TTN* gene variant.

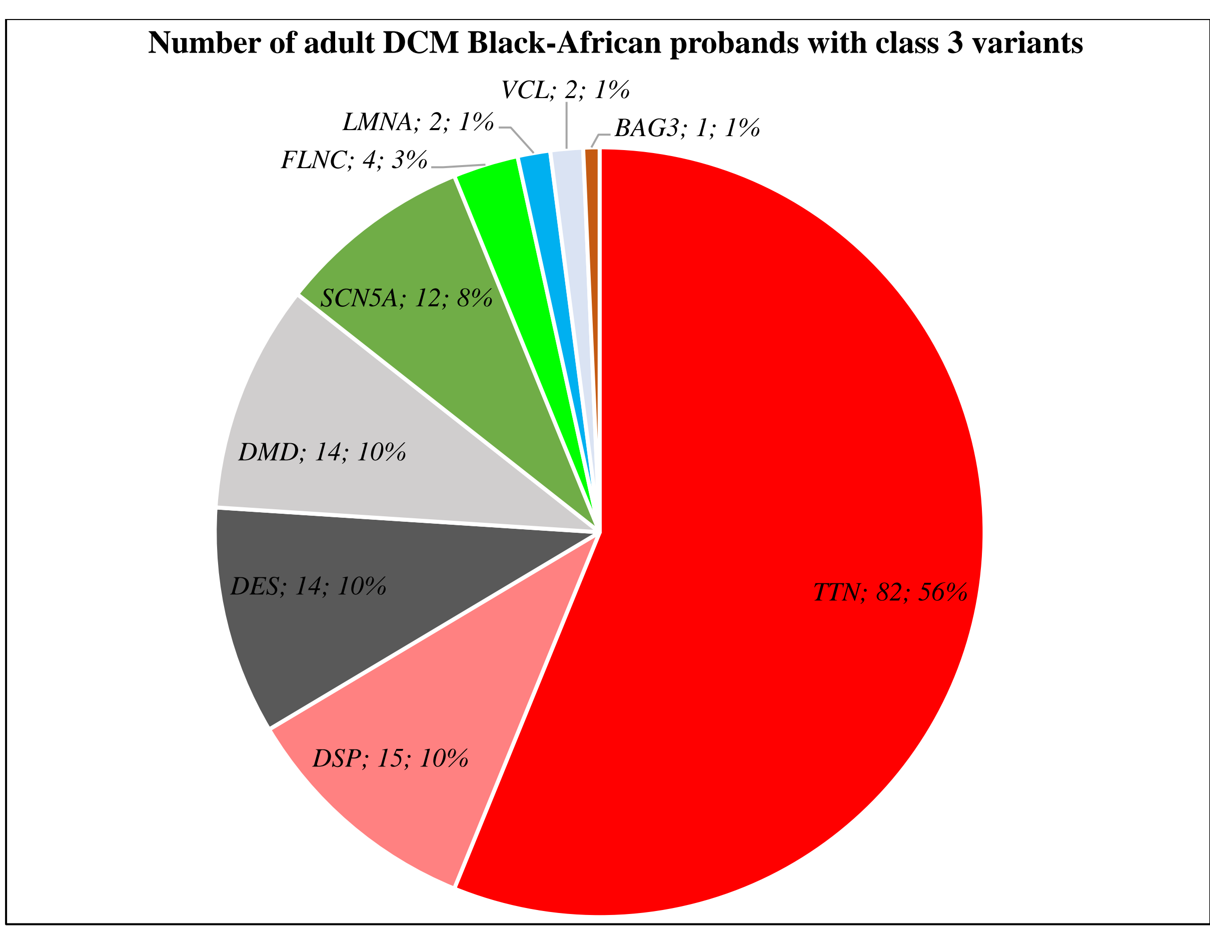


Figure 3.19: Adult DCM class 3 variant carriers of Black-African ancestry. *The number of BA adult-onset probands carrying only class 3 variants and the gene harbouring the class 3 variants.*

3.3.1.3.2.2.2 Class 3 variant carriers of Mixed ancestry (MA)

We found 76 MA adult-onset DCM probands carrying only class 3 variants: 57.9% (44/76) females and 42.1% (32/76) males. The 76 MA adult-onset DCM probands carried 17 genes harbouring 102 class 3 variants.

Of the 102 class 3 variants carried by the MA probands with only class 3 variants, 55.9% (57/102) were novel, and 44.1% (45/102) were reported class 3 variants. The 57 novel and 45 reported class 3 variants were carried by 80 and 68 adult-onset MA probands, respectively. Interestingly, only 44.7% (34/76) of the MA probands with only class 3 variants carried multiple variants, and 55.3% (42/76) had only one class 3 variant each. The majority, 92.2% (94/102), of the class 3 variants in the MA probands had a missense molecular consequence. The remaining eight variants were four frameshift, two splice-donor, one inframe_deletion and one splice-acceptor variant.

Figure 3.20 presents the eight genes, with most class 3 variants found in the MA probands relevant to causing DCM. Figure 3.20 showed that most MA probands carrying only class 3 variants had a *TTN* gene variant.

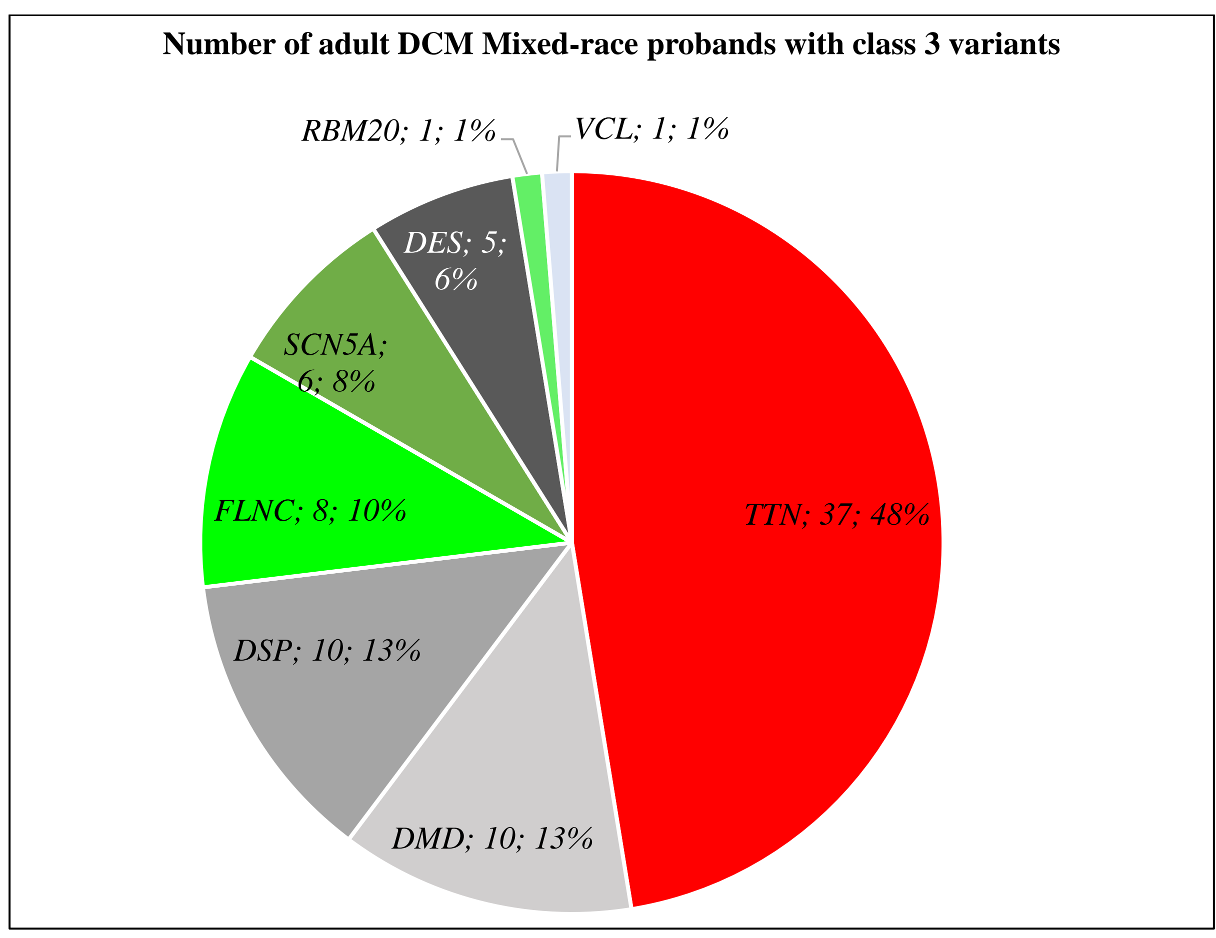


Figure 3.20: Adult DCM class 3 variant carriers of Mixed ancestry. *The number of MA adult-onset probands carrying only class 3 variants and the gene harbouring the class 3 variants.*

3.3.1.3.2.2.3 Class 3 variant carriers of White ancestry

We found 10 White adult-onset DCM probands carrying only class 3 variants: 70% (7/10) males and 30% (3/10) males. The 10 White adult-onset DCM probands had six genes harbouring 17 class 3 variants. The four genes were *TTN* found in seven probands, and the *FLNC*, *RBM20*, and *SCN5A* genes were each carried by one proband.

Of the 17 class 3 variants carried by the White probands with only class 3 variants, 58.8% (10/17) were reported, and 41.2% (7/17) were novel class 3 variants. The ten reported and seven novel class 3 variants were carried by 8 and 5 adult-onset White probands, respectively. Only 42.8% (3/7) of the White probands with only class 3 variants carried multiple variants, and 70% (7/10) carried only one class 3 variant each. The majority, 88.2% (15/17), of the class

3 variants in the White probands had a missense molecular consequence. The remaining two variants were an inframe deletion and a splice-acceptor variant.

3.3.1.3.3 Summary of classes 1 to 5 variants carriers

Figure 3.21 summarises the number of adult DCM probands in this study according to the ACMG five-tier variants classification system categorised by the diagnostic yield.

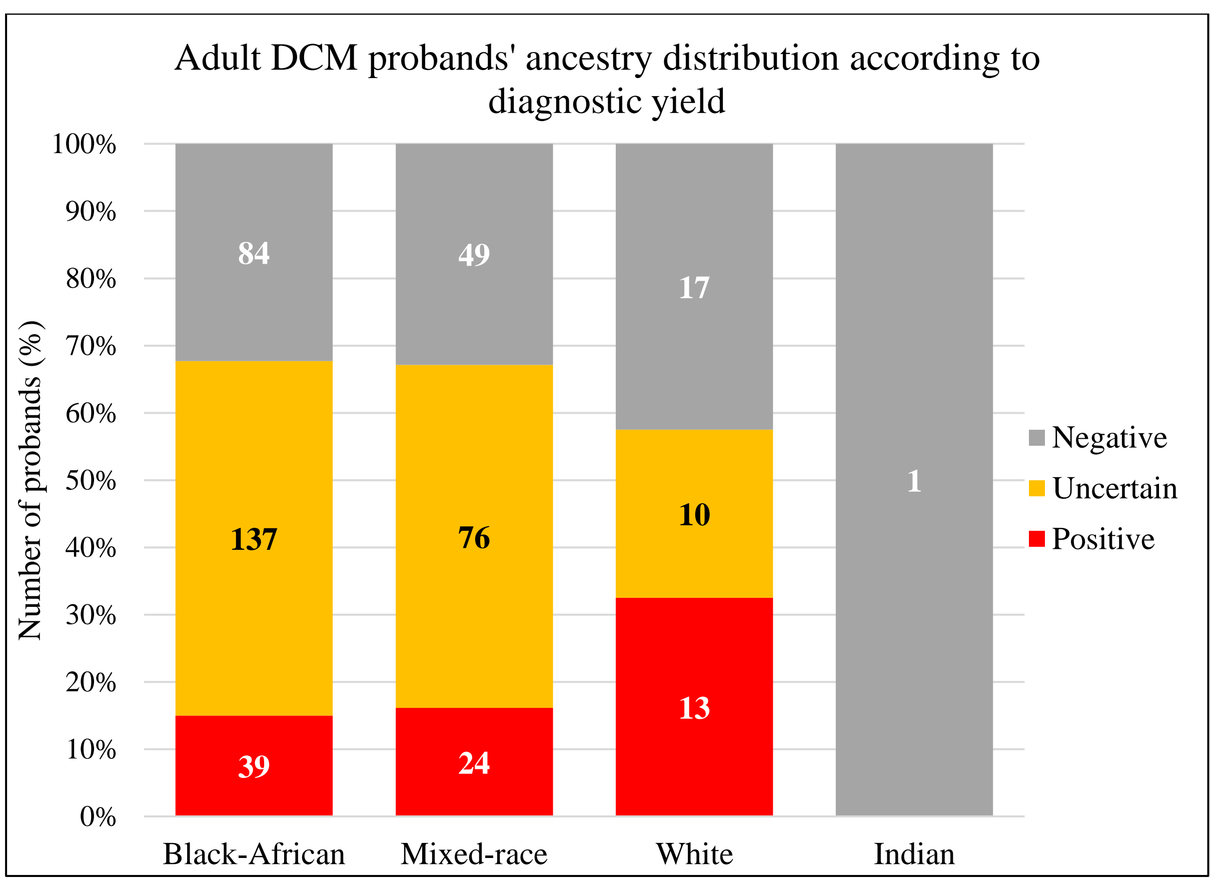


Figure 3.21: Summary of adult DCM classes 1 to 5 variants carriers. *The adult probands' diagnostic yield is categorised according to the cohort population ancestry.*

We performed targeted sequencing and analysis on our adult-onset cohort of 450 probands, shown in Figure 3.22. We found 76 probands carrying class 4 or class 5 variants: 51 probands carrying class 5 variants, and 25 probands carrying class 4 variants. Also, we reported that of the 76 probands, 13 had PPCM. Further, 223 probands carried only class 3 variants and 151 probands carried class 1 or 2 variants. We presented the genetic findings for the classes 3 to 5 variant carriers separately for the DCM and PPCM. The genetic findings of classes 1 and 2 variant carriers are beyond the scope of this thesis.

However, the baseline data for the 151 genotype-negative probands showed they were diagnosed at a mean (SD) age of 36.3 (12.4) years. Most, 65.6% (99/151), of the genotype-negative probands had non-familial DCM, while 25.8% (39/151) reported familial DCM and 8.6% (13/151) had no data for the disease aetiology. Also, 56.3% (85/151) of the probands had female sex assigned at birth. We also presented in Table 3.3 that the 151 consisted of 84 BA, 49 MA, 17 WA, and one IA proband.

3.3.2 Peripartum cardiomyopathy (PPCM)

3.3.2.1 Baseline characteristics

As so much controversy surrounds PPCM, we also wanted to address PPCM on its own as a separate disease entity from DCM. Of the adult DCM probands recruited, 30.2% (136/450) were diagnosed with PPCM phenotype, as shown in Table 3.3. Figure 3.22A shows that of the 136 PPCM probands, 93 were BA, 42 were MA, and one was WA female.

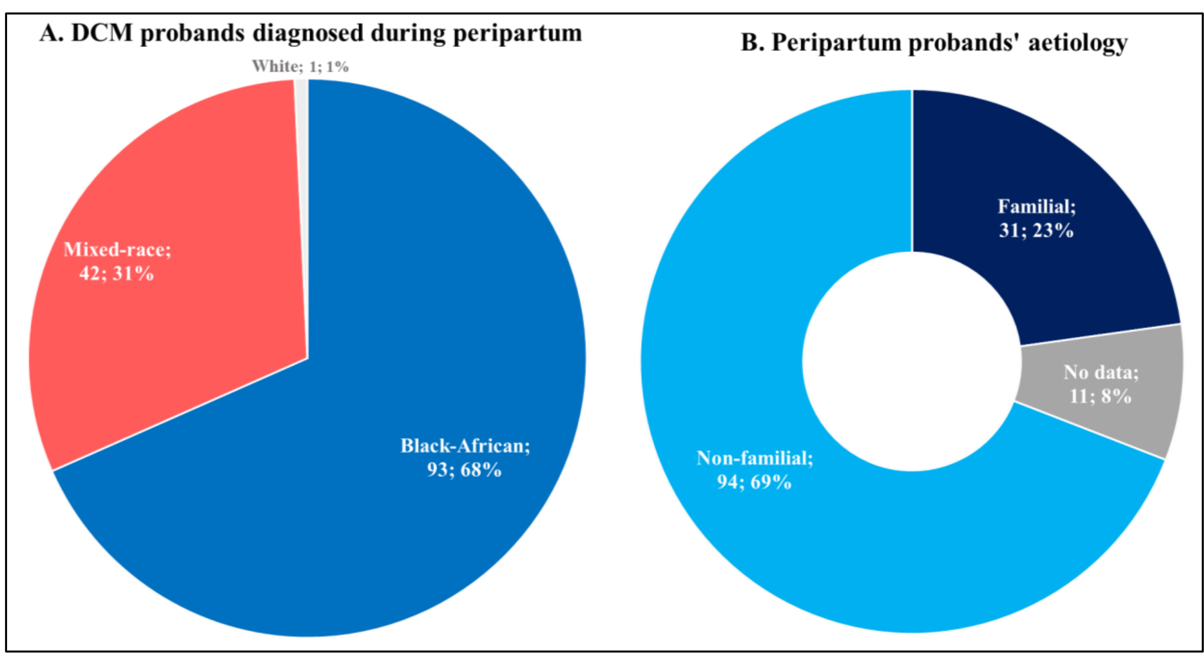


Figure 3.22: PPCM probands family history. *Pie chart showing the population distribution at diagnosis for the 136 PPCM probands.*

The 136 PPCM probands had a mean age of 29.6 years (SD 7.0, range 19-46). They were about seven years younger when compared to the overall adult DCM probands' age of diagnoses (mean = 35.6 years). When we looked at their age across the ancestry, we found that the BA and MA populations were diagnosed at almost the same age (mean 29.7, SD 7.6 and mean 29.4, SD 5.4 years, respectively), while only one 20-years-old WA proband was recorded in Figure 3.23 A. The data in Figure 3.23 B shows that about 70% of the PPCM were due to non-familial causes, whereas about 22% of the BA and MA probands had familial DCM.

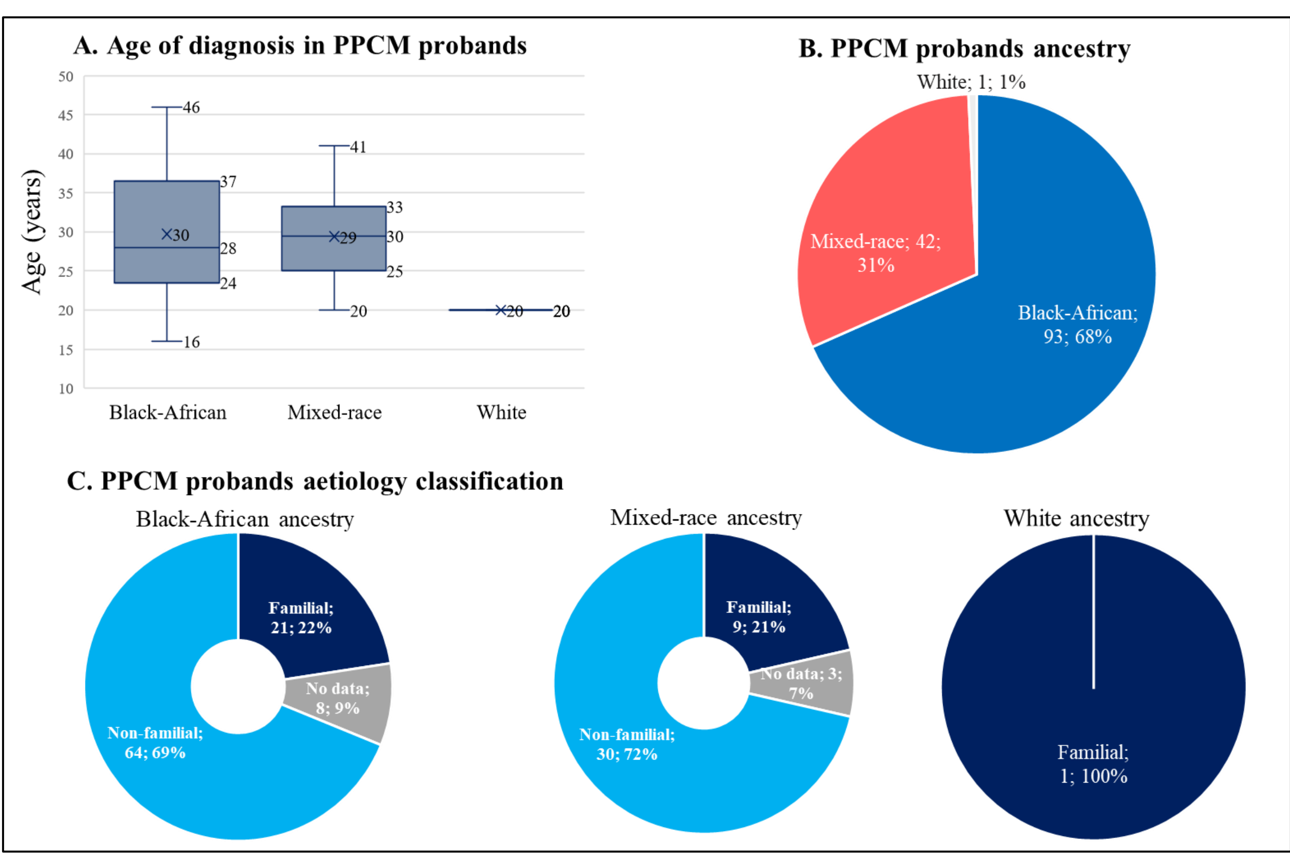


Figure 3.23: PPCM probands' age, sex assigned at birth and aetiology distribution. *Graphs showing comparisons between PPCM probands regarding age and diagnostic classification.* ***A.*** *Population groups vs. age of diagnosis;* ***B.*** *Population groups vs. diagnostic classification.*

3.3.2.2 Targeted sequencing of PPCM probands

Targeted sequencing found 13 (9.6%) genotype-positive PPCM probands: seven had class 5 variants, and four carried class 4 (Table 3.7). In contrast, 85 probands carried uncertain (class 3) variants, and 38 probands had genotype-negative (class 1 or 2) variants. We further analyse the class 3-5 variants and present the findings in the next section, starting with class 4 and 5 variant findings.

Table 3.7: Baseline results for the PPCM cohort according to ACMG classification

LABEL	LEVELS	TOTAL	POSITIVE	UNCERTAIN	NEGATIVE
DIAGNOSTIC YIELD, N (%)		136	13 (9.6)	85 (62.5)	38 (27.9)
AGE, YEARS	Mean (SD)	29.6 (7.0)	28.9 (8.0)	30.2 (7.0)	28.3 (6.6)
LVEF, %	Mean (SD)	30.1 (11.3)	23.9 (7.4)	30.9 (10.6)	30.6 (13.6)
NYHA	Class I	3 (2.4)	1 (8.3)	2 (2.6)	0 (0.0)
	Class II	29 (23.2)	4 (33.3)	14 (17.9)	11 (31.4)
	Class III	67 (53.6)	3 (25.0)	45 (57.7)	19 (54.3)
	Class IV	18 (14.4)	2 (16.7)	12 (15.4)	4 (11.4)
	ND	8 (6.4)	2 (16.7)	5 (6.4)	1 (2.9)
FAMILIAL	No	94 (69.1)	8 (61.5)	59 (69.4)	27 (71.1)
	Yes	31 (22.8)	4 (30.8)	19 (22.4)	8 (21.1)
	No data	11 (8.1)	1 (7.7)	7 (8.2)	3 (7.9)
ANCESTRY	Black-African	93 (68.4)	9 (69.2)	58 (68.2)	26 (68.4)
	Indian	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
	Mixed	42 (30.9)	4 (30.8)	26 (30.6)	12 (31.6)
	White	1 (0.7)	0 (0.0)	1 (1.2)	0 (0.0)

VALUES ARE MEAN ± 1SD OR N (%) BASED ON NON-MISSING VALUES OF THE TOTAL ANALYSED: 136. POSITIVE-CLASS 4 & 5, UNCERTAIN-CLASS 3, AND NEGATIVE-CLASS 1&2. LVEF-LEFT VENTRICULAR EJECTION FRACTION, NYHA-NEW YORK HEART ASSOCIATION.

3.3.2.2.1 Probands with pathogenic variants (class 4 and 5)

3.3.2.2.1.1 Baseline characteristics

Of the 9.6% (13/136) PPCM probands positive, nine were BA and four were MA probands. The 13 PPCM probands carrying class 4 and 5 variants were diagnosed at a mean age of 28.9 years (SD 8.0, range 19-46) in Table 3.8. They were about eight years younger when compared to the overall adult-onset DCM age of diagnoses for probands with positive variants (mean = 36.4 years). Of the 13 probands, seven carried class 5 variants, 71.4% (5/7) were BA, and 28.6% (2/7) were MA, while six PPCM probands with class 4 variants, 66.7% (4/6) were BA, and 33.3% (2/6) were MA. Table 3.8 shows that the class 5 variants in the PPCM probands were in two genes (*TTN* and *BAG3*) and the class 4 variants in two genes (*TTN* and *MYH7*). We then analyse the class 4 and 5 variants according to their ancestry to compare the findings with the DCM probands.

3.3.2.2.1.2 Diagnostic findings

3.3.2.2.1.2.1 Class 4 and 5 variants carriers of Black-African ancestry: PPCM cohort

We identified nine BA probands with class 4 and 5 variants. Five probands carried four class 5 variants: three were novel, and one reported variant, as shown in Table 3.8. The four class 5 variants were in *TTN* (3/4) and *BAG3* (1/4). We noted that 75% (3/4) of the class 5 variants were found in *TTN*, and the only reported variant, *TTN* c.95008C>T, recurred in 40% (2/5) of

the PPCM probands of BA. In comparison, the three novel variants we found had various molecular consequences: frameshift, stop-gained and splice-acceptor variants.

The other four probands carried four class 4 variants: two novel and two reported variants. The four class 4 variants were in *TTN* (3/4) and *MYH7* (1/4) genes. We noted that 75% (3/4) of the class 4 variants found in the *TTN* gene were two stop_gained and one splice_donor variant, while the *MYH7* novel variant had a missense molecular consequence.

3.3.2.2.1.2.2 Class 4 and 5 variants carriers of Mixed ancestry: PPCM cohort

We found four MA probands with class 4 and 5 variants. Two probands carried reported class 5 stop_gained variants in *TTN* and *BAG3* genes. The other two probands had two class 4 variants found in the *TTN* gene: one a reported stop_gained and one novel frameshift variant.

Table 3.8: Class 4 and 5 variants (n=12) in the PPCM probands (n=13) spanning two population groups

PROBAND_ID	ANCESTRY	AGE	GENE	PROTEIN CHANGE	cDNA CHANGE	CONSEQUENCES	EVIDENCE	dbSNP_ID	ZYGOSITY	ACMG
BLACK-AFRICAN ANCESTRY PROBANDS (N=5) WITH CLASS 5 VARIANTS										
10100396	Black-African	35	*TTN*	p.Arg31670Ter	c.95008C>T	stop_gained	Reported	rs1322596650	Het	Class 5
106304571	Black-African	20	*TTN*	p.Arg31670Ter	c.95008C>T	stop_gained	Reported	rs1322596650	Het	Class 5
10100007	Black-African	19	*BAG3*	p.Trp36Ter	c.107G>A	stop_gained	Novel	None	Het	Class 5
10100402	Black-African	25	*TTN*	Splice Site	c.9704-2A>G	splice_acceptor	Novel	None	Het	Class 5
106304556	Black-African	46	*TTN*	p.Asp24091GlufsTer8	c.72269_72272dup	frameshift	Novel	None	Het	Class 5
BLACK-AFRICAN ANCESTRY PROBANDS (N=4) WITH CLASS 4 VARIANTS										
106304528	Black-African	32	*TTN*	Splice Site	c.48638+2T>G	splice_donor	Reported	None	Het	Class 4
401303008	Black-African	25	*TTN*	p.Arg15898Ter	c.47692C>T	stop_gained	Reported	rs775186117	Het	Class 4
10100338	Black-African	41	*TTN*	p.Glu28397Ter	c.85189G>T	stop_gained	Novel	None	Het	Class 4
106304542	Black-African	25	*MYH7*	p.Pro152Arg	c.455C>G	missense	Novel	None	Het	Class 4
MIXED ANCESTRY PROBANDS (N=2) WITH CLASS 5 VARIANTS										
10100065	Mixed	27	*TTN*	p.Arg20626Ter	c.61876C>T	stop_gained	Reported	rs72646846	Het	Class 5
10100344	Mixed	32	*BAG3*	p.Arg309Ter	c.925C>T	stop_gained	Reported	rs869248137	Het	Class 5
MIXED ANCESTRY PROBANDS (N=2) WITH CLASS 4 VARIANTS										
10100713	Mixed	22	*TTN*	p.Arg28779Ter	c.86335C>T	stop_gained	Reported	rs1060500525	Het	Class 4
10100082	Mixed	27	*TTN*	p.Ile14101SerfsTer40	c.42300del	frameshift	Novel	None	Het	Class 4

3.3.2.2.2 Probands with variants of uncertain significance (class 3)

3.3.2.2.2.1 Baseline characteristics

Eighty-five PPCM probands carried only class 3 variants, and when we explored their age at diagnosis, we found no difference (p=0.33) between the three population groups (Figure 3.24). However, the BA probands' median age of 29 years (mean=30.2) was higher than that of 27 years for the probands with class 4 and 5 variants.

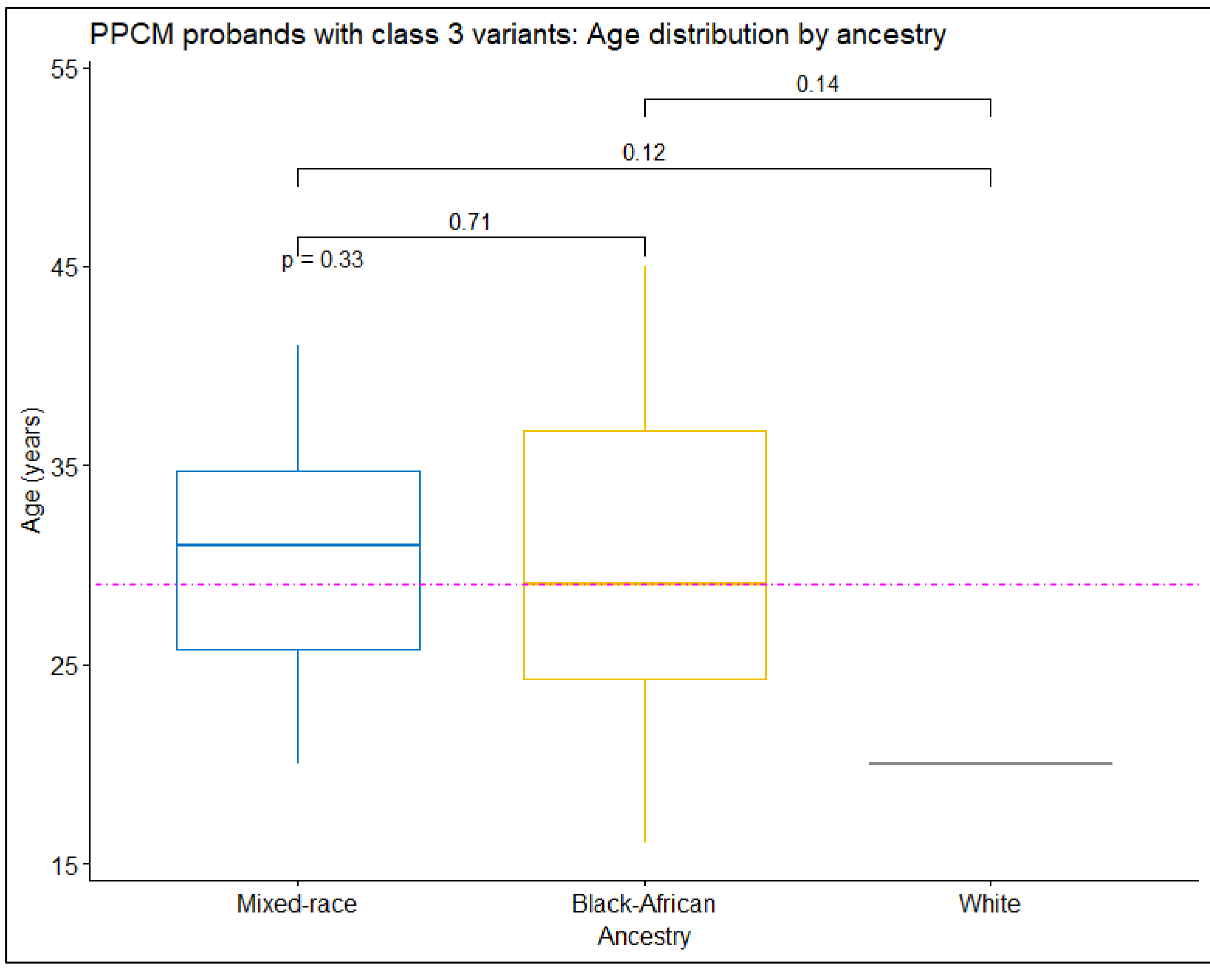


Figure 3.24: PPCM class 3 variant carriers according to ancestry. *The age-to-ancestry summary for 85 PPCM probands with class 3 variants. The magenta line indicates the median age of 27 years for the 13 positive probands.*

3.3.2.2.2.2 Diagnostic findings

3.3.2.2.2.2.1 Class 3 variants carriers' genetics findings: PPCM

SS Table 2 provides details for the class 3 genes and variants identified; however, we analysed the results regarding their onset age compared to the various self-reported ancestry and sex. We also summarised the findings in SS Table 2 for the PPCM probands with variants of uncertain significance according to their self-reported ancestry.

3.3.2.2.2.2.2 Class 3 variants carriers of Black-African ancestry: PPCM cohort

We found 58 BA PPCM probands carrying only class 3 variants; 75 class 3 variants harboured in 20 genes.

Of the 75 class 3 variants carried by the BA PPCM probands with only class 3 variants, 62.7% (47/75) were novel and 37.3% (28/75) were reported class 3 variants. The 47 novel and 28 reported class 3 variants were carried by 65 and 54 PPCM probands, respectively. We also found that only 50% (29/58) of the BA probands with only class 3 variants carried multiple variants, and the other half (29/58) carried only one class 3 variant each.

The majority, 89.3% (67/75), of the class 3 variants in the BA probands had a missense molecular consequence. The remaining eight variants were five frameshifts, two inframe_deletion and one missense&splice_region variant. Figure 3.25 presents the nine genes with most of the class 3 variants found in the BA probands that are relevant to cause PPCM. Figure 3.25 showed that most PPCM BA probands carried only *TTN* and *DSP* gene class 3 variants.

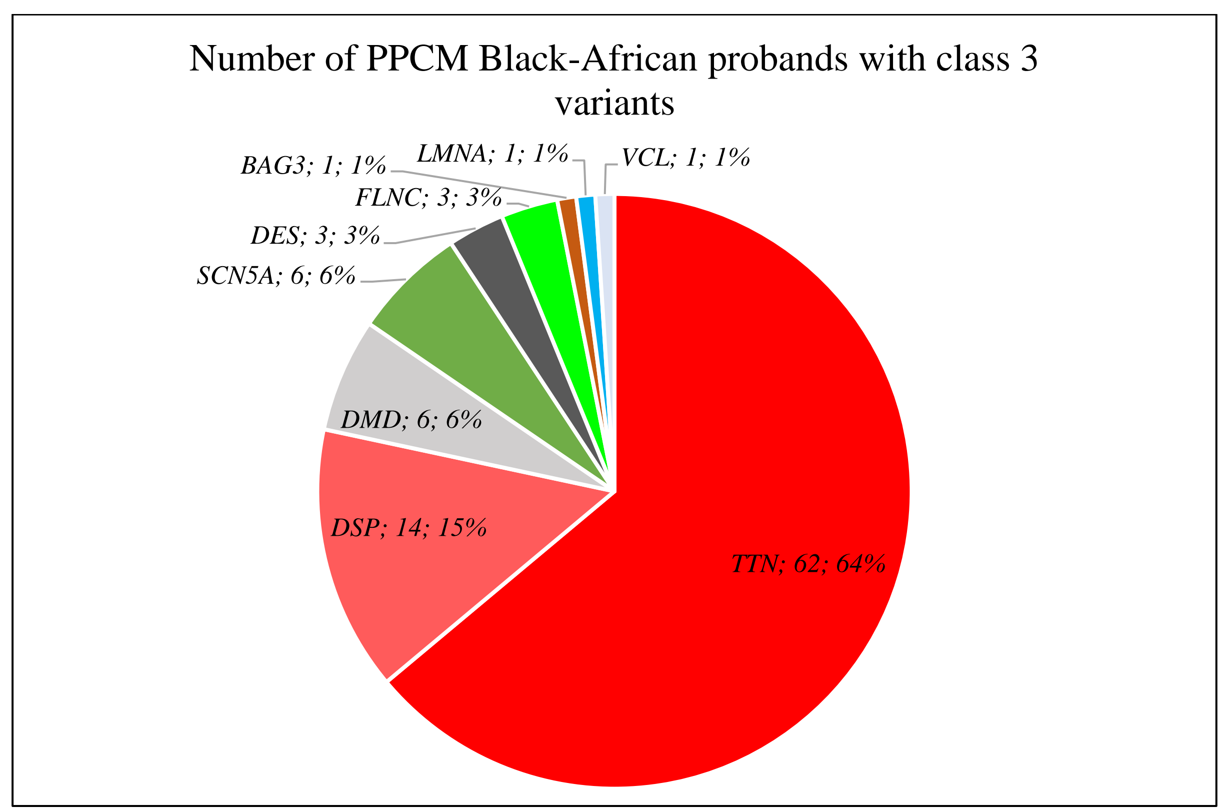


Figure 3.25: PPCM class 3 variant carriers of Black-African ancestry. *The number of BA PPCM probands carrying only class 3 variants and the gene harbouring the class 3 variants.*

3.3.2.2.2.2.3 Class 3 variant carriers of Mixed ancestry: PPCM cohort

We found 26 MA adult-onset DCM probands carrying only class 3 variants. The 26 MA adult-onset DCM probands carried 43 class 3 variants in 13 genes.

Of the 43 class 3 variants carried by the MA probands with only class 3 variants, 67.4% (29/43) were novel and 32.6% (14/43) were reported class 3 variants. The 29 novel and 14 reported class 3 variants were carried by 34 and 18 adult-onset MA PPCM probands, respectively. Only 42.3% (11/26) of the MA probands with only class 3 variants carried multiple variants, and 53.8% (14/26) carried only one class 3 variant each.

The majority, 95.3% (41/43), of the class 3 variants in the MA probands had a missense molecular consequence, while two were frameshift. Figure 3.26 presents the nine genes with most of the class 3 variants found in the MA probands that are relevant to cause PPCM. Figure 3.26 showed that most PPCM MA probands carried only *TTN* and *DMD* gene class 3 variants.

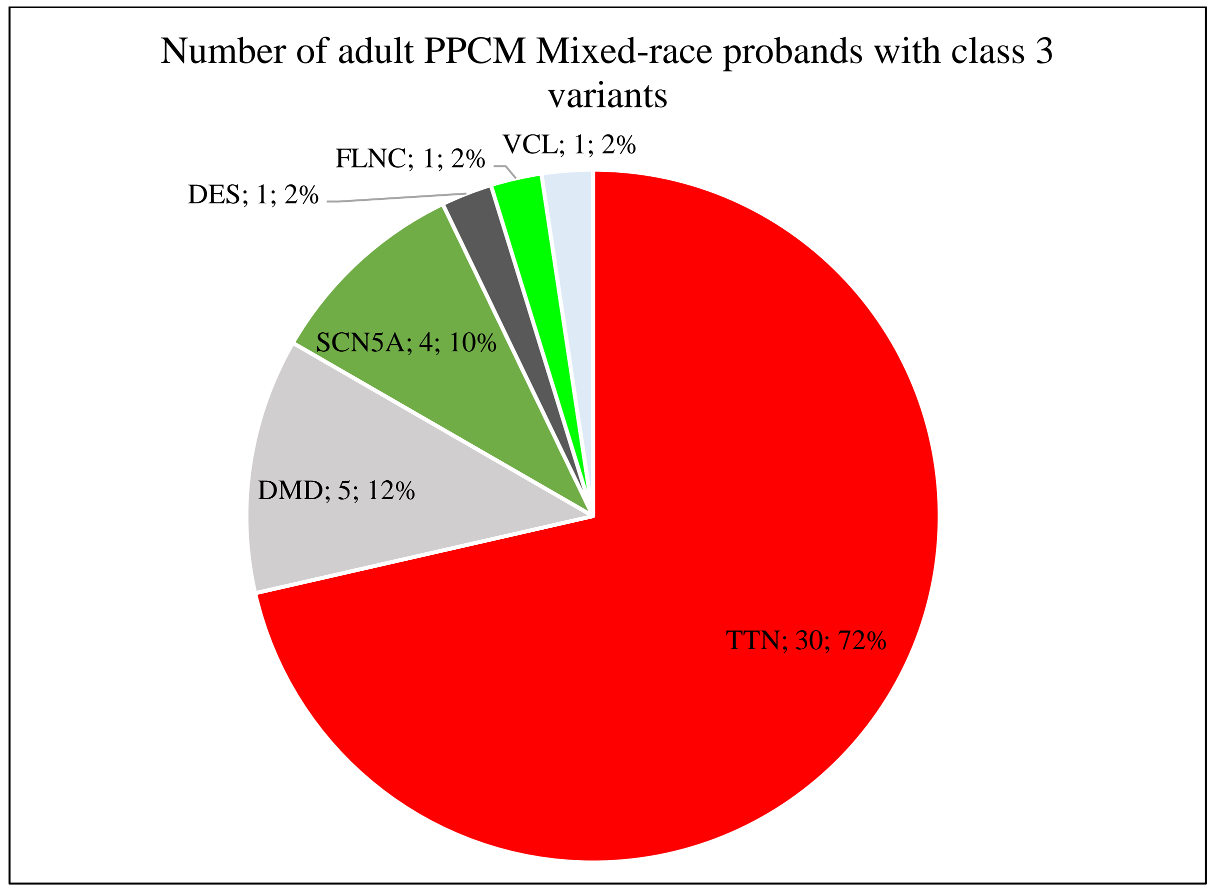


Figure 3.26: PPCM class 3 variant carriers of Mixed ancestry. *The number of MA adult-onset probands carrying only class 3 variants and the gene harbouring the class 3 variants.*

3.3.2.2.2.2.4 Class 3 variant carrier of White ancestry: PPCM cohort

We found one White PPCM proband carrying a novel *TTN* missense class 3 variant.

3.3.2.2.3 Summary of classes 1 to 5 variants carriers for the PPCM

We recruited 136 probands with PPCM and identified CMO-causing variants in 9.6% (13/136) of the probands. Of the 13 PPCM probands with class 4 and 5 variants, 69.2% (9/13) were BA, and 30.8% (4/13) were MA. There was neither a proband of White nor Indian ancestry. The PPCM probands carried class 4 and 5 variants in only three genes (*BAG3*, *MYH7* and *TTN*). The diagnostic yields we summarised in Figure 3.27 according to the probands' ancestry show that about 70% of the PPCM probands were of BA and 30% of MA.

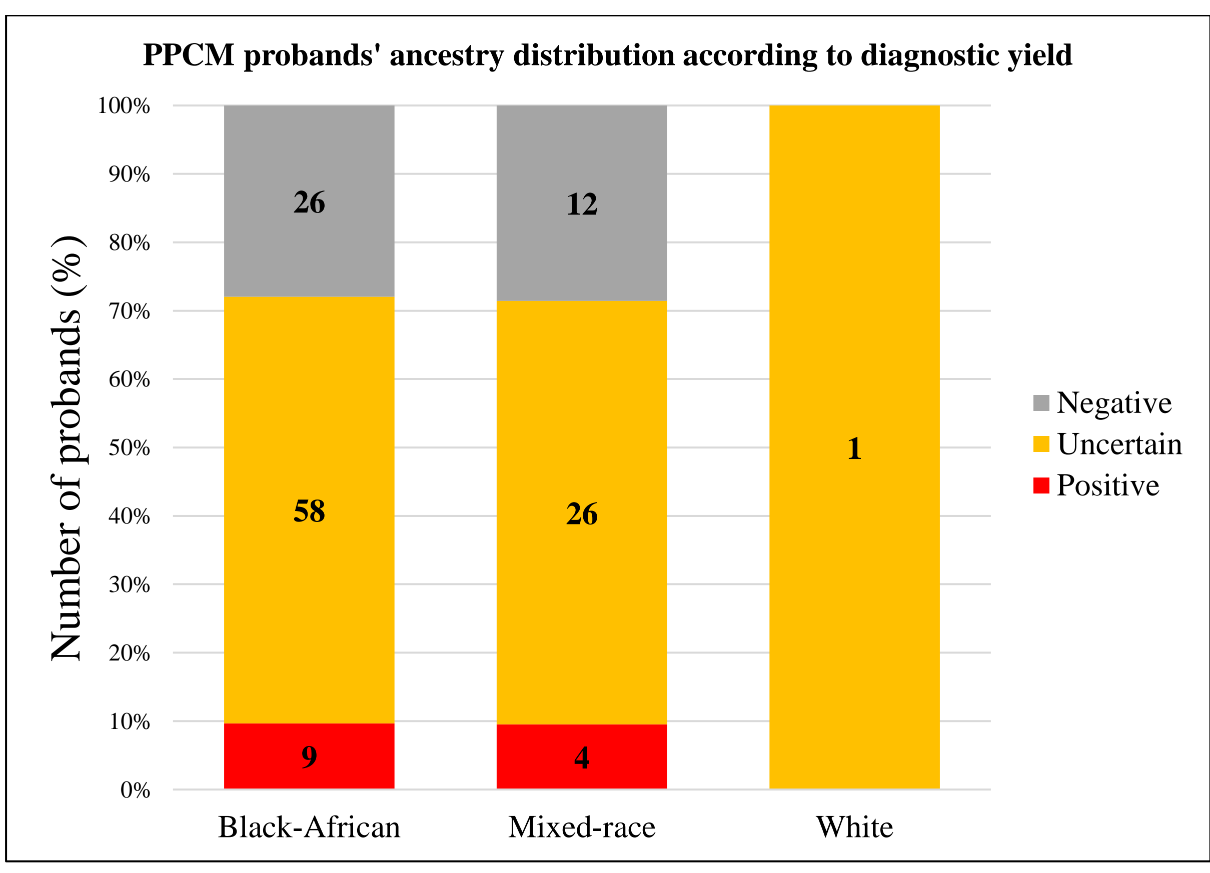


Figure 3.27: Summary of PPCM classes 1 to 5 variants carriers. The PPCM probands' ACMG/AMP variants classification is *categorised according to the cohort population ancestry based on the diagnostic yield.*

3.3.3 Paediatric DCM probands

3.3.3.1 Baseline characteristics for the paediatric DCM cohort

In this study, we enrolled 89 unrelated paediatric probands clinically diagnosed with DCM and classified them into two DCM subtypes (75.3% dilated and 24.7% non-compaction). The paediatrics clinical data analysis revealed that their average age at diagnosis was 3.1 years (SD 3.5, range <13), with 55% (49/89) of the probands recorded as female sex at birth. Their declared-reported ancestry was 69.7% Black-African, 28.1% Mixed, and 2.2% White ancestry probands, as shown in Table 3.10. No FH data was available for the paediatric probands at diagnosis.

We further analysed the baseline variation per population group in our paediatric DCM cohort, Figure 3.28. The 62 paediatric BA probands were diagnosed at a mean age of 3.4 years, as shown in Figure 3.28 A, with 54.8% (34/62) female (Figure 3.28 B). In comparison, the paediatric MA probands were, on average, the youngest group with a mean age of 2.2 years. Also, 52% (13/25) of MA probands recruited were females. We also recruited two female WA paediatric probands diagnosed at ages 6 and 12. No proband of the Indian ancestry was recruited.

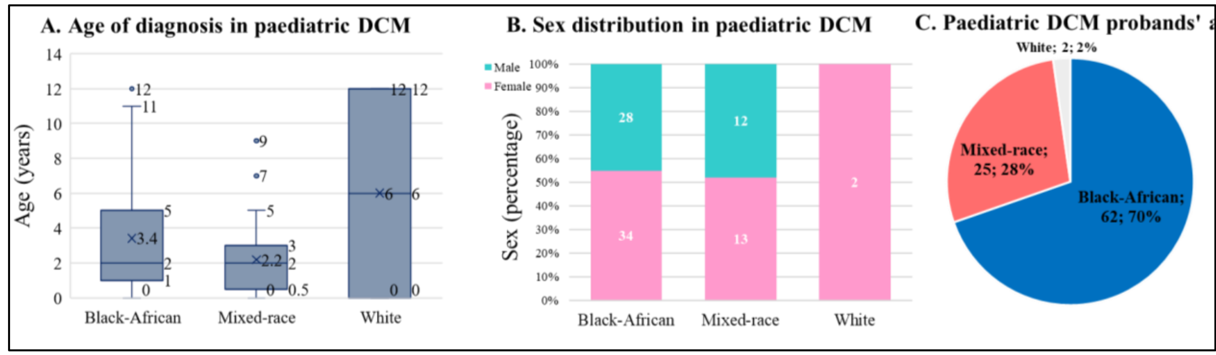


Figure 3.28: Paediatric DCM probands' age, sex assigned at birth and aetiology distribution. *Graphs showing comparisons between paediatric DCM population groups regarding age, sex assigned at birth and diagnostic classification. Population groups vs age of diagnosis in HCM; B. Population groups vs sex distribution; C. Population groups vs ancestry classification; D. Population groups vs aetiology/diagnostic classification.*

3.3.3.2 Targeted sequencing for the paediatric DCM cohort

Of the 89 paediatric DCM participants recruited, the targeted NGS strategy identified variants in all of them; however, we found only two probands with a genotype-positive result. The genotype-positive result yielded a diagnostic rate of 2.2% (2/89) for our paediatric probands.

The two probands were BA individuals: one female and one male, as shown in Table 3.9. In comparison, 57 paediatrics had only class 3 variants, while 30 carried class 1 or 2.

Table 3.9: Baseline results for the paediatric DCM probands cohort

LABEL	LEVELS	TOTAL	POSITIVE	UNCERTAIN	NEGATIVE
DIAGNOSTIC YIELD, N (%)		89	2 (2.2)	57 (64.0)	30 (33.7)
AGE AT DIAGNOSIS, YEARS	Mean (SD)	3.1 (3.5)	2.0 (1.4)	2.8 (3.3)	3.7 (3.9)
SEX RECORDED AT BIRTH	Female	49 (55.1)	1 (50.0)	32 (56.1)	16 (53.3)
	Male	40 (44.9)	1 (50.0)	25 (43.9)	14 (46.7)
ANCESTRY	Black-African	62 (69.7)	2 (100.0)	39 (68.4)	21 (70.0)
	Mixed	25 (28.1)	0 (0.0)	17 (29.8)	8 (26.7)
	White	2 (2.2)	0 (0.0)	1 (1.8)	1 (3.3)
DCM SUBTYPES	Dilated	67 (75.3)	1 (100.0)	44 (77.2)	22 (73.3)
	Non-Compaction	22 (24.7)	1 (100.0)	13 (22.8)	8 (26.7)

3.3.3.2.1 Diagnostic findings

3.3.3.2.1.1 Paediatric probands with pathogenic (class 4 and 5) variants

The paediatric probands positive for ACMG class 4 and 5 variants were two in total, 2.2% (2/89). Table 3.9 shows their baseline data analysis, indicating they had disease onset at a mean age of 2.0 years (SD 1.4, range 1-3). The two paediatric probands were of BA, one female and the other a male.

The two probands' genetic result is shown in Table 3.10, indicating that they each carried one variant. The female proband carried a novel *DMD* c.1149+1G>T splice_donor class 5 variants, while the male proband had an *MYBPC3* c.3409T>G missense class 4 variant.

Table 3.10: Class 4 and 5 variants found in the paediatric DCM cohort

PROBAND_ID	SEX	ANCESTRY	AGE	GENE	PROTEIN CHANGE	cDNA CHANGE	CONSEQUENCES	ZYGOSITY	ACMG
102302024	Male	Black-African	3	*MYBPC3*	p.Phe1137Val	c.3409T>G	missense	Het	Class 4
105303524	Female	Black-African	1	*DMD*	Splice Site	c.1149+1G>T	splice_donor	Het	Class 5

The variant in the *DMD* gene was a novel heterozygous c.1149+1G>T mutation with a CADD score of 31. The variant was absent in the gnomAD and ClinVar databases. Four Class 4 and 5 variants have been curated in the ClinVar at the same chromosomal position (amino acid residue) as the *DMD* c.1149+1G>T variant. When visualised in Alamut^{TM} Visual Plus, the affected amino acid was conserved across multiple species. Further, splice site predictor tools in Alamut scored very high above the threshold (SSF=79.02, MaxEnt=7.20, NNSPLICE=0.93 and GeneSplicer=2.42). The new nucleotide caused a loss of a 9-mer on the 5' end and a 6-mer on the 3' end of the splice donor site, thus leading to the loss of the nearest natural splice donor site.

The *MYBPC3* c.3409T>G (p.Phe1137Val) novel heterozygous was absent in the gnomAD and ClinVar databases. The *MYBPC3* c.3409T>G (p.Phe1137Val) exon 31 variant (Align GVGD: Class C45, CADD: 28.9, Mutation taster: Deleterious, PolyPhen2: Probably damaging, SIFT: Deleterious) was assigned a class 4 pathogenicity. The affected amino acid is in an evolutionarily conserved region across multiple species.

3.3.3.3 Paediatric probands with variants of unknown significance (class 3)

We identified 57 (39 BA, 17 MA and one WA) paediatric probands as carrying only class 3 variants, as shown in Table 3.10.

Clinical characteristics

Baseline data analysis of the 57 paediatric probands carrying only class 3 variants showed they were diagnosed at a mean age of 2.8 years (SD 3.3, range ≤11) in Table 3.9. There was thus no significant difference between the overall paediatric DCM age of diagnoses and probands positive for only class 3 variants.

When we separated the uncertain probands group into females (56.1%) and males (43.9%), we also noted an insignificant age difference ($p=0.25$). However, when we proceeded to separate the females from the males according to their ancestry, we found that the female probands of BA and males of MA were comparatively younger, as shown in Figure 3.29.

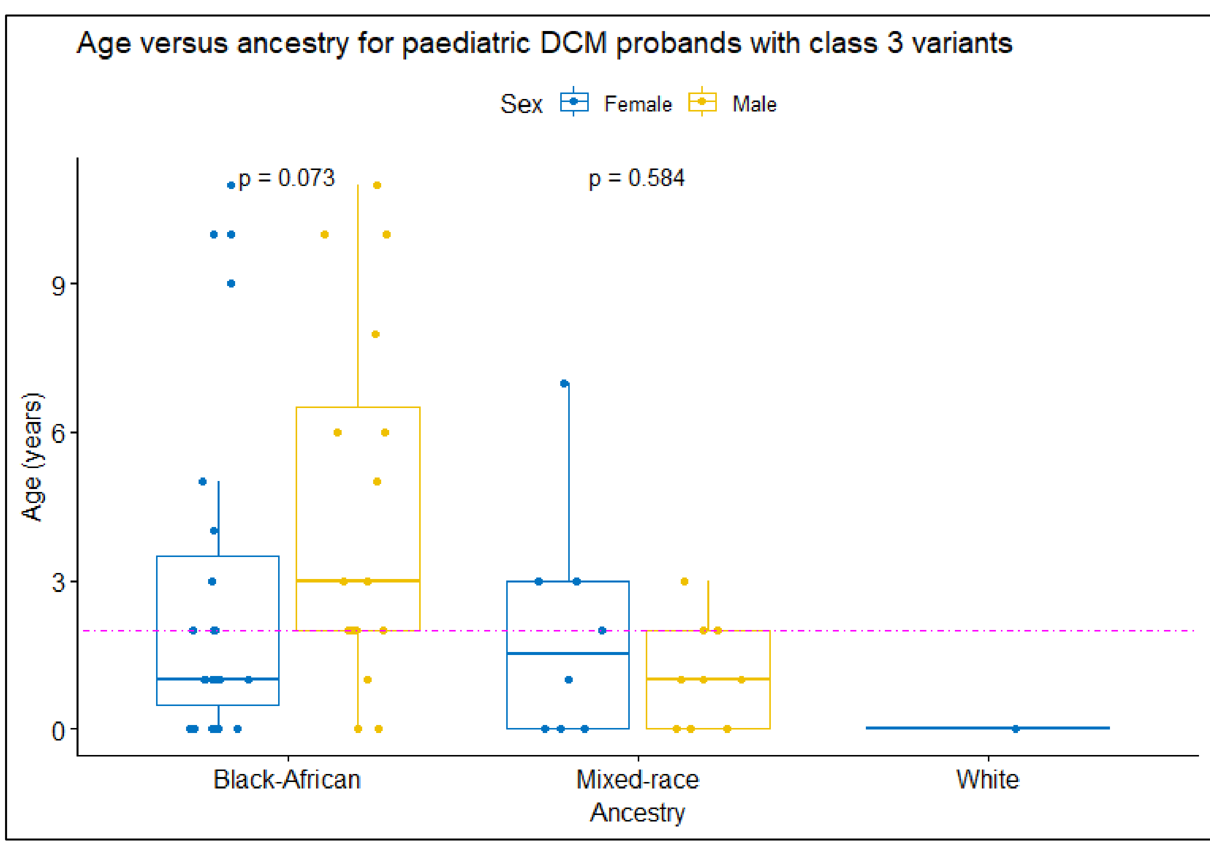


Figure 3.29: Paediatric DCM class 3 variant distribution according to sex. *The age distribution according to population ancestry for paediatric DCM probands carrying only class 3 variants grouped by sex assigned at birth. The magenta line indicates the 57 probands' median age of 2 years.*

Also, when we separated the probands into groups of novel and reported class 3 variant carriers, there was no significant difference (p-value=0.058). However, we noted that those with novel class 3 variants had a low median age within the BA and MA population groups.

3.3.3.3.1 Genetics findings

These 57 paediatric probands carried 24 genes harbouring 66 class 3 variants listed in SS Table 2. We curated the 66 variants according to the ACMG/AMP guidelines for the adult DCM. Of these, 9% (6/66) were truncating, and the rest were missense variants.

The ACMG guidelines predicted four class 3 variants (three reported and one novel) that barely met the class 4 variants classification of being as LP in 7.0% (4/57) of the paediatric probands. Of the four paediatric probands, 75% (3/4) were MA paediatric probands, and one was a BA.

One of the four LP class 3 variants was novel variant carried by the BA proband: *MYH7* c.2952del frameshift variant. Except for the one *MYH7* frameshift variant, all the other class 3

variants had a missense molecular consequence, and Table 3.11 shows the four class 3 heterozygous variants.

Table 3.11: Class 3 variants of interest in the paediatric DCM probands

PROBAND_ID	SEX	ANCESTRY	AGE	GENE	PROTEIN CHANGE	CDNA CHANGE	EVIDENCE	dbSNP_ID	ACMG
10300009	Male	Mixed	0	*MYH7*	p.Gly178Arg	c.532G>A	Reported	rs730880156	Class 3
102302020	Male	Mixed	1	*MYH7*	p.Cys538Arg	c.1612T>C	Reported	rs397516116	Class 3
102302030	Male	Mixed	1	*SCN5A*	p.Arg971Cys	c.2911C>T	Reported	rs61737825	Class 3
105303503	Male	Black-African	2	*MYH7*	p.Leu985TrpfsTer8	c.2952del	Novel	None	Class 3

We have further summarised the findings for the paediatric probands with class 3 variants only according to their self-reported ancestry for the class 3 variants listed in SS Table 2.

3.3.3.3.1.1 Class 3 variants carriers of Black-African ancestry: paediatric cohort

We found 39 BA paediatric DCM probands carrying only class 3 variants: 59% (23/39) females and 41% (16/39) males. The 39 BA paediatric DCM probands carried 59 class 3 variants harboured by 20 genes.

Of the 59 class 3 variants carried by the BA paediatric probands with only class 3 variants, 57.6% (34/59) were novel and 42.4% (25/59) were reported class 3 variants. The 34 novel and 25 reported class 3 variants were carried by 30 and 29 paediatric BA probands, respectively. We also found that 66.7% (26/39) of the paediatric BA probands with only class 3 variants carried multiple variants, and 33.3% (13/39) carried only one class 3 variant each.

The majority, 91.5% (54/59), of the class 3 variants in the paediatric BA probands had a missense molecular consequence. The remaining five variants were three frameshifts and two inframe_deletion.

The relevant class 3 variants in the paediatric BA probands were found in eight genes, as shown in Figure 3.29. Figure 3.29 shows that most BA probands carrying only class 3 variants had a *TTN* gene variant. The second common gene was the *DSP*; however, 72.7% (8/11) of the paediatric probands with the *DSP* gene carried a compound heterozygous reported c.4022G>A and novel c.2033A>G variants shown in SS Table 2.

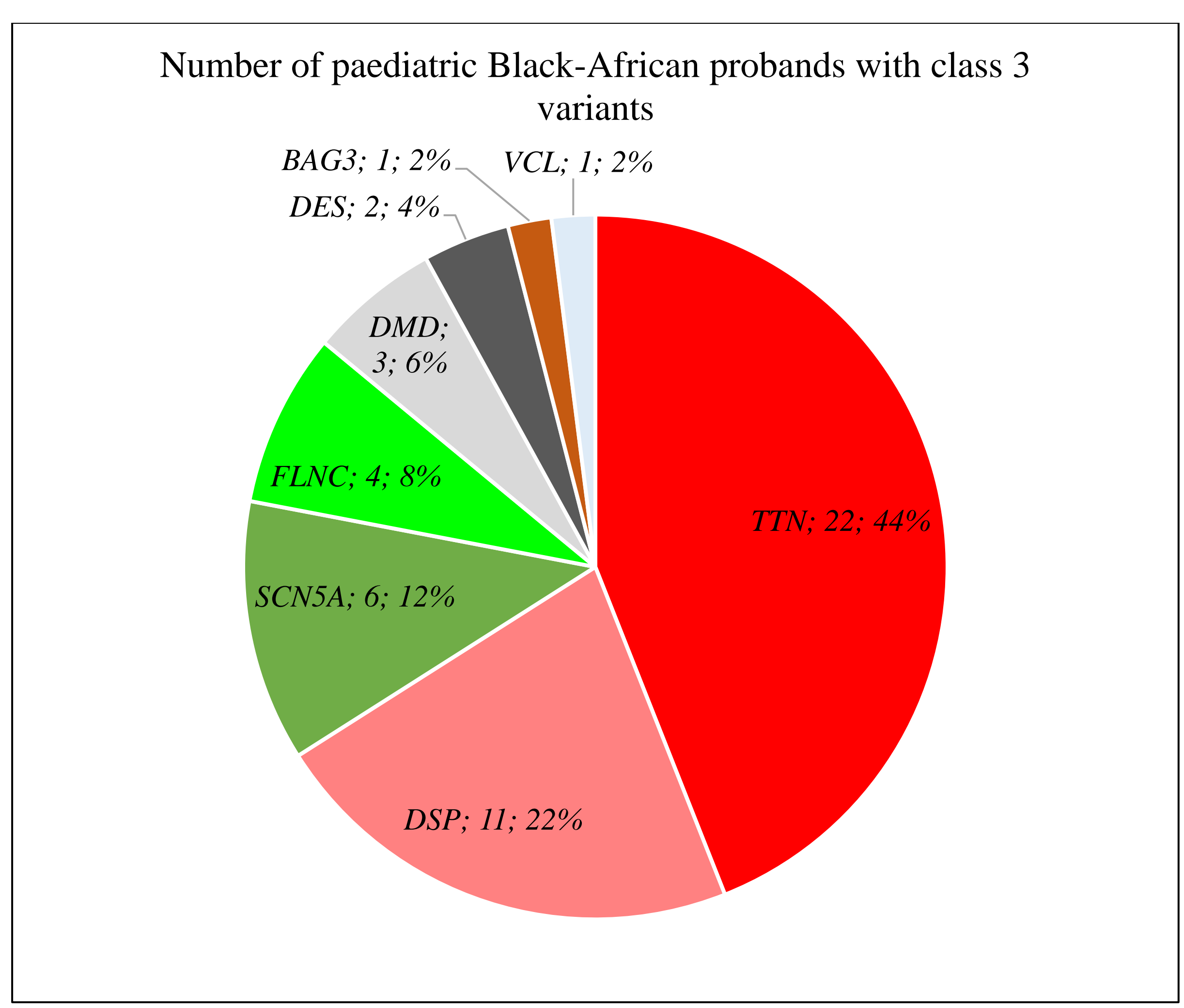


Figure 3.30: Paediatric DCM class 3 variant carriers of Black-African ancestry. *The number of BA paediatric probands identified with gene harbouring class 3 variants.*

3.3.3.3.1.2 Class 3 variants carriers of Mixed ancestry: paediatric cohort

We found 17 MA paediatric DCM probands carrying only class 3 variants: 52.9% (8/17) males and 47.1% (8/17) females. The 17 MA paediatric DCM probands carried 25 class 3 variants harboured by ten genes.

Of the 25 class 3 variants carried by the MA paediatric probands with only class 3 variants, 56% (14/25) were novel and 44% (11/25) were reported class 3 variants. The 14 novel and 11 reported class 3 variants were carried by 13 and 14 paediatric MA probands, respectively. We also found that 64.7% (11/17) of the paediatric MA probands with only class 3 variants carried multiple variants, and 35.3% (6/17) carried only one class 3 variant each.

The majority, 96% (24/25), of the class 3 variants in the paediatric MA probands had a missense molecular consequence, and one was a frameshift variant.

We found four relevant genes with class 3 variants in the paediatric MA probands, as shown in Figure 3.31. Figure 4.31 shows that most MA probands carrying only class 3 variants had a *TTN* gene variant. Like in the MA probands, the second common gene was the *DSP,* and all the MA paediatric probands with the *DSP* gene carried the compound heterozygous reported c.4022G>A and novel c.2033A>G variants shown in SS Table 2.

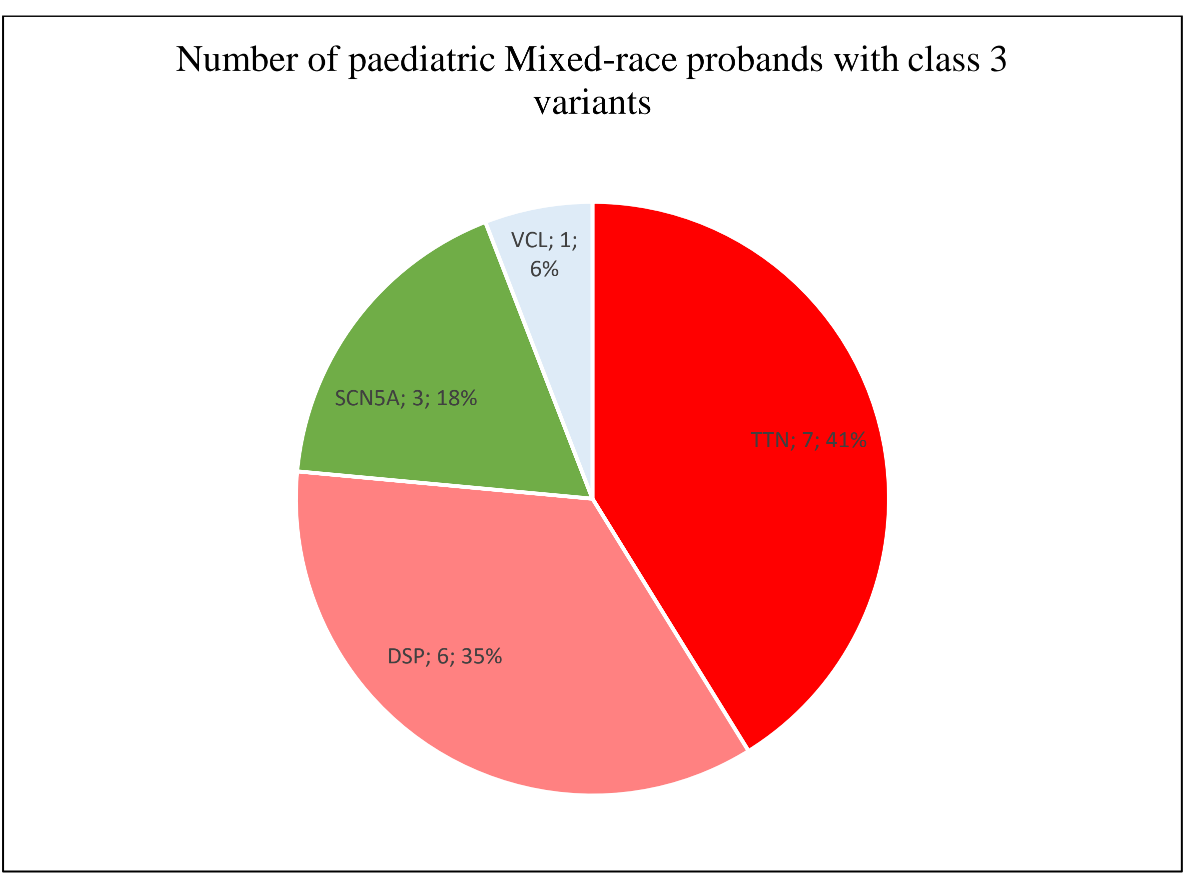


Figure 3.31: Paediatric DCM class 3 variant carriers of Mixed ancestry. *The MA paediatric probands were identified with gene harbouring class 3 variants.*

3.3.3.3.1.3 Class 3 variants carriers of White ancestry: paediatric cohort

We identified two paediatric probands of WA carrying only class 3 variants: females less than one year old. One carried a reported inframe *RBM20* c.141_143del, and the other carried a reported missense *TTN* c.44281C>T variant.

3.3.4 Possible founder class 4 and 5 variants in the DCM cohort

We found multiple probands that carried ten possible founder class 4 and 5 variants, as shown in Figure 3.32. The possible founder variants occurred in probands of specific ancestry. These variants included the *TTN* c.95008C>T found in 14 BA, *TTN* c.87624C>A found in five WA (and one MA) and *LMNA* c.568C>T we identified in 4 MA probands. The remaining seven variants were each carried by two probands.

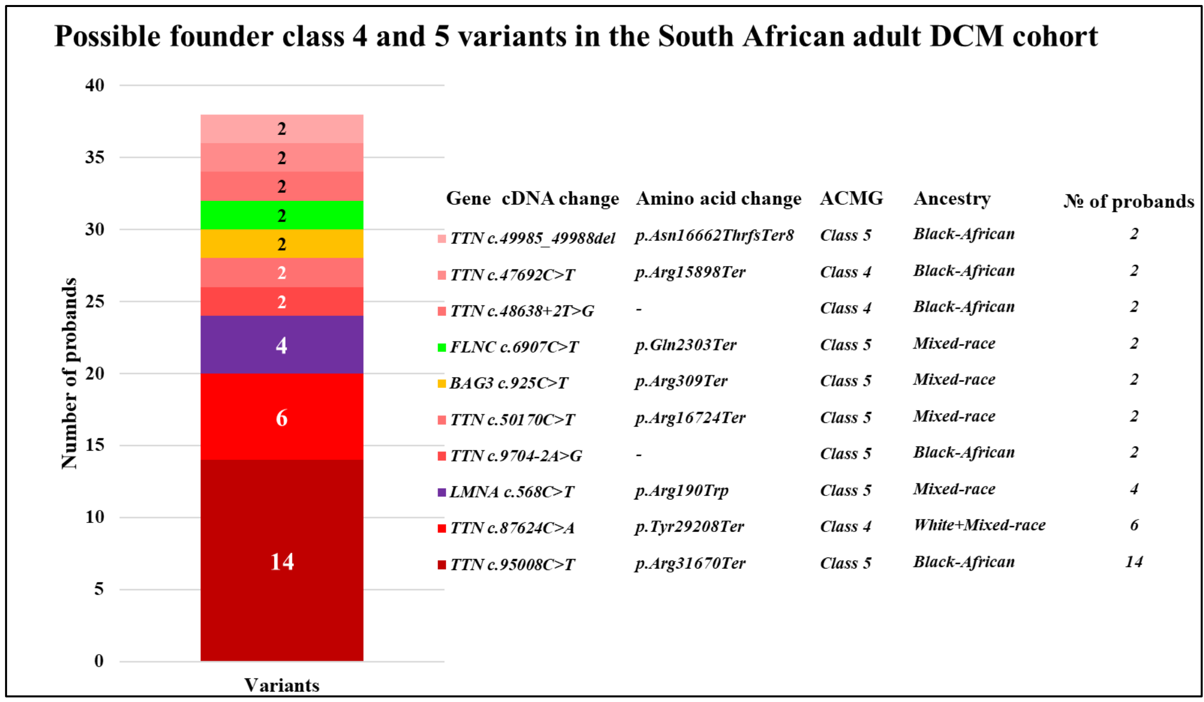


Figure 3.32: Adult DCM possible founder variants. *The recurring class 4 and 5 variants were identified in multiple adult-onset DCM probands.*

3.4 **DISCUSSION**

In Africa, the genetics of dilated cardiomyopathy is not well documented (53, 60). Through this first multicentre research called IMHOTEP, we provide unique insights into the genetic basis of DCM. In this thesis, targeted NGS data for 539 (females=55.8%) unrelated DCM probands living in South Africa were analysed. Globally, no study has provided such huge NGS data for African DCM patients living on the continent (37, 57). The probands recruited span various ancestry, including 322 (59.7%) Black-African, 174 (32.3%) Mixed, 42 (7.8%) White, and one (0.2%) Indian ancestry proband (s). These probands were from low-resource semi-urban to urban areas of South Africa, and we hypothesise gene-environment interactions could be contributing to their disease (110, 159). Indeed, other authors have shown that low economic status is associated with DCM disease outcomes such as HF (183). Notably, 92% of the probands were Black-African and Mixed DCM patients. It is important to note that the demographic characteristics of these DCM probands at the time of diagnosis closely resemble those of previous, more extensive South African clinical cardiovascular disease (CVD) studies (184) and the general population of South Africa. As per the Statistics South Africa census in 2022, the majority of citizens were Black (80.9%), approximately half of the population were female (51.1%), and the majority were young to middle-aged (between 15 to 59 years old).

Of the 539 probands, 450 (83.5%) were adults (and adolescents, ≥13 years), while 89 (16.5%) were paediatric cases (<13 years), as shown in Table 3.2. Our adult probands' FH findings at diagnosis were heterogeneous. Like previous DCM findings from South Africa (57), only 28.4% (128/450) of our adult probands had familial disease. We infer that the low identification of individuals with familial DCM, particularly among Black-African ancestry (BA) probands, could be due to small pedigree size as concise genealogical records are available for the BA subgroup (185). Also, in this DCM cohort, the probands were diagnosed as having one of the three subtypes of DCM: dilated, non-compaction, and peripartum. The findings for the adult-onset and paediatric DCM probands are discussed separately in sections 3.4.1 and 3.4.2, respectively.

3.4.1 Adult-onset DCM findings

3.4.1.1 Baseline findings for adult-onset DCM cohort

We recruited 450 adults (and adolescents ≥13 years) clinically diagnosed with DCM (dilate and peripartum forms) and determined their genetic characteristics. The DCM probands had a mean age of 35.6 years at diagnosis, which is like what has been reported in South Africa for patients presenting with DCM (184). However, in this study, the DCM probands of WA were about six years older than those of the other ancestry. Although the WA probands cohort was the oldest, they were younger than their European counterparts, with a mean ≈ median age of 49 in a study with a similar design (186). While this study flagged age differences, the sex distribution between the various ancestry showed a slight female predominance in BA and male predominance in the MA and WA. The age and sex difference led us to hypothesise that being a BA female or a WA male is associated with the onset age disparity. This research shows that 63% of the BA probands self-identified as females and 70% of the WA probands as males. In addition, about 54% (136/252) of all the females (mainly BA females) were diagnosed with DCM during peripartum. The female finding is like most reports in the literature, in which the authors emphasised that African ancestry is a risk factor for peripartum DCM with a prevalence of 1 in 1000 in South Africa (160).

The adult probands' self-reported ancestry at diagnosis revealed that 57.8% were BA, 33.1% MA, 8.9% WA, and <1% were IA probands. The high percentage of Black South Africans diagnosed with DCM is expected, as the demographic of the Western Cape province of South Africa as of 2021, where a majority of the DCM probands were recruited, was 51% BA, 25% MA, 17% WA, 6% IA and 1% other. Much like the rest of South Africa, the Western Cape has a tumultuous migration history, and the population movements have led to a distinctive combination of genetics, languages, cultures, poverty, and traditions (27, 70). The population movement has particularly affected the citizens of BA, who continue to experience poverty, limited access to advanced medical care, and related social issues (25). Thus, this study's findings could advise policymakers as our results have flagged young onset age and ancestry as contributing factors to the DCM phenotype in South Africa.

We have discussed the genetics findings according to the ancestry per ACMG classification to understand further the disparities shown by the baseline clinical data.

3.4.1.2 Targeted sequencing findings for adult-onset DCM cohort

We used a comprehensive routine targeted NGS panel testing and identified DCM-causing variants in 76 probands (39 BA, 24 MA and 13 WA). While 223 (137 BA, 76 MA and ten WA) DCM probands carried only class 3 variants, and 151 (84 BA, 49 MA, 17 WA and one IA) DCM probands had only class 1 and 2 variants. The genotyped findings yielded a positive rate of 16.9% (76/450), with 83.1% (374/450) having an uncertain or unknown genetic cause. We first described the probands who tested positive for class 4 and 5 variants and then discussed class 3 variants and their potential modifying effects on DCM. The analysis and discussion of class 1 and 2 variants are beyond the scope of this thesis.

3.4.1.2.1 Pathogenic (class 4 and 5) variants findings

3.4.1.2.1.1 Class 4 and 5 variant carriers' clinical findings

We found that the probands positive for targeted sequencing (ACMG class 4 and 5 variant carriers) had a mean age of 36.5 years. There was thus no significant onset age difference between the overall DCM age of diagnosis and genotype-positive probands. Very few studies have reported young adults' onset of DCM at a mean age of about 38 (187, 188). Notably, 36.5 years is a very young age for DCM onset compared to European data, with a mean ≈ median age of 49 in a study with a similar design (186). Also, studies from the USA (187), the UK (142), Norway (95), and the Netherlands (189) have reported much older mean onset ages of about 50 years at clinical diagnosis. Given that our probands were recruited from low-resource semi-urban to urban areas, we hypothesised that gene-environment interactions could be contributing to their disease (110, 159). Indeed, low economic status is associated with early onset of DCM disease outcomes such as HF (183).

Interestingly, when we compared our genotype-positive probands for ACMG class 4 and 5 per ancestry, we noted a significant difference (p=0.00076) within the MA probands. Notably, the MA had the most heterogenous genotype-positive result. There was an insignificant difference within the BA (p=0.47413) and WA (p=0.88625) probands. We further investigated and found that the MA probands carrying class 4 variants were significantly younger (p=0.021), about ten years younger than their counterparts with class 5 variants. We hypothesise that the MA probands' diverse genetic aetiology played a role. For example, they carried mainly sarcomere and nuclear envelop protein variants, the association of which has been demonstrated with the early onset of DCM with fatalities (137). Moreover, most of the MA probands in this study

were recruited from around the mother city of Cape Town, known for harbouring people of various ancestry over the past 300 years. Similar to the MA class 4 variant carriers, the BA probands with class 5 variants mainly carried sarcomere protein variants and were younger than their class 4 variant carriers' counterparts. Although the early onset of DCM has been associated with the BA (141, 190), the association of class 4 or 5 genetic variants has remained understudied in this population group.

Furthermore, we found no age differences between our WA probands and the international cohorts. Even though some White South Africans may share ancestral links with some European cohorts (169), our WA probands had many variants never reported in Europe. The genetic makeup of White South Africans of European descent may differ from that of White people in Europe due to gene-environment interactions. These interactions might have affected the disease progression in our probands.

Besides differences in ancestry and age at diagnosis, we observed distinct trends based on sex. Unlike at diagnosis, where more females (56%) were enrolled, our genotype-positive probands mainly were males (57.8%). Similarly, when we proceeded to separate the class 4 and 5 variant carriers according to their self-reported ancestry, we found that the positive MA and WA probands were mainly males (66.7% [16/24]) and (76.9% [10/13]), respectively. However, the BA probands mainly were females (53.8% [21/39]). Notably, our young males (46.1% [35/76]) and peripartum females (17.1% [13/76]) were the most affected. Globally, male sex and pregnancy are known modifiers of DCM disease, especially amongst those younger than 45 years (191, 192). We hypothesise our young males are more susceptible to developing DCM due to changes in sex hormones and gene-environment interactions (137) because Figure 4.7 revealed that they were more likely to carry a class 5 variant, whereas females (the majority diagnosed during pregnancy) were mainly class 4 variant carriers. Sex hormones such as testosterone in males (137) and oestrogens in pregnant females (163) are known risk factors for adult-onset DCM. In healthy individuals, testosterone and oestrogens bind to their receptors on myocytes and cardiac vascular endothelial cells to alter the functions of platelets and immune cells. Immune cells and platelet dysfunction influence cardiac inflammation, myocyte remodelling, and thrombosis involved in DCM. Most inflammatory cells that drive ventricle enlargement predisposing DCM patients to thromboembolic events are typical in males (137). Males also have lower levels of oestrogen receptors in the arteries than females, and activating oestrogen receptors by 17β-oestradiol prevents cardiomyocyte apoptosis and reduces cardiac

hypertrophy and fibrosis in women. Further, decreased levels of 17β-oestradiol predispose primarily female patients with DCM to HF (193).

Our adult-onset DCM baseline findings are unique and different from those of similar studies such as the UK Biobank DCM study (143), the ASPREE trial (82), the Dutch CMO study (191), and to name a few, in that a) we studied multiracial participants from the SSA region, b) we included all the DCM subtypes (dilated, left ventricular non-compaction and peripartum), c) we recruited participants from low-resource semi-urban to urban areas, and d) we diagnosed mainly young economically active probands. The adult DCM molecular genetic findings are discussed further under the headings of *class 4 and 5 variant carrier*s.

3.4.1.2.1.2 Class 4 and 5 variant carriers' genetic findings

The comprehensive routine targeted NGS panel testing identified 76 DCM probands, resulting in a cohort diagnostic rate of 16.9%, which is comparable to what has been reported by similar studies within the last three years (95, 114). However, the 16.9% is low compared to most published genetic testing yields for adult-onset DCM because the yield has historically ranged from 20% to >55% (142, 146). We also found that the yield for our probands of WA (32.5%) was within the international published ranges. Our MA and BA probands had a much lower yield than their White counterparts, with 16.1% and 15.0% yields, respectively. We hypothesise that the difference in the yields is due to clinical aetiologies. Our findings showed that most WA probands we recruited had familial DCM comparable to the literature for data mainly generated from Europe. We found this interesting because the gene panel we used was designed specifically to diagnose inherited CMO routinely in the UK, and our WA probands are generally of European descent. Nevertheless, our DCM cohort's probands showed a distinct genetic architecture compared to European published data (51).

The 76 genotyped positive probands carried 14 high-evidence DCM-causing genes (120) that harboured 51 (26 novel and 25 reported) class 4 and 5 variants. The 14 genes found encode for the sarcomere (*TTN*, *MYH7*, *MYBPC3* and *TNNT2*), nuclear envelope (*LMNA*, *RBM20* and *PRKAG2*), cell-junction (*SCN5A*, *JUP* and *PKP2*), Z-disc (*BAG3*) and other cytoskeletal structure (*FLNC*, *DMD* and *FHL1*) proteins. Our finding aligns with most studies showing that genes encoding sarcomere, Z-disc, or cytoskeleton proteins are the leading cause of familial DCM (58). Similarly, 33% (25/76) of our class 4 and 5 variant carriers had familial DCM. To

better understand the pathomechanisms of these DCM-causing genes and variants, we have described the mutations based on the cellular location of the genes.

3.4.1.2.1.2.1 The sarcomere (TTN, MYH7, MYBPC3 and TNNT2) genes' variants

Class 4 and 5 variants altering the sarcomere protein structures (common in our BA class 5 and MA class 4 variant carriers) are associated with cardiac conduction, ventricular arrhythmia, and high heart transplantation rate in advanced and heritable DCM (57, 90). This study's four sarcomere genes (*TTN*, *MYH7*, *MYBPC3* and *TNNT2*) accounted for 68.6% (35/51) of the class 4 and 5 variants. We, therefore, attributed our study's primary genetic aetiology of adult-onset DCM to sarcomere genetic alterations. However, the *TTN* is the predominant sarcomere gene in this cohort.

3.4.1.2.1.2.1.1 The titin (*TTN*) gene class 4 and 5 variants (n=29)

No proband carried two or more *TTN* gene class 4 or class 5 variants; however, we curated 29 *TTN* (12 class 4 and 17 class 5, and with a combined total of 16 novel and 13 reported) truncating variants in 52 probands (31 BA, 13 MA and eight WA probands).

We discuss the *TTN*tv that occur in >1 proband or variants reported to be highly penetrant/severe; however, we have briefly described all where necessary.

3.4.1.2.1.2.1.1.1 TTNtv c.95008C>T

The *TTN* c.95008C>T (p.Arg31670Ter) class 5 variant has only been reported once in the literature in the ASPREE trial study (82) and with four submissions in the ClinVar database. The stop_gained *TTN*tv in the A-band region causes severe early-onset DCM (110), which we have seen in our proband with this variant. Although our BA probands were diagnosed at age <65 with or without HF history, the ASPREE study in which the participants were all asymptomatic with age >70 years reported the *TTN* c.95008C>T variant in one proband (82). The demographic differences (e.g., age and ancestry) between our cohort and the European cohort suggest that our participants might have other DCM disease modifiers contributing to their disease (110, 159).

3.4.1.2.1.2.1.1.2 TTNtv c.47692C>T

The *TTN*tv c.47692C>T (p.Arg15898Ter) class 4 has been reported in non-African DCM patients (194). The *TTN*tv c.47692C>T variant carriers in the literature have had mild phenotype; however, the probands in this study had a severe early onset DCM (145). We hypothesised that other DCM modifiers, for example, cardiac toxic risk factors such as alcohol and other substance abuse, had triggered the disease (25). These risk factors are known in our general population and have been shown to worsen DCM outcomes, especially in *TTN* variant carriers (57).

3.4.1.2.1.2.1.1.3 TTNtv c.50170C>T

The *TTN*tv c.50170C>T (p.Arg16724Ter) stop_gained class 5 variant is a well-known cause of DCM, as it has been reported in a cohort study in a 17-year-old proband (83). This variant has also been reported as pathogenic three times in the ClinVar database. The variant causes severe early-onset DCM in a 17-year-old proband (83), which we have seen in two MA probands with this variant.

3.4.1.2.1.2.1.1.4 TTNtv c.87624C>A

The *TTN*tv c.87624C>A (p.Tyr29208Ter) stop-gained class 4 variant has previously been reported (169, 195) in two unrelated patients. This variant has also been reported as LP two times in the ClinVar database. The variant causes mild early-onset DCM in the patients (169, 195), which we have also seen in our six probands. However, the six probands were recruited from the same geographic region with a common culture. Half of these probands had familial DCM at diagnosis, and most were recruited from the same study site. We hypothesise that these six probands may also have a titinopathy due to a possible founder variant effect (196).

3.4.1.2.1.2.1.1.5 TTNtv c.9704-2A>G

The splice_acceptor *TTN*tv c.9704-2A>G class 5 variant is a novel mutation. This study identified the *TTN* c.9704-2A>G truncating variant in two unrelated probands diagnosed with early-onset DCM, and one was diagnosed during peripartum.

3.4.1.2.1.2.1.1.6 TTNtv c.49985_49988del

The frameshift *TTN*tv c.49985_49988del (p.Asn16662ThrfsTer8) class 5 variant is a novel mutation. The was found in two unrelated probands diagnosed with early-onset mild DCM.

Most of the other *TTN* variants are novel class 4 and 5 variants. More than half of the probands with this variant were recruited from the same study site, and maybe their DCM has a genetic aetiology due to a common ancestor. Hence, we hypothesise that the probands with non-familial *TTN*-related DCM possibly have complex pathomechanisms involved in their disease (196). The probands with non-familial DCM may probably have had the disease triggered by environmental or cardiac toxic risk factors such as alcohol and other substance abuse. These risk factors are known in our general population and have been shown to worsen DCM outcomes, especially in *TTN* variant carriers (57).

3.4.1.2.1.2.1.1.7 The natural titin (TTN) gene

The *TTN* gene encodes for a giant elastic protein called titin, which connects the Z-disc to the M-line and other sarcomere proteins. The prominent role of titin is to generate force for contraction in striated muscle tissues such as the heart muscle. The titin protein is expressed constitutively in the heart muscle (197), and class 4 or 5 variants in the *TTN* gene sequence may affect the protein's expression level or change its regular structural conformation, impairing its sarcomere function. Poor-quality (e.g., shortened or misfolded) titin protein product will cause sarcomere protein haploinsufficiency or disrupt various signalling pathways or myocyte disarray and collagen deposition (148, 168, 198). These disruptions impair sarcomere force generation, increase susceptibility to arrhythmias and abnormal remodelling, or disrupt cardiomyocyte alignment, promoting abnormal electrical conduction (167). If these disruptions are left untreated, they can lead to arrhythmias, systolic dysfunction, and death.

The standard type of *TTN* class 4 or 5 variants associated with DCM are truncating mutations, and very few are missense (199). Missense variants are single point mutations in the nucleotide chain of coding DNA sequence that result in an amino acid change in the protein sequence, and truncating variants are mutations in the nucleotide chain of coding DNA that introduce a premature termination codon in the protein-coding sequence. The estimated prevalence of DCM-related *TTN* class 4 and 5 variants ranges from about 15% to 25%, based on several studies from Europe and North America (95). Although most *TTN* variants we identified were missense, they are classes 1 to 3, whereas most DCM-causing ones are truncating variants.

Titin truncating variants (*TTN*tv) have consistently been the predominant cause of DCM, with an estimated prevalence of about 12% to 45% (95).

We found that 68.4% (52/76) of our genotype-positive probands had DCM-related TTNtv variants, which is a higher carrier rate than most studies have reported for DCM-related TTNtv patients (95, 142). Our high percentage of DCM-related *TTN*tv could be due to established risk factors such as male sex, young onset age, pregnancy, African ancestry, and possible modifiers such as gene-environment interactions (141, 190). Testosterone in males (137) and oestrogens in pregnant females (163) are known risk factors that increase the chance of an early onset of DCM. Testosterone and oestrogen affect immune and platelet cells by binding to receptors on heart and vascular cells. Males are more prone to DCM due to specific inflammatory cells when dysfunction in these cells leads to cardiac inflammation, muscle damage, and blood clotting (137). Notably, in our study, primarily young South African males and females (diagnosed during pregnancy) were DCM-related *TTN*tv class 4 and 5 variant carriers. Also, the prevalence of DCM during peripartum is higher amongst women of Black African descent globally (24, 160), and the reasons are not fully established.

*TTN*tv class 4 and 5 variants in the significant post-slice region of titin with very high exons PSI index (169) are known causes of most classified heritable DCM (195). A recent study demonstrated that variants in this region were capable of causing impaired autophagy in the myocytes (200). Autophagy defects cause loss of sarcomere protein quality control (201). Thus, causing DCM via a dominant negative mechanism (i.e., increased aggregated truncated peptides) rather than haploinsufficiency (177). Similarly, the 13 reported *TTN*tv we found were located in the very high exons PSI index A-band region of the gene, as shown in Figure 4.33. Specifically, four reported variants were proximal to the I-band post-slice region, and three were proximal to the M-line. Two of the variants proximal to the M-line were the recurring reported *TTN*tv c.95008C>T (p.Arg31670Ter) (82) and c.87624C>A (p.Tyr29208Ter) (169) variants found in 18.4% (14/76) and 7.9% (6/76) of the probands, respectively. The other four reported *TTN*tv were proximal to the I-band post-slice region known to cause DCM via haploinsufficiency (110). It is important to note that the functional role played by some of these 13 reported variants in *TTN*tv-related adult-onset DCM is inconclusive (110).

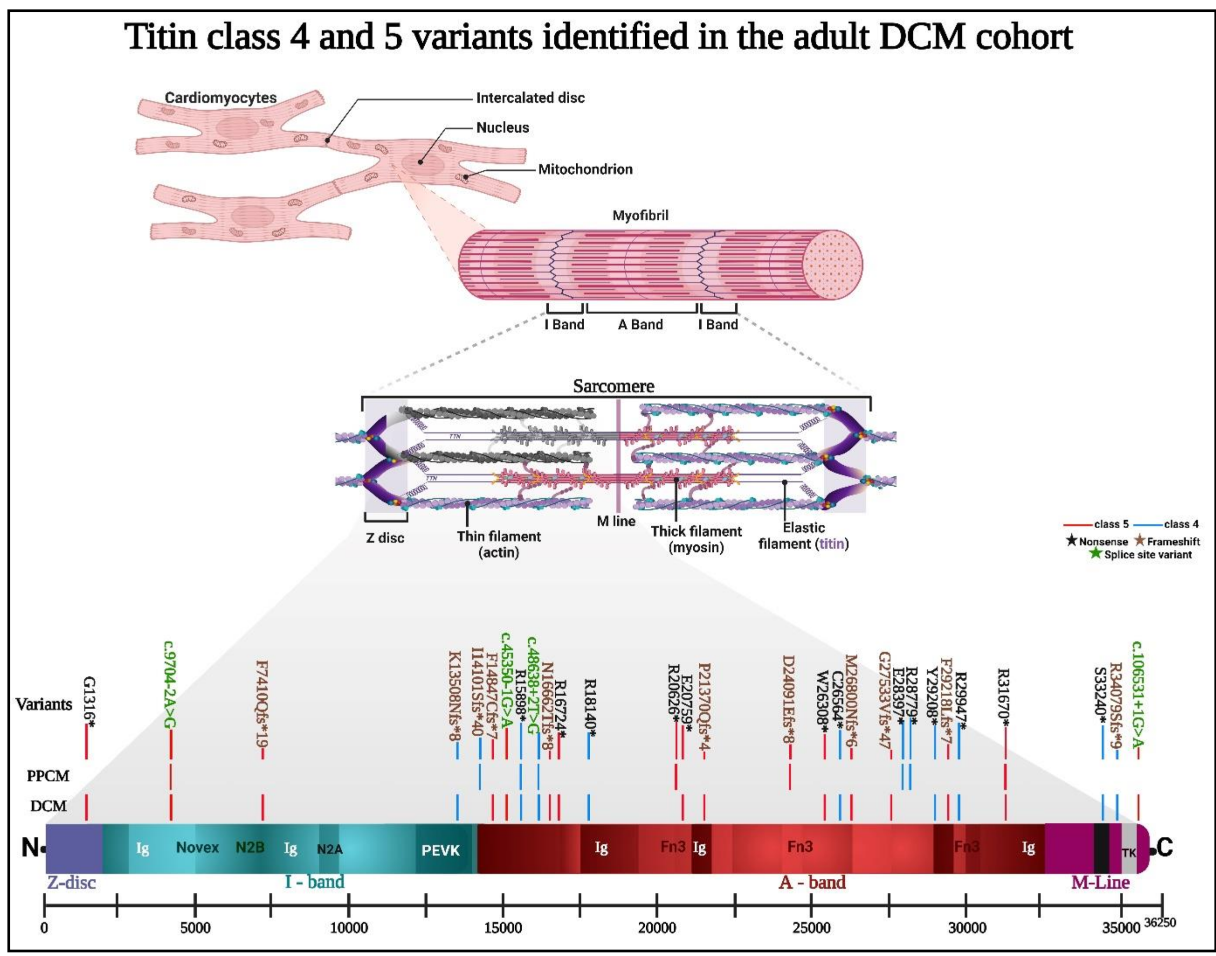


Figure 3.33: Titin's class 4 and 5 variants identified in the adult DCM cohort. *Class 4 -5 variants were identified in the adult-onset DCM. Stop-gained variants have a black font, frameshift variants in brown and splice site variants in green. The blue strips represent class 4 variants, and class 5 variants are in red. Drawn by P. Ndibangwi using Microsoft 365 PowerPoint.*

We also found 55.1% (16/29) of the *TTN*tv as novel located in distinct regions of the titin protein associated with DCM (202, 203), as shown in Figure 3.32. Three novel variants were proximal to the Z-disc and I-band pre-slice region, six proximal to the I-band post-slice region, and seven distal to the I-band post-slice region. Variants in these regions are shown to generate truncated titin peptides of varying sizes in the adult human heart (90). Firstly, the three novel variants identified outside the A-band region (the Z-disc and I-band pre-slice regions) raise many questions because the impact of variants in this region is unknown (204). Of particular note is that the proximal Z-disc and I-band pre-slice region supports sarcomere integrity (205). Class 4 and 5 variants in the region are highly associated with mitochondrial protein acetylation. Acetylation in this region promotes the titin isoform switch, and a switch from N2BA to N2B will stiffen this elastic portion of the protein spring (203). As a result, sarcomere integrity would be lost. Secondly, the other six novel variants were identified proximal to the I-band post-slice region where the mechanism of *TTN*tv-related DCM is associated with

haploinsufficiency (167, 168). Haploinsufficiency occurs when truncated mRNA undergoes nonsense mediated decay (NMD), resulting in a 50% reduction of the wild-type protein. In the sarcomere, there will be a decrease in the titin content and its force-generating capacity, thereby increasing sarcomere instability (206). Thirdly, we found seven novel variants distal to the I-band post-slice region. Titin's A-band region is known for causing adult-onset DCM, and the current mechanism of A-band *TTN*tv-related DCM is associated with dominant negative (110). That is when highly soluble truncated peptides weakly bond to their wild-type protein structures. The truncated peptides will dissociate and form aggregates in the cell, such as titin aggregates (110, 201). Titin protein aggregates may lead to sarcomere quality control defects that might further prevent the dissolution process, reducing the wildtype protein content and increasing sarcomere content instability (110, 195, 205).

These studies show that sarcomere protein force-generating capacity insufficiency is our cohort's common cause of *TTN*tv-related adult-onset DCM. Hence, we recommend that further research on antifibrotic effects (drug) should be considered.

*3.4.1.2.1.2.1.1.8 TTN*tv *c.89839C>T* and *MYBPC3 c.2176C>T*

The *TTN*tv c.89839C>T (p.Arg29947Ter) stop_gained variant has been reported in ClinVar associated with DCM. The variant carrier in this study had an early onset familial DCM similar to findings attributed to *TTN*tv (204). The proband also carried the *MYBPC3* c.2176C>T (p.Arg726Cys) missense variant previously reported in a cohort of HCM patients (207).

The human wildtype *MYBPC3* gene encodes for myosin-binding protein C, cardiac-type (cMyBP-C) protein found in the cross-bridge-bearing zone (C region) of A bands in sarcomere. The cMyBP-C protein regulates the interactions between the sarcomere's thick (myosin) and thin (actin) filaments. Its crucial role in maintaining the structural arrangement of sarcomere and mechanical response to cellular interactions triggered by changes in Ca^{2+} concentration is also essential. Alterations in the *MYBPC3* exon 23 coding DNA sequence may start a series of chained reactions in the sarcomere that may cause structural disarray and remodelling or abnormal Ca^{2+} handling or disrupt the sarcomere function.

Although the prevalence of class 4 and 5 variants in *MYBPC3* was approximately 4% in DCM patients with a wide range of phenotypes, (143) there is currently limited evidence for an association of MYBPC3 with DCM (120). (Most carriers of this *MYBPC3*:c.2176C>T (p.Arg726Cys) missense variant have had a range of penetrance and were mainly males

between 40 and 50 years old (77). Notably, most *MYBPC3*-related heritable DCM cases are due to frameshift variants causing haploinsufficiency (143). However, the *TTN*tv c.89839C>T in this proband is expected to play a crucial role in the pathomechanism of the phenotype presented. Besides, the pathomechanism of haploinsufficiency often reported in missense *MYBPC3*-related heritable DCM patients is less understood (143).

3.4.1.2.1.2.1.2 The beta-myosin heavy chain (*MYH7*) gene class 4 and 5 variants (n=4)

We identified four *MYH7* variants in this cohort; two have previously been reported in the ClinVar database.

3.4.1.2.1.2.1.2.1 MYH7: c.5791-2A>T and c.5792del

The *MYH7* c.5791-2A>T splice_acceptor has been reported once in the ClinVar, and the c.5792del variant is novel. Both variants surround another novel *MYH7* c.5791G>T (p.Gly1931Cys) pathogenic variant that has been reported in a family with severe biventricular apical hypertrophy (210). Also, patients' clinical manifestation, disease severity, and prognosis of monoallelic double variants illustrated in a cis-manner in *MYH7* are thought to be better (147).

3.4.1.2.1.2.1.2.2 MYH7 c.3748C>G

The *MYH7* c.3748C>G (p.Arg1250Gly) has been reported three times in the ClinVar data; however, there are no published findings for the mutation. Another study reported the Arg1250 amino acid change in an African American family with DCM casualties (211). Similarly, the *MYH7* c.3748C>G (p.Arg1250Gly) in our male MA proband diagnosed at age 22 years is on the same amino acid in the LMM domain associated with DCM (149). The *MYH7* c.3748C>G variant can affect the protein's charge and size at position 1250 by inducing conformational changes (147, 149).

3.4.1.2.1.2.1.2.3 MYH7 c.455C>G

The *MYH7* c.455C>G (p.Pro152Arg) class 4 variant is a novel missense mutation. The variant is in a highly conserved mutation hotspot on exon 5 of the *MYH7* HMM head domain associated with early onset of CMO (149). Similarly, our *MHY7* c.455C>G DCM patient had known published clinical features, for example, early onset age with HF complications mainly being

ventricular arrhythmias (149). It is essential to consider the broad variability of phenotype expression in *MHY7*-related DCM as we understand that this variant is the cause of the observed phenotype in our proband (143).

The *MYH7* gene encodes for beta-myosin heavy chain protein (motor protein) interacting with actin-based motor molecules to generate the force essential for cardiac muscle contraction. The proteolysis of beta-myosin heavy chain produces two protein domains: light meromyosin (LMM) and heavy meromyosin (HMM) (149). The HMM can be further cleaved into two globular sub-fragments (S1) and one rod-shaped sub-fragment (S2). The beta-myosin heavy chain protein also maintains the structural and organisational integrity of the cardiac sarcomere structure by arranging and aligning the thin and thick filaments within the sarcomere for efficient myocardial contraction (147, 212). The beta-myosin heavy chain protein's crucial roles help the sarcomere structure adapt to various physiological and pathological stresses (213).

Pathophysiological studies have reported that some diseases, such as the DCM, are caused by variants in the *MYH7* gene sequence that alter the expected protein structure (137, 214). *MYH7* missense variants cause a gain-of-function by producing poison peptides that disrupt the formation of functional sarcomeres, leading to increased contractility and delayed relaxation (147, 149). *MYH7*-related DCM is characterised by cardiac systolic dysfunction commonly caused by sarcomere disarray, impaired calcium ions (Ca2+) handling, abnormal myocardial contraction, or other factors (149). *MYH7*-related DCM is genetically heterogeneous, and systolic dysfunction is most common in patients with variants in the protein's LMM domain (137, 149). These authors also linked variants in the LMM domain to the later onset of DCM more than variants in other regions. The prevalence of *MYH7*-related DCM ranges from 4% to 20% in adult-onset patients (98), and in this study, the prevalence is 7.8%. The missense variants in the LMM domain were the common cause of *MYH7*-related DCM in this study—the schematic representation of the *MYH7* gene and its DCM-related variants, as shown in Figure 3.34.

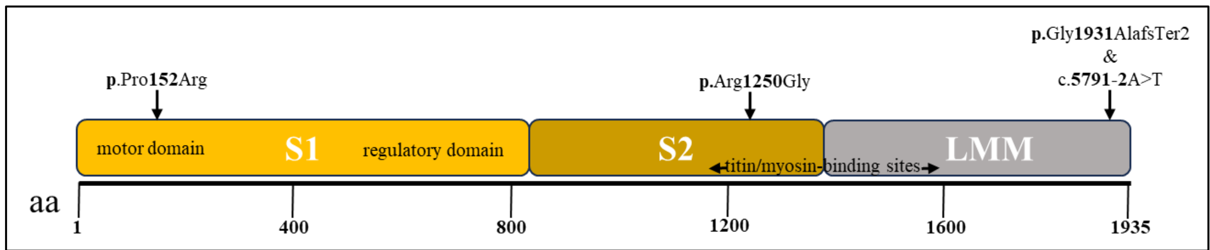


Figure 3.34: Schematic structure of *MYH7* with *DCM-related variants' location*. *The DCM-related MYH7 variants' locations were identified in the IMHOTEP DCM cohort and drawn by P. Ndibangwi using Microsoft 365 PowerPoint.*

3.4.1.2.1.2.1.3 The troponin type 2 (*TNNT2*) gene class 5 variant

3.4.1.2.1.2.1.3.1 TNNT2 c.421C>T

Similar to multiple reports in the literature, the missense *TNNT2* c.421C>T (p.Arg141Trp) variant carrier in this study had a FH of HF classified with severe early-onset DCM (215, 216).

Naturally, the *TNNT2* gene encodes for a heart muscle-specific troponin type 2 protein (cTnT), which forms part of a tropomyosin-binding subunit of the troponin complex structure. cTnT protein in the adult troponin protein complex is highly conserved in the tropomyosin binding domain. The troponin protein complex is located on the thin filament of the sarcomere and regulates myocardial contraction in response to fluctuation in intracellular concentration of calcium ions (Ca^{2+}) (217). Alterations in the *TNNT2* gene nucleotide sequences have been reported to cause DCM (217, 218), and the reported pathomechanism of *TNNT2*-related DCM is due to impaired sarcomere force generation caused by improper Ca^{2+} handling leading to cardiomyocyte stress or remodelling. For example, the *TNNT2* c.421C>T (p.Arg141Trp) variant is reported to cause a significant reduction in myofilament Ca^{2+} sensitivity for force development, thus leading to irregular heart muscle contraction (217). We, therefore, suggest that the *TNNT2* c.421C>T variant has caused the severe complex familial DCM phenotype we observed in our patient by impairing Ca^{2+} handling that triggers a remodelling process leading to ventricular cavity enlargement and wall thinning that ends up as a systolic dysfunction (17, 217).

3.4.1.2.1.2.2 The nuclear envelope (LMNA and RBM20) genes' variants

3.4.1.2.1.2.2.1 The lamin a/c (*LMNA*) gene class 4 and 5 variants (n=2)

Figure 3.35 shows the *LMNA* class 4 and 5 variants on the schematic diagram of the gene.

3.4.1.2.1.2.2.1.1 LMNA c.568C>T

Unlike the 6% prevalence reported in Norway, the *LMNA* c.568C>T (p.Arg190Trp) variant was identified in 16.7% of the MA probands in this DCM cohort (64). In the Norwegian cohort, 20% of the *LMNA* patients required heart transplantation, and similarly, our patients with the *LMNA* variants had severe phenotypes at presentation (219). The occurrence of *LMNA* c.568C>T variant in multiple young probands suggests that the manifestation of the disease has been exacerbated by factors such as a lack of good medical care quality (25, 183). Also, our LMNA c.568C>T variant carriers had the same ancestry and shared a common culture and heritage.

3.4.1.2.1.2.2.1.2 LMNA c.1622G>A

The c.1622G>A variant was also published in the literature (220) and again reported with established aetiology for familial DCM (221). Other studies have also shown that *LMNA* is a predominant nuclear envelope gene with a 27% heart transplantation rate (173).

The variants that alter the nuclear protein's structure are associated with severe DCM in 30% to 40% of the disease patients (95). Individuals with nuclear envelope variants, specifically *LMNA*, present more with ventricular arrhythmias in about 50% of patients (173), resulting in a high heart transplantation rate in about 20% of patients or risk of SCD (64). The development of ventricular arrhythmias has been demonstrated using animal and hiPSC experimental models showing disorganised cytoplasmic actin filaments and gene dysregulation caused by mutant lamin A protein (222, 223). In the hiPSC experiment, the hiPSC-derived cardiomyocytes from a patient with dilated cardiomyopathy showed abnormal nuclear morphology and specific disruptions in peripheral chromatin. The disrupted regions were enriched for transcriptionally active genes and had a lower frequency of contact with LAMIN B1. The mutant cardiomyocytes disrupted lamin A-chromatin interactions, which were enriched with genes related to non-myocyte lineages and these genes were correlated with higher expression in those lineages (223). Similarly, the myocardium from patients with LMNA variants showed abnormal expression of non-myocyte pathways (223).

The *LMNA* gene encodes for a component of the two-dimensional matrix proteins, a fibrous layer located on the nucleoplasm side next to the inner nuclear membrane of cardiomyocytes. These proteins (lamin A and lamin C) stabilise the cardiomyocyte's nucleus and may interact with the chromatin structure. The *LMNA* gene also plays a role in signalling and

mechanotransduction, nuclear stability during cell division, gene regulation and nuclear envelope integrity. Class 4 and 5 variants altering the *LMNA* gene's product have been associated with various genetic diseases called laminopathies (220, 224), for example, Emery-Dreifuss muscular dystrophy, limb-girdle muscular dystrophy and DCM. The DCM-related *LMNA* c.568C>T (224) and c.1622G>A (220, 222) variants we have identified are classified with functional molecular deleterious effects for familial DCM.

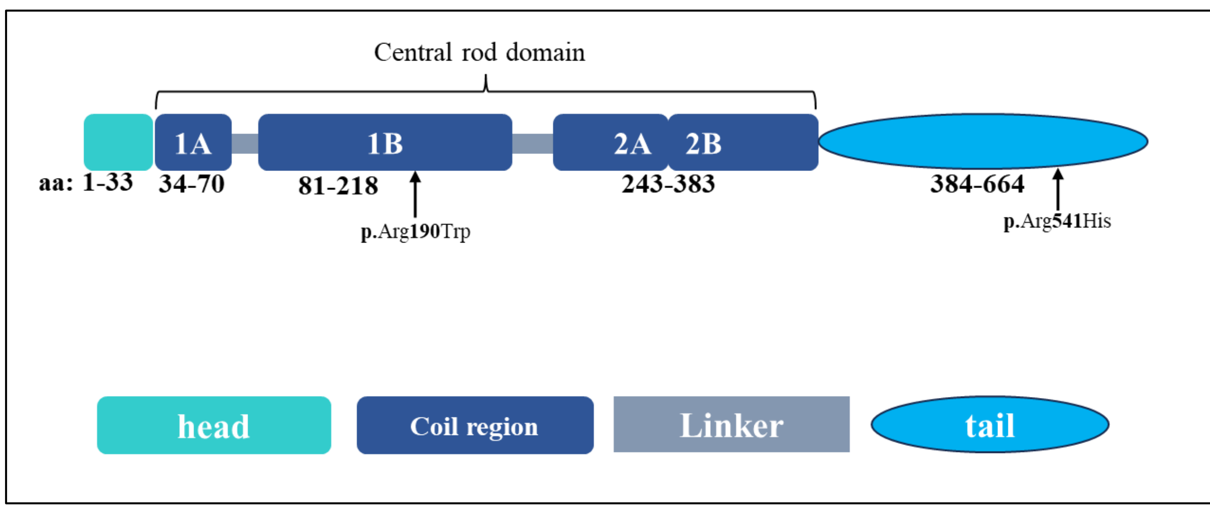


Figure 3.35: Schematic structure of *LMNA* with DCM-related variants' location. *DCM-related LMNA variants' location identified in the IMHOTEP DCM cohort, drawn by P. Ndibangwi using Microsoft 365 PowerPoint.*

In contrast to other studies where probands with *LMNA* missense variants had late onset at above 45 (220, 224), our probands were diagnosed before age 45. This phenotype suggests that the early onset age in our probands and their severe disease may be due to other risk factors such as socioeconomic status. Patients diagnosed with DCM outcomes such as HF at an early age might have phenotypes exacerbated by poverty (64). Low economic status is associated with DCM outcomes in low- and middle-income countries (LMICs), specifically in patients with lower education levels who have had a greater risk of cardiovascular events and mortality (183). For example, 23% of South African youth (specifically males) are not in employment, education, or training: a rate that is three times that of the UK, 5.4 times of Germany, 1.3 times that of Brazil, and 2.5 times of Malaysia (25). Thus, the majority of unemployed individuals in the LMICs have shown poor health-seeking habits, increased exposure to substance abuse and infectious diseases, and less access to good medical care quality and nutrition (25, 64).

3.4.1.2.1.2.2.2 The RNA-binding motif protein 20 (*RBM20*) gene class 4 and 5 variants

3.4.1.2.1.2.2.2.1 RBM20 c.1900C>T (p.Arg634Trp)

The *RBM20* c.1900C>T (p.Arg634Trp) class 5 missense variant has segregated with familial DCM (225, 226). Pathogenic missense variants in *RBM20* account for 2% to 6% of familial DCM and exhibit early onset and severe clinical expression (227). Like with the *LMNA* gene, *RBM20* variant carriers also have a high heart transplantation rate of 12%, but at a remarkably young age (mean 28.5 years) or risk sudden death (173). Similar to most published findings, the adolescent proband in this study had a severe DCM phenotype at presentation (228).

The wildtype *RBM20* gene encodes for an RNA-binding motif protein, as shown in Figure 3.36, which regulates the alternative splicing of about 30 specific mRNA sequences encoding for sarcomeric proteins, such as the *TTN* gene pre-mRNA sequence (229). The binding of RBM20 protein to particular mRNA sequences may influence the inclusion or exclusion of specific exons during splicing, thereby modulating the different isoforms of sarcomeric proteins, such as titin. For example, by regulating the alternate splicing of titin mRNA, the RBM20 may switch the titin isoform from N2BA to N2B, thereby influencing the stiffness of the elastic portion of the sarcomere protein (203). Thus, impaired alternate splicing activities of the *RBM20* gene may impact the sarcomere contractility, resulting in its dysfunction and triggering the development of CMO and HF (228). Moreover, *RBM20* pathogenic variants are associated with the titin isoforms switch, greater titin compliance, and irregular calcium homeostasis, leading to DCM development (203).

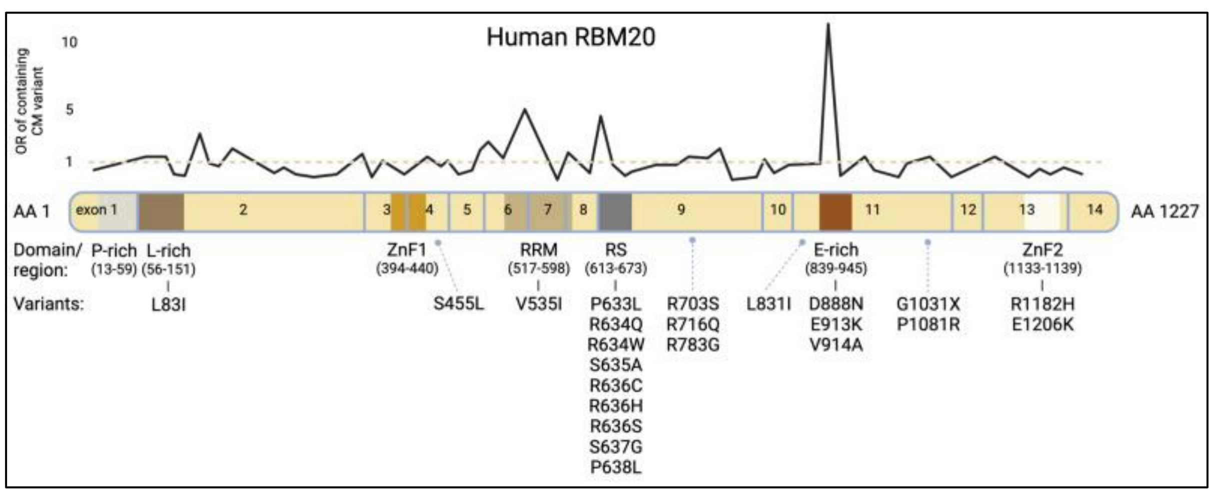


Figure 3.36: Schematic protein structure of human RBM20 with corresponding exons. *The black line chart displays the odds ratio (OR) for variant observation within the respective sections in a cardiomyopathy population vs. the general population (Genome Aggregation Database [gnomAD]) (227).*

3.4.1.2.1.2.3 The Z-disc (BAG3 and FLNC) gene variants

Class 4 and 5 variants in Z-disc proteins can cause DCM, hypertrophic and arrhythmogenic cardiomyopathies, which may lead to HF, arrhythmias, and sudden death (230). One Z-disc gene (*BAG3*) associated with DCM was found in this cohort (120). We found four adults (one BA, one WA, and two MA probands) carrying three *BAG3* variants, as shown in Figure 3.37.

3.4.1.2.1.2.3.1 The BCL2-associated athanogene 3 (*BAG3*) gene class 4 and 5 variants

3.4.1.2.1.2.3.1.1 BAG3 c.925C>T (p.Arg309Ter)

The *BAG3* c.925C>T (p.Arg309Ter) stop_gained variant has been reported eight times in the ClinVar database. As other studies have shown (231, 232), the variant segregates with familial DCM, showing high penetrance and early age of onset in our cohort.

3.4.1.2.1.2.3.1.2 BAG3 c.214G>T (p.Glu72Ter)

The c.214G>T (p.Glu72Ter) is a novel variant, and the expected mRNA produced by the mutant DNA sequence might be targeted for nonsense-mediated decay (NMD) (230). Hence, the disease mechanism might be haploinsufficiency or a dominant negative effect (233).

3.4.1.2.1.2.3.1.3 BAG3 c.107G>A (p.Trp36Ter)

The BAG3 protein domain harbouring this novel variant is known for causing DCM (83). The mRNA produced by the novel *BAG3* c.107G>A (p.Trp36Ter) variant might be targeted for NMD, and the expected mechanism of the disease may be haploinsufficiency (233).

The *BAG3* gene encodes for one of three Z-disc BCL2-associated athanogene (BAG) proteins, a molecular chaperone that modulates autophagy and protein homeostasis in the human heart muscle (230). The BAG proteins compete with Hip for binding to the Hsc70/Hsp70 ATPase domain to regulate protein folding and promote substrate release during proteostasis (201). All three BAG proteins bind with high affinity to the ATPase domain of Hsc70 and inhibit its chaperone activity in a Hip-repressible manner. The BAG3 protein also regulates autophagy, cell survival and apoptosis to maintain the structural integrity of the myocardium in response to cellular stress, such as heat shock and oxidative and mechanical stress (230, 233). Plausible pathogenic variants in the *BAG3* gene that can impair these processes from functioning would lead to cardiac diseases such as familial DCM (231), myofibrillar myopathy, or myocarditis

(233). Over 86% of the reported *BAG3* variants of clinical significance in literature are truncating mutations (stop_gained and frameshift), whereas the majority of those with uncertain significance are missense (32, 231, 233). About 30.1% of patients with an incidence of 5.1% per year had clinical endpoints, including cardiac death, heart transplant, LV assist device, aborted SCD, and severe ventricular arrhythmia (234). Similarly, the three *BAG3* pathogenic variants reported in this study were stop_gained mutations found in probands with severe DCM (234). In addition, our probands were diagnosed with severe DCM within the variable age of onset from 18 years to 64 years reported in the literature (233).

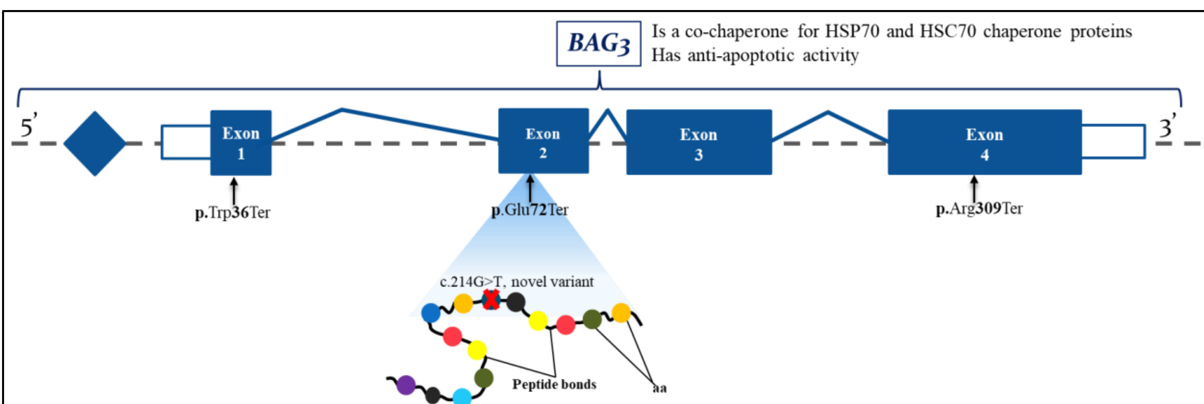


Figure 3.37: Schematic structure of *BAG3* with DCM-related variants' location. *Schematic structure of the functional domains of the BAG3 protein, drawn by P. Ndibangwi using Microsoft 365 PowerPoint.*

3.4.1.2.1.2.3.2 The Filamin C (*FLNC*) gene class 4 and 5 variants

3.4.1.2.1.2.3.2.1 FLNC c.4021C>T

The *FLNC* c.4021C>T (p.Arg1341Ter) is reported only in the ClinVar database and is in the gene's domains associated with a high risk of SCD (235). The truncating variants are expected to create a shortened, non-functioning filamin-C protein causing *FLNC*-related DCM via LoF and haploinsufficiency pathomechanism with severe phenotype (107, 235). Notably, the proband was young and diagnosed with severe DCM, as commonly reported in patients at an average age of 39.7 years (107). We also found the *FLNC* c.4021C>T variant in the BA proband's affected brother discussed in section 3.3.1.3.1.3 as *family segregation analysis*.

3.4.1.2.1.2.3.2.2 FLNC c.6907C>T

The *FLNC* c.6907C>T (p.Gln2303Ter) is reported in the literature (107) and, like other stop_gained *FLNC* variants in the ROD2 domain, has been associated with severe DCM. The

truncating variants in the ROD2 domain are expected to create a non-functioning filamin-C protein causing *FLNC*-related DCM via LoF pathomechanism with severe phenotype (107, 235). Notably, the variant carriers in our cohort were young and diagnosed with severe DCM (107) and shared a common culture and geographical region.

3.4.1.2.1.2.3.2.3 FLNC c.6006dup

Like other *FLNC* LoF variant products, the novel splice_acceptor *FLNC* c.6006dup (p.Ile2003AspfsTer35) can cause severe DCM (235). The clinical key features of DCM due to the *FLNC*-related DCM-causing variant include normal to low voltage ECG, subepicardial enhancement on image and risk of ventricular arrhythmia (235, 236).

The *FLNC* gene (one of three related gamma filamin genes) encodes the filamin-C protein, essential in crosslinking the sarcomere actin filaments to the Z-disc. The filamin-C protein further intercalates the Z-disc of the sarcomere onto the cell membrane to form a complex cytoskeletal structure (237). This complex structure lies on the load-bearing sites of cardiomyocytes to maintain cardiomyocytes' structural stability and mediate crucial cell signalling pathways (237).

DCM-related *FLNC* diseases are mainly caused by truncating variants, as shown in Figure 3.38 (107). Experimental findings have shown sarcomere disarray, impaired mechanical properties within the cardiomyocytes, disruption of signal transduction and cardiomyocyte apoptosis and remodelling in most *FLNC* patients (235, 238). These pathomechanisms impair myocardial function, causing myofibrillar myopathy, muscular phenotypes, or DCM. *FLNC* disease-causing variants that cause LoF of the gene product are associated with ACM and DCM (235, 236). Also, *FLNC*-related DCM ranged from a malignant clinical course to a severe phenotype with a risk of SCD (107), and the severe DCM phenotype clinical course has been reported as due to haploinsufficiency. The prevalence of severe *FLNC*-related DCM is < 9% (97, 143), and in this study, it is reported to be 5.9%, the variants identified in two BA and two MA probands.

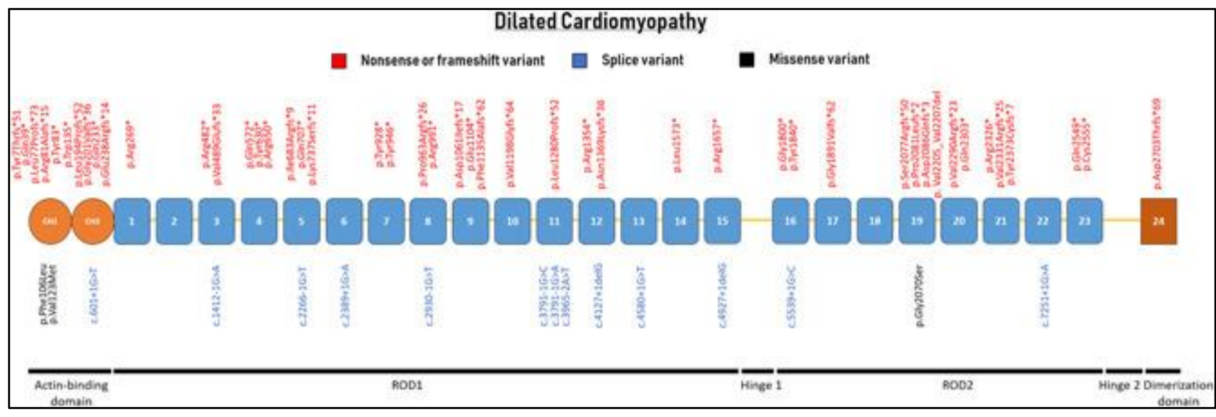


Figure 3.38: Schematic representation of the FLNC gene with its protein-coding domains. *Numbers inside the boxes refer to the Ig-like domains of filamin C. Above and below the schematic are variants associated with DCM annotated at the protein level (107).*

3.4.1.2.1.2.4 The cytoskeletal structure (DMD) gene variant

3.4.1.2.1.2.4.1 The dystrophin (*DMD*) gene class 5 variant

3.4.1.2.1.2.4.1.1 DMD c.1603-2A>G

The functional experiments to illustrate the molecular consequence of the c.1603-2A>G variant were unavailable; however, the impact of the c.1603-2A>C variant is explained to be due to the skipping of exon 14-15 (239, 240). The exon skipping is expected to result in a partially functioning DMD protein, probably a milder phenotype (240).

The *DMD* gene encodes for dystrophin protein, a large protein of about 2 Mb that intercalates the inner cytoskeletal organelles with the extracellular matrix structures. The dystrophin forms complex structures, especially glycoprotein, and dystrophin-glycoprotein complexes are involved in stabilising sarcolemma, signalling events, transmitting synapses, and central nervous system activity. *DMD*-related disease-causing variants might cause Duchenne muscular dystrophy, Becker muscular dystrophy and CMO. *DMD*-related adult-onset CMO patients have had an early onset disease associated with mild skeletal muscle findings (114) and, in some cases, have progressed rapidly without muscular dystrophy (69). *DMD*-related DCM is inherited in an X-linked pattern, and about <0.5% of the patients have presented with the severe phenotype (95).

3.4.1.2.1.2.5 Other genes' (PKP2, JUP and SCN5A) variants

We found cell-cell junctions and cell-membranes class 4 and 5 variants in three genes (*PKP2*, *JUP* and *SCN5A*). Desmosomes and other cell membrane proteins are essential structures in cells constantly under mechanical stress. In the myocardium, these proteins maintain cardiomyocytes' integrity (through adhesion forces) and cell-to-cell ion transport. Variants that alter these protein structures have been associated with VT, ventricular arrhythmia, and CMO. For example, five desmosome genes have been described with a prevalence of about 13% in end-stage DCM probands (241), and we think that our findings suggest possible ACM and DCM overlap in this cohort, which we recommend should be further investigated.

3.4.1.2.1.2.5.1 The plakophilin-2 (*PKP2*) gene

The *PKP2* c.2509del (p.Ser837ValfsTer94) frameshift variant is reported in the literature (242). Class 4 and 5 variants in the *PKP2* gene are commonly associated with arrhythmogenic right ventricular cardiomyopathy (ARVC); however, *PKP2* pathogenic variant carriers have had known aetiology of DCM, such as ventricular arrhythmias (98). The *PKP2* gene has been reported amongst five other desmosomal genes with class 4 and 5 variants in 13% of end-stage DCM patients undergoing heart transplantation (241). Generally, *PKP2*-related CMOs are severe cases, and often, the patients would require heart transplants, or SCD would be inevitable.

The *PKP2* class 5 heterozygous frameshift c.2509del variant we reported is known to cause arrhythmogenic ventricular dysplasia (242). The *PKP2* c.2509del variant is expected to create a longer protein product; however, the predicted protein product's molecular consequences have not been determined using a function experiment. Most carriers of this variant in literature have presented a range of penetrance and various ARVC phenotypes, such as biventricular involvement (243). Similarly, our proband was diagnosed with arrhythmias in idiopathic DCM.

A normal *PKP2* gene encodes for the desmosomal plakophilin-2 protein found primarily in the heart muscle's cardiomyocyte-to-cardiomyocyte boundaries, nucleus, and cytoskeleton. The plakophilin-2 protein comprises armadillo repeats linking the desmoplakin to desmoglein-2 protein filaments to keep the desmosomes intact. Alterations in the *PKP2* gene sequence will generate poor-quality plakophilin-2 protein with weakened bonds linking it to the desmoplakin and desmoglein-2 protein filaments that keep adjacent cardiomyocytes intact. Such changes decrease the cardiomyocyte-to-cardiomyocyte tightness and increase detachment (244). The

impaired cardiomyocyte-to-cardiomyocyte adhesion will increase electrical impulse imbalances and induce immune cell infiltration, resulting in abnormal structural remodelling, tachycardia, and ventricular fibrillation.

3.4.1.2.1.2.5.2 The junction plakoglobin (*JUP*) gene

The *JUP* c.960dup (p.Asn321Ter) variant affects an evolutionarily conserved region of the *JUP* gene, and the mRNA produced might be targeted for NMD, resulting in a haploinsufficiency pathomechanism (245). Exon 6 is transcribed in all *JUP* transcripts and expressed in the head, heart, and other tissues. The *JUP* gene encodes for a desmosomal plakoglobin (γ-catenin) protein found primarily in cell-to-cell adherens junctions. The plakoglobin connects the cytoskeletal actin structure to the cadherin in the adherens junctions in the heart muscle. The complex structure stabilises the desmosomes to maintain the structural and mechanical integrity of the heart muscle. Plakoglobin is the only known adherens constituent common to sub-membranous plaques of both desmosomes and intermediate junctions. Altering the plakoglobin structure may result in a loss of mechanical stability, signalling and adherens regulation in the heart muscle, which might cause multiple cardiovascular system disorders. For example, *JUP*-related diseases include Naxos, ARVC and DCM disorders. *JUP*-related adult-onset DCM is uncommon (246), and *JUP*-related DCM patients are usually diagnosed with life-threatening phenotypes characterised by left ventricular or biventricular involvement (247).

3.4.1.2.1.2.5.3 The sodium channel protein type 5 subunit alpha (*SCN5A*) gene

Literature has reported that 2% to 10% of DCM patients are *SCN5A* variant carriers (143), and <1% of familial DCM cases have been documented carrying *SCN5A* missense variants (137). The *SCN5A* c.3683 A>G (Tyr1228Cys) exon 21 variant and c.4012C>A (p. Leu1338Ile) exon 23 variant affect the protein in regions conserved across multiple species and in humans. Pathogenic variants in the cytoplasmic ion transport domain have caused prolonged QT (LQT) syndrome type 3, Brugada syndrome type 1 and severe DCM in an AD manner (248, 249). Similarly, we hypothesise that our probands' heart muscles might be unable to maintain normal cardiac electrical stability due to these variants; however, due to unavailable functional data, we have assigned class 3 LP status to these variants pending further reclassification.

The *SCN5A* gene encodes for an alpha subunit of the cardiac sodium ion (Na^+) channel transmembrane protein, mediating the voltage-dependent Na+ permeability of excitable cardiomyocyte membranes. It thus generates action potential and maintains the propagation of the action potential for synchronous depolarisation and repolarisation to keep a normal heart rhythm for efficient electrical conduction of the heart muscles during mechanical activity. Inefficient regulation of the excitability of cardiomyocytes caused by variations in the *SCN5A* gene sequence may lead to various cardiac electrical disorders and arrhythmias. Notably, excitable cardiomyocytes' membrane action potential is detectable with an ECG and cardiac disorders such as heart block on ECG (or late potential on signal-average) are often found in *SCN5A*-related DCM (250). Most *SCN5A*-related DCM cases have required heart transplants or had a SCD.

3.4.1.2.1.2.6 Novel class 4 and 5 variants found in the adult-onset DCM cohort

Half of the DCM-causing variants (51.0%, 26/51) we identified were novel and harboured by high evidence of DCM-causing genes (120). The proportion of novel class 4 and 5 variants was high in this study compared to international studies with a similar design (188). The high proportion of novel variants may be explained in that *i*) maybe the human genome reference sequence we used for variants calling contained very few complete African exomes, *ii*) the genetic determinants of the African adult-onset DCM are largely unstudied (37, 57), and *iii*) no DCM GWAS have been conducted to determine gene loci associated with the phenotype (251). More novel class 4 and 5 variants would likely be identified in this cohort because we have used stringent ACMG/AMP and ClinGen criteria targeting DCM disease to assign class 3 to most novel variants (35).

Table 3.5 and Figure 3.8 notably revealed that the most prevalent gene is *TTN*, harbouring 56.9% of class 4 and 5 variants. In addition, the *MYH7* was also prevalent amongst the BA and *LMNA* in the MA probands. To better understand these targeted NGS results, we have discussed the findings under the headings class 5 and *class 4 variants*. Also, we considered class 3 variants as of unknown significance variants and discussed them as *class 3 variants*.

3.4.1.2.1.3 Relatives screened for class 4 and 5 variants

3.4.1.2.1.3.1 Family of 10100344 and 10100312 with BAG3 c.925C>T variant

We screened a known highly penetrant *BAG3* (p.Arg309Ter) c.925C>T variant in the relatives' available gDNA samples of two families (231). In the first family, the proband (II:4), his daughter (III:2) and two grandchildren (IV:1 and IV:2) were positive (+/-) for the *BAG3* c.925C>T variant, as shown in the family pedigree in Figure 3.14. Participant III:2 was diagnosed at the age of 55 years old with DCM, and IV:2 was diagnosed with DCM at the age of 26 years old. The c.925C>T has also been reported in other cohorts with variable onset ages, such as in this family (231, 233, 234). In the second family, the proband (ID 10100344) and a clinically affected 22-year-old daughter were also positive for the variant. The *BAG3* c.925C>T variant has been segregated with familial DCM (231). Literature has also shown that the variant carriers have had an early onset before 37 years (231, 232). However, onset is delayed in some affected individuals, such as our 55-year-old taking hypertensive medication such as beta-blockers and has lived without requiring hospitalisation (231). In contrast, other carriers at first presentation had severe phenotypes requiring heart transplants, exacerbated in others by postpartum CMO or viral infections such as HIV (233).

3.4.1.2.1.3.2 Family of 10100323 with TNNT2 c.421C>T variant

The relatives' available gDNA samples were screened for a reported *TNNT2* (p.Arg141Trp) c.421C>T missense variant (17). The proband's younger at-risk sister (II:4), suspected of having DCM at the age of 39 years, was found carrying this variant, Figure 3.15. There was no sample for their deceased mother, who had a history of heart disease.

We have described the natural history of *TNNT2* above and its pathomechanism for DCM, and we, therefore, hypothesise *TNNT2* (p.Arg141Trp) c.421C>T missense variant might have caused their DCM (17, 217). Our 10100323 family might have had a familial CMO because the *TNNT2* c.421C>T variant is located in the tropomyosin BD 1 domain of the protein. Pathogenic missense variants in the tropomyosin BD 1 domain have been shown to cause CMO phenotype by impairing Ca^{2+} handling, triggering a remodelling process (252). The significant reduction in myofilament Ca^{2+} sensitivity for force development caused cardiomyocyte remodelling, leading to irregular heart muscle contraction (217). The cardiomyocyte remodelling process preceded the heart's ventricular cavity enlargement and wall thinning, leading to left ventricular systolic dysfunction (LVEF =30% in our patient).

3.4.1.2.1.3.3 Family of 10100351 with FLNC c.4021C>T variant

The *FLNC (*p.Arg1341Ter) c.4021C>T nonsense variant was carried by a 22-year-old female proband (II:4) that was validated, and relatives were screened (Figure 3.16). The affected older brother (II:1), clinically diagnosed with DCM at age 29 years, was positive for the variant. However, the older brother was diagnosed a few months after the proband, and we suggest the proband's early onset might have been exacerbated by the human immunodeficiency viral infection (253) or because a possible genetic modifier, *TTN* c.34408A>G (p.Lys11470Glu) variant, is present in this individual. Interestingly, although their parents' gDNA samples were unavailable for screening, the main variant *FLNC* c.4021C>T has been reported in an infant diagnosed with *FLNC*-related CMO (254).

Truncating variants in the *FLNC* gene cause protein haploinsufficiency in documented DCM patients (235), and some carriers have developed severe or overlapping phenotypes of dilated and left-dominant ACM with frequent premature sudden death complications (235). We hypothesise that similar pathomechanism underlies the disease in the ID 10100351 family. However, this may be the first case to show the variable expressivity of the variant *FLNC* c.4021C>T variant in CMO.

3.4.1.2.1.3.4 Family of 101304303 with SCN5A c.589G>T variant

The proband (II:2) carrying the *SCN5A (*p.Asp197Tyr) c.589G>T missense variant was diagnosed at age 18 years and is a trio of family members whose at-risk parents were screened, Figure 3.17. The proband's father (I:2), who was 43 years old at the time of referral for familial DCM clinical assessment, tested positive for the variant. Neither affected relative carried any other suspected variant of interest that may be associated with heritable CMO.

We assigned *SCN5A* c.589G>T (p.Asp197Tyr) exon 5 variant (Align GVGD: Class C65, CADD: 29.8, Mutation taster: Deleterious, PolyPhen2: Probably damaging, SIFT: Deleterious) to a class 4 (35). Although we assigned the variant to class 4, it was not included in Table 3.3 because it has been reported as a class 3 in *SCN5A*-related conditions, i.e., Brugada syndrome and CMO (255). In another study, the same amino acid residue, p.Asp197His, was assigned to class 5, and it segregated with severe DCM (256). In this study, the amino acid changes from Aspartate (acidic molecular) to Tyrosine (polar molecule), not the Histidine (basic). The *SCN5A* c.589G>T variant affects the gene sequence in a cytoplasmic ion transport domain region conserved in humans and across multiple species (248). Most *SCN5A* missense variants

in this domain have caused prolonged QT (LQT) syndrome type 3, Brugada syndrome type 1 and severe DCM in an AD manner (248, 249). Therefore, we hypothesise that the affected individuals' heart muscles might be unable to maintain normal cardiac electrical stability during mechanical stress, leading to the observed phenotype (248, 256).

3.4.1.2.2 Class 3 variant findings

About 50% (223/450) of the adult-onset DCM probands had only class 3 variant (s): 61.4% (137/223) BA, 34.1% (76/223) MA and 4.5% (10/223) WA probands. Moreover, females accounted for 61% (136/223) of these probands with only class 3 variants, and most of the female probands were diagnosed with PPCM. Females diagnosed during pregnancy and of African descent have been explained above as being risk factors associated with DCM (257).

We identified 213 variants in 33 CMO disease-causing genes carried by the probands with only class 3 variants. To better understand if these class 3 variants have modifier effects, we described them according to the various ancestry within the cohort. Of the 213 variants, 12 barely meet the class 4 variants' criteria, which we have classified as class 3 due to lack of evidence to support LP status, as reported in Table 3.7.

3.4.1.2.2.1 Class 3 variant carriers of Black-African ancestry (BA)

In this study, we reported 137 BA adult-onset DCM probands carrying only class 3 variants; the majority were females, 64.2% (88/137). Notably, 66% (58/88) of these females were diagnosed as having peripartum DCM. Although there was no significant difference between the overall adult-onset DCM sex recorded at diagnoses and probands carrying only class 3 variants, it is essential to note that 66% of the BA probands with only class 3 variants had PPCM. It is important because the genetics of PPCM are unknown, and perhaps the adult-onset DCM genes panel testing we used was ineffective for our BA PPCM probands. Also, only 13.1% (18/137) of these BA probands were diagnosed with familial DCM, and we think that the genetic aetiologies of the disease are unknown in the other probands. Further, the 137 BA adult-onset DCM probands were young adult-onsets with a mean age of 33.7 years at diagnosis, probably because most of the probands were of economically active age and may afford frequent hospital visits for their disease to be classified early. The 137 adult-onset BA probands carried 130 class 3 variants identified in 25 CMO-associated genes, as shown in Figure 4.19.

Most, 90.8% (118/130) of the class 3 variants had missense molecular consequences, and 58.5% (76/130) were novel, with many found in the *TTN* gene. The most frequently identified gene in the BA probands was *TTN,* carried by 59.9% (82/137) of the probands with only class 3 variants, and the *TTN* gene had a class 3 variants prevalence of 53.8% (70/130). Except for one reported splice-donor variant, all the *TTN* class 3 variants were missense mutations. The natural *TTN* gene harbours many missense variants that are well tolerated in the general population, thus making assigning disease-causing status difficult without functional work. The *TTN* gene encodes for a sarcomere giant elastic protein called titin, described above. Also, the other most frequent genes we identified in the adult-onset BA probands that may cause DCM were *SCN5A* in 12 and *ALPK3* in 11 probands carrying only class 3 variants. These two genes are associated with each other in that they interact with one another and the sarcomere proteins to maintain cardiomyocytes' structural stability and integrity. Some of the class 3 variants these genes harboured might modify our BA population's DCM phenotypes. For example, the *MHY7* c.2183C>T (p.Ala728Val) missense variant we found in ten BA and MA probands with DCM has been reported in a single family with HCM (258). Perhaps this may signify that the variants' dose contributed to their phenotypes. In SS Table 2, out of the 88 BA probands with novel class 3 variants, 38 (43.2%) carried multiple unknown class 3 variants. However, majority were *TTN* missense variants.

3.4.1.2.2.2 Class 3 variant carriers of Mixed ancestry (MA)

We also reported 76 MA adult-onset DCM probands carrying only class 3 variants; the majority were females, 57.9% (44/76). Although there was no significant difference between the overall MA adult-onset DCM probands sex distribution recorded at diagnoses and MA probands carrying only class 3 variants, it is important to note that 59.1% (26/44) of the MA probands with only class 3 variants had PPCM. It is important because the genetics of PPCM are unknown, and perhaps the adult-onset DCM genes panel testing we used was ineffective for our PPCM probands. Also, only 19.7% (15/76) of these MA probands were diagnosed with familial DCM, which indicates that the cause of their disease is unknown genetic aetiologies. Further, the 76 MA adult-onset DCM probands were young adult-onsets with a mean age of 35.4 years at diagnosis, probably because most were economically active and could afford frequent hospital visits. The 76 adult-onset MA probands carried 102 class 3 variants identified in 20 CMO-associated genes, as shown in Figure 4.20.

We found that 92.2% (94/102) of the class 3 variants carried by the MA probands had missense molecular consequences, and 55.9% (57/102) were novel, with many identified in the *TTN* gene. Like in the BA probands, the most frequently reported gene in the MA probands with only class 3 variants was *TTN,* carried by 48.7% (37/76) of the MA probands and in the MA probands, the *TTN* gene had a class 3 variant prevalence of 45.1% (46/102). Six of the 37 MA probands carried splice-site, and 31 carried missense *TTN* class 3 variants. We reported two of the splice-site variants as likely pathogenic in Table 3.7. Also, the other most frequent genes we identified that may modify DCM phenotype were *FLNC* and *MYBPC3* genes, each in eight MA probands carrying only class 3 variants. These two genes are associated with each other in that they interact to maintain the cytoskeletal structural integrity and intercalate the sarcomere to stabilise it during mechanical stress.

3.4.1.2.2.3 Class 3 variant carriers of White ancestry (WA)

Unlike in the BA and MA probands, we reported 10 WA adult DCM probands carrying only class 3 variants in this study, with the majority being males, 70% (7/10). Only one proband (male) was diagnosed with familial DCM, and one female was diagnosed as having peripartum DCM, which indicates that the cause of their disease is unknown genetic aetiologies. Further, the 10 WA adult-onset DCM probands were much older, with a mean age of 47.1 years at diagnosis. The 10 WA probands carried 17 class 3 variants identified in six CMO-associated genes, as shown in Figure 4.21.

Most, 88.2% (15/17) of the class 3 variants had missense molecular consequences, and 58.8% (10/17) were reported with many in the *TTN* gene. Similarly, most White adult-onset probands, like their BA and MA counterparts with only class 3 variants, carried *TTN* variants with a prevalence of 70.6% (12/17). Also, except for two probands, one with a reported splice-acceptor and the other with a novel Inframe-deletion variant, the White probands carried mainly *TTN* class 3 missense variants. We also reported five missense variants found in the *ALPK3*, *FLNC*, *RBM20*, *SCN5A*, and *TMEM43* genes in one White proband each. These genes are associated with each other in that they are involved in cardiac energy generation and transduction. We think the class 3 variants harboured by these genes possibly modify the DCM phenotypes in our WA probands.

3.4.1.2.3 Class 1 and 2 variant carriers' findings

All the class 1 or class 2 variants were excluded from the analysis because the study of genetic polymorphism is beyond the scope of this thesis. Genetic polymorphism is the occurrence of multiple common variants at a specific position in the cDNA sequence of a gene within a population. They may be a single nucleotide polymorphism (SNP), insertion or deletion, or any other genetic variant and a modifier for the severity of Mendelian diseases, such as the DCM (147). We found 151 adult-onset DCM probands as class 1 and class 2 variant carriers and have listed these variants to be explored in another IMHOTEP study.

3.4.2 Peripartum cardiomyopathy (PPCM) findings

3.4.2.1 Baseline findings for adult-onset PPCM cohort

DCM female patients diagnosed during pregnancy were classified as peripartum cardiomyopathy (PPCM) probands because the clinical entity affects women of childbearing age more during pregnancy or in the early post-partum period (160). We reported 136 PPCM probands with an average of 29.6 years (SD 7.0 years) at diagnosis. Of the 136 probands, 13 (nine BA and four MA) carried classes 4 and 5 variants with no significant age difference compared to the entire PPCM probands. When we look at the 13 probands' distribution according to the various ancestry, we find that only probands of Black African descent carried the classes 4 and 5 variants identified. Like in other studies, DCM patients of Black African descent are more likely to have PPCM. That is, epidemiological reports have shown that Haitian and Nigerian women of childbearing have the highest incidence rate of PPCM compared to Japanese women with the lowest prevalence (257).

When we considered the combined total of females with genotype-positive results in our adult-onset DCM cohort, these PPCM probands formed 40.6% (13/32) of the females. The 40.6% genotype-positive rate is higher than the 22% clinical diagnostic aetiology reported in Figure 4.3. We think that PPCM in our cohort is a distinct clinical entity (158, 159) specifically because some authors have identified in more than 95% of PPCM patients that ECG abnormalities are risk factors (259).

3.4.2.2 Targeted sequencing findings for the PPCM cohort

3.4.2.2.1 Class 4 and 5 variants found in PPCM probands

The 13 PPCM probands (seven carrying class 5 and six carrying class 4 variants) yielded a diagnostic rate of 9.6% for this clinical entity. International PPCM studies with participants of Black African descent have yielded similar results ranging from 10% to 17.1% (158, 159). Unlike in the literature, our PPCM probands carried three genes harbouring 12 variants: six class 5 variants in *BAG3* and *TTN* and six class 4 variants in *MYH7* and *TTN* genes presented in Table 3.8. Except for the *MYH7* missense variant, the *BAG3* and *TTN* mutations were truncating variants. These 12 variants are in sarcomere genes whose protein product forms the mechanical motor section of the complex structure used during heart muscle contraction. Reported variants that cause diseases by altering sarcomere protein structures are linked to cardiac conduction disorders, ventricular arrhythmia, and high heart transplantation rate in

heritable DCM. The molecular consequences of these variants we have identified in our PPCM probands were described under the headings *the sarcomere genes* and *the Z-disc genes*.

3.4.2.2.2 Class 3 variants found in the PPCM probands

We reported 85 (58 BA, 26 MA and one White) PPCM probands carrying only class 3 variants in Table 3.7. We showed that these probands form 63% (85/135) of the female probands in this DCM cohort of participants carrying only class 3 variants. When we proceeded to analyse the 85 PPCM in comparison to the other 50 DCM female probands, we observed that the PPCM probands were, on average, younger with a mean age of 30.2 years (SD 7.0) compared to 34.6 years (SD 12.1) for the 50 probands at diagnosis. However, when we compared their disease onset ages with those in international studies, we found similar average ages ranging from 30 to 34 years at diagnosis (158, 159).

We also found that 95.3% (81/85) of our PPCM probands had non-familial disease aetiologies, and non-familial DCM was the reported FH. Although being Black has been explained to be a risk factor for developing PPCM, only 4.7% of familial PPCM probands reported were patients of BA. We then describe the class 3 variants finding according to the various population subtypes.

3.4.2.2.2.1 PPCM probands with class 3 variants of Black-African ancestry (BA)

The targeted NGS data analysis revealed that 68.2% (58/93) of BA probands carried only class 3 variants for this population group. The 58 BA PPCM probands carried 75 class 3 variants identified in 20 genes, Figure 4.27. Of the 58 BA probands, 29 (50%) carried only one class 3 variant each, while the remaining probands carried multiple class 3 variants. We could not analyse the yield of the 50% multiple class 3 variants carrier rate with the literature due to the nature of data in publication. However, we observed that 37.3% (28/75) of the class 3 variants found were reported in the literature as variants of uncertain significance. We also observed that most BA probands were from very small families, with only 6.9% (4/58) having a familial PPCM. This finding was similar to most clinical reports for adult-onset DCM from South Africa in that the patients were reported with idiopathic disease (15, 60). We also observed in the BA probands carrying novel class 3 variants that the commonly affected genes were *DMD, MYH7* and *ANKRD1*. These gene alterations are common causes of intercalated disc disruption and ventricular arrhythmias in DCM patients (260). Almost all the BA PPCM probands with

reported class 3 variants carried *TTN* and *DSP* genes missense variants. C-terminal *DSP* variants are often associated with LV-dominant disease (120), while *TTN* missense variants are rare causes of DCM. There might be unknown mutation loci in these five genes (*TTN*, *SCN5A*, *MYH7*, *FLNC*, and *ALPK3*), which may be identified using polygenic risk scoring in another study. We think so because a recent study demonstrated that variant load differences increase with DCM severity in patients carrying class 3 variants (187), although not yet known in Black Africans.

3.4.2.2.2.2 PPCM probands with class 3 variants of Mixed ancestry (MA)

The targeted NGS data for this population group revealed that 26 MA probands carried only class 3 variants. There were 43 class 3 variants harboured in 13 genes: 67.4% (29/43) were novel, and 32.6% (14/43) were reported variants presented in Figure 4.28. Like the BA PPCM probands, many of these MA probands were from very small families; however, none had a positive FH of HF. In the MA probands carrying novel class 3 variants, we also observed that the commonly affected genes were *TTN, SCN5A*, and *LAMP2*. In contrast, almost all the reported class 3 variants were the *TTN* genes' missense variants. Similar to the BA probands, we think there might be unknown mutation loci in these three genes (*TTN*, *SCN5A*, and *LAMP2*) carried by the MA probands.

3.4.3 Paediatric DCM findings

3.4.3.1 Baseline findings for the paediatric DCM cohort

We recruited 89 paediatric probands clinically diagnosed with DCM and determined their genetic characteristics. The paediatric DCM probands had a mean age of 3.1 years (SD 3.5, range <12) at diagnosis. The paediatrics' self-reported ancestry revealed that 70% had BA, 28% MA, and 2% had WA. Table 3.10 showed that at enrolment, most of the probands, 55% (49/89), were recorded as female sex.

We used a comprehensive routine targeted NGS panel testing and identified disease-causing variants in two paediatric DCM probands. The genotype-positive paediatric probands had yielded a positive rate of 2.2%, and 97.8% (87/89) of the paediatric probands had an unknown disease-causing variant. The positive rate of 2.2% is very low compared to what has recently been reported by similar studies. Globally, the paediatric DCM yields have ranged from 19% to slightly <45% (96, 132).

The two genotype-positive paediatric probands consisted of one female and one male. In this study, more females than males were clinically diagnosed at enrolment, contrasting with similar studies where more males than females are affected (96). The male proband was a 3-year-old clinically diagnosed with an enlarged left ventricular chamber. In contrast, the female proband was a one-year-old diagnosed with unspecified left ventricular non-compaction. Non-compaction is a typical feature of paediatric DCM (261). The following section discusses the genetic findings of our two paediatric DCM probands.

3.4.3.2 Targeted sequencing findings for paediatric DCM cohort

3.4.3.2.1 Class 4 and 5 variant carriers' findings

Our paediatric probands' yield is very low, and the two variants we have reported are not entirely convincing. We therefore hypothesise that the *MYBPC3* c.3409T>G class 4 variant in the male proband, and the female proband with a *DMD* c.1149+1G>T class 5 variant might be due to syndromic CMO (96, 262). Literature has shown that paediatric CMO has diverse causes, including infections, metabolic diseases, syndromic and non-syndromic familial, and idiopathic CMO. Furthermore, published paediatric genetic testing results vary widely depending on the phenotype, age of presentation, and the type of testing conducted.

Likewise, our results have differed from the majority of findings in the literature. For example, the *MYBPC3* c.3409T>G variants in the MyBP-C functional protein domain, fibronectin-type III, is encoded by exon 31, and class 4 and 5 variants in this domain are associated with an early onset of DCM (263). *MYBPC3*-related paediatric DCM cases have a high risk of major adverse cardiac events with more severe features (96, 132). However, *MYBPC3*-related paediatric DCM is unknown for the African population (136).

Our paediatric cohort's low yield may also be linked to novel or recently discovered recessive genes (e.g., *JPH2* gene LoF in DCM) not covered by the panel that could explain some of the negative cases (43, 264). The IMHOTEP DCM paediatric yield of 2% is relatively low compared to 29% in Australia (132), 32% in Finland (43), and 19% to 35% in the USA (42, 96) studies that have used larger panels. In addition, variants in some established DCM-causing genes lacked evidence of pathogenicity. For example, five class 3 variants identified in *MYH7*, *SCN5A* and *DSP* genes met borderline criteria for being classified under class 4. The genes found (Table 3.12) are highly associated with paediatric DCM (96), and they encode for the sarcomere (*MYH7*), (cell-junction (*SCN5A*) and other cytoskeletal structure (*DSP*) proteins. Furthermore, none of the five variants was considered a possible founder mutation. We, therefore, recommended these variants for further analysis, including screening in our in-house adults' control or the probands' relatives.

3.4.3.2.2 Class 3 variants findings

About 64% (57/89) of the paediatric DCM probands had only class 3 variant (s). Mostly females, 56% (32/57), and BA, 68% (39/57) probands carried the class 3 variants. The demographic differences observed in this study have been reported in other studies with similar designs (96). However, the high yield of class 3 variants may be because we have used an established Eurocentric adult DCM panel. Perhaps it may be because the genetic determinants of paediatrics are mainly unknown (262). We identified 76 class 3 variants harboured in 24 genes; however, none of the variants was considered of interest because of the current ACMG classification rules applied. Thus, we recommend that some of the variants be reclassified in the future.

3.4.4 Possible founder variants in the DCM cohort

A founder variant is a disease-causing variant carried by several patients in a specific population due to the pathogenic mutation in their common ancestor. To our knowledge, no DCM-causing founder variant has been reported in Africa; however, very few DCM founder variants have been reported globally (196). Interestingly, we have reported seven (n=7) possible founder class 4 and class 5 variants in this study: four in the *TTN* gene and one each in *LMNA*, *BAG3*, and *FLNC* genes.

3.4.4.1 Possible class 4 and 5 founder variants in the adult-onset DCM cohort

The class 4 and 5 variants with the highest prevalence included the *TTN* c.95008C>T (18.4%), *LMNA* c.568C>T (5.3%), *BAG3* c.925C>T (2.6%), *FLNC* c.6907C>T (2.6%), *TTN* c.50170C>T (2.6%), and *TTN* c.9702-2A>G (2.6%). The other possible DCM-related founder variants were *TTN* c.9702-2A>G identified in BA, *FLNC* c.6907C>T in MA, and *TTN* c. 50170C>T in other MA probands, and these probands were recruited at our Cape Town site, Figure 4.33. Citizens of Cape Town and, to an extent, the Western Cape province of South Africa have lived a distinct lifestyle rich in diversity, heritage, and cultures. Finally, the *LMNA* c.568C>T and *BAG3* c.925C>T variants carried by the probands of MA have been previously reported separately in White (Caucasian) cohorts. We understand that Whites from Europe have shared ancestry with South African MA citizens, and we have suggested our team invite the authors of the European studies to collaborate in another research to trace the possible common ancestor. Furthermore, five WA and one MA probands were identified carrying a class 4 *TTN* c.87624C>A stop-gained variant. Hence, we recommend further investigating *TTN* c.87624C>A as a possible founder variant.

We, therefore, think that these seven variants are possible founder mutations and recommend that a haplotype analysis and genealogical tracing for a common ancestor should be performed to ascertain the status of each of these variants.

3.4.5 Digenic and compound heterozygous pathogenic variants

Digenic inheritance describes an individual carrying two Class 4 and 5 variants in different genes, whereas compound heterozygous explains when the two are in the same gene. Digenic (139) and compound heterozygous (265) inheritance of two or more heterozygous variants

have been well-characterised in DCM probands. Most reported probands have had a spectrum of DCM featuring early onset and severe phenotypes (266). Three probands carrying two Class 4 and 5 variants each were identified in this study. The Class 4 and 5 variants were found in two and one proband (s) carried as digenic and compound heterozygous variants, respectively.

3.4.5.1 The digenic variants: *TTN* c.95008C>T and *SCN5A* c.4012C>A, and the *TTN* c.89839C>T and *MYBPC3* c.2176C>T

We found digenic heterozygous variants in two adult-onset DCM probands. The first proband carried *TTN* c.95008C>T reported stop-gained class 5 variant and *SCN5A* c.4012C>A novel missense class 4 variant. The proband is a female BA patient diagnosed at age 64 years with an unknown cause of DCM. We reported the *TTN* c.95008C>T variant in multiple BA probands in this study, and we think that the novel *SCN5A* c.4012C>A might have exacerbated this proband's severe late onset of DCM because the novel variant affects exon 23 that has been reported in patients with cardiac arrhythmias (248).

The second proband was a male MA participant diagnosed at 27 years of age with familial DCM. He carried two report class 4 variants, the *TTN* c.89839C>T and *MYBPC3* c.2176C>T mutations. Both variants were found online in the ClinVar database; however, there is no established literature to demonstrate the effects of these variants. We think both variants' compounded effects have triggered a severe phenotype in our proband.

Digenic inheritance of variants, with one of the variants being in the *TTN* gene (A-band), is reported in 2.3% of DCM cases (139), and the authors demonstrated with a digenic mouse model that the absence of the *TTN*tv variant unaffected sarcomere insufficiency. Further, the clinical data in the study revealed that patients with multiple variants have the same disease severity as those with one variant each. However, other findings reported that the severity of *TTN*tv-related disease with digenic inheritance is worsened by the second gene's variant, amongst other factors (23, 267, 268). Similarly, our digenic probands with *TTN*tv-related disease had severe DCM; for example, the 27-year-old male was diagnosed with an NYHA functional class III phenotype and had an LVEF of 12%.

3.4.5.2 The compound heterozygous variants: *MYH7* c.5792del and c.5791-2A>T

One of our adult DCM probands carried *MYH7* (c.5792del and c.5791-2A>T) compound heterozygous variants which we could not determine if there are in trans. The affected amino acid residue has been reported recently in a CMO proband (210), and similarly, probands with severe DCM that carried *MYH7* compound heterozygous variants have been described (171, 269). Besides, 4% to 20% of DCM cases in the literature are attributed to *MYH7* heterozygous variants (98). The *MYH7* c.5791-2A>T splice acceptor and c.5792del (p.Gly1931AlafsTer2) frameshift variants are on the intron 39 and exon 40 boundary. This coding region of the *MYH7* gene is evolutionarily conserved in humans and other vertebrates, and molecular tools predict both variants to disrupt the canonical splice acceptor site of five amino acids proximal to the gene's natural stop codon. However, the resultant RNA is unexpected to undergo NMD. Functional studies contradicting these variants' role as an aetiology of familial CMO are currently unavailable. So far, much is unknown about *MYH7* gene LoF pathomechanisms in DCM (149). The current evidence suggests that these two variants can exacerbate DCM via a dominant negative mechanism.

3.4.6 Polygenic risk scores

Polygenic risk score (PRS) is an *in-silico* tool used to estimate an individual's genetic risk of a trait or disease. The PRS for DCM genetic liability has been calculated recently, and the authors used 36,000 (98.2% European, 0.8% South Asian, 0.3% African and 0.3% East Asian diaspora) participants for the scoring (270). The authors used the PRS to identify 28 loci associated with cardiovascular phenotype disease-causing genes, including *BAG3*, *FLNC* and *TTN*. Knowing that *TTN*tv-related DCM is common, the authors further investigated the influence of the PRS in 59 *TTN*tv carriers (120). They found that the *TTN*tv-related DCM carriers had a 7.2-mL increase in LVESV and a 2.6% reduction in LVEF for the PRS. LVEF is a valuable parameter for the clinical characterisation of DCM. Notably, we reported that young probands with *TTN*tv-related DCM had very low LVEF percentages. Reduced LVEF has been detected in *TTN*tv-related DCM carriers using PRS in European cohorts. However, it is important to note that the PRS for DCM was calculated in a cohort consisting of 0.3% African participants (270).

PRS analysis can identify the genetic basis of rare traits, like DCM, and its accuracy depends on population genomic data (271). Therefore, the current PRS for DCM could be inaccurate in our population because African data was unavailable when the scores were computed. The

limited use of PRS in Africa is attributed to the lack of enough GWAS studies on traits associated with them (271). Few genetic studies have emerged from the African continent compared to the high-income countries (37). For example, only one study has used PRS to identify severe ultrarare mutations in genes that are critical to schizophrenia in the South African Xhosa population (117). Also, in a reviewed study, the authors identified only five articles for PRS in the SSA region five, and there was no record of a GWAS for the DCM (271).

The lack of data from the African continent (37), in turn, raised some debate, especially concerning the use of Eurocentric genetic findings such as PRS for citizens on the continent to assign risk to alleles (251). Specifically, because African populations possess the richest human genetic diversity, they are responsible for complex traits in all populations (272). For example, African heterozygosity at coding sequence 31.7 Mb of the human genome, calculated in 10-kb intervals, was approximately 1.25 proportion greater than among non-Africans (117). Therefore, we recommend that future IMHOTEP genetics studies include the determination of PRS for the various forms of CMO.

3.5 **CONCLUSION**

In this study, 16.9% of adult and 2.2% of paediatric probands had DCM-causing variants, and *TTN* truncating variants were the most frequent. The yield information is relevant in changing the management plan for these patients. Although our yield is low, there are a few possible explanations. Firstly, we used the standards set for practice at the OMGL to diagnose the inherited forms of CMO routinely using targeted NGS. Secondly, maybe because we recruited participants from public tertiary hospitals, patients suspected of having heart disease are referred to tertiary hospitals. Finally, it may be a recruitment bias, as well-to-do patients are often seen first at private hospitals. Otherwise, the higher yield in literature implies the authors have used the ACMG/Clin-Gen variants investigation guidelines differently. It is precisely due to these differences in the variant investigation that, over the past five years, periodic re-evaluation of the variant classification guidelines has been recommended (87). However, the adult yield per the probands' ancestry shows that the WA had a 32.5% positive rate, the MA probands yielded 16.1%, and the BA probands had a 15.0% class 4 and 5 variants yield. Our findings showed that the panel we used is less effective for Black South African DCM patients who are commonly carrying *TTN* variants.

Our genotype-positive DCM probands are diagnosed at a very young age compared to the data in the literature. We also reported that many of the DCM probands were of BA and MA. When we look at their self-reported sex distribution, we found that most BA probands were females, whereas most MA and WA probands were males. This may be due to changes in sex hormones such as testosterone in males (137) and oestrogens in pregnant females (163), known risk factors for adult-onset DCM. We hypothesised that being a young Black South African diagnosed with DCM disease might have an underlying genetic determinant. We also suggest that the young female probands diagnosed during peripartum as having DCM might have had a distinct idiopathic clinical entity.

The comprehensive panel testing revealed 14 genes, but four were predominant DCM-causing genes in our cohort. The four genes encode the sarcomere (i.e., *TTN* gene), z-disk (i.e., *BAG3* gene), nuclear envelope (i.e., *LMNA* gene), and cytoskeleton (i.e., *FLNC* gene) proteins. *TTN* is this DCM study's most prevalent sarcomere gene, with a carrier rate of 68.4%. However, *TTN* variants were more common in the BA probands compared to their MA and WA counterparts. The *TTN* mutations were truncating variants identified mainly in the A-band region of the gene. We also identified five *TTN* recurring variants, and four of the variants were carried by BA probands. Our findings supported the findings in the literature that the A-band

region of the sarcomere protein might be a mutation hotspot in our South African DCM cohort. Three of our adult probands carried multiple class 4 or 5 variants, and the BA probands had a very high recurring variants rate.

This study identified ten recurring DCM-causing variants; for example, the *TTN* c.95008C>T occurred in 14 BA probands, *BAG3* c.925C>T in two MA probands, *LMNA* c.568C>T in four MA probands and *TTN* c.87624C>A in one MA and five WA probands. Besides, we screened all available relatives of the genotype-positive probands and informed the genetic counsellor of our findings. Our clinical team is helping the affected or "at-risk" relatives to consider lifestyle adjustments and further motivate more individuals to enrol on the organ transplant list. For example, two carriers of the deadly *LMNA* c.568C>T variant have enrolled in the heart transplant list during this study. Notably, unlike other patients, our *LMNA* c.568C>T variant carriers were below 45 years at diagnosis. In addition to the early onset, their severe phenotype possibly has been exacerbated by their socioeconomic status. Indeed, findings in the literature have shown that low economic status is strongly associated with HF in LMICs.

Finally, this study comprised 35% familial and 65% non-familial genotype-positive probands, and 60% of the *TTN*tv carriers were non-familial. Most probands of WA had familial DCM, while many of the BA and MA probands had non-familial DCM. Notably, the MA probands had a heterogeneous genotype-positive result as they only shared some genes with BA or WA probands. Also, most MA probands carried mainly genes associated with high heart transplantation rates, such as *LMNA*. However, we have informed the clinical team about our research results. They have reached out to the probands who tested positive, as well as their relatives, to provide them with genetic counselling and to initiate case and contact tracing. Additionally, we created specific awareness programs about heritable DCM for local communities where we found the highest number of genotype-positive probands. In some cases, we also identified possible founder variants.

Chapter 4: Hypertrophic Cardiomyopathy (HCM)

4.1 **INTRODUCTION**

4.1.1 Clinical background

Hypertrophic cardiomyopathy (HCM) is a common monogenic disease inherited in an autosomal dominant pattern (273). The primary feature of HCM is the thickening of the heart muscle without any other underlying conditions such as hypertension, valvular heart disease, or amyloidosis (Figure 1.3). On histological specimens, a typical HCM heart muscle would show disorganised myocardial architecture, with the myocytes generally being in disarray due to the loss of normal parallel alignment. Typically, widely distributed disarray involves more than 10% of the myocardium. Autopsy findings have revealed extensive myocyte disarray with increased interstitial fibrosis in HCM-related SCD patients.

4.1.2 Signs and symptoms

The most common symptom of HCM is dyspnoea, which can be triggered by physical activity (274). Other symptoms may include chest pain, fainting, stroke symptoms, or sudden death (273). The HCM clinical diagnosis shown in Figure 1.9 section 1.1.1.11 has involved the use of *i*) 12-lead ECG to detect left axis deviation, repolarisation changes, abnormal Q-waves and deep T-wave inversions used for screening, *ii*) echocardiography to reveal LV wall thickness > 15mm, and *iii*) diastolic dysfunction on Doppler imaging.

4.1.3 HCM disease burden

HCM can occur at any age; however, it is most commonly diagnosed in people before their mid-30s (133). The incidence of HCM is between 1 in 200 and 1 in 500 people in high-income countries (274, 275), with an annual mortality rate of approximately 0.5% and 5-year mortality of 2.5% (276). HCM in childhood is estimated to have an incidence rate ranging from 0.002% to 0.005%, with a prevalence of 0.029% and adult prevalence of 0.2% (51).

In certain populations in the SSA region, HCM affects around 4% of the paediatric and nearly 50% of the adults presenting with symptoms of the disease (14). The prevalence of HCM in

the South African adult population is around 49.9%, and there is genetic heterogeneity (156). HCM is a condition that can affect individuals of all ancestry and can vary in severity (156, 277, 278). While some patients may not experience any symptoms, others may require a left ventricular assist device implantation or a heart transplant as soon as they are diagnosed (277, 278). HCM can be especially dangerous for young athletes, as it may result in HF or SCD (279).

4.1.4 Aetiology of HCM

HCM is a CVD with an established genetic basis often inherited in an AD pattern and with variable penetrance (280, 281). Familial HCM accounts for 50% of cases, while the rest are assumed to be spontaneous mutations (282). Most HCM genetic aetiologies have been linked to monogenic phenotypes attributed mainly to sarcomere gene variations (16).

4.1.4.1 Monogenic HCM

The "single gene hypothesis" for HCM suggests that a single genetic variant in an HCM-causing gene accounts for all aspects of the observed phenotype (283). So far, researchers have described over 92%-99% of monogenic class 4 or 5 variants in eight robustly validated sarcomeric protein-coding genes known for causing HCM (284, 285). Of these eight genes, the *MYBPC3* and *MYH7* carry about 30-40% and 20-30% of class 4 and 5 variants considered monogenic, respectively (66, 156). Notably, some of the *MYBPC3* and *MYH7* monogenic class 4 or 5 variants have also been described as founder variants, and the literature findings have shown that HCM is a complex phenotype for one variant to define every phenotypic aspect of the disease. For example, the *MYBPC3* Ala797Thr South African subpopulation founder variant accounts for ≈25% of cases (67), and the *MYBPC3* Gln1061Ter and *MYH7* Arg1053Gln founder variants account for about 24% of HCM in Finland (286) have caused variable phenotypes. Suggesting that these variants could activate other pathways leading to increased contractility with HCM risk and exacerbating patient phenotypes differently (182). Generally, HCM patients with founder variants have had poorer lifelong outcomes, such as in the Japanese population, where some sarcomere founder variant carriers had more prominent arrhythmogenic events (287). Other founder variants (in *TNNI3* and *TNNT2* genes), such as in the Singaporean HCM patients of predominantly Chinese ancestry, have displayed reduced

penetrance with variable disease onset (285). Though these HCM disease-causing variant carriers have had mainly single sarcomere gene variation, under 1% of the patients have carried two or more variants and, to a lesser extent, with monogenic class 4 and 5 variants in non-sarcomere genes (288, 289).

Although non-sarcomere genes analysis has yielded little to no evidence for causing heritable HCM (284), those with plausibly robust aetiological evidence of HCM have included *TNNC1*, *PLN*, *FLNC*, *FHL1*, and *CSRP3* (113, 284). Class 4 and 5 variants in these non-sarcomeric genes, such as the CSRP3, have often needed more focus on functional work to demonstrate the observed phenotype (290). That could be because some patients carrying the non-sarcomeric variants of interest have presented with HCM at a relatively older age (291). However, it is salient to distinguish subtypes of HCM, such as an X-linked inheritance, from phenocopy conditions, for example, the Anderson–Fabry disease (23). The distinction is significant because phenocopy conditions have occurred in syndromic conditions such as the Noonan syndrome and storage diseases such as the Anderson–Fabry disease that has presented similar phenotype-genotype correlation with late-onset HCM (291).

4.1.4.2 Left ventricular HCM

Left ventricular (LV) hypertrophy has characterised the diagnosis of HCM, specifically, adults with LV wall thickness ≥13 mm on echocardiogram or CMR and children with LV wall thickness more than two standard deviations greater than the predicted mean (z-score >2) (292). LV hypertrophy in HCM patients has been unexplained solely by loading conditions, and other factors, such as genetic variations, are known to have contributed significantly to the development of some HCM phenotypes. Notably, the Human Gene Mutation Database has linked variants in over 51 genes to various HCM phenotypes (113), and researchers have associated 16 with complex LV hypertrophy amongst these genes. Remarkably, class 4 and 5 variants in *MYBPC3* and *MHY7* genes have caused only LV-HCM in about 70% of the patients, and those with complex LV hypertrophy in HCM, e.g., Noonan syndrome and Anderson–Fabry disease have carried *LAMP2*, *GAA*, *GLA*, *PRKAG2,* and *TTR* variants (277). Other disease-causing variants in genes such as *TNNI3* have caused LV mass and wall thickness with HCM onset at a younger age, and patients have had a higher risk of SCD (293). Furthermore, individuals of BA in South Africa diagnosed with LV-HCM are often younger,

less likely to carry sarcomere variants and have had worse symptoms compared to patients of WA (30).

4.1.4.3 Dominant and Recessive HCM

Heritable HCM has occurred predominantly as an AD disorder commonly caused by non-truncating variants (282). However, AR modes of inheritance have been described (294), especially in patients carrying *TRIM63 (*or *ALPK3*) gene variants (295). In Africa, paediatric HCM researchers from Egypt have reported AR as the most common mode of inheritance (296). Another study of infantile HCM has highlighted variant-specific molecular differences found in *MYL2*-associated CMO with AR mode of inheritance (282). These studies' findings suggest that the AR inheritance pattern could be more frequent in HCM and might have been exacerbated by possible modifiers.

Many HCM patients lack clear genetic causes due to the complicated genotype-phenotype correlation, making it challenging to define pathogenicity for class 3 variants (297). However, the different genetic causes of HCM can lead to either obstructive or non-obstructive phenotypes (273).

4.1.5 Types of HCM

There are two types of HCM: obstructive, accounting for 70% of cases, and non-obstructive (298). The majority of patients with non-obstructive (thickened left ventricle) disease may remain asymptomatic or develop only minor symptoms, maintaining average life expectancy and a benign clinical course (133). Symptoms and risks are more common in obstructive (thickened septum) HCM compared to non-obstructive HCM; meanwhile, apical HCM is linked to a higher risk of atrial fibrillation and stroke than non-apical HCM (297). Approximately 25% of patients with HCM have had an obstruction at rest (299) and are more prone to syncope in response to changes in preload and afterload caused by hypovolemia, severe anaemia, use of vasodilators, and diuretics (273). Also, about 10% to 20% of patients with obstructive HCM have a lifetime risk of SCD due to ventricular arrhythmia (299). Furthermore, patients with familial HCM who carry variants that destabilise myosin conformations are at higher risk of HF and arrhythmias (213).

4.1.6 Genetics of HCM

HCM is caused by multiple mutations, mainly found in the sarcomeric proteins (300). The disease is marked by myocardial fibre disarray caused by variants primarily found in eight definitive sarcomere genes (300). Of these genes, the *MYBPC3* and *MYH7* variants account for 70% of patients diagnosed with familial HCM (16). The most relevant HCM disease-causing genes are summarised in Table 4.1.

Table 4.1: Prevalence of published adult HCM-related prevalent genes

GENE SYMBOL	FUNCTION	PREVALENCE (%)	REF: PMID
MYBPC3	Sarcomere	20 - 66.7	35653365
MYH7	Sarcomere	15 - 25	33673806
TNNT2	Sarcomere	5 - 20	36264615
TPM1	Sarcomere	Rare - 8	33673806
TNNI3	Sarcomere	Rare - 5.9	35653365
FLNC	Actin crosslinking	0.5 - 2	28356264
GLA	Lysosome	Rare - 2	33673806
MYL2	Sarcomere	Rare - 2	33673806
ACTC1	Sarcomere	Rare - 1	33673806
MYL3	Sarcomere	Rare - 1	33673806

4.1.7 HCM genotype-phenotype correlation

The genes associated with HCM are clustered across the cardiomyocytes within the sarcomere, sarcoplasmic reticulum, nucleus, and other cytoplasmic structures (264, 300). The cellular locations and functions of proteins encoded by HCM-associated genes are summarised in Figure 4.1 (264, 284).

It has been observed that most variants in *MYBPC3*, which is the most common cause of genetic HCM, result in frameshift mutations that decrease the total quantity of functional protein (90). The pathomechanism of most HCM diseases caused by *MYBPC3* variants is reported as haploinsufficiency caused by a LoF mechanism (301). On the other hand, most variants in *MYH7*, the second most common cause of HCM, are AD and primarily result in single amino acid changes. These changes are reported to cause gain-of-function via a dominant-negative polypeptide mechanism (90, 282). These pathomechanisms are being targeted for HCM therapeutics.

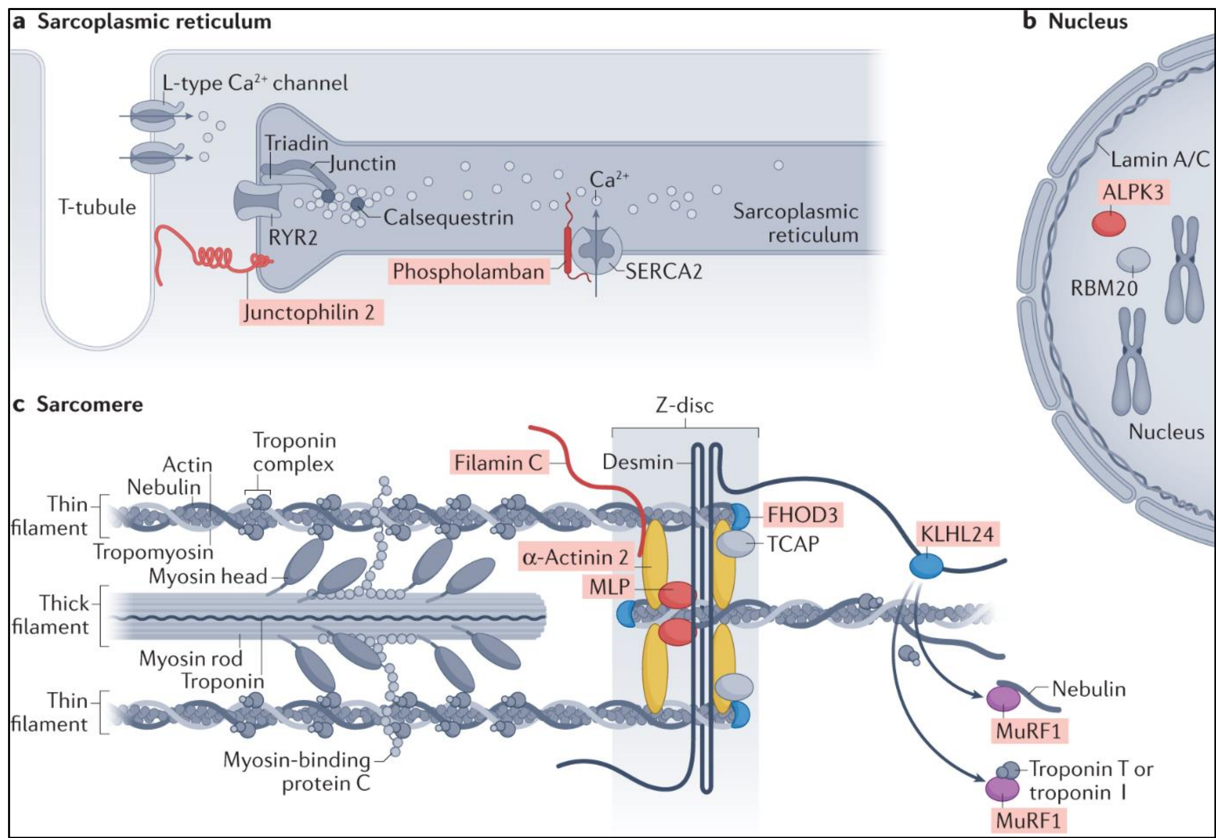


Figure 4.1: Hypertrophic cardiomyopathy proteins: cellular locations and functions. ***a***) *Proteins involved in* Ca^{2+} *regulation at the transverse tubule (T-tubule)–sarcoplasmic reticulum complex.* ***b***) *Nuclear proteins involved in transcriptional regulation.* ***c***) *Proteins of the sarcomere, including important proteins of the Z-disc and ubiquitin ligase complex. The non-sarcomeric proteins are highlighted in red. These include α-actinin 2 (encoded by ACTN2), α-protein kinase 3 (encoded by ALPK3), muscle LIM protein (MLP; encoded by CSRP3), FH1/FH2 domain-containing protein 3 (FHOD3; encoded by FHOD3), filamin C (encoded by FLNC), junctophilin 2 (encoded by JPH2), Kelch-like protein 24 (KLHL24; encoded by KLHL24), phospholamban (encoded by PLN), and muscle-specific RING finger protein 1 (MuRF1; encoded by TRIM63). Other proteins listed are RBM20, RNA-binding protein 20; RYR2, ryanodine receptor 2; SERCA2, sarcoplasmic/endoplasmic Ca2+ ATPase 2; and TCAP, telethonin (264).*

4.1.8 Therapeutics

Therapeutic strategies could differ if the disease results from haploinsufficiency caused by LoF or gain-of-function via a dominant-negative polypeptide mechanism (90). The general management principle for HCM involves using drug therapy (e.g., Beta-blockers, verapamil, diltiazem low-dose diuretic in patients with LVEF value ≥50%), mechanical circulatory support, or heart transplantation to manage the symptoms. In most HCM patients, the left ventricular outflow tract obstruction or SCD risk prediction is monitored for phenotype-specific management while on angiotensin-converting enzyme inhibitor or angiotensin receptor blocker (51, 174). For example, a recently discovered compound, mavacamten (MYK-461), has been found to act as an allosteric inhibitor of cardiac myosin adenosine triphosphatase (ATPase) at the sarcomere level. This results in a decreased hypercontractility

of cardiomyocytes and promotes cardiac muscle relaxation (133). The positive impact of mavacamten on HCM patients was initially demonstrated in the phase 2 PIONEER-HCM trial (302).

Gene therapy techniques such as genome editing, exon skipping, allele-specific silencing, spliceosome-mediated RNA trans-splicing, and gene replacement have been developed to remove genetic defects (303). Most of these methods have been tested for their effectiveness and efficiency in animal or human-induced pluripotent stem cell models of HCM with encouraging outcomes (303). The scientists examined the potential of exon skipping as a gene therapy for Mybpc3-targeted knock-in mice using antisense oligonucleotides (303, 304). The mutant mice had a homozygous G>A variant on the last nucleotide of *Mybpc3* exon 6 in an evolutionary conserved 5′ splice_donor site sequence of the gene. The alteration results in different aberrant mRNAs. By using specifically designed antisense oligonucleotides and adeno-associated virus serotype 9-mediated delivery to skip both exons 5 and 6 of Mybpc3, the researchers were able to reduce cardiac dysfunction and prevent the development of left ventricular hypertrophy in newborn mice (304, 305). Although the therapeutic effect was temporary, this proof-of-concept study is a significant breakthrough towards developing a causal therapy for HCM (304).

4.1.9 HCM Genetic Modifiers

Heritable HCM is a complex phenotype characterised by monogenic and oligogenic variants, genetic heterogeneity, and unclear phenotype overlaps, especially in patients with known modifiers in *MYH7, TTR* and noncoding regions that regulate gene expression (187, 306). Also, an expanded analysis could detect new possible clinically relevant modifiers genes. Notably, the pathogenicity status of many variants has changed, especially when functional evidence is later available. Hence, the variants were categorised as class 3 as they cannot completely explain the phenotype and we currently lack the evidence to have it reclassified as a class 4.

4.2 **METHODS**

The methods for this thesis are discussed in Chapter 2 and again briefly described in Section 4.2 for the HCM cohort.

4.2.1 Ethical compliance

Ethical clearance was obtained from the Human Research Ethics Committees (HREC) at the (FHS, (UCT (main study IMHOTEP registry, HREC 766/2014 and HREC 009/2020). All participants were informed before they provided a signed consent (and permission) form.

4.2.2 Study population

Sixty-seven (n = 67) unrelated participants clinically diagnosed with HCM were recruited from February 2015 to October 2021 at six sites in South Africa (Appendix J). At least one cardiologist examined the probands at the participating site. The records were again re-evaluated in Cape Town by a team of at least one cardiologist, a research fellow, and a clinic doctor. The clinical examination included baseline assessment, complete medical history, pedigree, ECG and transthoracic echocardiography. The clinical diagnostic criteria for HCM were unexplained LVH in any myocardial segment defined as a wall thickness of ≥15mm in adults, or ≥13mm with a FH of HCM, and disproportionate thickening of the interventricular septum compared to the posterior wall (51). The inclusion and exclusion criteria are described in section 2.1.2 of chapter two.

4.2.3 DNA Extraction and quality control

All our genomic DNAs were extracted from the probands' peripheral whole-blood or buffy leukocytes following the published manufacturer's methods for Puregene Blood Core Kit C (Appendix C) or PAXgene Blood DNA kit (Appendix D), respectively. We assessed the integrity, quality, and quantity of the gDNA samples according to the SOPs in Appendices E and I. We measured and computed the purity and concentrations with a calibrated NanoDrop™ 2000 spectrophotometer (Thermo SCIENTIFIC) operating with software v2.6 on a Windows 10 computer. At the same time, the samples' intactness was checked with a 1% agarose gel.

The gel electrophoresis was programmed to run at 110V for 75 minutes. The gel was visualised in a Uvitec Xplorer D55 gel doc running on an Xplorer 1D software version 15.08 (Uvitec Cambridge, Cambridge, UK).

4.2.4 Targeted Sequencing and Variant Calling

A custom-designed list of 38 CMO disease-causing genes was created by the (OMGL in the UK. The custom designed chip contained 38 specific genes including: *ACTC1*, *ACTN2*, *CSRP3*, *GLA*, *LAMP2*, *MYBPC3*, *MYH7*, *MYL2*, *MYL3*, *PRKAG2*, *TNNI3*, *PLN*, *TNNC1*, *TNNT2*, *TPM1*, *TTR*, and other HCM-related genes. The complete genes list is attached as Appendix F.

All gDNA samples underwent targeted resequencing, initially performed using the Haloplex kit but subsequently replaced with the Twist NGS kit in September 2021. The data generated were mapped to the human reference genome (GRCh37/hg19) and transformed into variant call format (VCF) using the Genome Analysis Toolkit (GATK) v3.8-1-0 and Platypus v 0.8.1. We then annotated the VCF files with Ensembl Variants Effect Predictor (VEP) and analysed them.

Through a specific set of filtering criteria, we selected variants for further analysis. The variants were classified according to the refined ACMG guidelines criteria as class 1 (B), class 2 (LB), class 3 (VUS), class 4 (LP) and class 5 (P). We selected variants (1) with a gnomAD minor allele frequency (MAF) of <1% (2), variants that were predicted by bioinformatics tools (i.e., SIFT and PolyPhen2) to have damaging deleterious or unknown protein function. Alamut™ Visual Plus version 1.5.1 and ClinVar (2022-12-17) were used to analyse the variants further. We also looked at the variant caller scores for Haplocaller and Platypus for threshold values above 400 and 700, respectively, as variants below these thresholds are artefacts or polymorphisms. We also visualised the BAM files for probands carrying class 4 or 5 variants in the Alamut, as this could identify possible false positive variants. All identified variants of interest were validated using Sanger sequencing.

4.2.5 Biostatistical analysis:

The data was analysed using R and RStudio (version 4.1.2); results were presented as $\bar{x} = \pm 1\delta$ for numerical variables and total counts with percentages for categorical variables.

4.3 **RESULTS**

4.3.1 Baseline characteristics for the adult-onset HCM cohort

Sixty-seven (n=67) unrelated probands clinically diagnosed with HCM were enrolled in this study and comprised 89.6% (60/67) adult-onset and 10.4% (7/67) paediatrics HCM (Table 4.2). The panel results for the adult (≥ 13 years) and paediatric (< 13 years) are presented separately.

Table 4.2: Baseline results for the HCM cohort, including adults and paediatric probands

LABEL	LEVELS	TOTAL IMHOTEP	TOTAL HCM ADULT-ONSET	TOTAL HCM PAEDIATRICS
NUMBER OF PROBANDS WITH GENETICS DATA ANALYSED, N (%)		690	60 (8.7)	7 (1.0%)
AGE AT DIAGNOSIS (YEARS)	Mean (SD)	31.7 (16.4)	41.3 (13.9)	3.6 (4.0)
LEFT VENTRICULAR EJECTION FRACTION (%)	Mean (SD)	36.1 (19.6)	70.7 (8.7)	-
NEW YORK HEART ASSOCIATION	Class I	83 (15.7)	24 (42.9)	-
	Class II	162 (30.6)	24 (42.9)	-
	ND	10 (1.9)	0 (0.0)	-
	Class III	223 (42.1)	7 (12.5)	-
	Class IV	52 (9.8)	1 (1.8)	-
FAMILIAL	No	405 (58.7)	35 (58.3)	-
	Yes	125 (18.1)	21 (35.0)	-
	No data	160 (23.2)	4 (6.7)	-
SEX ASSIGNED AT BIRTH	Female	364 (52.8)	21 (35.0)	0 (0.0)
	Male	326 (47.2)	39 (65.0)	7 (100.0)
ANCESTRY (	Black-African	381 (55.2)	15 (25.0)	6 (85.7)
	Mixed	221 (32.0)	34 (56.7)	0 (0.0)
	White	83 (12.0)	11 (18.3)	1 (14.3)
	Indian	5 (0.7)	0 (0.0)	0 (0.0)

VALUES ARE MEAN ± 1SD OR N (%) BASED ON NON-MISSING VALUES OF THE TOTAL ANALYSED.

Table 4.2 presents the baseline clinical data, which indicates that the average age was 41.3 years (SD = 13.9 and range = 16-75 years) at the time of diagnosis. Around 65% of the recorded probands were male. Moreover, over 85% of the probands were classified as NYHA functional class I and II. As shown in Figure 4.2 A, their self-reported ancestry found 56.7% Mixed, 25.0% Black-African and 18.3% White probands. The FH results in Figure 4.2 B indicate that 58% (35/60) of HCM probands had non-familial HCM, 35% (21/60) had familial HCM, and no data was available for 7% (4/60) of the probands' HF medical history.

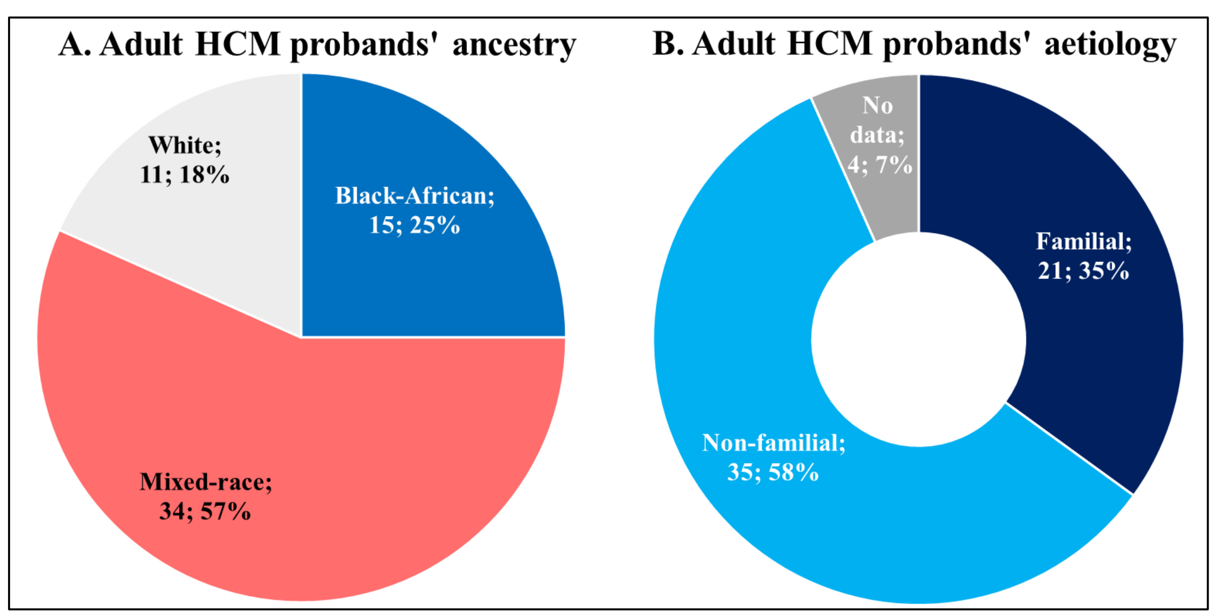


Figure 4.2: Adult HCM probands family history.

4.3.2 Baseline characteristics for the adult-onset HCM population groups

As South Africa has such a diverse population, we took the opportunity to investigate any potential differences in age, sex assigned at birth, and diagnostic classification across the three dominant populations of South Africa: Black-Africa, White and Mixed ancestry. We observed distinct differences in sex and aetiology across the three population groups (Figure 4.3 A-D). However, we used caution in the interpretation as some population groups were in small numbers.

When we analysed the representative populations within the HCM cohort, we found that the population group with the most significant number of probands was the MA group, where 34 probands were enrolled (Figure 4.3 C). These patients were slightly older than the other two groups (Black African and White ancestries) with a mean age of 44.9, SD 13.4 years (Figure 4.3 A), and 62% (21/34) were male (Figure 4.3 B).

In comparison, the 15 BA HCM probands recruited had a mean age of 37.6 (SD=12.3) years, with more males at enrolment, with 73% (11/15). Also, 40.0% (6/15) of the BA probands had familial HCM aetiologies (Figure 4.3 D).

We also recruited 11 WA probands, and the data showed that they had a mean age of 35.3 (SD 14.8) years, with 64% (7/11) assigned male sex at birth (Figure 2C). Only 36% (4/44) of the probands had familial aetiologies, while almost two-thirds had non-familial HCM (Figure 4.3 D).

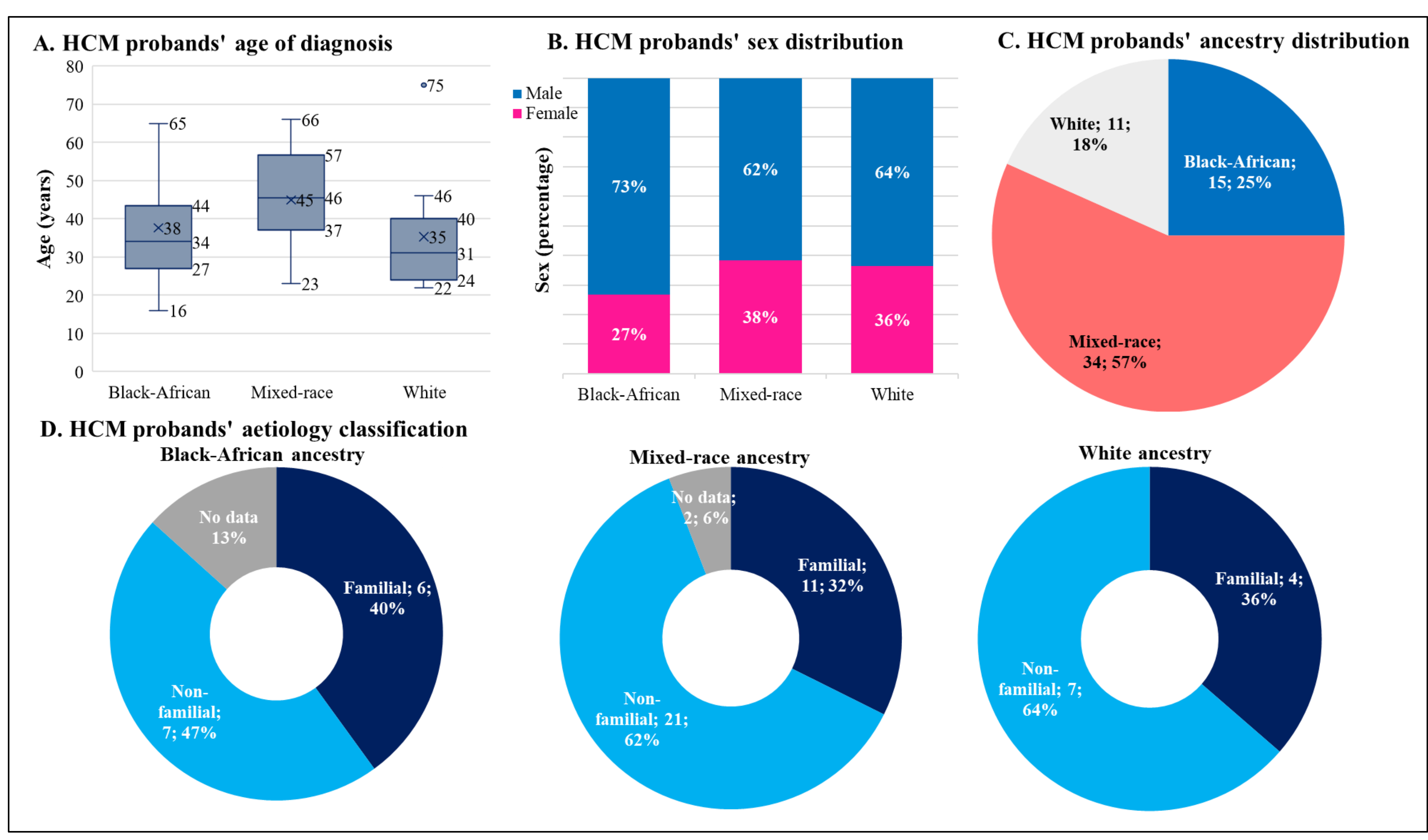


Figure 4.3: Adult HCM probands' age, sex assigned at birth and aetiology distribution. Graphs showing comparisons between adult-onset HCM population groups in terms of age, sex assigned at birth and diagnostic classification. *A. Population groups vs age of diagnosis in HCM; B. Population groups vs sex distribution; C. Population groups vs ancestry classification; D. Population groups vs aetiology/diagnostic classification.*

4.3.3 Targeted sequencing for adult-onset HCM probands

All variants reported in this HCM cohort were curated using the refined ACMG classification system 2018. Table 4.3 shows the baseline characteristics of ACMG classes 1-5 variants for the adult-onset HCM cohort. Whereas SS Table 4 shows the fitted logistic regression model for adult HCM probands. This chapter also delves into these variants under the various ACMG class headings. According to the ACMG guidelines, class 5 is P, class 4 is LP, class 3 is a VUS, class 2 is LB, and class 1 is B.

Table 4.3: Baseline results for the adult-onset HCM cohort according to ACMG classification

LABEL	LEVELS	TOTAL	POSITIVE	UNCERTAIN	NEGATIVE	P-VALUE
HCM diagnostic yield, n (%)		60	14 (23.3)	24 (40.0)	22 (36.7)	
AGE	Mean (SD)	41.3 (13.9)	33.9 (11.1)	41.5 (15.2)	45.9 (12.3)	0.036
LVEF	Mean (SD)	70.7 (8.7)	71.2 (7.1)	70.0 (10.7)	71.2 (7.2)	0.889
NYHA						0.642
	Class I	24 (42.9)	6 (46.2)	10 (43.5)	8 (40.0)	0.109
	Class II	24 (42.9)	4 (30.8)	10 (43.5)	10 (50.0)	0.552
	Class III	7 (12.5)	2 (15.4)	3 (13.0)	2 (10.0)	0.813
	Class IV	1 (1.8)	1 (7.7)	0 (0.0)	0 (0.0)	0.992
FAMILIAL*						0.644
	Yes	21 (35.0)	4 (28.6)	7 (29.2)	10 (45.5)	0.827
	No	35 (58.3)	9 (64.3)	16 (66.7)	10 (45.5)	0.154
	No data	4 (6.7)	1 (7.1)	1 (4.2)	2 (9.1)	0.930
SEX*						0.398
	Male	39 (65.0)	7 (50.0)	17 (70.8)	15 (68.2)	0.694
	Female	21 (35.0)	7 (50.0)	7 (29.2)	7 (31.8)	0.134
ANCESTRY*						0.049
	Mixed	34 (56.7)	8 (57.1)	9 (37.5)	17 (77.3)	0.058
	Black-African	15 (25.0)	2 (14.3)	10 (41.7)	3 (13.6)	0.032
	White	11 (18.3)	4 (28.6)	5 (20.8)	2 (9.1)	0.908

VALUES ARE MEAN ± 1SD OR N (%) BASED ON NON-MISSING VALUES OF THE TOTAL ANALYSED:60. POSITIVE-CLASS 4 & 5, UNCERTAIN-CLASS 3, AND NEGATIVE-CLASS 1&2: LVEF-LEFT VENTRICULAR EJECTION FRACTION, NYHA-NEW YORK HEART ASSOCIATION. ASTERISKS (*)-SELF-REPORTED.

Targeted sequencing found 23.3% (14/60) adult-onset HCM probands carrying class 4 and/or class 5 variants: 12 probands carrying class 5 and two probands carrying class 4 variants. Meanwhile, 40.4% (24/60) probands carried only class 3 variants, and 36.7% (22/60) probands carried class 1 and/or 2 variants. We grouped the variant classes for the result presentation and discussion: classes 4 and 5 had positive genetic results, class 3 had uncertain genetic results, and classes 1 and 2 were negative, as shown in Figure 4.4A. To understand the genetic architecture of South Africa’s diverse populations, we also looked at the HCM ACMG classification yields across the probands' ancestry, Figure 4.4B.

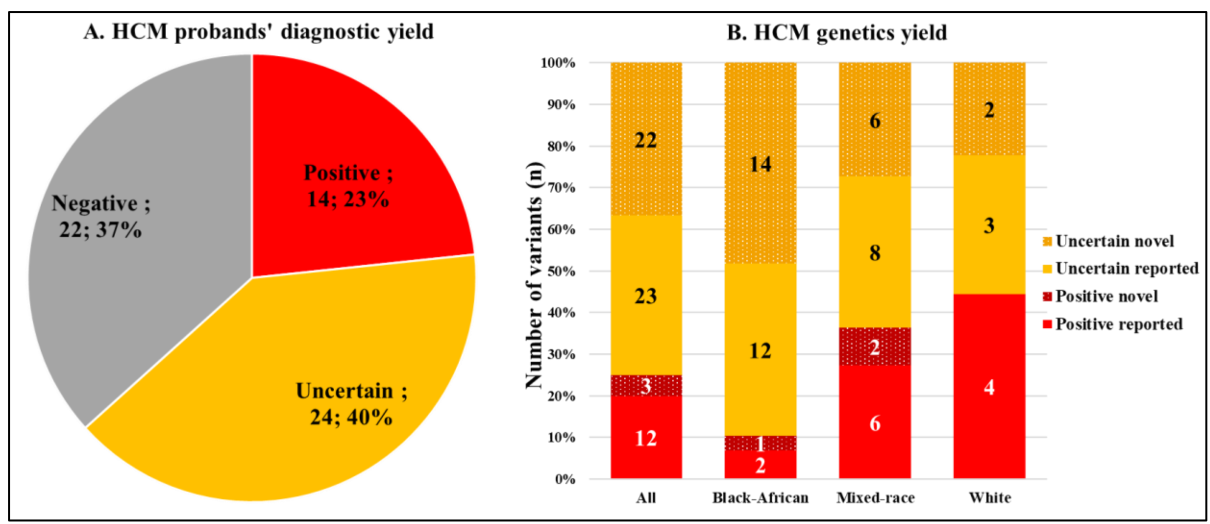


Figure 4.4: Adult HCM probands' diagnostic yield. ***A.*** *Pie chart indicating the diagnostic yield for the adult HCM cohort,* ***B****. Bar graph showing the count of probands with variants grouped as novel vs reported within the HCM cohort.*

Figure 4.4B presents the number of novel and reported variants we found in our HCM cohort. Figure 4.4B indicates the low positive yields across our HCM cohort and the large yield of class 3 variants, especially in the BA sub-population group. We classified 75% (45/60) of the mutations as variants of unknown significance. We first described the positive (pathogenic; class 4 and 5) variants and followed this up with a report on variants of uncertain significance (class 3) and their possible modifying effects on HCM. Class 1 and 2 (Negative) variants are beyond the scope of this thesis as we attempted to focus on the variants that have a higher likelihood of causing HCM; thus, we present and briefly discuss the 22 probands with negative findings.

4.3.3.1 Probands with pathogenic (class 4 and 5) variants

We analysed the positive probands for ACMG class 4 and 5 variants and found that 23.3% (14/60) carried HCM-causing variants: one of the probands carried two class 5 variants, as listed in Table 4.4.

Table 4.4: Class 4 and 5 variants (n=13) in the adult-onset HCM probands (n=14) spanning three population groups

PROBAND_ID	SEX	ANCESTRY	AGE	FAMILIAL	GENE	PROTEIN CHANGE	CDNA CHANGE	CONSEQUENCES	EVIDENCE	dbSNP_ID	ZYGOSITY	ACMG
101301575	Female	Black-African	24	Familial	*TNNI3*	p.Leu144His	c.431_432delinsAT	missense	Reported	None	Het	Class 5
					TNNI3	p.Leu144Gln	c.431T>A	missense	Reported	rs121917760	Het	Class 5
401303027	Male	Black-African	33	Non-familial	*MYBPC3*	Splice Site	c.821+1G>A	splice_donor	Reported	rs397516073	Het	Class 5
10100083	Female	Mixed	64	Non-familial	*MYH7*	p.Arg204His	c.611G>A	missense	Reported	rs397516260	Het	Class 5
10100502	Male	Mixed	29	Familial	*MYH7*	p.Arg204His	c.611G>A	missense	Reported	rs397516260	Het	Class 5
10100527	Male	Mixed	25	Non-familial	*ALPK3*	p.Arg1607Ter	c.4819C>T	stop_gained	Reported	rs1344615174	Het	Class 5
10100715	Female	Mixed	38	Familial	*MYBPC3*	p.Glu258Lys	c.772G>A	missense	Reported	rs397516074	Het	Class 5
101301381	Female	Mixed	37	Non-familial	*MYH7*	p.Ala797Thr	c.2389G>A	missense	Reported	rs3218716	Het	Class 5
10100517	Male	Mixed	36	Non-familial	*MYBPC3*	p.Ser478Ter	c.1433C>A	stop_gained	Novel	None	Het	Class 5
10100066	Male	White	25	Non-familial	*MYBPC3*	p.Trp792ValfsTer41	c.2373dup	frameshift	Reported	rs397515963	Het	Class 5
10100518	Male	White	46	Familial	*GLA*	p.Pro259ArgfsTer5	c.774_775del	frameshift	Reported	rs869312398	Hom	Class 5
10100535	Female	White	23	Non-familial	*MYH7*	p.Arg249Gln	c.746G>A	missense	Reported	rs3218713	Het	Class 5
101301393	Female	White	31	Non-familial	*MYH7*	p.Ala797Thr	c.611G>A	missense	Reported	rs3218716	Het	Class 5
10100505	Male	Mixed	38	Non-familial	*DSP*	p.Lys1892del	c.5676_5678del	inframe_deletion	Novel	rs768845836	Het	Class 4
10100536	Female	Mixed	25	No data	*MYH7*	p.Arg1420Trp	c.4258C>T	missense	Reported	rs145213771	Het	Class 4

4.3.3.1.1 Baseline characteristics

As our preliminary data analysis had shown that the number of probands carrying the class 4 and 5 variants differed by ancestry and age, we separated the data analysis according to the three ancestries: 57.1% (8/14) of the probands were MA, 28.6% (4/14) proband were White and 14.3% (2/14) of the probands were BA.

We explored the age of the adult-onset HCM probands that were class 4 and 5 positive and found no statistically significant difference (p=0.43) between the three population groups (Figure 4.5A). However, the Mixed probands were older, with a much higher median age of 37.

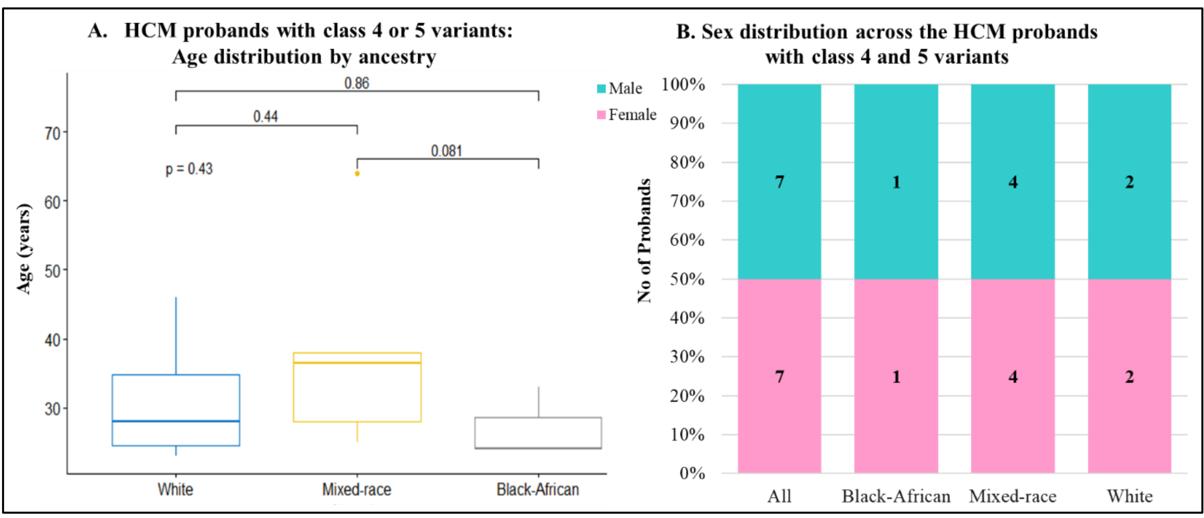


Figure 4.5: Age and sex distribution for adult-HCM class 4 and 5 variant carriers. ***A.*** *The age-to-ancestry summary for adult HCM probands with class 4 or 5 variants,* ***B.*** *The proportion of adult-onset HCM-positive probands with class 4 or 5 variants according to sex assigned at birth***.**

We further grouped the probands carrying the class 4 and 5 variants into various ancestry and noted that the eight positive MA probands had a mean age of 36.5 years (SD 11.6; range 25-64), which was younger than the HCM cohort mean (SD) age of 41.3 (13.9) years (Table 4.3). The diagnostic classification showed that 62.5% (5/8) showed non-familial or idiopathic diagnoses without genetic testing and FH information, and 25% (2/8) were classified as familial.

For the two BA HCM probands with class 4 or 5 variants, one was a female diagnosed with the familial disease at 24 years, and the other was a male diagnosed with non-familial HCM at 33 years old, Figure 4.5B.

For the four White HCM probands, they had a mean age of 31.3 years (SD 9.0); we recruited two females and two males. The diagnostic classification showed that 75% (3/4) had a non-familial or idiopathic diagnosis without genetic testing and FH information, and one proband had a familial classification, Figure 4.5B.

4.3.3.1.2 Diagnostic findings

The 14 (8 MA, 4 WA and 2 BA) adult HCM probands carried 13 class 4 and 5 variants spanning six genes. We observed that *MYBPC3* gene variants were present in all three ancestries. However, *MYH7*, the most prominent gene, was prevalent in WA and MA probands at 50% prevalence each.

4.3.3.1.2.1 Class 4 and 5 variants carriers' genetic characteristics

Overall, we found 13 class 4 and 5 variants; most of the mutations were in *MHY7* [30.8% (4/13)] and *MYBPC3* [[30.8% (4/13)] genes accounting for 61.5% (8/13) of the variants. Two *MYH7* variants recur in two probands, increasing the total to 15, as shown in Figure 4.6B. The *GLA*, *DSP*, *ALPK3* and *TNNI3* genes were each carried by one proband. Besides, the *TNNI3* gene harboured two class 5 variants carried as compound heterozygous variants by one proband.

The sunburst in Figure 4.6C shows the genetic distribution for the class 4 and 5 variants, indicating that the *MYH7* mutations identified were reported missense variants. For the four variants in the *MYBPC3* gene, we found a combination of reported/novel missense, splice donor and stop-gain variants. The two *TNNI3* variants were both reported class 5 missense variants. Most of the variants (60%) identified were missense mutations, and the remaining 40% consisted of two stop-gained, two frameshifts, one inframe_deletion and one splice-donor variant.

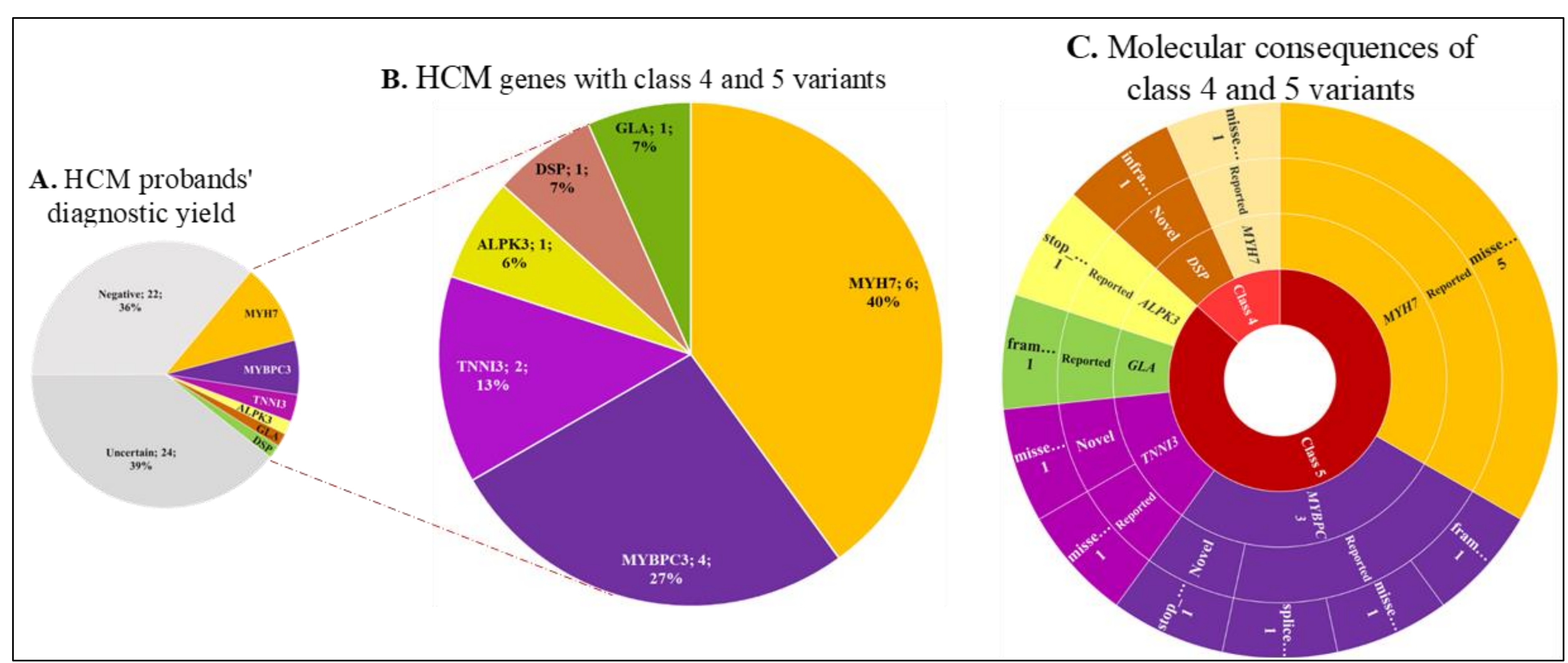


Figure 4.6: Genes found in adult-onset HCM cohort with class 4 and/or class 5 variants.

Among the 60 HCM probands initially recruited, the NGS panel yielded a 23.3% (14/60) diagnostic rate. Further, the probands carrying class 4 and 5 variants yielded 36.4% (4/11) for the WA probands, 23.5% (8/34) for MA and 13.3% (2/15) for the BA probands. We also noticed that *MYBPC3* occurred across the three ancestries; however, the most prominent gene, *MYH7*, was found in the MA and WA probands with a prevalence of 50%, respectively. Furthermore, the MA probands had a heterogeneous genetic aetiology. Hence, we further analysed the genes with class 4 and 5 variants identified through targeted sequencing per the probands' ancestry.

4.3.3.1.2.2 Class 4 and 5 variants carriers of Mixed ancestry (MA)

Of the 34 MA adult-onset HCM probands we recruited, we found eight (50% females) carrying class 4 (n=2) and 5 (n=5) variants. The MA probands carried seven class 4 and 5 variants harboured by the *MYH7* [50% (4/8)], *MYBPC3* [25% (2/8)], *ALPK3* [12.5% (1/8)] and *DSP* [12.5% (1/8)] genes, as shown in Figure 4.7D.

4.3.3.1.2.2.1 MYH7 c.611G>A

The missense *MYH7* c.611G>A (p.Arg204His) class 5 variant recurred in two MA probands diagnosed at the ages of 29 and 64, both classified as NYHA class I, with an ejection fraction of 80% and 67%, respectively. There was no FH of disease in one proband; however, targeted sequencing identified the *MYH7* c.611G>A class 5 variant as the likely cause. The NM_000257.4 (exon 7):c.611G>A alteration alters the amino acid composition, specifically from Arginine to Histidine (with similar polarity and positive charge) at residue 204. The c.611G>A variant is in the S1 sub-fragment motor domain.

4.3.3.1.2.2.2 MYH7 c.2389G>A

Proband (ID 101301381) is a 37-year-old MA female classified as NYHA class III, with an ejection fraction of 70%. There was no FH of disease in the proband; however, targeted sequencing identified the *MYH7* c.2389G>A (p.Ala797Thr) class 5 variant as the likely cause of disease.

The variant was also carried by a 31-year-old WA female classified as NYHA class II. The *MYH7* (exon 21) c.2389G>A variant results in an amino acid change from a non-polar Alanine

to a polar Threonine at residue 797. Substituting a non-polar, hydrophobic aliphatic amino acid with a polar, hydrophilic, uncharged amino acid (p.Ala797Thr) significantly changes the *MYH7* exon 21 physicochemical properties. The c.2389G>A variant is in the S1 sub-fragment regulatory domain of the *MYH7* gene.

4.3.3.1.2.2.3 MYH7 c.4258C>T

Proband (ID 10100536) is a 25-year-old MA female who tested positive for the *MYH7* c.4258C>T (p.Arg1420Trp) class 4 variant as the likely cause of disease. *MYH7* c.4258C>T missense variant in exon 31 causes an amino acid substitution that replaces an Arginine with Tryptophan at residue 1420. This change significantly alters the amino acid properties, as a non-polar, hydrophobic amino acid with an aromatic side chain replaced the polar and hydrophilic basic amino acid. Also, computational prediction tools and conservation analysis are consistent with pathogenicity.

4.3.3.1.2.2.4 MYBPC3 c.1433C>A

Proband (ID 10100517) is a 36-year-old MA male classified as NYHA class I, with an ejection fraction of 69%. There was no FH of disease in one proband; however, targeted sequencing identified the *MYBPC3* c.1433C>A (p.Ser478Ter) stop_gained class 5 variant as the likely cause of disease. The novel NM_000256.3 (*MYBPC3*):c.1433C>A variant creates a termination codon on amino acid residue 478 in an evolutionary conserved region. The replaced polar Serine amino acid and its surrounding residues are highly conserved across 12 species accessed. The proband (10100517) we identified carrying this variant was diagnosed with non-familial HCM, and the variant pathogenicity is consistent with the phenotype.

4.3.3.1.2.2.5 MYBPC3 c.772G>A

Proband (ID 10100715) is a 38-year-old MA female who had been classified as NYHA class I with a FH of disease. Targeted sequencing identified the *MYBPC3* c.772G>A (p.Glu258Lys) missense class 5 variant as the likely cause of the disease. The variant caused a replacement of a negatively charged polar Glutamic acid with a positively charged polar Lysine residue, resulting in a shift in the physiochemical dynamics. The variant RNA analysis indicates that

this missense change alters splicing in the last codon of *MYBPC3* exon 6 and might result in a truncated or no protein product or exon skipping.

4.3.3.1.2.2.6 ALPK3 c.4819C>T

We reported a 25-year-old MA male clinically diagnosed with HCM carrying a heterozygous *ALPK3* c.4819C>T (p.Arg1607Ter) stop_gain class 5 variant. There was no FH of disease in one proband classified as NYHA class I, with an ejection fraction of 73%. The variant has been reported three times in the ClinVar database, and the mutation is expected to result in a no protein due to NMD.

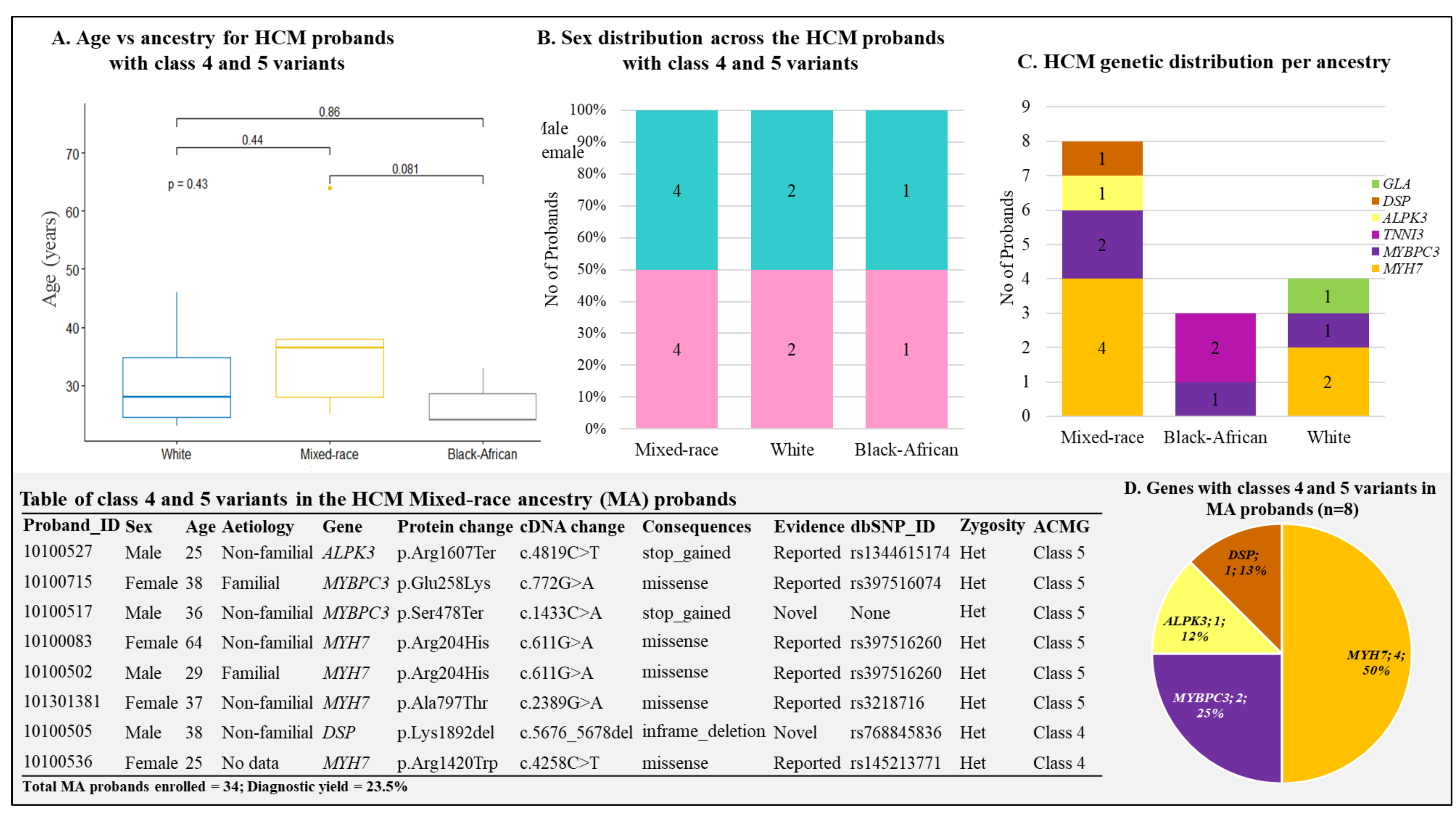


Table of class 4 and 5 variants in the HCM Mixed-race ancestry (MA) probands

Proband_ID	Sex	Age	Aetiology	Gene	Protein change	cDNA change	Consequences	Evidence	dbSNP_ID	Zygosity	ACMG
10100527	Male	25	Non-familial	*ALPK3*	p.Arg1607Ter	c.4819C>T	stop_gained	Reported	rs1344615174	Het	Class 5
10100715	Female	38	Familial	*MYBPC3*	p.Glu258Lys	c.772G>A	missense	Reported	rs397516074	Het	Class 5
10100517	Male	36	Non-familial	*MYBPC3*	p.Ser478Ter	c.1433C>A	stop_gained	Novel	None	Het	Class 5
10100083	Female	64	Non-familial	*MYH7*	p.Arg204His	c.611G>A	missense	Reported	rs397516260	Het	Class 5
10100502	Male	29	Familial	*MYH7*	p.Arg204His	c.611G>A	missense	Reported	rs397516260	Het	Class 5
101301381	Female	37	Non-familial	*MYH7*	p.Ala797Thr	c.2389G>A	missense	Reported	rs3218716	Het	Class 5
10100505	Male	38	Non-familial	*DSP*	p.Lys1892del	c.5676_5678del	inframe_deletion	Novel	rs768845836	Het	Class 4
10100536	Female	25	No data	*MYH7*	p.Arg1420Trp	c.4258C>T	missense	Reported	rs145213771	Het	Class 4

Total MA probands enrolled = 34; Diagnostic yield = 23.5%

Figure 4.7: Adult HCM probands of Mixed ancestry. *A. The age-to-ancestry summary for class 4 and 5 variant carriers. B. The sex-to-ancestry summary class 4 and 5 variant carriers, C. The genetic findings summary for the ACM cohort, D. Table shows the details of class 4 and 5 variants.*

4.3.3.1.2.3 Class 4 and 5 variants carriers of White ancestry (WA)

Of the 11 WA adult-onset HCM probands we recruited, we found four (50% females) carrying class 5 variants. The four probands carried four reported class 5 variants in *MYH7* [50% (2/4)], *MYBPC3* [25% (1/4)] and *GLA* [25% (1/4)] genes, as shown in Figure 4.8.

4.3.3.1.2.3.1 MYH7 c.746G>A

Proband (ID 10100535) is a 23-year-old WA female classified as NYHA class II, with an ejection fraction of 64%. There was no FH of disease in the proband; however, targeted sequencing identified the missense *MYH7* c.746G>A (p.Arg249Gln) class 5 variant as the likely cause of disease. The *MYH7* c.746G>A variant in exon 9 causes a change of amino acid from a positively charged polar Arginine to a polar Glutamine residue at position 249, located in a well-established functional domain. Notably, the amino acids have similar polarities. The c.746G>A variant is in the S1 sub-fragment motor domain.

4.3.3.1.2.3.2 MYBPC3 c.2373dup

Proband (ID 10100066) is a 25-year-old WA male classified as NYHA class III with an ejection fraction of 83%. The proband had no FH of the disease; however, targeted sequencing identified the frameshift *MYBPC3* c.2373dup (p.Trp792ValfsTer41) class 5 variant as the likely cause of disease. The c.2373dup frameshift variant disrupts the functional fibronectin type III domain of the *MYBPC3* gene by replacing Tryptophan with Valine at residue 792. The disruption creates a termination codon downstream from the variant at the 41st amino acid residue.

4.3.3.1.2.3.3 GLA c.774_775del

Proband (ID 10100518) is a 46-year-old WA male classified as NYHA class II, with an ejection fraction of 74%. The proband had a FH of disease. Targeted sequencing identified the frameshift *GLA* c.774_775del (p.Pro259ArgfsTer5) class 5 variant as the likely cause of disease. This variant disrupts the glycoside hydrolase superfamily domain of the *GLA* exon 5 by replacing Proline with Arginine. The disruption creates a termination codon at the fifth amino acid residue downstream from position 259. The affected C-terminus of the *GLA* exon 5 is a mutation hotspot associated with HCM and HCM phenocopy conditions such as Fabry disease.

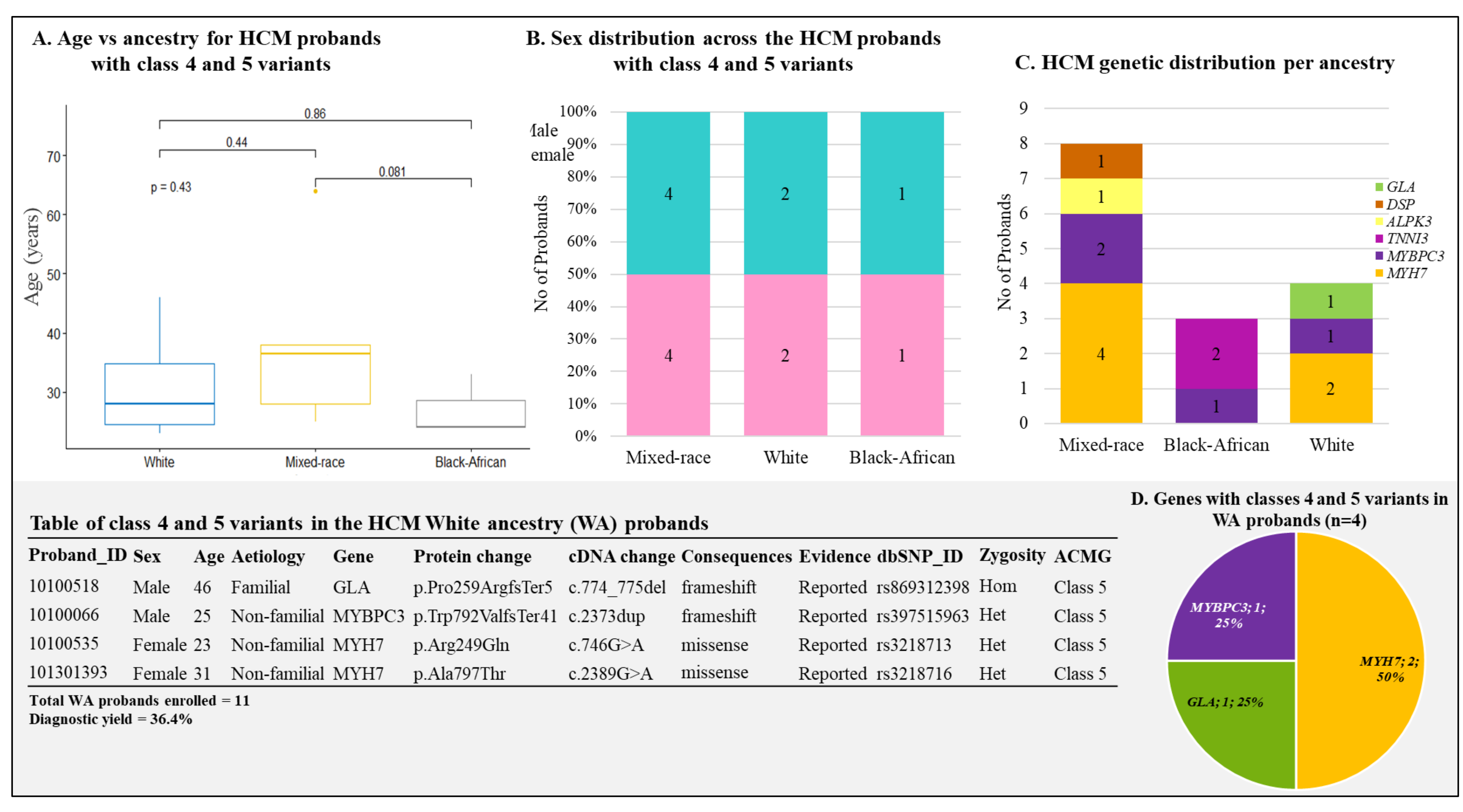


Table of class 4 and 5 variants in the HCM White ancestry (WA) probands

Proband_ID	Sex	Age	Aetiology	Gene	Protein change	cDNA change	Consequences	Evidence	dbSNP_ID	Zygosity	ACMG
10100518	Male	46	Familial	GLA	p.Pro259ArgfsTer5	c.774_775del	frameshift	Reported	rs869312398	Hom	Class 5
10100066	Male	25	Non-familial	MYBPC3	p.Trp792ValfsTer41	c.2373dup	frameshift	Reported	rs397515963	Het	Class 5
10100535	Female	23	Non-familial	MYH7	p.Arg249Gln	c.746G>A	missense	Reported	rs3218713	Het	Class 5
101301393	Female	31	Non-familial	MYH7	p.Ala797Thr	c.2389G>A	missense	Reported	rs3218716	Het	Class 5

Total WA probands enrolled = 11
Diagnostic yield = 36.4%

Figure 4.8: Adult HCM probands of White ancestry. ***A.*** *The age-to-ancestry summary for class 4 and 5 variant carriers.* ***B.*** *The sex-to-ancestry summary class 4 and 5 variant carriers, C. The genetic findings summary for the ACM cohort, D. Table shows the details of class 4 and 5 variants.*

4.3.3.1.2.4 Class 4 and 5 variants carriers of Black-African ancestry (BA)

Of the 15 BA adult-onset HCM probands we recruited, we found two (50% females) carrying class 5 variants. The two BA probands carried three class 5 variants (two reported and one novel) in *TNNI3* [67% (2/3)] and *MYBPC3* [33% (1/3)] genes, as shown in Figure 4.9.

4.3.3.1.2.4.1 TNNI3 c.431T>A and c.431_432delinsAT

Proband (ID 101301575) is a 24-year-old BA female classified as NYHA class IV, with an ejection fraction of 61%. The proband had a FH of HF. Targeted sequencing identified the missense *TNNI3* c.431T>A (p.Leu144Gln) and c.431_432delinsAT (p.Leu144His) class 5 compound heterozygous variants as the likely cause of disease.

4.3.3.1.2.4.2 MYBPC3 c.821+1G>A

Proband (ID 401303027) is a 33-year-old BA male classified as NYHA class I, with an ejection fraction of 64%. The proband had no FH of the disease; however, targeted sequencing identified the splice_donor *MYBPC3* c.821+1G>A class 5 variant as the likely cause.

The *MYBPC3* c.821+1G>A splice_donor variant we identified in a 33-year-old male diagnosed with non-familial HCM has been submitted 21 times in ClinVar with casualties' reports, including a sudden infant death syndrome. The *MYBPC3* c.821+1G>A canonical splice site variant is predicted to affect mRNA splicing, resulting in a significantly altered protein due to either exon skipping, shortening or the inclusion of intronic material. The variant is also predicted to cause loss of exon 7 alone or exons 7 and 8, with both transcripts capable of creating frameshifts and premature stop codons in exon 9. The molecular consequence of this variant also involved the loss of titin and myosin-binding sites.

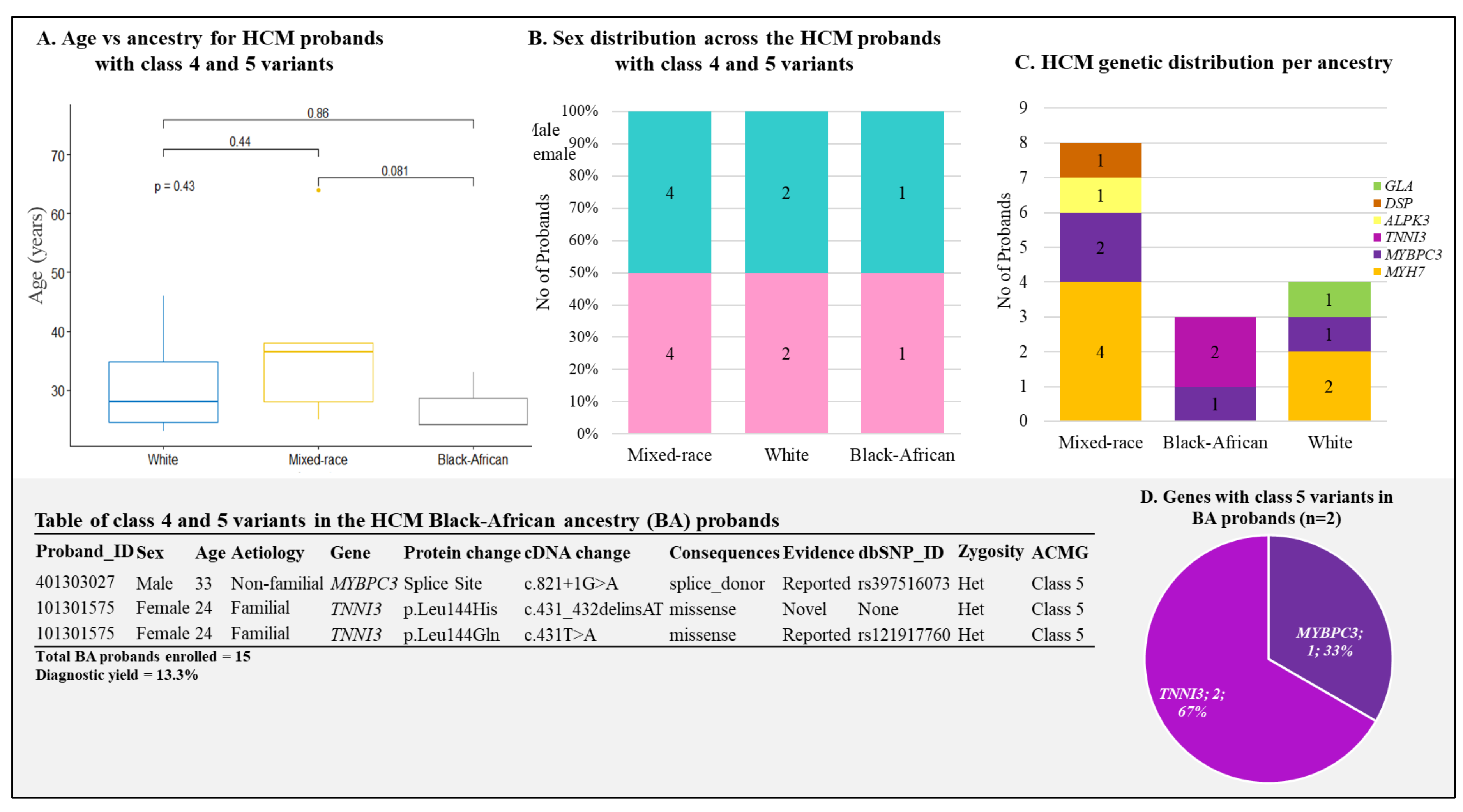


Table of class 4 and 5 variants in the HCM Black-African ancestry (BA) probands

Proband_ID	Sex	Age	Aetiology	Gene	Protein change	cDNA change	Consequences	Evidence	dbSNP_ID	Zygosity	ACMG
401303027	Male	33	Non-familial	*MYBPC3*	Splice Site	c.821+1G>A	splice_donor	Reported	rs397516073	Het	Class 5
101301575	Female	24	Familial	*TNNI3*	p.Leu144His	c.431_432delinsAT	missense	Novel	None	Het	Class 5
101301575	Female	24	Familial	*TNNI3*	p.Leu144Gln	c.431T>A	missense	Reported	rs121917760	Het	Class 5

Total BA probands enrolled = 15
Diagnostic yield = 13.3%

Figure 4.9: Adult HCM probands of Black-African ancestry. *A. The age-to-ancestry summary for class 4 and 5 variant carriers. B. The sex-to-ancestry summary class 4 and 5 variant carriers, C. The genetic findings summary for the ACM cohort, D. Table shows the details of class 4 and 5 variants.*

4.3.3.1.3 Familial segregation of class 4 and 5 variants

Two first-degree relatives of one proband (ID 101301393) were screened and found negative for the *MYH7* c.2389G>A missense variant.

4.3.3.2 Probands with variants of uncertain significance (class 3)

Twenty-four adult-onset HCM probands carried only class 3 variants: ten MA, nine BA, and five WA probands, as shown in Table 4.3. The 24 probands carried 38 class 3 variants listed in SS Table 2. The class 3 variants were further analysed, and we found that 50% (19/38) were novel. Of the 19 novel class 3 variants, 94.7% (18/19) were missense variants carried by 54.2% (13/24) of probands. Of some probands, 20.8% (5/24) had both novel and reported class 3 variants. We also noted that some class 3 variants favoured a class 4 category, as summarised in the next section.

Two of the 38 class 3 variants of potential interest favour the likely pathogenic class 4, but we had insufficient evidence according to the ACMG classification for HCM. These two variants were carried by 8.3% (2/24) of the class 3 variant carriers, as shown in Table 4.5. These variants were predicted *in silico* to have deleterious molecular consequences by SIFT, PolyPhen2 and REVEL database. However, other researchers have submitted the variants in the ClinVar database as class 3 due to insufficient evidence to support a class 4 or 5 pathogenicity. In the next section, we present the clinical characteristics of adult-onset HCM probands with only class 3 variants.

Table 4.5: Class 3 variants of interest in the HCM probands

GENE	PROTEIN CHANGE	cDNA CHANGE	CONSEQUENCES	EVIDENCE	dbSNP_ID	ACMG	PROBAND ID
MYBPC3	p.Gly416Ser	c.1246G>A	missense	Reported	rs371513491	Class 3	10100522
MYBPC3	p.Ala534Val	c.1601C>T	missense	Reported	rs374349666	Class 3	101301447

4.3.3.2.1 Baseline characteristics

Data analysis of the adult HCM probands carrying only class 3 variants showed they were diagnosed at a mean age of 41.5 years (SD 15.2, range 16-66) in Table 4.3. There was thus no significant difference between the overall adult-onset HCM age of diagnoses and probands positive for only class 3 variants. However, when we separated the class 3 variant carriers into ancestry groups, we noted a considerable age difference between the MA probands [mean age

49.7 years (SD 13.1, range 23-66)] in Figure 4.10. Figure 4.10 shows that the MA probands with class 3 variants were older than the BA and WA probands.

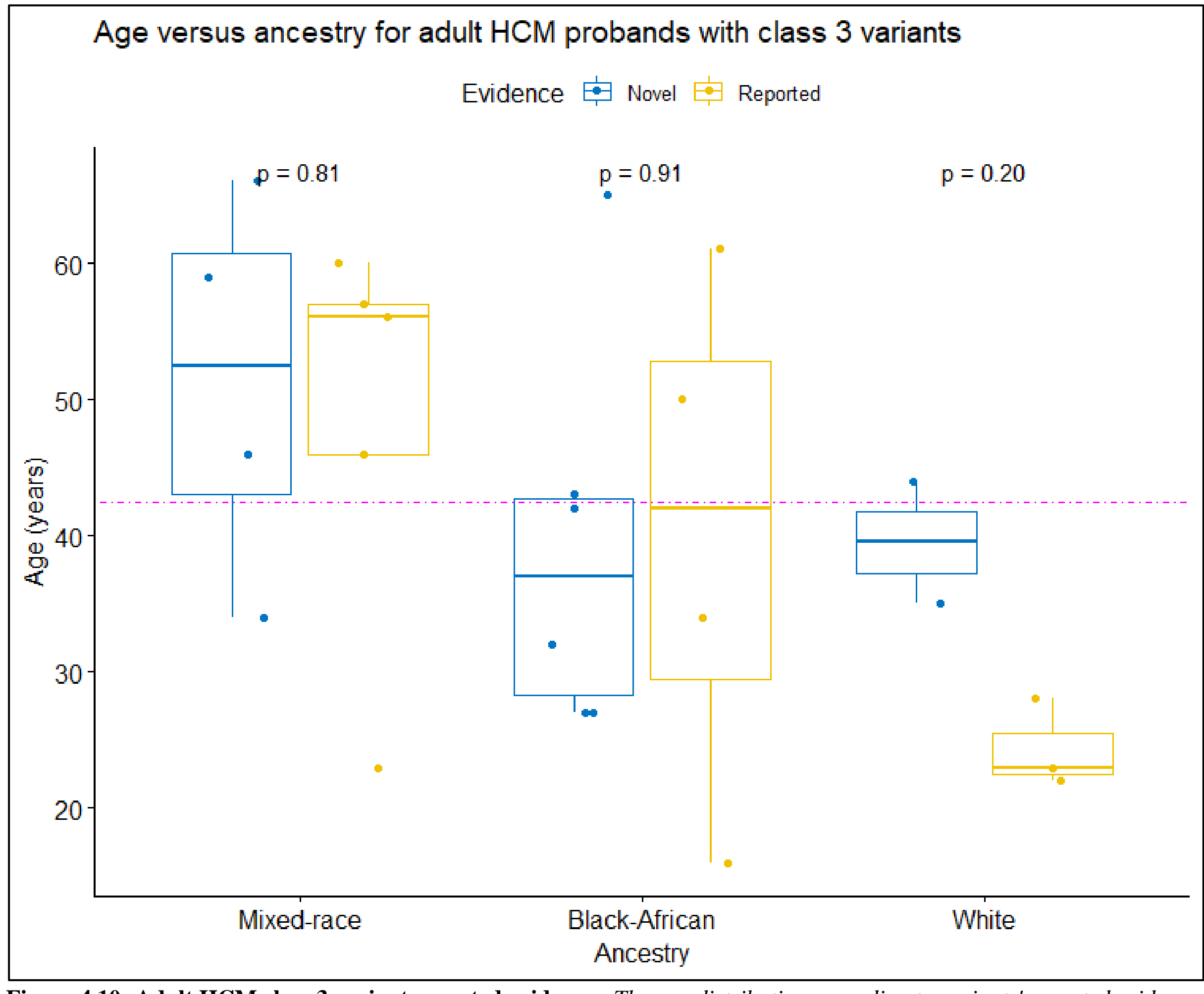


Figure 4.10: Adult HCM class 3 variant reported evidence. *The age distribution according to variants' reported evidence for adult-onset HCM probands carrying only class 3 variants.*

We also noted a 1:2.4 ratio of female to male probands carrying class 3 variants. When we separated the class 3 variant carriers into females and males, we noted an insignificant age difference within the various ancestries in Figure 4.11. However, Figure 4.11 shows that the female HCM probands carrying class 3 variants were younger than the males with the same ancestry. We then proceeded to investigate the genetic variants found per ancestry.

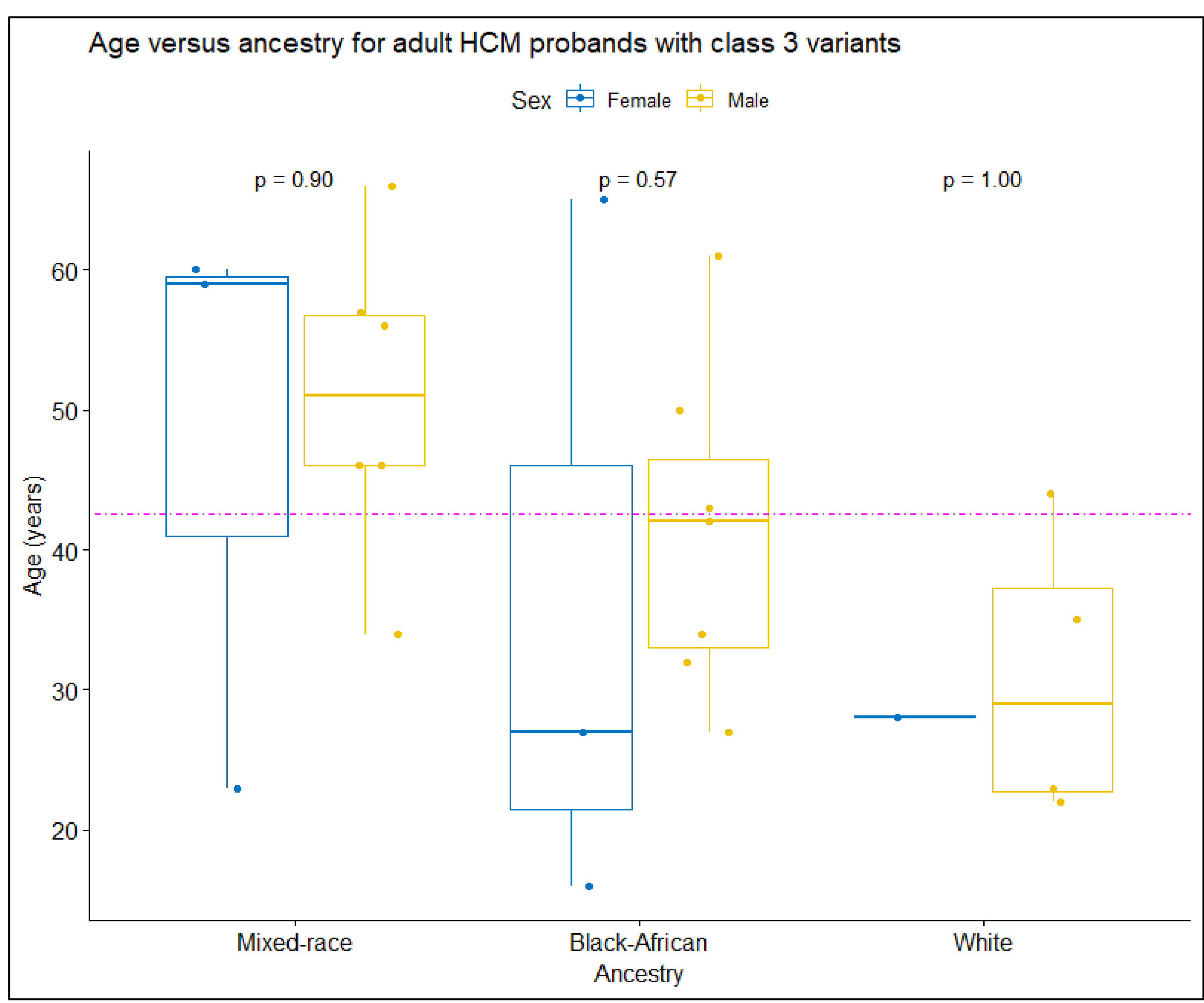


Figure 4.11: Adult HCM class 3 variant distribution according to sex. *The age distribution according to sex assigned at birth for the various ancestry adult-onset HCM probands carrying only class 3 variants.*

4.3.3.2.2 Diagnostic findings

The SS Table 2 summarises findings for the HCM class 3 variant carriers; however, we have briefly presented the results according to the various population groups in this section.

4.3.3.2.2.1 Class 3 variants carriers of Black-African ancestry (BA)

We found 10 BA adult-onset HCM probands carrying only class 3 variants: 30% (3/10) females and 70% (7/10) males. The 10 BA probands carried eight genes harbouring 21 class 3 variants; 80% (8/10) of the probands carried multiple variants. Of the 21 class 3 variants carried by the BA probands, 57.1% (12/21) were novel and 42.9% (9/21) were reported class 3 variants. The majority, 90.5% (19/21), of the class 3 variants carried by the BA probands had a missense molecular consequence. The remaining two variants were one frameshift and one splice-donor variant.

Eight genes harboured the class 3 variants we found in the BA probands; however, only four HCM-related genes are shown in Figure 4.12. Figure 4.12 shows that *MYBPC3* and *ALPK3* were the predominant HCM-related genes with class 3 variants.

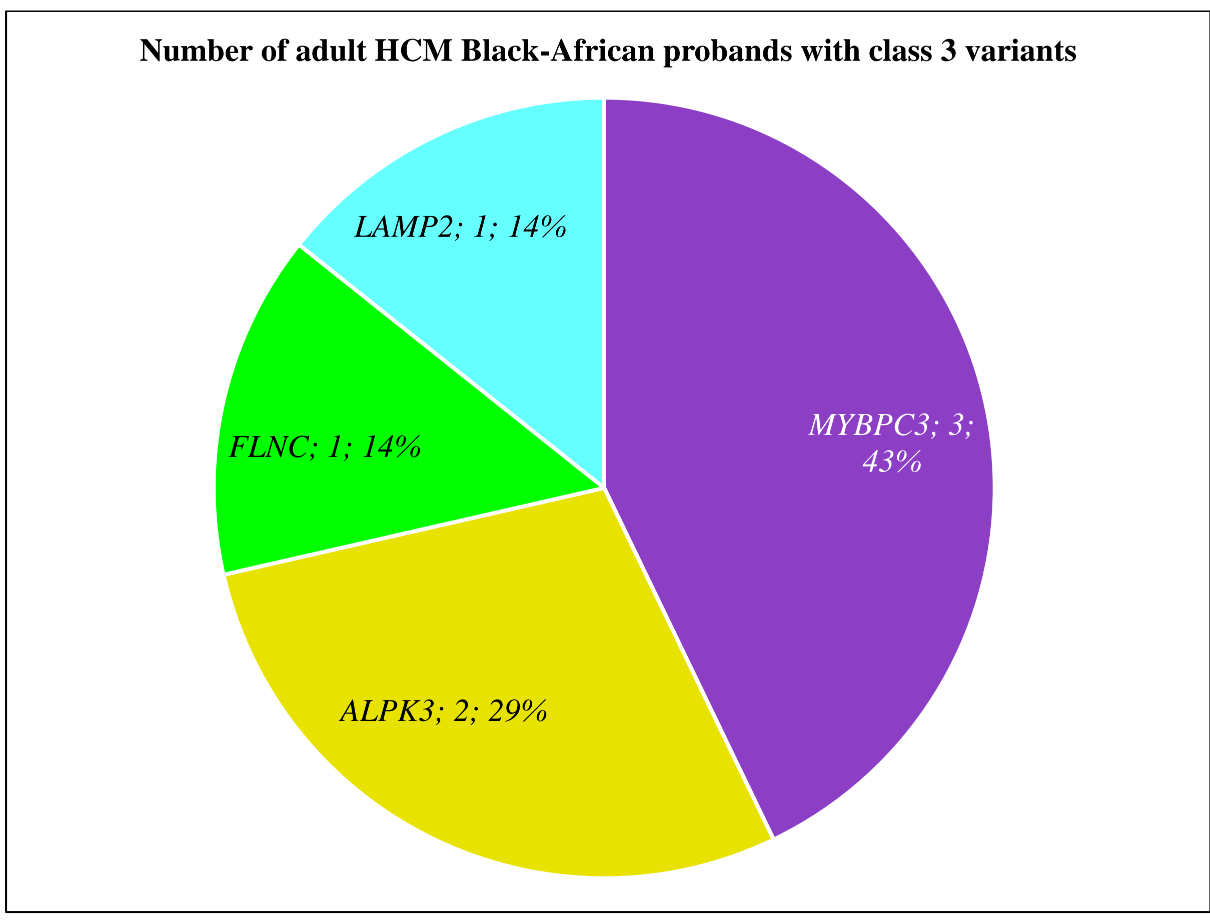


Figure 4.12: Adult HCM class 3 variant carriers of Black-African ancestry.

4.3.3.2.2.2 Class 3 variants carriers of Mixed ancestry (MA): HCM cohort

We found nine MA adult-onset HCM probands carrying only class 3 variants: 33.3% (3/9) females and 66.7% (6/9) males. The nine MA probands carried seven genes harbouring 14 class 3 variants; 55.5% (5/9) of the probands carried multiple variants. Of the 14 class 3 variants carried by the MA probands, 42.9% (6/14) were novel and 57.1% (8/14) were reported class 3 variants. The majority, 85.7% (12/14), of the class 3 variants carried by the BA probands had a missense molecular consequence. The remaining two variants were one frameshift and one splice-donor variant.

However, only two genes (*MYBPC3* and *ALPK3*) harboured HCM relevant variants. We found the *MYBPC3* c.1601C>T and c.713G>A variants in two probands and the ALKP3 c.904G>A variant in one proband.

4.3.3.2.2.3 Class 3 variants carriers of White ancestry (WA): HCM cohort

We found five WA adult-onset HCM probands carrying class 3 variants: four males and one female. The five WA probands carried three genes harbouring five class 3 variants. Three of the five class 3 variants carried by the WA probands were reported, and two were novel. Unlike the BA and MA class 3 variant carriers, none of the five WA probands carried multiple class 3 variants. Four class 3 variants in the WA probands had a missense molecular consequence, and one was a frameshift variant.

However, only the *FLNC* gene harboured HCM-relevant variants. The *FLNC* c.4636G>A missense variant was carried by one proband.

4.3.3.3 Summary of classes 1 to 5 variants carriers in our HCM cohort

The number of adult-onset HCM probands with variants found according to the ACMG five-tier classification system is summarised in Figure 4.13.

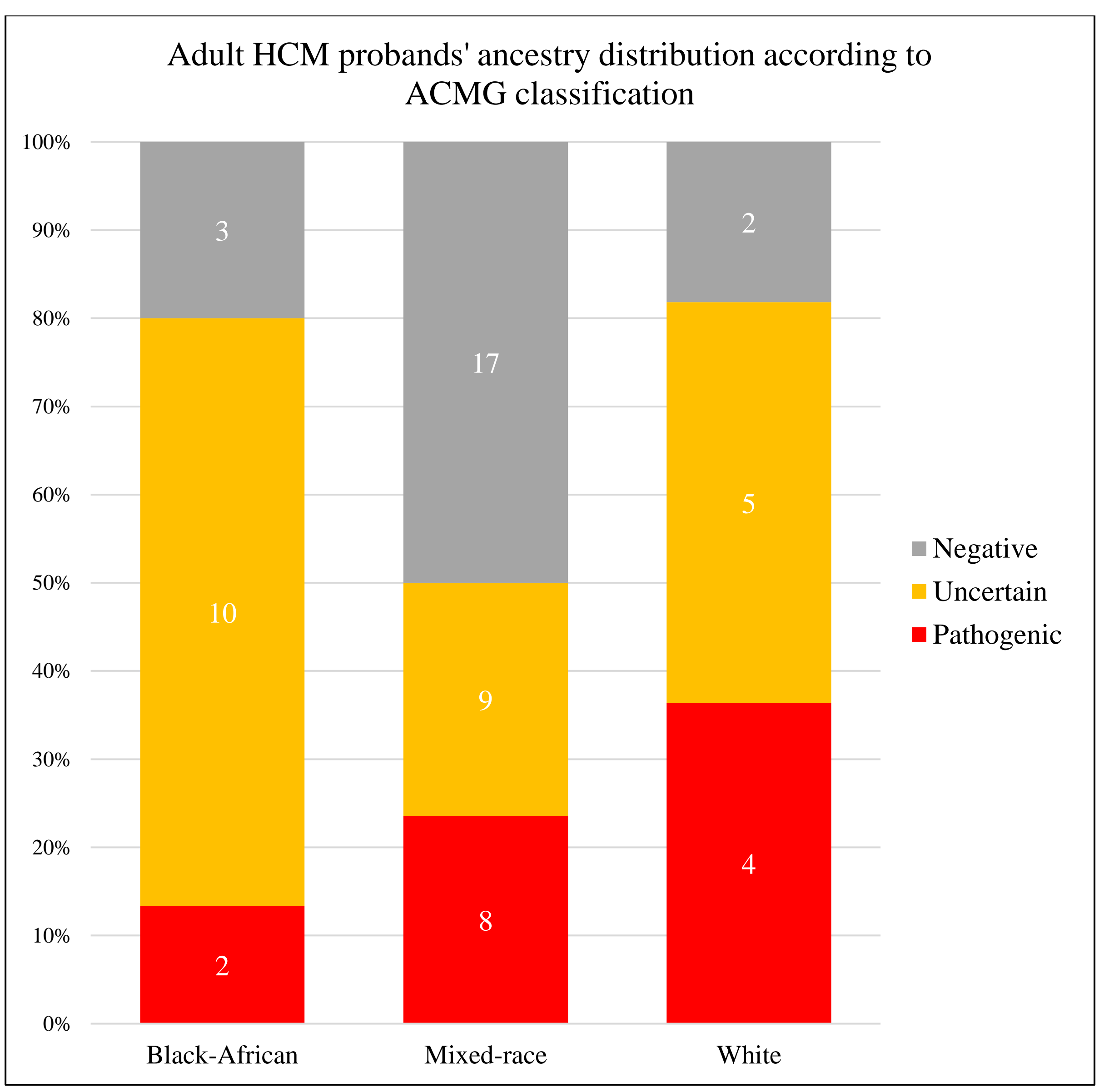


Figure 4.13: Summary of adult HCM classes 1 to 5 variants carriers. *The adult-onset HCM probands were categorised according to the cohort's population ancestry.*

Targeted sequencing was performed on our adult-onset cohort of 60 probands, as shown in Figure 4.13, and we found 14 probands carrying class 4 or 5 variants: 12 probands carrying class 5 variants and two probands carrying class 4 variants. While 24 probands carried class 3 variants and 22 probands carried class 1 and/or 2 variants.

We found six genes harbouring 13 class 4 or 5 variants; 23.1% (3/13) were novel, and 76.9% (10/13) were reported. Amongst the reported variants, 15.4% (2/13) were recurring mutations. Also, we found two class 5 variants in one proband. The variants prevalence ranged from 7.7% (1/13) to 30.8% (4/13) for the 14 HCM probands (Table 4.6).

Table 4.6: Summary of combined class 4 and 5 variants in our adult HCM cohort.

GENE	CLASS 4 VARIANTS		CLASS 5 VARIANTS		TOTAL	PREVALENCE
SYMBOL	**Novel**	**Reported**	**Novel**	**Reported**	**n**	**(%)**
MYH7	0	1	0	3	4	30.8
MYBPC3	0	0	1	3	4	30.8
TNNI3	0	0	1	1	2	15.4
ALPK3	0	0	0	1	1	7.7
DSP	1	0	0	0	1	7.7
GLA	0	0	0	1	1	7.7
TOTAL	1	1	2	9	13	

4.3.4 Paediatric HCM probands

4.3.4.1 Baseline characteristics for paediatric HCM cohort

In this study, we enrolled seven unrelated paediatric probands clinically diagnosed with HCM. Their average age at diagnosis was 3.6 years (SD 4, range <10) and the probands were all assigned male sex at birth. Self-reported ancestry found 86% (6/7) BA and 14% (1/7) WA probands. The clinical characteristics are shown below in Table 4.7.

Table 4.7: Baseline results for the paediatric HCM probands cohort

LABEL	LEVELS	TOTAL	UNCERTAIN	NEGATIVE
DIAGNOSTIC YIELD, N (%)		7	4 (57.1)	3 (42.9)
AGE AT ENROLMENT, YEARS	Mean (SD)	3.6 (4.0)	0.7 (0.6)	5.8 (4.2)
SEX RECORDED AT BIRTH	Male	7 (100)	3 (42.9)	4 (57.1)
ANCESTRY	Black-African	6 (85.7)	3 (100.0)	3 (75.0)
	White	1 (14.3)	0 (0.0)	1 (25.0)

4.3.4.2 Targeted sequencing for paediatric HCM probands

The analysed targeted sequencing data disclosed three paediatric probands with class 3 and four probands with class 1-2 variants. There was no paediatric proband positive for the class 4 or 5 variant in our HCM cohort.

4.3.4.3 Paediatric probands with variants of uncertain significance (class 3)

There were three BA male paediatric probands carrying only class 3 variants in this cohort, as shown in SS Table 2. The NGS data showed that the class 3 variant carriers were diagnosed at a median age of about one year, Table 4.7. The probands were all BA males, and notably, there was no significant difference between the overall paediatric HCM sex or age at clinical diagnoses and the four probands with class 3 variants. The four probands carried five missense class 3 variants, four novel and one reported. The novel variants were harboured by the *DMD* and *TTN,* whereas the reported [c.104222C>T (p.Thr34741Ile)] was in the *TTN* gene.

4.3.5 Possible founder variants in the HCM cohort

Multiple probands carrying the same class 5 variant were uncommon in our HCM cohort of 67 probands. However, we identified two probands with the same class 5 variants. The class 5 variants were found in the *MYH7* gene (i.e., c.611G>A and c.2389G>A). Two MA probands carried the *MYH7* c.611G>A (p.Arg204His) missense variants: one female and one male. Similarly, two female probands (one MA and one WA) carried the *MYH7* c.2389G>A (p.Ala797Thr) missense variant.

4.4 **DISCUSSION**

This study used a comprehensive targeted NGS panel to identify variants causing HCM in a South African cohort of 67 unrelated probands, including 60 adult-onset and seven paediatric cases. The probands were of various ancestry, including Black-African, Mixed, Indian, and White ancestry. We recruited the probands from public tertiary cardiac clinics serving patients from low-resource semi-urban and urban areas of South Africa. The adult-onset and paediatric HCM findings are discussed separately.

4.4.1 Adult-onset HCM probands' baseline findings

We enrolled 60 adult HCM probands diagnosed at a mean age of 41.3 years (SD 13.9, range 16-75 years). Our HCM probands cohort had a similar mean onset age at diagnosis compared to some reported global average onset ages of about 44 years in studies with similar designs (156, 307). However, the literature's baseline age of ≤45 years has been reported mostly in larger cohorts with sample sizes >1200 participants (95, 192).

We found that 65% (39/60) of the probands were male, consistent with self-reported sex in literature (308). As per published findings, female proband relatives are more likely to test positive for gene variants found in the family at a younger age (95, 308).

Unlike more extensive HCM studies consisting of predominantly (>85%) WA participants (95, 278), our probands' self-reported ancestry revealed that 56.7% were of MA, 25% of BA and 18% of WA. Our sample size for participants of WA was limited compared to the literature, mainly from European countries with common ancestry to White South Africans (67, 156). It is plausible that through genetic testing and genealogical tracing, families of White people in South Africa, particularly those with European ancestry, have been notified of their disease status and may have made informed health decisions. Nevertheless, the results of our MA and BA probands indicate that a Black African heritage may increase the likelihood of developing HCM, a conclusion that is consistent with current research (156). Indeed, when we looked further at our probands FH at diagnosis, we found that 35% (21/60) had familial HCM consisting of 40% (6/15) BA, 36% (4/11) White and 32% (11/34) MA probands.

4.4.2 Adult-onset HCM probands' targeted sequencing findings

Fourteen (23.3%) of the 60 adult HCM probands (8 MA, 4 WA, and 2 BA) had pathogenic variants. The remaining 46 probands (76.7%) had no pathogenic variants: 24 with (10 BA, 9 MA and 5 WA) only class 3 variants and 22 with (17 MA, 3 BA and 2 WA) only class 1 and 2 variants, as shown in Table 4.3. We first discussed the findings of the probands who tested positive for class 4 and 5 variants, followed by a brief report on class 3 variants and their probable modifying effects on HCM.

Amino acid changes

Pathogenic variants tend to destabilise proteins and their interactions. It has been shown that genetic variants causing diseases often result in significant changes to the physicochemical properties of proteins at the amino acid sequence level (309). These changes include alterations in charge, hydrophobicity, and geometry. Structural analysis of disease-causing variants, often observed in the human population, reveals similar trends. Specifically, it has been observed that these variants tend to cause more changes in the hydrogen bond network and salt bridges than harmless amino acid mutations (309). Both experimental and computational thermodynamic data indicate that these disease-causing variants tend to destabilise proteins and their interactions. The following discussion examined some of these variants and how they led to HCM. We aim to explore a potential correlation between genotype and HCM phenotype.

4.4.2.1 Pathogenic (class 4 and 5) variants

4.4.2.2 Class 4 and 5 variant carriers' clinical findings

Table 4.3 shows a positive rate of 23.3% (14/60) for genotyped findings, comparable to recent similar studies with global yields ranging from 20% to over 65% (42, 192). The mean age of these probands is about eight years lower than (p-value=0.036) our HCM cohort mean age at diagnosis. On average, the 14 probands were diagnosed at a younger age of 33.9 years (SD 11.1, range 23-64 years), which differs from what is published in the literature. Most HCM patients carrying pathogenic variants in the literature are about 50 years old (95, 307).

Similar to the literature, a higher proportion of our female probands, 33.3% (7/21), than males, 17.9% (7/39), had pathogenic variants (16, 95). It is important to note that the family members of female probands are more likely to test positive for the mutation found in the proband at a younger age (308). Unfortunately, we could not investigate this further in our study due to the

unavailability of samples from the relatives. Our findings contradict previous studies that suggest HCM-causing variants are primarily found in older females and younger males (307).

Our study also revealed that 71.4% of participants with class 4 or 5 variants (8 MA and 2 BA probands) were of Black African descent. Prior research has indicated that this subpopulation group is more likely to develop HCM (156). Even though Figure 5 unequivocally demonstrated that genetic causes affect male and female probands equally, a higher proportion of young Black African females were more vulnerable to HCM-causing variants. Contrarily, the literature reports a higher incidence of HCM in more youthful males and older females (310). Besides, Black ancestry has been identified as a predictor for HCM with worse outcomes, especially in carriers of sarcomere class 4 and 5 variants (278).

4.4.2.3 Class 4 and 5 variants carriers' genetic findings

The class 4 and 5 variants were identified in six HCM disease-causing genes (120). The six genes identified encode for the cardiomyocyte's sarcomere (*MYH7*, *MYBPC3* and *TNNI3*), cytoskeleton (*ALPK3* and *GLA*), and desmosome (*DSP*) proteins. Like in most published studies, HCM-causing variants are mainly found in the proteins encoded by eight core sarcomere genes (i.e., *MYH7*, *MYBPC3*, *TNNT2*, *TNNI3*, *TPM1*, *ACTC1, MYL2* and *MYL3*) and many variants have been found in the *MYBPC3* and *MYH7* genes (90). Similarly, *MYH7* and *MYBPC3* pathogenic variants accounted for 66.7% of the genotype-positive probands in this study.

4.4.2.3.1 The sarcomere genes' (*MYH7*, *MYBPC3* and *TNNI3*) variants

The sarcomere genes *MYH7*, *MYBPC3*, and *TNNI3* can be inherited in an AD manner, affecting both males and females equally in some families (278, 308). However, the proportion of female probands we reported with class 4 or 5 variants was higher in our cohort [33.3% (7/21) females vs 17.9% (7/39) males]. The main known sarcomere (*MYH7* and *MYBPC3*) genes' related HCM class 4 or 5 variants we curated have often remodelled the affected patients' myocardium (311). The myocardial ventricular walls thickened without cavity enlargement, and most *MYH7* than *MYBPC3*-related HCM patients usually have a higher rate of developing myocardial fibrosis (312). Myocardial fibrosis frequently leads to affected patients requiring a heart transplant or risking SCD. Notably, we reported 11 probands of various ancestry with sarcomere genes known for causing familial HCM (*MYH7*, *MYBPC3*, and *TNNI3*), and 90.9%

(10/11) of them had disease onset before age 38. Unexpectedly, 63.6% (7/11) were females, and one of the female probands carried two *TNNI3* compound heterozygous variants.

4.4.2.3.1.1 **The *MYH7* gene**

We reported four class 4 and 5 variants identified in the *MYH7* gene. *MYH7* gene is the second well-known genetic cause of HCM, with prevalence ranging from 15% to 25% in adult HCM patients (16). In this study, the common cause of HCM was *MYH7* variants, with a carrier rate of 42.9% (6/14). Like other findings, over 83% (5/6) of our *MYH7* variant carriers were female probands (307). Notably, *MYH7*-related HCM is characterised by diastolic dysfunction due to abnormal myocardial thickened ventricular walls and impaired contractile function (283). Diastolic dysfunction is most common in end-stage HCM patients associated with missense variants (307).

We described the *MYH7*:c.611G>A (313), c.746G>A (280), c.2389G>A (314) and c.4258C>T (315) missense variants we found that have been previously reported in the literature.

4.4.2.3.1.1.1 MYH7 c.611G>A variant

Despite the similarity in the polarity of both amino acids in the *MYH7* c.611G>A (p.Arg204His) variant, the alteration is in the globular head domain of the *MYH7* gene, accounting for part of 35% of HCM *MYH7*-related variants (313). The variant has been mentioned 16 times in the ClinVar database, the carriers have had various ancestry, and the patients have presented with severe early-onset HCM (156).

4.4.2.3.1.1.2 MYH7 c.746G>A

This *MYH7* c.746G>A (p.Arg249Gln) variant has been reported in several HCM patients with robust segregation data for HCM (156). The variant has been mentioned 17 times in the ClinVar database and multiple studies (316). For example, in a three-generation pedigree study, the affected members carrying the *MYH7* c.746G>A either died suddenly or required an urgent heart transplant (280).

4.4.2.3.1.1.3 MYH7 c.2389G>A

The c.2389G>A (p.Ala797Thr) variant pathogenicity is known, and the mutation has caused a more severe HCM traced to a White South African ancestor (314). The c.2389G>A variant carriers are at risk of SCD, mostly as adolescents or young adults (67).

4.4.2.3.1.1.4 MYH7 c.4258C>T

This variant has been reported in over ten patients in the literature and is characterised by early onset, severe phenotypes, and left ventricular hypertrophy (315, 317). Another amino acid change (p.Arg1420Gln) at this position has also been reported in SCD patients diagnosed with HCM patients (318).

In comparison, the *MYH7* c.4258C>T class 4 variant alters the *MYH7* gene's light meromyosin (LMM) domain, where it interacts with the titin and myosin-binding protein C3 sites. The alteration is expected to reduce LMM domain flexibility and impair sarcomere contractile function. These pathomechanisms are commonly described in *MHY7*-related HCM patients with later onset (147). However, our probands had an early onset of the disorder, and we think *MYH7* gene-environmental interactions might have triggered their early onset.

Similarly, our *MYH7*-related HCM probands all had missense variants located on the meromyosin domain of the gene. Most class 4 or 5 missense variants in the *MYH7* gene's heavy meromyosin domain have been reported as an aetiology of HCM compared to other regions (312). Specifically, missense variants in the S1 (head) and S2 (hinge) sub-fragments of the heavy meromyosin domain are the common cause of *MYH7*-related HCM (147). Figure 4.14 illustrates that the identified variants were distributed across the three domains of the *MYH7* gene.

Pathogenic variants in the S1 motor domain are expected to cause structural abnormalities between the myosin and Adenosine triphosphate (ATP) binding site, impairing energy utilisation. Meanwhile, variants in the S1 regulatory domain might cause structural abnormalities between the myosin and actin filaments that impair sarcomere sliding. Structural alteration of the S1 sub-fragments in the site of ATPase activity and cardiac α-actin binding site is the primary pathomechanisms of HCM-related *MYH7* missense variants (147). Our probands with the missense variants in the S1 domain had onset before age 38, consistent with literature for severe early-onset familial HCM (319).

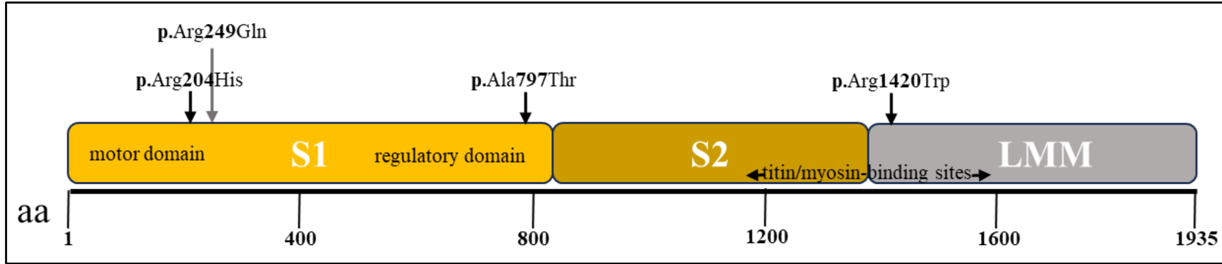


Figure 4.14: Schematic structure of *MYH7* with HCM-related variants' location. *The cardiac β-myosin heavy chain 7 is divided into three regions: the sub-fragment 1 (S1), sub-fragment 2 (S2) and light meromyosin (LMM). The S1 includes a motor and a regulatory domain. The S2 and LMM interact with the myosin-binding protein C3 and titin. Drawn by P. Ndibangwi using Microsoft 365 PowerPoint.*

4.4.2.3.1.2 The *MYBPC3* gene

We identified four variants in the *MYBPC3* gene in BA, MA, and WA probands, including three truncating (two reported (c.2373dup (320) frameshift and c.821+1G>A (321) splice_donor) and a novel (c.1433C>A stop-gained)), and one reported missense c.772G>A variant (322).

MYBPC3 gene is the most common cause of genetic HCM, and its prevalence has ranged from about 20% to 66.7% in adult-onset HCM patients (95). In this study, the carrier rate of the *MYBPC3* gene class 4 and 5 variants was 30.8% (4/13). The carriers were mainly young males, unlike our *MYH7*-related HCM cases, where most carriers were females. Probands with *MYBPC3* variants are more likely to be male and develop LV hypertrophy at a younger age (66). Our young male probands had three *MYBPC3* class 5 truncating variants: variant as schematically represented in Figure 4.15.

4.4.2.3.1.2.1 MYBPC3 c.2373dup

The reported c.2373dup (p.Trp792ValfsTer41) variant is a known Dutch founder mutation associated with late-onset familial HCM (182); however, the affected proband we reported was diagnosed at age 25. The prevalence is 46% in the Dutch population (66). The cMyBP-C C6 functional fibronectin type III disruption results in a characteristic thickened ventricular wall that leads to diastolic abnormalities or cavity obstruction (320). Left ventricular outflow tract obstruction is common in *MYPBC3* variant carriers, and the patients usually receive surgical septal myectomy to relieve the symptoms (323).

4.4.2.3.1.2.2 MYBPC3 c.821+1G>A

The c.821+1G>A splice_donor variant has been reported to segregate with HCM in several families, and the patients have mainly developed the phenotype before turning 30 years old (85, 324) and like in these families, our proband with the c.821+1G>A (a BA male) developed severe HCM at 33 years old. The c.821+1G>A variant is in the phosphorylation motif domain of the gene. Phosphorylation of cardiac MyBP-C protein regulates cardiac contraction through modulation of actin-myosin interactions mediated by the protein's amino-terminal (N')-region (C0-C2 domains, 358 amino acids) (325), as shown in *Figure 4.15*. During myocardial injury, dephosphorylation of cMyBP-C leads to the cleavage of the C0-C1f region consisting of 271 amino acids (325). Variants such as the *MYBPC3* c.821+1G>A splice_donor that disrupts the phosphorylation motif could lead to sarcomere structural stability due to cMyBP-C haploinsufficiency (326).

Subsequently, it leads to contractile dysfunction with a characteristic thickened ventricular wall that could lead to cavity obstruction (326). *MYBPC3*-related HCM is usually characterised by diastolic dysfunction due to profound LV hypertrophy, and cavity obstruction has been detected in about 25% of HCM cases, especially in *MYBPC3* truncating variant carriers (327).

4.4.2.3.1.2.3 MYBPC3 c.1433C>A

The c.1433C>A variant is in the immunoglobin-like C type 3 (C3) domain of the *MYBPC3* gene region associated with familial HCM (246). Over 50% of *MYBPC3* truncating variants have caused HCM through LoF or haploinsufficiency (299). The c.1433C>A variant mutant shortened gene product is expected to undergo NMD, causing sarcomere haploinsufficiency and resulting in ventricular hypercontractility (213). Evidence from human myectomy samples has shown a 24% lower cMyBP-C content in the heart tissues with *MYBPC3* variants (323). We, therefore, hypothesise that the termination codon on the C3 domain introduced by the c.1433C>A variant will diminish the cMBP-C protein's titin and myosin LMM binding regions and cause sarcomere haploinsufficiency.

4.4.2.3.1.2.4 MYBPC3 c.772G>A

In cultured murine cardiomyocytes, transgenic rescue experiments have demonstrated altered contractile kinetics and reduced contractile force due to the presence of the c.772G>A (p.Glu258Lys) variant, independent of any effects on splicing (328). Additionally, a mouse

model has been created with this variant, which exhibits the characteristic symptoms of HCM (305), and functional RNA analyses have demonstrated that nucleotide alteration on the affected amino acid residue is pathogenic (322). The authors proposed that phosphorylation and protein-protein interactions could be the primary mechanism of c.772G>A related HCM. Our patient with the *MYBPC3* c.772G>A variant, like other *MYBPC3*-related HCM patients with missense variants, has impaired relaxation on spectral and tissue Doppler imaging (307). These patients, however, are expected to have a milder phenotype (322).

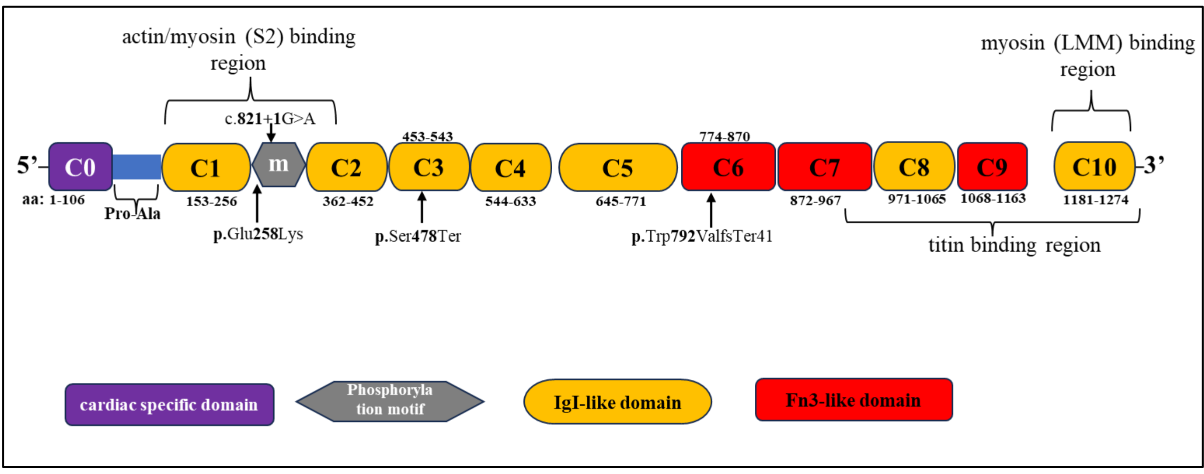


Figure 4.15: Schematic structure of *MYBPC3* with HCM-related variants' location. *Cardiac MyBP-C consists of 8 immunoglobins (Ig) like and 3 fibronectin (Fn3) domains with binding sites for actin, myosin and titin. The N-terminal C0 is specific for cMyBPC, as is the MyBPC-like motif between C1 and C2, which contains 3 phosphorylation sites. Drawn by P. Ndibangwi using Microsoft 365 PowerPoint.*

4.4.2.3.1.3 The *TNNI3* gene

4.4.2.3.1.3.1 TNNI3 c.431T>A and c.431_432delinsAT compound heterozygous variants

Compound heterozygous (two different class 4 or 5 variants in the same gene) is reported in about 2% to 23.6% of HCM patients and generally causes a more severe phenotype with early onset age (210, 297). We found one (1.7%) proband carrying *TNNI3* c.431T>A and c.431_432delinsAT compound heterozygous variants. Our clinic team diagnosed the proband with severe familial HCM at the age of 24. At diagnosis, the proband self-identified as a female of BA. Her clinical features were similar to those previously described in another South African proband carrying the c.431_432delinsAT (p.Leu144His) missense variant associated with focal ventricular hypertrophy (329). Meanwhile, the c.431T>A (p.Leu144Gln) has been reported to be pathogenic in patients with idiopathic RCM (330). Like these cases, our proband presented with a severe disorder, reporting symptoms of dyspnoea, and had sinus rhythm on ECG with LV hypertrophy, MWT of 10mm on ECHO, and NYHA functional class IV HF.

The *TNNI3* gene encodes for a thin filament protein complex component in the sarcomere called troponin subunit I3. The *TNNI3* gene product is specific to the heart and inhibits the actin-myosin interactions to mediate cardiac muscle relaxation. During cardiac muscle relaxation, the troponin I main binding domain with amino acids sequence from codon 137 to 148 regulates Ca^{2+} sensitivity and inhibits the cardiac muscle's actin-myosin ATPase activity. Loss of *TNNI3* actin-myosin ATPase inhibitory activity prevents myosin and actin from detaching and, as a result, impairing energy utilisation, which leads to contractile dysfunction (331). Diastolic dysfunction is a clinical feature of HCM, and HCM-related *TNNI3* class 4 or 5 variants have accounted for up to 6% of mild to severe HCM cases (95, 246). We, therefore, hypothesise that our proband's severe presentation could be attributed to these *TNNI3* compound heterozygous variants.

4.4.2.3.2 The cytoskeleton genes' (*ALPK3* and *GLA*) variants

Cardiomyocytes' cytoskeletal proteins are involved in several cardiac functions, including structural stability and intracellular communication. Class 4 or 5 variants altering the cytoskeletal protein structures are associated with cardiovascular system disorders, such as cardiac conduction, ventricular arrhythmia, and CMOs. In this study, we reported two cytoskeletal structure genes (*ALPK3* and *GLA*) whose variants have been associated with HCM.

4.4.2.3.2.1 The *ALPK3* gene

The α-kinase 3 (*ALPK3*) is a protein-coding gene whose function is relatively unknown, although its protein product is predicted to participate in cardiomyocyte differentiation. This pseudo-kinase is localised to the cell's nuclear envelope and M-band to enable ATP binding activity (289). Biallelic class 4 or 5 truncating variants (tv) that cause the lack of *ALPK3* gene product enzymatic activity are commonly associated with infant mixed DCM/HCM phenotypes that rapidly progress into HCM within one year (332). Infant-onset HCM is common in *ALPK3*tv carriers, and *ALPK3* heterozygous variant carriers have generally presented with similar clinical features that include mixed apical subtypes of HCM and RV involvement with late-onset (333). About 13.5% of HCM-related heterozygous *ALPK3*tv carriers have had late onset, although often reported in young adults (334).

4.4.2.3.2.1.1 ALPK3 c.4819C>T

In this study, we reported a 25-year-old MA male clinically diagnosed with HCM carrying a heterozygous NM_020778.4 (*ALPK3*) c.4819C>T (p.Arg1607Ter) stop-gained variant. The c.4819C>T variant is reported in ClinVar without a known function work. However, heterozygous *ALPK3*-tv are generally pathogenic, and many have segregated with clinical HCM features (335). The pathomechanism of pathogenic heterozygous *ALPK3*-tv is LoF, and carriers have developed a severe cardiac phenotype with extensive myocardial fibrosis leading to HF (335, 336).

4.4.2.3.2.2 The *GLA* gene

The *GLA* gene encodes the lysosomal α-galactosidase A, an enzyme that hydrolyses glycosphingolipids such as globotriaosylceramide (GL-3) in various tissues and organs, including the heart's myocardium. Class 4 or 5 variants that cause the lack of *GLA* gene product enzymatic activity led to GL-3 accumulation in the lysosomes. GL-3 accumulation usually causes Fabry disease, an X-linked lysosomal storage disease (337). The clinical manifestations of Fabry disease include cardiac involvement with irreversible fibrosis, LV hypertrophy typically associated with instances of diastolic dysfunction and late gadolinium enhancement.

4.4.2.3.2.2.1 GLA c.774_775del

In this study, we identified a 46-year-old White male with a known GLA c.774_775del (p.Pro259ArgfsTer5) frameshift variant, which is irrelevant in HCM. However, we have reported the variant as a possible cause of Fabry disease (337). Patients carrying mild *GLA* mutations with ≥20% residual *in vitro* activity usually present with later-onset disease and a symptom spectrum different from the classically affected patients (338). However, a severe phenotype caused by a 2bp *GLA* c.718_719delAA (p.Lys240Glufs*9) deletion has been reported in a 12-year-old male patient clinically diagnosed with HCM with no FH of SCD (339). Furthermore, the prevalence of *GLA*-related HCM is <2% and the reported cases had similar clinical findings of atrial fibrillation and non-sustained VT (16), as observed in our proband (10100518).

As shown in Figure 4.16, this frameshift variant introduced a stop codon in a highly conserved and mutation hotspot on exon 5 of the *GLA* gene. The *GLA* mutation hotspot is associated with

HCM and Fabry disease (337, 340). However, only one case has been reported in the literature with this c.774_775del variant (337).

Figure 4.16: Schematic structure of *GLA* with HCM-related variant's location. The position of the p.Pro259ArgfsTer5 mutation is indicated (marked with an arrow). *TIM, triosephosphate isomerase.*

4.4.2.4 Variants of unknown significance (class 3)

Forty per cent (24/60) of the adult-onset HCM probands had only class 3 variant (s). The 24 probands had a mean (SD) age of 41.5 (15.2) years at enrolment, Table 4.3. The mean age of these probands is about eight years higher than the mean age of their class 4 or 5 variants carriers' counterparts at diagnosis. Moreover, males accounted for 71% of the class 3 variants carriers, and the proband of BA and MA were the most affected.

We have reported 38 class 3 variants in 12 CMO disease-causing genes, which we have further analysed to determine how much they could affect the severity of HCM phenotypes. Modifier genes are not causal but can affect the severity of HCM phenotypes (307). Previous studies show that genetic background can influence cardiac hypertrophic response (341). Clinical observations of patients with HCM have shown significant variability in phenotypic expression of hypertrophy, sudden death, and cardiac failure due to the influence of modifier genes. Although HCM is a classic monogenic disorder, the final phenotype is the product of the causal variant, modifier genes, and environmental factors (342). Identifying and characterising the role of modifier genes is essential for designing specific and comprehensive therapy for HCM.

4.4.2.4.1 The Black-African ancestry (BA) HCM probands with class 3 variants only

We reported 10 BA adult-onset HCM probands carrying only class 3 variants in this study, and the majority were males, 70% (7/10). There was thus no significant difference between the overall adult-onset HCM sex recorded at diagnoses and probands carrying only class 3 variants. Also, only three (30%) of these BA probands were diagnosed with familial HCM, and the genetic aetiologies of the disease are unknown in the other probands, primarily due to their small family sizes. Further, the 10 BA adult-onset HCM probands were young adult-onsets with a mean age of 39.7 years at diagnosis, probably because most of the probands were of economically active age and could afford frequent hospital visits for their disease to be

classified early. The ten adult-onset BA probands carried 21 class 3 variants identified in eight CMO-associated genes, as shown in Figure 4.12.

Most, 61.9% (13/21) of the class 3 variants had missense molecular consequences and 69.2% (9/13) were novel. The novel variants were found most frequently in the *TTN* and *DSP* genes; however, both genes are irrelevant in HCM.

4.4.2.4.2 The Mixed ancestry (MA) HCM probands with class 3 variants

We reported nine MA adult-onset HCM probands carrying only class 3 variants in this study, and the majority were males, 66.7% (6/9). Although there was no significant difference between the overall MA adult-onset HCM probands sex distribution recorded at diagnoses and those carrying only class 3 variants, it is essential to note that they were, on average, older (mean age = 49.7 years). Also, only two (22.2%) of these MA probands were diagnosed with familial HCM, which indicates that the cause of their disease is unknown genetic aetiologies. The nine adult-onset MA probands carried 14 class 3 variants identified in seven CMO-associated genes.

We found that 85.7% (12/14) of the class 3 variants carried by the MA probands had missense molecular consequences in seven genes. Unlike the BA probands, these probands had a broad spectrum of genes; however, we think that none of these variants has influenced the HCM severity. For example, the *ALPK3* c.904G>A variant favoured a LB based on rule BP4 of the ACMG classification system.

4.4.2.4.3 The White ancestry HCM probands with class 3 variants

Unlike in the BA and MA probands, we reported five WA adult-onset HCM probands carrying only class 3 variants in this study, with the majority being males, 80% (4/5). Only two probands were diagnosed with familial HCM, but the five WA probands carried five class 3 variants identified in three CMO-associated genes.

Apart from one frameshift variant, all the other class 3 variants had missense molecular consequences. Unlike the BA and MA probands, these probands had very few genes; however, only two variants found in the *FLNC* and *ALPK3* genes might have modified the reported phenotypes, such as the *FLNC* c.4636G>A. Out of these two variants, we concluded that the *FLNC* c.4636G>A variant, which barely met the criteria for class 4 variants, could be an HCM

phenotype modifier. It is worth noting that the *FLNC* c.4636G>A has segregated with CMO phenotypes in an Italian family with members clinically diagnosed with restrictive and hypertrophic CMO (107). The c.4636G>A (p.Gly1546Ser) missense variant replaces Glycine, which is a neutral and non-polar amino acid, with Serine, which is a neutral and polar amino acid at codon 1546 of the FLNC protein sequence. Although there is insufficient data to classify this variant as pathogenic, most *FLNC* missense variants in the ROD1 domain are enriched in HCM and RCM phenotypes (97, 343, 344).

4.4.2.5 Possible founder variants in the HCM cohort

Two *MYH7* missense class 5 variants [c.611G>A (p.Arg204His) and c.2389G>A (p.Ala797Thr)] we identified are possible founder variants. During the literature search, we found that the c.2389G>A carried by two females of WA and MA has been described as a South African founder mutation linked to a common ancestor (67).

Meanwhile, the c.611G>A, first reported in our population in a male proband of Indian ancestry (3), was carried by a female and a male proband of MA. The c.611G>A variant has been reported five times in the gnomAD exome database in individuals with Asian and non-Finish European ancestry. Given South Africa's colonial history, we recommend investigating this variant as a possible founder mutation.

4.4.2.6 Polygenic risk scores

Polygenic risk score (PRS) is an *in-silico* method used to estimate an individual's genetic risk of a trait or disease. Recently, the PRS for HCM has been calculated using genome-wide association studies (GWAS) and multi-trait analysis data of the UK Biobank participants. The authors found that a significant proportion (37%) of the phenotype variability for the 1733 HCM-disease variant carriers studied could be predicted with the PRS, and they identified 16 loci associated with HCM (182). These findings were later evaluated for HCM risk in both population-based and hospital-based settings to evaluate the contribution of genetic and clinical factors to HCM risk (345). Both studies demonstrated, using data from the UK Biobank participants, that PRS substantially improved HCM risk prediction beyond that achieved with clinical factors (182, 345). Noticeably, the UK Biobank participants of African ancestry were <1% used for the scoring, and there is no record of GWAS for HCM patients from the African continent. This raised some concerns, especially regarding using the PRS for HCM cases on

the African continent. Therefore, PRS should be considered as a further study to expand on the findings of this thesis.

4.4.3 Paediatric HCM

4.4.3.1 Baseline findings for the paediatric HCM cohort

Seven paediatric probands diagnosed with HCM were included in this study to determine their genetic characteristics. The paediatric probands had a mean age of 3.6 years (SD 4.0, range <12) at diagnosis. The paediatrics' self-reported ancestry revealed that 86% and 14% had BA and WA, respectively. All the participants were identified as male during the enrolment process, as shown in Table 4.7. We found this concerning because most male patients diagnosed with HCM in childhood have a significantly higher lifetime risk of developing incident left ventricular systolic dysfunction (LVSD) (278). LVSD impairs the heart's contractile function, reducing the heart's ability to pump blood effectively. Environmental factors, age, and gender can impact the development and progression of LVSD in familial HCM.

4.4.3.2 Targeted sequencing findings for paediatric HCM cohort

We identified only class 3 variants in 42.9% (3/7) of paediatric probands using a comprehensive NGS panel; no proband had a class 4 or 5 variant. Our positive rate of zero per cent is unlike those of other paediatric HCM cohorts with similar study designs that have yielded about 30% (96). Although similar panels have yielded positive results for paediatric (42), we think our paediatric genetic aetiologies were unidentifiable through the targeted NGS strategy because of the small sample size.

The three paediatric probands with only class 3 variants were all males of BA. The first proband carried three novel *TTN* missense class 3 variants: c.11122C>G, c.46468C>A, and c.58199T>C. The second proband carried a reported class 3 *TTN* c.104222C>T variant and the third proband carried a novel *DMD* c.2428A>C variant. These variants were not further investigated for the following reasons: 1) missense variants in these genes are less likely to cause paediatric HCM; 2) the variants were classified as ACMG class 3, but they were likely to be benign; 3) none of the genes were one of the primary HCM-causing genes (313).

4.5 **CONCLUSION**

Our adult-onset HCM class 4 or 5 variants carriers diagnosed using targeted NGS with a mean (SD) age of 33.9 (11.1) years were about a decade younger than their global counterparts. Unlike reported in the literature, our female probands carrying class 4 or 5 variants were much younger, with a mean age of 34.6 years at diagnosis.

Most of our adult probands were of Black African descent at enrolment. The genetic yield in our HCM cohort ranged from a minimum of 13.3% in BA probands to 34.6% in probands of WA, with the MA probands yielding 23,5%. Our adult HCM probands combined yield was 23.3%. The *MYBPC3* and *MYH7* accounted for 61.5% (8/13) of the class 4 and 5 variants. The *MYH7* c.611G>A and c.2389G>A missense variants, each recurring in two probands. We confirmed *TNNI3* compound heterozygous variants in a 24-year-old female BA proband diagnosed with familial HCM.

Finally, seven paediatric and 46 adult probands had negative or unknown genetic results. Of these probands, we hypothesised that the *FLNC*: c.4636G>A class 3 variant has affected the severity of the HCM phenotype in the affected proband. Our finding highlights that in South Africa, a significant proportion of novel genetic variants, specifically class 3, may require further evaluation to determine if they are pathogenic or benign.

Chapter 5: Restrictive Cardiomyopathy (RCM)

5.1 **INTRODUCTION**

5.1.1 Clinical background

Restrictive cardiomyopathy (RCM) is a potentially lethal cardiomyopathy that is characterised by diastolic dysfunction caused by abnormal muscle relaxation and myocardial stiffness, resulting in restrictive filling of the ventricles (346, 347). Diastolic dysfunction is frequently accompanied by left atrial or bi-atrial enlargement, normal ventricular size, and systolic function (347). The ventricular dysfunction could also cause infiltrative, non-infiltrative, storage and endomyocardial disorders, the latter reflecting endocardial pathology (e.g., fibrosis, fibroelastosis, thrombosis) that impairs diastolic function. RCM can vary according to pathogenesis, clinical presentation, diagnostic evaluation, treatment, and prognosis (346, 347).

5.1.2 Signs and symptoms

The most common symptoms of RCM are dyspnoea, palpitations, syncope, or thromboembolic events, often triggered by physical activity. Other symptoms might include chest pain, syncope or symptoms of stroke or sudden death. RCM patients usually receive a functional diagnosis, and some of the clinical investigations shown in Figure 1.9 in section 1.1.1.11 have required a high index of suspicion and rely heavily on echocardiography and cardiac imaging to distinguish the different disease mechanisms. For example, heritable RCM is characterised by near-normal-sized left ventricles with enhanced stiffness and enlarged atria due to increased ventricle end-diastolic pressure.

5.1.3 RCM disease burden

RCM is the most uncommon type of cardiomyopathy, accounting for 2% to 5% of paediatric cardiomyopathy cases. However, RCM has the worst outcome among cardiomyopathies, with survival rates at 1, 2, and 5 years after diagnosis reported to be 82%, 80%, and 68%, respectively. It has been reported that without a heart transplant, the median survival is two and a half years. RCM remains a largely unexplored disease with limited treatment options (347).

Although the global prevalence of heritable RCM is unknown, limited data from Africa suggests that the incidence of RCM in the SSA region is about 4% and 27.7% in the paediatric and adult certain populations, respectively (14). However, the prevalence in the South African population is unknown, and the reported patients have had similar genetic causes to those in North American and European people, characterised by the features of HCM or DCM (348). The said South African patients have had a range of modes of inheritance patterns, and their genetic characteristics are often unknown at clinical diagnosis until determined using candidate genes panel testing.

5.1.4 Aetiology of RCM

RCM can stem from various causes, such as idiopathic or familial, which can be inherited through AD, AR, or X-linked patterns. Additionally, RCM can be secondary to a systemic disorder, where the disease can be classified as infiltrative or non-infiltrative RCM. Infiltrative types include amyloidosis, sarcoidosis, Gaucher's disease, and Hurler syndrome, while non-infiltrative forms include idiopathic RCM and scleroderma (347, 349). RCM can also be caused by storage diseases such as Fabry's disease, glycogen storage disorders, and hemochromatosis. Other factors such as endomyocardial fibrosis, hyper-eosinophilic syndromes, metastatic tumours, radiation, sickle cell disease, and certain medications like anthracyclines, chemotherapeutic agents, ergotamine, and chloroquine can also damage the endocardium (349).

It is worth noting that patients under the age of 30 with RCM are commonly associated with a wide range of systemic disorders, while those over 65 years of age with RCM are more likely to be linked to amyloidosis, hemochromatosis, and radiation-induced endomyocardial damage (347, 350).

5.1.5 Types of RCM

The overlap in morphological and physiological abnormalities, genetic mutations, and the dynamic nature of the disease make it challenging to categorise cardiomyopathies (347). All classification systems have two primary functions: (i) offering nosographic schemes in which diseases are categorized based on their unique characteristics and (ii) proposing criteria that aid

in the diagnostic process. A single classification scheme that serves both functions well is rare (46).

The different forms of RCM are classified based on myocardial histology, genetic basis, and the nature of restriction. RCM is a condition that different disease mechanisms can cause. These mechanisms can be grouped into four main categories, which are used to classify the various disorders that can lead to RCM: (i) interstitial fibrosis and intrinsic myocardial dysfunction, (ii) infiltration of extracellular spaces, (iii) accumulation of storage material within cardiomyocytes, or (iv) endomyocardial fibrosis as shown in Figure 5.1 (46).

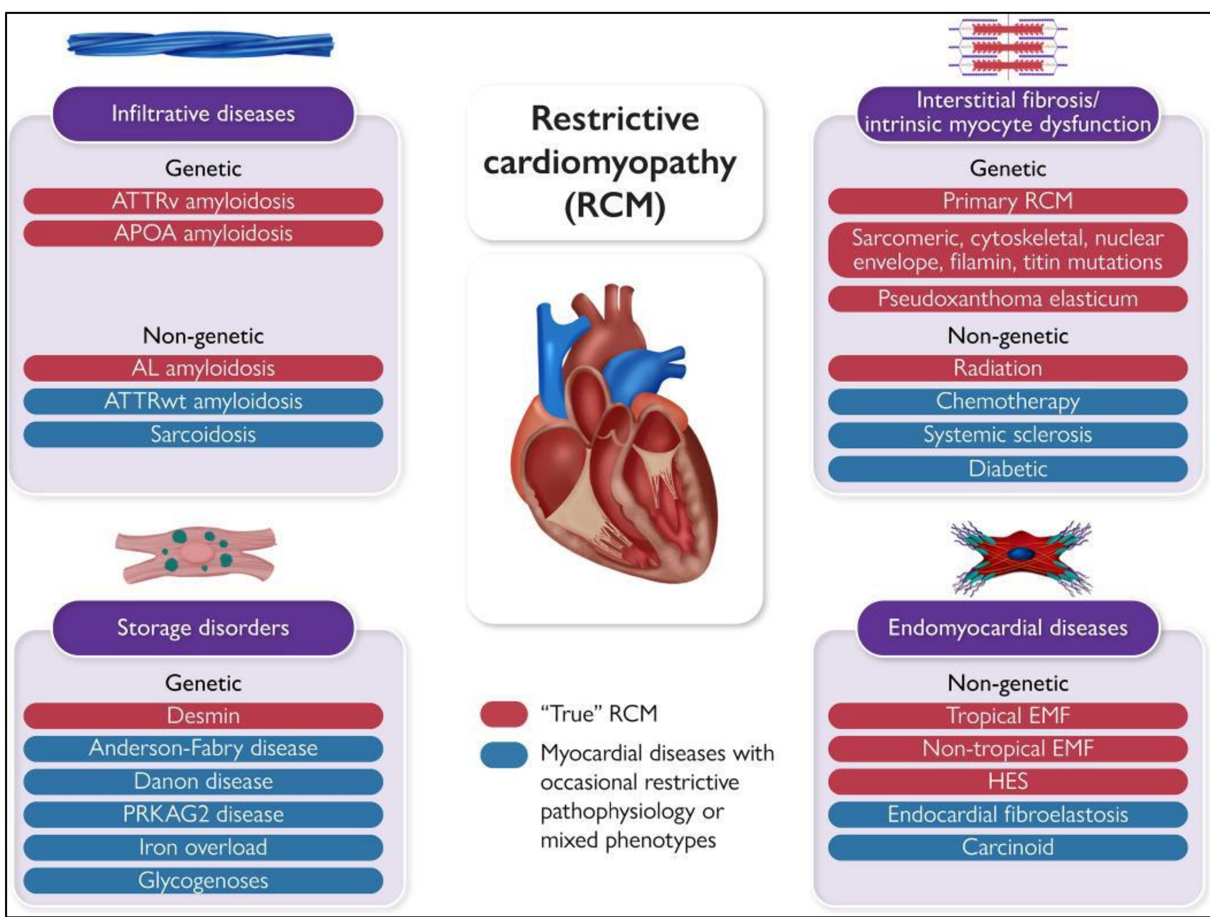


Figure 5.1: Overview of RCM proposed classification. *APOA, apolipoprotein A; ATTR, amyloid transthyretin (v, variant; wt, wild-type); AL, amyloid light-chain; EMF, endomyocardial fibrosis; HES, hypereosinophilic syndrome; PRKAG2, protein kinase AMP-activated non-catalytic subunit gamma 2* (46).

RCM can present differently in each patient, with congestive HF, atrial fibrillation, elevated systemic and pulmonary venous pressures, and a potential need for a heart transplant being common indicators. Bi-atrial dilatation is typically the main pathological finding, resulting from increased filling pressures and characterised by a normal ventricular cavity and size but

with an abnormal, rigid, rubbery texture of the endocardium. Microscopic changes in the myocardium may include nonspecific findings such as myocyte hypertrophy, focal or diffuse perimyocytic fibrosis, and focal myofiber disarray.

5.1.6 Genetics of RCM

Over the past few decades, advancements in molecular genetics and improved imaging techniques have expanded diagnostic capabilities and understanding of the pathophysiology underlying RCM (347).

Genetic RCM is generally characterized by a near-normal-sized left ventricle with enhanced stiffness and enlarged atria due to increased ventricle end-diastolic pressure. The disease is combined with an abnormal filling pattern and thus belongs to the diastolic diseases. Systolic function, at least at the beginning of the disease, is near normal but might be reduced at later stages. Sometimes mild hypertrophy is also observed, making the diagnostic distinction between RCM and HCM difficult (351).

With the advent of molecular genetics and improvements in diagnostic modalities, the understanding of cardiomyopathies shifted to include modifier genes and causative genetic mutations in order to explain pathologic mechanisms. Patients with RCM have been shown to have multiple genetic mutations, which are primarily found in the genes encoding cytoskeletal and sarcomeric proteins, which are regulated by mechanotransduction and mechanosensing proteins (348). In genetically based RCM, approximately 30% of cases are familial, with inheritance usually recorded as autosomal dominant (331, 347). The genes that have been linked to RCM are *TNNT2*, *TNNI3*, *MYBPC3*, *MYH7*, *MYL2*, *MYL3*, *DES*, *TTN*, *BAG3*, *LNMA*, and *FLNC* (Table 5.1).

Cell lines and animal model studies have revealed a few mechanisms of pathogenesis, including increased Ca^{2+} sensitivity, sarcomere disruption, and protein aggregates. More research on the pathophysiology of RCM is required in order to develop innovative therapeutic approaches that can reverse RCM phenotypes (347). Currently, the molecular genetic mechanisms of RCM are less understood, especially in our population, compared to the reported findings summarised (352).

Table 5.1: Prevalence of published adult RCM-related genes

GENE SYMBOL	FUNCTION	PREVALENCE (%)	REF: PMID
MYH7	Sarcomere	2 - 10	35026164
FLNC	Actin crosslinking	1 - 8	31245841
TNNI3	Sarcomere	Rare - 2	27662471
MYBPC3	Sarcomere	<1	27662471
TNNC1	Sarcomere	<1	27662471
TNNT2	Sarcomere	<1	27662471
ACTN2	Z-disc	<1	27662471

The use of the HCM candidate genes panel to determine the genetic aetiology of RCM in most cascade families has yielded a genotype-positive rate of about 54% in most cohorts (348). However, genes in the sarcomere, desmosome, and cytoskeletal have accounted for most (≈38%) reported RCM-causing variants (353). Notably, nontruncating variants (mostly in *FLNC*) have triggered several severe RCM (343, 354), and classes 4-5 variants in *FLNC*, *MYBPC3*, and *TNNI3* genes have dominated adult heritable and sporadic cases (353). Literature has revealed that *TNNI3* missense variants have been a common mechanism for the disorder, especially in paediatrics (355). Furthermore, the AR mode of inheritance with mutations in *TNNI3* has occurred more often in adult patients (356).

5.1.7 RCM genotype-phenotype correlation

A smaller percentage (30%) of primary adult RCM cases are due to genetic determinants; however, the aetiology is often unclear (347). Some patients with RCM diagnosed with unknown causes carry the genetic aetiology of the disease in an AD pattern and may have the genes found in HCM or DCM phenotypes. Examples include variants in the 19 genes encoding for cardiomyocytes sarcomeric and cytoskeletal proteins commonly affecting cardiac contractility, as shown in Figure 5.2 (352).

Phenotype overlaps are common in patients diagnosed with familial RCM, especially with HCM, DCM, and LVNC phenotypes (347). The overlaps have occurred more when sarcomeric disease-causing genes are involved; however, several genes whose proteins are not directly involved in contractile function have been reported in many RCM patients, for example, desmin, filamin C (344) and crystallin B genes (352). Although the knowledge of the molecular mechanism of heritable RCM is patchy, patients diagnosed with molecular genetic determinants have had several clinical outcomes (331). Generally, the interplay between

sarcomeric or associated proteins damages the myocyte dynamics, resulting in diastolic dysfunction, impaired structural stability, myocyte-to-myocyte miscommunication and increased ventricular stiffness (264).

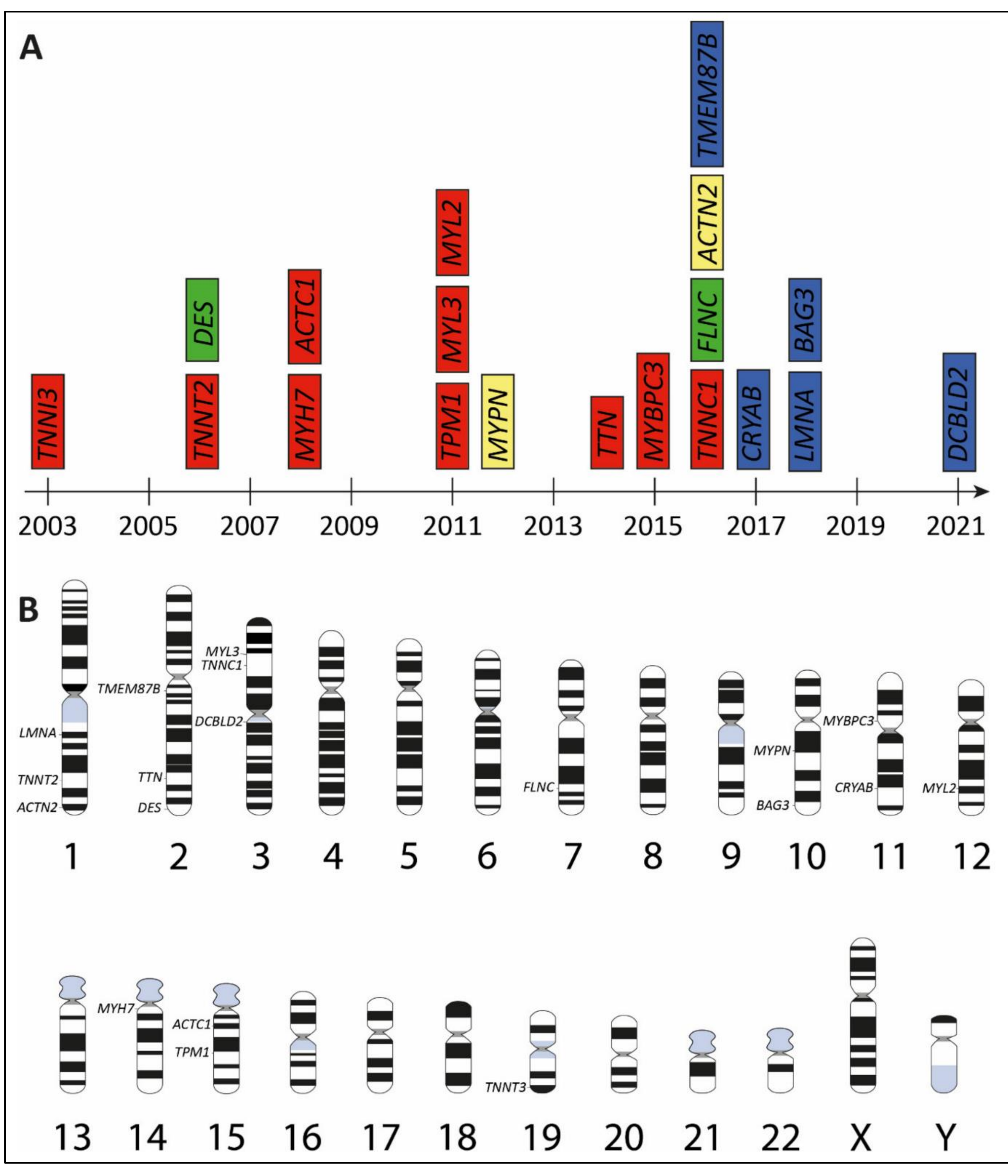


Figure 5.2: Overview of RCM-related genes. **(A)** *Genes associated with restrictive cardiomyopathy (RCM) according to the year of discovery. Different subcellular localisations are colour-coded (red = sarcomere; green = cytoskeleton; yellow = Z-disc and blue = others).* ***(B)*** *Chromosomal location of RCM-associated genes (352).*

5.1.8 Therapeutics

Treatment of RCM involves addressing the root cause, if it is known, and managing HF. Medications such as Angiotensin-converting enzyme (ACE) inhibitors, beta-blockers and diuretics play a vital role in managing heart-related health conditions (174, 175). Diuretics are the primary treatment for reducing volume overload (346). However, managing volume status can be challenging in patients with RCM, as they rely on high filling pressures to maintain cardiac output, and excessive diuresis may cause tissue hypoperfusion (346, 357). The use of β-blockers or calcium channel blockers to increase filling time or manage arrhythmias should be carefully introduced, as some patients may not tolerate them. ACE inhibitors and angiotensin II receptor blockers may also be considered, but their benefits are not well established and may not be well tolerated. Anticoagulation is necessary for patients with atrial fibrillation, mural thrombus, or evidence of systemic embolization, and it may be helpful for most patients due to the tendency for thrombus formation in the left atrial appendage (346, 358). Advanced HF therapies, such as heart transplantation, may be helpful for selected patients. Left ventricular assist device (LVAD) therapy may be particularly suitable for patients with RCM as a bridge to transplantation or definitive therapy (346).

Despite advances in animal models and cell line genetic therapies, primary RCM remains one of the least studied cardiomyopathies. The objective of gene therapy for RCM disease is to cease the production of the pathogenic protein and aim for organ recovery, leading to better quality of life and prolonged survival (346). For example, a proof concept for a therapeutic approach to reduce transthyretin production in the peripheral heart has been achieved by targeting messenger RNA (mRNA) transcribed from the disease-causing gene (359). The researchers demonstrated that the deposition of hepatocyte-derived transthyretin amyloid in the peripheral heart that leads to amyloidosis or RCM could be reduced in both mutant and nonmutant forms of transthyretin.

5.1.9 RCM Genetic Modifiers

Although the prevalence of RCM is unknown, studies have identified approximately 19 genes that contribute to the disease (352). These genes, which encode for sarcomere, cytoskeleton, or Z-disc proteins, are located on autosomes and typically show mutations inherited in an AD mode or appear as *de novo*. It is worth noting that RCM shares genetic overlap with other forms of cardiomyopathies, and additional genetic modifiers may influence various phenotypical

differences (352, 360). For example, some rare *TTN* missense variants may have a modifier effect on the RCM phenotype, suggesting a recessive inheritance pattern (360). Factors such as intronic polymorphisms or mutations that impact gene expression or interactions with protein control components can cause protein aggregations that disrupt sarcomeric structure and lead to contractile dysfunction (331). These factors can ultimately result in increased ventricle stiffness and a greater risk of malignant arrhythmia, and the effects can be further amplified by cell-to-cell imbalances (331, 352).

5.2 **METHODS**

The methods discussed in Chapter 2 are again briefly described in Section 6.2 for the RCM cohort.

5.2.1 Ethical Compliance

We obtained ethical clearance from the Human Research Ethics Committees (HREC) at the (FHS (UCT (HREC 009/2020). The participating South Africa and Mozambique centres provided additional clearance for the main IMHOTEP study. All participants were informed before signing consent (and assent) to join the IMHOTEP study.

5.2.2 Study Population

Forty-three (43) unrelated probands clinically diagnosed with RCM were recruited from February 2018 to October 2021. The probands received clinical examinations performed by at least a cardiologist at the clinics participating in this study across South Africa and Mozambique. The clinical investigations included baseline assessment, complete medical history, pedigree, ECG and transthoracic echocardiography. For the clinical diagnostic criteria, the European Society of Cardiology.

5.2.3 DNA Extraction and Quality Control

Genomic DNAs were extracted from probands' buffy or peripheral whole-blood leukocytes following the published manufacturer's methods for Puregene Blood Core Kit C (Appendix C) or PAXgene Blood DNA kit (Appendix D), respectively. We assessed the integrity, quality, and quantity of the gDNA samples using a calibrated NanoDrop^{TM} 2000 spectrophotometer (Thermo SCIENTIFIC) with software v2.6 on a Windows 10 computer. We used the values to compute the various concentrations. We checked the samples' intactness with a 1% agarose gel electrophoresis programmed to run at 110V for 75 minutes. Each gel was visualised in a Uvitec Xplorer D55 gel doc running on an Xplorer 1D software version 15.08 (Uvitec Cambridge, Cambridge, UK). A dilution was made for a portion of good-quality samples to a volume of 200 uL with a concentration of 120ng/uL. The aliquots of 75uL each were couriered to the OMGL for targeted NGS.

5.2.4 Targeted Sequencing and Variant Calling

We utilised a list of 38 CMO disease-causing genes used by the (OMGL, UK, to diagnose the disease. Using the Haloplex and Twist NGS kits, OMGL sequenced the gDNA of the probands for 38 specific genes associated with DCM and HCM including *ACTC1*, *ACTN2*, *BAG3*, *CSRP3*, *DES*, *DSP*, *DMD*, *FLNC*, *GLA*, *LAMP2*, *LMNA*, *MYBPC3*, *MYH7*, *MYL2*, *MYL3*, *PLN*, *PRKAG2*, *RBM20*, *SCN5A*, *TNNI3*, *TNNC1*, *TNNT2*, *TPM1*, *TTR*, *TTN*, *VCL*, and other DCM and HCM-related genes. The complete list of genes is attached in Appendix F. The gDNA samples were sequenced, targeting these genes at the OMGL using an NGS method. Our Oxford team transformed the data generated into Variant Call Format (VCF) using GATK v3.8-1-0 and Platypus v 0.8.1, and we annotated the vcf files with Ensembl Variants Effect Predictor (VEP) in Cape Town by our local team for analysis.

Through a specific set of filtering criteria, we selected variants for further analysis. We chose 1) variants with a gnomAD minor allele frequency (MAF) of <1% and 2) variants predicted by bioinformatics tools, SIFT and PolyPhen2 to have damaging deleterious or unknown protein function. Alamut™ Visual Plus version 1.5.1 and ClinVar (2022-12-17) were used to analyse the variants further. We further classified the variants according to the refined ACMG guidelines criteria as class 1 (B), class 2 (LB), class 3 (VUS), class 4 (LP) and class 5 (P). We also looked at the variant caller scores for Haplocaller and Platypus for threshold values above 400 and 700, respectively. The researchers of this study visualised all the BAM files for probands carrying class 4 to 5 variants in the Alamut. All identified variants of interest were validated using Sanger sequencing.

5.2.5 Biostatistical Analysis:

We have analysed the descriptive data in R and RStudio (version 4.1.2); the results were presented as $\bar{x} = \pm 1\delta$ for numerical variables and total counts with percentages for categorical variables.

5.3 **RESULTS**

This cohort consisted of 43 unrelated adult probands (≥ 13 years) clinically diagnosed with RCM, and we have reported their data findings in section 6.3.

5.3.1 Baseline characteristics for the adult RCM cohort

Through IMHOTEP, we enrolled 43 unrelated adult (≥ 13 years) probands clinically diagnosed with RCM in this study. We analysed data for the 43 probands and their baseline result in Table 5.2.

Table 5.2: Baseline results for the RCM cohort

LABEL	LEVELS	TOTAL IMHOTEP	TOTAL RCM
NUMBER OF PROBANDS WITH GENETICS DATA ANALYSED, N (%)		690	43 (6.2)
AGE AT DIAGNOSIS (YEARS)	Mean (SD)	31.7 (16.4)	33.0 (14.7)
LEFT VENTRICULAR EJECTION FRACTION (%)	Mean (SD)	36.1 (19.6)	60.8 (16.5)
NEW YORK HEART ASSOCIATION CLASSIFICATION	Class I	83 (15.7)	3 (9.7)
	Class II	162 (30.6)	18 (58.1)
	ND	10 (1.9)	0 (0.0)
	Class III	223 (42.1)	9 (29.0)
	Class IV	52 (9.8)	1 (3.2)
FAMILIAL	No	359 (52.0)	23 (53.5)
	Yes	171 (24.8)	8 (18.6)
	No data	160 (23.2)	12 (27.9)
SEX ASSIGNED AT BIRTH	Female	364 (52.8)	28 (65.1)
	Male	326 (47.2)	15 (34.9)
ANCESTRY	Black-African	381 (55.2)	35 (81.4)
	Mixed	221 (32.0)	5 (11.6)
	White	83 (12.0)	2 (4.7)
	Indian	5 (0.7)	1 (2.3)

VALUES ARE MEAN ± 1SD OR N (%) BASED ON NON-MISSING VALUES OF THE TOTAL ANALYSED.

Overall, for RCM, the baseline clinical data showed that the average age at diagnosis was 33 years (SD 14.7, range 16-67), with 65.1% (28/43) of the probands self-identified as females. Their self-reported ancestry found 81.4% (35/43) BA, 11.6% (5/43) MA, 4.7% (2/43) WA and 2.3% (1/43) IA probands, as shown in Figure 5.3A. Further analysis of the baseline data presented in Figure 5.3B shows that 53.5% (23/43) of the probands reported having non-familial RCM at diagnosis, 27.9% (12/43) of the probands had no FH data, and 18.6% (8/43) had familial RCM.

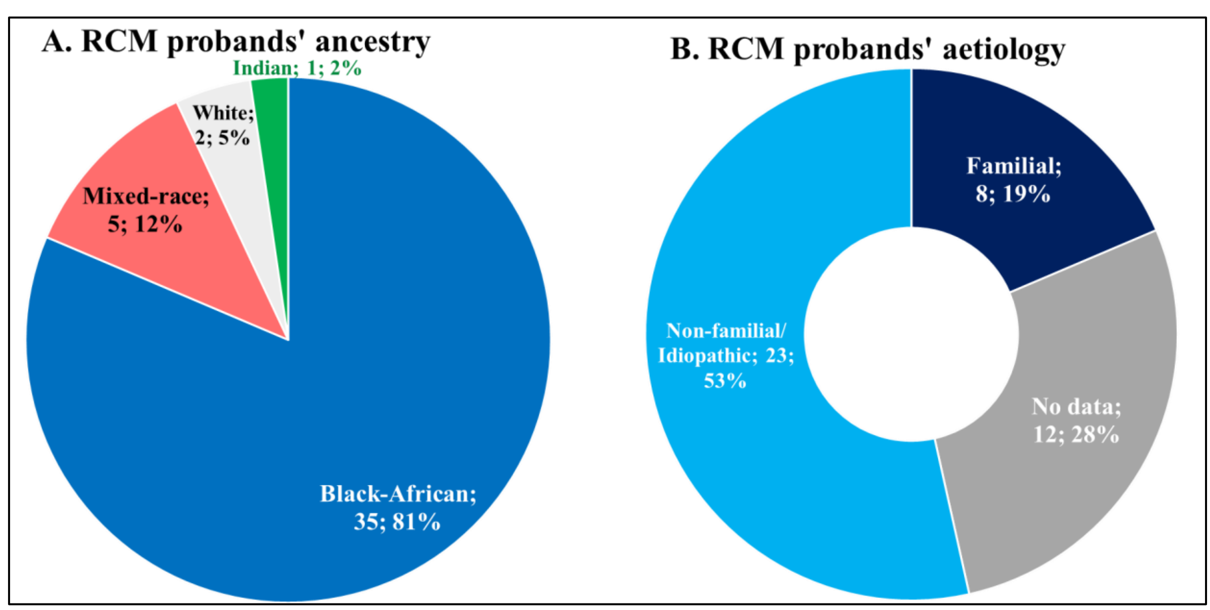


Figure 5.3: RCM probands family history.

5.3.2 Baseline characteristics for the adult RCM according to probands' ancestry

As this is a pilot study, and because South Africa has such a diverse population, we took the opportunity to investigate any potential differences in age, sex assigned at birth, and diagnostic classification across the three dominant populations of South Africa: Black Africa, White and Mixed. However, we used caution in the interpretation because some population groups were in small numbers. We separated the data analysis according to four population groups: at 81.4% (35/43), the Black-African group was the predominant population group in our cohort, the Mixed group followed this at 11.6% (5/43), the White population at 4.6% (2/43) and one Indian proband at 2.3% (1/43), Figure 5.4B.

The 35 BA probands were all recruited from our Mozambique recruitment centre; they were young, with a mean age of 30 years (SD 14.2 years, range 16-63), (Figure 5.4A) with 62.9% (22/36) being female (Figure 5.4B). The aetiology showed that 54.3% (19/35) were non-familial, and only 14% (5/36) had a familial aetiology for their HF history (Figure 5.4D).

In comparison, the 11% (5/43) MA probands recruited were all males with a mean age of 30 (SD 13.9, range 19-57 years). The aetiology showed that 40% (2/5) were non-familial, 40% (2/5) had familial aetiology, and 20% (1/5) had no data for the cause or HF history.

We also recruited 5% (2/43) WA probands (Figure 5.4A), and the data showed that they had a mean age of 48.5 years (SD 18.5, range 30-67), (Figure 5.4B), (which was older than the other

two population groups). The two probands were male and female, and the aetiology showed they had non-familial RCM (Figure 5.4D).

Finally, we recruited one male IA proband, who was diagnosed at the age of 30 with familial RCM.

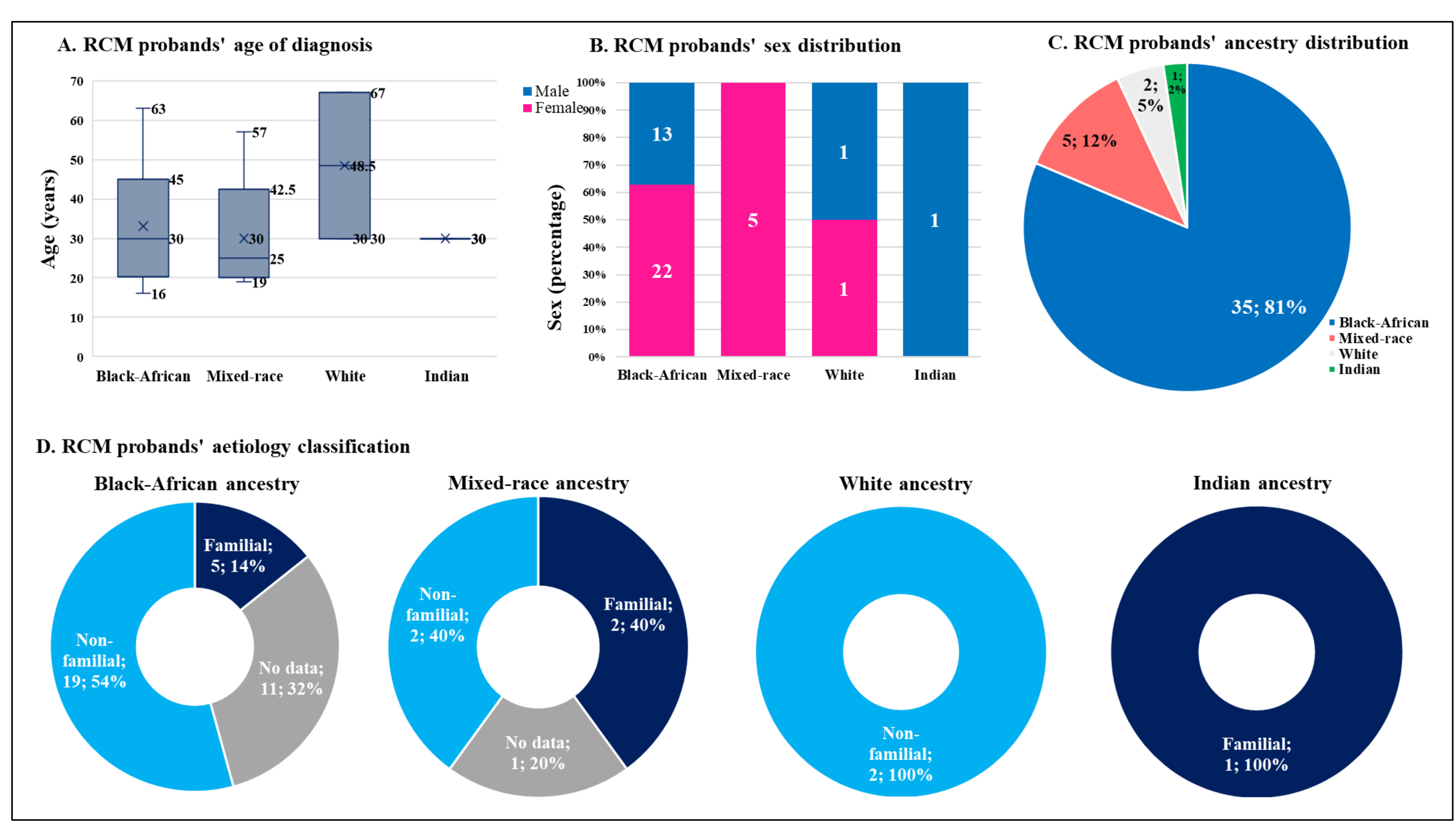


Figure 5.4: RCM probands' age, sex assigned at birth and aetiology distribution. *Graphs showing comparisons between adult-onset RCM population groups regarding age, sex assigned at birth and diagnostic classification.* ***A.*** *Population groups vs age of diagnosis in RCM;* ***B.*** *Population groups vs sex distribution;* ***C.*** *Population groups vs ancestry classification;* ***D.*** *Population groups vs aetiology/diagnostic classification.*

5.3.3 Targeted sequencing result for the adult RCM cohort

We analysed targeted NGS data for the 43 RCM probands and presented the result according to the ACMG guidelines in Table 5.3. Whereas SS Table 5 shows the fitted logistic regression model for adult RCM probands. ACMG classes 4 and 5 variants are positive, class 3 is uncertain, and classes 1 and 2 are negative targeted NGS results. In the following section, we present the genetic result of class 3, 4 and 5 variants only because the class 1 and 2 (negative) variants are beyond the scope of this thesis.

Table 5.3: Baseline results for the RCM cohort according to ACMG classification

LABEL	LEVELS	TOTAL	POSITIVE	UNCERTAIN	NEGATIVE	P-VALUE
RCM diagnostic yield, n (%)		43	4 (9.3)	21 (48.8)	18 (41.9)	
AGE	Mean (SD)	33.0 (14.7)	26.0 (3.9)	35.8 (14.7)	31.2 (15.9)	0.386
LVEF	Mean (SD)	60.8 (16.5)	53.7 (8.1)	60.1 (18.5)	63.2 (15.8)	0.669
NYHA						0.887
	Class I	3 (9.7)	0 (0.0)	1 (6.7)	2 (15.4)	0.571
	Class II	18 (58.1)	2 (66.7)	8 (53.3)	8 (61.5)	0.485
	Class III	9 (29.0)	1 (33.3)	5 (33.3)	3 (23.1)	0.327
	Class IV	1 (3.2)	0 (0.0)	1 (6.7)	0 (0.0)	0.994
FAMILIAL*						0.863
	Yes	8 (18.6)	1 (25.0)	5 (23.8)	2 (11.1)	0.178
	No	23 (53.5)	2 (50.0)	10 (47.6)	11 (61.1)	0.270
	No data	12 (27.9)	1 (25.0)	6 (28.6)	5 (27.8)	0.448
SEX*						0.115
	Female	28 (65.1)	4 (100.0)	15 (71.4)	9 (50.0)	0.065
	Male	15 (34.9)	0 (0.0)	6 (28.6)	9 (50.0)	0.083
ANCESTRY*						0.002
	Black-African	35 (81.4)	1 (25.0)	19 (90.5)	15 (83.3)	0.400
	Mixed	5 (11.6)	3 (75.0)	1 (4.8)	1 (5.6)	0.347
	White	2 (4.7)	0 (0.0)	0 (0.0)	2 (11.1)	0.995
	Indian	1 (2.3)	0 (0.0)	1 (4.8)	0 (0.0)	0.997

VALUES ARE MEAN ± 1SD OR N (%) BASED ON NON-MISSING VALUES OF THE TOTAL ANALYSED:43. POSITIVE-CLASS 4 & 5, UNCERTAIN-CLASS 3, AND NEGATIVE-CLASS 1&2: LVEF-LEFT VENTRICULAR EJECTION FRACTION, NYHA-NEW YORK HEART ASSOCIATION. ASTERISKS (*)-SELF-REPORTED.

Overall, targeted sequencing identified 9% (4/43) RCM probands carrying four class 4 and 5 variants, 48.8% (21/43) carrying 29 class 3 variants and 41.9% (18/43) of the probands with class 1 and 2 variants. The RCM diagnostic yield and the number of variants identified within various classes 3, 4, and 5 are summarised in Figures 5.5A and 5.5B, respectively. Figure 5.5B also presents an overview of the distribution of novel and reported variants across classes 3, 4 and 5 for this RCM cohort. This graph highlights that in our RCM cohort, a significant proportion of variants, especially class 3, may require further evaluation to determine if they are P, B or possible modifiers. Figure 5.5B also illustrates the low class 4 and 5 variants yield across the subpopulation groups and the large yield of class 3 variants, especially in both the BA and MA population groups. We first investigated the probands that were class 4 and 5

variant carriers (positive) and followed this up with a report on class 3 variants (uncertain) and their possible modifying effects on RCM.

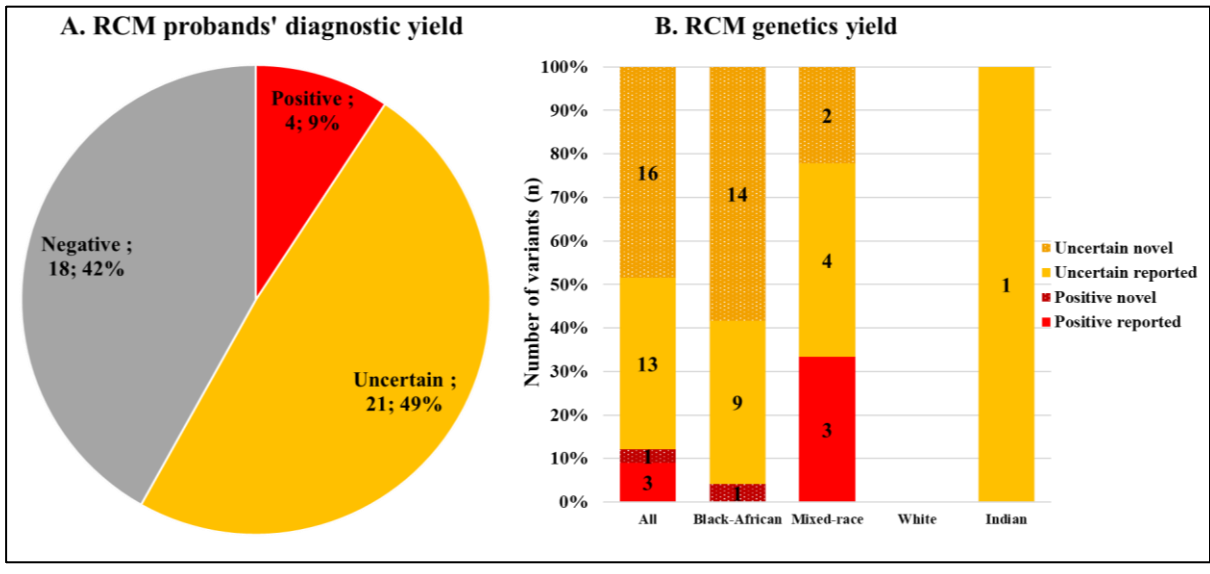


Figure 5.5: RCM probands' diagnostic yield. ***A.*** *Pie chart indicating the diagnostic yield for the adult RCM cohort,* ***B.*** *Bar graph showing the count of probands with variants grouped as novel vs reported within the RCM cohort.*

5.3.3.1 Positive (class 4 and 5) variants in the adult-onset RCM cohort

Overall, we report a genetic yield of 9% (4/43) as we found only four class 4 and 5 variants in four probands: three Mixed and one Black-African proband. We note that 91% of the cohort was mutation negative or uncertain: 39 probands included 21 with only class 3 variants and 18 with only class 1-2 variants. Although most (81.4%) of the RCM probands enrolled were of BA, about 97% (34/35) did not carry any class 4 or 5 variant.

5.3.3.1.1 Class 4 and 5 variants carriers' clinical characteristics

When we explored the combined baseline data for the probands positive group (class 4 and 5 variants carriers), we found that the three probands with class 5 were diagnosed at the ages of 21, 25 and 28 years old, and the one proband with a class 4 variant was 30 years old. Their combined mean (SD) age at diagnosis was 26.0 (3.9), much younger compared to the cohort's mean (SD) age of 33.0 (14.7) years. We also found that the four probands were females: three MA probands and one BA proband, as shown in Figure 5.6. The three MA probands carried

class 5 variants, and one (33.3%) had a familial RCM at diagnosis. In contrast, the one BA proband carrying the class 4 variant was diagnosed with a non-familial RCM.

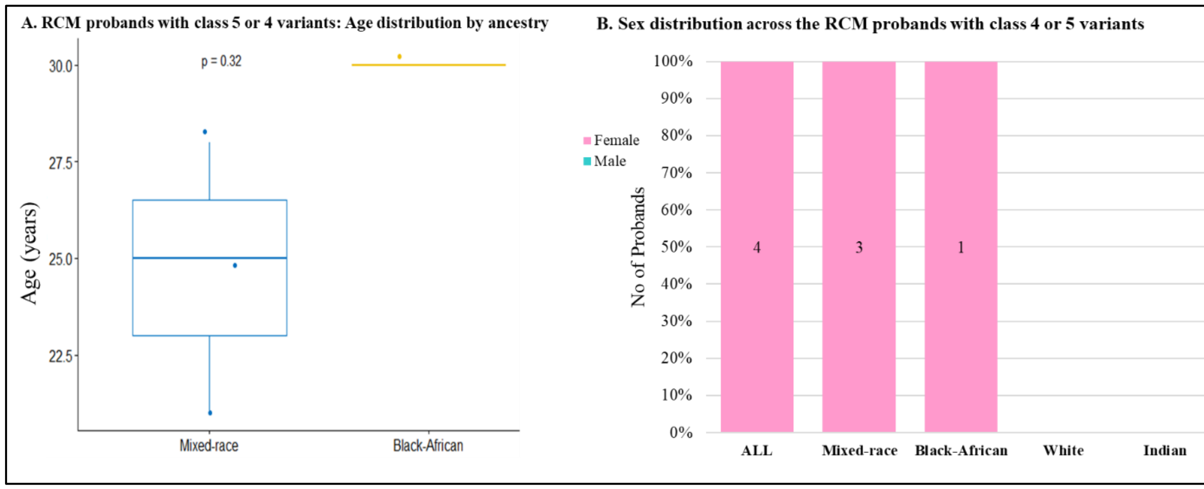


Figure 5.6: Age and sex distribution for RCM class 4 and 5 variant carriers. ***A.*** *Summary of sex assigned at birth,* ***B.*** *Distribution of males and females across the populations in RCM.*

5.3.3.1.2 Class 4 and 5 variants carriers' genetic results

The sunburst in Figure 5.7 shows the genetic distribution for the class 4 and 5 variants for the RCM probands. When we proceeded to look at the number of probands per gene shown in Figure 5.7 B, we noticed that the carrier rate of class 4 and 5 variants ranged from 50% in *DES* to 25% in the *FLNC* and *MHY7* genes. We then analysed the variant according to the population group.

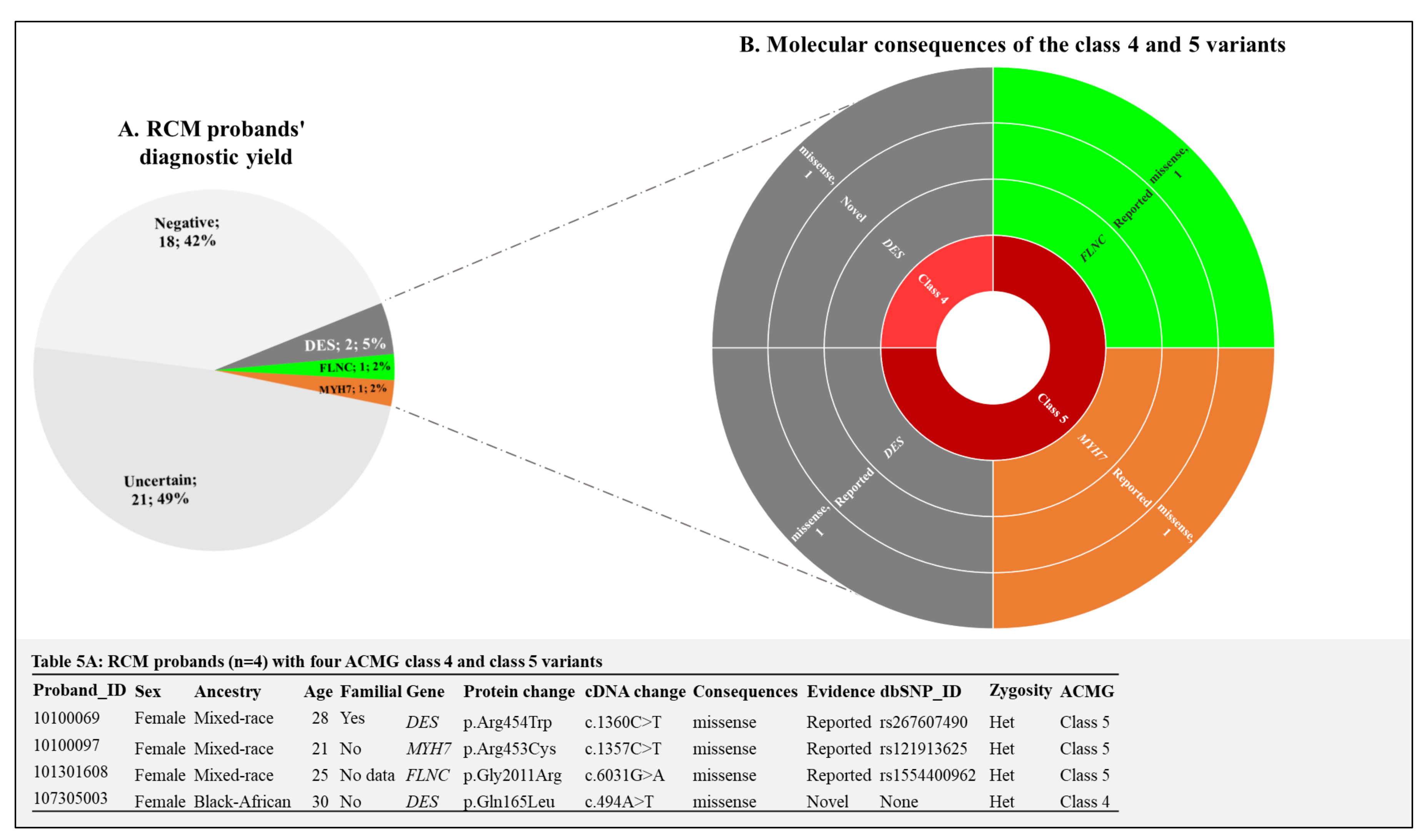


Table 5A: RCM probands (n=4) with four ACMG class 4 and class 5 variants

Proband_ID	Sex	Ancestry	Age	Familial	Gene	Protein change	cDNA change	Consequences	Evidence	dbSNP_ID	Zygosity	ACMG
10100069	Female	Mixed-race	28	Yes	*DES*	p.Arg454Trp	c.1360C>T	missense	Reported	rs267607490	Het	Class 5
10100097	Female	Mixed-race	21	No	*MYH7*	p.Arg453Cys	c.1357C>T	missense	Reported	rs121913625	Het	Class 5
101301608	Female	Mixed-race	25	No data	*FLNC*	p.Gly2011Arg	c.6031G>A	missense	Reported	rs1554400962	Het	Class 5
107305003	Female	Black-African	30	No	*DES*	p.Gln165Leu	c.494A>T	missense	Novel	None	Het	Class 4

Figure 5.7: Genes found in RCM cohort with class 4 and/or class 5 variants. ***A.*** *Targeted NGS result; symbols of genes with class 4 or 5 variants; Uncertain-probands with only class 3 variants; Negative-probands with only classes 1 &2 variants for adult-onset RCM cohort,* ***B.*** *Sunburst chart for the three genes found in adult-onset RCM cohorts with class 4 or 5 variants.* **Table 5A:** Adult-onset RCM probands (n=4) with four ACMG class 4 and 5 variants.

5.3.3.1.2.1 Class 4 and 5 variants carriers of Mixed ancestry (MA)

Of the five MA probands recruited, we found three (60%) probands positive for three reported class 5 variants. The probands were female and were diagnosed at ages 21, 25, and 28, respectively. Two of the three probands showed a non-familial or idiopathic diagnosis without genetic testing and FH information (Figure 5.7, Table 5A). The three class 5 variants they carried are missense mutations in *DES* c.1360C>T, *FLNC* c.6031G>A and *MYH7* c.1357C>T genes.

5.3.3.1.2.1.1 DES c.1360C>T

Proband (ID 10100069) is a 28-year-old MA female classified as NYHA class II, with an ejection fraction of 61%. There was a FH of the disease and targeted sequencing identified the known *DES* c.1360C>T (p.Arg454Trp) variant. This variant alters the carboxy-terminal tail domain of desmin to create an abnormal protein.

5.3.3.1.2.1.2 MYH7 c.1357C>T

Proband (ID 10100097) is a 21-year-old female classified as NYHA class III, with an ejection fraction of 45%. The proband had no FH of HF and targeted sequencing identified the known *MYH7* c.1357C>T (p.Arg453Cys) variant. The *MYH7* c.1357C>T variant replaces an Arginine amino acid residue 453 with Cysteine in the S1 regulatory domain. Arginine is an amino acid with a positive charge and a relatively large side chain, while Cysteine is a neutral amino acid with a small side chain. Cysteine's ability to form disulfide bonds can cause structural changes in the S1 domain.

5.3.3.1.2.1.3 FLNC c.6031G>A

Proband (ID 101301608) is a 25-year-old female of MA carrying an *FLNC* c.6031G>A (p.Gly2011Arg) variant. The variant is in the evolutionary conserved R18 Ig-loop of the ROD 2 domain in exon 37 of the filamin protein. The variant has been found to be disease-causing by several *in silico* prediction tools (e.g., Align GVGD: Class C65, CADD: 27.9, Mutation taster: Deleterious, PolyPhen2: Probably damaging, SIFT: Deleterious).

5.3.3.1.2.2 Class 4 and 5 variants carriers of Black-African ancestry (BA)

Out of 35 BA probands recruited, we found only one female proband (2.9%) diagnosed at 30 years with a novel missense class 4 variant in the desmin (*DES*) gene, whereas 87.1% (34/35) of the BA probands were negative.

5.3.3.1.2.2.1 DES c.494A>T (p. Gln165Leu) variant

Proband (ID 107305003) is a 30-year-old female classified as NYHA class II, with an ejection fraction of 55%. There was no FH of the disease, and targeted sequencing identified the novel *DES* c.494A>T (p.Gln165Leu) variant likely to cause the disease. The variant was classified as LP based on several *in silico* prediction tools ratings (e.g., Align GVGD: Class C65, CADD: 29.3, Mutation taster: Deleterious, PolyPhen2: Probably damaging and SIFT: Deleterious, REVEL score = 0.78). The c.494A>T variant replaces a polar Glutamine molecule with a non-polar Leucine molecule at amino acid position 165. The physicochemical difference between the two amino acids is high, and the polar Glutamine molecule has remained evolutionarily conserved in 12 species of vertebrates assessed. The *DES* c.494A>T variant is expected to alter the head domain of desmin to create an abnormal protein prone to misfolding and aggregation.

Summarily, the RCM cohort yields the least positive targeted sequencing results, as shown in Figure 5.8. the targeted sequencing found 60% (3/5) MA and 2.9% of the BA probands positive. The IA probands had uncertain while the WA probands had negative targeted sequencing results.

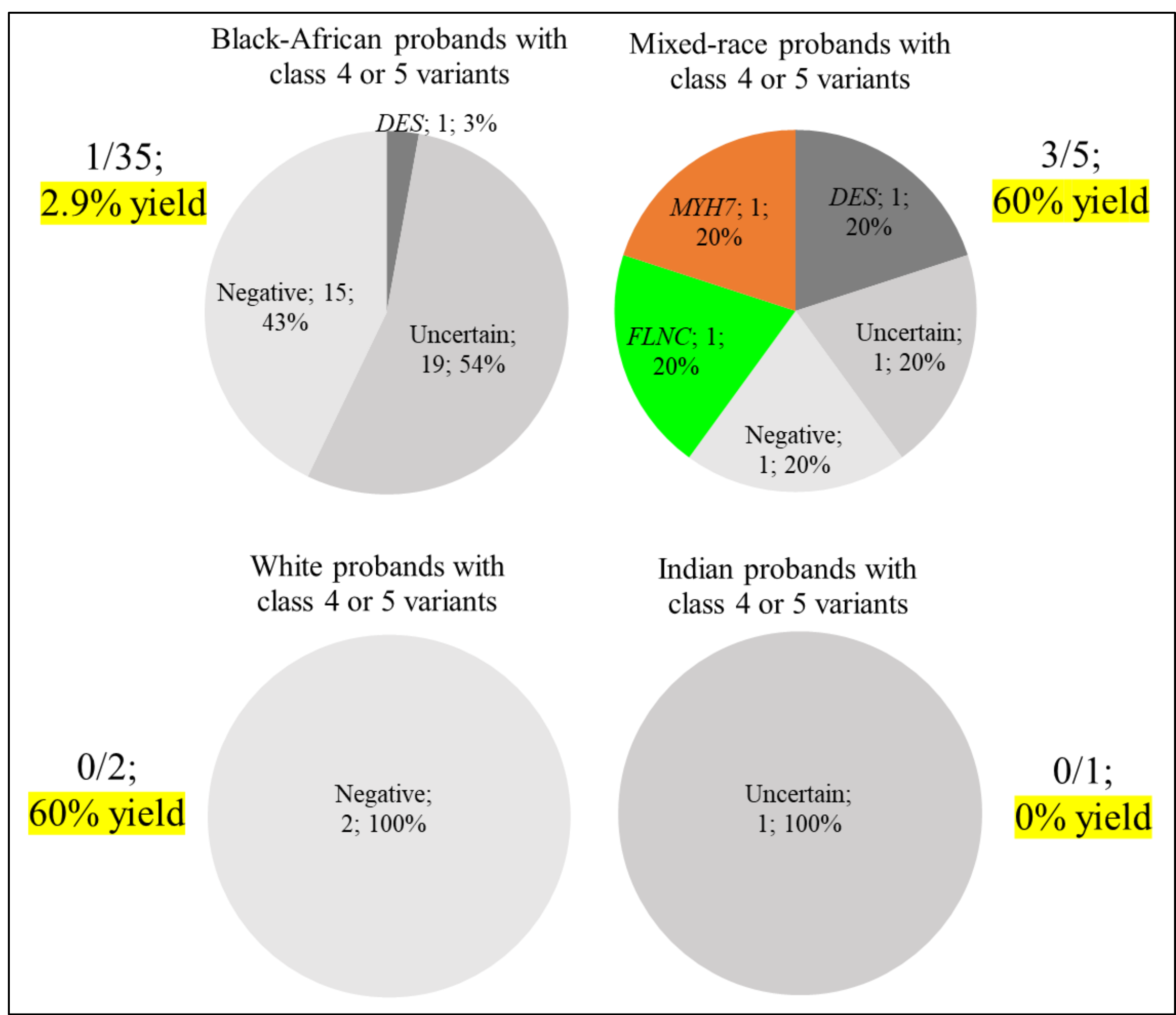


Figure 5.8: Summary of class 4 and 5 diagnostic yields for the RCM population groups.

5.3.3.1.3 Family segregation of class 4 and 5 variants

First-degree relatives of class 4 and 5 variant carriers were identified using pedigree charts and screened for the genetic mutations found in the proband. We screened 11 first-degree relatives of two genotype-positive RCM probands (ID 10100069 and 101301608), as shown in Figures 5.9 and 5.10.

5.3.3.1.3.1 Family of proband ID 10100069

We screened the father (I:2) and brother (II:2) of the proband (II:3) with ID 10100069, carrying *DES* c.1360C>T class 5 variant. The mutation screening result and the summarised family's medical history in Figure 5.9 show that the affected brother (II:2) also carried the *DES* c.1360C>T variant.

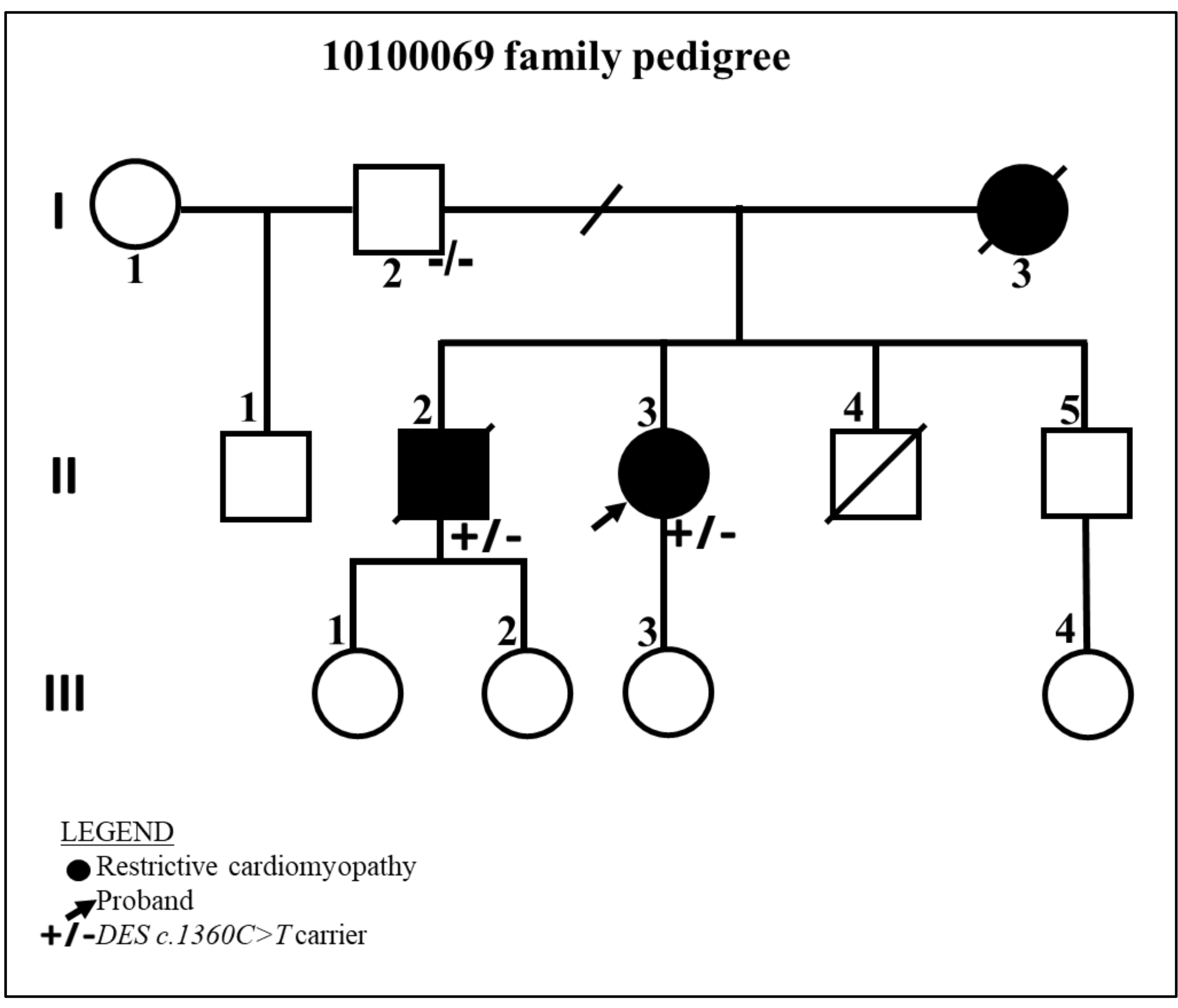


Figure 5.9: Family segregation of *DES* c.1360C>T variant. *The proband is marked with an arrow, 10100069.1. Two relatives were screened. 10100069.2 is positive for DES, and 10100069.3 is negative for DES.*

5.3.3.1.3.2 Family of proband ID 101301608

The proband (IV:5) with study ID 101301608 and six relatives were included for targeted NGS. Through targeted NGS analysis, we found that the proband and four affected relatives (II:3, III:3, III:5, III:7) were positive for a reported *FLNC* c.6031G>A. Also, III:2 without the *FLNC* c.6031G>A variant carried a *TTN* c.31207+2T>C class 5 splice-donor heterozygous variant. Figure 5.10 shows the 101301608-family pedigree chart. Our genetic counsellor will contact the genotype-positive relatives to relay the information and guide the affected individuals.

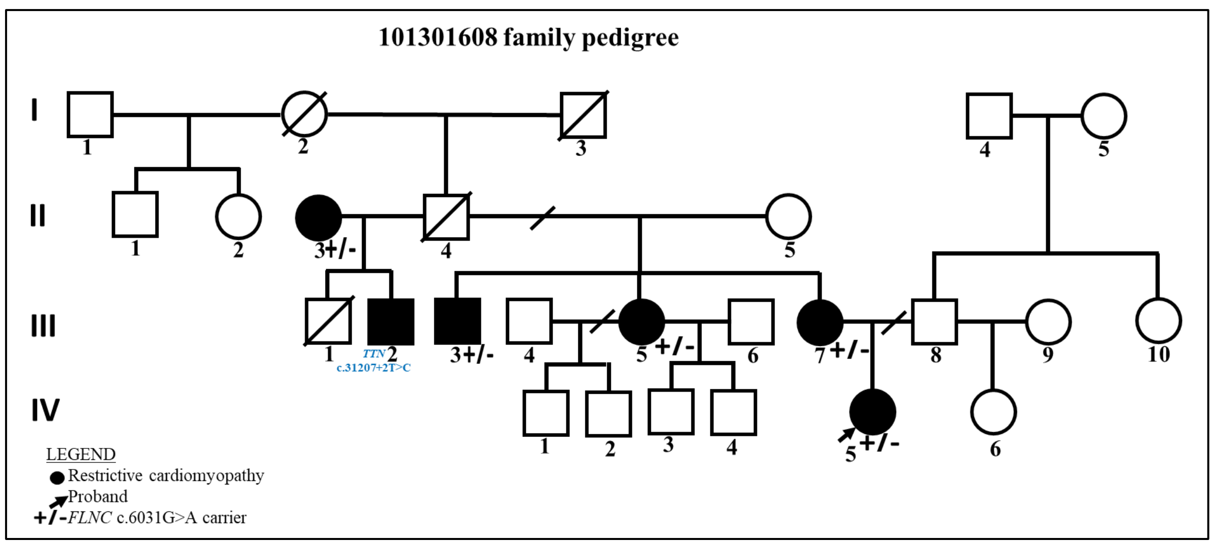


Figure 5.10: Family segregation of *FLNC* c.6031G>A variant. *Pedigree chart for proband ID 101301608 and relatives. The proband found with FLNC c.6031G>A is marked with an arrow.*

5.3.3.2 Probands with variants of uncertain significance (class 3)

5.3.3.2.1 Baseline characteristics

We found 21 RCM probands carrying only class 3 variants: 19 BA, one MA, and one Indian proband, as shown in Table 6.2. Data analysis for the probands carrying only class 3 variants showed that they were diagnosed at a mean age of 35.8 years (SD 14.7, range 16-63). There was thus no significant difference between the overall adult-onset RCM age of diagnoses and probands positive for only class 3 variants. We noted that about 67% (14/21) of the RCM class 3 variant carriers were BA female probands. Hence, we looked into the genes with class 3 variants if being a female of BA is a possible modifier for RCM.

5.3.3.2.2 Diagnostic findings

5.3.3.2.3 Class 3 variants carriers' genetics findings

We are aware of the dangers of over-interpreting class 3 variants but attempted to do so as we wanted to identify "hot" variants that could act as modifiers of RCM. Targeted resequencing identified 21 RCM probands carrying eight genes harboring 29 class 3 variants listed in SS Table 2.

We found that 55.2% (16/29) of the class 3 variants identified were novel mutations, with 33.3% (7/21) of the probands carrying both novel and reported VUSs. Except for one *TTN*

c.32471-1G>A (PVS1 Strong, PM2 Moderate) splice acceptor variant, the class 3 variants were missense mutations. The *TTN* c.32471-1G>A variant was found together with *MYH7* c.2282C>A (PM2 Moderate, PP3 Moderate) missense variant in one female proband of MA diagnosed at 57 years as having familial RCM. We described these two variants as the potential cause of RCM in this proband.

The 16 novel missense class 3 variants identified were in four (*DMD*, *DSP*, *SCN5A* and *TTN*) genes. The *TTN* gene harbours the majority, 68.8% (11/16), of the novel class 3 variants, followed by *DMD* with three and the other two genes had one novel variant each. For details of all the class 3 variants information, including the genomic coordinates, SIFT, PolyPhen2, CADD, population frequencies, clinical annotation, and others, see Supplementary Table 1.

5.3.3.3 Summary of ACMG classes 1-5 variants for the adult-onset RCM

The number of RCM probands found with or without variants classified according to the ACMG five-tier system is summarised in Figure 5.11. We found four probands carrying class 4 or 5 variants: three probands carrying class 5 variants and one proband carrying class 4 variants. While 21 probands carried only class 3 variants and 18 probands carried class 1 and/or 2 variants.

We found only three genes (*DES*, *FLNC* and *MYH7*) in the four probands, with the *DES* harbouring two variants. The four variants were missense mutations, and three were reported in the literature.

Figure 5.11 also shows that the uncertain and negative genetic results were high in the BA probands, while most MA probands had positive results. Also, there is a high rate of recurring class 3 amongst the probands of BA, with the 19 BA probands carrying 23 class 3 variants. Besides, the BA and MA probands had a high yield of class 3 variants.

No further analysis was required for the class 1-2 variants identified in the adult RCM probands.

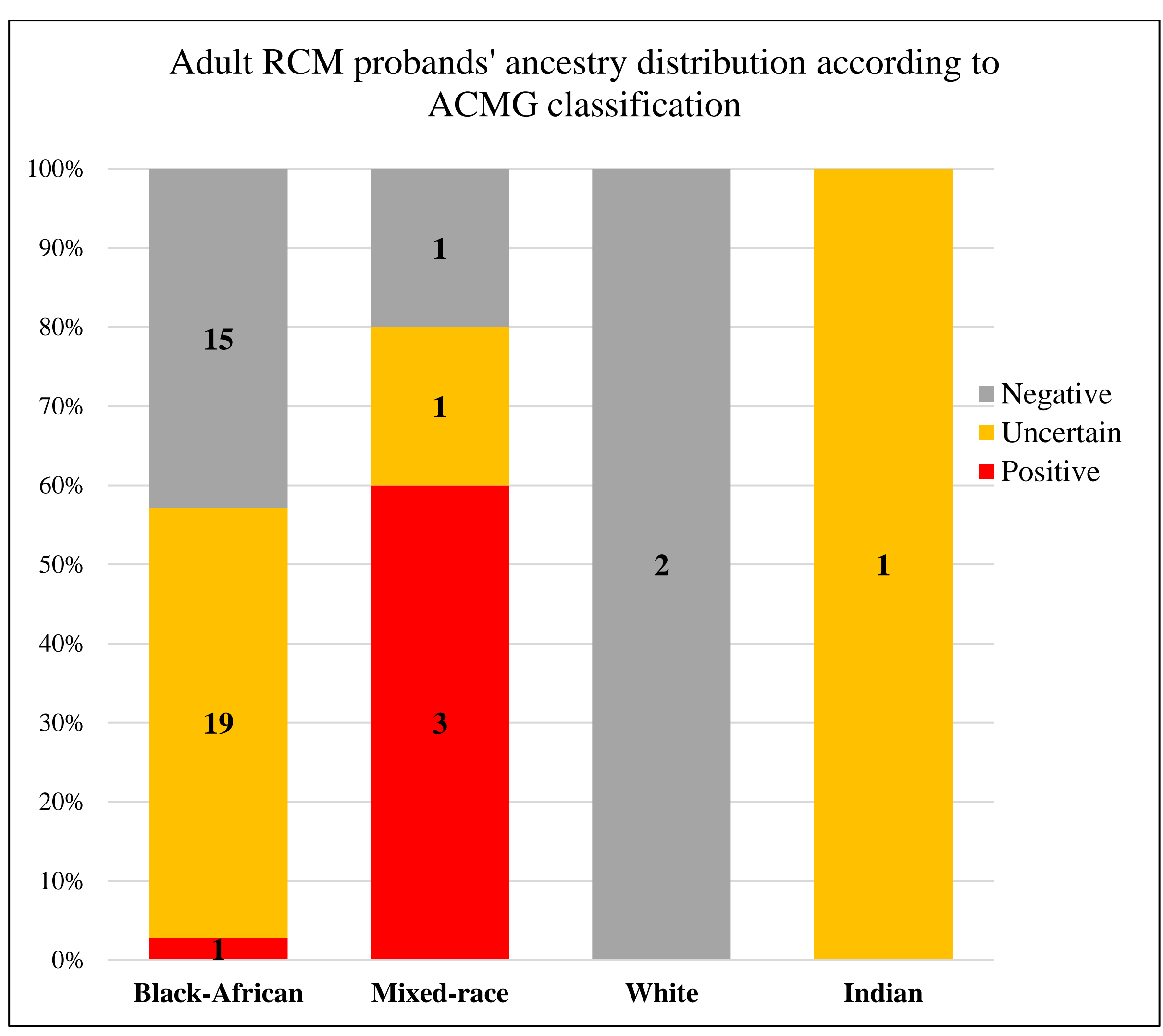


Figure 5.11: Summary of adult RCM classes 1 to 5 variants carriers. *The adult-onset RCM probands were categorised according to the cohort's population ancestry. Positive-class 4 & 5, Uncertain-class 3, and Negative-class 1&2.*

5.4 **DISCUSSION**

In this RCM cohort, we used a comprehensive routine targeted sequencing method to identify potential disease-causing variants in an RCM cohort of 43 unrelated adult probands (≥ 13 years). The participants were of various ancestry, including 35 Black Africans (Mozambique), five Mixed (South Africa), two White (South Africa), and one Indian (South Africa) ancestry individuals suspected of having heart complications and referred to the public tertiary cardiac clinics.

The participants with RCM recruited from Mozambique were all found to have endomyocardial fibrosis (EMF), a disease frequently observed in equatorial regions and SSA (56, 361). EMF is a common cause of RCM, and a population screening in an endemic area of Mozambique suggests that it may have a broader disease spectrum and higher prevalence than previously known (362). In equatorial regions, EMF is responsible for about 20% of cases of HF and 15% of cardiac deaths in equatorial Africa (46). RCM caused by EMF typically arises in areas with limited resources, where poverty plays a significant role in the development of the disease. As a result, diagnosis is often delayed, compounded by limited access to proper medical and surgical treatment (361).

Environmental factors such as warm climate, poverty, diet, allergens, and viral and parasitic infections may impact the development of EMF (138). However, the cause of EMF has been debated, with various factors suggested, including parasitic infection (e.g., chronic helminths infection), malnutrition, a protein-poor diet, magnesium deficiency, high vitamin D, serotonin, cassava plant toxicity (which contains linamarin, a compound that releases cyanide), eosinophilia, immunological and genetic factors (56, 363). EMF also has a hereditary component linked to environmental factors (56). The risk factors elicit an inflammatory process leading to progressive endomyocardial damage and scarring (46, 138). The endomyocardial damage and scarring may occur in one or both ventricles, leading to restrictive ventricular filling (56, 363).

In this chapter, we describe the aetiology and genetics of RCM in South Africa and Mozambique.

5.4.1 RCM probands' baseline findings

Among the 43 RCM probands initially recruited, the NGS panel yielded a 9.3% (4/43) positive rate of 60% (3/5) for the Mixed, 3% (1/36) for the Black-African and 0% for the White and Indian ancestry probands. The low yield supports previous findings that familial RCM is a rare heart muscle (346), specifically amongst the probands of BA in this study. The low yield among Black-African probands suggests that the RCM disease burden could be worsened by environmental factors such as limited access to specialised medical care (361, 363). The disease's low yield among BA participants (97%) possibly reflects the EMF phenotype in this cohort (37), and it is possible that the genetic architecture of this cohort is distinct from typical RCM or that the environmental factors are more prevalent (56). Generally, the various ethnic groups in Africa with the most diverse genomes are underrepresented in almost all genetic databases (127).

The mean age at diagnosis for our 43 RCM probands was 33 years, which is an early age of onset for this rare form of adult CMO. Very few studies have reported on adult-onset RCM, which typically occurs around the age of 55 at the time of diagnosis (191). Many participants in this cohort were Black Africans, and the early onset of RCM suggests that there are other contributing factors to the phenotype. Like most RCM cases in the literature (42), 65.1% (28/43) of our RCM participants were recorded as females. Hence, we hypothesised that our probands might have developed the phenotype earlier but had milder signs and symptoms (10, 11).

We found four genotype-positive probands of the 43 RCM participants (three MA and one BA). Our yield is low compared to the adult RCM global rates of about 25% to 54% (191, 348). Unlike in the literature, where most of these adult cohorts with genetic results are from European RCM patients, our probands are mostly endomyocardial fibrosis (EMF) cases. Endomyocardial fibrosis is an important cause of RCM worldwide, which classically presents in childhood and adolescence (348, 361). Few extreme presentations in infancy and older adults have also been reported (96, 191).

Our four genotype-positive probands had a mean age of 26 years at diagnosis. We comprehend that this is a very young age for an RCM cohort compared to what is known for the adult phenotype globally, which is about 50 years for genotype-positive adult RCM patients (191). As the baseline data has shown in other studies, we also found that our four probands were females (42, 132). Some of these authors have also reported a bimodal distribution with the

second peak among women of childbearing age. However, several studies found no sex-specific preponderance (96), whereas others show conflicting higher rates among women in Uganda (364) and among men in Mozambique (361) at diagnosis.

5.4.2 RCM probands' targeted sequencing findings

We used a comprehensive routine targeted NGS panel testing and identified RCM-causing variants in four probands, yielding a diagnostic rate of 9.3% for this cohort. Also, we found 90.7% (39/43) of the cohort to have either negative or uncertain genetic results. The positive result is low compared to what has recently been reported by similar studies with a global yield of > 20% (42, 191). Many participants with RCM were recruited from Mozambique, where they had been diagnosed with EMF.

The gene panel we used for sequencing was specifically designed for inherited common forms of CMOs in European cohorts. Although we conducted variant investigations cautiously, EMF is a common cause of RCM with a less established genetic basis (56, 362). We hypothesise that the interaction between unknown genes, variants, and environmental factors triggers inflammatory processes that cause progressive damage and scarring of the endomyocardium (46, 138). The heart muscle damage and scarring lead to contractile dysfunction and stiffness of the ventricles in the participants we studied (331).

5.4.2.1 Pathogenic (class 4 and 5) variants in the adult-onset RCM

The four (three MA and one BA) probands carried four disease-causing variants, with 75% (3/4) occurring in cytoskeleton genes (*DES* and *FLNC*) and 25% (1/4) in sarcomere gene (*MYH7*). Similar to the literature, RCM-causing variants are mainly found in the cytoskeleton and sarcomere proteins (344, 347).

5.4.2.1.1 The cytoskeleton (*DES* and *FLNC*) genes

Most class 4 and 5 variants altering the cytoskeletal protein structures are associated with cardiovascular system disorders, including cardiac conduction disorders, ventricular arrhythmias, and CMOs (236).

Cardiomyocytes' cytoskeletal proteins network proteins are involved in several cardiac functions, including structural integrity and signalling. Abnormal changes in these proteins' structure could result in the protein forming aggregates within the cardiomyocytes. Malformed proteins can often lead to a weakened cytoskeleton structure, which can cause the death of cardiomyocytes and remodelling.

In this RCM cohort, we identified three class 4 and 5 variants in *DES* (*DES*:c.494A>T (p.Gln165Leu) and c.1360C>T (p.Arg454Trp)) and *FLNC* (*FLNC*:c.6031G>A (p.Gly2011Arg)) genes that could impair cytoskeletal integrity (365, 366) and are discussed below.

5.4.2.1.1.1 The desmin (DES) gene class 4 and 5 variants

5.4.2.1.1.1.1 DES c.494A>T (p.Gln165Leu) variant

The resultant abnormal desmin protein might increase myofibrils' susceptibility to degeneration, disrupting the cytoarchitecture (347, 352). The disruption causes damage and fibrosis in heart muscle cells, leading to stiffness and impaired function during relaxation. Therefore, we hypothesise that the *DES* c.494A>T variant is a potential cause of the observed phenotype in our 30-year-old patient.

Pathogenic variants in the disease-causing desmin (*DES*) gene cause desminopathy and are often *de novo* with most responsible for primary RCM, particularly in patients with familial disease (300, 367). Desminopathy leads to severe skeletal muscle disorders with the formation of inclusion bodies, weakening of the cytoskeletal desmin intermediate filament, disruption of subcellular organisation of organelles and myofibrillar degradation (368). In a meta-analysis study, up to 50%, 60%, and 74% of *DES* class 4 or 5 variant carriers had cardiomyopathy, cardiac conduction disease, and skeletal myopathy, respectively (369).

The involvement of the heart in desminopathy often leads to a high number of complications, such as high-degree conduction blocks, sustained VT, cardiac insufficiency, cardioembolic stroke, or sudden death (367). In many cases, treatment with a pacemaker or implantable cardioverter is required, and in severe cases, heart transplantation may be necessary, such as in 60% and 10% of cases, respectively, reported (367, 370).

5.4.2.1.1.1.2 DES c.1360C>T (p.Arg454Trp) variant

The *DES* c.1360C>T is documented in several patients clinically diagnosed with life-threatening RCM (366, 371). This variation has been observed to cause an earlier onset of symptoms, usually before adulthood, and a severe cardiac phenotype with high penetrance (366). As a result, patients may require a pacemaker or implantable cardioverter defibrillator (ICD) at an earlier age. The c.1360C>T variant in an *in vitro* study impairs the biomechanical properties of desmin and alters myocyte metabolism by affecting the myoblast cytoskeleton (365). Also, abnormalities in the intercalated disc were detected via myocardial tissue immunohistochemistry, including an absence of desmin and reduced expression of desmoplakin, plakophilin 2, and connexin 43 (372).

Like in other families with desminopathy in the literature, there is significant clinical diversity, with incomplete penetrance and variation among patients and families carrying the same *DES* variant (365, 367). Although several RCM patients have been diagnosed with the *DES* c.1360C>T variant (373), we have added significant information for this variant expression.

The wildtype *DES* gene encodes for desmin protein, a muscle-specific class III intermediate filament that crosslinks myofibrils to each other and the plasma cell membrane. The primary function of desmin is to maintain the sarcomere structural integrity by crosslinking the Z-discs to form the myofibrils. Equally important, desmin links the sarcomere to the sarcolemma, nucleus, and mitochondria, thus strengthening the myocardial fibres during activity (371), as shown in Figure 5.12. Classes 4-5 variants capable of altering desmin protein would lead to sarcomere disarray and impaired cytoskeletal integrity. Specifically, abnormal desmin protein aggregation within the cardiomyocytes could cause fibrosis and stiffening that disrupt signal transduction or cause cardiomyocyte apoptosis and remodelling (347).

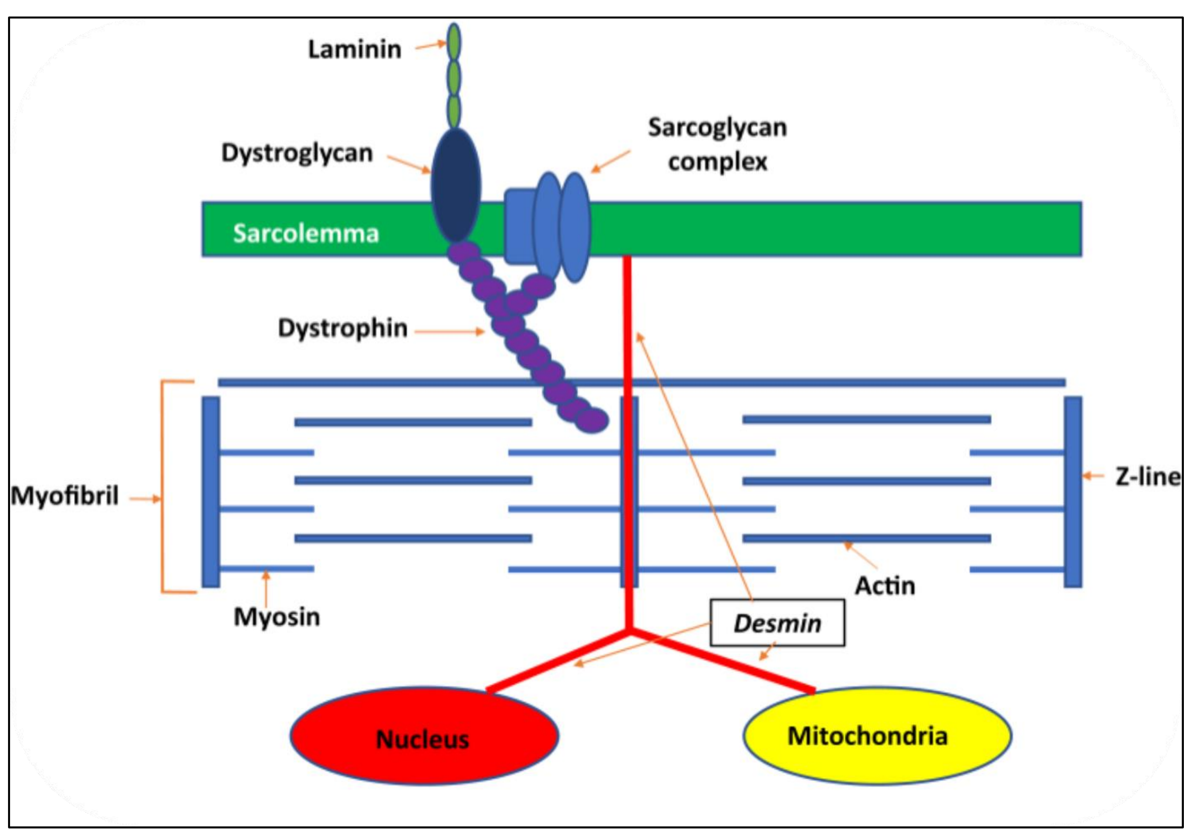


Figure 5.12: Schematic representation of the *DES* gene. *Schematic representation of desmin and other proteins involved in myocyte cytoarchitecture (371).*

5.4.2.1.1.2 The filamin C (*FLNC*) gene class 5 variant

5.4.2.1.1.2.1 FLNC c.6031G>A

The missense *FLNC* c.6031G>A (p.Gly2011Arg) variant is linked to hypertrophic and arrhythmias-related CMOs (107). For example, the reported p.Gly2011Glu variant (on the same codon as p.Gly2011Arg) in a French patient with severe *FLNC*-related structural heart defects and CMO has segregated with disease (374). Notably, missense variants in the *FLNC* ROD 2 domain keep the actin-binding protein domain dissociated by mechanical force, innately lengthening the protein product and making it prone to misfolding (237). The misfolded protein will form aggregates that alter filamin-sarcomere protein interaction, causing cardiomyocyte stress and apoptosis and impairing cardiac diastolic function (331). The main features in RCM-affected hearts include myocardial fibrosis and stiffening, as observed in our proband.

We therefore screened and found four affected relatives positive for the *FLNC* c.6031G>A in this proband's (ID101301608) family. These family members had wide-ranging mixed phenotypes of RCM, HCM and other phenocopy conditions such as the Noonan-like syndrome. The affected members showed clinical symptoms of cardiac failure such as webbed neck,

prominent café au lait on arm, leg and back scoliosis. A heritable RCM phenotype was found in this family, characterised by myocardial ventricular stiffness and hypertrophy, which we have shown to be associated with the *FLNC* c.6031G>A variant.

The *FLNC* gene (one of three related gamma filamin genes) encodes a filamin-C protein that crosslinks the sarcomere actin filaments to the Z-disc in the cytoskeleton of tissues, including the myocardium. The cardiac filamin-C intercalates the Z-discs of the sarcomere onto the cell membrane to form a complex cytoskeletal structure. This complex structure lies on the load-bearing sites of cardiomyocytes to maintain cardiomyocytes' structural stability and mediate crucial cell signalling functions. Specifically, missense variants altering this complex protein's physiological structure would lead to sarcomere disarray and loss of filamin-C protein from the intercalated disc. Thus, impairing the mechanical properties of the protein within the cardiomyocytes and disrupting signal transduction could cause cardiomyocyte apoptosis and remodelling. These pathomechanisms will collectively impair myocardial function and cause stiffening of the myocardium due to its loss of function.

5.4.2.1.2 The sarcomere (*MYH7*) gene variant

5.4.2.1.2.1 *MYH7* c.1357C>T (p.Arg453Cys) variant

The *MYH7* c.1357C>T (p.Arg453Cys) class 5 variant has been documented in several structural heart defects and HCM patients (156, 375). These studies confirmed the highly penetrant *MYH7* c.1357C>T variant in severe HF patients and SCD cases. Also, a functional study has shown that the variant will cause sarcomere disarray and hypo-contractility, impairing the myocardial diastolic function (376).

The *MYH7* gene encodes for a prominent sarcomere component called beta-myosin heavy chain protein, predominantly expressed in the ventricles and essential for cardiac contraction. Alterations in the *MYH7* gene will either increase or decrease the beta-myosin heavy chain protein expression level (331), and a high expression level will make the sarcomere proteins unstable and prone to degradation. In contrast, a low expression level will impair relaxation due to contraction dysfunction. The resulting *MYH7*-related diastolic dysfunction is a characteristic feature of RCM (147), with a prevalence ranging from 2% to 10% (96).

5.4.2.2 Variants of uncertain significance (class 3) found in our adult RCM cohort

All human genomes contain thousands of VUS, and data to support or contradict the clinical significance of a variant can emerge over time. Our current interpretations of VUSs rely heavily on variant frequency in putatively healthy controls in these most European population-scale databases. As these databases expand to become more representative of human genetic diversity, a considerable proportion of VUSs and putatively diagnostic variants are found to be present at frequencies incompatible with causing severe genetic disorders (377). The dangers of over-interpreting genetic variants can be substantial, particularly in Africa, where there is a severe lack of Afrocentric control databases; this is a huge concern as this makes VUS interpretation nearly impossible.

We are aware of the dangers of over-interpreting VUSs but attempted to do so as the VUSs could act as modifiers of disease or, if enough evidence is found, the variants could become a "hot" VUS. We interpreted the class 3 variants in this RCM cohort with caution.

We identified 21 adult RCM probands with class 3 variants, of which 91% (19/21) were of BA, and 73.7% (14/19) were females. Our class 3 variant carriers' demographic finding of the majority being Black females is similar to those of studies from the Northern American continent with participants of African descent (42, 96). However, our data includes new demographic information because American researchers have primarily published studies on paediatric cohorts. In contrast, we have reported on adults, particularly Black-African females of childbearing age. Although these variants have not been robustly linked to RCM, we hypothesise they may have contributed to the disease in our cohort, as RCM can occur as a mixed phenotype. Most of our participants had idiopathic disease at diagnosis.

Most class 3 variants were missense mutations, 96.6% (28/29), except for the *TTN* c.32471-1G>A splice acceptor. Notably, the *TTN* c.32471-1G>A and *MYH7* c.2282C>A variants were likely disease-causing. Interestingly, one proband carried both variants, suggesting possible digenic heterozygosity variants. We further described these two variants found in these genes with known mechanisms for CMOs as possible RCM genetic causes.

5.4.2.3 Digenic variants: *TTN* . and *MYH7* c.2282C>A (p.Thr761Asn) variants

The published *TTN* c.32471-1G>A (378) and *MYH7* c.2282C>A (156) variants are associated with various CMO phenotypes. The documented CMO-related phenotypes are characterised by thickened and stiffened heart muscles (156, 378). Although the *TTN* c.32471-1G>A in the

A-band region of titin protein is published as pathogenic in neurodevelopmental disorders (378), we hypothesise that it might not cause the observed RCM phenotype on its own. Titin A-band truncating variant generally causes the development of cardiac arrhythmias (167, 379); however, the impact of *TTN* c.32471-1G>A is not established. We, therefore, hypothesise that the presence of *MYH7* c.2282C>A variant could be intensifying the sarcomere protein damage in the S1 regulatory myosin binding site of the β-myosin heavy chain protein motor domain that generates force for heart muscle contraction (380). Alterations in this motor domain of the *MYH7* gene have been reported to increase heart muscle stiffness and impair its contractility, leading to cardiac arrhythmias and HF (90). Further, Ntusi and others have described the *MYH7* c.2282C>A (p.Thr761Asn) variant in another South African proband with HCM (156). Moreover, *TTN*tv and *MYH7* digenic mutations increase stress and sarcomere insufficiency (139). A specific feature of RCM is protein aggregation, which leads to contractile dysfunction and stiffness when caused by sarcomeric or sarcomere-associated protein variants (331).

5.5 **CONCLUSION**

The findings from this cohort support previous reports that adult RCM is a rarer form of CMO, and this study achieved a routine diagnostic yield of 9.3% (4/43). We have limited data at this stage to compare the sexes between our cohort's paediatrics and adult probands.

The majority of the probands at diagnosis and the class 4 and 5 variant carriers were females of childbearing age. The participants of BA or MA were commonly affected; however, many carried class 3 variants. Also, we found one proband with two likely genetic modifying variants. No specific gene was prevalent in this RCM cohort because the *MYH7*, *FLNC* and *DES* genes were the primary disease-causing genes harbouring the classes 4-5 variants.

The challenges of RCM in the tropics and SSA highlight the need for increased awareness, improved healthcare infrastructure, and targeted interventions in regions affected by endomyocardial fibrosis. By addressing these issues, researchers can work towards reducing the burden of this disease and improving outcomes for individuals living in low-resource settings impacted by endomyocardial fibrosis.

Chapter 6: Arrhythmogenic Cardiomyopathy (ACM)

6.1 **INTRODUCTION**

6.1.1 Clinical background

Arrhythmogenic cardiomyopathy (ACM) is an inheritable form of (CMO that is characterised clinically by ventricular arrhythmias and (SCD at a young age (381, 382). Although, in its original description, the disease was believed to predominantly involve the right ventricle, biventricular and left-dominant variants, in which the myocardial lesions affect in parallel or even mostly the left ventricle, is nowadays commonly observed (383). The histopathology has revealed islands of surviving myocytes, interspersed with fibrous and fatty tissue in affected individuals, with the right side of their heart appearing like yellowish/whitish parchment due to fibrofatty infiltration and scaring in mild to severe cases (384), as shown in Figure 6.1. The fibrofatty infiltration and scaring process starts as a wavefront phenomenon from the sub-epicardium and extends to the endocardium to become transmural, resulting in cardiomyocyte death (384). The fibro-fatty replacement of the myocardium disrupts the intraventricular conduction of the electric impulse, which results in the delay and onset of the re-entry phenomena, constituting the mechanism of ventricular arrhythmias (382).

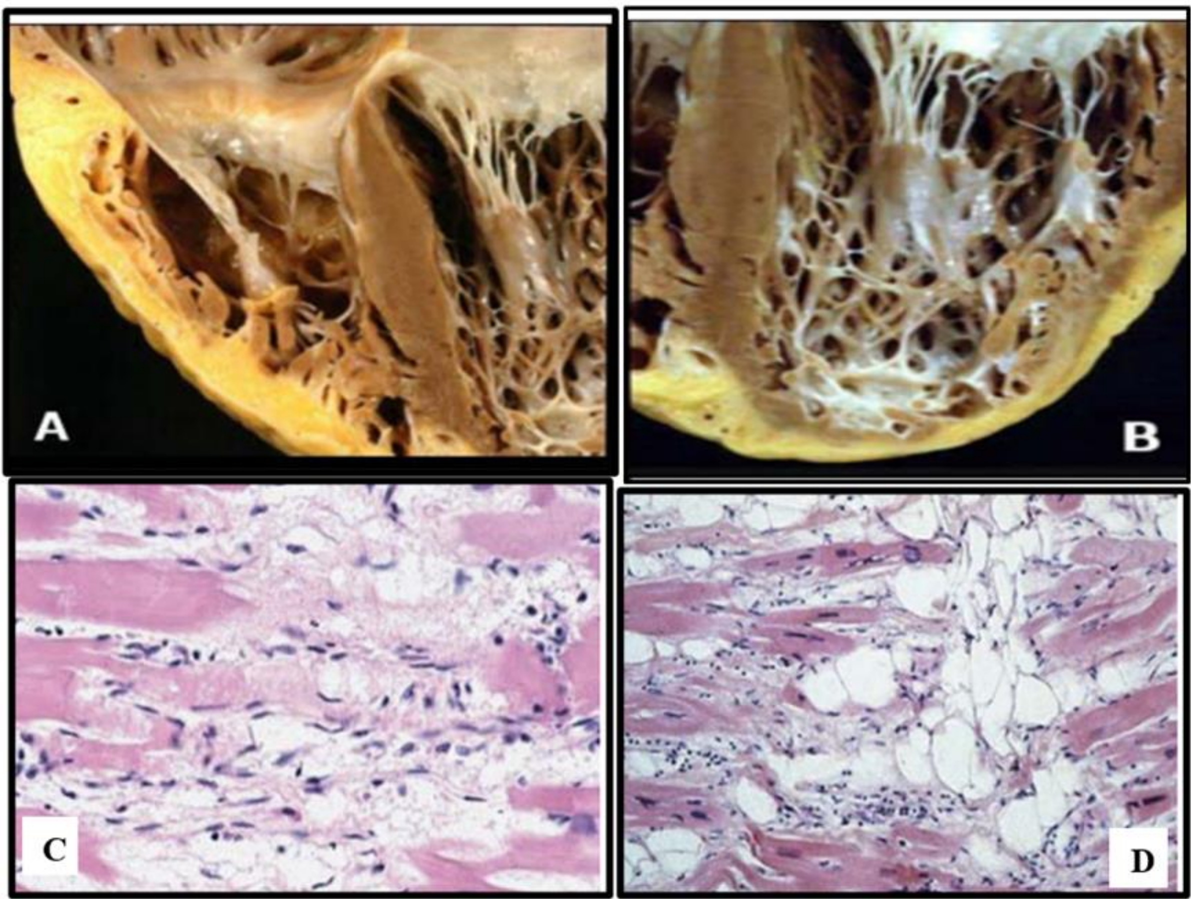


Figure 6.1: A heart sample from an ACM patient. ***A&B,*** *Note the biventricular involvement at a gross examination—ongoing myocyte death with* ***C.*** *early fibrosis and* ***D.*** *adipocytes infiltration (384).*

A definitive diagnosis requires remarkable fat replacement and replacement-type fibrosis and myocyte degenerative changes (384). The myocardial pathological changes may be focal in 20% and diffuse in 80% of patients. Interestingly, most ACM patients have had grossly normal left ventricular cavity walls and septum (385, 386).

6.1.2 Signs and symptoms

The most common symptom of ACM is palpitation, which can occur at rest but is often triggered by physical activity (51). Other symptoms might include chest pain, light-headedness or blackouts and fainting caused by irregular heart rhythms. The clinical diagnosis has involved the use of i) signal-averaged ECG to detect late right ventricular (RV) potentials and epsilon waves, ii) echocardiography to reveal enlargement and hypokinetic of the RV, iii) cardiac MRI to detect fatty infiltration of the RV free wall, and iv) RV angiography considered a gold

standard is used for the definitive diagnosis of ACM and transvenous biopsy of the right ventricle can be highly specific for (ARVC (385, 387).

6.1.3 ACM disease burden

The definitive identification of the disease phenotype is usually challenging, and notably, ACM affects the young (<35 years) and physically active individuals, especially athletes. Most ACM patients have acquired the disease in an autosomal dominant (AD) inheritance pattern (388), and most studies have reported ACM at a prevalence rate of 1:2000-5000 (but with limited data from Africa) CMO patients, noting that about 40% of the affected individuals were asymptomatic. Also, about 4% to 16% of ACM patients have presented with HF, and SCD has occurred in 20% of patients exacerbated by physical exercise (389).

The incidence of ACM in the (SSA region has reached 1.7% and 5.6% for the paediatric and adult populations, respectively (14). The epidemiology of ACM in South Africa is largely unknown, and only a few published studies, especially from our research group in Cape Town, have linked HF to ACM in about 6% of the probands (155).

6.1.4 Aetiology of ACM

The most common cause of ACM is a genetic defect of desmosomal genes, although there are other genetic and non-genetic causes (390). The disease comprises a spectrum of conditions of different aetiologies involving the right ventricle, the left ventricle, or both, either genetic or non-genetic (391). The common denominator of all these conditions is the prominent presence of non-ischaemic ventricular myocardial scarring and scar-related ventricular arrhythmias (391). All ACM conditions are associated with a higher risk of (SCD because myocardial fibrosis acts as a substrate of malignant ventricular arrhythmias (387, 391).

6.1.5 Types of Arrhythmogenic Cardiomyopathy (ACM)

6.1.5.1 Arrhythmogenic right ventricular cardiomyopathy (ARVC)

In the right dominant ARVC, a dilated right ventricle with regional wall motion abnormalities with little left ventricular involvement is observed (392). As the disease progresses, myocardial

remodelling may move from the subepicardial layers to the transmural space, and myocardial wall thinning can be seen on macroscopic examination (393). The sub-tricuspid region and the RV outflow tract are especially vulnerable to this remodelling process, which can form aneurysms (29, 394). Patients with ARVC may show obvious signs of fibrosis, fatty infiltration, or inflammation in their cardiac histopathology (392).

ARVC is a rare (RV dysfunction characterised by RV fibrofatty infiltration. It is a leading cause of HF in young individuals, especially athletes, with a prevalence ranging from 1:5000 to 1:2000 and a slight male sex predominance (29, 395). ARVC patients have had variable expressivity, pathological pathways, and recovery rates, and its specific diagnostic criteria have remained a unique evolving clinical investigative approach (385). The current ACM working task force (387) has established risk stratification markers, including age, male sex, cardiac arrest, and VT (396). Currently, molecular genetic analyses have formed part of the diagnosis (85), and most individuals diagnosed using genetic techniques have presented with risk stratification markers related to penetrance. Most genetics findings have shown that the intercalated disc consisting of the adherens junction, desmosome, gap junction, and ion channels has harboured many variants known for causing ARVC (394).

6.1.5.2 Left-dominant arrhythmogenic cardiomyopathy (LDAC)

ACM is a heart muscle disease with left ventricular phenotypes in some patients (397). Left-dominant arrhythmogenic cardiomyopathy (LDAC) is a distinct form of ACM. It is characterised by early occurrence of LV involvement, where arrhythmias precede gross structural alterations. Global RV function is preserved initially. The structural and electrocardiographic (ECG) findings are left-sided analogues to those observed in ARVC/D (392).

It is usual to misdiagnose this under-recognised clinical entity (398, 399), especially as some heritable LDAC patients have presented with a phenotype that overlaps with other forms of CMOs (400). For example, LDAC is often misdiagnosed as DCM in some patients, and notably, a modified genes panel utilised for DCM clinical diagnosis would identify the disease-causing variants in most LDAC patients (401).

6.1.5.3 Biventricular arrhythmogenic cardiomyopathy (BiVAC)

Biventricular ACM is characterised by the early involvement of both ventricles (392). The disease progression is characterised by impaired systolic function and dilation of both

ventricles, leading to the clinical features of congestive HF. In comparison to dilated cardiomyopathy (DCM), where both ventricles are involved, ventricular arrhythmias of both (LBBB and right bundle branch block (RBBB) configurations are present at an early stage, originating in the left ventricle (392).

The current clinical task force criteria used in diagnosing ARVC might fail to detect BiVAC before death (393). Most BiVAC patients diagnosed with AD (402) or autosomal recessive (AR) mode of inheritance have received an incorrect diagnosis at some point (403). Misdiagnosing BiVAC as DCM, for example, has happened in about 8-12% of ACM patients carrying *DSP* classes 4-5 variants (404). Contrarily, classes 4-5 variants in *TMEM43 (*non-desmosomal gene) and, to a lesser extent, in the *DSG2 gene have been identified in BiVAC patients, especially those* taking competitive sports (405).

6.1.6 Genetics of ACM

Common ACM-causing genes are clustered across the cardiomyocyte cell membrane and are fundamental to intercardiomyocyte adhesion and signalling activities (394). The various cellular components implicated in ACM phenotypes are represented in Figure 6.2 and occur in constituents of the i) desmosome (desmocollin 2, desmoglein 2, junction plakoglobin, plakophilin 2 and desmoplakin) ii) adherens junction (cadherin 2 and catenin-α3) iii) calcium handling (phospholamban, ryanodine receptor 2, etc.) iv) Intracellular structural proteins (desmin, titin and filamin C) v) sodium channel and transforming growth factor-β3 and vi) nuclear envelope (transmembrane protein 43, lamin A) (394).

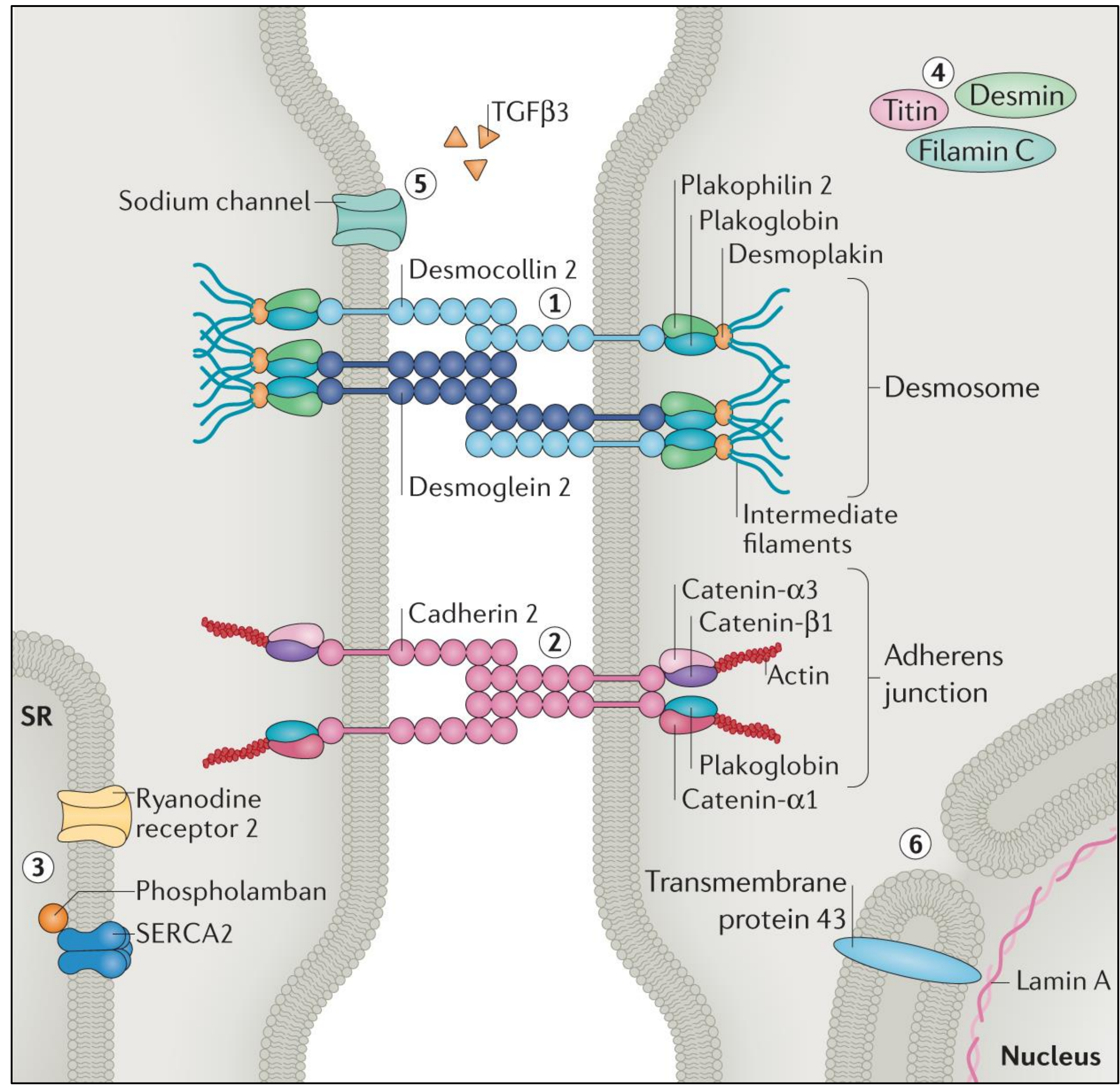


Figure 6.2: Protein components with reported ACM-causing variants. *Categories of protein in which ACM-causing variants occur are labelled accordingly. (1) desmosomal proteins, including desmocollin 2, desmoglein 2, junction plakoglobin, plakophilin 2 and desmoplakin. (2) Components of the adherens junction, including cadherin 2 and catenin-α3. (3) Contributors to calcium handling, including phospholamban and ryanodine receptor 2 located in the membrane of the sarcoplasmic reticulum (SR). (4) Intracellular structural proteins, including desmin, titin and filamin C. (5) The sodium channel and transforming growth factor-β3 (TGFβ3). (6) Nuclear envelope proteins transmembrane protein 43 and lamin A. SERCA2, sarcoplasmic/endoplasmic reticulum calcium ATPase 2 (394).*

6.1.7 ACM genotype-phenotype correlation

The majority of the ACM disease-causing genes encode for the desmosomal plaque proteins and variants in the genes (consisting of plakophilin 2 (*PKP2*) (in 10%–45% of patients), desmoplakin (*DSP*) (10%–15%), desmoglein 2 (*DSG2*) (7%–10%), and desmocollin-2 (*DSC2*) (2%)) have been reported that cause heritable disease predominantly in the European's population (406). Amongst the ACM disease-causing genes is a plausible intercalated disc

candidate gene, *CDH2* (407), which was initially reported in a South African family with ARVC/D surfaced in our study population. Alterations in *CDH2* and other genes' protein products expressed in the ventricles would yield different clinical features that have caused a variety of ACM phenotypes ranging from mild to severe conditions (408). Figure 6.3 overviews ACM's three main classified heritable subtypes (392, 402).

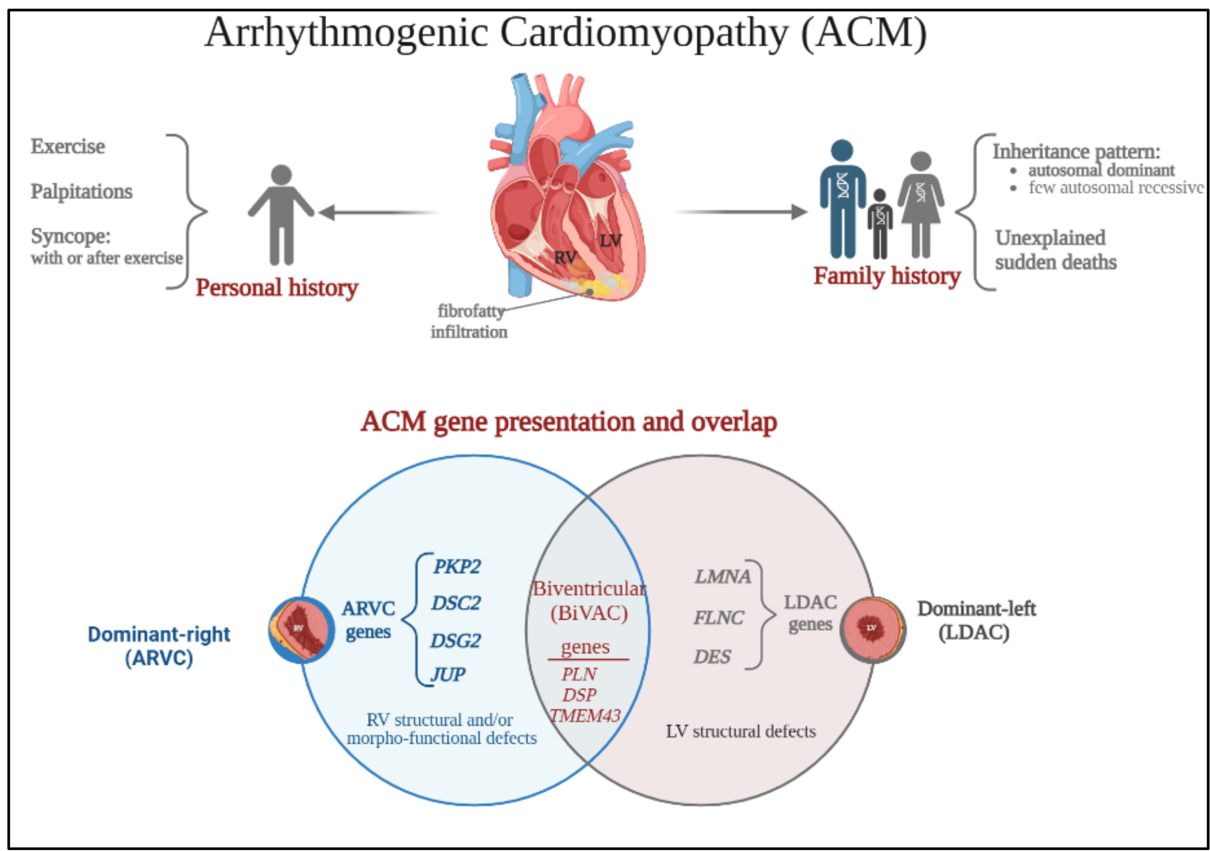


Figure 6.3: ACM genes presentation and overlap. *The classified subtypes of heritable ACM were drawn by P. Ndibangwi using BioRender.*

Notably, variants in the most common genes, for example, *PKP2* (31%) and *DSG2* (29%), have accounted for most of the age-related ARVC and in some patients, *DSP* variants (12%) have caused ARVC triggered by physical exercise (95, 396, 404). Heritable ARVC is the most common subtype of ACM; however, the genetic determinants are less established in South Africa compared to European countries such as the Netherlands (250). Recent LDAC genetic testing has found that most patients carry classes 4 or 5 variants in the *DSP* (18), *MYBPC3* and *PKP2* (5, 21) genes; however, much has remained unknown about these phenotype genetic determinants. Most genotype-positive ACM patients have had only one class 4 or 5 variant; however, about 2% to 4% of ACM *patients* carrying multiple variants (with at least one

desmosomal gene variant) have yielded a positive genetic diagnosis of *BiVAC* (250). Additionally, multiple disease-causing variants or those of unknown pathogenicity can add complexity and contribute to disease development and/or have modifier effects (85).

6.1.8 Therapeutics

The current management approach for patients with ACM includes exercise restriction, β-blocker therapy, implantable cardioverter-defibrillator insertion, catheter ablation, and, in some cases, surgical interventions (396). The treatments available for ACM mainly aim to manage the symptoms, enhance cardiac function, and address the condition's underlying causes (409), as shown in Figure 6.4.

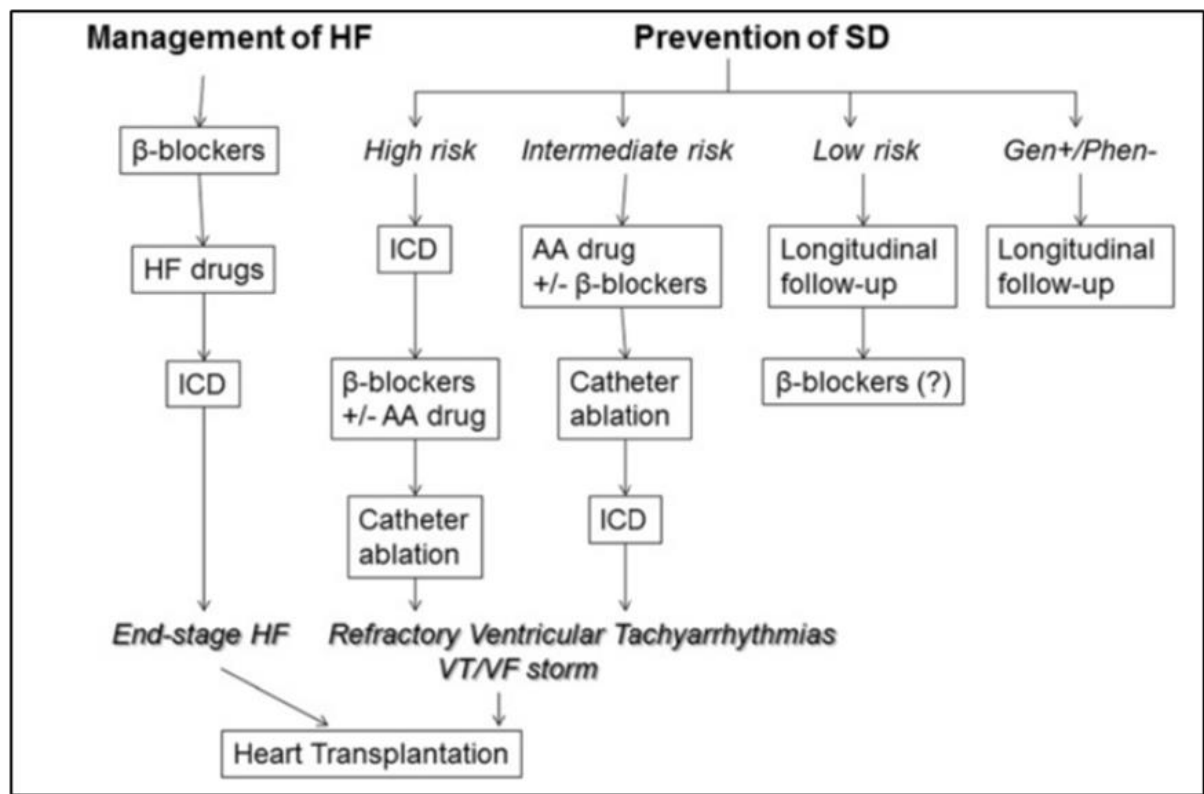


Figure 6.4: A flow chart showing the clinical treatment for patients affected by ACM. *On the left is the management of heart failure (HF); on the right is the management of ventricular arrhythmias and prevention of sudden death (SD) according to the SD risk category. Abbreviations. AA: antiarrhythmic ICD: Implantable cardioverter defibrillator; VT: ventricular tachycardia; VF: ventricular fibrillation; Gen+/Phen-: Genotype positive/phenotype negative (409).*

Below is a summary of the essential therapeutic strategies for ACM:

6.1.8.1 Medications

Various antiarrhythmic drugs are used, such as sodium blockers, β-blockers, sotalol, amiodarone, and verapamil, alone or in combination (409, 410). Various authors have reported varying efficacy rates with antiarrhythmic drugs for treating both inducible and non-inducible VT in ACM patients. Among the available drugs, sotalol is considered the most effective, with a recommended dosage of 320-480 mg/day. However, amiodarone, while less effective, is also an option, although it has a higher risk of extracardiac side effects during long-term follow-up. β-blockers, on the other hand, are only effective in non-inducible patients. It is recommended to avoid co-administration of more than one drug. However, there is conflicting data regarding the effectiveness of empiric antiarrhythmic drugs. Some studies have shown a higher efficacy of amiodarone, while others have demonstrated the inefficacy of antiarrhythmic drugs against ICD intervention and SD. There may be differences in study populations, therapeutic approach, drug doses, and follow-up duration that explain the divergence between treatment options (409).

Medications play a vital role in managing heart-related health conditions and symptoms management. For example, ACE inhibitors and ARBs help dilate blood vessels, decrease afterload, and improve the heart's pumping ability, while diuretics can help alleviate fluid retention and reduce oedema by increasing urine output (174, 175).

6.1.8.2 Device Therapy

Implantable Cardioverter-Defibrillator (ICD) is a medical device that can be implanted in individuals at risk of life-threatening arrhythmias. It is designed to monitor and regulate heart rhythms, thus preventing sudden cardiac arrest (175).

The 2015 International Task Force (ITF) Consensus Statement on the treatment of ACM has categorised risk into three levels based on estimated annual rates of malignant arrhythmic events (Figure 6.5). The high-risk category's estimated risk of major arrhythmic events has been determined using data on mortality rates associated with previous events, and specific risk factors (apex of the pyramid) is >10% per year, in the intermediate-risk category (mid-of pyramid) ranges from 1% to 10% per year, and in the low-risk category (base of the pyramid) is <1% per year (391).

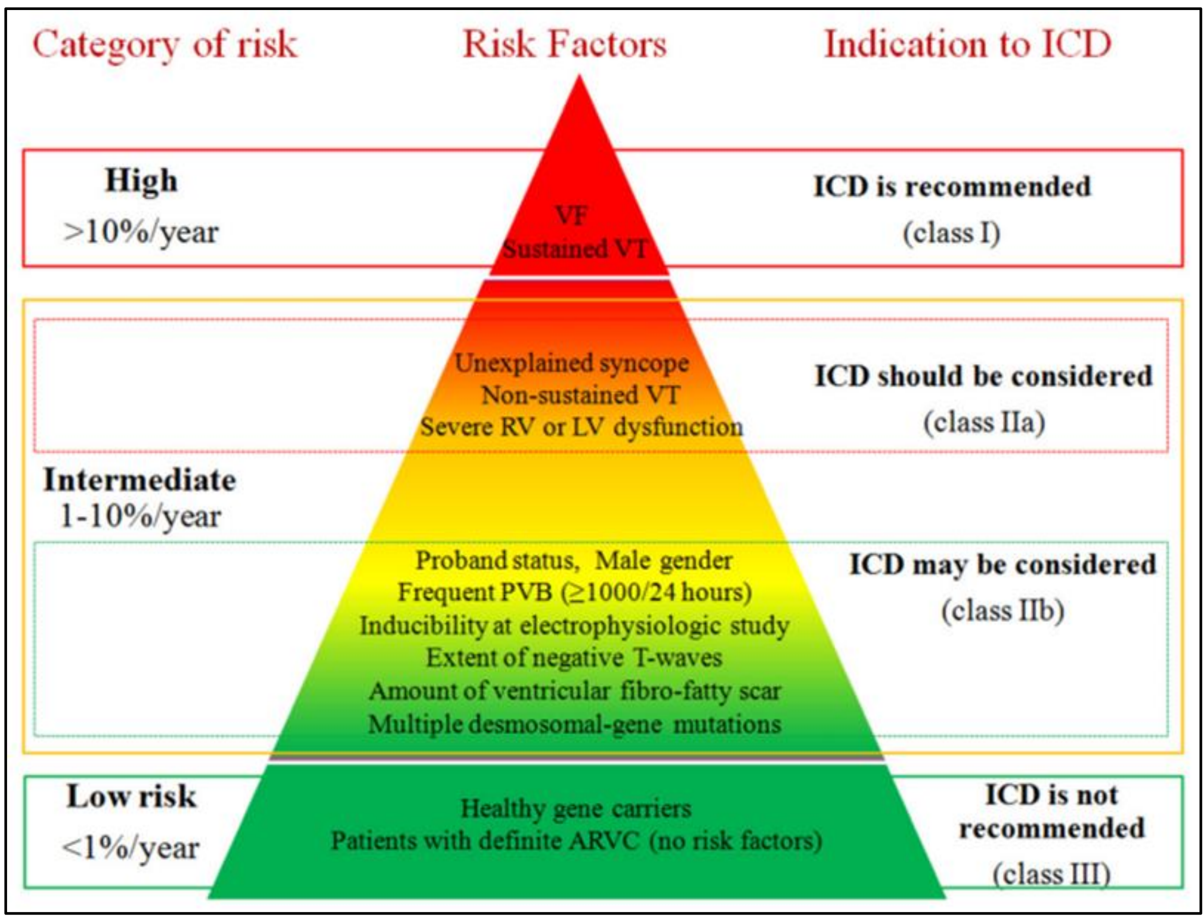


Figure 6.5: Implantable cardioverter-defibrillator therapy risk pyramid for ACM. *LV indicates left ventricle; PVB, premature ventricular beats; RV, right ventricle; VF, ventricular fibrillation; and VT, ventricular tachycardia (391).*

Cardiac Resynchronization Therapy (CRT) is a specialised treatment that involves using a special pacemaker to coordinate the contractions of the heart's ventricles. CRT helps improve overall cardiac function, benefiting individuals suffering from HF or other cardiac conditions (176).

6.1.8.3 Lifestyle modifications

Lifestyle modifications to improve cardiovascular health include regular exercise, dietary changes, and quitting smoking. A heart-healthy diet, low in salt and saturated fats, and personalised exercise recommendations are often prescribed (51). Quitting smoking is essential for overall heart health.

6.1.8.4 Management of underlying causes

Identifying and addressing the underlying causes of ACM, including inherited mutations, viral infections, autoimmune disorders, or metabolic conditions, is crucial to managing the condition effectively (51, 174).

6.1.8.5 Heart transplantation

In cases of severe ACM that are unresponsive to other treatments, heart transplantation may be considered as a final option (51).

6.1.8.6 Genetic counselling

Genetic counselling could be recommended for cases with a genetic component to assess risk to family members and provide guidance on preventive measures (51).

6.1.8.7 Gene therapy

Gene therapy is a form of treatment for ACM that aims to address the underlying genetic factors contributing to the disease. Gene therapy can be done by introducing, modifying, or replacing genes to correct or compensate for genetic defects (141, 178). Gene therapy is being explored as a safe and feasible option for treating ACM. Researchers are investigating different strategies, including gene delivery to improve heart function and RNA-based approaches to regulate gene expression. Adeno-associated viruses are the most commonly used viral vectors for gene delivery (411). However, gene therapy for ACM faces challenges such as immune responses to viral vectors and off-target effects. Clinical trials are evaluating the safety and effectiveness of gene therapies in humans.

6.1.9 ACM genetic modifiers

ACM is a genetic condition that can present with a range of symptoms, from mild to severe, including SCD (412). It is common for cases to be isolated to a single patient without a FH, and about 35% to 50% have no pathogenic variant (368, 413). Those carrying ACM-associated variants might not show any disease expression, making identifying the genetic cause challenging, indicating that other factors and modifiers could be involved (368).

The majority of ACM-related variants are found in genes that encode proteins in the cardiac desmosomes located within the intercalated disc. However, other genetic factors also contribute to the disease process, although their precise role is not yet fully understood (352, 414). For example, a combination of a stop_gained variant and a large deletion in the *DSG2* gene has

been reported in a patient with a non-familial phenotype suggestive of an AR inheritance pattern (414). Variants in *AKAP9* and *DLG1* genes have also been identified as possible genetic modifiers of arrhythmic risk and phenotype severity in cardiomyopathies (85).

6.2 **METHODS**

The methods discussed in Chapter 2 are briefly described in Section 5.2 for the ACM cohort.

6.2.1 Ethical Compliance:

Ethical clearance with reference number **HREC 009/2020** was obtained for this study from the Human Research Ethics Committees (HREC) at the (FHS (UCT. All participants were informed before they provided a signed consent form in the main study IMHOTEP registry (HREC 766/2014).

6.2.2 Study Population:

Forty-one unrelated probands clinically diagnosed with ACM were recruited from February 2018 to October 2021. The probands were examined by cardiologists at various participating clinics across South Africa and reviewed at the leading recruitment site at Groote Schuur Hospital in Cape Town. The clinical examination included baseline assessment, medical history, pedigree, ECG, and transthoracic echocardiography. Our clinical team used the European Society of Cardiology guidelines for the clinical diagnostic criteria, and the primary IMHOTEP study protocol has detailed inclusion and exclusion criteria (Appendix A).

6.2.3 DNA Extraction and Quality Control

We extracted genomic DNAs from the probands' buffy or peripheral whole-blood leukocytes following the published manufacturer's methods for Puregene Blood Core Kit C (Appendix C) or PAXgene Blood DNA kit (Appendix D), respectively. The purity and concentrations of the gDNA samples were computed with a calibrated NanoDropTM 2000 spectrophotometer (Thermo SCIENTIFIC) with software v2.6 on a Windows 10 computer. We checked the samples' intactness using a 1% agarose gel. The gel electrophoresis was programmed to run at 110V for 75 minutes, and we used a Uvitec Xplorer D55 gel doc running on an Xplorer 1D software version 15.08 (Uvitec Cambridge, Cambridge, UK) to visualise the gels.

6.2.4 Targeted Sequencing and Variant Calling

A custom-designed list of 38 CMO disease-causing genes was created by the (OMGL in the UK. The custom designed chip contained 38 specific genes including: *CDH2*, *DSP*, *PKP2*, *LMNA*, *PLN*, *DES*, *PRKAG2*, *DSC2*, *SCN5A*, *DSG2*, *JUP*, *TMEM43* and other ACM-related genes. The complete genes list is attached as Appendix F. All gDNA samples underwent targeted resequencing, initially performed using the Haloplex kit but subsequently replaced with the Twist NGS kit in September 2021. The data generated were mapped to the human reference genome (GRCh37/hg19) and transformed into variant call format (VCF) using the Genome Analysis Toolkit (GATK) v3.8-1-0 and Platypus v0.8.1. The VCF files were annotated with Ensembl Variants Effect Predictor (VEP) and analysed.

Through a specific set of filtering criteria, we selected variants for further analysis. The variants were classified according to the refined ACMG guidelines criteria as class 1 (B), class 2 (LB), class 3 (VUS), class 4 (LP) and class 5 (P). Variants were selected (1) with a gnomAD minor allele frequency (MAF) of <1%, (2) variants that were predicted by bioinformatics tools, SIFT and PolyPhen2 to have damaging deleterious or unknown protein function probably. Alamut™ Visual Plus version 1.5.1 and ClinVar (2022-12-17) were used to analyse the variants further. We also looked at the variant caller scores for Haplocaller and Platypus for threshold values above 400 and 700, respectively, as variants below these thresholds are artefacts or polymorphisms. The BAM files for probands carrying class 4 to 5 variants were also visualised in the Alamut as this could identify possible false positive variants. All identified variants of interest were validated using Sanger sequencing.

6.2.5 Biostatistical Analysis:

The data was analysed in Microsoft Excel 365 or R and RStudio (version 4.1.2); results were presented as $\bar{x} = \pm 1\delta$ for numerical variables and total counts with percentages for categorical variables.

6.3 **RESULTS**

6.3.1 Baseline characteristics for the adult-onset ACM cohort

We recruited and enrolled 41 consecutive unrelated ACM probands (≥ 13 years) through IMHOTEP, Table 6.1.

Table 6.1: Baseline results for the ACM cohort

LABEL	LEVELS	TOTAL IMHOTEP COHORT	ACM ADULT-ONSET PROBANDS
NUMBER OF PROBANDS WITH GENETICS DATA ANALYSED, N (%)		690	41 (5.9)
AGE AT DIAGNOSIS (YEARS)	Mean (SD)	31.7 (16.4)	40.6 (13.3)
LEFT VENTRICULAR EJECTION FRACTION (%)	Mean (SD)	36.1 (19.6)	60.2 (11.0)
NEW YORK HEART ASSOCIATION CLASSIFICATION	Class I	83 (15.7)	28 (77.8)
	Class II	162 (30.6)	7 (19.4)
	ND	10 (1.9)	1 (2.8)
	Class III	223 (42.1)	-
	Class IV	52 (9.8)	-
FAMILIAL	No	405 (58.7)	27 (65.9)
	Yes	125 (18.1)	14 (34.1)
	No data	160 (23.2)	5 (12.2)
SEX ASSIGNED AT BIRTH	Female	364 (52.8)	14 (34.1)
	Male	326 (47.2)	27 (65.9)
ANCESTRY	Black-African	381 (55.2)	3 (7.3)
	Mixed	221 (32.0)	8 (19.5)
	White	83 (12.0)	27 (65.9)
	Indian	5 (0.7)	3 (7.3)

VALUES ARE MEAN ± 1SD OR N (%) BASED ON NON-MISSING VALUES OF THE TOTAL ANALYSED.

They presented at an average age of 40.6 years (SD 13.3, range 15-75), while 65.9% (27/41) of the probands were assigned male sex at birth. Figure 6.6 indicates the probands' self-reported ancestry, showing that 65.9% were WA, 19.5% MA, 7.3% BA, and 7.3% IA. Regarding aetiology classification based on clinical definitions, 65.9% (27/41) of this cohort showed non-familial or idiopathic diagnosis without genetic testing and FH information. Also, 34.1% (14/41) of the probands had familial, and 12.2% (5/41) had no data for FH.

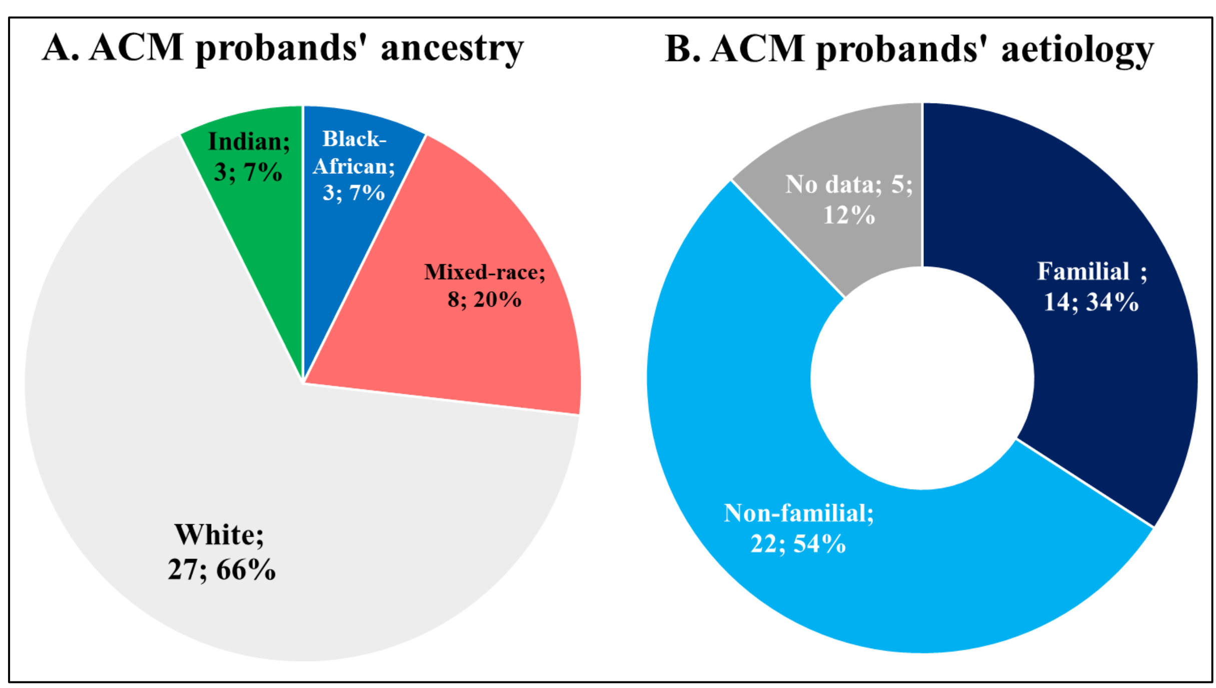


Figure 6.6: ACM probands family history.

6.3.2 Baseline characteristics for the adult-onset ACM population groups

Due to the prevalence of the (WA probands in this cohort, along with our understanding of South Africa's diverse population, we conducted an in-depth investigation to explore potential variances in age, sex assigned at birth, and diagnostic classification within the four major demographic groups: Black-African, Mixed, Indian, and White. We observed notable sex differences and diagnostic classification across the various population groups, as illustrated in Figure 7.3. Nevertheless, we approached our analysis with prudence, considering our sample's limited representation of specific population groups.

The 27 ACM probands of WA enrolled had a mean age of 41 (SD 14.7; range 15-75 years). Of these probands, 70.4% (19/27) were assigned male sex at birth. Diagnostic classification showed that 51.9% (12/27) had a non-familial or idiopathic diagnosis without genetic testing or FH information, while 33.3% (9/27) were classified as familial, as illustrated in Figure 7.3 D.

Within the ACM cohort, we identified eight participants of MA ancestry, with an average age of 38 years (SD 7.8; range 34-46). Their mean age was younger than the average age of the ACM cohort, which was 40.6 years. Of the probands, 62.5% (5/8) were male, another 62.5%

(5/8) had familial causes, and 37.5% (3/8) had non-familial or idiopathic aetiology. It is worth noting that clinical classification was made before genetic testing and FH information. For further details, please refer to Figure 6.7.

For the three Black-African ACM probands enrolled, they had a mean age of 33 years (SD 17.9), which was younger than the ACM cohort average of 40.6 years; 100% (3/3) of the probands were assigned female sex at birth, and the diagnostic classification showed that all were non-familial or idiopathic in the absence of genetic testing and FH information, Figure 6.7.

For the three Indian ACM probands enrolled, they had a mean age of 39 years (SD 12.6); 100% (3/3) of the probands were assigned male sex at birth, and the diagnostic classification showed that 66% (2/3) had a non-familial or idiopathic diagnosis in the absence of genetic testing and FH information. One proband had an unknown classification, Figure 6.7.

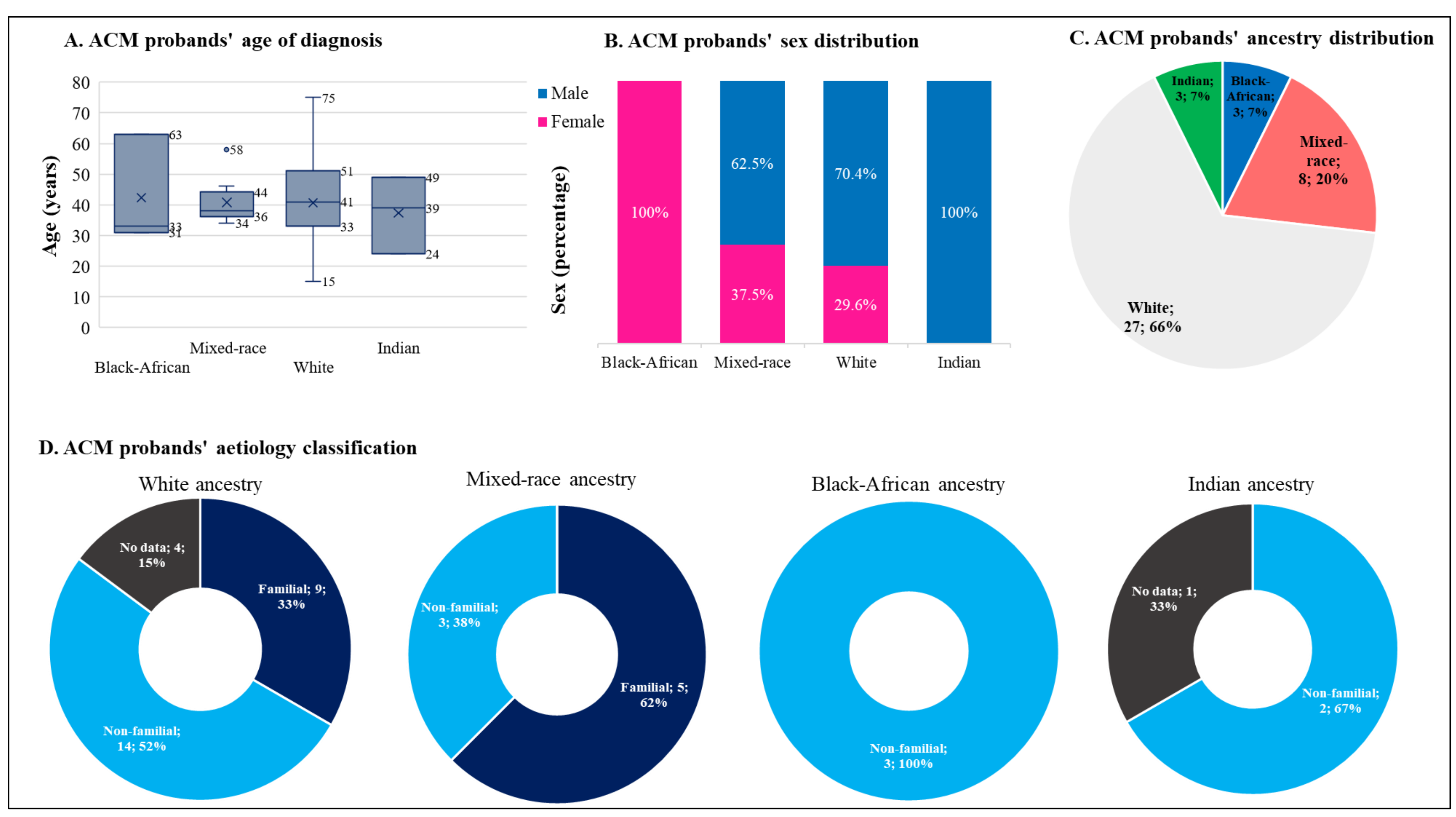


Figure 6.7: ACM probands' age, sex assigned at birth and aetiology distribution. *A. Population groups vs age of diagnosis in ACM; B. Population groups vs sex distribution; C. Population groups vs ancestry classification; D. Population groups vs aetiology/diagnostic classification.*

6.3.3 Targeted sequencing for adult-onset ACM probands

Table 6.2 presents targeted sequencing results and baseline characteristics for the 41 ACM probands in this cohort. Whereas SS Table 6 shows the fitted logistic regression model for adult ACM probands. In this context, the probands with class 4 and 5 variants have positive genetic results, those with class 3 have uncertain genetic results, and those with class 1 and 2 have negative genetic results.

Table 6.2: Baseline results for the ACM cohort according to ACMG classification

LABEL	LEVELS	TOTAL	POSITIVE	UNCERTAIN	NEGATIVE	P-VALUE
ACM diagnostic yield, n (%)		41	12 (29.3)	16 (39.0)	13 (31.7)	
AGE	Mean (SD)	40.6 (13.3)	35.9 (15.2)	43.8 (15.7)	40.9 (5.9)	0.303
LVEF	Mean (SD)	60.2 (11.0)	60.1 (16.9)	58.7 (8.3)	62.6 (7.1)	0.744
NYHA						0.507
	Class I	28 (77.8)	8 (72.7)	11 (78.6)	9 (81.8)	0.065
	Class II	7 (19.4)	3 (27.3)	3 (21.4)	1 (9.1)	0.365
	ND	1 (2.8)	0 (0.0)	0 (0.0)	1 (9.1)	0.994
FAMILIAL*						0.858
	Yes	14 (34.1)	5 (41.7)	4 (25.0)	5 (38.5)	0.292
	No	22 (53.7)	6 (50.0)	10 (62.5)	6 (46.2)	0.593
	No data	5 (12.2)	1 (8.3)	2 (12.5)	2 (15.4)	0.865
SEX*						0.302
	Female	14 (34.1)	2 (16.7)	7 (43.8)	5 (38.5)	0.292
	Male	27 (65.9)	10 (83.3)	9 (56.2)	8 (61.5)	0.692
ANCESTRY*						0.510
	White	27 (65.9)	9 (75.0)	9 (56.2)	9 (69.2)	0.994
	Mixed	8 (19.5)	2 (16.7)	3 (18.8)	3 (23.1)	0.994
	Indian	3 (7.3)	1 (8.3)	1 (6.2)	1 (7.7)	0.994
	Black-African	3 (7.3)	0 (0.0)	3 (18.8)	0 (0.0)	0.994

VALUES ARE MEAN ± 1SD OR N (%) BASED ON NON-MISSING VALUES OF THE TOTAL ANALYSED:41. POSITIVE-CLASS 4 & 5, UNCERTAIN-CLASS 3, AND NEGATIVE-CLASS 1&2: LVEF-LEFT VENTRICULAR EJECTION FRACTION, NYHA-NEW YORK HEART ASSOCIATION. ASTERISKS (*)-SELF-REPORTED.

Overall, targeted sequencing found 29.3% (12/41) ACM probands with ACM-causing variants: nine probands were positive for class 5 and three probands were positive for class 4 variants (Table 6.2). Meanwhile, 39% (16/41) probands carried class 3 variants, and 31.7% (13/41) probands carried only class 1 and 2 variants. This ACM diagnostic yield highlights that in South Africa, a significant portion of probands (39%) with variants assigned to class 3 may require further evaluation to determine if they are P or B.

We also note that 39 variants were found across ACM classes 3-5: 14 classes 4 and 5, and 25 class 3. Figure 6.8 presents an overview of the distribution of novel and reported variants across different population groups. We also noted that 28.2% (11/39) of the variants were found in classes 4 and 5, with the larger percentage being reported, and 64.1% (25/39) of the variants

were clustering in class 3, which are variants of unknown significance, with 40% being recorded as novel.

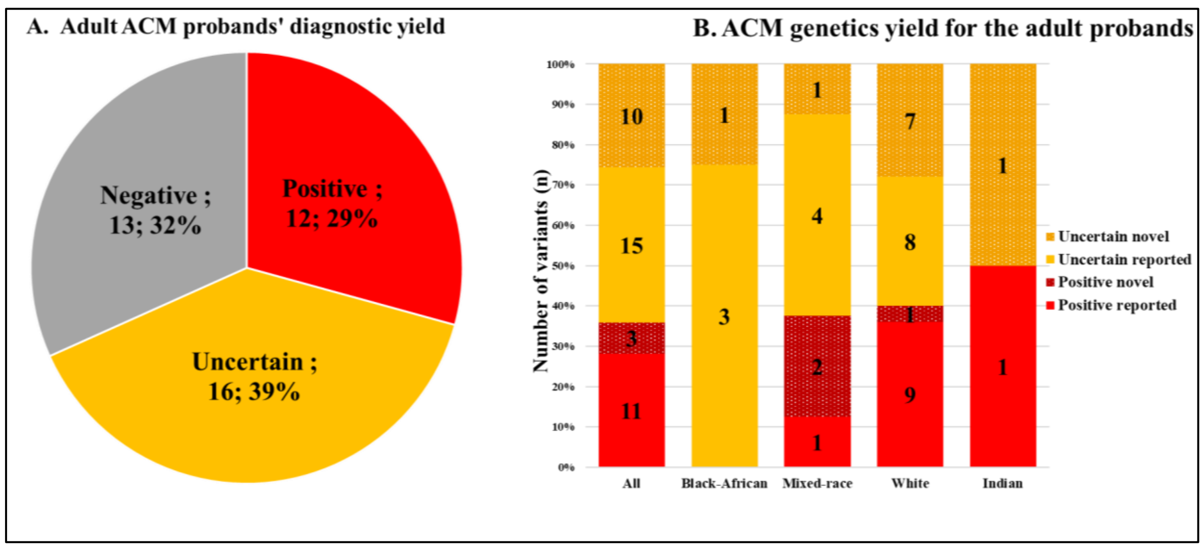


Figure 6.8: ACM probands' diagnostic yield. *A. Pie chart indicating the diagnostic yield for the adult ACM cohort, B. Bar graph showing the count of probands with variants grouped as novel vs reported within the ACM cohort.*

This thesis comprehensively examines the impacts of class 4 and 5 variants on ACM. We also touch upon the modifying effects of class 3 variants on ACM through a succinct report. It is important to note that our research excludes classes 1 and 2 variants as they have a lower likelihood of causing ACM.

6.3.3.1 Probands with pathogenic (class 4 and 5) variants

6.3.3.1.1 Baseline characteristics

Through targeted sequencing, we identified 12 ACM probands with 11 pathogenic variants: 75% (9/12) of the probands were White, 17% (2/12) of the probands were Mixed, and 8% (1/12) probands were Indian. We found no class 4 and 5 variants in the BA probands (Table 6.3). The 12 probands had a mean (SD) age of 35.9 (14.5), about five years younger than the combined mean age of 40.6 (13.3) years at diagnosis. We also found that only two probands were females, and for the rest of the class 4 and 5 variant carriers, 83.3% (10/12) were males.

Table 6.3: Class 4 and 5 variants (n=11) in the adult-onset ACM probands (n=12) spanning three population groups

PROBAND_ID	SEX	ANCESTRY	AGE	AETIOLOGY	GENE	PROTEIN CHANGE	CDNA_CHANGE	CONSEQUENCES	EVIDENCE	dbSNP_ID	ZYGOSITY	ACMG
10100091*	Male	White	19	Familial	*PKP2*	p.His733ProfsTer8	c.2198_2202del	frameshift	Reported	rs759179184	Het	Class 5
10100091*	Male	White	19	Familial	*ALPK3*	p.Gln901Ter	c.2701C>T	stop_gained	Novel	None	Het	Class 4
10100105	Male	Mixed	46	Familial	*VCL*	Splice Site	c.2949+2T>C	splice_donor	Novel	None	Het	Class 4
10100130	Male	White	62	Non-familial	*PKP2*	p.Ala570ValfsTer7	c.1709del	frameshift	Reported	rs397517005	Het	Class 4
10100131	Female	White	26	No data	*PKP2*	p.Arg388Trp	c.1162C>T	missense	Reported	rs766209297	Hom	Class 5
10100137	Male	White	15	Non-familial	*CDH2*	p.Asp407Asn	c.1219G>A	missense	Reported	rs568089577	Het	Class 4
10100142	Male	Indian	49	Non-familial	*PKP2*	p.Ala326ArgfsTer29	c.968_975dup	frameshift	Reported	rs1555148032	Het	Class 5
10100155	Male	White	33	Non-familial	*PKP2*	Splice Site	c.2489+1G>A	splice_donor	Reported	rs111517471	Het	Class 5
10100317#	Male	Mixed	36	Familial	*TTN*	p.Gly27533ValfsTer47	c.82598del	frameshift	Reported	None	Het	Class 5
10100317#	Male	Mixed	36	Familial	*RBM20*	p.Glu1199Ter	c.3595G>T	stop_gained	Novel	rs777768807	Het	Class 4
10100390	Male	White	40	Non-familial	*FLNC*	p.Gln357Ter	c.1069C>T	stop_gained	Reported	None	Het	Class 5
101301408	Male	White	37	Familial	*PKP2*	p.Arg388Trp	c.1162C>T	missense	Reported	rs766209297	Het	Class 5
101301411	Male	White	15	Non-familial	*PKP2*	p.Arg388Trp	c.1162C>T	missense	Reported	rs766209297	Hom	Class 5
101301539	Female	White	53	Familial	*PKP2*	p.Arg388Trp	c.1162C>T	missense	Reported	rs766209297	Het	Class 5

*One proband (10100091) has a *PKP2* (class 5) and an additional novel class 4, *ALPK3* p.Glu1199Ter; c.3595G>T stop_gained variant while the other one proband# (10100317) has a *TTN* (class 5) and an additional novel class 4, *RBM20* p.Gln901Ter; c.2701C>T variant.

6.3.3.1.2 Diagnostic findings

6.3.3.1.2.1 Class 4 and 5 variants carriers' genetic characteristics

The 12 probands (nine White, two Mixed and one Indian) carried a total of 11 disease-causing variants, with 45.5% (5/11) occurring in the desmosomal gene (*PKP2*), 27.3% (3/11) in the cytoskeleton genes (*ALPK3*, *FLNC*, and *VCL*), 9.1% (1/11) in the adherens junction gene (*CDH2*), 9.1% (1/11) in the sarcomere gene (*TTN*), and 9.1% (1/11) in the nuclear gene (*RBM20*).

Of the seven genes with the 11 variants identified, we found *PKP2* to be the dominant gene, occurring in 67% (8/12) of the probands, followed by 9% (1/12) in *FLNC*, 8% (1/12) *TTN*, 8% (1/12) *VCL* and 8% (1/12) *CDH2*, Figure 6.9.

To gain a more accurate understanding of the genetic underpinnings within the South African populations, we examined the class 4 and 5 (section 6.3.3.1.2) and class 3 (section 6.3.3.2.2) variants within the context of the dominant population groups in our ACM cohort: White, Mixed, Indian and Black-African.

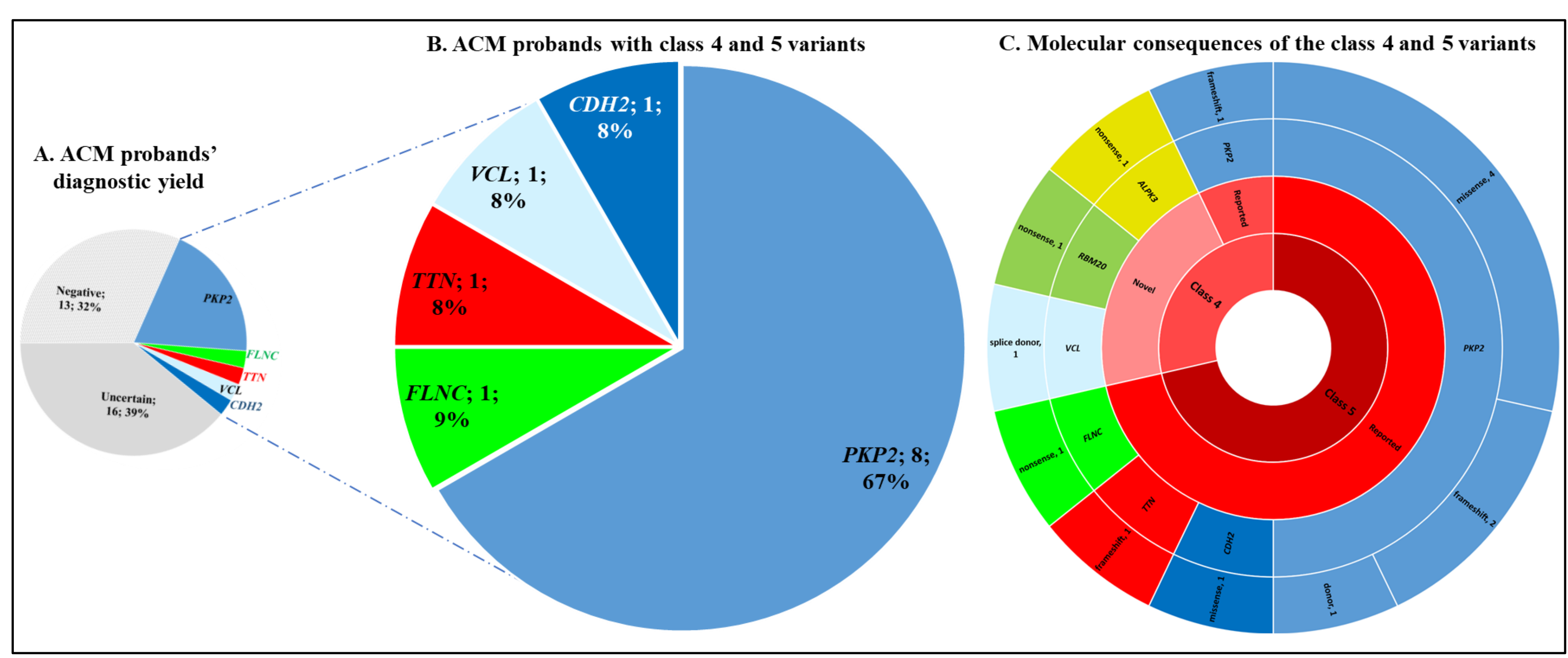


Figure 6.9: Genes found in ACM cohort with class 4 and/or class 5 variants.

6.3.3.1.2.1.1 *Class 4 and 5 variant carriers of White ancestry (WA)*

Out of the 27 probands with WA probands we recruited, 33% (9/27) tested positive for class 4 or 5 variants. The remaining participants were either negative or had variants of uncertain significance. Notably, the nine WA probands constituted 75% (9/12) of the positive probands in our ACM cohort, as shown in Figure 6.10.

The WA probands had a much younger age of onset [mean=31.9 (SD 12.1, range 15-62)] (Figure 6.10 A) compared to the ACM cohort age of onset [mean of 40.6 years (median=36.5 years)]; 78% (7/9) probands were males at birth (Figure 6.10 B).

The nine WA probands carried ten variants harboured by four genes, 70% (7/10) of which occurred in the *PKP2* gene.

6.3.3.1.2.1.1.1 PKP2 c.1162C>T

The *PKP2* c.1162C>T (p.Arg388Trp) missense variant recurred in four WA probands. We report the variant in mildly to severely affected ACM patients expressed in either homozygous or heterozygous alleles. The c.1162C>T point mutation disrupts plakophilin-2 at codon 388. The disruption is predicted to impair cell adhesion, increase electrical impulse imbalances, and induce immune cell infiltration, resulting in abnormal structural remodelling. In the homozygous state, the *PKP2* c.1162C>T variant is expected to be more severe.

6.3.3.1.2.1.1.2 PKP2 c.2198_2202del and ALPK3 c.2701C>T

We identified one proband (ID 10100091) with two variants: a reported class 5 *PKP2* c.2198_2202del (p.His733ProfsTer8) frameshift and novel *ALPK3* c.2701C>T class 4 stop_gained variant. The 19-year-old WA male proband was diagnosed with familial disease classified as NYHA class I, with an ejection fraction of 55%.

6.3.3.1.2.1.1.3 PKP2 c.2489+1G>A

Proband (ID 10100155) is a 33-year-old BA male classified as NYHA class I with no FH of heart disease. Targeted sequencing identified the *PKP2* c.2489+1G>A splice_donor class 5 variant as the disease cause. The splice_donor variant is expected to disrupt plakophilin-2 exon 12 by altering the 3' splicing site. The resultant protein is expected to cause the myocardium's

structural instability, predisposing patients to cardiac arrhythmias leading to early ventricular tachyarrhythmias.

6.3.3.1.2.1.1.4 PKP2 c.1709del

One male proband (ID 10100130) of WA diagnosed at the age of 62 and classified as NYHA class II, with an ejection fraction of 82%, had the *PKP2* c.1709del (p.Ala570ValfsTer7) frameshift variant. The *PKP2* c.1709del variant is on a highly conserved region of the armadillo domain of the plakophilin-2 protein, and the alteration can create a shortened or no protein product.

6.3.3.1.2.1.1.5 CDH2 c.1219G>A

Proband ID 10100137 is a 15-year-old male of WA classified as NYHA class I, with an ejection fraction of 58% and no history of heart disease. Targeted sequencing identified the proband's missense *CDH2* c.1219G>A (p.Asp407Asn) variant. The variant is located in the cadherin exon 9 domain and is predicted to interfere with normal cell adhesion processes. The alteration results in the replacement of aspartic acid, which is negatively charged with neutral asparagine, leading to a loss of charge at the mutated position. Additionally, the larger side chain of Asparagine can impact the size and cadherin-like domain structure of the protein, which may impact the structure and function of the N-cadherin protein. Any modifications to the adhesive properties of N-cadherin can cause faulty cell-to-cell interactions in the heart muscle.

6.3.3.1.2.1.1.6 FLNC c.1069C>T

Proband (ID 10100390) is a 40-year-old WA male who had been classified as NYHA class II, with an ejection fraction of 65%; there was no FH of heart disease. Targeted sequencing identified the *FLNC* c.1069C>T (p.Gln357Ter) stop_gained class 5 variant as the likely cause of disease. The *FLNC* c.1069C>T disease-causing variant is located in exon 7 on the ROD1 Ig-like domain of the filamin-C protein. Class 4 or 5 variants in the ROD1 region are capable of causing structural instability of the cardiomyocytes, predisposing the myocardium to stress.

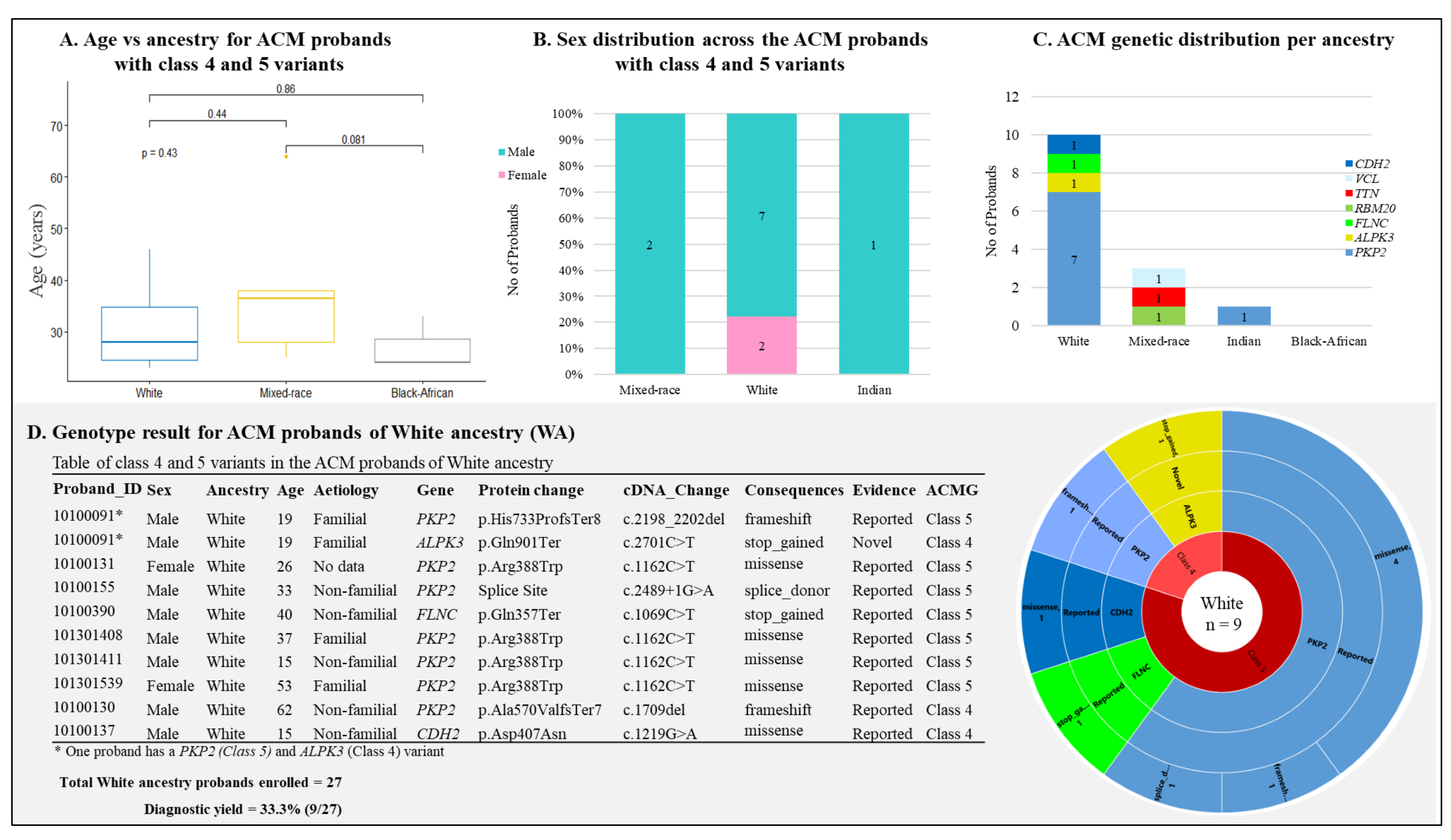


Table of class 4 and 5 variants in the ACM probands of White ancestry

Proband_ID	Sex	Ancestry	Age	Aetiology	Gene	Protein change	cDNA_Change	Consequences	Evidence	ACMG
10100091*	Male	White	19	Familial	*PKP2*	p.His733ProfsTer8	c.2198_2202del	frameshift	Reported	Class 5
10100091*	Male	White	19	Familial	*ALPK3*	p.Gln901Ter	c.2701C>T	stop_gained	Novel	Class 4
10100131	Female	White	26	No data	*PKP2*	p.Arg388Trp	c.1162C>T	missense	Reported	Class 5
10100155	Male	White	33	Non-familial	*PKP2*	Splice Site	c.2489+1G>A	splice_donor	Reported	Class 5
10100390	Male	White	40	Non-familial	*FLNC*	p.Gln357Ter	c.1069C>T	stop_gained	Reported	Class 5
101301408	Male	White	37	Familial	*PKP2*	p.Arg388Trp	c.1162C>T	missense	Reported	Class 5
101301411	Male	White	15	Non-familial	*PKP2*	p.Arg388Trp	c.1162C>T	missense	Reported	Class 5
101301539	Female	White	53	Familial	*PKP2*	p.Arg388Trp	c.1162C>T	missense	Reported	Class 5
10100130	Male	White	62	Non-familial	*PKP2*	p.Ala570ValfsTer7	c.1709del	frameshift	Reported	Class 4
10100137	Male	White	15	Non-familial	*CDH2*	p.Asp407Asn	c.1219G>A	missense	Reported	Class 4

* One proband has a *PKP2 (Class 5)* and *ALPK3* (Class 4) variant

Figure 6.10: ACM probands of White ancestry. *A. The age-to-ancestry summary for class 4 and 5 variant carriers. The magenta line indicates the median age of 37 years, B. The sex-to-ancestry summary class 4 and 5 variant carriers, C. The genetic findings summary for the ACM cohort, D. Table and sunburst chart show the molecular consequences of class 4 and 5 variants.*

6.3.3.1.2.1.2 Class 4 and 5 variant carriers of Mixed ancestry (MA)

Participants of MA were a small group, and only 25% (2/8) of the probands carried class 4 or 5 variants, as shown in Figure 6.11.

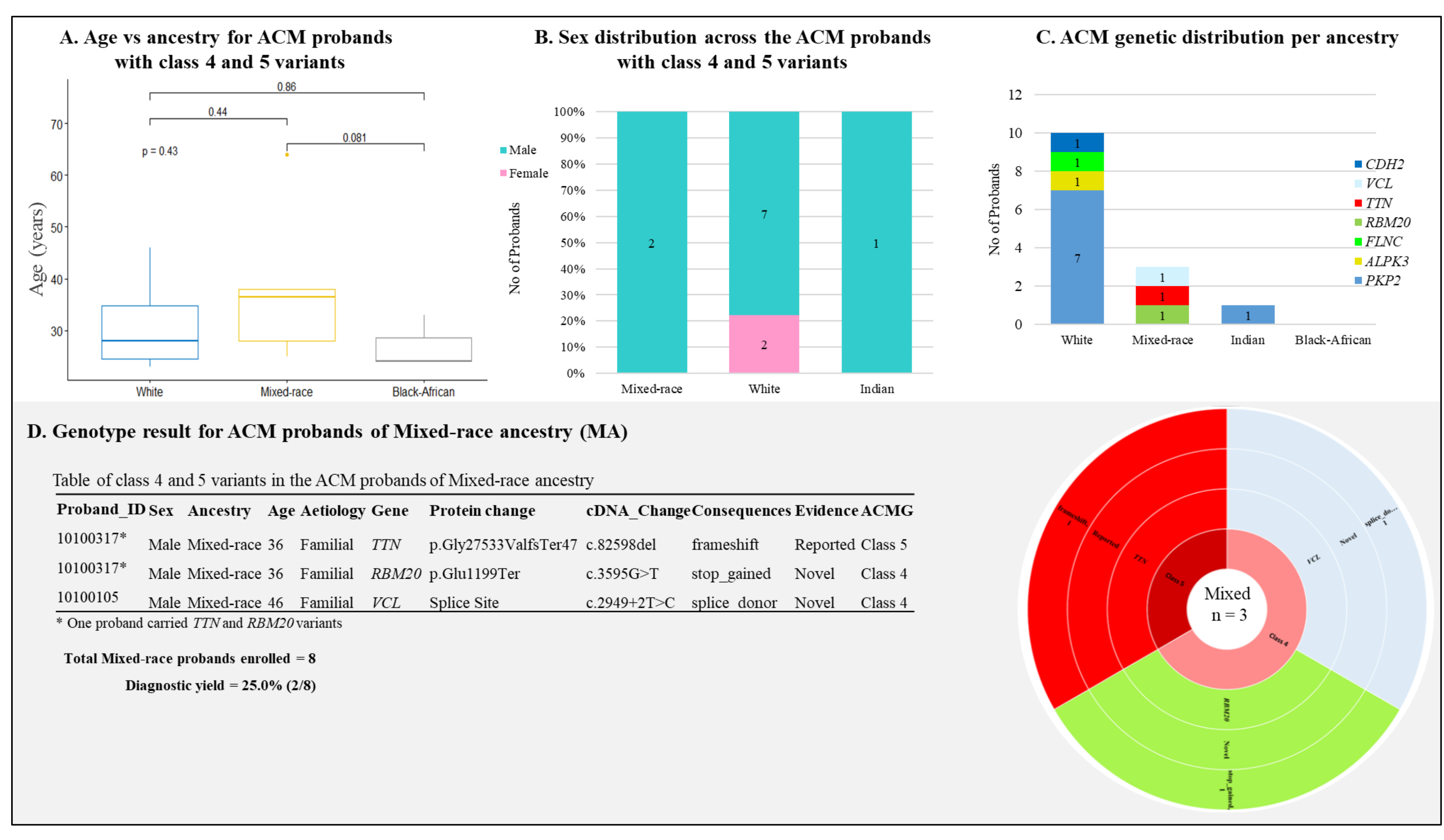
6.3.3.1.2.1.2.1 TTN c.82598del and RBM20 c.3595G>T

Proband (ID 10100317) is a 36-year-old MA male who had been classified as NYHA class II, with an ejection fraction of 20%; there was a FH of heart disease. Targeted sequencing identified the *TTN* c.82598del (p.Gly27533ValfsTer47) frameshift class 5 variant and the novel *RBM20* c.3595G>T (p.Glu1199Ter) stop_gain class 4 variant as the likely cause of disease. The novel *RBM20* c.3595G>T is on the 5' end of exon 40 in a highly conserved amino acid sequence region.

The *TTN* c.82598del (p.Gly27533Valfs*47) variant reported in this proband is on coding exon 327 of the *TTN* transcript NM_001267550.2 at chromosomal position 2:179428261 of the A-band region. The variant creates a translational frameshift sequence predicted to terminate at the 47th amino acid downstream. The *TTN* c.82598del variant is expected to produce a truncated mRNA that may undergo decay and result in titin haploinsufficiency or be misfolded and aggregated.

6.3.3.1.2.1.2.2 VCL c.2949+2T>C

Proband (ID 10100105) diagnosed with familial ACM at 46, classified as NYHA class II, with an ejection fraction of 67%, had a novel *VCL* c.2949+2T>C splice_donor variant. The c.2949+2T>C variant is predicted to disrupt RNA splicing at a highly conserved domain shared by vinculin and α-catenin. The splicing prediction tools showed that the c.2949+2T>C variant would alter the correct transcript splicing to create a truncated protein.



Proband_ID	Sex	Ancestry	Age	Aetiology	Gene	Protein change	cDNA_Change	Consequences	Evidence	ACMG
10100317*	Male	Mixed-race	36	Familial	*TTN*	p.Gly27533ValfsTer47	c.82598del	frameshift	Reported	Class 5
10100317*	Male	Mixed-race	36	Familial	*RBM20*	p.Glu1199Ter	c.3595G>T	stop_gained	Novel	Class 4
10100105	Male	Mixed-race	46	Familial	*VCL*	Splice Site	c.2949+2T>C	splice_donor	Novel	Class 4

Figure 6.11: ACM probands of Mixed ancestry. *A. The age-to-ancestry summary for class 4 and 5 variant carriers. The magenta line indicates the median age of 37 years, B. The sex-to-ancestry summary class 4 and 5 variant carriers, C. The genetic findings summary for the ACM cohort, D. Table and sunburst chart show the molecular consequences of class 4 and 5 variants.*

6.3.3.1.2.1.3 *Class 4 and 5 variant carriers of Indian ancestry (IA)*

6.3.3.1.2.1.3.1 PKP2 c.968_975dup

We recruited three participants of IA and identified one proband (33%) with a class 5 variant. The male proband diagnosed at the age of 49 with non-familial DCM classified as NYHA class I had a reported *PKP2* c.968_975dup (p.Ala326ArgfsTer29) frameshift variant. The variant replaced Alanine with Arginine at residue position 326, leading to a downstream premature stop codon on the 29th amino acid.

6.3.3.1.2.1.4 *Class 4 and 5 variant carriers of Black-African ancestry (BA)*

We noted that not many BA participants were diagnosed with ACM. Of the three BA probands recruited into the ACM cohort, no class 4 or 5 variant was found.

6.3.3.1.3 Familial segregation of class 4 and 5 variants

First-degree relatives of three (10100130, 10100390 and 10100142) probands were screened for the class 4 and 5 variants they carried using Sanger sequencing, and none of the relatives were found positive for the mutations.

6.3.3.2 Probands with variants of uncertain significance (class 3)

Sixteen ACM probands carried only class 3 variants: nine WA, three BA, three MA and one IA proband. The 16 probands had eight genes harbouring 25 class 3 variants listed in SS Table 2. The clinical characteristics of the ACM probands carrying only class 3 variants are presented in the next section. However, we identified three ACM-related class 3 variants of interest, which are presented in Table 6.4. The three reported class 3 variants are heterozygous missense mutations identified in *DSP* and *SCN5A* genes.

Table 6.4: Class 3 variants of interest in the ACM probands

Proband_ID	Sex	Ancestry	Age	Gene	Protein change	cDNA_Change	Evidence	ACMG
10100134	Male	White	40	*SCN5A*	p.Arg1195His	c.3584G>A	Reported	Class 3
10100104	Male	Mixed	34	*SCN5A*	p.Leu189Phe	c.565C>T	Reported	Class 3
10100133	Male	Mixed	37	*DSP*	p.Asn287Ser	c.860A>G	Reported	Class 3

6.3.3.2.1 Baseline characteristics

Data analysis of the ACM probands carrying only class 3 variants showed they were diagnosed at a mean age of 42.1 years (SD 16.7, range 15-75) in Table 6.2. There was thus no significant difference between the overall adult-onset ACM age of diagnoses and probands positive for only class 3 variants. We also found no significant difference in the age of ACM class 3 variant carriers across the various ancestries. However, half (8/16) of the class 3 variant carriers were male, and six were WA probands.

6.3.3.2.2 Diagnostic findings

We identified eight genes and 25 class 3 variants according to the ACMG/AMP guidelines in the 16 probands. In most of the class 3 variants, 60% (15/25) were reported, and 93.3% (14/15) were missense, while over a third (10/25) were novel. For complete details of the 25 class 3 variants information, including the genomic coordinates, SIFT, PolyPhen2, CADD, population frequencies, clinical annotation, and others, see Supplementary Table 1.

6.3.3.2.2.1 Class 3 variant carriers of White ancestry (WA)

We found nine class 3 variant carriers of WA in this ACM cohort: 66.7% (6/9) males and 33.3% females. The nine probands carried seven genes harbouring 15 class 3 variants. Interestingly, only 33.3% (3/9) of these probands had multiple class 3 variants.

Of the 15 class 3 variants, 53.3% (8/15) were reported in online resources, and 46.7% (7/15) were novel. The majority, 86.7% (13/15), of the class 3 variants had a missense molecular consequence, while the remaining two were frameshift variants.

Of the seven genes with class 3 variants, the *TTN* gene was found in 67% of the WA probands. However, there was no relevant ACM (or ARVC) gene, and we noted that an *SCN5A* c.3584G>A may be an ACM genetic modifier. The variant resulted in Arginine substitution by Histidine at residue 1195 out of the 2016 amino acids in the *SCN5A* coding DNA sequence.

6.3.3.2.2.2 Class 3 variant carriers of Mixed ancestry (MA): ACM cohort

We found three MA probands carrying class 3 missense variants: two males and one female. The first proband carried two reported variants: *DMD* c.3937A>Gand *DSP* c.5593A>T

variants. The second proband had a reported *DSP* c.860A>G variant, while the third carried two variants: a novel *TTN* c.64804G>T and a reported *SCN5A* c.565C>T variants. We also found that the *DSP* c.860A>G (p.Asn287Ser) and *SCN5A* c.565C>T (p.Leu189Phe) variants could be possible ACM modifiers.

6.3.3.2.2.3 Class 3 variant carriers of Black-African ancestry (BA): ACM cohort

We found three female BA probands carrying only class 3 variants. The first proband carried two reported missense variants: *TTN* c.1054G>A and *DES* c.935A>C variants. The second proband had a reported *ALPK3* c.904G>A missense variant, and the third one carried a novel *DSP* c.8508_8519del inframe_deletion variant.

6.3.3.2.2.4 Class 3 variant carriers of Indian ancestry (IA): ACM cohort

We found one Indian male proband carrying a novel *TTN* c.42509T>G missense variant.

6.3.3.3 Summary of ACMG classes 1 to 5 variants carriers: ACM cohort

Figure 6.12 shows the summarised genetic findings indicating the number of ACM probands with variants found according to the ACMG five-tier classification system.

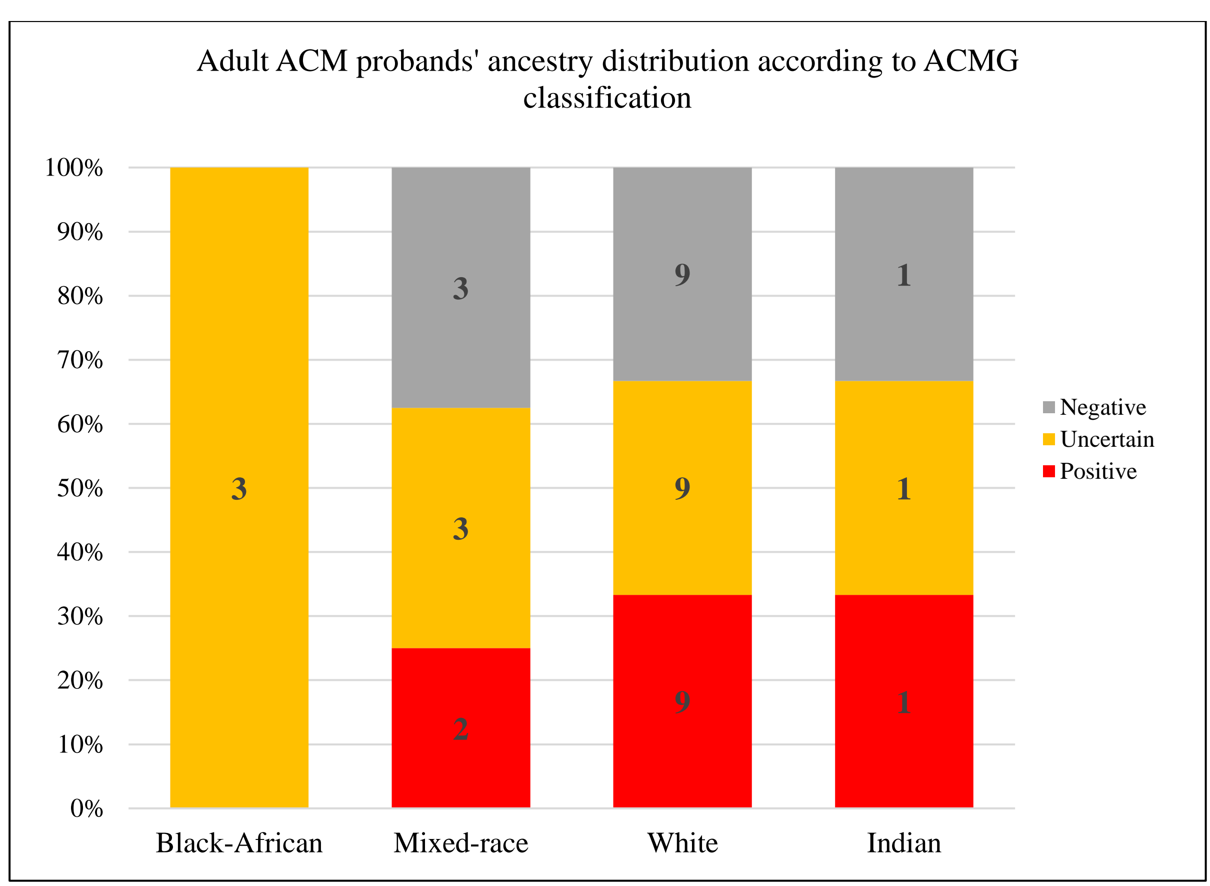


Figure 6.12: Summary of adult ACM classes 1 to 5 variants carriers. *The ACMG/AMP variants classification is categorised according to the cohort population ancestry.*

Due to the uniqueness of South Africa's population, we also explored the genetic yields across the three dominant population groups (Black, Mixed, Indian and White) to better understand the ACMG class distribution. Forty-one ACM probands were recruited, with the White population group being the most predominant with 27 probands, followed by the eight Mixed and three Black-African and Indian probands, respectively. Our data showed that the cohort was about 40.6 years old at the age of diagnosis, with a predominance of males.

Targeted sequencing was performed on our ACM cohort of 41 probands, and we found 12 probands carrying class 4 or class 5 variants: 10 probands carrying class 5 variants and two probands carrying class 4 variants. While 16 probands carried only class 3 variants and 13 probands carried class 1 or 2 variants.

Figure 6.8 summarises the ACM genetics yield across the cohort, indicating low class 4 and 5 variants yield and a significant yield for class 3 variants, with 71% of the probands being genotype negative/uncertain. Of note was the sizeable WA population in the ACM cohort, with 33.3% (9/27) class 4 and 5 and 33.3% (9/27) class 3 yields.

6.3.3.4 Possible founder variants in the ACM cohort

Four probands carried the known *PKP2* c.1162C>T (p.Arg388Trp) missense variant in both heterozygous and homozygous forms. The reported *PKP2* c.1162C>T variant is suggested in the literature as a possible South African founder ARVC-causing variant.

6.4 **DISCUSSION**

In Africa, the genetics of ACM is not well documented. Through our study called IMHOTEP, we determine genetic aetiology in 41 unrelated probands (≥ 13 years) using a targeted NGS panel test of 38 genes (Appendix F); the panel was designed for the routine clinical diagnosis of inherited CMO in the UK. Although this approach has identified disease-causing variants in established CMO genes only, the importance of the targeted NGS yield cannot be underestimated because ACM is a rare AD ventricular dysfunction that is a leading cause of SCD in young individuals, especially athletes.

6.4.1 ACM probands' baseline findings

At the time of diagnosis, our 41 probands had an average age of 40.6 years (SD 13.3), and 65.9% (27/41) of the cohort was male. Also, 65% of the participants self-reported as WA probands, 19.5% (8/41) were of MA, 7.3% (3/41) were of BA, and 7.3% (3/41) were of IA. Our initial data indicates that most probands with ACM were White males under the age of 41.

6.4.2 ACM probands' targeted sequencing findings

We used a comprehensive routine targeted NGS panel testing and a diagnostic yield of 29.3%, with 73.2% (29/41) of the ACM cohort having uncertain or negative genotype results. Of the 12 probands that were positive for class 4 and 5, nine were WA, three were MA, and one was IA, while 39% (16/41) probands were carrying class 3 variants, and 31.7% (13/41) probands were carrying only class 1 and 2 variants. The analysis and discussion of class 1 and 2 variants are beyond the scope of this thesis.

6.4.2.1 Pathogenic (class 4 and 5) variants

6.4.2.1.1 Class 4 and 5 variant carriers' clinical findings

The 12 class 4 and 5 variant carriers were younger, with a mean age of 35.9 years, compared to the cohort's average age of 40.6 years at diagnosis. Younger age of onset has been reported as a risk for ACM, with a global average age of about 45 years, especially in males (95, 327, 415). Generally, more young males have been diagnosed with ACM in European cohorts compared to the North American studies, wherein the proportion of patients of African descent

is higher (18, 191). Our findings were consistent with most literature: 83.3% (10/12) of genotype-positive probands are male.

Self-reported ancestry of the class 4 and 5 positive probands revealed that 75% (9/12) of our class 4 and 5 variant carriers were White at diagnosis. We also report that our age and sex assigned at birth findings are similar to that of most published research on ARVC (26, 29), whereby the disease occurs mainly in young male athletes. Thus, we suggest that an ACM with genetic aetiology in South Africa is more common in young male patients of WA (396). Specifically, we think ACM is common in young White males around the Western Cape province of South Africa because most of them participate in endurance physical activities such as rugby or mountain biking.

The positive yield is moderate compared to what has recently been reported by similar studies, with the global yields ranging from 17% to >60% (191, 192, 415). These 12 probands' relatives will benefit if screened on time because the IMHOTEP leading site offers care for both probands and asymptomatic genotype-positive relatives.

6.4.2.1.2 Class 4 and 5 variant carriers' genetic findings

Notably, we found multiple class 4 and 5 variants in 1.7% (2/12) of our genotype-positive probands, and it is worth noting that about 2% to 4% of familial ACM patients in literature are known multiple variants carriers (250). Further, we identified two pairs of digenic heterozygous variants in the sarcomere (*TTN*), nucleus (*RBM20*), desmosome (*PKP2*) and cytoskeleton (*ALPK3*) genes. The 11 ACM-causing variants are discussed in the following paragraphs, starting with the desmosome gene harbouring-most variants.

6.4.2.1.2.1 The desmosomal plakophilin-2 (*PKP2*) gene

Like in the literature, this study shows that 45.5% (5/11) of the ACM-causing variants occurred in the *PKP2* gene carried by 66.7% (8/12) of the genotype-positive probands (408). Similarly, literature has reported most desmosomal class 4 or 5 variants in the *PKP2* gene, with a prevalence ranging from 20% to 81% (95). As the introduction (section 6.1.7) mentions, *PKP2* is one of the primary desmosomal genes known to lead to ARVC.

Below, we describe the five *PKP2* variants found in eight probands of WA and IA.

6.4.2.1.2.1.1 PKP2 c.1162C>T (p.Arg388Trp)

The c.1162C>T (p.Arg388Trp) missense class 5 variant caused a LoF or haploinsufficiency of the *PKP2* gene protein product. The severe phenotype is due to abnormal plakophilin-2 protein structural remodelling exacerbated by physical exercise (382). The c.1162C>T variant has been reported to be a founder variant and has occurred within 11% of SA's ACM population (416). In the homozygous patients, the variant has a more severe phenotype exacerbated by physical exercise. The c.1162C>T variant caused a LoF or haploinsufficiency and, in its milder heterozygous form, allowed the mutation to be propagated within families (416). Our finding supports literature wherein ACM disease-causing classes 4-5 variants are mainly found in the proteins encoded by desmosomal gene (*PKP2*) (214, 242) and to a lesser extent in genes encoding for non-desmosomal proteins (e.g., *FLNC* and *TTN*) (250, 408).

6.4.2.1.2.1.2 PKP2 c.2198_2202del and ALPK3 c.2701C>T digenic heterozygous variants

The *PKP2* gene variants result in weakened connections between adjacent cardiomyocytes, leading to decreased tightness and increased detachment (244). Impaired cardiomyocyte adhesion can cause electrical imbalance, immune infiltration, structural remodelling, tachycardia, and ventricular fibrillation. Meanwhile, alterations in the *ALPK3* gene sequence can cause sarcomere M-band structural abnormalities, disrupt ion channel regulation, affect several signalling pathways, and promote abnormal cardiomyocyte remodelling and apoptosis (289).

The *PKP2* c.2198_2202del (p.His733ProfsTer8) class 5 variant is an established disease in a mutation hotspot known for causing right ventricular ACM. For example, a White male clinically diagnosed with ARVC by researchers in the UK who later died in sleep at the age of 37 years (393) was said to have died of this variant. The affected proband in this study also carried another mutation, a nonsense *ALPK3* c.2701C>T (p.Gln901Ter) class 4 variant predicted deleterious by *in silico* tools. Also, a pathogenic *ALPK3* truncating variant in the affected exon 6 has been functionally documented to promote abnormal cardiomyocyte remodelling and apoptosis (336). We hypothesise that the combined effect of the variants could be exacerbating the disease severity, especially as our proband was diagnosed as an adolescent compared to other patients carrying only the *PKP2* c.2198_2202del variant (250).

6.4.2.1.2.1.3 PKP2 c.968_975dup (p.Ala326ArgfsTer29)

PKP2 LoF variants are the primary cause of ARVC (250), and similarly, the frameshift *PKP2* c.968_975dup (p.Ala326ArgfsTer29) creates an eight bp duplication variant reported in the literature to have caused severe ARVC (417). The alteration mentioned once in the ClinVar database is in a mutation hotspot surrounded by several reported pathogenic delins variants where LoF is a known mechanism in the affected ACM patients (242).

6.4.2.1.2.1.4 PKP2 c.1709del (p.Ala570ValfsTer7)

The known *PKP2* c.1709del frameshift variant is also reported in the literature in ARVC patients of WA (418). Truncating variants LoF in plakophilin-2 protein products have not been functionally validated for ACM; however, *PKP2* haploinsufficiency pathomechanism is expected to cause ACM whereby the cardiomyocyte-to-cardiomyocyte connections are weakened (244). The weakened adhesion will increase neighbouring cardiomyocytes' detachment from each other and cause electrical imbalance or arrhythmias (419).

PKP2 splice donor c.2489+1G>A

The *PKP2* c.2489+1G>A variant is a known pathogenic canonical splice site donor mutation that has also been associated with early-onset ARVC (386). A haplotype study has suggested that this variant is a Dutch founder mutation (420). Although the variant has been reported in patients of various ancestry, the variant carrier in this study is a WA proband from the Cape Town region with a history of Dutch settlers (27).

Pathogenic variants that have disrupted this desmosome-coding gene (e.g., *PKP2* missense variants) generally create structural instability of the myocardium, predisposing patients to cardiac arrhythmias (386, 412). The *PKP2* gene, amongst other desmosome genes, encodes cell-to-cell membrane proteins that maintain cells' structural integrity in various tissues, including the myocardium. For example, the *PKP2* gene encodes the desmosomal plakophilin-2 protein found primarily in the heart muscle's cardiomyocyte-to-cardiomyocyte boundaries, nucleus, and cytoskeleton. The plakophilin-2 protein is composed of armadillo repeats, which connect the desmoplakin to desmoglein-2 protein filaments to maintain the integrity of the desmosomes. Alterations in the *PKP2* gene sequence will weaken the connection between adjacent cardiomyocytes, thereby decreasing the cardiomyocyte-to-cardiomyocyte tightness and increasing detachment (244). The impaired cardiomyocyte-to-cardiomyocyte adhesion

(Figure 6.2) will increase electrical impulse imbalances and induce immune cell infiltration, resulting in abnormal structural remodelling, tachycardia, and ventricular fibrillation (394). Arrhythmias are critical features of ACM, mainly ventricular fibrillation, commonly associated with early onset age.

6.4.2.1.2.2 The adherens junction gene (*CDH2*)

6.4.2.1.2.2.1 Cadherin 2 (CDH2) class 5 variant

6.4.2.1.2.2.1.1 CDH2 c.1219G>A (p.Asp407Asn)

The *CDH2* c.1219G>A (p.Asp407Asn) variant identified in a WA proband in this study was first published by Mayosi and others as a likely pathogenic ARVC-causing variant (407). It is worth noting that the exact pathomechanisms underlying the pathogenicity of this variant may vary, and the available information may be updated as scientific research progresses (391).

The *CDH2* gene encodes a classical cadherin 2, a cadherin superfamily member. Multiple transcript variants can be created through alternative splicing, with one of them encoding a preproprotein. This preproprotein undergoes proteolytic processing to form a calcium-dependent cell adhesion molecule and glycoprotein that plays a vital role in various biological processes. These processes include the development of the nervous system, establishing left-right asymmetry, and forming bone and cartilage. The calcium-dependent cell adhesion protein is responsible for homotypic cell-cell adhesion by dimerising with a *CDH2* chain from another cell. Cadherins also help in sorting different cell types and regulate neural stem cell quiescence by anchoring neural stem cells to ependymocytes in the adult subependymal zone. However, when *CDH2* is cleaved by MMP24, the anchorage is affected, leading to the modulation of neural stem cell quiescence (421).

6.4.2.1.2.3 Other genes

This study found two (~2%) possible digenic variants' carriers. The presence of a class 4 or 5 variant in each of two different genes, each of which contributes to the observed phenotype (digenic heterozygous), is reported in 2% to 4% of ACM patients (250). The first proband carried *PKP2* c.2198_2202del and *ALPK3* c.2701C>T discussed above, and the second proband had *TTN* c.82598del and *RBM20* c.3595G>T class 4 and 5 variants.

6.4.2.1.2.3.1 TTN c.82598del and RBM20 c.3595G>T digenic heterozygous variants

ACM patients carrying digenic variants have been reported to present with severe phenotypes and arrhythmias (250). Notably, the *TTN* and *RBM20* are dilated CMO phenotype-related genes. Generally, in CMO patients, *TTN*tv in the A-band region is the most common cause of DCM, and the pathomechanism causing ACM is unclear (205). However, the severe arrhythmias observed in our patient may have been worsened by the *RBM20* c.3595G>T, a novel class 4 variant. Most pathogenic variants in the *RBM20* gene are associated with ventricular arrhythmias in CMO (192). Unlike this study, one study has reported both genes in a severe DCM (422), which suggests that ACM could be stratified by affected genes, not only clinical phenotype (423).

The *TTN* gene encodes for a giant elastic protein called titin, which connects the Z-disc to the M-line and other sarcomere proteins. The prominent role of titin is to generate force for contraction in striated muscle tissues such as the myocardium, as described in section 3.7.1.4.1. On the other hand, the *RBM20* gene product (described in section 3.7.1.4.2) regulates the mRNA splicing of genes encoding key structural proteins such as titin involved in cardiac development. The *TTN* gene encodes various sizes of titin protein due to the alternative splicing regulatory mechanism of *RBM20* (376). *RBM20* coding DNA sequence variants can alter titin isoform expression level or change sarcomere protein assembly and organisation, affecting titin elasticity and sarcomere structural integrity. Also, *RBM20* variants may result in Ca^{2+} dysregulation due to disrupted cardiac excitation-contraction coupling pathways. Poor-quality (e.g., shortened or misfolded) *RBM20* gene product will cause sarcomere protein haploinsufficiency, disruption of various signalling pathways, or myocyte disarray and fibrosis (227).

6.4.2.1.2.3.2 FLNC c.1069C>T (p.Gln357Ter)

There is no evidence in the literature to associate this variant with ACM, although it has been submitted twice in the ClinVar database. We therefore hypothesise that the *FLNC* c.1069C>T variant would yield a shortened or no protein product, resulting in filamin-C LoF in the patient. *FLNC* LoF is a known pathomechanism for ACM characterised by an impaired myocardial ability to withstand the stress of contraction and relaxation (235).

Wildtype filamin-C (*FLNC*) is one of the three related gamma genes that encodes the filamin-C protein described in section 3.7.1.4.4. *FLNC* is a classical left (or bi-) ventricular ACM

disease-causing gene, and its truncating variants are associated with a high risk of ACM (and DCM) (235). Most researchers have associated *FLNC* disease-causing variants on the ROD 1 region of the protein with ACM phenotype (424). The clinical key features of ACM due to *FLNC* disease-causing variants include normal to low voltage EGC and subepicardial enhancement on the image (250).

6.4.2.1.2.3.3 VCL splice site donor c.2949+2T>C

The VCL truncated protein will lose its vinculin α-catenin domain, an actin filament linker found in all types of Adherens junctions (425). However, the *VCL* gene LoF is an unknown mechanism for adult CMO (151). More analysis of the VCL c.2949+2T>C variant is needed as it may cause the proband's heart muscle to lose its structure and suffer damage. This novel variant can also lead to electrical impulse imbalances and collagen deposition in myocardial tissue (172).

The natural vinculin (*VCL*) gene encodes a membrane-cytoskeletal protein associated with the cell-cell or cell-matrix junctions' linkage (426). Vinculin protein is important in anchoring the actin cytoskeletal complex to the integrin- and cadherin-containing adhesion molecules on the cell membrane (172). This role is essential for embryonic development to create integrin clusters, promote force generation and possibly regulate apoptosis. Alterations in the *VCL* gene sequence will weaken the adhesion molecules' connection between adjacent cardiomyocytes, thereby increasing electrical impulse imbalances and abnormal structural remodelling.

6.4.2.2 Uncertain significance (class 3) variants

Despite the complexity of implementing the ACMG classification guidelines for this form of CMO, this study identified 17 probands carrying only class 3 variants. Some of these class 3 variants are possible genetic modifiers of ACM because both genotype and environmental modifiers, such as sport, have played a considerable role in the progression of the disease (250, 415). We suggested further investigation of these class 3 variants in a matched-control cohort. These class 3 variant carriers yield 41.5% (17/41), about twice the proportion of those with class 4 or 5 variants (29.3%). The targeted NGS approach has boosted the detection of these probands with only class 3 variants. The identification of deleterious variants in multiple

affected family members could be used to identify "at-risk" and asymptomatic relatives in further experiments (139).

Contrary to the genotype-positive group with a mean age of 37.8 years, these 17 probands had a mean (SD) age of 42.1 (16.7) years at diagnosis (Table 6.2; i.e., about five years older). However, in the literature, ACM patients have mostly present with the disease at a much later age (327). Thus, we strongly recommend that some of the class 3 variants, such as the *DES* c.860A>G (p.Asn287Ser), should be further investigated.

What is also seen is that 40% (10/25) of the class 3 variants are novel. Even though the group size is small, the MA (and BA) group also shows a similar pattern to the majority WA group, with 17.6% of the probands having class 3 variants.

6.4.2.2.1 Genetic modifiers

6.4.2.2.1.1 The *SCN5A* c.3584G>A (p.Arg1195His) and c.565C>T (p.Leu189Phe) variants

Both *SCN5A* variants this study reports are carried by males diagnosed with non-familial phenotype. One MA proband diagnosed at the age of 34 with non-familial DCM had the *SCN5A*:c.565C>T, whereas the *SCN5A*:c.3584G>A was reported to be carried by a WA diagnosed at the age of 40.

The *SCN5A*:c565C>T (p.Leu189Phe) missense variant is in exon 5 of the ion transport domain of the *SCN5A* coding DNA sequence. Substituting Leucine with Phenylalanine in the ion transport domain of the protein can have several physiological effects. Both amino acids are hydrophobic, but Phenylalanine has a larger and more rigid aromatic side chain than Leucine. This alteration can influence the hydrophobicity and packing of the side chain within the protein's structure. Leucine is often found in hydrophobic cores, contributing to stability, and Phenylalanine can similarly participate in hydrophobic interactions. However, the amino acid substitution may modify the SCN5A protein's ion transport domain structure. Leucine residues are known for their involvement in protein-protein interactions, and the switch to Phenylalanine may impact these interactions due to differences in sidechain characteristics. Additionally, mutations in the affected amino acid residue 189 regions are associated with genetic disorders, emphasising the importance of understanding the mutation's context and potential implications for ACM disease mechanisms (256).

The *SCN5A*:c.3584G>A (p.Arg1195His) missense variant is in exon 20 of the sodium ion transport-associated domain of the *SCN5A* coding DNA sequence. Replacing an Arginine with a Histidine in a protein can have significant physiological effects. While Arginine has a positive charge, Histidine can be either positively or neutrally charged depending on the pH level of its surroundings. This change in charge can impact the protein's interactions with other molecules, ultimately affecting its overall function. The close pKa of Histidine to physiological levels allows it to participate in the protein's pH-dependent processes and proton transfer reactions. Additionally, the substitution may impact the structure of the sodium ion transport-associated domain, potentially affecting its stability or interactions (427). Since this domain is critical for proper functioning and is conserved across 12 different species, it is essential to understand this variant's potential implications for ACM disease mechanisms. Mutations around the amino acid residue 1195, which is affected by this substitution, are linked to cardiac arrhythmias, highlighting the need for careful consideration of this substitution's ramifications (248).

6.4.2.2.1.2 The *DES*: c.860A>G (p.Asn287Ser) variant

The D*ES*: c.860A>G (p.Asn287Ser) missense variant is in exon 4 of the intermediate filament protein domain of the *DES* coding DNA sequence. The altered Asparagine amino acid is in a highly conserved region, and its replacement with Serine can result in a range of physiological effects (369). Asparagine has a larger polar side chain capable of hydrogen bonding, while Serine has a smaller side chain. The p.Asn287Ser substitution can alter the hydrogen bonding patterns within the desmin protein, which may impact the local structure. The size and structure alteration may affect the protein's function, mainly because the Asparagine residue plays a significant role in catalysis, binding, or structural stability. Moreover, Asparagine residues are often targets for post-translational modifications such as glycosylation, and substituting Asparagine with Serine may remove a potential glycosylation site, thus impacting protein glycosylation patterns and functions. Lastly, The p.Asn287Ser substitution is linked to genetic cardiac disorders, underscoring the significance of understanding the variant's potential implications for ACM disease mechanisms.

Significantly, this study reported 30 class 3 variants in 11 CMO disease-causing genes, and the three modifiers described above barely met the class 4 variants' criteria. Additionally, over 33% (10/30) of the class 3 variants were in the desmosomal genes, and other studies have shown that desmosomal desmoplakin harbour modifier variants, which are often triggered by physical

exercise (193). However, experimental studies and structural analyses are essential for comprehensively understanding the physiological consequences of these class 3 variants.

6.4.2.3 Possible founder variants

The missense *PKP2* c.1162C>T variant was previously reported in four White South Africans diagnosed with ARVC sharing a common haplotype (416). The homozygous *PKP2* c.1162C>T variant carriers had a very early onset age and severe ARVC. Similar outcomes of very young onset age in patients carrying *PKP2* homozygous disease-causing variants have also been reported by other researchers (250, 417). The heterozygous carriers have a milder phenotype, which allows the variant to be passed down from generation to generation (416).

The impact of the *PKP2* c.1162C>T founder variant is enormous for these families, especially as it is an inherited condition. Also, the cost of genetic testing and counselling and the risk assessment to understand the full impact of the variant could be time-consuming and expensive. However, family members who want to know their status can consider lifestyle changes such as reproductive, dietary, and physical exercise planning. At least 30% to 50% of ACM cases have also been facilitated by non-genetic factors such as inflammation or exercise (382).

6.5 **CONCLUSION**

Heritable ACM in young adults is one of the leading genetic causes of CMO morbidity and mortality globally. It has been shown in this study that ACM in 29.3% of our young adult probands was caused by heritable disease-causing variants. We understand that our genotype-positive probands are moderately young because the global average age at diagnosis for adult ACM probands is greater than 40 years. Also, most of these probands were males, which is like what we know, whereby males and females are unequally affected by ACM disease-causing. Further, the probands carrying class 4 and 5 variants yielded 33.3% (9/27) for the WA, 25% (2/8) for the MA and 33.3% (1/3) for the IA probands.

Although targeted NGS might result in a lower carrier rate detection, it cannot be underestimated due to its clinical and research implications, as 12 probands had ACM-causing variants out of 41. Our diagnostic yield could be improved if additional information, such as segregation analysis for some class 3 variants, is considered. A review of the ACMG guidelines for class 3 variants may also be beneficial, as our current strategy is quite stringent. Despite the complexity of implementing the ACMG classification guidelines, we identified 16 (39%) probands with only class 3 variants, and this will facilitate the work of an expert team that will be assigned to perform genotype-phenotype correlations.

We found that the *PKP2* class 4 and 5 variants caused 66.7% of the heritable ACM in this cohort, and about 41.7% of the genotype-positive probands had familial ACM. Additionally, this study has shown that participants with class 4 and 5 *PKP2* disease-causing variants were frequently diagnosed with ACM. The finding fits nicely into the broader context of genetic ACM and the literature, and most importantly, possible avenues for future research should include the genealogy tracing of the *PKP2* c.1162C>T variants and potential recruitment of population controls.

The study has addressed the aim in that we have identified ACM-causing variants in 12 adult ACM probands using targeted NGS and an additional 16 probands as having class 3 variants.

Chapter 7: A summary of the IMHOTEP study

The purpose of this chapter is to provide an overview of the IMHOTEP genetics study results and description explored within the context of the specific cohorts in chapters 3 (DCM), 4 (HCM), 5 (RCM) and 6 (ACM).

Through this project, we sought to characterise the genetic determinants in probands clinically diagnosed with cardiomyopathy (CMO) using targeted NGS. We described a cohort comprising 690 (594 adult-onset and 96 paediatric) probands spanning four cohorts: DCM (n=539), HCM (n=67), RCM (n=43) and ACM (n=41).

7.1 **Baseline characteristics for the probands in the IMHOTEP study**

The cardiomyopathy cohort comprising 690 probands (594 adults and 96 paediatric) was recruited at various tertiary hospitals across South Africa. The probands were diagnosed with CMO of various clinical aetiologies, with 25% (171/690) having no reported familial history, 52% (359/690) having non-familial, and 23% (160/690) having no data. Notably, at enrolment, 20% (136/690) of these probands were diagnosed with peripartum CMO (PPCM).

We further analysed the baseline clinical data and found that overall, the mean age at diagnosis was 31.7 years, with 52.8% (364/690) of the probands recorded as female at birth. Self-reported ancestry found 55% Black-African, 32% Mixed, 12% White and 1% Indian probands.

7.1.1 Adult-onset Cardiomyopathies

All adult cohorts displayed a remarkably young mean age at diagnosis compared to the European populations with a similar study design (186). The European cohorts had an overall mean age of onset of 55 years, compared to South Africa (SA), where the overall mean age of onset was much younger at 31.7 years. This observation also held true for the four cohorts (Figure 7.1). When we further analysed the cohorts according to the sex assigned at birth, we noted that 52.8% of the adult IMHOTEP probands self-identified as females.

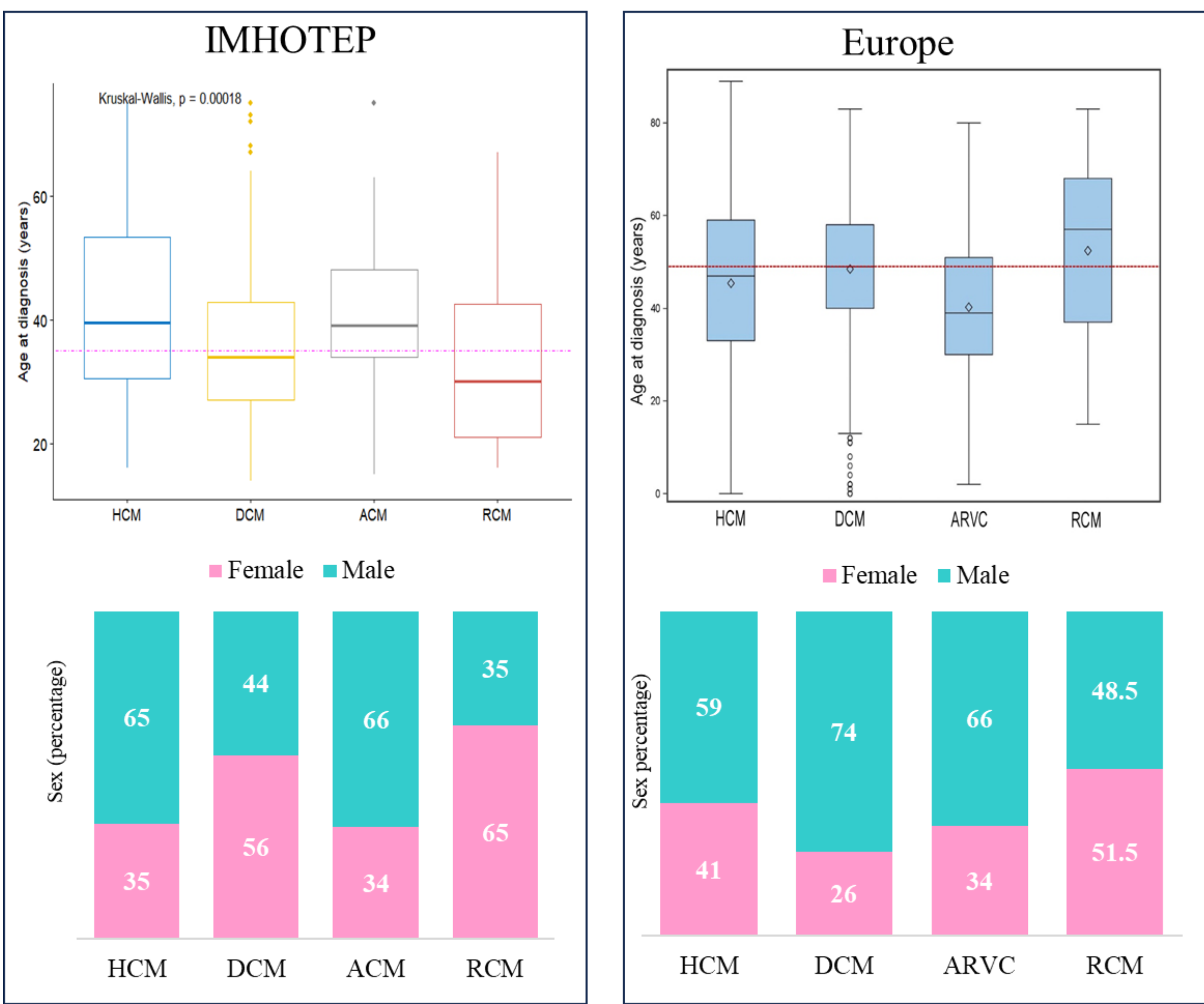


Figure 7.1: Adult-onset IMHOTEP probands' age and sex compared to a European cohort. *The adult-onset IMHOTEP probands' age and sex assigned at birth compared to a European cohort. ACM-arrhythmogenic cardiomyopathy, ARVC-arrhythmogenic right ventricular cardiomyopathy, DCM-dilated cardiomyopathy, HCM-hypertrophic cardiomyopathy, RCM-restrictive cardiomyopathy. The magenta line indicates the median age of 35 for the 594 adult probands (186).*

We then looked at the ancestry of the adults, and something that is uniquely South African is the various ethnic populations that make up this country. Through this pilot project, we compared the South African cohort to that of the European cohorts and were able to draw unique insights and conclusions. We found that 52.7% were Black-African, 33% were Mixed, 13% were White and <1% were Indian ancestry probands, Figure 7.2.

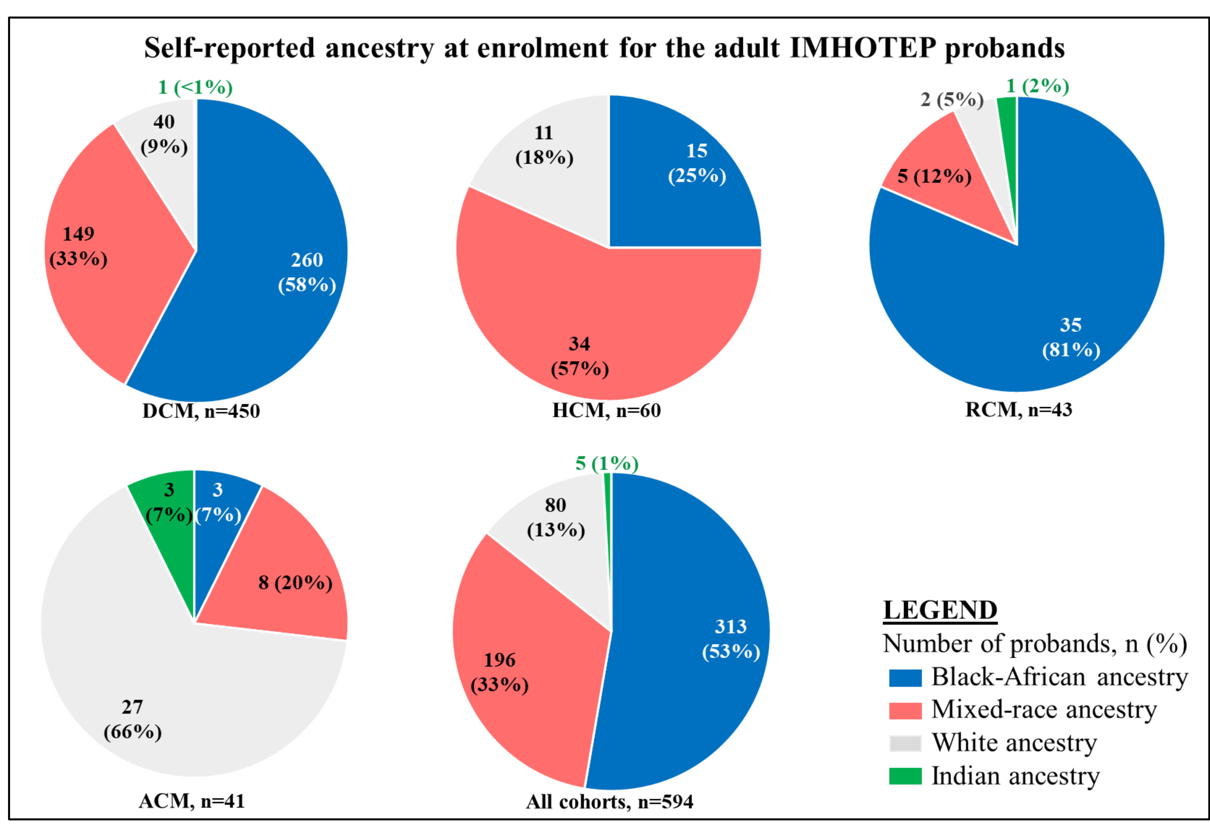


Figure 7.2: The various ancestry of adult IMHOTEP probands in this study.

7.1.1.1 Dilated Cardiomyopathy (DCM)

The DCM cohort constituted 75.8% (450/594) of the adult IMHOTEP study probands. The adult DCM probands had a mean age of 35.6 years at diagnosis compared to the European mean age of onset, which was about 50 years, and the genotype-positive probands had a mean age of 36.4 years. Similar average ages have been reported in South Africa for adult patients presenting with DCM (184); however, the probands' early onset ages are significantly low compared to the literature (186).

At baseline, 56% of the adult probands were female; however, after the genetic testing, 58% of the genotype-positive probands were males. As there were more females at baseline, we had a look to see if there was a genetic difference, but there was not. Also, the male-to-female ratio for the genotype-positive is the same as in the literature. We therefore hypothesise that more women were seen at enrolment due to peripartum cardiomyopathy, and their genetic landscape is unknown. The female preponderance could also be due to changes in sex hormones such as testosterone in males (137) and oestrogens in pregnant females (163), which are known risk factors for adult-onset DCM (explained in Chapter 3, discussion section 3.4.1.2.1.1).

The dominant populations recruited were BA (58%) and MA (33%). We are aware that recruitment bias could have skewed the DCM population results, but as these ratios were reflecting the reference population demographics of the Western Cape, which is 5:3:1:1 (Black Africa: Mixed; White; Indian), the values might reflect the true prevalence of cardiomyopathy in South Africa. Globally, BA adult participants are least represented in major studies, with about 1.6% in the UK (143) and 11.9% in the USA (192). The genetic results of the MA probands were unique, as they showed a combination of the BA and WA findings. Additionally, it was observed that most of the MA probands carried the same *LMNA* variant linked to SCD or the need for cardiac transplantation (219).

7.1.1.2 Peripartum Cardiomyopathy (PPCM)

The PPCM cohort constituted 30.2% (136/450) of the adult DCM study probands. The PPCM probands had a mean age of 29.6 years at diagnosis, and the genotype-positive probands had a mean age of 28.9 years. The genotype-positive probands' early onset ages are significantly low, especially when compared to the literature (158, 159). However, the factors contributing to the age difference between the genotype-positive and at diagnosis remains unknown.

BA (68%) and MA (31%) were the dominant populations recruited. This is expected because these ratios reflect the reference population demographics of the Western Cape. However, PPCM affects more Black females than other ethnic groups globally (159).

7.1.1.3 Hypertrophic Cardiomyopathy (HCM)

The HCM cohort constituted 10.1% (60/594) of the adult IMHOTEP study probands. The adult HCM probands had a mean age of 41.3 years at diagnosis compared to the European mean age of 45, and the genotype-positive probands had a mean age of 33.9 years. The genotype-positive probands' early onset ages are significantly low, especially compared to the literature (186). However, the factors contributing to the age difference between the genotype-positive and at diagnosis remains unknown.

At baseline, 65% of the adult probands were males.

The dominant populations recruited were MA (57%) and BA (25%). This was not expected because these ratios did not reflect the reference population demographics of the Western Cape. Globally, over 95% of adult HCM participants in the literature have been non-African patients (307, 327).

7.1.1.4 Restrictive Cardiomyopathy (RCM)

The RCM cohort constituted 7.2% (43/594) of the adult IMHOTEP study probands with a mean age of 33.0 years at diagnosis, and the genotype-positive probands had a mean age of 26.0 years. The European age of onset was about 56 years, compared to 33.0 years in the RCM cohort, which is much younger in this study. The genotype-positive probands' early onset ages are significantly low, especially compared to the literature (186). However, the factors contributing to the early onset age remain unknown.

At baseline, 65% of the probands were females, and all the genotype-positive probands were also females. Notably, about 90% of the RCM probands had endomyocardial fibrosis, a disease frequently diagnosed in Mozambique, where the majority of the RCM probands were recruited (56, 361).

The dominant populations recruited were BA (81%) and MA (12%); this is expected because these ratios reflected the reference population demographics of Maputo, Mozambique. Very small numbers (<15) of adolescent RCM participants of African descent are reported in the literature (42).

7.1.1.5 Arrhythmogenic Cardiomyopathy (ACM)

The ACM cohort constituted 6.9% (41/594) of the adult IMHOTEP study probands. We noted that our ACM cohort had approximately the same mean age of onset as that of the European (≈41 years); however, the genotype-positive probands were younger (mean age = 35.9). The genotype-positive probands' early onset ages are significantly low, especially compared to the literature (186). However, the factors contributing to the early onset age remain unknown. The reason for this has not been elucidated, but one reason could possibly be the convergence of communicable and NCDs where a myriad of external factors, such as infectious diseases, diet, and other environmental pressures, now impact NCDs. Other reasons could be modifying risk factors that predispose our patients to early onset of CMOs, such as gene-environment interactions or greater awareness and good healthcare-seeking behaviour.

At baseline, 65.9% of the adult probands were males, and after the genetic testing, 83.3% of the genotype-positive probands were also males.

The dominant populations recruited were WA (66%) and MA (20%); this is expected. Other researchers in our group have previously reported that 46% of ARVC participants in South

Africa are European descendants (155). Globally, <2% of adult ACM participants in the larger cohorts have been Black patients (98, 415). We are aware that recruitment bias could have skewed the ACM population results, but these ratios did not reflect the reference population demographics of the Western Cape.

7.1.2 Paediatric Cardiomyopathies

The paediatric cohort constituted 13.9% (96/690) of the IMHOTEP study probands: DCM (n=89) and HCM (n=7). Most of our paediatric probands had childhood CMOs (96). The sex distribution between the South African IMHOTEP paediatric probands and the USA cohort (42) showed that males and females were equally diagnosed with CMO (USA 51.3% vs SA 49.0%). At diagnosis, 70% of the IMHOTEP paediatric probands were of BA.

7.2 **Genetic findings for the IMHOTEP study**

This study revealed a combined positive genetics diagnostic yield of 15.7% (108/690) in the IMHOTEP study. The adult probands yielded 17.8% (106/594), as shown in Figure 7.3, and the paediatrics yielded 2.1% (2/96). For adults, CMO disease-causing genes with class 4 or 5 variants were found in 16.9% (76/450) of DCM probands, 23.3% (14/60) of HCM probands, 9.3% (4/43) of RCM probands, and 29.3% (12/41) of ACM probands. About half, 47.8% (284/690) of the IMHOTEP probands had uncertain genetic results, and 34.3% (204/594) yielded negative genetic results.

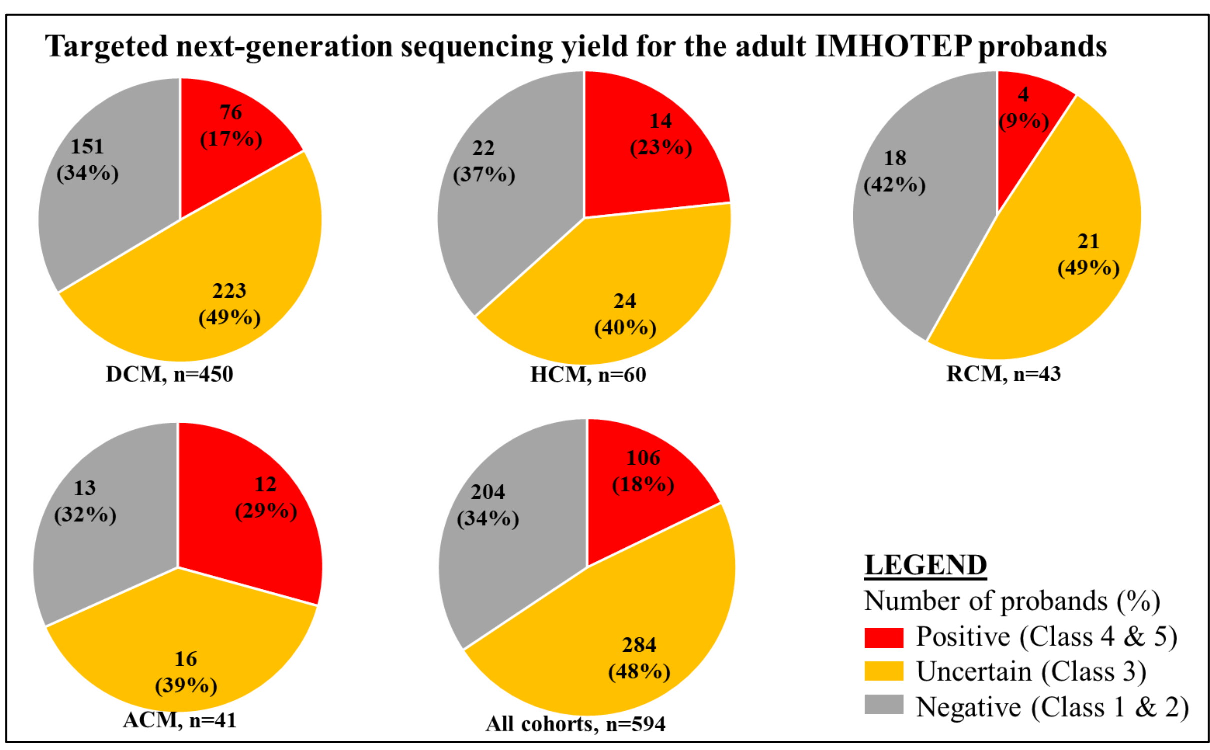


Figure 7.3: Genetics results for 594 adult IMHOTEP probands. *The result was categorised as Positive-ACMG/AMP classes 4 and 5 variant carriers, Uncertain-class 3 variant carriers, and Negative-classes 1 and 2 variant carriers. ACMG/AMP- the American College of Medical Genetics and Genomics/Association for Molecular Pathology.*

7.2.1 Adult-onset Cardiomyopathies

7.2.1.1 DCM probands carrying class 4 and 5 variants

Through IMHOTEP, we have shown that DCM is the most dominant form of adult cardiomyopathy in South Africa. We show that the adult DCM probands had a positive genetic yield of 16.9% (76/450). The 16.9% yield is comparable to other large cardiomyopathy studies in Norway, where the diagnostic yield was 14.1% (95), in the UK at 16% (143), and in Thailand

at 16.1% (299). However, many other extensive studies have reported DCM yields greater than 25% (17, 145, 146).

Our diagnostic result also differed across the various subpopulation groups in the IMHOTEP adult DCM cohort, with the BA yielding 15.0%, the MA at 16.1% and the WA at 32.5%. The low yield could signify that the panel of 38 cardiomyopathy genes was ineffective for African probands. Notably, the panel of genes we utilised was specifically designed for inherited CMO and used for routine diagnosis in the UK. The ACMG classification stringency could also contribute to our low yield.

The adult DCM probands had a wide range of genetic variants; however, 68.4% (52/76) of the probands were found to carry a class 4 or class 5 *TTN*tv. Amongst these *TTN*tv, the c.95008C>T (p.Arg31670Ter) stop_gained variant had a prevalence of 18% (14/78) and was the most recurrent as it was found in 14 probands. Additional recurrent variants were identified that would need further investigation to determine if they are true founder variants.

7.2.1.2 PPCM probands carrying class 4 and 5 variants

The positive genetics result for our combined PPCM cohort was 9.6% (13/136). The IMHOTEP adult HCM cohort (SA study) had a diagnostic yield of 23.3% compared to Germany; 10% to 16.4% (158, 159), Israel; 17.1% (158) and Denmark; 18.2% (158).

The genetic aetiologies of the adult PPCM probands were mainly due to sarcomere protein alteration. PPCM disorder was associated with two genes (*TTN* and *BAG3*) in 92.3% of the genotype-positive probands.

7.2.1.3 HCM probands carrying class 4 and 5 variants

The positive genetics result for our adult HCM cohort was 23.3% (14/60). The adult HCM diagnostic yield of 23.3% is high compared to Norway; 11.9% (95), Singapore; 16.5% (285) and France; 20.3% (428). However, the percentage of HCM cases in the adult IMHOTEP cohort was lower compared to the global average, with a yield of over 45% (429, 430).

The genetic aetiologies of the adult HCM probands were mainly due to sarcomere protein alteration. Generally, HCM disorders are mostly associated with eight sarcomere genes (e.g., *MYBPC3*). Figure 4.6 shows that the adult HCM probands with class 4 or 5 variants in *MYH7*

(40%) and *MYBPC3* (27%) carried about 71% of the disease-causing variants curated. This agrees with the range reported in similar published studies.

7.2.1.4 RCM probands carrying class 4 and 5 variants

The IMHOTEP RCM cohort (SA study) had a diagnostic yield of 9.3% (4/43) compared to the Netherlands; 25% (191). This was the cohort with the least yield. Most of these adult RCM probands had other causes of the disease, such as endomyocardial fibrosis, which might have exacerbated the RCM phenotype. As expected, no RCM-causing variants were identified.

Of the 36 BA recruited, only 3% (1/36) of the probands were found to have a *DES* class 5 variant, while 60% (3/5) of the MA had *DES*, *FLNC*, and *MYH7* class 4 and 5 variants. The rest of the RCM cohort were negative. The low yield could be due to the fact that the participants with RCM were recruited from Mozambique, where they had been diagnosed with EMF.

7.2.1.5 ACM probands carrying class 4 and 5 variants

Our adult probands findings showed that the IMHOTEP ACM cohort had the highest diagnostic yield, 29.3% (12/41). However, the ACM yield was low when compared to other studies from Switzerland with 59.5% (431), China; 56.8% (432) and the USA; 32% (417).

Like in the European cohorts, the desmosome protein complexes caused ACM genetic aetiologies in this study. Notably, in European cohorts, the desmosomal gene (e.g., *PKP2*) variants are common causes of ACM, especially those associated with the younger age of onset. In our ACM cohort, the *PKP2* variants accounted for disease in 62% of our ACM probands. Also, 50% of our WA ACM probands with a positive result carried the *PKP2* c.1162C>T class 5 variants. The c.1162C>T variant is a previously reported South African ACM founder mutation associated with mild to severe disease early onset.

7.2.1.6 Summary of classes 3 to 5 variants across the adult IMHOTEP probands

Figure 7.4 summarises the class 3 to 5 variants across the whole adult IMHOTEP cohorts. Overall, 21.5% (84/390) of the variants were class 4 and 5, whereas class 3 variants predominated at 78% (306/390). This large proportion of class 3 variants is repeated across all

four cohorts, underscoring the urgent need for reference population genetics data in South Africa. This missing data could lead to inaccurate classifications and interpretations of variants. It is also important to note that some global yields are presented as the total genotype-positive findings, whereas others are the routine genetics diagnostic yields, especially in countries such as the UK, where CMO genetic testing is performed routinely (142).

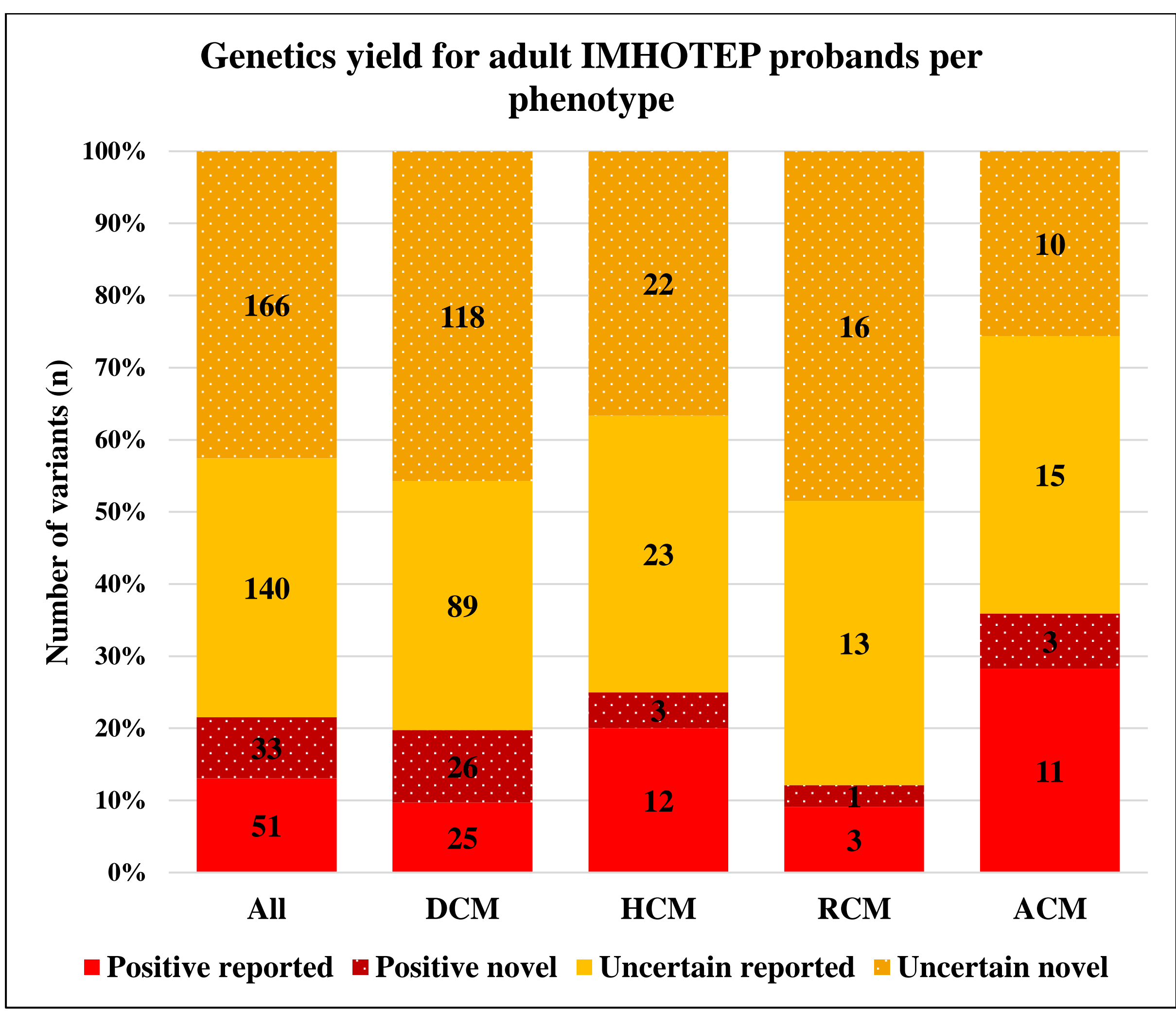
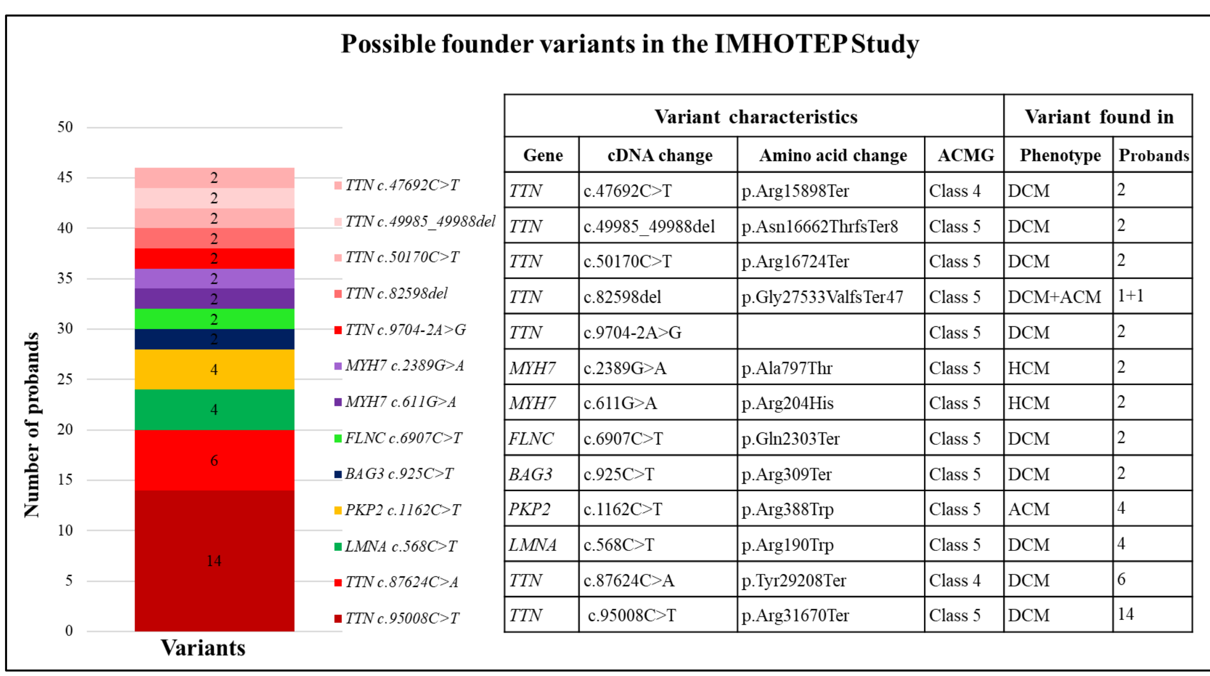


Figure 7.4: Genetics yield for adult IMHOTEP probands per phenotype.

The uniqueness of South Africa's populations is further demonstrated by the identification of 13 possible founder variants shown in Figure 7.5. These variants will have to be further explored to determine if they are indeed founder variants.



Variant characteristics				Variant found in	
Gene	cDNA change	Amino acid change	ACMG	Phenotype	Probands
TTN	c.47692C>T	p.Arg15898Ter	Class 4	DCM	2
TTN	c.49985_49988del	p.Asn16662ThrfsTer8	Class 5	DCM	2
TTN	c.50170C>T	p.Arg16724Ter	Class 5	DCM	2
TTN	c.82598del	p.Gly27533ValfsTer47	Class 5	DCM+ACM	1+1
TTN	c.9704-2A>G		Class 5	DCM	2
MYH7	c.2389G>A	p.Ala797Thr	Class 5	HCM	2
MYH7	c.611G>A	p.Arg204His	Class 5	HCM	2
FLNC	c.6907C>T	p.Gln2303Ter	Class 5	DCM	2
BAG3	c.925C>T	p.Arg309Ter	Class 5	DCM	2
PKP2	c.1162C>T	p.Arg388Trp	Class 5	ACM	4
LMNA	c.568C>T	p.Arg190Trp	Class 5	DCM	4
TTN	c.87624C>A	p.Tyr29208Ter	Class 4	DCM	6
TTN	c.95008C>T	p.Arg31670Ter	Class 5	DCM	14

Figure 7.5: The possible founder variants in the IMHOTEP study. *ACM-arrhythmogenic cardiomyopathy, DCM-dilated cardiomyopathy, HCM-hypertrophic cardiomyopathy.*

In summary (Figure 7.6), it is important to note that some of the global yields are the total genotype-positive findings, whereas others are the routine genetics diagnostic yields, especially in countries such as the UK, where CMO genetic testing is performed routinely (142).

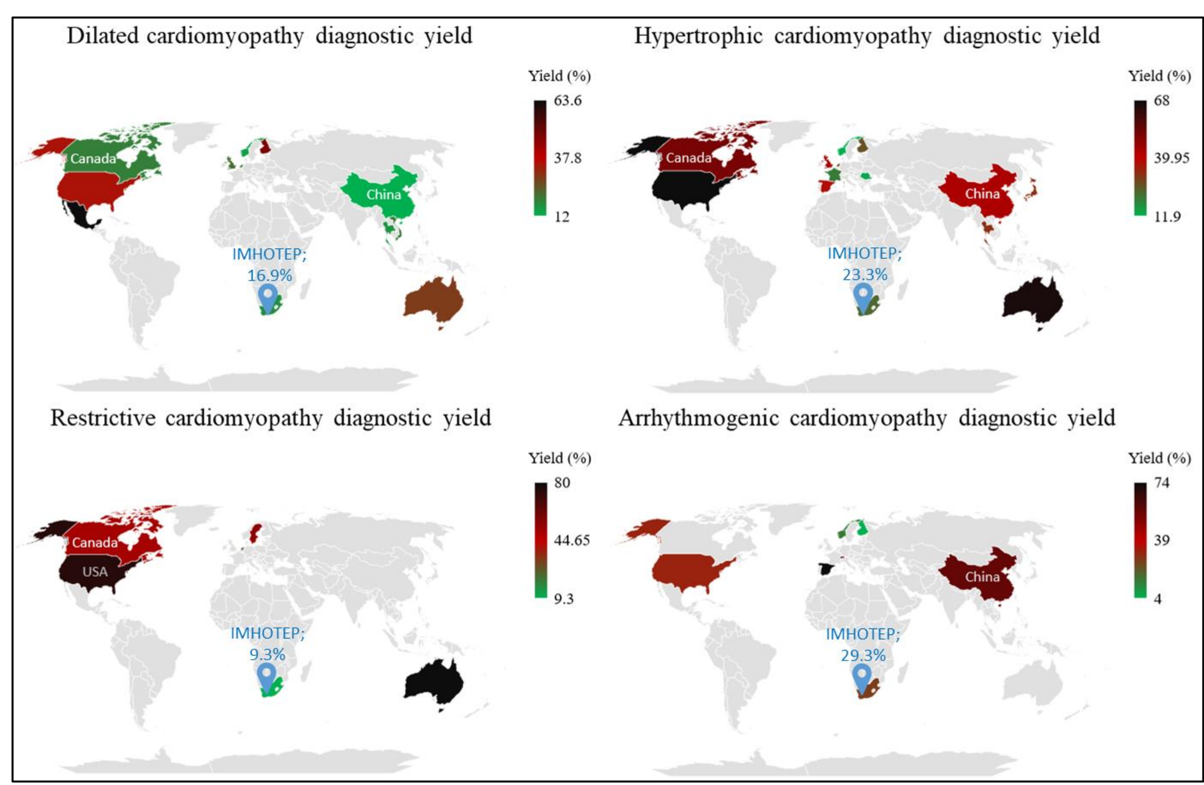


Figure 7.6: Map showing that our yields are much lower than in most studies.

7.2.2 Paediatric Cardiomyopathies

The paediatric DCM and HCM probands' diagnostic yields were 2.2% (2/89) and 0% (0/7), respectively. Like in similar published studies, genetic variants were rarely detected amongst the paediatric probands. For example, 24% of our paediatrics were infants < 1 year; however, only one paediatric proband had a pathogenic. The lower likelihood of finding a genetic cause in these probands could be possibly due to factors such as the vast genetic makeup of Black Africans or the specific genes targeted for NGS (96).

7.3 **CONCLUSION**

IMHOTEP is the largest cardiomyopathy registry in Africa, and through IMHOTEP, we have filled in key gaps in knowledge. However, this study highlights that there is still much work to be done as we only identified the genetic cause of disease in a small proportion of our patients. Overall, this yield is on the lower end of the spectrum compared to the rest of the world; however, the yields of the ACM and HCM cohorts were in the range of previously reported studies. We also note that the DCM probands in South Africa differ from the rest regarding patients' young age and aetiology. The findings from this study also support previous reports that adult RCM is a rarer form of CMO. We also explored the populations within South Africa, and even with the small cohort sizes, some trends emerged that will need further investigation to determine the accuracy of these preliminary results. We also noted the large percentage of class 3 variants (variants of uncertain significance), which could be due to a lack of genetic data on the reference populations in South Africa. We also investigated the paediatric cohort and only reported on a very small number of class 4 and 5 variants, while the rest of the cohort was negative.

The recommendation

Clinical practitioners at the grassroots level should include frequent case and contact tracing, awareness campaigns for lifestyle modification, for example, family planning and healthy lifestyle adoption, and individualised treatment and management plans.

The limitation

The participants in this study were recruited from public tertiary cardiac clinics serving patients from low-resource semi-urban and urban areas of South Africa, and the patients recruited might not accurately reflect the prevalence of the various CMOs in South Africa.

We did not screen the entire human exome with about 27000 genes and could have missed other class 4 and 5 variants. The human reference genome (GRCh37/hg19) that we used does not represent the entire human exome equally.

Reference

1. Roth GA, Forouzanfar MH, Moran AE, Barber R, Nguyen G, Feigin VL, et al. Demographic and epidemiologic drivers of global cardiovascular mortality. N Engl J Med. 2015;372 (14):1333-41.
2. Gouda HN, Charlson F, Sorsdahl K, Ahmadzada S, Ferrari AJ, Erskine H, et al. Burden of non-communicable diseases in sub-Saharan Africa, 1990-2017: results from the Global Burden of Disease Study 2017. Lancet Glob Health. 2019;7 (10):e1375-e87.
3. Amegah AK. Tackling the Growing Burden of Cardiovascular Diseases in Sub-Saharan Africa. Circulation. 2018;138 (22):2449-51.
4. Roth GA, Mensah GA, Johnson CO, Addolorato G, Ammirati E, Baddour LM, et al. Global Burden of Cardiovascular Diseases and Risk Factors, 1990-2019: Update From the GBD 2019 Study. J Am Coll Cardiol. 2020;76 (25):2982-3021.
5. Minja NW, Nakagaayi D, Aliku T, Zhang W, Ssinabulya I, Nabaale J, et al. Cardiovascular diseases in Africa in the twenty-first century: Gaps and priorities going forward. Front Cardiovasc Med. 2022;9:1008335.
6. Coniglio AC, Amos K, Shah SH. Genetic cardiomyopathies and clinical implications in heart transplant recipients: An opportunity for missed diagnoses. J Heart Lung Transplant. 2022;41 (9):1195-7.
7. McKenna WJ, Maron BJ, Thiene G. Classification, Epidemiology, and Global Burden of Cardiomyopathies. Circ Res. 2017;121 (7):722-30.
8. Sara JP, Salik I. Pediatric acute ischemic stroke and restrictive cardiomyopathy, a challenge for anesthesiologists. J Clin Anesth. 2018;54:87-8.
9. Jaouadi H, Kraoua L, Chaker L, Atkinson A, Delague V, Levy N, et al. Novel ALPK3 mutation in a Tunisian patient with pediatric cardiomyopathy and facio-thoraco-skeletal features. J Hum Genet. 2018;63 (10):1077-82.
10. Groenewegen A, Rutten FH, Mosterd A, Hoes AW. Epidemiology of heart failure. Eur J Heart Fail. 2020.
11. Keates AK, Mocumbi AO, Ntsekhe M, Sliwa K, Stewart S. Cardiovascular disease in Africa: epidemiological profile and challenges. Nat Rev Cardiol. 2017;14 (5):273-93.
12. Mandi DG, Bamouni J, Yameogo RA, Naibe DT, Kabore E, Kambire Y, et al. Spectrum of heart failure in sub-Saharan Africa: data from a tertiary hospital-based registry in the eastern center of Burkina Faso. Pan Afr Med J. 2020;36:30.
13. Agbor VN, Ntusi NAB, Noubiap JJ. An overview of heart failure in low- and middle-income countries. Cardiovasc Diagn Ther. 2020;10 (2):244-51.
14. Shaboodien G, Spracklen TF, Kamuli S, Ndibangwi P, Van Niekerk C, Ntusi NAB. Genetics of inherited cardiomyopathies in Africa. Cardiovasc Diagn Ther. 2020;10 (2):262-78.
15. Ntusi NBA, Wonkam A, Shaboodien G, Badri M, Mayosi BM. Frequency and clinical genetics of familial dilated cardiomyopathy in Cape Town: Implications for the evaluation of patients with unexplained cardiomyopathy. Samj S Afr Med J. 2011;101 (6):394-8.
16. Hathaway J, Helio K, Saarinen I, Tallila J, Seppala EH, Tuupanen S, et al. Diagnostic yield of genetic testing in a heterogeneous cohort of 1376 HCM patients. BMC Cardiovasc Disord. 2021;21 (1):126.
17. Carnevale A, Rosas-Madrigal S, Rosendo-Gutierrez R, Lopez-Mora E, Romero-Hidalgo S, Avila-Vazzini N, et al. Genomic study of dilated cardiomyopathy in a group of Mexican patients using site-directed next generation sequencing. Mol Genet Genomic Med. 2020:e1504.
18. Murray B, Hoorntje ET, Te Riele A, Tichnell C, van der Heijden JF, Tandri H, et al. Identification of sarcomeric variants in probands with a clinical diagnosis of arrhythmogenic right ventricular cardiomyopathy (ARVC). J Cardiovasc Electrophysiol. 2018;29 (7):1004-9.

19. Towbin JA, Lowe AM, Colan SD, Sleeper LA, Orav EJ, Clunie S, et al. Incidence, causes, and outcomes of dilated cardiomyopathy in children. JAMA. 2006;296 (15):1867-76.
20. Coughlin SS, Tefft MC, Rice JC, Gerone JL, Baughman KL. Epidemiology of Idiopathic Dilated Cardiomyopathy in the Elderly: Pooled Results from Two Case-Control Studies. American Journal of Epidemiology. 1996;143 (9):881-8.
21. Coughlin SS, Tefft MC. The Epidemiology of Idiopathic Dilated Cardiomyopathy in Women. Epidemiology. 1994;5 (4):449-55.
22. Dokainish H, Teo K, Zhu J, Roy A, AlHabib KF, ElSayed A, et al. Global mortality variations in patients with heart failure: results from the International Congestive Heart Failure (INTER-CHF) prospective cohort study. Lancet Glob Health. 2017;5 (7):e665-e72.
23. Cahill TJ, Ashrafian H, Watkins H. Genetic cardiomyopathies causing heart failure. Circulation Research. 2013;113 (6):660-75.
24. Sliwa K, Bauersachs J, Arany Z, Spracklen TF, Hilfiker-Kleiner D. Peripartum cardiomyopathy: from genetics to management. Eur Heart J. 2021;42 (32):3094-102.
25. Mngoma NF, Ayonrinde OA. Mental distress and substance use among rural Black South African youth who are not in employment, education or training (NEET). Int J Soc Psychiatry. 2023;69 (3):532-42.
26. Shaboodien G, Maske C, Wainwright H, Smuts H, Ntsekhe M, Commerford PJ, et al. Prevalence of myocarditis and cardiotropic virus infection in Africans with HIV-associated cardiomyopathy, idiopathic dilated cardiomyopathy and heart transplant recipients : a pilot study : cardiovascular topic. Cardiovascular Journal Of Africa. 2013;24 (6):218-23.
27. Petersen DC, Libiger O, Tindall EA, Hardie RA, Hannick LI, Glashoff RH, et al. Complex patterns of genomic admixture within southern Africa. PLoS Genet. 2013;9 (3):e1003309.
28. Heradien M, Goosen A, Moolman-Smook JC, Brink PA. Race and gender representation of hypertrophic cardiomyopathy or long QT syndrome cases in a South African research setting. Cardiovasc J Afr. 2007;18 (5):312-5.
29. DeWitt ES, Chandler SF, Hylind RJ, Beausejour Ladouceur V, Blume ED, VanderPluym C, et al. Phenotypic Manifestations of Arrhythmogenic Cardiomyopathy in Children and Adolescents. J Am Coll Cardiol. 2019;74 (3):346-58.
30. Eberly LA, Day SM, Ashley EA, Jacoby DL, Jefferies JL, Colan SD, et al. Association of Race With Disease Expression and Clinical Outcomes Among Patients With Hypertrophic Cardiomyopathy. JAMA Cardiol. 2020;5 (1):83-91.
31. Nel S, Khandheria BK, Libhaber E, Peters F, Ferreira Dos Santos C, Matioda H, et al. Prevalence and significance of isolated left ventricular non-compaction phenotype in normal black Africans using echocardiography. Int J Cardiol Heart Vasc. 2020;30:100585.
32. Fernlund E, Osterberg AW, Kuchinskaya E, Gustafsson M, Jansson K, Gunnarsson C. Novel Genetic Variants in BAG3 and TNNT2 in a Swedish Family with a History of Dilated Cardiomyopathy and Sudden Cardiac Death. Pediatr Cardiol. 2017;38 (6):1262-8.
33. Walsh R, Mazzarotto F, Whiffin N, Buchan R, Midwinter W, Wilk A, et al. Quantitative approaches to variant classification increase the yield and precision of genetic testing in Mendelian diseases: the case of hypertrophic cardiomyopathy. Genome Med. 2019;11.
34. Savarese G, Becher PM, Lund LH, Seferovic P, Rosano GMC, Coats A. Global burden of heart failure: A comprehensive and updated review of epidemiology. Cardiovasc Res. 2022.
35. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med. 2015;17 (5):405-24.
36. Glanzmann B, Jooste T, Ghoor S, Gordon R, Mia R, Mao J, et al. Human whole genome sequencing in South Africa. Scientific reports. 2021;11 (1).

37. Fatumo S, Chikowore T, Choudhury A, Ayub M, Martin AR, Kuchenbaecker K. A roadmap to increase diversity in genomic studies. Nat Med. 2022;28 (2):243-50.
38. Goodwin JF. The frontiers of cardiomyopathy. Br Heart J. 1982;48 (1):1-18.
39. Golbus JR, Puckelwartz MJ, Dellefave-Castillo L, Fahrenbach JP, Nelakuditi V, Pesce LL, et al. Targeted Analysis of Whole Genome Sequence Data to Diagnose Genetic Cardiomyopathy. Circ-Cardiovasc Gene. 2014;7 (6):751-U71.
40. Watkins H, Ashrafian H, Redwood C. Inherited cardiomyopathies. New Engl J Med. 2011;364 (17):1643-56.
41. Elliott P, Andersson B, Arbustini E, Bilinska Z, Cecchi F, Charron P, et al. Classification of the cardiomyopathies: a position statement from the european society of cardiology working group on myocardial and pericardial diseases. European Heart Journal. 2007;29 (2):270-6.
42. Ware SM, Wilkinson JD, Tariq M, Schubert JA, Sridhar A, Colan SD, et al. Genetic Causes of Cardiomyopathy in Children: First Results From the Pediatric Cardiomyopathy Genes Study. J Am Heart Assoc. 2021;10 (9):e017731.
43. Vasilescu C, Ojala TH, Brilhante V, Ojanen S, Hinterding HM, Palin E, et al. Genetic Basis of Severe Childhood-Onset Cardiomyopathies. J Am Coll Cardiol. 2018;72 (19):2324-38.
44. Povysil G, Chazara O, Carss KJ, Deevi SVV, Wang Q, Armisen J, et al. Assessing the Role of Rare Genetic Variation in Patients With Heart Failure. JAMA Cardiol. 2021;6 (4):379-86.
45. Maron BJ, Desai MY, Nishimura RA, Spirito P, Rakowski H, Towbin JA, et al. Diagnosis and Evaluation of Hypertrophic Cardiomyopathy. Journal of the American College of Cardiology. 2022;79 (4):372-89.
46. Rapezzi C, Aimo A, Barison A, Emdin M, Porcari A, Linhart A, et al. Restrictive cardiomyopathy: definition and diagnosis. European Heart Journal. 2022;43 (45):4679-93.
47. Elliott PM, Anastasakis A, Borger MA, Borggrefe M, Cecchi F, Charron P, et al. 2014 ESC Guidelines on diagnosis and management of hypertrophic cardiomyopathy The Task Force for the Diagnosis and Management of Hypertrophic Cardiomyopathy of the European Society of Cardiology (ESC). European Heart Journal. 2014;35 (39):2733-+.
48. Lipshultz SE, Law YM, Asante-Korang A, Austin ED, Dipchand AI, Everitt MD, et al. Cardiomyopathy in Children: Classification and Diagnosis: A Scientific Statement From the American Heart Association. Circulation. 2019;140 (1):e9-e68.
49. Pinto YM, Elliott PM, Arbustini E, Adler Y, Anastasakis A, Bohm M, et al. Proposal for a revised definition of dilated cardiomyopathy, hypokinetic non-dilated cardiomyopathy, and its implications for clinical practice: a position statement of the ESC working group on myocardial and pericardial diseases. Eur Heart J. 2016;37 (23):1850-8.
50. Arbustini E, Narula N, Tavazzi L, Serio A, Grasso M, Favalli V, et al. The MOGE (S) classification of cardiomyopathy for clinicians. J Am Coll Cardiol. 2014;64 (3):304-18.
51. Arbelo E, Protonotarios A, Gimeno JR, Arbustini E, Barriales-Villa R, Basso C, et al. 2023 ESC Guidelines for the management of cardiomyopathies. Eur Heart J. 2023;44 (37):3503-626.
52. Salemi VMC, Mohty D, Altavila SLLd, Melo MDTd, Kalil R, Bocchi EA. Insights into the Classification of Cardiomyopathies: Past, Present, and Future Directions. Clinics. 2021;76.
53. Sliwa K, Damasceno A, Mayosi BM. Epidemiology and etiology of cardiomyopathy in Africa. Circulation. 2005;112 (23):3577-83.
54. Ware JS, Amor-Salamanca A, Tayal U, Govind R, Serrano I, Salazar-Mendiguchía J, et al. Genetic Etiology for Alcohol-Induced Cardiac Toxicity. Journal of the American College of Cardiology. 2018;71 (20):2293-302.

55. Yuyun MF, Sliwa K, Kengne AP, Mocumbi AO, Bukhman G. Cardiovascular Diseases in Sub-Saharan Africa Compared to High-Income Countries: An Epidemiological Perspective. Glob Heart. 2020;15 (1):15.
56. Mocumbi AO, Stothard JR, Correia-de-Sá P, Yacoub M. Endomyocardial Fibrosis: an Update After 70 Years. Current Cardiology Reports. 2019;21 (11).
57. Tsabedze N, Ramsay M, Krause A, Wells Q, Mpanya D, Manga P. The genetic basis for adult-onset idiopathic dilated cardiomyopathy in people of African descent. Heart Fail Rev. 2023;28 (4):879-92.
58. Chan K, Harper AR, Ashrafian H, Yavari A. Cardiomyopathies. Medicine. 2018;46 (10):606-17.
59. Fairweather D, Beetler DJ, Musigk N, Heidecker B, Lyle MA, Cooper LT, Jr., et al. Sex and gender differences in myocarditis and dilated cardiomyopathy: An update. Front Cardiovasc Med. 2023;10:1129348.
60. Ntusi NAB, Sliwa K. Impact of Racial and Ethnic Disparities on Patients With Dilated Cardiomyopathy: JACC Focus Seminar 7/9. J Am Coll Cardiol. 2021;78 (25):2580-8.
61. Tayal U, Ware JS, Lakdawala NK, Heymans S, Prasad SK. Understanding the genetics of adult-onset dilated cardiomyopathy: what a clinician needs to know. Eur Heart J. 2021;42 (24):2384-96.
62. Taylor CN, Lau ES. Sex Differences in Cardiomyopathy. Curr Cardiovasc Risk. 2022;16 (11):159-70.
63. Argirò A, Ho C, Day SM, van der Velden J, Cerbai E, Saberi S, et al. Sex-Related Differences in Genetic Cardiomyopathies. Journal of the American Heart Association. 2022;11 (9).
64. Hasselberg NE, Haland TF, Saberniak J, Brekke PH, Berge KE, Leren TP, et al. Lamin A/C cardiomyopathy: young onset, high penetrance, and frequent need for heart transplantation. Eur Heart J. 2018;39 (10):853-60.
65. Kumar S, Androulakis AFA, Sellal J-M, Maury P, Gandjbakhch E, Waintraub X, et al. Multicenter Experience With Catheter Ablation for Ventricular Tachycardia in Lamin A/C Cardiomyopathy. Circulation: Arrhythmia and Electrophysiology. 2016;9 (8).
66. Adalsteinsdottir B, Burke M, Maron BJ, Danielsen R, Lopez B, Diez J, et al. Hypertrophic cardiomyopathy in myosin-binding protein C (MYBPC3) Icelandic founder mutation carriers. Open Heart. 2020;7 (1):e001220.
67. Moolman-Smook JC, De Lange WJ, Bruwer EC, Brink PA, Corfield VA. 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. 1999;65 (5):1308-20.
68. Hastings R, de Villiers CP, Hooper C, Ormondroyd L, Pagnamenta A, Lise S, et al. Combination of Whole Genome Sequencing, Linkage, and Functional Studies Implicates a Missense Mutation in Titin as a Cause of Autosomal Dominant Cardiomyopathy With Features of Left Ventricular Noncompaction. Circ Cardiovasc Genet. 2016;9 (5):426-35.
69. Rosenbaum AN, Agre KE, Pereira NL. Genetics of dilated cardiomyopathy: practical implications for heart failure management. Nat Rev Cardiol. 2020;17 (5):286-97.
70. Choudhury A, Ramsay M, Hazelhurst S, Aron S, Bardien S, Botha G, et al. Whole-genome sequencing for an enhanced understanding of genetic variation among South Africans. Nature Communications. 2017;8 (1).
71. Retshabile G, Mlotshwa BC, Williams L, Mwesigwa S, Mboowa G, Huang Z, et al. Whole-Exome Sequencing Reveals Uncaptured Variation and Distinct Ancestry in the Southern African Population of Botswana. The American Journal of Human Genetics. 2018;102 (5):731-43.
72. Pereira L, Mutesa L, Tindana P, Ramsay M. African genetic diversity and adaptation inform a precision medicine agenda. Nature Reviews Genetics. 2021;22 (5):284-306.

73. Rotimi CN, Bentley AR, Doumatey AP, Chen G, Shriner D, Adeyemo A. The genomic landscape of African populations in health and disease. Human Molecular Genetics. 2017;26 (R2):R225-R36.
74. Baker JL, Shriner D, Bentley AR, Rotimi CN. Pharmacogenomic implications of the evolutionary history of infectious diseases in Africa. The Pharmacogenomics Journal. 2016;17 (2):112-20.
75. Chimusa ER, Meintjies A, Tchanga M, Mulder N, Seoighe C, Soodyall H, et al. A genomic portrait of haplotype diversity and signatures of selection in indigenous southern African populations. PLoS Genet. 2015;11 (3):e1005052.
76. Oliver E, Oliver WH. The Colonisation of South Africa: A unique case. HTS Teologiese Studies / Theological Studies. 2017;73 (3).
77. Jarcho JA, McKenna W, Pare JA, Solomon SD, Holcombe RF, Dickie S, et al. Mapping a gene for familial hypertrophic cardiomyopathy to chromosome 14q1. N Engl J Med. 1989;321 (20):1372-8.
78. Ingles J, Macciocca I, Morales A, Thomson K. Genetic Testing in Inherited Heart Diseases. Heart Lung Circ. 2020;29 (4):505-11.
79. Cirino AL, Harris S, Lakdawala NK, Michels M, Olivotto I, Day SM, et al. Role of Genetic Testing in Inherited Cardiovascular Disease: A Review. JAMA Cardiol. 2017;2 (10):1153-60.
80. Musunuru K, Hershberger RE, Day SM, Klinedinst NJ, Landstrom AP, Parikh VN, et al. Genetic Testing for Inherited Cardiovascular Diseases: A Scientific Statement From the American Heart Association. Circ Genom Precis Med. 2020;13 (4):e000067.
81. Coll M, Perez-Serra A, Mates J, Del Olmo B, Puigmule M, Fernandez-Falgueras A, et al. Incomplete Penetrance and Variable Expressivity: Hallmarks in Channelopathies Associated with Sudden Cardiac Death. Biology (Basel). 2017;7 (1).
82. Lacaze P, Sebra R, Riaz M, Ingles J, Tiller J, Thompson BA, et al. Genetic variants associated with inherited cardiovascular disorders among 13,131 asymptomatic older adults of European descent. Npj Genom Med. 2021;6 (1):51.
83. Ramchand J, Wallis M, Macciocca I, Lynch E, Farouque O, Martyn M, et al. Prospective Evaluation of the Utility of Whole Exome Sequencing in Dilated Cardiomyopathy. J Am Heart Assoc. 2020;9 (2):e013346.
84. Manrai AK, Funke BH, Rehm HL, Olesen MS, Maron BA, Szolovits P, et al. Genetic Misdiagnoses and the Potential for Health Disparities. N Engl J Med. 2016;375 (7):655-65.
85. Forleo C, D'Erchia AM, Sorrentino S, Manzari C, Chiara M, Iacoviello M, et al. Targeted next-generation sequencing detects novel gene-phenotype associations and expands the mutational spectrum in cardiomyopathies. PLoS One. 2017;12 (7):e0181842.
86. Abdallah AM, Carlus SJ, Al-Mazroea AH, Alluqmani M, Almohammadi Y, Bhuiyan ZA, et al. Digenic Inheritance of LAMA4 and MYH7 Mutations in Patient with Infantile Dilated Cardiomyopathy. Medicina (Kaunas). 2019;55 (1).
87. Hershberger RE, Givertz MM, Ho CY, Judge DP, Kantor PF, McBride KL, et al. Genetic evaluation of cardiomyopathy: a clinical practice resource of the American College of Medical Genetics and Genomics (ACMG). Genet Med. 2018;20 (9):899-909.
88. Gacita AM, Fullenkamp DE, Ohiri J, Pottinger T, Puckelwartz MJ, Nobrega MA, et al. Genetic Variation in Enhancers Modifies Cardiomyopathy Gene Expression and Progression. Circulation. 2021;143 (13):1302-16.
89. Schuldt M, Johnston JR, He H, Huurman R, Pei J, Harakalova M, et al. Mutation location of HCM-causing troponin T mutations defines the degree of myofilament dysfunction in human cardiomyocytes. J Mol Cell Cardiol. 2021;150:77-90.
90. Lehman SJ, Crocini C, Leinwand LA. Targeting the sarcomere in inherited cardiomyopathies. Nature Reviews Cardiology. 2022.

91. Weldy CS, Ashley EA. Towards precision medicine in heart failure. Nat Rev Cardiol. 2021;18 (11):745-62.
92. Weintraub RG, Semsarian C, Macdonald P. Dilated cardiomyopathy. Lancet. 2017;390 (10092):400-14.
93. Eldemire R, Taylor MRG, Mestroni L. Understanding the role of titin in dilated cardiomyopathy. International Journal of Cardiology. 2020;316:186-7.
94. Taylor M, Graw S, Barnes C, Slavov D, Brun F, Pinamonti B, et al. Titin Mutations Cause Arrhythmogenic Right Ventricular Cardiomyopathy. Circulation. 2010;122 (21).
95. Stava TT, Leren TP, Bogsrud MP. Molecular genetics in 4408 cardiomyopathy probands and 3008 relatives in Norway: 17 years of genetic testing in a national laboratory. Eur J Prev Cardiol. 2022;29 (13):1789-99.
96. Ware SM, Bhatnagar S, Dexheimer PJ, Wilkinson JD, Sridhar A, Fan X, et al. The genetic architecture of pediatric cardiomyopathy. Am J Hum Genet. 2022;109 (2):282-98.
97. Ader F, De Groote P, Reant P, Rooryck-Thambo C, Dupin-Deguine D, Rambaud C, et al. FLNC pathogenic variants in patients with cardiomyopathies: Prevalence and genotype-phenotype correlations. Clin Genet. 2019;96 (4):317-29.
98. Bourfiss M, van Vugt M, Alasiri AI, Ruijsink B, van Setten J, Schmidt AF, et al. Prevalence and Disease Expression of Pathogenic and Likely Pathogenic Variants Associated With Inherited Cardiomyopathies in the General Population. Circ Genom Precis Med. 2022;15 (6):e003704.
99. Mullis KB. The unusual origin of the polymerase chain reaction. Sci Am. 1990;262 (4):56-61, 4-5.
100. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci U S A. 1977;74 (12):5463-7.
101. Hu T, Chitnis N, Monos D, Dinh A. Next-generation sequencing technologies: An overview. Hum Immunol. 2021;82 (11):801-11.
102. Bentley DR, Balasubramanian S, Swerdlow HP, Smith GP, Milton J, Brown CG, et al. Accurate whole human genome sequencing using reversible terminator chemistry. Nature. 2008;456 (7218):53-9.
103. Collins FS, Fink L. The Human Genome Project. Alcohol Health Res World. 1995;19 (3):190-5.
104. Sanghvi RV, Buhay CJ, Powell BC, Tsai EA, Dorschner MO, Hong CS, et al. Characterizing reduced coverage regions through comparison of exome and genome sequencing data across 10 centers. Genetics in Medicine. 2018;20 (8):855-66.
105. Spracklen TF, Keavney B, Laing N, Ntusi N, Shaboodien G. Modern genomic techniques in the identification of genetic causes of cardiomyopathy. Heart. 2022;108 (23):1843-50.
106. Gibbon VE. African ancient DNA research requires robust ethics and permission protocols. Nature Reviews Genetics. 2020;21 (11):645-7.
107. Verdonschot JAJ, Vanhoutte EK, Claes GRF, Helderman-van den Enden A, Hoeijmakers JGJ, Hellebrekers D, et al. A mutation update for the FLNC gene in myopathies and cardiomyopathies. Hum Mutat. 2020;41 (6):1091-111.
108. Bang M-L, Centner T, Fornoff F, Geach AJ, Gotthardt M, McNabb M, et al. The Complete Gene Sequence of Titin, Expression of an Unusual ≈700-kDa Titin Isoform, and Its Interaction With Obscurin Identify a Novel Z-Line to I-Band Linking System. Circulation Research. 2001;89 (11):1065-72.
109. Roncarati R, Viviani Anselmi C, Krawitz P, Lattanzi G, von Kodolitsch Y, Perrot A, et al. Doubly heterozygous LMNA and TTN mutations revealed by exome sequencing in a severe form of dilated cardiomyopathy. Eur J Hum Genet. 2013;21 (10):1105-11.
110. Santiago CF, Huttner IG, Fatkin D. Titin-related Cardiomyopathy: Is it a Distinct Disease? Curr Cardiol Rep. 2022;24 (9):1069-75.

111. Mak TSH, Lee YK, Tang CS, Hai JSH, Ran XR, Sham PC, et al. Coverage and diagnostic yield of Whole Exome Sequencing for the Evaluation of Cases with Dilated and Hypertrophic Cardiomyopathy. Scientific reports. 2018;8.
112. Robyns T, Willems R, Van Cleemput J, Jhangiani S, Muzny D, Gibbs R, et al. Whole exome sequencing in a large pedigree with DCM identifies a novel mutation in RBM20. Acta Cardiol. 2020;75 (8):748-53.
113. Thomson KL, Ormondroyd E, Harper AR, Dent T, McGuire K, Baksi J, et al. Analysis of 51 proposed hypertrophic cardiomyopathy genes from genome sequencing data in sarcomere negative cases has negligible diagnostic yield. Genetics in Medicine. 2019;21 (7):1576-84.
114. Stroeks SLVM, Hellebrekers DMEI, Claes GRF, Tayal U, Krapels IPC, Vanhoutte EK, et al. Clinical impact of re-evaluating genes and variants implicated in dilated cardiomyopathy. Genetics in Medicine. 2021.
115. Buchanan J, Blair E, Thomson KL, Ormondroyd E, Watkins H, Taylor JC, et al. Do health professionals value genomic testing? A discrete choice experiment in inherited cardiovascular disease. Eur J Hum Genet. 2019;27 (11):1639-48.
116. Cirino AL, Lakdawala NK, McDonough B, Conner L, Adler D, Weinfeld M, et al. A Comparison of Whole Genome Sequencing to Multigene Panel Testing in Hypertrophic Cardiomyopathy Patients. Circ-Cardiovasc Gene. 2017;10 (5).
117. Gulsuner S, Stein DJ, Susser ES, Sibeko G, Pretorius A, Walsh T, et al. Genetics of schizophrenia in the South African Xhosa. Science. 2020;367 (6477):569-73.
118. Spracklen TF, Kasher PR, Kraus S, Botha TL, Page DJ, Kamuli S, et al. Identification of a POLG Variant in a Family With Arrhythmogenic Cardiomyopathy and Left Ventricular Fibrosis. Circ Genom Precis Med. 2021;14 (1):e003138.
119. Sammani A, Baas AF, Asselbergs FW, Te Riele A. Diagnosis and Risk Prediction of Dilated Cardiomyopathy in the Era of Big Data and Genomics. J Clin Med. 2021;10 (5).
120. Jordan E, Peterson L, Ai T, Asatryan B, Bronicki L, Brown E, et al. Evidence-Based Assessment of Genes in Dilated Cardiomyopathy. Circulation. 2021;144 (1):7-19.
121. Hershberger RE, Cowan J, Jordan E, Kinnamon DD. The Complex and Diverse Genetic Architecture of Dilated Cardiomyopathy. Circ Res. 2021;128 (10):1514-32.
122. Burke MA, Cook SA, Seidman JG, Seidman CE. Clinical and Mechanistic Insights Into the Genetics of Cardiomyopathy. Journal of the American College of Cardiology. 2016;68 (25):2871-86.
123. Clarke SL, Assimes TL, Tcheandjieu C. The Propagation of Racial Disparities in Cardiovascular Genomics Research. Circulation: Genomic and Precision Medicine. 2021;14 (5).
124. Landry LG, Rehm HL. Association of Racial/Ethnic Categories With the Ability of Genetic Tests to Detect a Cause of Cardiomyopathy. JAMA Cardiol. 2018;3 (4):341-5.
125. Popejoy AB, Fullerton SM. Genomics is failing on diversity. Nature. 2016;538 (7624):161-4.
126. Jordan E, Kinnamon DD, Haas GJ, Hofmeyer M, Kransdorf E, Ewald GA, et al. Genetic Architecture of Dilated Cardiomyopathy in Individuals of African and European Ancestry. JAMA. 2023;330 (5):432-41.
127. Peterson RE, Kuchenbaecker K, Walters RK, Chen CY, Popejoy AB, Periyasamy S, et al. Genome-wide Association Studies in Ancestrally Diverse Populations: Opportunities, Methods, Pitfalls, and Recommendations. Cell. 2019;179 (3):589-603.
128. Singh RR. Target Enrichment Approaches for Next-Generation Sequencing Applications in Oncology. Diagnostics. 2022;12 (7).
129. Vos TA, C.; Arora, M.; et al.; G. B. D. Disease, Injury I, Prevalence C. Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. Lancet. 2016;388 (10053):1545-602.

130. Quiat D, Witkowski L, Zouk H, Daly KP, Roberts AE. Retrospective Analysis of Clinical Genetic Testing in Pediatric Primary Dilated Cardiomyopathy: Testing Outcomes and the Effects of Variant Reclassification. J Am Heart Assoc. 2020;9 (11):e016195.
131. Arola A, Jokinen E, Ruuskanen O, Saraste M, Pesonen E, Kuusela AL, et al. Epidemiology of idiopathic cardiomyopathies in children and adolescents. A nationwide study in Finland. Am J Epidemiol. 1997;146 (5):385-93.
132. Bagnall RD, Singer ES, Wacker J, Nowak N, Ingles J, King I, et al. Genetic Basis of Childhood Cardiomyopathy. Circ Genom Precis Med. 2022;15 (6):e003686.
133. El Hadi H, Freund A, Desch S, Thiele H, Majunke N. Hypertrophic, Dilated, and Arrhythmogenic Cardiomyopathy: Where Are We? Biomedicines. 2023;11 (2).
134. and NBN, Mayosi BM. <ntusi2009.pdf>. 2009.
135. Choudhry S, Puri K, Denfield SW. An Update on Pediatric Cardiomyopathy. Curr Treat Options Cardiovasc Med. 2019;21 (8):36.
136. Mehaney DA, Haghighi A, Embaby AK, Zeyada RA, Darwish RK, Elfeel NS, et al. Molecular analysis of dilated and left ventricular noncompaction cardiomyopathies in Egyptian children. Cardiol Young. 2022;32 (2):295-300.
137. Schultheiss H-P, Fairweather D, Caforio ALP, Escher F, Hershberger RE, Lipshultz SE, et al. Dilated cardiomyopathy. Nature Reviews Disease Primers. 2019;5 (1):32.
138. Morales A, Hershberger R. Clinical Application of Genetic Testing in Heart Failure. Curr Heart Fail Rep. 2017;14 (6):543-53.
139. Stroeks S, Lunde IG, Hellebrekers D, Claes GRF, Wakimoto H, Gorham J, et al. Prevalence and Clinical Consequences of Multiple Pathogenic Variants in Dilated Cardiomyopathy. Circ Genom Precis Med. 2023:e003788.
140. Miura K, Nakagawa H, Toyoshima H, Kodama K, Nagai M, Morikawa Y, et al. Environmental Factors and Risk of Idiopathic Dilated Cardiomyopathy-A Multi-Hospital Case-Control Study in Japan. Circulation Journal. 2004;68 (11):1011-7.
141. Hofmeyer M, Haas GJ, Jordan E, Cao J, Kransdorf E, Ewald GA, et al. Rare Variant Genetics and Dilated Cardiomyopathy Severity: The DCM Precision Medicine Study. Circulation. 2023;148 (11):872-81.
142. Mazzarotto F, Tayal U, Buchan RJ, Midwinter W, Wilk A, Whiffin N, et al. Reevaluating the Genetic Contribution of Monogenic Dilated Cardiomyopathy. Circulation. 2020;141 (5):387-98.
143. Shah RA, Asatryan B, Sharaf Dabbagh G, Aung N, Khanji MY, Lopes LR, et al. Frequency, Penetrance, and Variable Expressivity of Dilated Cardiomyopathy-Associated Putative Pathogenic Gene Variants in UK Biobank Participants. Circulation. 2022;146 (2):110-24.
144. Trachtenberg BH, Jimenez J, Morris AA, Kransdorf E, Owens A, Fishbein DP, et al. TTR variants in patients with dilated cardiomyopathy: An investigation of the DCM Precision Medicine Study. Genet Med. 2022;24 (7):1495-502.
145. Nguyen TV, Tran Vu MT, Do TNP, Tran THN, Do TH, Nguyen TMH, et al. Genetic Determinants and Genotype-Phenotype Correlations in Vietnamese Patients With Dilated Cardiomyopathy. Circ J. 2021;85 (9):1469-78.
146. Asselbergs FW, Sammani A, Elliott P, Gimeno JR, Tavazzi L, Tendera M, et al. Differences between familial and sporadic dilated cardiomyopathy: ESC EORP Cardiomyopathy & Myocarditis registry. ESC Heart Fail. 2021;8 (1):95-105.
147. Gao Y, Peng L, Zhao C. MYH7 in cardiomyopathy and skeletal muscle myopathy. Mol Cell Biochem. 2023.
148. Chmielewski P, Truszkowska G, Kowalik I, Rydzanicz M, Michalak E, Sobieszczanska-Malek M, et al. Titin-Related Dilated Cardiomyopathy: The Clinical Trajectory and the Role of Circulating Biomarkers in the Clinical Assessment. Diagnostics (Basel). 2021;12 (1).

149. de Frutos F, Ochoa JP, Navarro-Penalver M, Baas A, Bjerre JV, Zorio E, et al. Natural History of MYH7-Related Dilated Cardiomyopathy. J Am Coll Cardiol. 2022;80 (15):1447-61.
150. Rampersaud E, Siegfried JD, Norton N, Li D, Martin E, Hershberger RE. Rare variant mutations identified in pediatric patients with dilated cardiomyopathy. Prog Pediatr Cardiol. 2011;31 (1):39-47.
151. Akinrinade O, Lesurf R, Genomics England Research C, Lougheed J, Mondal T, Smythe J, et al. Age and Sex Differences in the Genetics of Cardiomyopathy. J Cardiovasc Transl Res. 2023.
152. Wanert C, El Louali F, Al Dybiat S, Nguyen K, Zaffran S, Ovaert C. Genetic profile and genotype-phenotype correlations in childhood cardiomyopathy. Arch Cardiovasc Dis. 2023.
153. Ilina MV, Kepron CA, Taylor GP, Perrin DG, Kantor PF, Somers GR. Undiagnosed heart disease leading to sudden unexpected death in childhood: a retrospective study. Pediatrics. 2011;128 (3):e513-20.
154. van Waning JI, Caliskan K, Hoedemaekers YM, van Spaendonck-Zwarts KY, Baas AF, Boekholdt SM, et al. Genetics, Clinical Features, and Long-Term Outcome of Noncompaction Cardiomyopathy. J Am Coll Cardiol. 2018;71 (7):711-22.
155. Fish M, Shaboodien G, Kraus S, Sliwa K, Seidman CE, Burke MA, et al. Mutation analysis of the phospholamban gene in 315 South Africans with dilated, hypertrophic, peripartum and arrhythmogenic right ventricular cardiomyopathies. Scientific reports. 2016;6:22235.
156. Ntusi NA, Shaboodien G, Badri M, Gumedze F, Mayosi BM. Clinical features, spectrum of causal genetic mutations and outcome of hypertrophic cardiomyopathy in South Africans. Cardiovasc J Afr. 2016;27 (3):152-8.
157. Reza N, Packard E, Goli R, Chowns JL, Owens AT, Arany Z, et al. Clinical Predictors of Referral for and Yield of Genetic Testing in Peripartum Cardiomyopathy. JACC Heart Fail. 2023.
158. Goli R, Li J, Brandimarto J, Levine LD, Riis V, McAfee Q, et al. Genetic and Phenotypic Landscape of Peripartum Cardiomyopathy. Circulation. 2021;143 (19):1852-62.
159. Ware JS, Li J, Mazaika E, Yasso CM, DeSouza T, Cappola TP, et al. Shared Genetic Predisposition in Peripartum and Dilated Cardiomyopathies. N Engl J Med. 2016;374 (3):233-41.
160. Viljoen C, Hoevelmann J, Sliwa K. Peripartum cardiomyopathy: risk factors and predictors of outcome. Curr Opin Cardiol. 2023;38 (3):223-32.
161. Azibani F, Pfeffer TJ, Ricke-Hoch M, Dowling W, Pietzsch S, Briton O, et al. Outcome in German and South African peripartum cardiomyopathy cohorts associates with medical therapy and fibrosis markers. ESC Heart Fail. 2020;7 (2):512-22.
162. van Spaendonck-Zwarts KY, Posafalvi A, van den Berg MP, Hilfiker-Kleiner D, Bollen IA, Sliwa K, et al. Titin gene mutations are common in families with both peripartum cardiomyopathy and dilated cardiomyopathy. Eur Heart J. 2014;35 (32):2165-73.
163. Hoes MF, Arany Z, Bauersachs J, Hilfiker-Kleiner D, Petrie MC, Sliwa K, et al. Pathophysiology and risk factors of peripartum cardiomyopathy. Nat Rev Cardiol. 2022.
164. Li S, Zhang C, Liu N, Bai H, Hou C, Song L, et al. Titin-truncating variants are associated with heart failure events in patients with left ventricular non-compaction cardiomyopathy. Clin Cardiol. 2019;42 (5):530-5.
165. Richard P, Ader F, Roux M, Donal E, Eicher JC, Aoutil N, et al. Targeted panel sequencing in adult patients with left ventricular non-compaction reveals a large genetic heterogeneity. Clin Genet. 2018.

166. Hirono K, Hata Y, Miyao N, Okabe M, Takarada S, Nakaoka H, et al. Increased Burden of Ion Channel Gene Variants Is Related to Distinct Phenotypes in Pediatric Patients With Left Ventricular Noncompaction. Circ Genom Precis Med. 2020;13 (4):e002940.
167. Huang G, Bisaria A, Wakefield DL, Yamawaki TM, Luo X, Zhang JA, et al. Titin-truncating variants in hiPSC cardiomyocytes induce pathogenic proteinopathy and sarcomere defects with preserved core contractile machinery. Stem Cell Reports. 2023;18 (1):220-36.
168. Fomin A, Gartner A, Cyganek L, Tiburcy M, Tuleta I, Wellers L, et al. Truncated titin proteins and titin haploinsufficiency are targets for functional recovery in human cardiomyopathy due to TTN mutations. Sci Transl Med. 2021;13 (618):eabd3079.
169. Roberts AM, Ware JS, Herman DS, Schafer S, Baksi J, Bick AG, et al. Integrated allelic, transcriptional, and phenomic dissection of the cardiac effects of titin truncations in health and disease. Sci Transl Med. 2015;7 (270):270ra6.
170. Mohammed F, Chidgey M. Desmosomal protein structure and function and the impact of disease-causing mutations. J Struct Biol. 2021;213 (3):107749.
171. Kolokotronis K, Kuhnisch J, Klopocki E, Dartsch J, Rost S, Huculak C, et al. Biallelic mutation in MYH7 and MYBPC3 leads to severe cardiomyopathy with left ventricular noncompaction phenotype. Hum Mutat. 2019;40 (8):1101-14.
172. Bays JL, DeMali KA. Vinculin in cell-cell and cell-matrix adhesions. Cell Mol Life Sci. 2017;74 (16):2999-3009.
173. Kayvanpour E, Sedaghat-Hamedani F, Amr A, Lai A, Haas J, Holzer DB, et al. Genotype-phenotype associations in dilated cardiomyopathy: meta-analysis on more than 8000 individuals. Clin Res Cardiol. 2017;106 (2):127-39.
174. McDonagh TA, Metra M, Adamo M, Gardner RS, Baumbach A, Böhm M, et al. 2021 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure. European Heart Journal. 2021;42 (36):3599-726.
175. Kadish A, Dyer A, Daubert JP, Quigg R, Estes NAM, Anderson KP, et al. Prophylactic Defibrillator Implantation in Patients with Nonischemic Dilated Cardiomyopathy. New Engl J Med. 2004;350 (21):2151-8.
176. Pescariu SA, Șoșdean R, Mircea MN, Ionac A, Pescariu S. The Challenges of Cardiac Resynchronization Therapy in Patients with Dilated Cardiomyopathy and Persistent Left Superior Vena Cava: A Case Report and Concise Literature Review. Biomedicines. 2023;11 (4).
177. Romano R, Ghahremani S, Zimmerman T, Legere N, Thakar K, Ladha FA, et al. Reading Frame Repair of TTN Truncation Variants Restores Titin Quantity and Functions. Circulation. 2022;145 (3):194-205.
178. Gabisonia K, Recchia FA. Gene Therapy for Heart Failure: New Perspectives. Curr Heart Fail Rep. 2018;15 (6):340-9.
179. Li T, Yang Y, Qi H, Cui W, Zhang L, Fu X, et al. CRISPR/Cas9 therapeutics: progress and prospects. Signal Transduction and Targeted Therapy. 2023;8 (1).
180. Ishida H, Narita J, Ishii R, Suginobe H, Tsuru H, Wang R, et al. Clinical Outcomes and Genetic Analyses of Restrictive Cardiomyopathy in Children. Circ Genom Precis Med. 2023;16 (4):382-9.
181. Morrow WR. Cardiomyopathy and heart transplantation in children. Curr Opin Cardiol. 2000;15 (4):216-23.
182. Tadros R, Francis C, Xu X, Vermeer AMC, Harper AR, Huurman R, et al. Shared genetic pathways contribute to risk of hypertrophic and dilated cardiomyopathies with opposite directions of effect. Nat Genet. 2021;53 (2):128-34.
183. Rosengren A, Smyth A, Rangarajan S, Ramasundarahettige C, Bangdiwala SI, AlHabib KF, et al. Socioeconomic status and risk of cardiovascular disease in 20 low-income, middle-income, and high-income countries: the Prospective Urban Rural Epidemiologic (PURE) study. Lancet Glob Health. 2019;7 (6):e748-e60.

184. Sliwa K, Wilkinson D, Hansen C, Ntyintyane L, Tibazarwa K, Becker A, et al. Spectrum of heart disease and risk factors in a black urban population in South Africa (the Heart of Soweto Study): a cohort study. Lancet. 2008;371 (9616):915-22.
185. Flaquer A, Strauch K. A comparison of different linkage statistics in small to moderate sized pedigrees with complex diseases. BMC Res Notes. 2012;5:411.
186. Charron P, Elliott PM, Gimeno JR, Caforio ALP, Kaski JP, Tavazzi L, et al. The Cardiomyopathy Registry of the EURObservational Research Programme of the European Society of Cardiology: baseline data and contemporary management of adult patients with cardiomyopathies. Eur Heart J. 2018;39 (20):1784-93.
187. Puckelwartz MJ, Pesce LL, Dellefave-Castillo LM, Wheeler MT, Pottinger TD, Robinson AC, et al. Genomic Context Differs Between Human Dilated Cardiomyopathy and Hypertrophic Cardiomyopathy. J Am Heart Assoc. 2021;10 (7):e019944.
188. Xiao L, Wu D, Sun Y, Hu D, Dai J, Chen Y, et al. Whole-exome sequencing reveals genetic risks of early-onset sporadic dilated cardiomyopathy in the Chinese Han population. Sci China Life Sci. 2022;65 (4):770-80.
189. Stroeks S, Hellebrekers D, Claes GRF, Tayal U, Krapels IPC, Vanhoutte EK, et al. Clinical impact of re-evaluating genes and variants implicated in dilated cardiomyopathy. Genet Med. 2021;23 (11):2186-93.
190. Ni H, Jordan E, Kinnamon DD, Cao J, Haas GJ, Hofmeyer M, et al. Screening for Dilated Cardiomyopathy in At-Risk First-Degree Relatives. J Am Coll Cardiol. 2023;81 (21):2059-71.
191. Alimohamed MZ, Johansson LF, Posafalvi A, Boven LG, van Dijk KK, Walters L, et al. Diagnostic yield of targeted next generation sequencing in 2002 Dutch cardiomyopathy patients. Int J Cardiol. 2021;332:99-104.
192. Dellefave-Castillo LM, Cirino AL, Callis TE, Esplin ED, Garcia J, Hatchell KE, et al. Assessment of the Diagnostic Yield of Combined Cardiomyopathy and Arrhythmia Genetic Testing. JAMA Cardiol. 2022;7 (9):966-74.
193. Poller W, Haas J, Klingel K, Kuhnisch J, Gast M, Kaya Z, et al. Familial Recurrent Myocarditis Triggered by Exercise in Patients With a Truncating Variant of the Desmoplakin Gene. Journal of the American Heart Association. 2020;9 (10).
194. Minoche AE, Horvat C, Johnson R, Gayevskiy V, Morton SU, Drew AP, et al. Genome sequencing as a first-line genetic test in familial dilated cardiomyopathy. Genetics in Medicine. 2019;21 (3):650-62.
195. Herman DS, Lam L, Taylor MR, Wang L, Teekakirikul P, Christodoulou D, et al. Truncations of titin causing dilated cardiomyopathy. N Engl J Med. 2012;366 (7):619-28.
196. Jansen M, Baas AF, van Spaendonck-Zwarts KY, Ummels AS, van den Wijngaard A, Jongbloed JDH, et al. Mortality Risk Associated With Truncating Founder Mutations in Titin. Circ Genom Precis Med. 2019;12 (5):e002436.
197. Di Feo MF, Lillback V, Jokela M, McEntagart M, Homfray T, Giorgio E, et al. The crucial role of titin in fetal development: recurrent miscarriages and bone, heart and muscle anomalies characterise the severe end of titinopathies spectrum. J Med Genet. 2023.
198. McAfee Q, Chen CY, Yang Y, Caporizzo MA, Morley M, Babu A, et al. Truncated titin proteins in dilated cardiomyopathy. Sci Transl Med. 2021;13 (618):eabd7287.
199. Dominguez F, Lalaguna L, Martinez-Martin I, Piqueras-Flores J, Rasmussen TB, Zorio E, et al. Titin Missense Variants as a Cause of Familial Dilated Cardiomyopathy. Circulation. 2023;147 (22):1711-3.
200. Zhou J, Ng B, Ko NSJ, Fiedler LR, Khin E, Lim A, et al. Titin truncations lead to impaired cardiomyocyte autophagy and mitochondrial function in vivo. Hum Mol Genet. 2019;28 (12):1971-81.

201. Dorsch LM, Schuldt M, Knezevic D, Wiersma M, Kuster DWD, van der Velden J, et al. Untying the knot: protein quality control in inherited cardiomyopathies. Pflugers Arch. 2019;471 (5):795-806.
202. Kim YG, Ha C, Shin S, Park JH, Jang JH, Kim JW. Enrichment of titin-truncating variants in exon 327 in dilated cardiomyopathy and its relevance to reduced nonsense-mediated mRNA decay efficiency. Front Genet. 2023;13.
203. Loescher CM, Hobbach AJ, Linke WA. Titin (TTN): from molecule to modifications, mechanics, and medical significance. Cardiovasc Res. 2022;118 (14):2903-18.
204. Franaszczyk M, Chmielewski P, Truszkowska G, Stawinski P, Michalak E, Rydzanicz M, et al. Titin Truncating Variants in Dilated Cardiomyopathy - Prevalence and Genotype-Phenotype Correlations. PLoS One. 2017;12 (1):e0169007.
205. Tabish AM, Azzimato V, Alexiadis A, Buyandelger B, Knoll R. Genetic epidemiology of titin-truncating variants in the etiology of dilated cardiomyopathy. Biophys Rev. 2017;9 (3):207-23.
206. Zaunbrecher RJ, Abel AN, Beussman K, Leonard A, von Frieling-Salewsky M, Fields PA, et al. Cronos Titin Is Expressed in Human Cardiomyocytes and Necessary for Normal Sarcomere Function. Circulation. 2019;140 (20):1647-60.
207. Garcia-Castro M, Coto E, Reguero JR, Berrazueta JR, Alvarez V, Alonso B, et al. [Mutations in sarcomeric genes MYH7, MYBPC3, TNNT2, TNNI3, and TPM1 in patients with hypertrophic cardiomyopathy]. Rev Esp Cardiol. 2009;62 (1):48-56.
208. Ripoll-Vera T, Perez Luengo C, Borondo Alcazar JC, Garcia Ruiz AB, Sanchez Del Valle N, Barcelo Martin B, et al. Sudden cardiac death in persons aged 50 years or younger: diagnostic yield of a regional molecular autopsy program using massive sequencing. Rev Esp Cardiol (Engl Ed). 2021;74 (5):402-13.
209. Iglesias M, Ripoll-Vera T, Perez-Luengo C, Garcia AB, Moyano S, Canos JC, et al. Diagnostic Yield of Genetic Testing in Sudden Cardiac Death with Autopsy Findings of Uncertain Significance. J Clin Med. 2021;10 (9).
210. Rodriguez-Lopez R, Garcia-Planells J, Martinez-Matilla M, Perez-Garcia C, Garcia Banacloy A, Guzman Lujan C, et al. Homozygous Pro1066Arg MYBPC3 Pathogenic Variant in a 26Mb Region of Homozygosity Associated with Severe Hypertrophic Cardiomyopathy in a Patient of an Apparent Non-Consanguineous Family. Life (Basel). 2022;12 (7).
211. Dellefave LM, Pytel P, Mewborn S, Mora B, Guris DL, Fedson S, et al. Sarcomere mutations in cardiomyopathy with left ventricular hypertrabeculation. Circ Cardiovasc Genet. 2009;2 (5):442-9.
212. Petropoulou E, Soltani M, Firoozabadi AD, Namayandeh SM, Crockford J, Maroofian R, et al. Digenic inheritance of mutations in the cardiac troponin (TNNT2) and cardiac beta myosin heavy chain (MYH7) as the cause of severe dilated cardiomyopathy. Eur J Med Genet. 2017;60 (9):485-8.
213. Toepfer CN, Garfinkel AC, Venturini G, Wakimoto H, Repetti G, Alamo L, et al. Myosin Sequestration Regulates Sarcomere Function, Cardiomyocyte Energetics, and Metabolism, Informing the Pathogenesis of Hypertrophic Cardiomyopathy. Circulation. 2020;141 (10):828-42.
214. Kolokotronis K, Pluta N, Klopocki E, Kunstmann E, Messroghli D, Maack C, et al. New Insights on Genetic Diagnostics in Cardiomyopathy and Arrhythmia Patients Gained by Stepwise Exome Data Analysis. J Clin Med. 2020;9 (7).
215. Lu Q-W, Morimoto S, Harada K, Du C-K, Takahashi-Yanaga F, Miwa Y, et al. Cardiac troponin T mutation R141W found in dilated cardiomyopathy stabilizes the troponin T–tropomyosin interaction and causes a Ca2+ desensitization. Journal of Molecular and Cellular Cardiology. 2003;35 (12):1421-7.
216. van der Velden J, Stienen GJM. Cardiac Disorders and Pathophysiology of Sarcomeric Proteins. Physiol Rev. 2019;99 (1):381-426.

217. Hershberger RE, Pinto JR, Parks SB, Kushner JD, Li D, Ludwigsen S, et al. Clinical and functional characterization of TNNT2 mutations identified in patients with dilated cardiomyopathy. Circ Cardiovasc Genet. 2009;2 (4):306-13.
218. Jeon SB, Kim H, Chun KH, Oh J, Kwon C, Choi HK, et al. Human induced pluripotent stem cell line YCMi007-A generated from a dilated cardiomyopathy patient with a heterozygous dominant c.613C > T (p. Arg205Trp) variant of the TNNT2 gene. Stem Cell Res. 2023;67:103048.
219. Cuenca S, Ruiz-Cano MJ, Gimeno-Blanes JR, Jurado A, Salas C, Gomez-Diaz I, et al. Genetic basis of familial dilated cardiomyopathy patients undergoing heart transplantation. The Journal of Heart and Lung Transplantation. 2016;35 (5):625-35.
220. Dittmer TA, Sahni N, Kubben N, Hill DE, Vidal M, Burgess RC, et al. Systematic identification of pathological lamin A interactors. Mol Biol Cell. 2014;25 (9):1493-510.
221. Wang S, Peng D. Case series: LMNA-related dilated cardiomyopathy presents with reginal wall akinesis and transmural late gadolinium enhancement. ESC Heart Fail. 2020;7 (5):3179-83.
222. Dialynas G, Flannery KM, Zirbel LN, Nagy PL, Mathews KD, Moore SA, et al. LMNA variants cause cytoplasmic distribution of nuclear pore proteins in Drosophila and human muscle. Hum Mol Genet. 2012;21 (7):1544-56.
223. Shah PP, Lv W, Rhoades JH, Poleshko A, Abbey D, Caporizzo MA, et al. Pathogenic LMNA variants disrupt cardiac lamina-chromatin interactions and de-repress alternative fate genes. Cell Stem Cell. 2021;28 (5):938-54 e9.
224. Bhattacharjee P, Banerjee A, Banerjee A, Dasgupta D, Sengupta K. Structural alterations of Lamin A protein in dilated cardiomyopathy. Biochemistry. 2013;52 (24):4229-41.
225. Brauch KM, Karst ML, Herron KJ, de Andrade M, Pellikka PA, Rodeheffer RJ, et al. Mutations in ribonucleic acid binding protein gene cause familial dilated cardiomyopathy. J Am Coll Cardiol. 2009;54 (10):930-41.
226. Sedaghat-Hamedani F, Haas J, Zhu F, Geier C, Kayvanpour E, Liss M, et al. Clinical genetics and outcome of left ventricular non-compaction cardiomyopathy. Eur Heart J. 2017;38 (46):3449-60.
227. Koelemen J, Gotthardt M, Steinmetz LM, Meder B. RBM20-Related Cardiomyopathy: Current Understanding and Future Options. J Clin Med. 2021;10 (18).
228. Pantou MP, Gourzi P, Gkouziouta A, Tsiapras D, Zygouri C, Constantoulakis P, et al. Phenotypic Heterogeneity within Members of a Family Carrying the Same RBM20 Mutation R634W. Cardiology. 2018;141 (3):150-5.
229. Guo W, Schafer S, Greaser ML, Radke MH, Liss M, Govindarajan T, et al. RBM20, a gene for hereditary cardiomyopathy, regulates titin splicing. Nat Med. 2012;18 (5):766-73.
230. Zushin PH, Zhou Y, Li A, Ashley EA, Wheeler MT, Wu JC. Generation of two human iPSC lines with Exon 3 mutations in BCL2-Associated Athanogene 3 (BAG3) from dilated cardiomyopathy patients. Stem Cell Res. 2023;67:103019.
231. Chami N, Tadros R, Lemarbre F, Lo KS, Beaudoin M, Robb L, et al. Nonsense mutations in BAG3 are associated with early-onset dilated cardiomyopathy in French Canadians. Can J Cardiol. 2014;30 (12):1655-61.
232. Seidel F, Holtgrewe M, Al-Wakeel-Marquard N, Opgen-Rhein B, Dartsch J, Herbst C, et al. Pathogenic Variants Associated With Dilated Cardiomyopathy Predict Outcome in Pediatric Myocarditis. Circ Genom Precis Med. 2021;14 (4):e003250.
233. Qu HQ, Feldman AM, Hakonarson H. Genetics of BAG3: A Paradigm for Developing Precision Therapies for Dilated Cardiomyopathies. J Am Heart Assoc. 2022;11 (23):e027373.
234. Dominguez F, Cuenca S, Bilinska Z, Toro R, Villard E, Barriales-Villa R, et al. Dilated Cardiomyopathy Due to BLC2-Associated Athanogene 3 (BAG3) Mutations. J Am Coll Cardiol. 2018;72 (20):2471-81.

235. Ortiz-Genga MF, Cuenca S, Dal Ferro M, Zorio E, Salgado-Aranda R, Climent V, et al. Truncating FLNC Mutations Are Associated With High-Risk Dilated and Arrhythmogenic Cardiomyopathies. J Am Coll Cardiol. 2016;68 (22):2440-51.
236. Akhtar MM, Lorenzini M, Pavlou M, Ochoa JP, O'Mahony C, Restrepo-Cordoba MA, et al. Association of Left Ventricular Systolic Dysfunction Among Carriers of Truncating Variants in Filamin C With Frequent Ventricular Arrhythmia and End-stage Heart Failure. JAMA Cardiol. 2021.
237. Mao Z, Nakamura F. Structure and Function of Filamin C in the Muscle Z-Disc. Int J Mol Sci. 2020;21 (8).
238. Brodehl A, Ferrier RA, Hamilton SJ, Greenway SC, Brundler MA, Yu W, et al. Mutations in FLNC are Associated with Familial Restrictive Cardiomyopathy. Hum Mutat. 2016;37 (3):269-79.
239. Tuffery-Giraud S, Saquet C, Thorel D, Disset A, Rivier F, Malcolm S, et al. Mutation spectrum leading to an attenuated phenotype in dystrophinopathies. Eur J Hum Genet. 2005;13 (12):1254-60.
240. Diegoli M, Grasso M, Favalli V, Serio A, Gambarin FI, Klersy C, et al. Diagnostic work-up and risk stratification in X-linked dilated cardiomyopathies caused by dystrophin defects. J Am Coll Cardiol. 2011;58 (9):925-34.
241. Garcia-Pavia P, Syrris P, Salas C, Evans A, Mirelis JG, Cobo-Marcos M, et al. Desmosomal protein gene mutations in patients with idiopathic dilated cardiomyopathy undergoing cardiac transplantation: a clinicopathological study. Heart. 2011;97 (21):1744-52.
242. Gerull B, Heuser A, Wichter T, Paul M, Basson CT, McDermott DA, et al. Mutations in the desmosomal protein plakophilin-2 are common in arrhythmogenic right ventricular cardiomyopathy. Nat Genet. 2004;36 (11):1162-4.
243. Refaat MM, Hassanieh S, Ballout JA, Zakka P, Hotait M, Khalil A, et al. Non-familial cardiomyopathies in Lebanon: exome sequencing results for five idiopathic cases. BMC Med Genomics. 2019;12 (1):33.
244. Costa S, Cerrone M, Saguner AM, Brunckhorst C, Delmar M, Duru F. Arrhythmogenic cardiomyopathy: An in-depth look at molecular mechanisms and clinical correlates. Trends Cardiovasc Med. 2021;31 (7):395-402.
245. Kirchner F, Schuetz A, Boldt LH, Martens K, Dittmar G, Haverkamp W, et al. Molecular insights into arrhythmogenic right ventricular cardiomyopathy caused by plakophilin-2 missense mutations. Circ Cardiovasc Genet. 2012;5 (4):400-11.
246. Walsh R, Thomson KL, Ware JS, Funke BH, Woodley J, McGuire KJ, et al. Reassessment of Mendelian gene pathogenicity using 7,855 cardiomyopathy cases and 60,706 reference samples. Genetics in Medicine. 2017;19 (2):192-203.
247. Karvonen V, Harjama L, Helio K, Kettunen K, Elomaa O, Koskenvuo JW, et al. A novel desmoplakin mutation causes dilated cardiomyopathy with palmoplantar keratoderma as an early clinical sign. J Eur Acad Dermatol Venereol. 2022.
248. Peters S, Thompson BA, Perrin M, James P, Zentner D, Kalman JM, et al. Arrhythmic Phenotypes Are a Defining Feature of Dilated Cardiomyopathy-Associated SCN5A Variants: A Systematic Review. Circ Genom Precis Med. 2022;15 (1):e003432.
249. Larsen MK, Christiansen SL, Hertz CL, Frank-Hansen R, Jensen HK, Banner J, et al. Targeted molecular genetic testing in young sudden cardiac death victims from Western Denmark. Int J Legal Med. 2020;134 (1):111-21.
250. Murray B, James CA. Genotype-phenotype Correlates in Arrhythmogenic Cardiomyopathies. Curr Cardiol Rep. 2022;24 (11):1557-65.
251. Kamiza AB, Toure SM, Vujkovic M, Machipisa T, Soremekun OS, Kintu C, et al. Transferability of genetic risk scores in African populations. Nat Med. 2022;28 (6):1163-6.
252. Ezekian JE, Clippinger SR, Garcia JM, Yang Q, Denfield S, Jeewa A, et al. Variant R94C in TNNT2-Encoded Troponin T Predisposes to Pediatric Restrictive Cardiomyopathy

and Sudden Death Through Impaired Thin Filament Relaxation Resulting in Myocardial Diastolic Dysfunction. J Am Heart Assoc. 2020;9 (5):e015111.
253. Papamanoli A, Muncan B, Yoo J, Psevdos G, Kalogeropoulos AP. Human Immunodeficiency Virus Infection-Associated Cardiomyopathy and Heart Failure. Journal of personalized medicine. 2022;12 (11).
254. Ritter A, Bedoukian E, Berger JH, Copenheaver D, Gray C, Krantz I, et al. Clinical utility of exome sequencing in infantile heart failure. Genet Med. 2020;22 (2):423-6.
255. Ciconte G, Monasky MM, Santinelli V, Micaglio E, Vicedomini G, Anastasia L, et al. Brugada syndrome genetics is associated with phenotype severity. Eur Heart J. 2021;42 (11):1082-90.
256. Kean AC, Helm BM, Vatta M, Ayers MD, Parent JJ, Darragh RK. Clinical characterisation of a novel SCN5A variant associated with progressive malignant arrhythmia and dilated cardiomyopathy. Cardiol Young. 2019;29 (10):1257-63.
257. Isogai T, Kamiya CA. Worldwide Incidence of Peripartum Cardiomyopathy and Overall Maternal Mortality. Int Heart J. 2019;60 (3):503-11.
258. Blair E, Price SJ, Baty CJ, Ostman-Smith I, Watkins H. Mutations in cis can confound genotype-phenotype correlations in hypertrophic cardiomyopathy. J Med Genet. 2001;38 (6):385-8.
259. Tibazarwa K, Lee G, Mayosi B, Carrington M, Stewart S, Sliwa K. The 12-lead ECG in peripartum cardiomyopathy. Cardiovascular Journal of Africa. 2012;23 (6):322-9.
260. Njoroge JN, Mangena JC, Aribeana C, Parikh VN. Emerging Genotype-Phenotype Associations in Dilated Cardiomyopathy. Curr Cardiol Rep. 2022;24 (9):1077-84.
261. Schultze-Berndt A, Kuhnisch J, Herbst C, Seidel F, Al-Wakeel-Marquard N, Dartsch J, et al. Reduced Systolic Function and Not Genetic Variants Determine Outcome in Pediatric and Adult Left Ventricular Noncompaction Cardiomyopathy. Front Pediatr. 2021;9:722926.
262. Al-Hassnan ZN, Almesned A, Tulbah S, Alakhfash A, Alhadeq F, Alruwaili N, et al. Categorized Genetic Analysis in Childhood-Onset Cardiomyopathy. Circ Genom Precis Med. 2020;13 (5):504-14.
263. van Waning JI, Moesker J, Heijsman D, Boersma E, Majoor-Krakauer D. Systematic Review of Genotype-Phenotype Correlations in Noncompaction Cardiomyopathy. J Am Heart Assoc. 2019;8 (23):e012993.
264. Walsh R, Offerhaus JA, Tadros R, Bezzina CR. Minor hypertrophic cardiomyopathy genes, major insights into the genetics of cardiomyopathies. Nat Rev Cardiol. 2022;19 (3):151-67.
265. Duvekot JC, Baas AF, Volker-Touw CML, Bikker H, Schroer C, Breur J. Early Lethal Noncompaction Cardiomyopathy in Siblings With Compound Heterozygous RYR2 Variant. Can J Cardiol. 2021;37 (11):1864-6.
266. Collyer J, Xu F, Munkhsaikhan U, Alberson NF, Orgil BO, Zhang W, et al. Combining whole exome sequencing with in silico analysis and clinical data to identify candidate variants in pediatric left ventricular noncompaction. Int J Cardiol. 2022;347:29-37.
267. Chen Q, Zheng W, Xu H, Yang Y, Song Z, Yuan L, et al. Digenic Variants in the TTN and TRAPPC11 Genes Co-segregating With a Limb-Girdle Muscular Dystrophy in a Han Chinese Family. Front Neurosci. 2021;15:601757.
268. Waldmuller S, Schroeder C, Sturm M, Scheffold T, Imbrich K, Junker S, et al. Targeted 46-gene and clinical exome sequencing for mutations causing cardiomyopathies. Mol Cell Probes. 2015;29 (5):308-14.
269. Blankenburg R, Hackert K, Wurster S, Deenen R, Seidman JG, Seidman CE, et al. beta-Myosin heavy chain variant Val606Met causes very mild hypertrophic cardiomyopathy in mice, but exacerbates HCM phenotypes in mice carrying other HCM mutations. Circ Res. 2014;115 (2):227-37.

270. Pirruccello JP, Bick A, Wang M, Chaffin M, Friedman S, Yao J, et al. Analysis of cardiac magnetic resonance imaging in 36,000 individuals yields genetic insights into dilated cardiomyopathy. Nat Commun. 2020;11 (1):2254.
271. Adam Y, Sadeeq S, Kumuthini J, Ajayi O, Wells G, Solomon R, et al. Polygenic Risk Score in African populations: progress and challenges. F1000Res. 2022;11:175.
272. McClellan JM, Lehner T, King MC. Gene Discovery for Complex Traits: Lessons from Africa. Cell. 2017;171 (2):261-4.
273. Brignole M, Cecchi F, Anastasakis A, Crotti L, Deharo JC, Elliott PM, et al. Syncope in hypertrophic cardiomyopathy (part II): An expert consensus statement on the diagnosis and management. Int J Cardiol. 2023;370:330-7.
274. Maron BJ, Gardin JM, Flack JM, Gidding SS, Kurosaki TT, Bild DE. 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. 1995;92 (4):785-9.
275. Semsarian C, Ingles J, Maron MS, Maron BJ. New Perspectives on the Prevalence of Hypertrophic Cardiomyopathy. Journal of the American College of Cardiology. 2015;65 (12):1249-54.
276. Maron BJ, Longo DL. Clinical Course and Management of Hypertrophic Cardiomyopathy. New Engl J Med. 2018;379 (7):655-68.
277. Marstrand P, Han L, Day SM, Olivotto I, Ashley EA, Michels M, et al. Hypertrophic Cardiomyopathy With Left Ventricular Systolic Dysfunction: Insights From the SHaRe Registry. Circulation. 2020;141 (17):1371-83.
278. Abou Alaiwi S, Roston TM, Marstrand P, Claggett BL, Parikh VN, Helms AS, et al. Left Ventricular Systolic Dysfunction in Patients Diagnosed With Hypertrophic Cardiomyopathy During Childhood: Insights From the SHaRe Registry. Circulation. 2023;148 (5):394-404.
279. Dass S, Cochlin LE, Suttie JJ, Holloway CJ, Rider OJ, Carden L, et al. Exacerbation of cardiac energetic impairment during exercise in hypertrophic cardiomyopathy: a potential mechanism for diastolic dysfunction. European Heart Journal. 2015;36 (24):1547-54.
280. Rosenzweig A, Watkins H, Hwang DS, Miri M, McKenna W, Traill TA, et al. Preclinical diagnosis of familial hypertrophic cardiomyopathy by genetic analysis of blood lymphocytes. N Engl J Med. 1991;325 (25):1753-60.
281. Hudson J, Sturm AC, Salberg L, Brennan S, Quinn GP, Vadaparampil ST. Disclosure of diagnosis to at-risk relatives by individuals diagnosed with hypertrophic cardiomyopathy (HCM). J Community Genet. 2019;10 (2):207-17.
282. Marian AJ, Braunwald E. Hypertrophic Cardiomyopathy: Genetics, Pathogenesis, Clinical Manifestations, Diagnosis, and Therapy. Circ Res. 2017;121 (7):749-70.
283. Ashrafian H, Redwood C, Blair E, Watkins H. Hypertrophic cardiomyopathy: A paradigm for myocardial energy depletion. Trends in Genetics. 2003;19 (5):263-8.
284. Walsh R, Buchan R, Wilk A, John S, Felkin LE, Thomson KL, et al. Defining the genetic architecture of hypertrophic cardiomyopathy: re-evaluating the role of non-sarcomeric genes. European Heart Journal. 2017;38 (46):3461-8.
285. Pua CJ, Tham N, Chin CW, Walsh R, Khor CC, Toepfer CN, et al. Genetic Studies of Hypertrophic Cardiomyopathy in Singaporeans Identify Variants in TNNI3 and TNNT2 that Are Common in Chinese Patients. Circ Genom Precis Med. 2020.
286. Jaaskelainen P, Vangipurapu J, Raivo J, Kuulasmaa T, Helio T, Aalto-Setala K, et al. Genetic basis and outcome in a nationwide study of Finnish patients with hypertrophic cardiomyopathy. ESC Heart Fail. 2019;6 (2):436-45.
287. Nakashima Y, Kubo T, Sugiura K, Ochi Y, Takahashi A, Baba Y, et al. Lifelong Clinical Impact of the Presence of Sarcomere Gene Mutation in Japanese Patients With Hypertrophic Cardiomyopathy. Circ J. 2020;84 (10):1846-53.

288. Zhang M, Sun X, Wu G, Wang D, Wang L, Zhang C, et al. Effect of Cis-Compound Variants in MYH7 on Hypertrophic Cardiomyopathy With a Mild Phenotype. Am J Cardiol. 2022;167:104-10.
289. Agarwal R, Wakimoto H, Paulo JA, Zhang Q, Reichart D, Toepfer C, et al. Pathogenesis of Cardiomyopathy Caused by Variants in ALPK3, an Essential Pseudokinase in the Cardiomyocyte Nucleus and Sarcomere. Circulation. 2022;146 (22):1674-93.
290. Lipari M, Wypasek E, Karpinski M, Tomkiewicz-Pajak L, Laino L, Binni F, et al. Identification of a variant hotspot in MYBPC3 and of a novel CSRP3 autosomal recessive alteration in a cohort of Polish patients with hypertrophic cardiomyopathy. Pol Arch Intern Med. 2020;130 (2):89-99.
291. Salazar-Mendiguchia J, Barriales-Villa R, Lopes LR, Ochoa JP, Rodriguez-Vilela A, Palomino-Doza J, et al. The p. (Cys150Tyr) variant in CSRP3 is associated with late-onset hypertrophic cardiomyopathy in heterozygous individuals. Eur J Med Genet. 2020;63 (12):104079.
292. Chen X, Zhai S, Xu H, Fan C, Yang Y, Hang F, et al. Mid-term outcomes of biventricular obstruction and left ventricular outflow tract obstruction after surgery correction in child and adolescent patients with hypertrophic cardiomyopathy. Plos One. 2018;13 (2).
293. Fahed AC, Nemer G, Bitar FF, Arnaout S, Abchee AB, Batrawi M, et al. Founder Mutation in N-terminus of Cardiac Troponin I Causes Malignant Hypertrophic Cardiomyopathy. Circ Genom Precis Med. 2020.
294. Wolf CM. Hypertrophic cardiomyopathy: genetics and clinical perspectives. Cardiovasc Diagn Ther. 2019;9 (Suppl 2):S388-S415.
295. Lipov A, Jurgens SJ, Mazzarotto F, Allouba M, Pirruccello JP, Aguib Y, et al. Exploring the complex spectrum of dominance and recessiveness in genetic cardiomyopathies. Nature Cardiovascular Research. 2023;2 (11):1078-94.
296. Darwish RK, Haghighi A, Seliem ZS, El-Saiedi SA, Radwan NH, El-Gayar DF, et al. Genetic study of pediatric hypertrophic cardiomyopathy in Egypt. Cardiol Young. 2020:1-7.
297. Zhou N, Weng H, Zhao W, Tang L, Ge Z, Tian F, et al. Gene-echocardiography: refining genotype-phenotype correlations in hypertrophic cardiomyopathy. Eur Heart J Cardiovasc Imaging. 2023;25 (1):127-35.
298. Prinz C, Farr M, Hering D, Horstkotte D, Faber L. The Diagnosis and Treatment of Hypertrophic Cardiomyopathy. Deutsches Ärzteblatt international. 2011.
299. Trachoo O, Yingchoncharoen T, Ngernsritrakul T, Iemwimangsa N, Panthan B, Klumsathian S, et al. Genomic findings of hypertrophic and dilated cardiomyopathy characterized in a Thai clinical genetics service. PLoS One. 2022;17 (9):e0267770.
300. Ingles J, Goldstein J, Thaxton C, Caleshu C, Corty EW, Crowley SB, et al. Evaluating the Clinical Validity of Hypertrophic Cardiomyopathy Genes. Circulation: Genomic and Precision Medicine. 2019;12 (2).
301. Yang Q, Sanbe A, Osinska H, Hewett TE, Klevitsky R, Robbins J. A mouse model of myosin binding protein C human familial hypertrophic cardiomyopathy. Journal of Clinical Investigation. 1998;102 (7):1292-300.
302. Heitner SB, Jacoby D, Lester SJ, Owens A, Wang A, Zhang D, et al. Mavacamten Treatment for Obstructive Hypertrophic Cardiomyopathy. Annals of Internal Medicine. 2019;170 (11).
303. Prondzynski M, Mearini G, Carrier L. Gene therapy strategies in the treatment of hypertrophic cardiomyopathy. Pflugers Arch. 2019;471 (5):807-15.
304. Gedicke-Hornung C, Behrens-Gawlik V, Reischmann S, Geertz B, Stimpel D, Weinberger F, et al. Rescue of cardiomyopathy through U7snRNA-mediated exon skipping in Mybpc3-targeted knock-in mice. EMBO Molecular Medicine. 2013;5 (7):1128-45.

305. Vignier N, Schlossarek S, Fraysse B, Mearini G, Kramer E, Pointu H, et al. Nonsense-mediated mRNA decay and ubiquitin-proteasome system regulate cardiac myosin-binding protein C mutant levels in cardiomyopathic mice. Circ Res. 2009;105 (3):239-48.
306. Robinson P, Liu X, Sparrow A, Patel S, Zhang YH, Casadei B, et al. Hypertrophic cardiomyopathy mutations increase myofilament Ca2+ buffering, alter intracellular Ca2+ handling, and stimulate Ca2+-dependent signaling. Journal of Biological Chemistry. 2018;293 (27):10487-99.
307. Lakdawala NK, Olivotto I, Day SM, Han L, Ashley EA, Michels M, et al. Associations Between Female Sex, Sarcomere Variants, and Clinical Outcomes in Hypertrophic Cardiomyopathy. Circ Genom Precis Med. 2021;14 (1):e003062.
308. Ricci F, Banihashemi B, Pirouzifard M, Sundquist J, Sundquist K, Sutton R, et al. Familial risk of dilated and hypertrophic cardiomyopathy: a national family study in Sweden. ESC Heart Fail. 2023;10 (1):121-32.
309. Petukh M, Kucukkal TG, Alexov E. On human disease-causing amino acid variants: statistical study of sequence and structural patterns. Hum Mutat. 2015;36 (5):524-34.
310. Siontis KC, Ommen SR, Geske JB. Sex, Survival, and Cardiomyopathy: Differences Between Men and Women With Hypertrophic Cardiomyopathy. J Am Heart Assoc. 2019;8 (21):e014448.
311. Herrera-Rodriguez DL, Totomoch-Serra A, Rosas-Madrigal S, Luna-Limon C, Marroquin-Ramirez D, Carnevale A, et al. Genes frequently associated with sudden death in primary hypertrophic cardiomyopathy. Arch Cardiol Mex. 2020;90 (1):58-68.
312. Huang M, Zhang J, Song C, Wang S, Zhou Z, Wang H, et al. SARC Gene Mutation Is Associated With Myocardial Fibrosis Measured by Histopathology and Cardiac Magnetic Resonance in Patients With Hypertrophic Cardiomyopathy. J Am Heart Assoc. 2023;12 (5):e028293.
313. Richard P, Charron P, Carrier L, Ledeuil C, Cheav T, Pichereau C, et al. Hypertrophic cardiomyopathy: distribution of disease genes, spectrum of mutations, and implications for a molecular diagnosis strategy. Circulation. 2003;107 (17):2227-32.
314. Moolman JC, Brink PA, Corfield VA. Identification of a novel Ala797Thr mutation in exon 21 of the beta-myosin heavy chain gene in hypertrophic cardiomyopathy. Hum Mutat. 1995;6 (2):197-8.
315. Van Driest SL, Jaeger MA, Ommen SR, Will ML, Gersh BJ, Tajik AJ, et al. Comprehensive analysis of the beta-myosin heavy chain gene in 389 unrelated patients with hypertrophic cardiomyopathy. J Am Coll Cardiol. 2004;44 (3):602-10.
316. Nagyova E, Radvanszky J, Hyblova M, Simovicova V, Goncalvesova E, Asselbergs FW, et al. Targeted next-generation sequencing in Slovak cardiomyopathy patients. Bratisl Lek Listy. 2019;120 (1):46-51.
317. Patel AP, Dron JS, Wang M, Pirruccello JP, Ng K, Natarajan P, et al. Association of Pathogenic DNA Variants Predisposing to Cardiomyopathy With Cardiovascular Disease Outcomes and All-Cause Mortality. JAMA Cardiol. 2022;7 (7):723-32.
318. Lahrouchi N, Raju H, Lodder EM, Papatheodorou E, Ware JS, Papadakis M, et al. Utility of Post-Mortem Genetic Testing in Cases of Sudden Arrhythmic Death Syndrome. J Am Coll Cardiol. 2017;69 (17):2134-45.
319. Homburger JR, Green EM, Caleshu C, Sunitha MS, Taylor RE, Ruppel KM, et al. Multidimensional structure-function relationships in human beta-cardiac myosin from population-scale genetic variation. Proc Natl Acad Sci U S A. 2016;113 (24):6701-6.
320. Michels M, Soliman OI, Kofflard MJ, Hoedemaekers YM, Dooijes D, Majoor-Krakauer D, et al. Diastolic abnormalities as the first feature of hypertrophic cardiomyopathy in Dutch myosin-binding protein C founder mutations. JACC Cardiovasc Imaging. 2009;2 (1):58-64.

321. Erdmann J, Raible J, Maki-Abadi J, Hummel M, Hammann J, Wollnik B, et al. Spectrum of clinical phenotypes and gene variants in cardiac myosin-binding protein C mutation carriers with hypertrophic cardiomyopathy. J Am Coll Cardiol. 2001;38 (2):322-30.
322. Gajendrarao P, Krishnamoorthy N, Selvaraj S, Girolami F, Cecchi F, Olivotto I, et al. An Investigation of the Molecular Mechanism of Double cMyBP-C Mutation in a Patient with End-Stage Hypertrophic Cardiomyopathy. J Cardiovasc Transl Res. 2015;8 (4):232-43.
323. Marston S, Copeland O, Jacques A, Livesey K, Tsang V, McKenna WJ, et al. Evidence From Human Myectomy Samples That MYBPC3 Mutations Cause Hypertrophic Cardiomyopathy Through Haploinsufficiency. Circulation Research. 2009;105 (3):219-U39.
324. Niimura H, Bachinski LL, Sangwatanaroj S, Watkins H, Chudley AE, McKenna W, et al. Mutations in the gene for cardiac myosin-binding protein C and late-onset familial hypertrophic cardiomyopathy. N Engl J Med. 1998;338 (18):1248-57.
325. Lynch TL, Kumar M, McNamara JW, Kuster DWD, Sivaguru M, Singh RR, et al. Amino terminus of cardiac myosin binding protein-C regulates cardiac contractility. Journal of Molecular and Cellular Cardiology. 2021;156:33-44.
326. Toepfer CN, Wakimoto H, Garfinkel AC, McDonough B, Liao D, Jiang JM, et al. Hypertrophic cardiomyopathy mutations in MYBPC3 dysregulate myosin. Science Translational Medicine. 2019;11 (476).
327. Nafissi NA, Abdulrahim JW, Kwee LC, Coniglio AC, Kraus WE, Piccini JP, et al. Prevalence and Phenotypic Burden of Monogenic Arrhythmias Using Integration of Electronic Health Records With Genetics. Circ Genom Precis Med. 2022;15 (5):e003675.
328. De Lange WJ, Grimes AC, Hegge LF, Spring AM, Brost TM, Ralphe JC. E258K HCM-causing mutation in cardiac MyBP-C reduces contractile force and accelerates twitch kinetics by disrupting the cMyBP-C and myosin S2 interaction. J Gen Physiol. 2013;142 (3):241-55.
329. Mouton JM, Pellizzon AS, Goosen A, Kinnear CJ, Herbst PG, Brink PA, et al. Diagnostic disparity and identification of two TNNI3 gene mutations, one novel and one arising de novo, in South African patients with restrictive cardiomyopathy and focal ventricular hypertrophy. Cardiovasc J Afr. 2015;26 (2):63-9.
330. Mogensen J, Kubo T, Duque M, Uribe W, Shaw A, Murphy R, et al. Idiopathic restrictive cardiomyopathy is part of the clinical expression of cardiac troponin I mutations. J Clin Invest. 2003;111 (2):209-16.
331. Cimiotti D, Budde H, Hassoun R, Jaquet K. Genetic Restrictive Cardiomyopathy: Causes and Consequences-An Integrative Approach. Int J Mol Sci. 2021;22 (2).
332. Caglayan AO, Sezer RG, Kaymakcalan H, Ulgen E, Yavuz T, Baranoski JF, et al. ALPK3 gene mutation in a patient with congenital cardiomyopathy and dysmorphic features. Cold Spring Harb Mol Case Stud. 2017;3 (5).
333. Cheawsamoot C, Phokaew C, Chetruengchai W, Chantranuwat P, Puwanant S, Tongsima S, et al. A Pathogenic Variant in ALPK3 Is Associated With an Autosomal Dominant Adult-onset Hypertrophic Cardiomyopathy. Circ Genom Precis Med. 2020;13 (6):e003127.
334. Herkert JC, Verhagen JMA, Yotti R, Haghighi A, Phelan DG, James PA, et al. Expanding the clinical and genetic spectrum of ALPK3 variants: Phenotypes identified in pediatric cardiomyopathy patients and adults with heterozygous variants. Am Heart J. 2020;225:108-19.
335. Lopes LR, Garcia-Hernandez S, Lorenzini M, Futema M, Chumakova O, Zateyshchikov D, et al. Alpha-protein kinase 3 (ALPK3) truncating variants are a cause of autosomal dominant hypertrophic cardiomyopathy. Eur Heart J. 2021;42 (32):3063-73.
336. Phelan DG, Anderson DJ, Howden SE, Wong RC, Hickey PF, Pope K, et al. ALPK3-deficient cardiomyocytes generated from patient-derived induced pluripotent stem cells and mutant human embryonic stem cells display abnormal calcium handling and establish that ALPK3 deficiency underlies familial cardiomyopathy. Eur Heart J. 2016;37 (33):2586-90.

337. Eng CM, Resnick-Silverman LA, Niehaus DJ, Astrin KH, Desnick RJ. Nature and frequency of mutations in the alpha-galactosidase A gene that cause Fabry disease. Am J Hum Genet. 1993;53 (6):1186-97.
338. Lukas J, Scalia S, Eichler S, Pockrandt A-M, Dehn N, Cozma C, et al. Functional and Clinical Consequences of Novel α-Galactosidase A Mutations in Fabry Disease. Human Mutation. 2016;37 (1):43-51.
339. Cecconi M, Parodi MI, Formisano F, Spirito P, Autore C, Musumeci MB, et al. Targeted next-generation sequencing helps to decipher the genetic and phenotypic heterogeneity of hypertrophic cardiomyopathy. International Journal of Molecular Medicine. 2016;38 (4):1111-24.
340. Pasqualim G, dos Santos BA, Giugliani R, Matte U. Simple and efficient screening of patients with Fabry disease with high resolution melting. Clinical Biochemistry. 2018;53:160-3.
341. Marian AJ. Modifier genes for hypertrophic cardiomyopathy. Curr Opin Cardiol. 2002;17 (3):242-52.
342. Chen Y, Xu F, Munkhsaikhan U, Boyle C, Borcky T, Zhao W, et al. Identifying modifier genes for hypertrophic cardiomyopathy. J Mol Cell Cardiol. 2020;144:119-26.
343. Roldan-Sevilla A, Palomino-Doza J, de Juan J, Sanchez V, Dominguez-Gonzalez C, Salguero-Bodes R, et al. Missense Mutations in the FLNC Gene Causing Familial Restrictive Cardiomyopathy. Circ Genom Precis Med. 2019;12 (3):e002388.
344. Gaudreault N, Ruel LJ, Henry C, Schleit J, Lague P, Champagne J, et al. Novel filamin C (FLNC) variant causes a severe form of familial mixed hypertrophic-restrictive cardiomyopathy. Am J Med Genet A. 2023;191 (6):1508-17.
345. Biddinger KJ, Jurgens SJ, Maamari D, Gaziano L, Choi SH, Morrill VN, et al. Rare and Common Genetic Variation Underlying the Risk of Hypertrophic Cardiomyopathy in a National Biobank. JAMA Cardiol. 2022;7 (7):715-22.
346. Muchtar E, Blauwet LA, Gertz MA. Restrictive Cardiomyopathy: Genetics, Pathogenesis, Clinical Manifestations, Diagnosis, and Therapy. Circ Res. 2017;121 (7):819-37.
347. Chintanaphol M, Orgil BO, Alberson NR, Towbin JA, Purevjav E. Restrictive cardiomyopathy: from genetics and clinical overview to animal modeling. Rev Cardiovasc Med. 2022;23 (3).
348. Kostareva A, Kiselev A, Gudkova A, Frishman G, Ruepp A, Frishman D, et al. Genetic Spectrum of Idiopathic Restrictive Cardiomyopathy Uncovered by Next-Generation Sequencing. PLoS One. 2016;11 (9):e0163362.
349. Pereira NL, Grogan M, Dec GW. Spectrum of Restrictive and Infiltrative Cardiomyopathies. Journal of the American College of Cardiology. 2018;71 (10):1130-48.
350. Albakri A. Restrictive cardiomyopathy: A review of literature on clinical status and meta-analysis of diagnosis and clinical management. Pediatric Dimensions. 2018;3 (2).
351. Tariq M. Importance of genetic evaluation and testing in pediatric cardiomyopathy. World Journal of Cardiology. 2014;6 (11).
352. Brodehl A, Gerull B. Genetic Insights into Primary Restrictive Cardiomyopathy. Journal of Clinical Medicine. 2022;11 (8).
353. Wu W, Lu CX, Wang YN, Liu F, Chen W, Liu YT, et al. Novel Phenotype-Genotype Correlations of Restrictive Cardiomyopathy With Myosin-Binding Protein C (MYBPC3) Gene Mutations Tested by Next-Generation Sequencing. J Am Heart Assoc. 2015;4 (7).
354. Tucker NR, McLellan MA, Hu D, Ye J, Parsons VA, Mills RW, et al. Novel Mutation in FLNC (Filamin C) Causes Familial Restrictive Cardiomyopathy. Circ Cardiovasc Genet. 2017;10 (6).

355. Ding WH, Han L, Xiao YY, Mo Y, Yang J, Wang XF, et al. Role of Whole-exome Sequencing in Phenotype Classification and Clinical Treatment of Pediatric Restrictive Cardiomyopathy. Chin Med J (Engl). 2017;130 (23):2823-8.
356. Pantou MP, Gourzi P, Gkouziouta A, Armenis I, Kaklamanis L, Zygouri C, et al. A case report of recessive restrictive cardiomyopathy caused by a novel mutation in cardiac troponin I (TNNI3). BMC Med Genet. 2019;20 (1):61.
357. Newman LS, Rose CS, Bresnitz EA, Rossman MD, Barnard J, Frederick M, et al. A Case Control Etiologic Study of Sarcoidosis. American Journal of Respiratory and Critical Care Medicine. 2004;170 (12):1324-30.
358. Trachtenberg BH, Hare JM. Inflammatory Cardiomyopathic Syndromes. Circulation Research. 2017;121 (7):803-18.
359. Coelho T, Adams D, Silva A, Lozeron P, Hawkins PN, Mant T, et al. Safety and Efficacy of RNAi Therapy for Transthyretin Amyloidosis. New Engl J Med. 2013;369 (9):819-29.
360. Akinrinade O, Helio T, Lekanne Deprez RH, Jongbloed JDH, Boven LG, van den Berg MP, et al. Relevance of Titin Missense and Non-Frameshifting Insertions/Deletions Variants in Dilated Cardiomyopathy. Scientific reports. 2019;9 (1):4093.
361. Beaton A, Mocumbi AO. Diagnosis and Management of Endomyocardial Fibrosis. Cardiol Clin. 2017;35 (1):87-98.
362. Mocumbi AO, Ferreira MB, Sidi D, Yacoub MH. A population study of endomyocardial fibrosis in a rural area of Mozambique. N Engl J Med. 2008;359 (1):43-9.
363. Gowda SN, Ali H-J, Hussain I. Overview of Restrictive Cardiomyopathies. Methodist DeBakey Cardiovascular Journal. 2022;18 (2):4-16.
364. Freers J, Mayanja-Kizza H, Ziegler JL, Rutakingirwa M. Echocardiographic diagnosis of heart disease in Uganda. Trop Doct. 1996;26 (3):125-8.
365. Bar H, Goudeau B, Walde S, Casteras-Simon M, Mucke N, Shatunov A, et al. Conspicuous involvement of desmin tail mutations in diverse cardiac and skeletal myopathies. Hum Mutat. 2007;28 (4):374-86.
366. Dias R, Aguiar TC. DES c.1360C>T: A Rare Desmin Variant Causing Early Distal Myopathy and Cardiomyopathy. Cureus. 2023;15 (3):e36368.
367. Batonnet-Pichon S, Behin A, Cabet E, Delort F, Vicart P, Lilienbaum A. Myofibrillar Myopathies: New Perspectives from Animal Models to Potential Therapeutic Approaches. J Neuromuscul Dis. 2017;4 (1):1-15.
368. Patel V, Asatryan B, Siripanthong B, Munroe PB, Tiku-Owens A, Lopes LR, et al. State of the Art Review on Genetics and Precision Medicine in Arrhythmogenic Cardiomyopathy. Int J Mol Sci. 2020;21 (18).
369. van Spaendonck-Zwarts KY, van Hessem L, Jongbloed JDH, de Walle HEK, Capetanaki Y, van der Kooi AJ, et al. Desmin-related myopathy. Clinical Genetics. 2011;80 (4):354-66.
370. Wahbi K, Béhin A, Charron P, Dunand M, Richard P, Meune C, et al. High cardiovascular morbidity and mortality in myofibrillar myopathies due to DES gene mutations: a 10-year longitudinal study. Neuromuscular Disorders. 2012;22 (3):211-8.
371. Oomen A, Jones K, Yeates L, Semsarian C, Ingles J, Sy RW. Rare desmin variant causing penetrant life-threatening arrhythmic cardiomyopathy. HeartRhythm Case Rep. 2018;4 (7):318-23.
372. Otten E, Asimaki A, Maass A, van Langen IM, van der Wal A, de Jonge N, et al. Desmin mutations as a cause of right ventricular heart failure affect the intercalated disks. Heart Rhythm. 2010;7 (8):1058-64.
373. Parrott A, Khoury PR, Shikany AR, Lorts A, Villa CR, Miller EM. Investigation of de novo variation in pediatric cardiomyopathy. Am J Med Genet C Semin Med Genet. 2020;184 (1):116-23.

374. Ader F, Villard E, Ledeuil C, Charron P, Richard P. [Genotype-phenotype correlations of pathogenic variants in the FLNC gene]. Med Sci (Paris). 2018;34 Hors serie n degrees 2:39-41.
375. Zou Y, Wang J, Liu X, Wang Y, Chen Y, Sun K, et al. Multiple gene mutations, not the type of mutation, are the modifier of left ventricle hypertrophy in patients with hypertrophic cardiomyopathy. Mol Biol Rep. 2013;40 (6):3969-76.
376. Mosqueira D, Mannhardt I, Bhagwan JR, Lis-Slimak K, Katili P, Scott E, et al. CRISPR/Cas9 editing in human pluripotent stem cell-cardiomyocytes highlights arrhythmias, hypocontractility, and energy depletion as potential therapeutic targets for hypertrophic cardiomyopathy. Eur Heart J. 2018;39 (43):3879-92.
377. Joynt ACM, Axford MM, Chad L, Costain G. Understanding genetic variants of uncertain significance. Paediatrics & Child Health. 2022;27 (1):10-1.
378. Alotibi RS, Sannan NS, AlEissa M, Aldriwesh MG, Al Tuwaijri A, Akiel MA, et al. The diagnostic yield of CGH and WES in neurodevelopmental disorders. Front Pediatr. 2023;11:1133789.
379. Santiago CF, Huttner IG, Fatkin D. Mechanisms of TTNtv-Related Dilated Cardiomyopathy: Insights from Zebrafish Models. J Cardiovasc Dev Dis. 2021;8 (2).
380. Redwood CS, Moolman-Smook JC, Watkins H. Properties of mutant contractile proteins that cause hypertrophic cardiomyopathy. Cardiovascular Research. 1999;44 (1):20-36.
381. Matolweni LO, Bardien S, Rebello G, Oppon E, Munclinger M, Ramesar R, et al. Arrhythmogenic right ventricular cardiomyopathy type 6 (ARVC6): Support for the locus assignment, narrowing of the critical region and mutation screening of three candidate genes. BMC Medical Genetics. 2006;7.
382. Fan X, Yang G, Duru F, Grilli M, Akin I, Zhou X, et al. Arrhythmogenic Cardiomyopathy: from Preclinical Models to Genotype-phenotype Correlation and Pathophysiology. Stem Cell Rev Rep. 2023;19 (8):2683-708.
383. Migliore F, Mattesi G, Zorzi A, Bauce B, Rigato I, Corrado D, et al. Arrhythmogenic Cardiomyopathy—Current Treatment and Future Options. Journal of Clinical Medicine. 2021;10 (13).
384. Thiene G, Corrado D, Basso C. Arrhythmogenic right ventricular cardiomyopathy/dysplasia. Orphanet J Rare Dis. 2007;2 (1).
385. Marcus FI, McKenna WJ, Sherrill D, Basso C, Bauce B, Bluemke DA, et al. Diagnosis of arrhythmogenic right ventricular cardiomyopathy/dysplasia: proposed modification of the task force criteria. Circulation. 2010;121 (13):1533-41.
386. Kawano K, Kondo H, Takahashi M, Shinohara T, Ohno S, Horie M, et al. Arrhythmogenic right-ventricular cardiomyopathy with plakophilin-2 genetic variant concomitant with early manifestation of ventricular tachyarrhythmia: a case series. Eur Heart J Case Rep. 2022;6 (10):ytac397.
387. Corrado D, van Tintelen PJ, McKenna WJ, Hauer RNW, Anastastakis A, Asimaki A, et al. Arrhythmogenic right ventricular cardiomyopathy: evaluation of the current diagnostic criteria and differential diagnosis. Eur Heart J. 2020;41 (14):1414-29.
388. Matolweni LO, Bardien S, Rebello G, Oppon E, Munclinger M, Ramesar R, et al. Arrhythmogenic right ventricular cardiomyopathy type 6 (ARVC6): support for the locus assignment, narrowing of the critical region and mutation screening of three candidate genes. BMC Med Genet. 2006;7:29.
389. James CA, Calkins H. Arrhythmogenic Right Ventricular Cardiomyopathy: Progress Toward Personalized Management. Annu Rev Med. 2019;70:1-18.
390. Corrado D, Wichter T, Link MS, Hauer RNW, Marchlinski FE, Anastasakis A, et al. Treatment of Arrhythmogenic Right Ventricular Cardiomyopathy/Dysplasia. Circulation. 2015;132 (5):441-53.

391. Corrado D, Basso C, Judge DP. Arrhythmogenic Cardiomyopathy. Circulation Research. 2017;121 (7):784-802.
392. Saguner AM, Duru F, Brunckhorst C, Akdis D. Arrhythmogenic Cardiomyopathy: Electrical and Structural Phenotypes. Arrhythmia & Electrophysiology Review. 2016;5 (2).
393. Miles C, Finocchiaro G, Papadakis M, Gray B, Westaby J, Ensam B, et al. Sudden Death and Left Ventricular Involvement in Arrhythmogenic Cardiomyopathy. Circulation. 2019;139 (15):1786-97.
394. Austin KM, Trembley MA, Chandler SF, Sanders SP, Saffitz JE, Abrams DJ, et al. Molecular mechanisms of arrhythmogenic cardiomyopathy. Nat Rev Cardiol. 2019;16 (9):519-37.
395. Maron BJ, Pelliccia A. The heart of trained athletes: cardiac remodeling and the risks of sports, including sudden death. Circulation. 2006;114 (15):1633-44.
396. Krahn AD, Wilde AAM, Calkins H, La Gerche A, Cadrin-Tourigny J, Roberts JD, et al. Arrhythmogenic Right Ventricular Cardiomyopathy. JACC Clin Electrophysiol. 2022;8 (4):533-53.
397. Corrado D, Basso C. Arrhythmogenic left ventricular cardiomyopathy. Heart. 2022;108 (9):733-43.
398. Sen-Chowdhry S, Syrris P, Prasad SK, Hughes SE, Merrifield R, Ward D, et al. Left-dominant arrhythmogenic cardiomyopathy: an under-recognized clinical entity. J Am Coll Cardiol. 2008;52 (25):2175-87.
399. Graziosi M, Leone O, Foa A, Agostini V, Ditaranto R, Foroni M, et al. Postmortem diagnosis of left dominant arrhythmogenic cardiomyopathy: the importance of a multidisciplinary network for sudden death victims. "HIC mors gaudet succurere vitae". Cardiovasc Pathol. 2020;44:107157.
400. Piriou N, Marteau L, Kyndt F, Serfaty JM, Toquet C, Le Gloan L, et al. Familial screening in case of acute myocarditis reveals inherited arrhythmogenic left ventricular cardiomyopathies. ESC Heart Fail. 2020;7 (4):1520-33.
401. Tsuruta Y, Sueta D, Takashio S, Oda S, Sakamoto K, Kaikita K, et al. Left-dominant arrhythmogenic cardiomyopathy with a nonsense mutation in DSP. ESC Heart Fail. 2020.
402. Singh SM, Casey SA, Berg AA, Abdelhadi RH, Katsiyiannis WT, Bennett MK, et al. Autosomal-dominant biventricular arrhythmogenic cardiomyopathy in a large family with a novel in-frame DSP nonsense mutation. Am J Med Genet A. 2018;176 (7):1622-6.
403. Jimenez-Jaimez J, Lopez Moreno E, Barrio Lopez MT, Gonzalez-Molina M, Alvarez M, Tercedor L. A recessive inheritance pattern contributes to arrhythmogenic biventricular cardiomyopathy. Rev Esp Cardiol (Engl Ed). 2014;67 (9):772-4.
404. Akdis D, Saguner AM, Burri H, Medeiros-Domingo A, Matter CM, Ruschitzka F, et al. Clinical predictors of left ventricular involvement in arrhythmogenic right ventricular cardiomyopathy. Am Heart J. 2020;223:34-43.
405. AbdelWahab A, Gardner M, Parkash R, Gray C, Sapp J. Ventricular tachycardia ablation in arrhythmogenic right ventricular cardiomyopathy patients with TMEM43 gene mutations. J Cardiovasc Electrophysiol. 2018;29 (1):90-7.
406. Afzal MR, Evanson C, Cardona A, Rusk K, Weiss R. Arrhythmogenic Right Ventricular Cardiomyopathy. Encyclopedia of Cardiovascular Research and Medicine2018. p. 182-91.
407. Mayosi BM, Fish M, Shaboodien G, Mastantuono E, Kraus S, Wieland T, et al. Identification of Cadherin 2 (CDH2) Mutations in Arrhythmogenic Right Ventricular Cardiomyopathy. Circ Cardiovasc Genet. 2017;10 (2).
408. Paul M, Schulze-Bahr E. Arrhythmogenic right ventricular cardiomyopathy : Evolving from unique clinical features to a complex pathophysiological concept. Herz. 2020;45 (3):243-51.

409. Pilichou K, Thiene G, Bauce B, Rigato I, Lazzarini E, Migliore F, et al. Arrhythmogenic cardiomyopathy. Orphanet J Rare Dis. 2016;11 (1).
410. Wichter T, Paul TM, Eckardt L, Gerdes P, Kirchhof P, B�cker D, et al. Arrhythmogenic Right Ventricular Cardiomyopathy. Herz Kardiovaskul�re Erkrankungen. 2005;30 (2):91-101.
411. Qiu J, Hammond SL, Leek AN, Richman EH, Tjalkens RB. Cellular selectivity of AAV serotypes for gene delivery in neurons and astrocytes by neonatal intracerebroventricular injection. Plos One. 2017;12 (12).
412. Gerull B, Brodehl A. Insights Into Genetics and Pathophysiology of Arrhythmogenic Cardiomyopathy. Current Heart Failure Reports. 2021;18 (6):378-90.
413. van Lint FHM, Murray B, Tichnell C, Zwart R, Amat N, Lekanne Deprez RH, et al. Arrhythmogenic Right Ventricular Cardiomyopathy-Associated Desmosomal Variants Are Rarely De Novo. Circulation: Genomic and Precision Medicine. 2019;12 (8).
414. Brodehl A, Meshkov A, Myasnikov R, Kiseleva A, Kulikova O, Klauke B, et al. Hemi- and Homozygous Loss-of-Function Mutations in DSG2 (Desmoglein-2) Cause Recessive Arrhythmogenic Cardiomyopathy with an Early Onset. International Journal of Molecular Sciences. 2021;22 (7).
415. Groeneweg JA, Bhonsale A, James CA, te Riele AS, Dooijes D, Tichnell C, et al. Clinical Presentation, Long-Term Follow-Up, and Outcomes of 1001 Arrhythmogenic Right Ventricular Dysplasia/Cardiomyopathy Patients and Family Members. Circ Cardiovasc Genet. 2015;8 (3):437-46.
416. Watkins DA, Hendricks N, Shaboodien G, Mbele M, Parker M, Vezi BZ, et al. Clinical features, survival experience, and profile of plakophylin-2 gene mutations in participants of the Arrhythmogenic Right Ventricular Cardiomyopathy Registry of South Africa. Heart Rhythm. 2009;6 (11):S10-S7.
417. Carruth ED, Beer D, Alsaid A, Schwartz MLB, McMinn M, Kelly MA, et al. Clinical Findings and Diagnostic Yield of Arrhythmogenic Cardiomyopathy Through Genomic Screening of Pathogenic or Likely Pathogenic Desmosome Gene Variants. Circ-Genom Precis Me. 2021;14 (2):201-12.
418. Syrris P, Ward D, Asimaki A, Sen-Chowdhry S, Ebrahim HY, Evans A, et al. Clinical expression of plakophilin-2 mutations in familial arrhythmogenic right ventricular cardiomyopathy. Circulation. 2006;113 (3):356-64.
419. Carruth ED, Young W, Beer D, James CA, Calkins H, Jing L, et al. Prevalence and Electronic Health Record-Based Phenotype of Loss-of-Function Genetic Variants in Arrhythmogenic Right Ventricular Cardiomyopathy-Associated Genes. Circ Genom Precis Med. 2019;12 (11):e002579.
420. van Tintelen JP, Entius MM, Bhuiyan ZA, Jongbloed R, Wiesfeld AC, Wilde AA, et al. Plakophilin-2 mutations are the major determinant of familial arrhythmogenic right ventricular dysplasia/cardiomyopathy. Circulation. 2006;113 (13):1650-8.
421. Kelleher KJ, Sheils TK, Mathias SL, Yang JJ, Metzger Vincent T, Siramshetty Vishal B, et al. Pharos 2023: an integrated resource for the understudied human proteome. Nucleic Acids Research. 2023;51 (D1):D1405-D16.
422. Gaertner A, Bloebaum J, Brodehl A, Klauke B, Sielemann K, Kassner A, et al. The Combined Human Genotype of Truncating TTN and RBM20 Mutations Is Associated with Severe and Early Onset of Dilated Cardiomyopathy. Genes (Basel). 2021;12 (6).
423. Sielemann K, Elbeck Z, Gartner A, Brodehl A, Stanasiuk C, Fox H, et al. Distinct Myocardial Transcriptomic Profiles of Cardiomyopathies Stratified by the Mutant Genes. Genes (Basel). 2020;11 (12).
424. Kiselev A, Vaz R, Knyazeva A, Khudiakov A, Tarnovskaya S, Liu J, et al. De novo mutations in FLNC leading to early-onset restrictive cardiomyopathy and congenital myopathy. Hum Mutat. 2018;39 (9):1161-72.

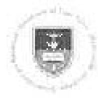
425. Yonemura S. Actin filament association at adherens junctions. J Med Invest. 2017;64 (1.2):14-9.
426. Peng X, Nelson ES, Maiers JL, DeMali KA. New insights into vinculin function and regulation. Int Rev Cell Mol Biol. 2011;287:191-231.
427. Yamagata K, Horie M, Aiba T, Ogawa S, Aizawa Y, Ohe T, et al. Genotype-Phenotype Correlation of SCN5A Mutation for the Clinical and Electrocardiographic Characteristics of Probands With Brugada Syndrome: A Japanese Multicenter Registry. Circulation. 2017;135 (23):2255-70.
428. Mirabel M, Damy T, Donal E, Huttin O, Labombarda F, Eicher JC, et al. Influence of centre expertise on the diagnosis and management of hypertrophic cardiomyopathy: A study from the French register of hypertrophic cardiomyopathy (REMY). Int J Cardiol. 2019;275:107-13.
429. Ho CY, Day SM, Ashley EA, Michels M, Pereira AC, Jacoby D, et al. Genotype and Lifetime Burden of Disease in Hypertrophic Cardiomyopathy: Insights from the Sarcomeric Human Cardiomyopathy Registry (SHaRe). Circulation. 2018;138 (14):1387-98.
430. van Velzen HG, Vriesendorp PA, Oldenburg RA, van Slegtenhorst MA, van der Velden J, Schinkel AFL, et al. Value of Genetic Testing for the Prediction of Long-Term Outcome in Patients With Hypertrophic Cardiomyopathy. Am J Cardiol. 2016;118 (6):881-7.
431. Costa S, Medeiros-Domingo A, Gasperetti A, Akdis D, Berger W, James CA, et al. Impact of Genetic Variant Reassessment on the Diagnosis of Arrhythmogenic Right Ventricular Cardiomyopathy Based on the 2010 Task Force Criteria. Circ Genom Precis Med. 2021;14 (1):e003047.
432. Chen L, Rao M, Chen X, Chen K, Ren J, Zhang N, et al. A founder homozygous DSG2 variant in East Asia results in ARVC with full penetrance and heart failure phenotype. Int J Cardiol. 2019;274:263-70.

Appendix

APPENDIX A- INCLUSION AND EXCLUSION CRITERIA

IMHOTEP study	
INCLUSION CRITERIA	**EXCLUSION CRITERIA**
Patients reside in South Africa • Cases diagnosed at autopsy are included • Idiopathic cardiomyopathies • Familial cardiomyopathies • Neuromuscular disorders with cardiac involvement • Non-familial or secondary causes of cardiomyopathy o Myocarditis (infective/toxin/immune) o Human Immunodeficiency Virus (HIV) o Drugs/Toxins o Pregnancy o Endocrine o Nutritional o Obesity o Alcohol o Tachycardiomyopathy o Eosinophilic o Kawasaki disease o Tako Tsubo cardiomyopathy o Amyloidosis o Autoimmune o Endomyocardial fibrosis o Carcinoid heart disease o Radiation o Metastatic Cancer	• Systemic arterial hypertension (>160/100mmHg documented and confirmed at repeated measures) • Coronary artery disease (obstruction >50% of the luminal diameter of a major branch) • Pericardial diseases • Congenital heart disease • Pulmonary disease with cor pulmonale • Valvular heart disease, including Rheumatic Heart Disease

APPENDIX B 1-HREC 009/2020



UNIVERSITY OF CAPE TOWN
Faculty of Health Sciences
Human Research Ethics Committee

Room G 50 Old Main Building
Groote Schuur Hospital
Observatory 7925
Email: hrec-enquiries@uct.ac.za
Website: www.health.uct.ac.za/fhs/research/humanethics/forms

07 February 2020

HREC REF: 009/2020

Prof N. Ntusi
Department of Medicine
J 46.53 OMB GSH

Dear Prof Ntusi

PROJECT TITLE: TARGETED RE-SEQUENCING OF A LARGE SOUTH AFRICA CARDIOMYOPATHY COHORT (PHD DEGREE - MR POLYCARP NDIBANGWI)

Thank you for submitting your new study to the Faculty of Health Sciences Human Research Ethics Committee (HREC) for review.

It is a pleasure to inform you that the HREC has **formally approved** the above-mentioned study subject.

Approval is granted for one year until the 28 February 2021.

Please submit a progress form, using the standardised Annual Report Form if the study continues beyond the approval period. Please submit a Standard Closure form if the study is completed within the approval period. (Forms can be found on our website: www.health.uct.ac.za/fhs/research/humanethics/forms)

The HREC acknowledges that the student: Mr Polycarp Ndibangwi will also be involved in this study.

Please note that for all studies approved by the HREC, the principal investigator **must** obtain appropriate institutional approval, where necessary, before the research may occur.

Please note that the ongoing ethical conduct of the study remains the responsibility of the principal investigator.

Please quote the HREC REF in all your correspondence

Yours sincerely

Signed by candidate

PROFESSOR M BLOCKMAN
CHAIRPERSON, FHS HUMAN RESEARCH ETHICS COMMITTEE
Federal Wide Assurance Number: FWA00001637.
Institutional Review Board (IRB) number: IRB00001938
NHREC-registration number: REC-210208-007

HREC Ref 009/2020
OL

APPENDIX B 2-HREC 009/2020 Renewal

FHS016: Annual Progress Report / Renewal

HREC office use only (FWA00001637; IRB00001938)
This serves as notification of annual approval, including any documentation described below.

☑ Approved	Annual progress report	Approved until /next renewal date	28/02/2024
☐ Not approved	See attached comments		
Signature Chairperson of the HREC/ Designee	Signed by candidate	Date Signed	13/2/2023

Note: Please email this form and supporting documents (if applicable) in a combined pdf file to hrec-enquiries@uct.ac.za.
Please clarify your plan for research-related activities during COVID-19 lockdown.
Please use the latest form found on our website.
http://www.health.uct.ac.za/fhs/research/humanethics/forms

HUMAN RESEARCH ETHICS COMMITTEE 13 FEB 2023 HEALTH SCIENCES FACULTY UNIVERSITY OF CAPE TOWN

Comments to PI from the HREC

Principal Investigator to complete the following:

1. Protocol information

Date (when submitting this form)	09 February 2023		
HREC REF Number	009/2020	Current Ethics Approval was granted until	28/02/2023
Protocol title	Targeted Resequencing of a Large South African Cardiomyopathy Cohort (Sub-study linked to HREC 788/2014)		
Protocol number (if applicable)			
Are there any sub studies linked to this study?	☐ Yes	✓ No	
If yes, could you please provide the HREC Reference number for all sub-studies? Note: A separate FHS016 must be submitted for each sub-study.			
Principal Investigator			

28 February 2022 Page 1 of 7 FHS016

(Note: Please complete the Closure form (FHS010) if the study is completed within the approval period)

APPENDIX C - Puregene Blood Core Kit C (Qiagen)

Purpose: DNA extraction protocols.

Gentra Puregene Blood Kit protocol – DNA purification from buffy coat

Note: exceptions to the protocols are written in bold

1. If the buffy coat preparation contains red blood cells, continue with step 2. Otherwise, pipet 3mL Cell Lysis Solution into a 15mLcentrifuge tube, add 150–250µL sample and continue with step 8.

2. Dispense 3 volumes of RBC Lysis Solution into a 15mL centrifuge tube (e.g., if processing 250µL buffy coat, dispense 750µL RBC Lysis Solution). Add 150–250µL buffy coat preparation.

3. Invert to mix and incubate for 10 min at room temperature (15–25°C). Invert again at least once during the incubation.

4. Centrifuge for 5 min at 2000 x g. (**Samples were centrifuged for 10 min**)

5. Carefully discard the supernatant by pipetting or pouring, leaving approximately 100–200µL of the residual liquid and the pellet.

6. Vortex the tube vigorously to resuspend the pellet in the residual liquid.

7. Add 3mL Cell Lysis Solution and pipet up and down or vortex vigorously to lyse the cells. Usually, no incubation is required; however, if cell clumps are visible, incubate at 37°C until the solution is homogeneous. Samples are stable in Cell Lysis Solution for at least two years at room temperature.

8. Optional: If RNA-free DNA is required, add 15µL RNase A Solution and mix by inverting 25 times. Incubate for 15 min at 37°C. Then, incubate the sample for 3 min on ice to quickly cool it. (**Not done**)

9. Add 1mL Protein Precipitation Solution and vortex vigorously for 20 s at high speed. (**1.5mL Protein Precipitation Solution was added**)

10. Centrifuge for 5 min at 2000 x g. (**Samples were centrifuged for 10 min**)

11. Pipette 3mL isopropanol into a clean 15mL centrifuge tube and add the supernatant from the previous step by pouring it carefully. (**6mL isopropanol was used**)

12. Mix by inverting gently 50 times. (**Samples were inverted until DNA was visible**)

13. Centrifuge for 3 min at 2000 x g. (**the visible DNA was removed by pipette and centrifuged for 1 min in a new 2mL microcentrifuge tube**)

14. Carefully discard the supernatant and drain the tube by inverting it on a clean piece of absorbent paper, taking care that the pellet remains in the tube.

15. Add 3mL of 70% ethanol and invert several times to wash the DNA pellet. (**1mL of 70% ethanol was added**)

16. Centrifuge for 1 min at 2000 x g.

17. Carefully discard the supernatant. Drain the tube on a clean piece of absorbent paper, taking care that the pellet remains in the tube. Allow to air dry for 5–10 min.

18. Add 300µL DNA Hydration Solution and vortex for 5 s at medium speed to mix. (**200 µl DNA Hydration Solution was added**)

19. Incubate at 65°C for 1hr to dissolve the DNA. (**Not done**)

20. Incubate at room temperature overnight.

APPENDIX D - PAXgene® Blood DNA Kit protocol

Before starting:

- Thaw frozen PAXgene Blood DNA Tubes in a wire rack at ambient temperature (18–25°C) for approximately 2hr or at 37°C in a water bath for approximately 15 minutes. Carefully invert the thawed PAXgene Blood DNA Tubes 10 times before beginning the procedure.
- Heat a heating block or water bath to 65°C for use in steps 8 and 17.
- Add 1.4mL Buffer BG4 (resuspension buffer) to lyophilised PreAnalytiX Protease. Dissolved PreAnalytiX Protease should be stored at 2–8°C or in aliquots at –20°C.
- For every sample, mix 5mL Buffer BG3 (digestion buffer), and 50µL reconstituted PreAnalytiX Protease. For example, to process ten samples, mix 50mL Buffer BG3 with 500µL PreAnalytiX Protease. The Buffer BG3–PreAnalytiX Protease mixture should be prepared immediately before the start of the procedure.

Procedure:

1. Pour all the blood from one PAXgene Blood DNA Tube into a Processing Tube containing 25mL Buffer BG1. Close the tube. To avoid cracking the blue lids of the Processing Tubes, do not overtighten them. Tighten the lid only until the first sign of resistance is felt. Mix by inverting the tube five times.
2. Centrifuge for 5 min at 2500 x g in a swing-out rotor.
3. Carefully discard the supernatant and place the tube in a rack.
4. Add 5mL Buffer BG2, close the tube and wash the pellet by vortexing vigorously for 5s.
5. Centrifuge for 3 min at 2500 x g in a swing-out rotor.
6. Carefully discard the supernatant and place the tube back in the rack.
7. Add 5mL Buffer BG3/PreAnalytiX Protease, close the tube and vortex for 20s at high speed.
8. Place the tube in a heating block or water bath and incubate at 65°C for 10 min.
9. Vortex again for 5s at high speed.
10. Add 5mL isopropanol (100%) and mix by inverting the tube at least 20 times until the white DNA strands clump visibly together.
11. Centrifuge for 3 min at 2500 x g.
12. Discard the supernatant and leave the tube inverted on a clean piece of absorbent paper for 1 min.
13. Add 5mL 70% (v/v) ethanol and vortex for 1s at high speed.
14. Centrifuge for 3 min at 2500 x g.
15. Discard the supernatant and leave the tube inverted on a clean piece of absorbent paper for at least 5 min.
16. Carefully dab the tube onto absorbent paper to remove ethanol from the rim and leave it inverted for a further 5 min to allow the DNA pellet to dry.
17. Add 1mL Buffer BG4 and dissolve the DNA by incubating for 1hr at 65°C in a heating block or water bath, followed by incubation overnight at room temperature.

APPENDIX E – Buffers and reagents: Agarose gel electrophoresis

Tris/Borate/EDTA (TBE) buffers

10X TBE (stock):

- 121.1g Tris (Glentham Life Sciences) (final concentration: 1 M)
- 61.8g Boric acid (AMRESCO, Solon, OH USA) (final concentration: 1 M)
- 7.4g EDTA (Glentham Life Sciences) (final concentration: 0.02 M)
- Made to a final volume of 1 L with sterile distilled water

1X TBE (working):

- Made by a 1:10 dilution of stock 10X TBE buffer with sterile distilled water (Milli-Q)

1 M Tris-HCl (pH 8.0):

- 121.1g Tris base (Glentham Life Sciences, Corsham, UK) in 800 ml distilled water
- Adjust the pH to 8.0 with concentrated HCl
- Make to a total volume of 1L with sterile distilled water

0.5 M EDTA (pH 8.0):

- 186.1g EDTA (Glentham Life Sciences) in 800 ml distilled water
- Adjust the pH to 8.0 with NaOH
- Make to a total volume of 1L with sterile distilled water

1X Tris/EDTA (TE) buffer:

- 1mL 1 M Tris-HCl (final concentration: 10 mM)
- 0.2mL 0.5 M EDTA (final concentration: 1 mM)
- 98.8mL sterile distilled water

Agarose gels

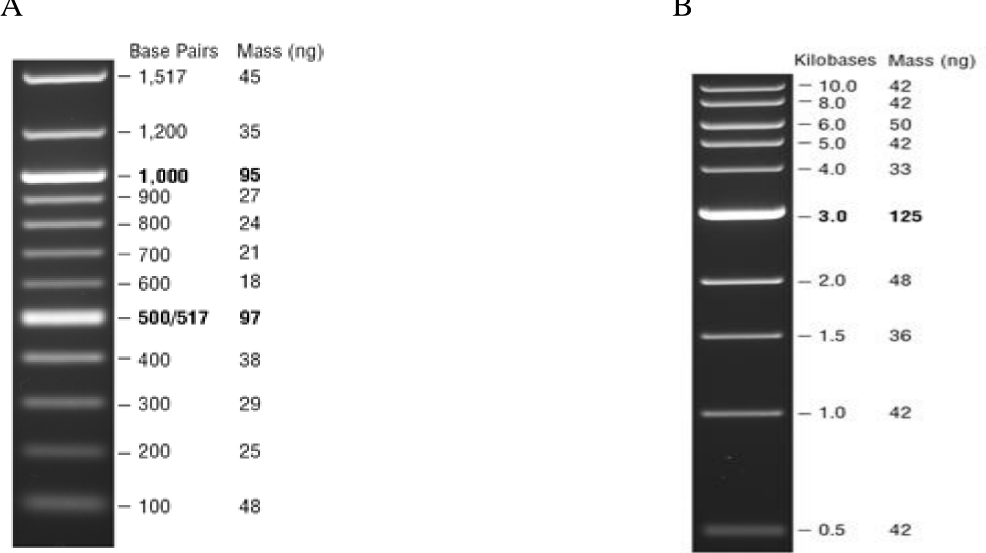
Loading dye:

- Loading dye mix:
    - Add 5µL GelRed® nucleic acid gel stain (Biotium),
    - To 1mL 5X Green GoTaq® Flexi buffer (Promega),
    - Mix gently before use
- Mixed with DNA samples at a ratio of 5:3 (dye: sample) before loading gel

1% agarose gel:

- 1.5g SeaKem® LE agarose (Lonza, Basel, Switzerland)
- 150mL 1X TBE

<u>DNA molecular weight markers</u>



DNA molecular weight markers. Shown are (A) the 100 bp ladder (NEB) and (B) the Quick-load® 1kb DNA ladder (NEB). Adapted from New England ® 1kb DNA ladder (NEB). Adapted from New England BioLabs product.

APPENDIX F –The targeted panel genes list

α-actin (*ACTC1*)

α-actinin 2 (*ACTN2*)

ankyrin repeat domain-containing protein 1 (*ANKRD1*)

BCL2-associated athanogene 3 (*BAG3*)

α-crystallin B chain (*CRYAB*)

cysteine and glycine-rich protein 3 (*CSRP3*)

desmin (*DES*)

dystrophin (*DMD*)

desmocollin 2 (*DSC2*)

desmoglein 2 (*DSG2*)

desmoplakin (*DSP*)

four-and-a-half LIM domains 1 (*FHL1*)

four-and-a-half LIM domains 2 (*FHL2*)

filamin C (*FLNC*)

α-galactosidase (*GLA*)

junction plakoglobin (*JUP*)

lysosome-associated membrane protein 2 (*LAMP2*)

lamin A/C (*LMNA*)

myosin-binding protein C (*MYBPC3*)

β-myosin heavy chain (*MYH7*)

regulatory light chain of myosin (*MYL2*)

essential light chain of myosin (*MYL3*)

plakophilin 2 (*PKP2*)

phospholamban (*PLN*)

AMP-activated protein kinase (*PRKAG2*)

RNA-binding motif protein 20 (*RBM20*)

sodium channel protein type 5 subunit alpha (*SCN5A*)

tafazzin (*TAZ*)

transmembrane protein 43 (*TMEM43*)

troponin C (*TNNC1*)

troponin I type 3 (*TNNI3*)

troponin T type 2 (*TNNT2*)

tropomyosin 1 (*TPM1*)

titin (*TTN*)

transthyretin (*TTR*)

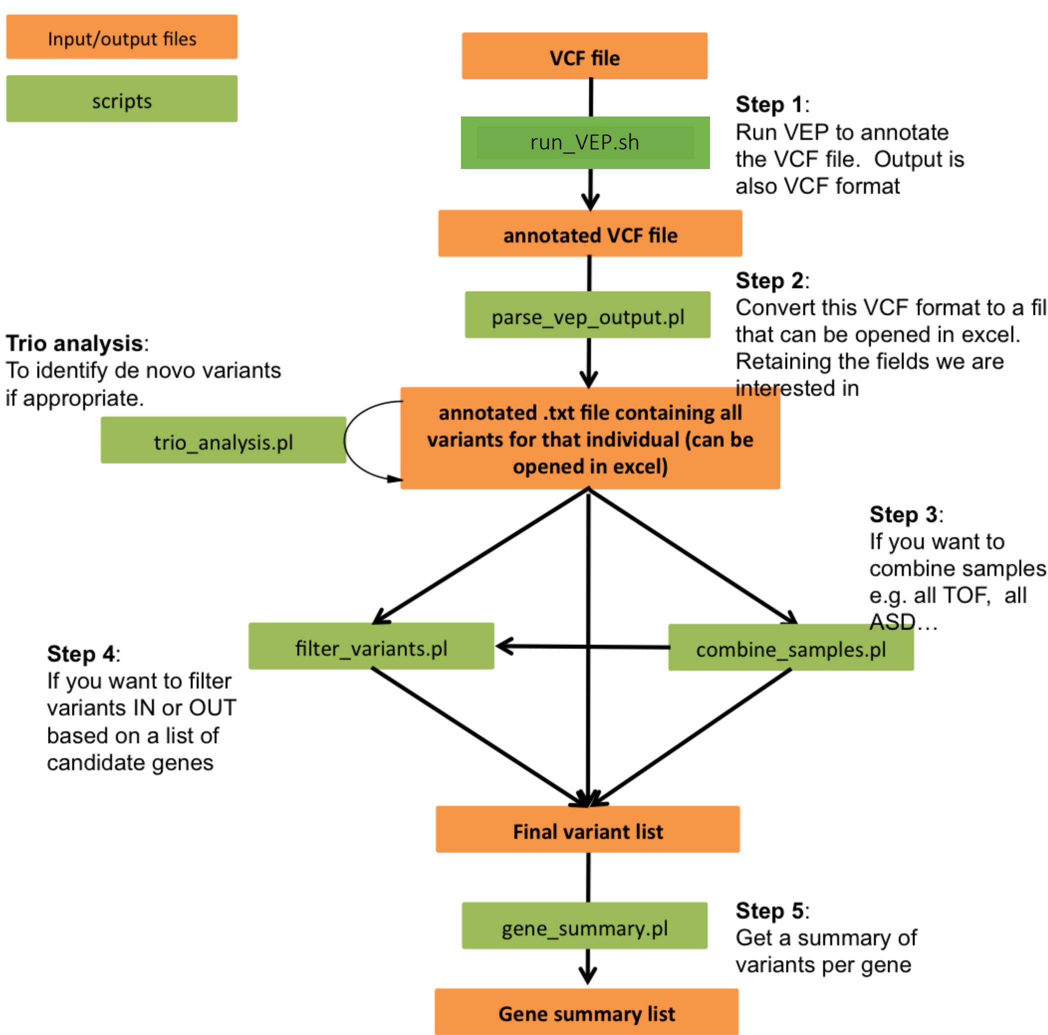
vinculin (*VCL*)

APPENDIX G – VCF file annotation (GRCh37) and analysis pipeline on UCT HPC cluster

This document describes how to go from GRCh37 VCF files to annotated files containing variants that can be from merged samples and/or filtered on gene candidate lists.

Please note: these instructions are for running on the UCT HPC cluster.

Schematic representation of the annotation and filtering pipeline:



Step 1:

The first stage of the process is to get the VCF files for each individual. These will typically be in gzipped format with the suffix '**.vcf.gz**'.

The script to run VEP on the UCT cluster is **/opt/exp_soft/vep/run_VEP.sh**.

You should take a copy of this script and copy it to your own home directory or wherever you wish to run the analysis.

Once copied to your local directory, run this on a single VCF file:

'ls

This will produce 'file1.annotated.vcf'

There are many options that can be set in the run_VEP.sh script. A full list of VEP parameters can be found here: https://www.ensembl.org/info/docs/tools/vep/script/vep_options.html

A general note on running the subsequent scripts on the UCT cluster

The following scripts can be run on the UCT cluster by submitting them to the compute nodes. This is done by wrapping the commands in a shellscript. The documentation for this can be found here: http://hpc.uct.ac.za/index.php/hpc-cluster/

Step 1 – running VEP – this has already been put into the script /opt/exp_soft/vep/run_VEP.sh.

For subsequent steps, you should put these commands into a script yourself, depending on what you want to do. For example, to run steps 2 and 4 on a single sample, you might have these commands in a shellscript called 'run_analysis.sh':

```
#!/bin/sh

#SBATCH --account=pathologycd

#SBATCH --partition=ada

#SBATCH --nodes=1 --ntasks=2

#SBATCH --time=3:00:00

#SBATCH --job-name="vep"

perl /opt/exp_soft/vep/scripts/parse_vep_output.pl ./file.annotated.vcf

perl /opt/exp_soft/vep/scripts/filter_variants.pl –sample ./file.annotated.txt –gene_list /path/to/TOF_gene_list.txt
```

To run it, type:

```
sbatch run_analysis.sh
```

To check the progress of your job, type 'squeue –u <username>'. Once it disappears from this list, it is finished (or potentially failed). You will also get the output file slurm-*12345*.out (depending on your job number). It can sometimes be useful to look here if something has gone wrong.

Step 2:

To convert this VCF formatted output file into a file that can be read by Excel (Again, this command should be put into a script for correct running on the cluster):

```
perl /opt/exp_soft/vep/scripts/perlparse_vep_output.pl file.annotated.vcf
```

Step 3:

If we repeat steps 1 and 2 for each of our VCF files, we may then wish to combine samples with similar phenotypes for a single analysis. To combine the variants of multiple samples, you first have to make a list of the files you want to combine.

The following command makes the list file:

ls file1.annotated.txt file2.annotated.txt file3.annotated.txt > input_list

The script can then be run:

perl combine_samples.pl –sample_list 145_list > 145_samples.annotated.txt

Step 4:

At this stage, you will have single or combined variant files. You may choose to filter them based on candidate gene lists.

perl /opt/exp_soft/vep/scripts/filter_variants.pl –sample file1.annotated.txt –gene_list TOF_gene_list.txt

This will only return variants in the gene list provided. If you want to return all variants **except** for the genes on the list (i.e. previously unknown genes), you can run it with an extra flag:

perl /opt/exp_soft/vep/scripts/filter_variants.pl –sample file1.annotated.txt –gene_list TOF_gene_list.txt –exclude

The output file will be called 'file1.annotated.filtered.txt' and can be opened in Excel.

Step 5:

If you want a gene-based summary of the variants:

perl /opt/exp_soft/vep/scripts/gene_summary.pl –input file.annotated.txt > file1.gene_summary.txt

Trio analysis:

This is an optional step if you have trio files and you want to look for de novo or homozygous recessive variants. It should be run after stage 2 has been run for each of the trio samples, and you have annotated.txt files for each.

perl /opt/exp_soft/vep/scripts/trio_analysis.pl –proband file_proband.annotated.txt –parent1 file_parent1.annotated.txt –parent2 file_parent2.annotated.vcf > file_proband_trio_analysis_annotated.txt

The first column of the output indicates if it is 'DENOVO' or 'HOMO_RECESSIVE', and the final two columns show the variant in the parents – obviously empty if de novo.

APPENDIX H – CVG SOP: Primer design

Primers are short oligonucleotides that are used in PCR to amplify target regions of DNA. For our purposes, PCR is typically used to validate the results of next-generation sequencing, perform segregation analysis in families, or screen candidate genes for pathogenic variants. When validating or performing segregation analysis, the target of PCR is usually a candidate pathogenic variant, and when screening a gene, the target is an entire exon.

Note: *Primer design uses the output from gene annotation (or the region of interest from the transcript in the Ensembl genebank). Make sure you have a recent annotation before starting the primer design.*

1. Find your target region in the annotation (as explained above, this may be an exon or a variant). Exons are labelled in the annotation, and variants can be found using the rs numbers in the explanatory lines.

2. Scan the genetic sequence ~40-100 bp on either side of the intended target for possible primer sequences. It is important to leave enough space on either side of the target region for a clear sequencing signal (the first ~20-40 bp of an electropherogram is often noisy). You may copy this sequence into a new Word document for work on

3. Look for primers that meet the criteria below:
   a. **Length:** 20-23 bp (can be 18-25 bp)
   At this length, primers are long enough to be specific but will still bind DNA at typical PCR annealing temperatures.

   b. **Melting temperature:** 55-60°C (can be 50-60°C)
   Melting temperature is defined as the temperature at which half of the DNA duplex is dissociated. During PCR, annealing typically occurs at 50-60°C. A primer with a melting temperature too high will be unlikely to bind during normal PCR annealing. If it is too low, the primer will be more likely to misprime

   c. **Melting temperature difference between both primers:** < 3°C (ideally less)
   If the melting temperatures are too disparate, the primers may not anneal at the same temperature in the reaction. It is crucial that both forward and reverse primers can anneal to DNA at the same temperature

   d. **GC content:** 45-65%
   G-C bonds contribute to the stable binding of primer to the DNA template. A GC content of 45-65% confers the optimum level of stability and usually falls within the desired melting temperature range. When the GC content is too low, this will affect the primer annealing temperatures and primer binding efficiency. When it's too high, it can increase the chances of secondary structure formation.

   e. **GC clamp:** 3` end
   Including a G or C base in the last three bases of the primer sequence helps promote specific binding at the 3` end of the primer, from which Taq polymerase incorporates dNTPs in chain elongation.

f. **A good sequence diversity with minimal sequence repeats**
A primer sequence that is too repetitive (e.g. ATATATAT) or with long runs of a single base (e.g. GGGGG) increases the chances of mispriming and primer non-specificity

g. **Minimal secondary structures**
Interactions between primers can reduce the amount of primer that is available to bind to your target region, reducing the yield of the reaction.

h. **PCR product size:** 50-300 bp (for HRM)
The amplicon length depends on the experimental purpose. For HRM, smaller products improve resolution and the ability to visualise genetic variation. For Sanger sequencing, longer products (< 800 bp) are acceptable.

4. Check each primer in OligoAnalyser (https://eu.idtdna.com/pages; you may need to make an account the first time). Paste each potential primer into the search box and check the parameters above as well as secondary structures (you only need to check the topmost structure). When checking the secondary structures, the ΔG measures the free energy of formation of that structure, i.e. the likelihood that the structure will form (lower values indicate increased likelihood):
   a. **Hairpin:** ΔG > -1; T_m < 50°C; fewer than 4 bonds
   A hairpin is formed by interactions within the primer; this should be avoided. Generally, a hairpin with a ΔG of -1 kcal/mol can be tolerated. The T_m will indicate if the structure is likely to melt under normal PCR temperatures and should be lower than the annealing temperature of the reaction. The more bonds are involved, the stronger the structure is likely to be

   b. **Self-dimer:** ΔG > -3.5; avoid 3` overhangs
   A primer self-dimer is formed by two primers of the same strand (e.g. two forward primers) interacting due to regions of homology; this should be avoided. Primer dimers will reduce the product yield. Generally, self-dimers with a ΔG of -3.5 kcal/mol can be tolerated. Avoid 3` overhangs, as Taq polymerase can bind to and extend these primer-dimers

   c. **Heterodimer:** ΔG > -3.5; avoid 3` overhangs
   A heterodimer is formed by interaction between sense and antisense primers due to regions of homology; this should be avoided. Primer dimers will reduce the product yield. Generally, heterodimers with a ΔG of -3.5 kcal/mol can be tolerated. Avoid 3` overhangs, as Taq polymerase can bind to and extend these primer-dimers

5. Check the specificity of each primer (to target your intended gene only) using BLASTn (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE_TYPE=BlastSearch). Paste each primer into the search box and search using the following conditions:
   a. **Organism:** Homo sapiens
   b. **Exclude:** models (XM/XP), uncultured/environmental sample sequences (check these)
   c. **Leave the rest as default**

6. Check the BLAST output; your genetic target should be a top hit. Check:
    a. **The query coverage:** should be 100%
    This is how much of your primer has a genetic match in the NCBI database; 100% of your primer should match the gene of interest.

    b. **The % identity:** should be 100%
    This is how much of the BLAST 'hit' matches your query sequence; this should be 100% because there should be no mismatches between your primer and the gene in the primer sequence.

    c. **The E-value:** lower scores indicate stronger hits
    This is the number of hits of a similar quality that may be expected to occur just by chance. As with statistics, the smaller the value, the more probable it is that your hit is not occurring by chance.

7. Optional: Check the POLAND predicted melt profile for the PCR amplicon (http://www.biophys.uni-duesseldorf.de/html/local/POLAND/poland.html):
    a. **The temperature of 50% probability vs sequence**
    This graph indicates the potential melting behaviour across your amplicon, allowing you to check regions that are predicted to melt easily and regions that may be more difficult to melt.

    b. **Differentiated hypochromicity at 260 and 282 nm (dA/dT) vs temperature**
    This indicates the potential HRM melt profile that your amplicon would generate

APPENDIX I- Reagents master mix 1x reaction

Reagent (stock)	**Final concentration/volume per reaction**
Forward primer (20µM)	0.8µM/0.5µL
Reverse primer (20µM)	0.8µM/0.5µL
dNTPs (20µM)	0.8µM/1µL
GoTaq Flexi buffer (5X)	1X/5 µL
GoTaq Polymerase (5U/ µL)	0.5U/0.1 µL
$MgCl_2$ (25mM)	1.5mM/3 µL
Evagreen (20X)	0.5X/0.5 µL
DNA	120ng/1 µL
dH20	13.4 µL
TOTAL volume	25 µL

CONDITION	**TEMPERATURE (TIME)**
Initial denaturation	95°C – 10 seconds
Denaturation Primer Annealing Template Elongation	95°C – 5 seconds 55°C – 10 seconds 50 cycles 72°C – 10 seconds
High-Resolution Melt	72 – 95°C (0.1°C increments)

Purification of the HRM product

REAGENT	QUANTITY/VOLUME PER REACTION
*Exonuclease I (*New England Biolabs)	1 U / 0.1 µL
*Shrimp Alkaline Phosphatase (*Promega)	2 U / 1 µL
HRM Product	5 µL
FINAL REACTION VOLUME	20 µL

CONDITION	TEMPERATURE (TIME)
Incubation	37°C – 1 hour
Deactivation	75°C – 15 min

Big-dye v3.1

REAGENT	FINAL CONCENTRATION / VOLUME PER REACTION
Forward Primer (20 µM)	2 mM / 2 µL
HRM product	3 µL
BigDye ® Terminator v3.1. Ready Reaction Mix (Applied Biosystems)	2 µL
5 X Sequencing Buffer (Applied Biosystems)	1 X
FINAL REACTION VOLUME	20 µL

CONDITION	TEMPERATURE (TIME)
Initial denaturation	96°C – 5 minutes
Denaturation Primer Annealing Template Elongation	96°C – 30 seconds 50°C – 15 seconds 25 cycles 60°C – 4 minutes

APPENDIX J- Recruitment sites

Recruitment sites (institutional HREC approval)	Groote Schuur Hospital (UCT), Red Cross War Memorial Children's Hospital (UCT), Tygerberg (SUN), Port Elizabeth (WSU), Mthatha (WSU), Bloemfontein (UFS), Mozambique

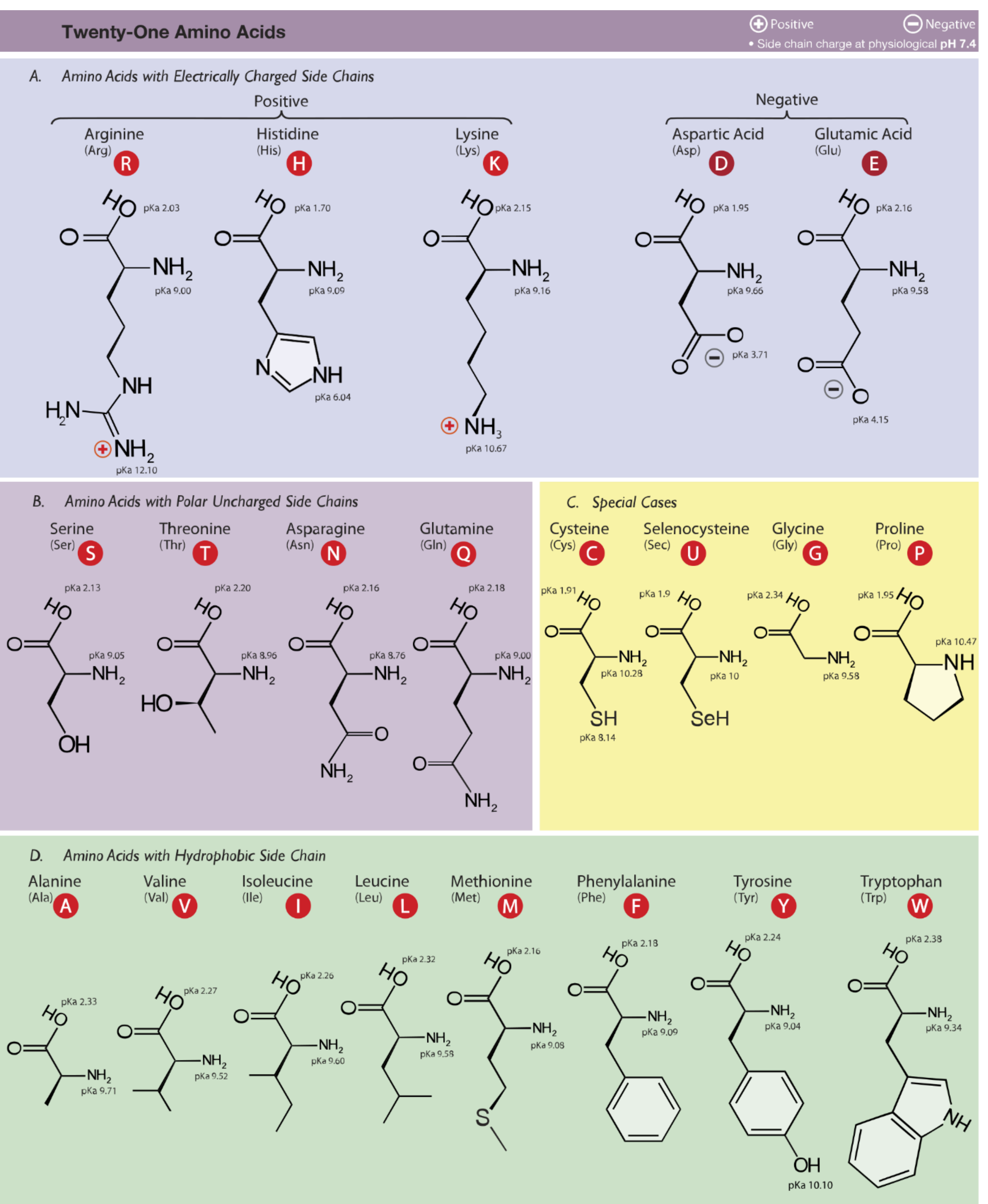
APPENDIX K – List of amino acids with abbreviations and classifications

Adapted from Zhang J. 2000. J Mol Evol, 50 (1):57. When assessing missense variation, amino acid replacements within a category are conservative, and between categories are considered radical, as adapted from Dagan T et al. 2002. Mol Evol Biol, 19 (7):1023.

Amino acid	Abbreviation	Charge	Volume and polarity
Alanine	A (Ala)	Neutral	Neutral, small
Arginine	R (Arg)	Positive	Polar, relatively large
Asparagine	N (Asn)	Neutral	Polar, relatively small
Aspartic acid	D (Asp)	Negative	Polar, relatively small
Cysteine	C (Cys)	Neutral	Special
Glutamic acid	E (Glu)	Negative	Polar, relatively small
Glutamine	Q (Gln)	Neutral	Polar, relatively small
Glycine	G (Gly)	Neutral	Neutral, small
Histidine	H (His)	Positive	Polar, relatively large
Isoleucine	I (Ile)	Neutral	Nonpolar, relatively small
Leucine	L (Leu)	Neutral	Nonpolar, relatively small
Lysine	K (Lys)	Positive	Polar, relatively large
Methionine	M (Met)	Neutral	Nonpolar, relatively small
Phenylalanine	F (Phe)	Neutral	Nonpolar, relatively large
Proline	P (Pro)	Neutral	Neutral, small
Serine	S (Ser)	Neutral	Neutral, small
Threonine	T (Thr)	Neutral	Neutral, small
Tryptophan	W (Trp)	Neutral	Nonpolar, relatively large
Tyrosine	Y (Tyr)	Neutral	Nonpolar, relatively large
Valine	V (Val)	Neutral	Nonpolar, relatively small

pK$_a$ Values for Common Alpha Amino Acids				
Amino Acid Type	**Amino Acid**	**α-COOH**	**α-NH_3^+**	**RH or RH^+**
Hydrophobic: Aliphatic	Glycine	2.34	9.60	
	Alanine	2.34	9.69	
	Valine	2.32	9.62	
	Leucine	2.36	9.68	
	Isoleucine	2.36	9.68	
	Proline	1.99	10.6	
	Methionine	2.28	9.21	
Hydrophobic: Aromatic	Phenylalanine	1.83	9.13	
	Tyrosine	2.2	9.11	10.07
	Tryptophan	2.38	9.39	
Hydrophilic: Polar Uncharged	Serine	2.21	9.15	
	Threonine	2.63	10.43	
	Cysteine	1.71	10.78	8.33
	Asparagine	2.02	8.8	
	Glutamine	2.17	9.13	
Hydrophilic: Acidic	Aspartic Acid	2.09	9.82	3.86
	Glutamic Acid	2.19	9.67	4.25
Hydrophilic: Basic	Arginine	2.17	9.04	12.48
	Histidine	1.82	9.17	6.00
	Lysine	2.18	8.95	10.53

https://wou.edu/chemistry/courses/online-chemistry-textbooks/ch450-and-ch451-biochemistry-defining-life-at-the-molecular-level/chapter-2-protein-structure/



https://wou.edu/chemistry/courses/online-chemistry-textbooks/ch450-and-ch451-biochemistry-defining-life-at-the-molecular-level/chapter-2-protein-structure/

APPENDIX L – IMHOTEP Phase I Variant Filtering SOP

Written by: P Ndibangwi

Variant investigation is an important step in the analysis of next-generation sequencing data. It involves the exclusion of most detected variants based on several criteria. For our purposes, many of these criteria are standard and should not change. The aim of filtering is usually to create a list of candidate variants for diagnostics and further analysis (e.g., segregation analysis or population screening).

Note: *Variant filtering uses the output files from the Variant Effect Predictor (VEP). Make sure you have these output files before filtering.*

1. Open the VEP text files using Microsoft Excel. These will typically include a filtered VEP file (against known disease genes, for example) or the total list of variants. Make sure that the gene symbol, exon, and intron columns are set to import using the text data format.

2. Turn on Excel's filter function (Data > Filter)

3. Filter the columns based on the following criteria:
   a. **Consequence:** select all frameshift, missense, nonsense, INDEL, splice donor/acceptor, transcript ablation/amplification and protein-altering variants. It is usually advisable to include Inframe insertions and deletions. *Applying this filter will exclude variants that are not in the exonic regions. It will also exclude variants unlikely to affect protein function, i.e., synonymous variants.*
   b. **gnomAD (ALL):** select MAF < 0.01 across ALL the gnomAD columns.
   This will exclude variants that are too common to cause disease. Make sure you include variants that do not have a recorded MAF (marked with a '–' in the output Excel sheet), as these may be rare.
   c. **PolyPhen2 and SIFT:** These are bioinformatic tools that score changes in amino acids (missense mutations) based on factors such as the chemical change in the amino acid and how well-conserved it is. For PolyPhen2, scores closer to 1 are predicted to be more likely to be disease-causing. For SIFT, scores closer to 0 are predicted to be more likely to be disease-causing. Note that only missense variants are scored, so absent values may indicate a truncating (frameshift or nonsense) variant; these MUST be included.

4. After or during the column filtering, you may want to copy the filtered data to a new Excel sheet or workbook for further analysis. You can then colour code variants; this will allow you to go back and check your filtering based on the following:
   a. **ACMG guidelines:** apply the ACMG five-tier classification rules or search the variants in online databases with built-in functions to classify variants according

to the recent ACMG five-tier classification system. Currently, common databases with this tool include Alamut Visual Plus (licenced), VarSome, ClinVar, CardioClassifier, and MASTERMIND. For example, Varsome (https://varsome.com/) looks at the predicted pathogenicity and assigns the ACMG rules. These human genomics online communities integrate many databases and use them to predict a variant's pathogenicity. On the home page, follow the on-screen prompts to subscribe, submit your queries and manually record the outputs.

b. **Literature Search:** Search the variants in all the National Center for Biotechnology Information databases (https://www.ncbi.nlm.nih.gov/) for published results to determine if the variant has been reported or not. Record the evidence as reported or novel.

c. **Novel variants:** For novel variants, **do an intensive search using current and relevant databases,** starting from checking the variant in MutationTaster (http://www.mutationtaster.org/), looking at the predicted disease-causing potential, frequency in gnomAD and 1000 Genomes, amino acid conservation and for missense variants also check in M-CAP (http://bejerano.stanford.edu/mcap/). With all the tools, follow the on-screen prompts to submit your queries and manually record the outputs.

d. **BAM files analysis:** To avoid validating artefacts, load the bam file of each sample with a (likely) disease-causing variant into a desktop bam file visualisation tool. Use the Integrative Genomics Viewer (download and install the software from https://software.broadinstitute.org/software/igv/home) or Alamut Visual Plus (licenced).

5. Compile a separate candidate list of reported and novel variants based on the abovementioned criteria. Priority is given to class 5 and 4 variants.

6. After filtering and compiling your list of candidate variants, the next step will be to validate these variants of interest. This involves PCR amplification followed by Sanger sequencing to determine whether they are real variants or false positives. Those variants that are real may then be screened in additional family members using similar techniques. These laboratory steps will require primers for PCR amplification and sequencing; refer to the gene annotation and primer design SOPs.

Supplementary Tables

SS Table 1: Table with ACMG/AMP rules for probands with class 4 and 5 variants

Probands with class 4 and 5 variants

ID	Sex	Ancestry	Age	NYHA	LVEF	Gene	Protein change	cDNA change	Consequences	ACMG	ACMG Abbreviation
ACM											
10100091	M	WA	19	Class I	55	*ALPK3*	p.Gln901Ter	c.2701C>T	stop_gained	Class 4	PVS1 Very Strong, PM2 Moderate
10100091	M	WA	19	Class I	55	*PKP2*	p.His733ProfsTer8	c.2198_2202del	frameshift	Class 5	PVS1 Very Strong, PP5 Strong, PM2 Moderate
10100105	M	MA	46	Class I	67	*FCL*	Splice Site	c.2949+2T>C	splice_donor	Class 4	PVS1 Very Strong, PM2 Strong
10100130	M	WA	62	Class II	82	*PKP2*	p.Ala570ValfsTer7	c.1709del	frameshift	Class 4	PVS1 Very Strong, PP5 Strong, PM2 Moderate
10100131	F	WA	26			*PKP2**	p.Arg388Trp	c.1162C>T	missense	Class 5	PP5 Strong, PM2 Supporting, PP3 Supporting
10100137	M	WA	15	Class I	58	*CDH2*	p.Asp407Asn	c.1219G>A	missense	Class 4	PVS1 Very Strong, PP5 Strong, PM2 Moderate
10100142	M	IA	49	Class I		*PKP2*	p.Ala326ArgfsTer29	c.968_975dup	frameshift	Class 5	PVS1 Very Strong, PM2 Strong, PP5 Moderate
10100155	M	WA	33	Class I		*PKP2*	Splice Site	c.2489+1G>A	splice_donor	Class 5	PVS1 Very Strong, PP5 Strong, PM2 Moderate
10100317	M	MA	36	Class II	20	*RBM20*	p.Glu1199Ter	c.3595G>T	stop_gained	Class 4	PVS1 Very Strong, PM2 Moderate
10100317	M	MA	36	Class II	20	*TTN*	p.Gly27533ValfsTer47	c.82598del	frameshift	Class 5	PVS1 Very Strong, PM2 Strong
10100390	M	WA	40	Class II	65	*FLNC*	p.Gln357Ter	c.1069C>T	stop_gained	Class 5	PVS1 Very strong, PM2 Strong, PP5 Supporting
101301408	M	WA	37	Class I	67	*PKP2*	p.Arg388Trp	c.1162C>T	missense	Class 5	PP5 Strong, PM2 Supporting, PP3 Supporting
101301411	M	WA	15	Class I	60	*PKP2**	p.Arg388Trp	c.1162C>T	missense	Class 5	PP5 Strong, PM2 Supporting, PP3 Supporting
101301539	F	WA	53	Class I	67	*PKP2*	p.Arg388Trp	c.1162C>T	missense	Class 5	PP5 Strong, PM2 Supporting, PP3 Supporting
DCM											
10100024	F	MA	45	Class III	51	*LMNA*	p.Arg190Trp	c.568C>T	missense	Class 5	PM1 Strong, PP5 Strong, PS3 Strong, PM5 Moderate, PM2 Supporting, PP2 Supporting, PP3 Supporting
10100043	M	WA	16	Class II	26	*LMNA*	p.Arg541His	c.1622G>A	missense	Class 5	PP5 Very Strong, PM1 Strong, PM2 Strong, PM5 Strong, PP2 Supporting, PP3 Supporting
10100045	M	MA	30	Class II	19	*TTN*	Splice Site	c.106531+1G>A	splice_donor	Class 5	PVS1 Very strong, PP5 Strong, PM2 Moderate
10100047	M	BA	44	Class II	22	*SCN5A*	p.Tyr1228Cys	c.3683A>G	missense	Class 4	PM2 Strong, PM1 Supporting, PP3 Supporting
10100305	F	BA	38	Class II	46	*TTN*	p.Glu20759Ter	c.62275G>T	stop_gained	Class 5	PVS1 Very Strong, PM2 Strong
10100306	M	WA	47	Class III	40	*TTN*	p.Gly1316Ter	c.3946G>T	stop_gained	Class 5	PVS1 Very Strong, PM2 Strong
10100309	F	BA	20	Class I	13	*TTN*	Splice Site	c.9704-2A>G	splice_acceptor	Class 5	PVS1 Very strong, PM2 strong
10100312	M	MA	60	Class IV	38	*BAG3*	p.Arg309Ter	c.925C>T	stop_gained	Class 5	PVS1 Very strong, PP5 Strong, PM2 Moderate
10100313	M	BA	24	Class II	15	*TTN*	p.Arg31670Ter	c.95008C>T	stop_gained	Class 5	PVS1 Very strong, PP5 Strong, PM2 Moderate
10100323	M	MA	27	Class II	30	*TNNT2*	p.Arg141Trp	c.421C>T	missense	Class 5	PP5 Strong, PS3 Strong, PM1 Moderate, PM2 Supporting, PP3 Supporting
10100328	M	WA	35	Class IV	17	*BAG3*	p.Glu72Ter	c.214G>T	stop_gained	Class 5	PVS1 Very Strong, PM2 Moderate, PP5 Supporting
10100329	F	MA	39	Class III	21	*FLNC*	p.Gln2303Ter	c.6907C>T	stop_gained	Class 5	PVS1 Very Strong, PM2 Strong, PP5 Moderate
10100340	M	WA	28	Class IV	18	*TTN*	p.Tyr29208Ter	c.87624C>A	stop_gained	Class 4	PVS1 Very Strong, PM2 Moderate, PP5 Moderate
10100345	M	BA	40	Class II	46	*TTN*	p.Phe7410GlnfsTer19	c.22227_22228del	frameshift	Class 5	PVS1 Very Strong, PM2 Moderate, PP5 Moderate
10100351	F	BA	22	Class III	14	*FLNC*	p.Arg1341Ter	c.4021C>T	stop_gained	Class 5	PVS1 Very strong, PP5 Strong, PM2 Moderate
10100363	M	WA	52	Class II	49	*PKP2*	p.Ser837ValfsTer94	c.2509del	frameshift	Class 5	PVS1 Very strong, PP5 Strong, PM2 Moderate
10100365	F	WA	28	Class III	16	*TTN*	p.Trp26308Ter	c.78923G>A	stop_gained	Class 5	PVS1 Very Strong, PM2 Strong, PP5 Moderate
10100375	M	MA	33	Class III	20	*TTN*	p.Arg16724Ter	c.50170C>T	stop_gained	Class 5	PVS1 Very strong, PP5 Strong, PM2 Moderate
10100376	M	MA	29	Class II	36	*TTN*	p.Arg16724Ter	c.50170C>T	stop_gained	Class 5	PVS1 Very strong, PP5 Strong, PM2 Moderate
10100378	M	MA	27	Class III	12	*TTN*	p.Arg29947Ter	c.89839C>T	stop_gained	Class 4	PVS1 Very strong, PM2 Moderate, PP5 Supporting
10100395	M	BA	37	Class III	14	*TTN*	p.Arg31670Ter	c.95008C>T	stop_gained	Class 5	PVS1 Very strong, PP5 Strong, PM2 Moderate
10100706	F	MA	44	Class III	35	*FHL1*	p.Thr152Ile	c.455C>T	missense	Class 5	PM1 Strong, PM2 Strong, PP3 Moderate
10100720	M	MA	30	Class II	19	*TTN*	p.Tyr29208Ter	c.87624C>A	stop_gained	Class 4	PVS1 Very Strong, PM2 Moderate, PP5 Moderate
10100724	M	WA	31	Class III	13	*TTN*	p.Met26800AsnfsTer6	c.80398dup	frameshift	Class 5	PVS1 Very Strong, PM2 Strong
10100742	M	MA	42	Class II	15	*TTN*	p.Pro21370GlnfsTer4	c.64109del	frameshift	Class 5	PVS1 Very Strong, PM2 Moderate
10100747	M	WA	73	Class III	35	*PRKAG2*	Splice Site	c.1585-2A>G	splice_acceptor	Class 5	PM2 Strong, PP3 Moderate
10100750	M	MA	37	Class III	12	*FLNC*	p.Gln2303Ter	c.6907C>T	stop_gained	Class 5	PVS1 Very Strong, PM2 Strong, PP5 Moderate
101301023	F	BA	26	Class II	41	*MYH7*	p.Gly1931AlafsTer2	c.5792del	frameshift	Class 4	PM2 Strong, PM1 Supporting, PP3 Supporting
101301023	F	BA	26	Class II	41	*MYH7*	Splice Site	c.5791-2A>T	splice_acceptor	Class 4	PVS1 Very Strong, PM2 Moderate
101301363	M	MA	34	Class III	11	*TTN*	p.Gly27533ValfsTer47	c.82598del	frameshift	Class 5	PVS1 Very Strong, PM2 Strong
101301367	F	BA	43	Class III	25	*TTN*	p.Arg31670Ter	c.95008C>T	stop_gained	Class 5	PVS1 Very strong, PP5 Strong, PM2 Moderate
101301368	F	BA	25	Class IV	15	*TTN*	p.Lys13508AsnfsTer8	c.40521_40527del	frameshift	Class 4	PVS1 Very Strong, PM2 Moderate
101301372	F	WA	50	Class III	51	*TTN*	p.Tyr29208Ter	c.87624C>A	stop_gained	Class 4	PVS1 Very Strong, PM2 Moderate, PP5 Moderate
101301389	M	MA	40	Class III	19	*TTN*	p.Phe14847CysfsTer7	c.44540_44541del	frameshift	Class 5	PVS1 Very Strong, PM2 Moderate
101301396	F	MA	38	Class II	46	*LMNA*	p.Arg190Trp	c.568C>T	missense	Class 5	PM1 Strong, PP5 Strong, PS3 Strong, PM5 Moderate, PM2 Supporting, PP2 Supporting, PP3 Supporting
101301409	M	MA	25	Class III	24	*TTN*	p.Arg18140Ter	c.54418C>T	stop_gained	Class 4	PVS1 Very Strong, PM2 Moderate, PP5 Moderate
101301418	M	MA	42			*LMNA*	p.Arg190Trp	c.568C>T	missense	Class 5	PM1 Strong, PP5 Strong, PS3 Strong, PM5 Moderate, PM2 Supporting, PP2 Supporting, PP3 Supporting
101301438	F	BA	35	Class II	20	*TTN*	p.Arg31670Ter	c.95008C>T	stop_gained	Class 5	PVS1 Very strong, PP5 Strong, PM2 Moderate
101301504	M	BA	46	Class III	32	*TTN*	p.Arg34079SerfsTer9	c.102234_102237del	frameshift	Class 4	PVS1 Very Strong, PM2 Strong, PP5 Supporting
101301508	F	BA	36	Class III	30	*TTN*	p.Arg15898Ter	c.47692C>T	stop_gained	Class 4	PVS1 Very strong, PP5 Strong, PM2 Moderate
101301528	M	BA	33	Class II	11	*TTN*	Splice Site	c.45350-1G>A	splice_acceptor	Class 5	PVS1 Very Strong, PM2 Strong

101301554	M	BA	16	Class IV	15	*RBM20*	p.Arg634Trp	c.1900C>T	missense	Class 5	PP5 Very Strong, PM1 Strong, PM5 Strong, PM2 Supporting, PP3 Supporting
101301563	M	MA	35			*LMNA*	p.Arg190Trp	c.568C>T	missense	Class 5	PM1 Strong, PP5 Strong, PS3 Strong, PM5 Moderate, PM2 Supporting, PP2 Supporting, PP3 Supporting
101301569	M	MA	30	Class III	11	*TTN*	p.Cys26564Ter	c.79692T>A	stop_gained	Class 4	PVS1 Very Strong, PM2 Moderate
101301717	F	BA	40			*TTN*	p.Phe29218LeufsTer7	c.87654del	frameshift	Class 5	PVS1 Very Strong, PM2 Strong
101304259	F	WA	49	Class I	62	*JUP*	p.Asn321Ter	c.960dup	frameshift	Class 4	PVS1 Very Strong, PM2 Strong
105303524	F	BA	1			*DMD*	Splice Site	c.1149+1G>T	splice_donor	Class 5	PVS1 Very Strong, PM2 Strong
106304529	M	WA	58	Class II	45	*TTN*	p.Tyr29208Ter	c.87624C>A	stop_gained	Class 4	PVS1 Very Strong, PM2 Moderate, PP5 Moderate
106304532	M	WA	26			*TTN*	p.Tyr29208Ter	c.87624C>A	stop_gained	Class 4	PVS1 Very Strong, PM2 Moderate, PP5 Moderate
106304543	F	BA	55	Class II	43	*TTN*	Splice Site	c.48638+2T>G	splice_donor	Class 4	PVS1 Very strong, PM2 strong
106304545	M	BA	51	Class IV	21	*TTN*	p.Arg31670Ter	c.95008C>T	stop_gained	Class 5	PVS1 Very strong, PP5 Strong, PM2 Moderate
106304552	M	BA	40			*TTN*	p.Arg31670Ter	c.95008C>T	stop_gained	Class 5	PVS1 Very strong, PP5 Strong, PM2 Moderate
106304557	M	BA	40			*TTN*	p.Arg31670Ter	c.95008C>T	stop_gained	Class 5	PVS1 Very strong, PP5 Strong, PM2 Moderate
106304560	F	BA	64			*SCN5A*	p.Leu1338Ile	c.4012C>A	missense	Class 4	PM1 Supporting, PM2 Supporting, PP3 Supporting
106304560	F	BA	64			*TTN*	p.Arg31670Ter	c.95008C>T	stop_gained	Class 5	PVS1 Very strong, PP5 Strong, PM2 Moderate
106304569	M	WA	42	Class III	19	*TTN*	p.Tyr29208Ter	c.87624C>A	stop_gained	Class 4	PVS1 Very Strong, PM2 Moderate, PP5 Moderate
106304575	M	BA	35	Class II	20	*DMD**	Splice Site	c.1603-2A>G	splice_acceptor	Class 5	PVS1 Very Strong, PM2 Strong, PP5 Supporting
106304577	M	BA	27	Class III	13	*TTN*	p.Arg31670Ter	c.95008C>T	stop_gained	Class 5	PVS1 Very strong, PP5 Strong, PM2 Moderate
106304582	M	BA	57	Class II	23	*TTN*	p.Arg31670Ter	c.95008C>T	stop_gained	Class 5	PVS1 Very strong, PP5 Strong, PM2 Moderate
106304597	M	MA	22	Class III	12	*MYH7*	p.Arg1250Gly	c.3748C>G	missense	Class 4	PM5 Moderate, PP3 Moderate, PM2 Supporting
401303009	F	BA	62	Class III	35	*TTN*	p.Asn16662ThrfsTer8	c.49985_49988del	frameshift	Class 5	PVS1 Very Strong, PM2 Strong
401303019	M	BA	41	Class III	31	*TTN*	p.Ser33240Ter	c.99719C>G	stop_gained	Class 4	PVS1 Very Strong, PM2 Moderate, PP5 Moderate
401303028	M	BA	30	Class III	19	*TTN*	p.Arg31670Ter	c.95008C>T	stop_gained	Class 5	PVS1 Very strong, PP5 Strong, PM2 Moderate
401303047	M	BA	48	Class IV	22	*FLNC*	p.Ile2003AspfsTer35	c.6006dup	splice_acceptor	Class 4	PVS1 Very Strong, PM2 Strong
LVNC											
10100438	M	BA	38	Class II	20	*TTN*	p.Arg31670Ter	c.95008C>T	stop_gained	Class 5	PVS1 Very strong, PP5 Strong, PM2 Moderate
10100702	M	BA	37	Class II	30	*TTN*	p.Asn16662ThrfsTer8	c.49985_49988del	frameshift	Class 5	PVS1 Very Strong, PM2 Strong, PP5 Supporting
102302024	M	BA	3			*MYBPC3*	p.Phe1137Val	c.3409T>G	missense	Class 4	PM2 Strong, PP3 Moderate
PPCM											
10100007	F	BA	19	Class III	35	*BAG3*	p.Trp36Ter	c.107G>A	stop_gained	Class 5	PVS1 Very strong, PM2 strong
10100065	F	MA	27	Class II	27	*TTN*	p.Arg20626Ter	c.61876C>T	stop_gained	Class 5	PP5 very stron, PVS1 Very strong
10100082	F	MA	27	Class II	20	*TTN*	p.Ile14101SerfsTer40	c.42300del	frameshift	Class 4	PVS1 Very strong, PM2 Moderate
10100338	F	BA	41	Class II	33	*TTN*	p.Glu28397Ter	c.85189G>T	stop_gained	Class 4	PVS1 Very strong, PM2 Moderate
10100344	F	MA	32	Class IV	27	*BAG3*	p.Arg309Ter	c.925C>T	stop_gained	Class 5	PVS1 Very strong, PP5 Strong, PM2 Moderate
10100396	F	BA	35	ND	31	*TTN*	p.Arg31670Ter	c.95008C>T	stop_gained	Class 5	PVS1 Very strong, PP5 Strong, PM2 Moderate
10100402	F	BA	25	ND	11	*TTN*	Splice Site	c.9704-2A>G	splice_acceptor	Class 5	PVS1 Very strong, PM2 strong
10100713	F	MA	22	Class III	22	*TTN*	p.Arg28779Ter	c.86335C>T	stop_gained	Class 4	PVS1 Very strong, PM2 Moderate, PP5 Supporting
106304528	F	BA	32	Class I	25	*TTN*	Splice Site	c.48638+2T>G	splice_donor	Class 4	PVS1 Very strong, PM2 strong
106304542	F	BA	25	Class II	23	*MYH7*	p.Pro152Arg	c.455C>G	missense	Class 4	PM2 Strong, PM1 Supporting, PP3 Supporting
106304556	F	BA	46			*TTN*	p.Asp24091GlufsTer8	c.72269_72272dup	frameshift	Class 5	PVS1 Very strong, PM2 strong
106304571	F	BA	20	Class IV	13	*TTN*	p.Arg31670Ter	c.95008C>T	stop_gained	Class 5	PVS1 Very strong, PP5 Strong, PM2 Moderate
401303008	F	BA	25	Class III	20	*TTN*	p.Arg15898Ter	c.47692C>T	stop_gained	Class 4	PVS1 Very strong, PP5 Strong, PM2 Moderate
HCM											
10100066	M	WA	25	Class III	83	*MYBPC3*	p.Trp792ValfsTer41	c.2373dup	frameshift	Class 5	PP5 Very Strong, PVS1 Very Strong, PM2 Moderate
10100083	F	MA	64	Class I	67	*MYH7*	p.Arg204His	c.611G>A	missense	Class 5	PM1 Strong, PM2 Supporting, PP3 Supporting, PP5 Supporting
10100502	M	MA	29	Class I	80	*MYH7*	p.Arg204His	c.611G>A	missense	Class 5	PM1 Strong, PM2 Supporting, PP3 Supporting, PP5 Supporting
10100505	M	MA	38	Class II	78	*DSP*	p.Lys1892del	c.5676_5678del	inframe_deletion	Class 4	PM2 Strong, PM4 Moderate, PP3 Supporting
10100517	M	MA	36	Class I	69	*MYBPC3*	p.Ser478Ter	c.1433C>A	stop_gained	Class 5	PVS1 Very Strong, PM2 Strong
10100518	M	WA	46	Class II	74	*GLA**	p.Pro259ArgfsTer5	c.774_775del	frameshift	Class 5	PVS1 Very Strong, PM2 Strong, PP5 Supporting
10100527	M	MA	25	Class I	73	*ALPK3*	p.Arg1607Ter	c.4819C>T	stop_gained	Class 5	PVS1 Very Strong, PM2 Moderate, PP5 Moderate
10100535	F	WA	23	Class II	64	*MYH7*	p.Arg249Gln	c.746G>A	missense	Class 5	PP5 Very Strong, PM1 Strong, PM2 Strong, PM5 Moderate, PP3 Supporting
10100536	F	MA	25			*MYH7*	p.Arg1420Trp	c.4258C>T	missense	Class 4	PP5 Very Strong, PM2 Supporting, PP3 Supporting
10100715	F	MA	38	Class I		*MYBPC3*	p.Glu258Lys	c.772G>A	missense	Class 5	PP5 Very Strong, PM2 Strong, PP3 Strong
101301381	F	MA	37	Class III	70	*MYH7*	p.Ala797Thr	c.2389G>A	missense	Class 5	PP5 Very Strong, PM1 Strong, PM5 Moderate, PM2 Supporting, BP4 Supporting
101301393	F	WA	31	Class II		*MYH7*	p.Ala797Thr	c.2389G>A	missense	Class 5	PP5 Very Strong, PM1 Strong, PM5 Moderate, PM2 Supporting, BP4 Supporting
101301575	F	BA	24	Class IV	61	*TNNI3*	p.Leu144His	c.431_432delinsAT	missense	Class 5	PM1 Strong, PM2 Strong, PM5 Strong, PP3 Supporting
101301575	F	BA	24	Class IV	61	*TNNI3*	p.Leu144Gln	c.431T>A	missense	Class 5	PM1 Strong, PM2 Strong, PM5 Strong, PP3 Supporting
401303027	M	BA	33	Class I	64	*MYBPC3*	Splice Site	c.821+1G>A	splice_donor	Class 5	PVS1 Very Strong, PP5 Strong, PM2 Moderate
RCM											
10100069	F	MA	28	Class II	61	*DES*	p.Arg454Trp	c.1360C>T	missense	Class 5	PM1 Strong, PS3 Strong, PP5 Moderate, PM2 Supporting, PP3 Supporting

10100097	F	MA	21	Class III	45	*MYH7*	p.Arg453Cys	c.1357C>T	missense	Class 5	PP5 Very Strong, PM1 Strong, PM5 Strong, PM2 Supporting, PP3 Supporting
101301608	F	MA	25			*FLNC*	p.Gly2011Arg	c.6031G>A	missense	Class 5	PP5 Very Strong, PP3 Moderate, PM2 Supporting
107305003	F	BA	30	Class II	55	*DES*	p.Gln165Leu	c.494A>T	missense	Class 4	PM5 Moderate, PP3 Moderate, PM2 Supporting

Asterisk (*)-Homozygous

SS Table 2: Table with ACMG/AMP rules for probands with class 3 variants

Probands with class 3 variants

ID	Sex	Ancestry	Age	NYHA	LVEF	Gene	Protein_change	cDNA_Change	Consequences	ACMG	ACMG Abbreviation
ACM											
10100050	M	WA	51	Class I	56	*TTN*	p.Gly8028Ala	c.24083G>C	missense	Class 3 LB	TTN missense not reviewed
10100104	M	MA	34	Class I	45	*SCN5A*	p.Leu189Phe	c.565C>T	missense	Class 3	PP3_Supporting, PM2_Supporting
10100104	M	MA	34	Class I	45	*TTN*	p.Val21602Leu	c.64804G>T	missense	Class 3 LB	TTN missense not reviewed
10100117	M	WA	38	Class I	67	*DSP*	p.Gln678Arg	c.2033A>G	missense	Class 3	PM2, PM1
10100117	M	WA	38	Class I	67	*DSP*	p.Arg1341His	c.4022G>A	missense	Class 3 LP	PM2 Strong, PM5 Supporting, PP3 Supporting
10100117	M	WA	38	Class I	67	*TTN*	p.Asp14861Asn	c.44581G>A	missense	Class 3 LB	TTN missense not reviewed
10100120	F	WA	15	Class I		*DMD*	p.Arg563Cys	c.1687C>T	missense	Class 3	BS1;BP4
10100133	M	MA	37	Class I	65	*DSP*	p.Asn287Ser	c.860A>G	missense	Class 3 LP	PM2 Strong, PM5 Supporting, PP3 Supporting
10100134	M	WA	40	Class I	64	*ALPK3*	p.Glu148GlnfsTer8	c.436_440dup	frameshift	Class 3 LP	PVS1 Strong, PM2 Moderate
10100134	M	WA	40	Class I	64	*FLNC*	p.Ser2546Thr	c.7636T>A	missense	Class 3	Conflicting
10100134	M	WA	40	Class I	64	*SCN5A*	p.Arg1195His	c.3584G>A	missense	Class 3	PP3_Supporting
10100134	M	WA	40	Class I	64	*TTN*	p.Arg30628Lys	c.91883G>A	missense	Class 3 LB	TTN missense not reviewed
10100145	M	WA	75	Class I	62	*TTN*	p.Gly17356Ser	c.52066G>A	missense	Class 3 LB	TTN missense not reviewed
10100146	F	BA	33	Class I	63	*DSP*	p.Ser2843_Arg2846del	c.8508_8519del	inframe_deletion	Class 3	PM2, PM1
10100152	F	WA	41	Class I	62	*ANKRD1*	p.Ala128LeufsTer8	c.382del	frameshift	Class 3	Gene not relevant in context
10100152	F	WA	41	Class I	62	*TTN*	p.Thr27130Ile	c.81389C>T	missense	Class 3 LB	TTN missense not reviewed
10100154	F	MA	58	Class I	67	*DMD*	p.Met1313Val	c.3937A>G	missense	Class 3	BP4
10100154	F	MA	58	Class I	67	*DSP*	p.Asn1865Tyr	c.5593A>T	missense	*Class 3 LB*	Borderline BS1 and borderline BP4
10100158	F	WA	61	Class I	53	*TTN*	p.Gly4809Ser	c.14425G>A	missense	Class 3 LB	TTN missense not reviewed
10100704	M	WA	48	Class II	46	*TTN*	p.Pro22865Leu	c.68594C>T	missense	Class 3 LB	TTN missense not reviewed
101301404	F	BA	63	Class II	66	*TTN*	p.Ala352Thr	c.1054G>A	missense	Class 3 LB	TTN missense not reviewed
101301404	F	BA	63	Class II	66	*DES*	p.Asp312Ala	c.935A>C	missense	Class 3 LP	PM2 Strong, PM1 Moderate, PP3 Supporting
101301517	F	BA	31	Class II	47	*ALPK3*	p.Val302Ile	c.904G>A	missense	Class 3	BA1
101304180	M	IA	24			TTN	p.Met14170Arg	c.42509T>G	missense	Class 3 LP	TTN missense not reviewed
101304184	M	WA	52			TTN	p.Lys2255Asn	c.6765A>C	missense	Class 3 LB	TTN missense not reviewed
DCM											
10100001	F	BA	28	Class III	33	*DMD*	p.Pro1604Arg	c.4811C>G	missense	Class 3	BP4
10100001	F	BA	28	Class III	33	*DSP*	p.Gln678Arg	c.2033A>G	missense	Class 3	PM2, PM1
10100001	F	BA	28	Class III	33	*DSP*	p.Arg1341His	c.4022G>A	missense	Class 3 LP	PM2 Strong, PM5 Supporting, PP3 Supporting
10100001	F	BA	28	Class III	33	*TTN*	p.Ile17045Thr	c.51134T>C	missense&splice_region	Class 3 LB	TTN missense not reviewed
10100010	M	BA	26	Class II	62	*TTN*	p.Arg30628Lys	c.91883G>A	missense	Class 3 LB	TTN missense not reviewed
10100012	M	MA	26	Class III	24	*TTN*	p.Arg30628Lys	c.91883G>A	missense	Class 3 LB	TTN missense not reviewed
10100015	F	MA	33	Class III	31	*MYH7*	p.Arg1677His	c.5030G>A	missense	Class 3	Conflicting
10100015	F	MA	33	Class III	31	*TTN*	p.Gly24659Ala	c.73976G>C	missense	Class 3 LB	TTN missense not reviewed
10100015	F	MA	33	Class III	31	*TTN*	p.Lys11470Glu	c.34408A>G	missense	Class 3 LB	TTN missense not reviewed
10100015	F	MA	33	Class III	31	*TTN*	p.Gln3708Glu	c.11122C>G	missense	Class 3 LB	TTN missense not reviewed
10100016	M	BA	40	Class III	19	*DSP*	p.Arg925Trp	c.2773C>T	missense	Class 3	Conflicting
10100021	M	WA	37	Class III	25	*TMEM43*	p.Arg328Trp	c.982C>T	missense	Class 3	Conflicting
10100021	M	WA	37	Class III	25	*TTN*	p.Arg16197Cys	c.48589C>T	missense	Class 3 LB	TTN missense not reviewed
10100021	M	WA	37	Class III	25	*TTN*	p.Phe15649Leu	c.46945T>C	missense	Class 3 LB	TTN missense not reviewed
10100021	M	WA	37	Class III	25	*TTN*	p.Cys9195Ser	c.27584G>C	missense	Class 3 LB	TTN missense not reviewed
10100028	M	BA	31	Class III	34	*TTN*	p.Lys11126Glu	c.33376A>G	missense	Class 3 LB	TTN missense not reviewed
10100030	F	BA	22	Class III	21	*TNNC1*	p.Met157Ile	c.471G>T	missense	Class 3	PM2, PM1
10100034	F	BA	53	Class III	21	*TTN*	p.Pro21332Thr	c.63994C>A	missense	Class 3 LB	TTN missense not reviewed
10100037	F	MA	41	Class IV	35	*DSP*	p.Val337Ala	c.1010T>C	missense	Class 3	PM2, PM1
10100038	F	BA	27	Class III	33	*MYL2*	p.Val61Met	c.181G>A	missense	Class 3	Conflicting
10100038	F	BA	27	Class III	33	*TTN*	p.Arg30628Lys	c.91883G>A	missense	Class 3 LB	TTN missense not reviewed
10100040	M	MA	63	Class II	19	*TTN*	Splice Site	c.32554+1G>T	splice_donor	Class 3 LP	PVS1 Very Strong, PM2 Moderate
10100044	M	BA	48	Class III	30	*PKP2*	p.Ile397Asn	c.1190T>A	missense	Class 3	Conflicting
10100053	F	BA	46	Class III	14	*TTN*	p.Ser19893Ala	c.59677T>G	missense	Class 3 LB	TTN missense not reviewed
10100059	M	MA	31	Class II	18	*DSP*	p.Gln678Arg	c.2033A>G	missense	Class 3	PM2, PM1
10100059	M	MA	31	Class II	18	*DSP*	p.Arg1341His	c.4022G>A	missense	Class 3 LP	PM2 Strong, PM5 Supporting, PP3 Supporting
10100061	F	BA	26	Class III	18	*TTN*	p.Gly19031Glu	c.57092G>A	missense	Class 3 LB	TTN missense not reviewed
10100061	F	BA	26	Class III	18	*TTN*	p.Asp9656Asn	c.28966G>A	missense	Class 3 LB	TTN missense not reviewed
10100067	M	MA	54	Class IV	19	*ALPK3*	p.Glu148GlnfsTer8	c.436_440dup	frameshift	Class 3 LP	PVS1 Strong, PM2 Moderate
10100073	M	BA	39	Class III	20	*TNNC1*	p.Arg147Cys	c.439C>T	missense	Class 3	PM2, PM1

10100076	M	MA	47	Class III	15	*LAMP2**	p.Asn119Ser	c.356A>G	missense	Class 3	PM2, PM1
10100090	M	BA	26	Class I	55	*ALPK3*	p.Glu148GlnfsTer8	c.436_440dup	frameshift	Class 3 LP	PVS1 Strong, PM2 Moderate
10100093	M	BA	52	Class III	20	*DES*	p.Arg315His	c.944G>A	missense	Class 3 LP	PM1 Moderate, PP3 Strong, PM2 Supporting
10100093	M	BA	52	Class III	20	*DSP*	p.Gln678Arg	c.2033A>G	missense	Class 3	PM2, PM1
10100093	M	BA	52	Class III	20	*DSP*	p.Arg1341His	c.4022G>A	missense	Class 3 LP	PM2 Strong, PM5 Supporting, PP3 Supporting
10100301	M	BA	37	Class II	19	*VCL*	p.Pro943ArgfsTer9	c.2828_2829del	frameshift	Class 3	Conflicting
10100303	M	MA	33	Class IV	34	*DMD*	p.Arg1183Gly	c.3547A>G	missense	Class 3	Conflicting
10100303	M	MA	33	Class IV	34	*DSP*	p.Asn408His	c.1222A>C	missense	Class 3	Conflicting
10100303	M	MA	33	Class IV	34	MYBPC3	p.Glu1179Lys	c.3535G>A	missense	Class 3	Conflicting
10100303	M	MA	33	Class IV	34	*TTN*	p.Thr34096Ile	c.102287C>T	missense	Class 3 LB	TTN missense not reviewed
10100304	F	MA	21	Class II	32	*TTN*	p.Asn32895_Glu32898delinsLys	c.98685_98693del	inframe_deletion	Class 3	small in frame deletions
10100307	M	MA	30	Class II	41	*FLNC*	p.Arg1434Cys	c.4300C>T	missense	Class 3	Conflicting
10100311	F	BA	59			*TTN*	p.Gly24659Ala	c.73976G>C	missense	Class 3 LB	TTN missense not reviewed
10100311	F	BA	59			*TTN*	p.Ser19893Ala	c.59677T>G	missense	Class 3 LB	TTN missense not reviewed
10100311	F	BA	59			*TTN*	p.Lys11470Glu	c.34408A>G	missense	Class 3 LB	TTN missense not reviewed
10100311	F	BA	59			*TTN*	p.Gln3708Glu	c.11122C>G	missense	Class 3 LB	TTN missense not reviewed
10100316	F	MA	42	Class III	51	*ALPK3*	p.Val302Ile	c.904G>A	missense	Class 3	BA1
10100316	F	MA	42	Class III	51	*TTN*	Splice Site	c.11254+2T>C	splice_donor	*Class 3 LB*	putative truncating variants not in 100% PSI regions
10100316	F	MA	42	Class III	51	*TTN*	p.Phe15649Leu	c.46945T>C	missense	Class 3 LB	TTN missense not reviewed
10100318	M	BA	38	Class III	29	*DSC2*	p.Gly811Val	c.2432G>T	missense	Class 3	PM2, PM1
10100318	M	BA	38	Class III	29	*DSP*	p.Gln678Arg	c.2033A>G	missense	Class 3	PM2, PM1
10100318	M	BA	38	Class III	29	*DSP*	p.Arg1341His	c.4022G>A	missense	Class 3 LP	PM2 Strong, PM5 Supporting, PP3 Supporting
10100318	M	BA	38	Class III	29	*MYH7*	p.Ala728Val	c.2183C>T	missense	Class 3 LP	PM1 Strong, PM2 Supporting, PP5 Supporting
10100321	M	MA	53	Class III	20	*MYBPC3*	p.Arg282Gln	c.845G>A	missense	Class 3	Conflicting
10100332	M	MA	34	Class III	27	*MYBPC3*	p.Ala534Val	c.1601C>T	missense	Class 3 LP	PM1 Moderate, PM2 Suppporting, PP3 Supporting
10100336	F	BA	40	Class III	16	*TTN*	p.Gly24659Ala	c.73976G>C	missense	Class 3 LB	TTN missense not reviewed
10100336	F	BA	40	Class III	16	*TTN*	p.Lys11470Glu	c.34408A>G	missense	Class 3 LB	TTN missense not reviewed
10100336	F	BA	40	Class III	16	*TTN*	p.Gln3708Glu	c.11122C>G	missense	Class 3 LB	TTN missense not reviewed
10100341	M	BA	19	Class IV	5	*DES*	p.Asp312Asn	c.934G>A	missense	Class 3 LP	PM2 Strong, PM1 Moderate, PP3 Supporting
10100349	M	WA	67			*TTN*	p.Arg3148Leu	c.9443G>T	missense	Class 3 LB	TTN missense not reviewed
10100354	M	MA	37	ND	27	*FLNC*	p.Gly406Ser	c.1216G>A	missense	Class 3	Conflicting
10100355	M	BA	49	Class II	21	*DSP*	p.Gln678Arg	c.2033A>G	missense	Class 3	PM2, PM1
10100355	M	BA	49	Class II	21	*DSP*	p.Arg1341His	c.4022G>A	missense	Class 3 LP	PM2 Strong, PM5 Supporting, PP3 Supporting
10100355	M	BA	49	Class II	21	*TTN*	p.Leu51Arg	c.152T>G	missense	Class 3 LB	TTN missense not reviewed
10100357	M	BA	38	Class II	15	*SCN5A*	p.Leu1308Phe	c.3922C>T	missense	Class 3	PM2, PM1
10100357	M	BA	38	Class II	15	TTN	p.Val19400Ala	c.58199T>C	missense	Class 3 LB	TTN missense not reviewed
10100357	M	BA	38	Class II	15	TTN	p.Leu15490Ile	c.46468C>A	missense	Class 3 LB	TTN missense not reviewed
10100357	M	BA	38	Class II	15	TTN	p.Gln3708Glu	c.11122C>G	missense	Class 3 LB	TTN missense not reviewed
10100358	F	BA	34	Class IV	38	*SCN5A*	p.Leu1308Phe	c.3922C>T	missense	Class 3	PM2, PM1
10100368	M	WA	43	Class II	46	RBM20	p.Phe920Cys	c.2759T>G	missense	Class 3	Conflicting
10100379	M	MA	31			*TPM1*	p.Glu243Lys	c.727G>A	missense	Class 3	Conflicting
10100379	M	MA	31			*TTN*	p.Ile14525Phe	c.43573A>T	missense	Class 3 LB	TTN missense not reviewed
10100384	M	BA	31	Class I	25	*TMEM43*	p.Arg328Trp	c.982C>T	missense	Class 3	Conflicting
10100387	M	BA	20	Class IV	14	*MYH7*	p.Arg1396Trp	c.4186C>T	missense	Class 3	Conflicting
10100387	M	BA	20	Class IV	14	*PKP2*	p.Ile397Asn	c.1190T>A	missense	Class 3	Conflicting
10100387	M	BA	20	Class IV	14	*TTN*	p.Ile32009Thr	c.96026T>C	missense&splice_region	Class 3 LB	TTN missense not reviewed
10100391	F	MA	53			*ALPK3*	p.Val302Ile	c.904G>A	missense	Class 3	BA1
10100391	F	MA	53			*FLNC*	p.Arg1434Cys	c.4300C>T	missense	Class 3	Conflicting
10100391	F	MA	53			*MYH7*	p.Ala728Val	c.2183C>T	missense	Class 3 LP	PM1 Strong, PM2 Supporting, PP5 Supporting
10100392	M	MA	34	Class I	55	*CDH2*	p.Gly668Glu	c.2003G>A	missense	Class 3	PM2, PM1
10200016	F	BA	4			*SCN5A*	p.Asp2003Asn	c.6007G>A	missense	Class 3 LP	PM2 Strong, PP3 Supporting
10200016	F	BA	4			*DMD*	p.Asp924His	c.2770G>C	missense	Class 3	Conflicting
10100433	F	BA	29	Class II	24	*PKP2*	p.Ile397Asn	c.1190T>A	missense	Class 3	Conflicting
10100436	M	BA	27	Class I	35	*DMD**	p.Ile766Val	c.2296A>G	missense	Class 3	Conflicting
10100436	M	BA	27	Class I	35	*TNNT2*	p.Arg158Pro	c.473G>C	missense	Class 3	Conflicting
10100439	M	WA	53	Class II	16	*TTN*	p.Trp17281Arg	c.51841T>C	missense	Class 3 LB	TTN missense not reviewed
10100444	M	MA	50	Class II	14	*ANKRD1*	p.Lys126ArgfsTer10	c.377del	frameshift	Class 3	Gene not relevant in context
10100444	M	MA	50	Class II	14	*TTN*	p.Pro2129Ala	c.6385C>G	missense	Class 3 LB	TTN missense not reviewed
10100700	M	BA	39	Class IV	29	*TTN*	p.Ser8409Tyr	c.25226C>A	missense	Class 3 LB	TTN missense not reviewed

10100701	M	BA	52			*TTN*	p.Glu11001Lys	c.33001G>A	missense	Class 3 LB	TTN missense not reviewed
10100708	M	BA	48	Class III	25	*SCN5A*	p.Arg1847His	c.5540G>A	missense	Class 3 LP	PM1 Moderate, PM2 Suppporting, PP3 Supporting
10100711	M	BA	18	Class III	17	*ALPK3*	p.Glu148GlnfsTer8	c.436_440dup	frameshift	Class 3 LP	PVS1 Strong, PM2 Moderate
10100711	M	BA	18	Class III	17	*TTN*	p.Ile16869Thr	c.50606T>C	missense	Class 3 LB	TTN missense not reviewed
10100712	M	BA	28	Class III	47	*MYH7*	p.Ala728Val	c.2183C>T	missense	Class 3 LP	PM1 Strong, PM2 Supporting, PP5 Supporting
10100712	M	BA	28	Class III	47	*SCN5A*	p.Leu1308Phe	c.3922C>T	missense	Class 3	PM2, PM1
10100714	M	BA	44	Class III	19	*MYH7*	p.Gly377Ser	c.1129G>A	missense	Class 3 LP	PM1 Strong, PM2 Supporting, PP3 Supporting
10100716	F	BA	18	Class III	42	*MYBPC3*	p.Glu1179Lys	c.3535G>A	missense	Class 3	Conflicting
10100721	F	MA	29	Class III	24	*DMD*	p.Arg204Thr	c.611G>C	missense	Class 3	Conflicting
10100723	M	BA	22	Class II	26	*TTN*	p.Lys11470Glu	c.34408A>G	missense	Class 3 LB	TTN missense not reviewed
10100726	M	MA	38	Class III	33	*SCN5A*	p.Arg971Cys	c.2911C>T	missense	Class 3 LP	PM2 Supporting, PP5 Supporting, PP3 Supporting
10100727	F	BA	20	Class III	23	*LAMP2*	p.Asp322Val	c.965A>T	missense	Class 3	PM2, PM1
10100728	M	BA	34	Class III	8	*TTN*	p.Gly19031Glu	c.57092G>A	missense	Class 3 LB	TTN missense not reviewed
10100728	M	BA	34	Class III	8	*TTN*	p.Lys11470Glu	c.34408A>G	missense	Class 3 LB	TTN missense not reviewed
10100735	F	BA	42	Class III	21	*MYH7*	p.Arg1396Trp	c.4186C>T	missense	Class 3	Conflicting
10100735	F	BA	42	Class III	21	*TTN*	p.Arg28111Lys	c.84332G>A	missense	Class 3 LB	TTN missense not reviewed
10100735	F	BA	42	Class III	21	*TTN*	p.Gly19031Glu	c.57092G>A	missense	Class 3 LB	TTN missense not reviewed
10100735	F	BA	42	Class III	21	*TTN*	p.Lys11470Glu	c.34408A>G	missense	Class 3 LB	TTN missense not reviewed
10100743	F	MA	21	Class I	64	*RBM20*	p.Tyr628His	c.1882T>C	missense&splice_region	Class 3	PM2, PM1
10100743	F	MA	21	Class I	64	*TTN*	p.Pro23850Ala	c.71548C>G	missense	Class 3 LB	TTN missense not reviewed
10100743	F	MA	21	Class I	64	*TTN*	p.Gly17656Arg	c.52966G>A	missense	Class 3 LB	TTN missense not reviewed
10100743	F	MA	21	Class I	64	*TTN*	p.Lys11470Glu	c.34408A>G	missense	Class 3 LB	TTN missense not reviewed
10100744	F	MA	40	Class IV	12	*DMD*	p.Ala2999Val	c.8996C>T	missense	Class 3 LP	PM2 Strong, PM1 Moderate
10100744	F	MA	40	Class IV	12	*DES*	p.Arg222His	c.665G>A	missense	*Class 3 LB*	Borderline BS1 and borderline BP4
10100746	M	MA	35	Class III	19	*DES*	p.Asp312Ala	c.935A>C	missense	Class 3 LP	PM2 Strong, PM1 Moderate, PP3 Supporting
10100746	M	MA	35	Class III	19	*TTN*	p.Gly23068Val	c.69203G>T	missense	Class 3 LB	TTN missense not reviewed
10200001	F	MA	1			*SCN5A*	p.Gln1832Glu	c.5494C>G	missense	Class 3	Conflicting
10200002	M	BA	0			*TTN*	p.Lys11470Glu	c.34408A>G	missense	Class 3 LB	TTN missense not reviewed
10200006	M	MA	3			*DSP*	p.Gln678Arg	c.2033A>G	missense	Class 3	PM2, PM1
10200006	M	MA	3			*DSP*	p.Arg1341His	c.4022G>A	missense	Class 3 LP	PM2 Strong, PM5 Supporting, PP3 Supporting
10200006	M	MA	3			*MYH7*	p.Ala728Val	c.2183C>T	missense	Class 3 LP	PM1 Strong, PM2 Supporting, PP5 Supporting
10200008	F	BA	11			*DSG2*	p.Ser884del	c.2649_2651del	inframe_deletion	Class 3	Conflicting
10200010	F	MA	0			*TTN*	p.Pro20108Leu	c.60323C>T	missense	Class 3 LB	TTN missense not reviewed
10200010	F	MA	0			*CDH2*	p.Arg697His	c.2090_2091inv	missense	Class 3	PM2, PM1
10200019	M	BA	3			*TTN*	p.Ser8409Tyr	c.25226C>A	missense	Class 3 LB	TTN missense not reviewed
10200020	F	MA	3			*DSP*	p.Gln678Arg	c.2033A>G	missense	Class 3	PM2, PM1
10200020	F	MA	3			*DSP*	p.Arg1341His	c.4022G>A	missense	Class 3 LP	PM2 Strong, PM5 Supporting, PP3 Supporting
10200020	F	MA	3			*FHL2*	p.Arg203Cys	c.607C>T	missense	Class 3	Conflicting
10200021	F	BA	0			*FLNC*	p.Gly1546Ser	c.4636G>A	missense	Class 3	Conflicting
10200022	M	MA	0			*ANKRD1*	p.Asn131LysfsTer2	c.392dup	frameshift	Class 3	Gene not relevant in context
10200026	M	BA	2			*DSP*	p.Gln678Arg	c.2033A>G	missense	Class 3	PM2, PM1
10200026	M	BA	2			*DSP*	p.Arg1341His	c.4022G>A	missense	Class 3 LP	PM2 Strong, PM5 Supporting, PP3 Supporting
10200026	M	BA	2			*TTN*	p.Gly24659Ala	c.73976G>C	missense	Class 3 LB	TTN missense not reviewed
10200026	M	BA	2			*TTN*	p.Lys11470Glu	c.34408A>G	missense	Class 3 LB	TTN missense not reviewed
10200026	M	BA	2			*TTN*	p.Gln3708Glu	c.11122C>G	missense	Class 3 LB	TTN missense not reviewed
10200028	M	MA	2			*DSP*	p.Gln678Arg	c.2033A>G	missense	Class 3	PM2, PM1
10200028	M	MA	2			*DSP*	p.Arg1341His	c.4022G>A	missense	Class 3 LP	PM2 Strong, PM5 Supporting, PP3 Supporting
10200028	M	MA	2			*TTN*	p.Asp14861Asn	c.44581G>A	missense	Class 3 LB	TTN missense not reviewed
10300007	F	MA	2			TMEM43	p.Arg328Trp	c.982C>T	missense	Class 3	Conflicting
10300007	F	MA	2			TTN	p.Arg16197Cys	c.48589C>T	missense	Class 3 LB	TTN missense not reviewed
10300007	F	MA	2			TTN	p.Phe15649Leu	c.46945T>C	missense	Class 3 LB	TTN missense not reviewed
10300007	F	MA	2			TTN	p.Cys9195Ser	c.27584G>C	missense	Class 3 LB	TTN missense not reviewed
101301351	M	BA	41	Class III	16	*TTN*	p.Arg19111Gln	c.57332G>A	missense	Class 3 LB	TTN missense not reviewed
101301351	M	BA	41	Class III	16	*TTN*	p.Ser8409Tyr	c.25226C>A	missense	Class 3 LB	TTN missense not reviewed
101301355	M	MA	55	Class III	16	TTN	p.Val24611Leu	c.73831G>T	missense	Class 3 LB	TTN missense not reviewed
101301355	M	MA	55	Class III	16	FLNC	p.Asn47Ser	c.140A>G	missense	Class 3	Conflicting
101301361	M	MA	22	Class III	30	*TTN*	p.Trp29581Cys	c.88743G>T	missense	Class 3 LB	TTN missense not reviewed
101301373	M	MA	30	Class III	13	*ALPK3*	p.Tyr241Asn	c.721T>A	missense	Class 3	BA1
101301373	M	MA	30	Class III	13	*DES*	p.Arg160Gln	c.479G>A	missense	Class 3 LP	PM2 Strong, PM1 Moderate, PP3 Supporting

101301379	M	WA	59	Class IV	17	*SCN5A*	p.Ser1744Thr	c.5231G>C	missense	Class 3LP	PM1 Strong, PM5 Moderate, PM2 Supporting, PP3 supporting
101301380	F	MA	29	Class III	32	*TTN*	Splice Site	c.32471-1G>A	splice_acceptor	Class 3LP	PVS1 Very Strong, PM2 Moderate
101301380	F	MA	29	Class III	32	*TTN*	p.Lys28450Asn	c.85350G>C	missense	Class 3LB	TTN missense not reviewed
101301380	F	MA	29	Class III	32	*TTN*	p.Ser19018Ala	c.57052T>G	missense	Class 3LB	TTN missense not reviewed
101301380	F	MA	29	Class III	32	*TTN*	p.Cys9226Tyr	c.27677G>A	missense	Class 3LB	TTN missense not reviewed
101301382	F	MA	26	Class III	38	*TTN*	Splice Site	c.32534+1G>T	splice_donor	Class 3LP	PVS1 Very Strong, PM2 Moderate, PP3 Supporting
101301384	M	BA	27	Class IV	39	LMNA	p.Arg427Cys	c.1279C>T	missense	Class 3LP	PM1 Strong, PM2 Supporting, PP2 Supporting, PP3 Supporting
101301384	M	BA	27	Class IV	39	TTN	p.Ala352Thr	c.1054G>A	missense	Class 3LB	TTN missense not reviewed
101301385	M	BA	57	Class III	10	*MYL2*	p.Val61Met	c.181G>A	missense	Class 3	Conflicting
101301385	M	BA	57	Class III	10	*TTN*	p.Ile18822Thr	c.56465T>C	missense	Class 3LB	TTN missense not reviewed
101301387	M	MA	40	Class III	12	*LAMP2*	p.Asp322Val	c.965A>T	missense	Class 3	PM2, PM1
101301391	F	MA	21	Class III	12	*MYBPC3*	p.Glu1179Lys	c.3535G>A	missense	Class 3	Conflicting
101301402	F	MA	48	Class III	20	TTN	p.Gly24659Ala	c.73976G>C	missense	Class 3LB	TTN missense not reviewed
101301402	F	MA	48	Class III	20	TTN	p.Ser6627Leu	c.19880C>T	missense	Class 3LB	TTN missense not reviewed
101301410	F	MA	43	Class III	45	*FLNC*	p.Lys2127Asn	c.6381G>C	missense	Class 3	PM2, PM1
101301410	F	MA	43	Class III	45	*MYBPC3*	p.Ala833Thr	c.2497G>A	missense	Class 3LP	PP5 Very Strong, PM2 Supporting, PP3 Supporting, BP6 Supporting
101301421	M	BA	44	Class IV	24	*DES*	p.Asp312Asn	c.934G>A	missense	Class 3LP	PM2 Strong, PM1 Moderate, PP3 Supporting
101301422	M	MA	68			*DSP*	p.Glu2400Lys	c.7198G>A	missense	Class 3	PM2, PM1
101301427	M	MA	42	Class III	20	*TTN*	p.Thr8852Lys	c.26555C>A	missense	Class 3LB	TTN missense not reviewed
101301435	M	BA	30	Class II	25	*DES*	p.Asp312Asn	c.934G>A	missense	Class 3LP	PM2 Strong, PM1 Moderate, PP3 Supporting
101301435	M	BA	30	Class II	25	*TTN*	p.Arg34949His	c.104846G>A	missense	Class 3LB	TTN missense not reviewed
101301435	M	BA	30	Class II	25	*TTN*	p.Leu29588Ser	c.88763T>C	missense	Class 3LB	TTN missense not reviewed
101301435	M	BA	30	Class II	25	*TTN*	p.Pro16152Ser	c.48454C>T	missense	Class 3LB	TTN missense not reviewed
101301435	M	BA	30	Class II	25	*TTN*	p.Ser9464Phe	c.28391C>T	missense	Class 3LB	TTN missense not reviewed
101301436	M	BA	45	Class III	21	*TTN*	p.Thr31293Ile	c.93878C>T	missense	Class 3LB	TTN missense not reviewed
101301436	M	BA	45	Class III	21	*TTN*	p.Val24611Leu	c.73831G>T	missense	Class 3LB	TTN missense not reviewed
101301436	M	BA	45	Class III	21	*TTN*	p.Glu11001Lys	c.33001G>A	missense	Class 3LB	TTN missense not reviewed
101301442	M	BA	20	Class II	22	TTN	p.Gly27897Asp	c.83690G>A	missense	Class 3LB	TTN missense not reviewed
101301446	M	BA	28	Class III	30	TTN	p.Thr23435Met	c.70304C>T	missense	Class 3LB	TTN missense not reviewed
101301446	M	BA	28	Class III	30	TTN	p.Lys11470Glu	c.34408A>G	missense	Class 3LB	TTN missense not reviewed
101301446	M	BA	28	Class III	30	TTN	p.Lys621Glu	c.1861A>G	missense	Class 3LB	TTN missense not reviewed
101301446	M	BA	28	Class III	30	TTN	p.Ala352Thr	c.1054G>A	missense	Class 3LB	TTN missense not reviewed
101301507	F	BA	40	Class III	36	*DSP*	p.Gln678Arg	c.2033A>G	missense	Class 3	PM2, PM1
101301507	F	BA	40	Class III	36	*DSP*	p.Arg1341His	c.4022G>A	missense	Class 3LP	PM2 Strong, PM5 Supporting, PP3 Supporting
101301509	M	BA	30	Class II	33	DES	p.Glu262Val	c.785A>T	missense	*Class 3 LB*	Borderline BS1 and borderline BP4
101301509	M	BA	30	Class II	33	ANKRD1	p.Lys132AsnfsTer4	c.396del	frameshift	Class 3	Gene not relevant in context
101301511	F	BA	43	Class III	48	TTN	p.Phe15649Leu	c.46945T>C	missense	Class 3LB	TTN missense not reviewed
101301511	F	BA	43	Class III	48	TTN	p.Ser6627Leu	c.19880C>T	missense	Class 3LB	TTN missense not reviewed
101301518	M	BA	38	Class II	21	*TTN*	p.Lys11470Glu	c.34408A>G	missense	Class 3LB	TTN missense not reviewed
101301518	M	BA	38	Class II	21	*TTN*	p.Lys621Glu	c.1861A>G	missense	Class 3LB	TTN missense not reviewed
101301518	M	BA	38	Class II	21	*TTN*	p.Ala352Thr	c.1054G>A	missense	Class 3LB	TTN missense not reviewed
101301523	M	MA	35	Class III	36	*DSP*	p.Asp468Asn	c.1402G>A	missense	Class 3LB	PM2 Strong , PM1 Moderate
101301527	M	MA	61	Class IV	19	*DSP*	p.Gln678Arg	c.2033A>G	missense	Class 3	PM2, PM1
101301527	M	MA	61	Class IV	19	*DSP*	p.Arg1341His	c.4022G>A	missense	Class 3LP	PM2 Strong, PM5 Supporting, PP3 Supporting
101301527	M	MA	61	Class IV	19	TTN	p.Gly24659Ala	c.73976G>C	missense	Class 3LB	TTN missense not reviewed
101301527	M	MA	61	Class IV	19	TTN	p.Ser6627Leu	c.19880C>T	missense	Class 3LB	TTN missense not reviewed
101301529	M	MA	59	Class I	40	DES	p.Arg375Trp	c.1123C>T	missense	Class 3LP	PM1 Moderate, PM2 Suppporting, PP3 Supporting
101301542	F	BA	43	Class II	56	*DES*	p.Glu262Val	c.785A>T	missense	Class 3	Borderline BS1 and borderline BP4
101301550	M	MA	21	Class I	59	TTN	p.Asp1232Val	c.3695A>T	missense	Class 3LB	TTN missense not reviewed
101301550	M	MA	21	Class I	59	DMD	p.Glu2910_Asn2912delinsValValA	c.8729_8734delinsTGGTCG	missense	Class 3LB	PM2 Strong, PM1 Moderate
101301558	M	BA	30	Class II	34	MYBPC3	p.Gly507Arg	c.1519G>A	missense	*Class 3 LB*	PM2 Strong, PP5 Strong, PM1 Moderate, PP3 Supporting
101301558	M	BA	30	Class II	34	DMD	p.Asp2112Tyr	c.6334G>T	missense	*Class 3 LB*	Borderline BS1 and borderline BP4
101301559	F	MA	40	Class II	32	SCN5A	p.Arg1847His	c.5540G>A	missense	Class 3LP	PM1 Moderate, PM2 Suppporting, PP3 Supporting
101301559	F	MA	40	Class II	32	DMD	p.Asp2112Tyr	c.6334G>T	missense	*Class 3 LB*	Borderline BS1 and borderline BP4
101301576	F	WA	0			TTN	p.Pro14761Ser	c.44281C>T	missense	Class 3LP	TTN missense not reviewed
101301576	F	WA	0			RBM20	p.Pro48del	c.141_143del	inframe_deletion	Class 3	Conflicting
101304303	M	MA	30			SCN5A	p.Asp197Tyr	c.589G>T	missense	Class 3LP	PM2 Strong, PM5 Moderate, PP3 Supporting
102302006	F	BA	0			TTN	p.Phe15649Leu	c.46945T>C	missense	Class 3LB	TTN missense not reviewed
102302006	F	BA	0			TTN	p.Val5230Ile	c.15688G>A	missense	Class 3LB	TTN missense not reviewed

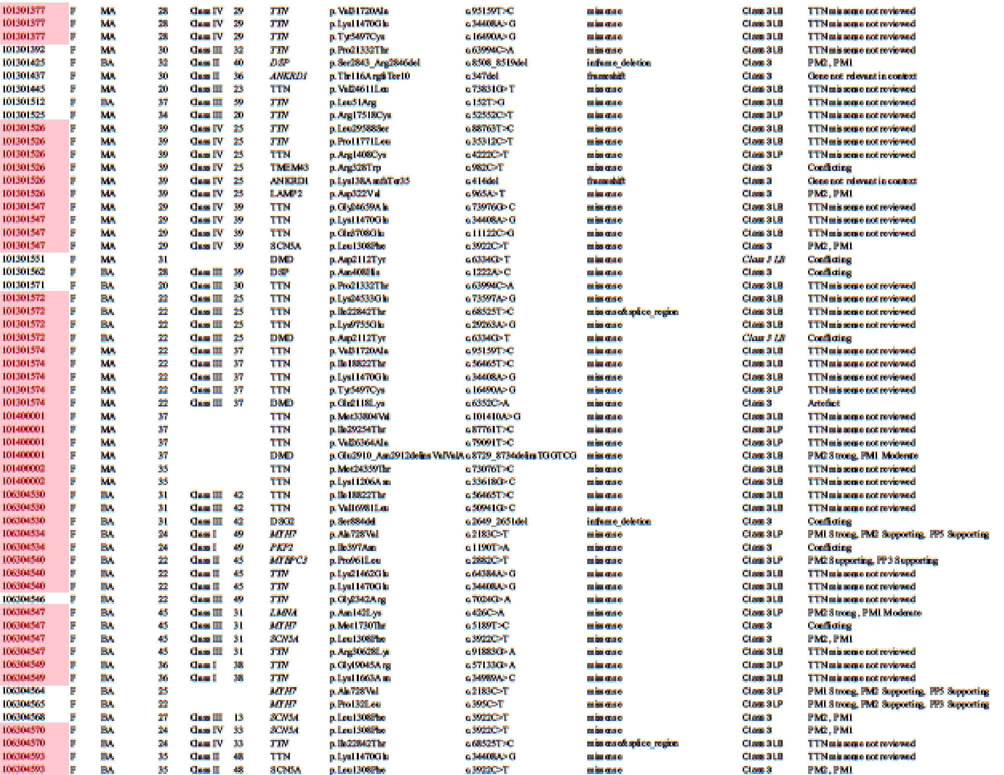
102302018	F	MA	0	*MYH7*	p.Ala728Val	c.2183C>T	missense	Class 3 LP	PM1 Strong, PM2 Supporting, PP5 Supporting
102302026	F	BA	1	*DSP*	p.Gln678Arg	c.2033A>G	missense	Class 3	PM2, PM1
102302026	F	BA	1	*DSP*	p.Arg1341His	c.4022G>A	missense	Class 3 LP	PM2 Strong, PM5 Supporting, PP3 Supporting
102302031	F	BA	1	*SCN5A*	p.Asp2003Asn	c.6007G>A	missense	Class 3 LP	PM2 Strong, PP3 Supporting
102302031	F	BA	1	*DSP*	p.Gln678Arg	c.2033A>G	missense	Class 3	PM2, PM1
102302031	F	BA	1	*DSP*	p.Arg1341His	c.4022G>A	missense	Class 3 LP	PM2 Strong, PM5 Supporting, PP3 Supporting
102302031	F	BA	1	*TTN*	p.Gly19031Glu	c.57092G>A	missense	Class 3 LB	TTN missense not reviewed
102302031	F	BA	1	*TTN*	p.Lys11470Glu	c.34408A>G	missense	Class 3 LB	TTN missense not reviewed
102302035	F	BA	3	*TTN*	p.Ser19893Ala	c.59677T>G	missense	Class 3 LB	TTN missense not reviewed
102302036	M	BA	0	DES	p.Glu262Val	c.785A>T	missense	Class 3	Borderline BS1 and borderline BP4
102302036	M	BA	0	BAG3	p.Thr568Ile	c.1703C>T	missense	Class 3	PVS1 , PM1 Moderate
102302038	M	BA	6	*TTN*	p.Glu11001Lys	c.33001G>A	missense	Class 3 LB	TTN missense not reviewed
102302039	F	BA	1	TTN	p.Ile32009Thr	c.96026T>C	missense	Class 3 LB	TTN missense not reviewed
102302039	F	BA	1	PKP2	p.Ile397Asn	c.1190T>A	missense	Class 3	Conflicting
102302041	F	MA	14	TTN	p.Val21602Leu	c.64804G>T	missense	Class 3 LB	TTN missense not reviewed
102302041	F	MA	14	TTN	p.Cys18472Gly	c.55414T>G	missense	Class 3 LB	TTN missense not reviewed
102302041	F	MA	14	DSP	p.Gln678Arg	c.2033A>G	missense	Class 3	PM2, PM1
102302041	F	MA	14	DSP	p.Arg1341His	c.4022G>A	missense	Class 3 LP	PM2 Strong, PM5 Supporting, PP3 Supporting
102302041	F	MA	14	MYH7	p.Asn977Ser	c.2930A>G	missense	Class 3 LP	PM2 Strong, PM1 Moderate
102302050	F	BA	0	*FLNC*	p.Phe726Ser	c.2177T>C	missense	Class 3	PM2, PM1
102302051	M	BA	3	TTN	p.Arg29578His	c.88733G>A	missense	Class 3 LB	TTN missense not reviewed
102302051	M	BA	3	LAMP2	p.Ala2980Thr	c.8938G>A	missense	Class 3	Conflicting
103302501	F	BA	2	*DMD*	p.Asp538Val	c.1613A>T	missense	Class 3	Conflicting
103302501	F	BA	2	*TTN*	p.Val24611Leu	c.73831G>T	missense	Class 3 LB	TTN missense not reviewed
103302504	F	BA	0	*CDH2*	p.Lys158Asn	c.474G>T	missense	Class 3	PM2, PM1
103302504	F	BA	0	*TTN*	Splice Site	c.11254+2T>C	splice_donor	*Class 3 LB*	putative truncating variants not in 100% PSI regions
105303501	F	BA	2	*ANKRD1*	p.Pro134GlnfsTer2	c.401del	frameshift	Class 3	Gene not relevant in context
105303501	F	BA	2	*FLNC*	p.Asn1294Ile	c.3881A>T	missense	Class 3	Conflicting
105303501	F	BA	2	*TTN*	p.Lys27133Gln	c.81397A>C	missense	Class 3 LB	TTN missense not reviewed
105303502	F	BA	1	*SCN5A*	p.Leu1308Phe	c.3922C>T	missense	Class 3	PM2, PM1
105303503	M	BA	2	MYH7	p.Leu985TrpfsTer8	c.2952del	frameshift	Class 3 LP	PM2 Strong, PM1 Moderate
105303503	M	BA	2	TTN*	p.Lys11470Glu	c.34408A>G	missense	Class 3 LB	TTN missense not reviewed
105303504	F	BA	1	*DSP*	p.Gln678Arg	c.2033A>G	missense	Class 3	PM2, PM1
105303504	F	BA	1	*DSP*	p.Arg1341His	c.4022G>A	missense	Class 3 LP	PM2 Strong, PM5 Supporting, PP3 Supporting
105303508	M	BA	11	DSC2	p.Gly811Val	c.2432G>T	missense	Class 3	PM2, PM1
105303508	M	BA	11	TTN	p.Val22404Leu	c.67210G>C	missense	Class 3 LB	TTN missense not reviewed
105303508	M	BA	11	TTN	p.Arg17196Thr	c.51587G>C	missense	Class 3 LP	TTN missense not reviewed
105303509	M	BA	2	*DMD*	p.Asp15His	c.43G>C	missense	Class 3	Conflicting
105303509	M	BA	2	*MYL3*	p.Ala2Pro	c.4G>C	missense	Class 3	Conflicting
105303511	M	BA	8	*TTN*	p.Ser8409Tyr	c.25226C>A	missense	Class 3 LB	TTN missense not reviewed
105303514	F	BA	9	*DSP*	p.Gln678Arg	c.2033A>G	missense	Class 3	PM2, PM1
105303514	F	BA	9	*DSP*	p.Arg1341His	c.4022G>A	missense	Class 3 LP	PM2 Strong, PM5 Supporting, PP3 Supporting
105303514	F	BA	9	*PKP2*	p.Ile397Asn	c.1190T>A	missense	Class 3	Conflicting
105303514	F	BA	9	*SCN5A*	p.Leu1308Phe	c.3922C>T	missense	Class 3	PM2, PM1
105303520	F	BA	2	*SCN5A*	p.Leu1308Phe	c.3922C>T	missense	Class 3	PM2, PM1
105303520	F	BA	2	*TTN*	Splice Site	c.11254+2T>C	splice_donor	*Class 3 LB*	putative truncating variants not in 100% PSI regions
105303520	F	BA	2	*TTN*	p.Pro20395Thr	c.61183C>A	missense	Class 3 LB	TTN missense not reviewed
105303525	F	BA	5	*DSP*	p.Gln678Arg	c.2033A>G	missense	Class 3	PM2, PM1
105303525	F	BA	5	*DSP*	p.Arg1341His	c.4022G>A	missense	Class 3 LP	PM2 Strong, PM5 Supporting, PP3 Supporting
105303525	F	BA	5	*TTN*	p.Ile35986Val	c.107956A>G	missense	Class 3 LB	TTN missense not reviewed
105303525	F	BA	5	*TTN*	p.Gly19031Glu	c.57092G>A	missense	Class 3 LB	TTN missense not reviewed
105303525	F	BA	5	*TTN*	p.Ile18822Thr	c.56465T>C	missense	Class 3 LB	TTN missense not reviewed
105303525	F	BA	5	*TTN*	p.Lys11470Glu	c.34408A>G	missense	Class 3 LB	TTN missense not reviewed
105303525	F	BA	5	*TTN*	p.Asp9656Asn	c.28966G>A	missense	Class 3 LB	TTN missense not reviewed
105303525	F	BA	5	*TTN*	p.Ala352Thr	c.1054G>A	missense	Class 3 LB	TTN missense not reviewed
105303551	F	BA	10	DES	p.Asp312Asn	c.934G>A	missense	Class 3 LP	PM2 Strong, PM1 Moderate, PP3 Supporting
105303551	F	BA	10	DSP	p.Gln678Arg	c.2033A>G	missense	Class 3	PM2, PM1
105303551	F	BA	10	DSP	p.Arg1341His	c.4022G>A	missense	Class 3 LP	PM2 Strong, PM5 Supporting, PP3 Supporting
105303561	F	MA	3	DSP	p.Gln678Arg	c.2033A>G	missense	Class 3	PM2, PM1

105303561	F	MA	3			DSP	p.Arg1341His	c.4022G>A	missense	Class 3 LP	PM2 Strong, PM5 Supporting, PP3 Supporting
105303562	M	BA	5			TTN	p.Gly24659Ala	c.73976G>C	missense	Class 3 LB	TTN missense not reviewed
105303562	M	BA	5			TTN	p.Pro23850Ala	c.71548C>G	missense	Class 3 LB	TTN missense not reviewed
105303562	M	BA	5			TTN	p.Lys11470Glu	c.34408A>G	missense	Class 3 LB	TTN missense not reviewed
105303562	M	BA	5			TTN	p.Gln3708Glu	c.11122C>G	missense	Class 3 LB	TTN missense not reviewed
105303562	M	BA	5			MYBPC3	p.Ser25Asn	c.74G>A	missense	Class 3	Conflicting
105303562	M	BA	5			MYL2	p.Val61Met	c.181G>A	missense	Class 3	Conflicting
105303564	M	BA	6			*DSP*	p.Gln678Arg	c.2033A>G	missense	Class 3	PM2, PM1
105303564	M	BA	6			*DSP*	p.Arg1341His	c.4022G>A	missense	Class 3 LP	PM2 Strong, PM5 Supporting, PP3 Supporting
105303565	M	BA	10			TTN	p.Val22640Leu	c.67918G>C	missense	Class 3 LB	TTN missense not reviewed
105303565	M	BA	10			DSP	p.Asp406Glu	c.1218C>G	missense	Class 3	PM2, PM1
105303565	M	BA	10			ANKRD1	p.Lys132AsnfsTer4	c.396del	frameshift	Class 3	Gene not relevant in context
105303572	F	MA	14			*TTN*	p.Gly24659Ala	c.73976G>C	missense	Class 3 LB	TTN missense not reviewed
105303572	F	MA	14			*TTN*	p.Lys11470Glu	c.34408A>G	missense	Class 3 LB	TTN missense not reviewed
105303572	F	MA	14			*TTN*	p.Gln3708Glu	c.11122C>G	missense	Class 3 LB	TTN missense not reviewed
105304501	F	BA	10			*DSP*	p.Ser2843_Arg2846del	c.8508_8519del	inframe_deletion	Class 3	PM2, PM1
105304501	F	BA	10			*TTN*	p.Arg30628Lys	c.91883G>A	missense	Class 3 LB	TTN missense not reviewed
106304502	F	MA	7			*DSP*	p.Gln678Arg	c.2033A>G	missense	Class 3	PM2, PM1
106304502	F	MA	7			*DSP*	p.Arg1341His	c.4022G>A	missense	Class 3 LP	PM2 Strong, PM5 Supporting, PP3 Supporting
106304502	F	MA	7			*TTN*	p.Thr34998Ile	c.104993C>T	missense	Class 3 LB	TTN missense not reviewed
106304502	F	MA	7			*TTN*	p.Ile20956Thr	c.62867T>C	missense	Class 3 LB	TTN missense not reviewed
106304504	M	BA	2			*VCL*	p.Val870Ile	c.2608G>A	missense	Class 3	Conflicting
106304504	M	BA	2			*TTN*	p.Gly25703Arg	c.77107G>A	missense	Class 3 LB	TTN missense not reviewed
106304504	M	BA	2			*TTN*	p.Arg24163Cys	c.72487C>T	missense	Class 3 LB	TTN missense not reviewed
106304533	M	MA	42	Class II	29	*TTN*	p.Trp29581Cys	c.88743G>T	missense	Class 3 LB	TTN missense not reviewed
106304535	F	BA	53	Class II	20	*DMD*	p.Asp15His	c.43G>C	missense	Class 3	Conflicting
106304535	F	BA	53	Class II	20	*MYBPC3*	p.Pro961Leu	c.2882C>T	missense	Class 3 LP	PM2 Moderate, PP3 Moderate
106304535	F	BA	53	Class II	20	*MYL3*	p.Ala2Pro	c.4G>C	missense	Class 3	Conflicting
106304548	F	BA	58	Class II	12	TTN	p.Thr31234Ile	c.93701C>T	missense	Class 3 LB	TTN missense not reviewed
106304548	F	BA	58	Class II	12	TTN	p.Gly24659Ala	c.73976G>C	missense	Class 3 LB	TTN missense not reviewed
106304548	F	BA	58	Class II	12	TTN	p.Lys11470Glu	c.34408A>G	missense	Class 3 LB	TTN missense not reviewed
106304551	M	WA	46			*FLNC*	p.Ser2546Thr	c.7636T>A	missense	Class 3	Conflicting
106304553	F	BA	17			*DES*	p.Asp312Ala	c.935A>C	missense	Class 3 LP	PM2 Strong, PM1 Moderate, PP3 Supporting
106304554	M	MA	18			*DSP*	p.Pro2618Ser	c.7852C>T	missense	Class 3 LP	PM2 Strong, PM5 Supporting, PP3 Supporting
106304554	M	MA	18			*FLNC*	p.Arg1434Cys	c.4300C>T	missense	Class 3	Conflicting
106304554	M	MA	18			*MYH7*	p.Ala728Val	c.2183C>T	missense	Class 3 LP	PM1 Strong, PM2 Supporting, PP5 Supporting
106304558	F	MA	34			*TTN*	Splice Site	c.32471-1G>A	splice_acceptor	Class 3 LP	PVS1 Very Strong, PM2 Moderate
106304558	F	MA	34			TTN	p.Lys28450Asn	c.85350G>C	missense	Class 3 LB	TTN missense not reviewed
106304558	F	MA	34			TTN	p.Ser19018Ala	c.57052T>G	missense	Class 3 LB	TTN missense not reviewed
106304558	F	MA	34			TTN	p.Cys9226Tyr	c.27677G>A	missense	Class 3 LB	TTN missense not reviewed
106304558	F	MA	34			DSP	p.Arg1341His	c.4022G>A	missense	Class 3 LP	PM2 Strong, PM5 Supporting, PP3 Supporting
106304558	F	MA	34			MYBPC3	p.Glu1179Lys	c.3535G>A	missense	Class 3	Conflicting
106304561	F	WA	39			*TTN*	p.Lys15556del	c.46666_46668del	inframe_deletion	Class 3	small in frame deletions
106304572	F	BA	22	Class III	20	*ACTN2*	p.Glu549Val	c.1646A>T	missense	Class 3 LP	PVS1 Very Strong, PM2 Supporting
106304572	F	BA	22	Class III	20	*FLNC*	p.Lys162Glu	c.484A>G	missense	Class 3 LP	PM2 Strong, PM1 Moderate
106304578	M	BA	47	Class IV	16	*TTN*	p.Thr21773Ile	c.65318C>T	missense	Class 3 LB	TTN missense not reviewed
106304578	M	BA	47	Class IV	16	*SCN5A*	p.Glu1832Gln	c.5494C>G	missense	*Class 3 LB*	Borderline BS1 and borderline BP4
106304580	F	BA	24	Class II	27	DES	p.Asp312Ala	c.935A>C	missense	Class 3 LP	PM2 Strong, PM1 Moderate, PP3 Supporting
106304581	F	BA	45	Class II	20	*LAMP2*	p.Asp322Val	c.965A>T	missense	Class 3	PM2, PM1
106304583	F	BA	20	Class I	39	*TTN*	p.Arg24888Thr	c.74663G>C	missense	Class 3 LB	TTN missense not reviewed
106304588	M	MA	39	Class III	15	*DSP*	p.Gln678Arg	c.2033A>G	missense	Class 3	PM2, PM1
106304591	F	BA	37	Class II	64	ALPK3	p.Val302Ile	c.904G>A	missense	Class 3	BA1
106304591	F	BA	37	Class II	64	LAMP2	p.Asp322Val	c.965A>T	missense	Class 3	PM2, PM1
106304595	M	MA	55	Class II	29	*TMEM43*	p.Arg328Trp	c.982C>T	missense	Class 3	Conflicting
106304596	M	BA	54			*TTN*	p.Leu24076Ser	c.72227T>C	missense	Class 3 LB	TTN missense not reviewed
106304596	M	BA	54			*SCN5A*	p.Leu1308Phe	c.3922C>T	missense	Class 3	PM2, PM1
106304598	M	WA	54	Class I	35	*TTN*	p.Arg24742His	c.74225G>A	missense	Class 3 LB	TTN missense not reviewed
106304598	M	WA	54	Class I	35	*ALPK3*	p.Gly220Trp	c.658G>T	missense	Class 3	BA1
106304599	F	MA	40	Class II	56	*FHL1*	p.Ser130Gly	c.388A>G	missense	Class 3	Conflicting

401300303	F	BA	48	Class IV	25	*TTN*	p.Glu5146Gly	c.15437A>G	missense	Class 3LB	TTN missense not reviewed
401300308	M	BA	30	Class III	22	*DES*	p.Asp312Asn	c.934G>A	missense	Class 3LP	PM2 Strong, PM1 Moderate, PP3 Supporting
401300308	M	BA	30	Class III	22	*TTN*	p.Arg34949His	c.104846G>A	missense	Class 3LB	TTN missense not reviewed
401300308	M	BA	30	Class III	22	*TTN*	p.Leu29588Ser	c.88763T>C	missense	Class 3LB	TTN missense not reviewed
401300308	M	BA	30	Class III	22	*TTN*	p.Pro16152Ser	c.48454C>T	missense	Class 3LB	TTN missense not reviewed
401300308	M	BA	30	Class III	22	*TTN*	p.Ser9464Phe	c.28391C>T	missense	Class 3LB	TTN missense not reviewed
401303001	M	BA	35	Class IV	25	*DES*	p.Asp312Asn	c.934G>A	missense	Class 3LP	PM2 Strong, PM1 Moderate, PP3 Supporting
401303001	M	BA	35	Class IV	25	*TTN*	p.Arg31185Cys	c.93553C>T	missense	Class 3LB	TTN missense not reviewed
401303001	M	BA	35	Class IV	25	*TTN*	p.Cys9195Ser	c.27584G>C	missense	Class 3LB	TTN missense not reviewed
401303001	M	BA	35	Class IV	25	*TTN*	p.Ile5736Thr	c.17207T>C	missense	Class 3LB	TTN missense not reviewed
401303003	F	BA	20	Class III	36	*TTN*	p.Ala34064Val	c.102191C>T	missense	Class 3LB	TTN missense not reviewed
401303003	F	BA	20	Class III	36	*TTN*	p.Val24611Leu	c.73831G>T	missense	Class 3LB	TTN missense not reviewed
401303003	F	BA	20	Class III	36	*TTN*	p.Lys11470Glu	c.34408A>G	missense	Class 3LB	TTN missense not reviewed
401303003	F	BA	20	Class III	36	*TTN*	p.Leu3554His	c.10661T>A	missense	Class 3LB	TTN missense not reviewed
401303004	M	BA	50	Class IV	29	*TTN*	p.Val24611Leu	c.73831G>T	missense	Class 3LB	TTN missense not reviewed
401303010	F	BA	33	Class III	34	*TTN*	p.Arg30628Lys	c.91883G>A	missense	Class 3LB	TTN missense not reviewed
401303013	M	MA	51	Class IV	25	TTN	p.Pro23850Ala	c.71548C>G	missense	Class 3LB	TTN missense not reviewed
401303013	M	MA	51	Class IV	25	TTN	p.Lys11470Glu	c.34408A>G	missense	Class 3LB	TTN missense not reviewed
401303013	M	MA	51	Class IV	25	TTN	p.Leu3554His	c.10661T>A	missense	Class 3LB	TTN missense not reviewed
401303017	M	BA	73	Class III	30	*TTN*	p.Arg30628Lys	c.91883G>A	missense	Class 3LB	TTN missense not reviewed
401303022	F	BA	25	Class III	25	*ALPK3*	p.Glu148GlnfsTer8	c.436_440dup	frameshift	Class 3LP	PVS1 Strong, PM2 Moderate
401303025	M	BA	25	Class III	28	*TTN*	p.Arg30628Lys	c.91883G>A	missense	Class 3LB	TTN missense not reviewed
401303029	F	WA	53	Class II	28	*TTN*	Splice Site	c.32471-1G>A	splice_acceptor	Class 3LP	PVS1 Very Strong, PM2 Moderate
401303029	F	WA	53	Class II	28	*TTN*	p.Lys28450Asn	c.85350G>C	missense	Class 3LB	TTN missense not reviewed
401303029	F	WA	53	Class II	28	*TTN*	p.Ser19018Ala	c.57052T>G	missense	Class 3LB	TTN missense not reviewed
401303029	F	WA	53	Class II	28	*TTN*	p.Cys9226Tyr	c.27677G>A	missense	Class 3LB	TTN missense not reviewed
401303030	M	BA	35	Class II	22	*TTN*	p.Asp35153Val	c.105458A>T	missense	Class 3LB	TTN missense not reviewed
401303030	M	BA	35	Class II	22	*DMD**	p.Ser810Arg	c.2428A>C	missense	Class 3	Conflicting
401303031	F	BA	32	Class III	23	*TTN*	p.Arg30628Lys	c.91883G>A	missense	Class 3LB	TTN missense not reviewed
401303034	M	BA	55	Class III	33	*TTN*	p.Phe15649Leu	c.46945T>C	missense	Class 3LB	TTN missense not reviewed
401303036	M	BA	47	Class III	23	*TTN*	p.Gly24659Ala	c.73976G>C	missense	Class 3LB	TTN missense not reviewed
401303036	M	BA	47	Class III	23	*TTN*	p.Lys11470Glu	c.34408A>G	missense	Class 3LB	TTN missense not reviewed
401303036	M	BA	47	Class III	23	*TTN*	p.Gln3708Glu	c.11122C>G	missense	Class 3LB	TTN missense not reviewed
401303036	M	BA	47	Class III	23	*ALPK3*	p.Val302Ile	c.904G>A	missense	Class 3	BA1
401303037	M	BA	32	Class III	12	*TTN*	p.Val24611Leu	c.73831G>T	missense	Class 3LB	TTN missense not reviewed
401303038	M	MA	36	Class III	38	*FLNC*	p.Ser625Cys	c.1874C>G	missense	Class 3	PM2, PM1
401303039	M	BA	27	Class IV	17	*ALPK3*	p.Val302Ile	c.904G>A	missense	Class 3	BA1
401303044	M	BA	27	Class IV	20	*MYBPC3*	p.Glu1179Lys	c.3535G>A	missense	Class 3	Conflicting
401303048	M	BA	55	Class III	20	*TTN*	p.Gly24659Ala	c.73976G>C	missense	Class 3LB	TTN missense not reviewed
401303048	M	BA	55	Class III	20	*DES*	p.Asp312Ala	c.935A>C	missense	Class 3LP	PM2 Strong, PM1 Moderate, PP3 Supporting
401303049	F	BA	47	Class II	25	*TTN*	p.Val379Met	c.1135_1143delinsATGAAAT(...	missense	Class 3LB	TTN missense not reviewed
401303049	F	BA	47	Class II	25	*TTN*	p.Arg34949His	c.104846G>A	missense	Class 3LB	TTN missense not reviewed
401303049	F	BA	47	Class II	25	*TTN*	p.Leu29588Ser	c.88763T>C	missense	Class 3LB	TTN missense not reviewed
401303049	F	BA	47	Class II	25	*TTN*	p.Pro16152Ser	c.48454C>T	missense	Class 3LB	TTN missense not reviewed
401303054	M	BA	25	Class III	14	DMD	p.Glu2910_Asn2912delinsValValAs...	c.8729_8734delinsTGGTCG	missense	Class 3LB	PM2 Strong, PM1 Moderate
401303054	M	BA	25	Class III	14	*TTN*	p.Arg30628Lys	c.91883G>A	missense	Class 3LB	TTN missense not reviewed
401303054	M	BA	25	Class III	14	*TTN*	p.Val21602Leu	c.64804G>T	missense	Class 3LB	TTN missense not reviewed
LVNC											
10200009	F	BA	1			*TPM1*	p.Asp14Gly	c.41A>G	missense	Class 3	Conflicting
10200009	F	BA	1			*TTN*	p.Lys11470Glu	c.34408A>G	missense	Class 3LB	TTN missense not reviewed
10200012	M	BA	1			*ALPK3*	p.Val302Ile	c.904G>A	missense	Class 3	BA1
10200012	M	BA	1			*PKP2*	p.Ile397Asn	c.1190T>A	missense	Class 3	Conflicting
10300009	M	MA	0			*MYH7*	p.Gly178Arg	c.532G>A	missense	Class 3LP	PM1 Strong, PM2 Strong, PP3 Strong
10300009	M	MA	0			*VCL*	p.Val870Ile	c.2608G>A	missense	Class 3	Conflicting
10300009	M	MA	0			*TTN*	p.Gly25703Arg	c.77107G>A	missense	Class 3LB	TTN missense not reviewed
10300009	M	MA	0			*TTN*	p.Arg24163Cys	c.72487C>T	missense	Class 3LB	TTN missense not reviewed
102301235	F	BA	0			FLNC	p.Arg334Cys	c.1000C>T	missense	Class 3	Conflicting
102302012	F	MA	0			*DSP*	p.Gln678Arg	c.2033A>G	missense	Class 3	PM2, PM1
102302012	F	MA	0			*DSP*	p.Arg1341His	c.4022G>A	missense	Class 3LP	PM2 Strong, PM5 Supporting, PP3 Supporting

102302012	F	MA	0			*MYH7*	p.Glu1693Lys	c.5077G>A	missense	Class 3	Conflicting
102302016	F	BA	0			*MYH7*	p.Arg1897Gly	c.5689C>G	missense	Class 3	PM2, PM1
102302017	M	MA	0			*ACTC1*	p.Pro34Arg	c.101C>G	missense	Class 3	PM2_supporting, PP3_supporting, PP2
102302017	M	MA	0			*TTN*	p.Val24611Leu	c.73831G>T	missense	Class 3 LB	TTN missense not reviewed
102302017	M	MA	0			*TTN*	p.Glu11001Lys	c.33001G>A	missense	Class 3 LB	TTN missense not reviewed
102302020	M	MA	1			*MYH7*	p.Cys538Arg	c.1612T>C	missense	Class 3 LP	PM1 Strong, PM2 Strong, PP3 Supporting
102302025	M	MA	2			*SCN5A*	p.Leu1308Phe	c.3922C>T	missense	Class 3	PM2, PM1
102302030	M	MA	1			*SCN5A*	p.Arg971Cys	c.2911C>T	missense	Class 3 LP	PM2 Supporting, PP5 Supporting, PP3 Supporting
102302033	F	BA	1			MYH7	p.Ala728Val	c.2183C>T	missense	Class 3 LP	PM1 Strong, PM2 Supporting, PP5 Supporting
102302033	F	BA	1			TTN	p.Gly24659Ala	c.73976G>C	missense	Class 3 LB	TTN missense not reviewed
102302033	F	BA	1			TTN	p.Val21602Leu	c.64804G>T	missense	Class 3 LB	TTN missense not reviewed
102302033	F	BA	1			TTN	p.Cys18472Gly	c.55414T>G	missense	Class 3 LB	TTN missense not reviewed
102302033	F	BA	1			TTN	p.Ser6627Leu	c.19880C>T	missense	Class 3 LB	TTN missense not reviewed
102302033	F	BA	1			SCN5A	p.Leu1308Phe	c.3922C>T	missense	Class 3	PM2, PM1
102302040	M	MA	1			TTN	p.Val24611Leu	c.73831G>T	missense	Class 3 LB	TTN missense not reviewed
102302040	M	MA	1			TTN	p.Glu11001Lys	c.33001G>A	missense	Class 3 LB	TTN missense not reviewed
102302040	M	MA	1			ACTC1	p.Pro34Arg	c.101C>G	missense	Class 3	PM2_supporting, PP3_supporting, PP2
105303574	M	BA	10			TTN	p.Val24611Leu	c.73831G>T	missense	Class 3 LB	TTN missense not reviewed
105303574	M	BA	10			TTN	p.Lys11470Glu	c.34408A>G	missense	Class 3 LB	TTN missense not reviewed
105303574	M	BA	10			TTN	p.Leu3554His	c.10661T>A	missense	Class 3 LB	TTN missense not reviewed
105303574	M	BA	10			DSP	p.Ser2843_Arg2846del	c.8508_8519del	inframe_deletion	Class 3	PM2, PM1
105303574	M	BA	10			MYBPC3	p.Glu1179Lys	c.3535G>A	missense	*Class 3 LB*	Borderline BS1 and borderline BP4
PPCM											
10100022	F	MA	23	Class III	27	*MYBPC3*	p.Glu1179Lys	c.3535G>A	missense	Class 3	Conflicting
10100023	F	BA	45	Class II	21	*TTN*	p.Ile32009Thr	c.96026T>C	missense&splice_region	Class 3 LB	TTN missense not reviewed
10100023	F	BA	45	Class II	21	*TTN*	p.Met1064Leu	c.3190A>T	missense	Class 3 LB	TTN missense not reviewed
10100025	F	BA	29	Class III	31	*DES*	p.Asp312Ala	c.935A>C	missense	Class 3 LP	PM2 Strong, PM1 Moderate, PP3 Supporting
10100025	F	BA	29	Class III	31	*VCL*	p.Pro943ArgfsTer9	c.2828_2829del	frameshift	Class 3	Conflicting
10100027	F	BA	38	Class III	19	*DMD*	p.Ser810Arg	c.2428A>C	missense	Class 3	Conflicting
10100032	F	BA	33	Class III	37	*ALPK3*	p.Val302Ile	c.904G>A	missense	Class 3	BA1
10100042	F	MA	25	Class II	39	*ACTN2*	p.Asn763Ser	c.2288A>G	missense	Class 3	PM2, PM1
10100051	F	MA	25	Class III	26	*TTN*	p.Gly19031Glu	c.57092G>A	missense	Class 3 LB	TTN missense not reviewed
10100051	F	MA	25	Class III	26	*TTN*	p.Lys11470Glu	c.34408A>G	missense	Class 3 LB	TTN missense not reviewed
10100051	F	MA	25	Class III	26	*TTN*	p.Asp9656Asn	c.28966G>A	missense	Class 3 LB	TTN missense not reviewed
10100051	F	MA	25	Class III	26	*TTN*	p.Ala352Thr	c.1054G>A	missense	Class 3 LB	TTN missense not reviewed
10100052	F	MA	41	Class III	31	*TTN*	p.Trp29873Gly	c.89617T>G	missense	Class 3 LB	TTN missense not reviewed
10100057	F	BA	23	Class IV	31	*TTN*	p.Ser8409Tyr	c.25226C>A	missense	Class 3 LB	TTN missense not reviewed
10100078	F	MA	34	Class III	22	*LAMP2*	p.Asp322Val	c.965A>T	missense	Class 3	PM2, PM1
10100080	F	BA	28	Class IV	25	*MYL2*	p.Val61Met	c.181G>A	missense	Class 3	Conflicting
10100084	F	BA	29	Class IV	21	*FLNC*	p.Val2205Met	c.6613G>A	missense	Class 3	PM2, PM1
10100084	F	BA	29	Class IV	21	*SCN5A*	p.Gly77Arg	c.229G>A	missense	Class 3	Conflicting
10100084	F	BA	29	Class IV	21	*TTN*	p.Gly35974Arg	c.107920G>A	missense	Class 3 LB	TTN missense not reviewed
10100084	F	BA	29	Class IV	21	*TTN*	p.Val34890Leu	c.104668G>C	missense	Class 3 LB	TTN missense not reviewed
10100085	F	MA	33	Class III	17	*MYBPC3*	p.Glu1179Lys	c.3535G>A	missense	Class 3	Conflicting
10100342	F	BA	26	Class III	12	*DSP*	p.Gln678Arg	c.2033A>G	missense	Class 3	PM2, PM1
10100342	F	BA	26	Class III	12	*DSP*	p.Arg1341His	c.4022G>A	missense	Class 3 LP	PM2 Strong, PM5 Supporting, PP3 Supporting
10100342	F	BA	26	Class III	12	*TTN*	p.Gly24659Ala	c.73976G>C	missense	Class 3 LB	TTN missense not reviewed
10100342	F	BA	26	Class III	12	*TTN*	p.Lys11470Glu	c.34408A>G	missense	Class 3 LB	TTN missense not reviewed
10100342	F	BA	26	Class III	12	*TTN*	p.Gln3708Glu	c.11122C>G	missense	Class 3 LB	TTN missense not reviewed
10100348	F	BA	41	Class IV	23	*PKP2*	p.Ile397Asn	c.1190T>A	missense	Class 3	Conflicting
10100353	F	BA	37	ND	30	TTN	p.Val24611Leu	c.73831G>T	missense	Class 3 LB	TTN missense not reviewed
10100353	F	BA	37	ND	30	TTN	p.Lys11470Glu	c.34408A>G	missense	Class 3 LB	TTN missense not reviewed
10100353	F	BA	37	ND	30	TTN	p.Leu3554His	c.10661T>A	missense	Class 3 LB	TTN missense not reviewed
10100353	F	BA	37	ND	30	TTN	p.Asn3026Ile	c.9077A>T	missense	Class 3 LB	TTN missense not reviewed
10100359	F	BA	41	Class III	32	*DMD*	p.Ser810Arg	c.2428A>C	missense	Class 3	Conflicting
10100359	F	BA	41	Class III	32	TTN	p.Pro21332Thr	c.63994C>A	missense	Class 3 LB	TTN missense not reviewed
10100359	F	BA	41	Class III	32	TTN	p.Gly19031Glu	c.57092G>A	missense	Class 3 LB	TTN missense not reviewed
10100359	F	BA	41	Class III	32	TTN	p.Lys11470Glu	c.34408A>G	missense	Class 3 LB	TTN missense not reviewed
10100360	F	BA	39	Class III	40	*DMD*	p.Ser810Arg	c.2428A>C	missense	Class 3	Conflicting

10100360	F	BA	39	Class III	40	*TTN*	p.Ser810Arg	c.2428A>C	missense	Class 3 LB	TTN missense not reviewed
10100382	F	MA	22	ND	13	*PKP2*	p.Ile397Asn	c.1190T>A	missense	Class 3	Conflicting
10100394	F	BA	26	Class II	21	*BAG3*	p.Pro234Ser	c.700C>T	missense	Class 3 LP	PVS1 , PM1 Moderate
10100398	F	BA	33	ND	59	*SCN5A*	p.Arg34His	c.101G>A	missense	Class 3	Conflicting
10100398	F	BA	33	ND	59	*TTN*	p.Pro5009Ser	c.15025C>T	missense	Class 3 LB	TTN missense not reviewed
10100399	F	BA	16	Class II	35	*ANKRD1*	p.Pro122LeufsTer4	c.365del	frameshift	Class 3	Gene not relevant in context
10100400	F	BA	27			*ANKRD1*	p.Glu130ArgfsTer6	c.388del	frameshift	Class 3	Gene not relevant in context
10100400	F	BA	27			*TTN*	Splice Site	c.11254+2T>C	splice_donor	*Class 3 LB*	putative truncating variants not in 100% PSI regions
10100400	F	BA	27			*TTN*	p.Ile21696Thr	c.65087T>C	missense	Class 3 LB	TTN missense not reviewed
10100403	F	BA	23	Class II	20	*ANKRD1*	p.Glu117AsnfsTer9	c.349del	frameshift	Class 3	Gene not relevant in context
10100404	F	BA	30	Class III	23	*ALPK3*	p.Val302Ile	c.904G>A	missense	Class 3	BA1
10100404	F	BA	30	Class III	23	*DSP*	p.Ser1623Cys	c.4868C>G	missense	Class 3	Conflicting
10100404	F	BA	30	Class III	23	*FLNC*	p.Gly846Gln	c.1037G>A	missense	Class 3	PM2, PM1
10100404	F	BA	30	Class III	23	*TTN*	p.Val24611Leu	c.73831G>T	missense	Class 3 LB	TTN missense not reviewed
10100404	F	BA	30	Class III	23	*TTN*	p.Lys11470Glu	c.34408A>G	missense	Class 3 LB	TTN missense not reviewed
10100404	F	BA	30	Class III	23	*TTN*	p.Leu3554His	c.10661T>A	missense	Class 3 LB	TTN missense not reviewed
10100407	F	MA	32	Class III	46	*SCN5A*	p.Gly9Asp	c.26G>A	missense	Class 3	PM2, PM1
10100409	F	BA	40	Class IV	22	*TAZ*	p.Pro8Ser	c.22C>T	missense	Class 3	PM2, PM1
10100412	F	BA	25	Class II	26	*JUP*	p.Arg86Trp	c.256C>T	missense	Class 3 LP	PM2 Strong, PM1 Moderate
10100412	F	BA	25	Class II	26	*TTN*	p.Gly19031Glu	c.57092G>A	missense	Class 3 LB	TTN missense not reviewed
10100412	F	BA	25	Class II	26	*TTN*	p.Lys11470Glu	c.34408A>G	missense	Class 3 LB	TTN missense not reviewed
10100412	F	BA	25	Class II	26	*TTN*	p.Asp9656Asn	c.28966G>A	missense	Class 3 LB	TTN missense not reviewed
10100412	F	BA	25	Class II	26	*TTN*	p.Glu6382Lys	c.19144G>A	missense	Class 3 LB	TTN missense not reviewed
10100412	F	BA	25	Class II	26	*TTN*	p.Ala352Thr	c.1054G>A	missense	Class 3 LB	TTN missense not reviewed
10100418	F	MA	33	Class IV	19	*TTN*	p.Glu34579Lys	c.103735G>A	missense	Class 3 LB	TTN missense not reviewed
10100418	F	MA	33	Class IV	19	*TTN*	p.Gly24659Ala	c.73976G>C	missense	Class 3 LB	TTN missense not reviewed
10100421	F	WA	20	Class II	22	*TTN*	p.Val24611Leu	c.73831G>T	missense	Class 3 LB	TTN missense not reviewed
10100422	F	BA	28	Class III	23	*DES*	p.Asp312Asn	c.934G>A	missense	Class 3 LP	PM2 Strong, PM1 Moderate, PP3 Supporting
10100422	F	BA	28	Class III	23	*DSP*	p.Gln678Arg	c.2033A>G	missense	Class 3	PM2, PM1
10100422	F	BA	28	Class III	23	*DSP*	p.Arg1341His	c.4022G>A	missense	Class 3 LP	PM2 Strong, PM5 Supporting, PP3 Supporting
10100422	F	BA	28	Class III	23	*MYBPC3*	p.Gly507Arg	c.1519G>A	missense	*Class 3 LP*	PM2 Strong, PP5 Strong, PM1 Moderate, PP3 Supporting
10100423	F	BA	32	ND	30	*DES*	p.Asp312Ala	c.935A>C	missense	Class 3 LP	PM2 Strong, PM1 Moderate, PP3 Supporting
10100423	F	BA	32	ND	30	*TMEM43*	p.Leu88Pro	c.263T>C	missense	Class 3	PM2, PM1
10100429	F	BA	40			*ALPK3*	p.Glu148GlnfsTer8	c.436_440dup	frameshift	Class 3 LP	PVS1 Strong, PM2 Moderate
10100430	F	MA	40	Class II	45	*DES**	p.Asp312Asn	c.934G>A	missense	Class 3 LP	PM2 Strong, PM1 Moderate, PP3 Supporting
10100432	F	BA	28	ND	45	*DSP*	p.Gln678Arg	c.2033A>G	missense	Class 3	PM2, PM1
10100432	F	BA	28	ND	45	*DSP*	p.Arg1341His	c.4022G>A	missense	Class 3 LP	PM2 Strong, PM5 Supporting, PP3 Supporting
10100432	F	BA	28	ND	45	*DSP*	p.Glu1740Lys	c.5218G>A	missense	*Class 3 LB*	Borderline BS1 and borderline BP4
10100432	F	BA	28	ND	45	*TTN*	p.Gly19031Glu	c.57092G>A	missense	Class 3 LB	TTN missense not reviewed
10100432	F	BA	28	ND	45	*TTN*	p.Lys14896Arg	c.44687A>G	missense	Class 3 LB	TTN missense not reviewed
10100432	F	BA	28	ND	45	*TTN*	p.Lys11470Glu	c.34408A>G	missense	Class 3 LB	TTN missense not reviewed
10100432	F	BA	28	ND	45	*TTN*	p.Pro11110Thr	c.33328C>A	missense	Class 3 LB	TTN missense not reviewed
10100442	F	BA	42	Class III	30	*TTN*	p.Ser8409Tyr	c.25226C>A	missense	Class 3 LB	TTN missense not reviewed
10100707	F	BA	38	Class II	44	*TTN*	p.Lys11470Glu	c.34408A>G	missense	Class 3 LB	TTN missense not reviewed
10100707	F	BA	38	Class II	44	*TTN*	p.Ser8409Tyr	c.25226C>A	missense	Class 3 LB	TTN missense not reviewed
10100729	F	MA	21	Class III	42	SCN5A	p.Glu1832Gln	c.5494C>G	missense	Class 3	Conflicting
10100729	F	MA	21	Class III	42	SCN5A	p.Ala29Val	c.86_87inv	missense	Class 3 LP	PM2 Strong, PM1 Moderate
10100729	F	MA	21	Class III	42	DSG2	p.Cys814Tyr	c.2441G>A	missense	Class 3	PM2, PM1
10100748	F	MA	37	Class III	47	*TTN*	p.Val21222Ala	c.63665T>C	missense	Class 3 LB	TTN missense not reviewed
10100749	F	MA	36	Class III	40	*FLNC*	p.Val1446Ala	c.4337T>C	missense	Class 3	PM2, PM1
101301354	F	BA	42	Class III	22	TTN	p.Phe15649Leu	c.46945T>C	missense	Class 3 LB	TTN missense not reviewed
101301354	F	BA	42	Class III	22	TTN	p.Leu15490Ile	c.46468C>A	missense	Class 3 LB	TTN missense not reviewed
101301354	F	BA	42	Class III	22	TTN	p.Glu3708Gln	c.11122C>G	missense	Class 3 LB	TTN missense not reviewed
101301354	F	BA	42	Class III	22	FLNC	p.Val1446Ala	c.4337T>C	missense	Class 3	PM2, PM1
101301375	F	MA	30	Class III	20	*PCL*	p.Lys1070Arg	c.3209A>G	missense	Class 3	Conflicting
101301375	F	MA	30	Class III	20	*DMD*	p.Leu239Phe	c.717G>T	missense	Class 3	Conflicting
101301376	F	MA	31	Class III	27	TTN	p.Gly632Arg	c.1894G>C	missense	Class 3 LB	TTN missense not reviewed
101301376	F	MA	31	Class III	27	DMD	p.Ala2999Val	c.8996C>T	missense	Class 3	Conflicting
101301376	F	MA	31	Class III	27	DMD	p.Glu2118Lys	c.6352C>A	missense	Class 3	Artefact



101301377	F	MA	28	Class IV	29	*TTN*	p.Val31720Ala	c.95159T>C	missense	Class 3 LB	TTN missense not reviewed
101301377	F	MA	28	Class IV	29	*TTN*	p.Lys11470Glu	c.34408A>G	missense	Class 3 LB	TTN missense not reviewed
101301377	F	MA	28	Class IV	29	*TTN*	p.Tyr5497Cys	c.16490A>G	missense	Class 3 LB	TTN missense not reviewed
101301392	F	MA	30	Class III	32	*TTN*	p.Pro21332Thr	c.63994C>A	missense	Class 3 LB	TTN missense not reviewed
101301425	F	BA	32	Class II	40	*DSP*	p.Ser2843_Arg2846del	c.8508_8519del	inframe_deletion	Class 3	PM2, PM1
101301437	F	MA	30	Class II	36	*ANKRD1*	p.Thr116ArgfsTer10	c.347del	frameshift	Class 3	Gene not relevant in context
101301445	F	MA	20	Class III	23	TTN	p.Val24611Leu	c.73831G>T	missense	Class 3 LB	TTN missense not reviewed
101301512	F	BA	37	Class III	59	*TTN*	p.Leu51Arg	c.152T>G	missense	Class 3 LB	TTN missense not reviewed
101301525	F	MA	34	Class III	20	*TTN*	p.Arg17518Cys	c.52552C>T	missense	Class 3 LP	TTN missense not reviewed
101301526	F	MA	39	Class IV	25	*TTN*	p.Leu29588Ser	c.88763T>C	missense	Class 3 LB	TTN missense not reviewed
101301526	F	MA	39	Class IV	25	*TTN*	p.Pro11771Leu	c.35312C>T	missense	Class 3 LB	TTN missense not reviewed
101301526	F	MA	39	Class IV	25	TTN	p.Arg1408Cys	c.4222C>T	missense	Class 3 LP	TTN missense not reviewed
101301526	F	MA	39	Class IV	25	TMEM43	p.Arg328Trp	c.982C>T	missense	Class 3	Conflicting
101301526	F	MA	39	Class IV	25	ANKRD1	p.Lys138AsnfsTer35	c.414del	frameshift	Class 3	Gene not relevant in context
101301526	F	MA	39	Class IV	25	LAMP2	p.Asp322Val	c.965A>T	missense	Class 3	PM2, PM1
101301547	F	MA	29	Class IV	39	TTN	p.Gly24659Ala	c.73976G>C	missense	Class 3 LB	TTN missense not reviewed
101301547	F	MA	29	Class IV	39	TTN	p.Lys11470Glu	c.34408A>G	missense	Class 3 LB	TTN missense not reviewed
101301547	F	MA	29	Class IV	39	TTN	p.Gln3708Glu	c.11122C>G	missense	Class 3 LB	TTN missense not reviewed
101301547	F	MA	29	Class IV	39	SCN5A	p.Leu1308Phe	c.3922C>T	missense	Class 3	PM2, PM1
101301551	F	MA	31			DMD	p.Asp2112Tyr	c.6334G>T	missense	*Class 3 LB*	Conflicting
101301562	F	BA	28	Class III	39	DSP	p.Asn408His	c.1222A>C	missense	Class 3	Conflicting
101301571	F	BA	20	Class III	30	TTN	p.Pro21332Thr	c.63994C>A	missense	Class 3 LB	TTN missense not reviewed
101301572	F	BA	22	Class III	25	TTN	p.Lys24533Glu	c.73597A>G	missense	Class 3 LB	TTN missense not reviewed
101301572	F	BA	22	Class III	25	TTN	p.Ile22842Thr	c.68525T>C	missense&splice_region	Class 3 LB	TTN missense not reviewed
101301572	F	BA	22	Class III	25	TTN	p.Lys9755Glu	c.29263A>G	missense	Class 3 LB	TTN missense not reviewed
101301572	F	BA	22	Class III	25	DMD	p.Asp2112Tyr	c.6334G>T	missense	*Class 3 LB*	Conflicting
101301574	F	MA	22	Class III	37	TTN	p.Val31720Ala	c.95159T>C	missense	Class 3 LB	TTN missense not reviewed
101301574	F	MA	22	Class III	37	TTN	p.Ile18822Thr	c.56465T>C	missense	Class 3 LB	TTN missense not reviewed
101301574	F	MA	22	Class III	37	TTN	p.Lys11470Glu	c.34408A>G	missense	Class 3 LB	TTN missense not reviewed
101301574	F	MA	22	Class III	37	TTN	p.Tyr5497Cys	c.16490A>G	missense	Class 3 LP	TTN missense not reviewed
101301574	F	MA	22	Class III	37	DMD	p.Gln2118Lys	c.6352C>A	missense	Class 3	Artefact
101400001	F	MA	37			TTN	p.Met33804Val	c.101410A>G	missense	Class 3 LB	TTN missense not reviewed
101400001	F	MA	37			TTN	p.Ile29254Thr	c.87761T>C	missense	Class 3 LP	TTN missense not reviewed
101400001	F	MA	37			TTN	p.Val26364Ala	c.79091T>C	missense	Class 3 LP	TTN missense not reviewed
101400001	F	MA	37			DMD	p.Glu2910_Asn2912delinsValValA	c.8729_8734delinsTGGTCG	missense	Class 3 LB	PM2 Strong, PM1 Moderate
101400002	F	MA	35			TTN	p.Met24359Thr	c.73076T>C	missense	Class 3 LB	TTN missense not reviewed
101400002	F	MA	35			TTN	p.Lys11206Asn	c.33618G>C	missense	Class 3 LB	TTN missense not reviewed
106304530	F	BA	31	Class III	42	TTN	p.Ile18822Thr	c.56465T>C	missense	Class 3 LB	TTN missense not reviewed
106304530	F	BA	31	Class III	42	TTN	p.Val16981Leu	c.50941G>C	missense	Class 3 LB	TTN missense not reviewed
106304530	F	BA	31	Class III	42	DSG2	p.Ser884del	c.2649_2651del	inframe_deletion	Class 3	Conflicting
106304534	F	BA	24	Class I	49	*MYH7*	p.Ala728Val	c.2183C>T	missense	Class 3 LP	PM1 Strong, PM2 Supporting, PP5 Supporting
106304534	F	BA	24	Class I	49	*PKP2*	p.Ile397Asn	c.1190T>A	missense	Class 3	Conflicting
106304540	F	BA	22	Class II	45	*MYBPC3*	p.Pro961Leu	c.2882C>T	missense	Class 3 LP	PM2 Supporting, PP3 Supporting
106304540	F	BA	22	Class II	45	*TTN*	p.Lys21462Glu	c.64384A>G	missense	Class 3 LB	TTN missense not reviewed
106304540	F	BA	22	Class II	45	*TTN*	p.Lys11470Glu	c.34408A>G	missense	Class 3 LB	TTN missense not reviewed
106304546	F	BA	22	Class III	49	*TTN*	p.Gly2342Arg	c.7024G>A	missense	Class 3 LB	TTN missense not reviewed
106304547	F	BA	45	Class III	31	*LMNA*	p.Asn142Lys	c.426C>A	missense	Class 3 LP	PM2 Strong, PM1 Moderate
106304547	F	BA	45	Class III	31	*MYH7*	p.Met1730Thr	c.5189T>C	missense	Class 3	Conflicting
106304547	F	BA	45	Class III	31	*SCN5A*	p.Leu1308Phe	c.3922C>T	missense	Class 3	PM2, PM1
106304547	F	BA	45	Class III	31	*TTN*	p.Arg30628Lys	c.91883G>A	missense	Class 3 LB	TTN missense not reviewed
106304549	F	BA	36	Class I	38	*TTN*	p.Gly19045Arg	c.57133G>A	missense	Class 3 LP	TTN missense not reviewed
106304549	F	BA	36	Class I	38	*TTN*	p.Lys11663Asn	c.34989A>C	missense	Class 3 LB	TTN missense not reviewed
106304564	F	BA	25			*MYH7*	p.Ala728Val	c.2183C>T	missense	Class 3 LP	PM1 Strong, PM2 Supporting, PP5 Supporting
106304565	F	BA	22			*MYH7*	p.Pro132Leu	c.395C>T	missense	Class 3 LP	PM1 Strong, PM2 Supporting, PP5 Supporting
106304568	F	BA	27	Class III	13	*SCN5A*	p.Leu1308Phe	c.3922C>T	missense	Class 3	PM2, PM1
106304570	F	BA	24	Class IV	33	*SCN5A*	p.Leu1308Phe	c.3922C>T	missense	Class 3	PM2, PM1
106304570	F	BA	24	Class IV	33	*TTN*	p.Ile22842Thr	c.68525T>C	missense&splice_region	Class 3 LB	TTN missense not reviewed
106304593	F	BA	35	Class II	48	TTN	p.Lys11470Glu	c.34408A>G	missense	Class 3 LB	TTN missense not reviewed
106304593	F	BA	35	Class II	48	SCN5A	p.Leu1308Phe	c.3922C>T	missense	Class 3	PM2, PM1

106304600	F	BA	32	Class III	22	*TTN*	p.Arg28111Lys	c.84332G>A	missense	Class 3 LB	TTN missense not reviewed
401300302	F	BA	25	Class III	39	DMD	p.Glu2910_Asn2912delinsValValA	c.8729_8734delinsTGGTCG	missense	Class 3 LB	PM2 Strong, PM1 Moderate
401300305	F	BA	37	Class IV	35	TTN	p.Thr25580Ala	c.76738A>G	missense	Class 3 LB	TTN missense not reviewed
401300305	F	BA	37	Class IV	35	TTN	p.Pro1207Thr	c.3619C>A	missense	Class 3 LB	TTN missense not reviewed
401300306	F	BA	26	Class III	28	*TTN*	p.Pro23850Ala	c.71548C>G	missense	Class 3 LB	TTN missense not reviewed
401300306	F	BA	26	Class III	28	*TTN*	p.Lys11470Glu	c.34408A>G	missense	Class 3 LB	TTN missense not reviewed
401300306	F	BA	26	Class III	28	*TTN*	p.Leu3554His	c.10661T>A	missense	Class 3 LB	TTN missense not reviewed
401303002	F	BA	21	Class IV	29	LAMP2	p.Asp322Val	c.965A>T	missense	Class 3	PM2, PM1
401303006	F	BA	38	Class III	25	*DSP*	p.Ser2843_Arg2846del	c.8508_8519del	inframe_deletion	Class 3	PM2, PM1
401303006	F	BA	38	Class III	25	*JUP*	p.Ser241Leu	c.722C>T	missense	Class 3	PM2, PM1
401303014	F	BA	32	Class III	37	*TTN*	p.Arg30628Lys	c.91883G>A	missense	Class 3 LB	TTN missense not reviewed
401303021	F	BA	29	Class III	23	*TTN*	p.Gly17891Asp	c.53672G>A	missense	Class 3 LB	TTN missense not reviewed
401303023	F	BA	19	Class III	20	*ALPK3*	p.Val302Ile	c.904G>A	missense	Class 3	BA1
401303024	F	BA	35	Class III	28	TTN	p.Glu14697Lys	c.44089G>A	missense	Class 3 LP	TTN missense not reviewed
401303024	F	BA	35	Class III	28	ALPK3	p.Val302Ile	c.904G>A	missense	Class 3	BA1
401303051	F	BA	22	Class II	37	TTN	p.Gly14590Glu	c.43769G>A	missense	Class 3 LB	TTN missense not reviewed
401303051	F	BA	22	Class II	37	DMD	p.Asp2112Tyr	c.6334G>T	missense	*Class 3 LB*	Conflicting
401303055	F	BA	22	Class III	15	TTN	p.Ala352Thr	c.1054G>A	missense	Class 3 LB	TTN missense not reviewed
401303055	F	BA	22	Class III	15	DMD	p.Gln2118Lys	c.6352C>A	missense	Class 3	Artefact
401303055	F	BA	22	Class III	15	DMD	p.Asp2112Tyr	c.6334G>T	missense	*Class 3 LB*	Conflicting
401303056	F	BA	31	Class III	27.6	DSP	p.Gln678Arg	c.2033A>G	missense	Class 3	PM2, PM1
401303056	F	BA	31	Class III	27.6	DSP	p.Arg1341His	c.4022G>A	missense	Class 3 LP	PM2 Strong, PM5 Supporting, PP3 Supporting
401303056	F	BA	31	Class III	27.6	DMD	p.Gln2118Lys	c.6352C>A	missense	Class 3	Artefact
401303058	F	BA	17	Class III	16.2	TTN	p.Tyr16386Cys	c.49157A>G	missense	Class 3 LB	TTN missense not reviewed
401303059	F	BA	35	Class III	40.7	TTN	p.Phe15649Leu	c.46945T>C	missense	Class 3 LB	TTN missense not reviewed
401303059	F	BA	35	Class III	40.7	TTN	p.Glu5146Gly	c.15437A>G	missense	Class 3 LB	TTN missense not reviewed
401303059	F	BA	35	Class III	40.7	DSP	p.Gln678Arg	c.2033A>G	missense	Class 3	PM2, PM1
401303059	F	BA	35	Class III	40.7	DSP	p.Arg1341His	c.4022G>A	missense	Class 3	Conflicting
HCM											
401303046	F	BA	65	Class II	82	*DSP*	p.Arg1341His	c.4022G>A	missense	Class 3 LP	PM2 Strong, PM5 Supporting, PP3 Supporting
10100072	M	MA	57	Class II	76	*SCN5A*	p.Glu1832Gln	c.5494C>G	missense	Class 3	Conflicting
10100072	M	MA	57	Class II	76	*TMEM43*	p.Arg328Trp	c.982C>T	missense	Class 3	Conflicting
10100098	M	BA	42	Class I	83	*TTN*	p.Thr29933Ile	c.89798C>T	missense	Class 3 LB	TTN missense not reviewed
10100098	M	BA	42	Class I	83	*TTN*	p.Leu20789Phe	c.62365C>T	missense	Class 3 LB	TTN missense not reviewed
10100098	M	BA	42	Class I	83	*TTN*	p.Pro18980Leu	c.56939C>T	missense	Class 3 LB	TTN missense not reviewed
10100510	M	BA	27	Class II	64	*TTN*	p.Ala19169Pro	c.57505G>C	missense	Class 3 LB	TTN missense not reviewed
10100512	F	MA	59	Class III	68	*TTN*	p.Ser33014Thr	c.99041G>C	missense	Class 3 LB	TTN missense not reviewed
10100512	F	MA	59	Class III	68	*TTN*	p.Arg20310His	c.60929G>A	missense	Class 3 LB	TTN missense not reviewed
10100513	F	BA	27	Class I	67	*DSP*	p.Gln678Arg	c.2033A>G	missense	Class 3	PM2, PM1
10100513	F	BA	27	Class I	67	*DSP*	p.Arg1341His	c.4022G>A	missense	Class 3	Conflicting
10100513	F	BA	27	Class I	67	*TTN*	p.Gly24659Ala	c.73976G>C	missense	Class 3 LB	TTN missense not reviewed
10100513	F	BA	27	Class I	67	*TTN*	p.Lys11470Glu	c.34408A>G	missense	Class 3 LB	TTN missense not reviewed
10100513	F	BA	27	Class I	67	*TTN*	p.Gln3708Glu	c.11122C>G	missense	Class 3 LB	TTN missense not reviewed
10100520	M	BA	43	Class I	68	*LAMP2**	p.Asp322Val	c.965A>T	missense	Class 3	PM2, PM1
10100522	M	BA	34	Class I	50	*ALPK3*	p.Glu148GlnfsTer8	c.436_440dup	frameshift	Class 3 LP	PVS1 Strong, PM2 Moderate
10100522	M	BA	34	Class I	50	FLNC	p.Val1641Ile	c.4921G>A	missense	Class 3	Conflicting
10100522	M	BA	34	Class I	50	*MYBPC3*	p.Gly416Ser	c.1246G>A	missense	Class 3 LP	PM2 Moderate, PP5 Supporting
10100528	M	BA	61	Class II	56	*ALPK3*	p.Glu148GlnfsTer8	c.436_440dup	frameshift	Class 3 LP	PVS1 Strong, PM2 Moderate
10100528	M	BA	61	Class II	56	*TTN*	Splice Site	c.11254+2T>C	splice_donor	*Class 3 LB*	putative truncating variants not in 100% PSI regions
10100533	F	MA	23	Class III	81	*ALPK3*	p.Val302Ile	c.904G>A	missense	Class 3	BA1
10100533	F	MA	23	Class III	81	*DES*	p.Glu262Val	c.785A>T	missense	*Class 3 LB*	Borderline BS1 and borderline BP4
10100538	M	BA	50	Class I	61	*CDH2*	p.Pro477Leu	c.1430_1431delinsTG	missense	Class 3	PM2_supporting; BP4
10100538	M	BA	50	Class I	61	*PKP2*	p.Ser757Tyr	c.2270C>A	missense	Class 3	PM2, PM1
10100539	M	MA	56	Class II	52	*SCN5A*	p.Arg971Cys	c.2911C>T	missense	Class 3 LP	PM2 Supporting, PP5 Supporting, PP3 Supporting
10100703	F	MA	60	Class II	73	*MYBPC3*	p.Arg238His	c.713G>A	missense	Class 3	Conflicting
10200004	M	BA	0			*DMD**	p.Ser810Arg	c.2428A>C	missense	Class 3	Conflicting
10200007	M	BA	1			*TTN*	p.Val19400Ala	c.58199T>C	missense	Class 3 LB	TTN missense not reviewed
10200007	M	BA	1			*TTN*	p.Leu15490Ile	c.46468C>A	missense	Class 3 LB	TTN missense not reviewed
10200007	M	BA	1			*TTN*	p.Gln3708Glu	c.11122C>G	missense	Class 3 LB	TTN missense not reviewed

101301029	M	WA	35	Class I	67	*TTN*	p.Ile18077Val	c.54229A>G	missense	Class 3 LB	TTN missense not reviewed
101301369	M	WA	22	Class II	70	*TTN*	p.Tyr8950Cys	c.26849A>G	missense	Class 3 LB	TTN missense not reviewed
101301371	M	MA	34	Class II	85	TTN	p.Gly24659Ala	c.73976G>C	missense	Class 3 LB	TTN missense not reviewed
101301371	M	MA	34	Class II	85	TTN	p.Ser6627Leu	c.19880C>T	missense	Class 3 LB	TTN missense not reviewed
101301390	F	WA	28	Class II	54	*FLNC*	p.Gly1546Ser	c.4636G>A	missense	Class 3	Conflicting
101301417	M	MA	66	Class I	67	*ANKRD1*	p.Lys138AsnfsTer35	c.414del	frameshift	Class 3	Gene not relevant in context
101301417	M	MA	66	Class I	67	*TTN*	Splice Site	c.32554+1G>T	splice_donor	Class 3 LP	PVS1 Very Strong, PM2 Moderate, PP3 Supporting
101301440	M	WA	23	Class I	71	*ALPK3*	p.Glu148GlnfsTer8	c.436_440dup	frameshift	Class 3 LP	PVS1 Strong, PM2 Moderate
101301447	M	MA	46	Class II	80	MYBPC3	p.Ala534Val	c.1601C>T	missense	Class 3 LP	PM1 Moderate, PM2 Suppporting, PP3 Supporting
101301501	M	MA	46	Class III	81	*SCN5A*	p.Leu1308Phe	c.3922C>T	missense	Class 3	PM2, PM1
101301561	M	WA	44	Class I		TTN	p.Thr1026Ile	c.3077C>T	missense	Class 3 LP	TTN missense not reviewed
102302007	M	BA	1			*TTN*	p.Thr34741Ile	c.104222C>T	missense	Class 3 LB	TTN missense not reviewed
106304536	F	BA	16			*MYBPC3*	p.Gly507Arg	c.1519G>A	missense	*Class 3 LB*	PM2 Strong, PP5 Strong, PM1 Moderate, PP3 Supporting
106304536	F	BA	16			*TTN*	p.Lys11470Glu	c.34408A>G	missense	Class 3 LB	TTN missense not reviewed
401303045	M	BA	32	Class I	83	TTN	p.Pro21332Thr	c.63994C>A	missense	Class 3 LB	TTN missense not reviewed
401303045	M	BA	32	Class I	83	TTN	p.Val19400Ala	c.58199T>C	missense	Class 3 LB	TTN missense not reviewed
401303045	M	BA	32	Class I	83	TTN	p.Leu15490Ile	c.46468C>A	missense	Class 3 LB	TTN missense not reviewed
401303045	M	BA	32	Class I	83	TTN	p.Gln3708Glu	c.11122C>G	missense	Class 3 LB	TTN missense not reviewed
401303045	M	BA	32	Class I	83	MYBPC3	p.Glu1179Lys	c.3535G>A	missense	Class 3	Conflicting
401303046	F	BA	65	Class II	82	*DSP*	p.Gln678Arg	c.2033A>G	missense	Class 3	PM2, PM1
RCM											
10100601	M	IA	30	Class III	43	*DES**	p.Gly456Arg	c.1366G>A	missense	Class 3 LP	PM2 Strong, PM1 Supporting, PP3 Supporting
101301407	F	MA	57	Class III	54	TTN	Splice Site	c.32471-1G>A	splice_acceptor	Class 3 LP	PVS1 Very Strong, PM2 Moderate
101301407	F	MA	57	Class III	54	TTN	p.Lys28450Asn	c.85350G>C	missense	Class 3 LB	TTN missense not reviewed
101301407	F	MA	57	Class III	54	TTN	p.Ser19018Ala	c.57052T>G	missense	Class 3 LB	TTN missense not reviewed
101301407	F	MA	57	Class III	54	TTN	p.Cys9226Tyr	c.27677G>A	missense	Class 3 LB	TTN missense not reviewed
101301407	F	MA	57	Class III	54	MYH7	p.Thr761Asn	c.2282C>A	missense	Class 3 LP	PM2 Moderate, PP3 Moderate
101301429	M	BA	56	Class II	45	*TTN*	p.Gly24659Ala	c.73976G>C	missense	Class 3 LB	TTN missense not reviewed
101301429	M	BA	56	Class II	45	*TTN*	p.Lys11470Glu	c.34408A>G	missense	Class 3 LB	TTN missense not reviewed
101301429	M	BA	56	Class II	45	*TTN*	p.Gln3708Glu	c.11122C>G	missense	Class 3 LB	TTN missense not reviewed
107305004	F	BA	24	Class II	60	SCN5A	p.Asp2003Asn	c.6007G>A	missense	Class 3 LP	PM2 Strong, PP5 Supporting
107305007	F	BA	50	Class I	67	DMD	p.Gln2118Lys	c.6352C>A	missense	Class 3	Artefact
107305007	F	BA	50	Class I	67	DMD	p.Asp2112Tyr	c.6334G>T	missense	*Class 3 LB*	Conflicting
107305008	F	BA	32	Class II	44	*DMD*	p.Glu2910_Asn2912delins ValValA	c.8729_8734delins TGGTCG	missense	Class 3 LB	PM2 Strong, PM1 Moderate
107305009	F	BA	21			DMD	p.Asp2112Tyr	c.6334G>T	missense	*Class 3 LB*	Conflicting
107305011	F	BA	19	Class II	70	DMD	p.Gln2118Lys	c.6352C>A	missense	Class 3	Artefact
107305013	F	BA	36	Class II	68	DMD	p.Asp2112Tyr	c.6334G>T	missense	*Class 3 LB*	Conflicting
107305013	F	BA	36	Class II	68	DMD	p.Gln2118Lys	c.6352C>A	missense	Class 3	Artefact
107305014	F	BA	20			TTN	p.Ala29469Glu	c.88406C>A	missense	Class 3 LB	TTN missense not reviewed
107305014	F	BA	20			TTN	p.Pro18263Leu	c.54788C>T	missense	Class 3 LB	TTN missense not reviewed
107305014	F	BA	20			TTN	p.Glu7838Lys	c.23512G>A	missense	Class 3 LB	TTN missense not reviewed
107305014	F	BA	20			DMD	p.Glu2910_Asn2912delins ValValA	c.8729_8734delins TGGTCG	missense	Class 3 LB	PM2 Strong, PM1 Moderate
107305017	M	BA	16			DMD	p.Gln2118Lys	c.6352C>A	missense	Class 3	Artefact
107305024	M	BA	51	Class III	14	TTN	p.Gly32961Ala	c.98882G>C	missense	Class 3 LB	TTN missense not reviewed
107305024	M	BA	51	Class III	14	TTN	p.Tyr20639Asp	c.61915T>G	missense	Class 3 LB	TTN missense not reviewed
107305024	M	BA	52	Class III	21	TTN	p.Gly32961Ala	c.98882G>C	missense	Class 3 LB	TTN missense not reviewed
107305024	M	BA	52	Class III	21	TTN	p.Tyr20639Asp	c.61915T>G	missense	Class 3 LB	TTN missense not reviewed
107305025	F	BA	40	Class IV	70	DSP	p.Gln678Arg	c.2033A>G	missense	Class 3	PM2, PM1
107305025	F	BA	40	Class IV	70	DSP	p.Arg1341His	c.4022G>A	missense	Class 3	Conflicting
107305026	M	BA	34			TTN	p.Pro21332Thr	c.63994C>A	missense	Class 3 LB	TTN missense not reviewed
107305027	F	BA	63	Class III	81	*DSP*	p.Pro2618Ser	c.7852C>T	missense	Class 3	Conflicting
107305028	M	BA	30	Class II	70	*TTN*	p.Arg30628Lys	c.91883G>A	missense	Class 3 LB	TTN missense not reviewed
107305032	F	BA	45	Class II	53	MYBPC3	p.Gly507Arg	c.1519G>A	missense	*Class 3 LB*	PM2 Strong, PP5 Strong, PM1 Moderate, PP3 Supporting
107305032	F	BA	45	Class II	53	TTN	p.Arg30628Lys	c.91883G>A	missense	Class 3 LB	TTN missense not reviewed
107305103	F	BA	37	Class II	76	TTN	p.Val30998Ala	c.92993T>C	missense	Class 3 LB	TTN missense not reviewed
107305103	F	BA	37	Class II	76	TTN	p.Lys16709Asn	c.50127G>C	missense	Class 3 LB	TTN missense not reviewed
107305103	F	BA	37	Class II	76	TTN	p.Asp10236Ala	c.30707A>C	missense	Class 3 LB	TTN missense not reviewed
107305108	F	BA	17	Class III	87	FLNC	p.Ala1247Val	c.3740C>T	missense	*Class 3 LB*	Conflicting
107305108	F	BA	17	Class III	87	DMD	p.Glu2910_Asn2912delins ValValA	c.8729_8734delins TGGTCG	missense	Class 3 LB	PM2 Strong, PM1 Moderate

SS Table 3: The fitted logistic regression model for adult DCM probands

Dependent: Panel Diagnosis Yield, n (%)		Positive	Uncertain	Negative	OR (univariable)	OR (multivariable)
Age	Mean (SD)	36.4 (11.8)	34.9 (11.5)	36.3 (12.4)	0.99 (0.97-1.01, p=0.505)	1.02 (0.99-1.04, p=0.247)
LVEF	Mean (SD)	25.7 (12.1)	28.8 (11.8)	28.0 (11.9)	1.02 (1.00-1.05, p=0.079)	1.02 (1.00-1.05, p=0.080)
NYHA	Class I	3 (10.7)	11 (39.3)	14 (50.0)	-	-

Dependent: Panel Diagnosis Yield, n (%)		Positive	Uncertain	Negative	OR (univariable)	OR (multivariable)
	Class II	25 (22.1)	45 (39.8)	43 (38.1)	0.42 (0.10-1.33, p=0.186)	0.46 (0.10-1.55, p=0.251)
	Class III	29 (14.0)	108 (52.2)	70 (33.8)	0.74 (0.17-2.28, p=0.634)	0.72 (0.16-2.42, p=0.624)
	Class IV	9 (18.0)	31 (62.0)	10 (20.0)	0.55 (0.11-2.03, p=0.397)	0.48 (0.09-1.99, p=0.334)
	ND	2 (22.2)	6 (66.7)	1 (11.1)	0.42 (0.06-3.66, p=0.389)	0.16 (0.02-1.55, p=0.090)
Familial	Yes	34 (26.6)	55 (43.0)	39 (30.5)	-	-
	No data	8 (18.6)	22 (51.2)	13 (30.2)	1.58 (0.69-3.97, p=0.297)	-
	No	34 (12.2)	146 (52.3)	99 (35.5)	**2.61 (1.53-4.45, p<0.001)**	**2.48 (1.39-4.43, p=0.002)**
Sex	Female	32 (12.7)	135 (53.6)	85 (33.7)	-	-
	Male	44 (22.2)	88 (44.4)	66 (33.3)	**0.51 (0.31-0.84, p=0.008)**	0.92 (0.46-1.79, p=0.810)
Ancestry	Black-African	39 (15.0)	137 (52.7)	84 (32.3)	-	-
	Mixed	24 (16.1)	76 (51.0)	49 (32.9)	0.92 (0.53-1.62, p=0.765)	0.94 (0.51-1.76, p=0.836)
	White	13 (32.5)	10 (25.0)	17 (42.5)	**0.37 (0.18-0.79, p=0.008)**	**0.39 (0.16-0.97, p=0.039)**
Subtype	Dilated	63 (20.1)	138 (43.9)	113 (36.0)	-	-
	Peripartum	13 (9.6)	85 (62.5)	38 (27.9)	**2.37 (1.30-4.66, p=0.008)**	**2.19 (0.92-5.38, p=0.079)**

Number in data frame = 450, Number in model = 405, Missing = 45, AIC = 360.3, C-statistic = 0.702, H&L = Chi-sq (8) 5.45 (p=0.708)

SS Table 4: The fitted logistic regression model for adult HCM probands

Dependent: Panel Diagnosis Yield, n (%)		Positive	Uncertain	Negative	OR (univariable)	OR (multivariable)
Age	Mean (SD)	33.9 (11.1)	41.5 (15.2)	45.9 (12.3)	1.06 (1.01-1.13, p=0.027)	1.06 (0.99-1.15, p=0.125)
LVEF	Mean (SD)	71.2 (7.1)	70.0 (10.7)	71.2 (7.2)	0.99 (0.91-1.07, p=0.820)	0.98 (0.87-1.09, p=0.723)
NYHA	Class I	6 (25.0)	10 (41.7)	8 (33.3)	-	-
	Class II	4 (16.7)	10 (41.7)	10 (41.7)	1.67 (0.41-7.43, p=0.480)	1.80 (0.28-13.34, p=0.538)
	Class III	2 (28.6)	3 (42.9)	2 (28.6)	0.83 (0.14-6.88, p=0.849)	1.30 (0.12-19.65, p=0.833)

Dependent: Panel Diagnosis Yield, n (%)		Positive	Uncertain	Negative	OR (univariable)	OR (multivariable)
	Class IV	1 (100.0)			0.00 (NA- ∞, p=0.994)	0.00 (NA-∞, p=0.994)
Familial	Yes	4 (19.0)	7 (33.3)	10 (47.6)	-	-
	No data	1 (25.0)	1 (25.0)	2 (50.0)	0.71 (0.07-16.36, p=0.786)	-
	No	9 (25.7)	16 (45.7)	10 (28.6)	0.68 (0.16-2.46, p=0.569)	0.25 (0.03-1.63, p=0.179)
Sex	Female	7 (33.3)	7 (33.3)	7 (33.3)	-	-
	Male	7 (17.9)	17 (43.6)	15 (38.5)	2.29 (0.67-7.93, p=0.185)	0.85 (0.10-5.92, p=0.871)
Ancestry	Black-African	2 (13.3)	10 (66.7)	3 (20.0)	-	-
	Mixed	8 (23.5)	9 (26.5)	17 (50.0)	0.50 (0.07-2.36, p=0.420)	0.28 (0.01-2.56, p=0.317)
	White	4 (36.4)	5 (45.5)	2 (18.2)	0.27 (0.03-1.74, p=0.183)	0.17 (0.01-2.53, p=0.233)
	Indian	0 (NaN)	0 (NaN)	0 (NaN)	-	-

Number in data frame = 60, Number in model = 52, Missing = 8, AIC = 62.3, C-statistic = 0.776, H&L = Chi-sq (8) 9.59 (p=0.295)

SS Table 5: The fitted logistic regression model for adult RCM probands

Dependent: Panel Diagnosis Yield, n (%)		Positive	Uncertain	Negative	OR (univariable)	OR (multivariable)
Age	Mean (SD)	26.0 (3.9)	35.8 (14.7)	31.2 (15.9)	1.05 (0.97-1.20, p=0.338)	1.03 (0.89-1.27, p=0.644)
LVEF	Mean (SD)	53.7 (8.1)	60.1 (18.5)	63.2 (15.8)	1.03 (0.96-1.11, p=0.433)	1.12 (0.95-1.62, p=0.338)
NYHA	Class II	2 (11.1)	8 (44.4)	8 (44.4)	0.00 (NA-∞, p=0.997)	0.00 (NA-∞, p=0.999)
	Class III	1 (11.1)	5 (55.6)	3 (33.3)	0.00 (NA-Inf, p=0.997)	0.00 (NA-∞, p=0.999)
	Class I		1 (33.3)	2 (66.7)	-	-
	Class IV		1 (100.0)		1.00 (0.00-Inf, p=1.000)	4.55 (0.00-Inf, p=1.000)
Familial	Yes	1 (12.5)	5 (62.5)	2 (25.0)	-	-
	No data	1 (8.3)	6 (50.0)	5 (41.7)	1.57 (0.06-44.30, p=0.762)	-
	No	2 (8.7)	10 (43.5)	11 (47.8)	1.50 (0.06-18.18, p=0.755)	0.78 (0.01-28.31, p=0.885)
Sex	Female	4 (14.3)	15 (53.6)	9 (32.1)	-	-
	Male	0 (0.0)	6 (40.0)	9 (60.0)	52394335.48 (0.00-NA, p=0.995)	131198588.90 (0.00-NA, p=0.997)
Ancestry	Black-African	1 (2.9)	19 (54.3)	15 (42.9)	-	-
	Mixed	3 (60.0)	1 (20.0)	1 (20.0)	**0.02 (0.00-0.22, p=0.004)**	0.05 (0.00-1.22, p=0.114)
	White	0 (0.0)	0 (0.0)	2 (100.0)	3401435.08 (0.00-NA, p=0.997)	5828922.08 (0.00-NA, p=0.999)
	Indian	0 (0.0)	1 (100.0)	0 (0.0)	3401435.10 (0.00-NA, p=0.998)	0.46 (0.00-NA, p=1.000)

Number in data frame = 43, Number in model = 31, Missing = 12, AIC = 32.6, C-statistic = 0.94, H&L = Chi-sq (8) 1.14 (p=0.997)

SS Table 6: The fitted logistic regression model for adult ACM probands

Dependent: Panel Diagnosis Yield, n (%)		Positive	Uncertain	Negative	OR (univariable)	OR (multivariable)
Age	Mean (SD)	35.9 (15.2)	43.8 (15.7)	40.9 (5.9)	1.04 (0.99-1.11, p=0.153)	1.12 (1.01-1.31, p=0.064)
LVEF	Mean (SD)	60.1 (16.9)	58.7 (8.3)	62.6 (7.1)	1.00 (0.92-1.08, p=0.985)	0.94 (0.83-1.04, p=0.257)
NYHA	Class I	8 (28.6)	11 (39.3)	9 (32.1)	-	-
	Class II	3 (42.9)	3 (42.9)	1 (14.3)	0.53 (0.10-3.21, p=0.470)	**0.02 (0.00-0.58, p=0.059)**
	ND			1 (100.0)	6260544.32 (0.00-NA, p=0.995)	-
Familial	Yes	5 (35.7)	4 (28.6)	5 (35.7)	-	-
	No data	1 (20.0)	2 (40.0)	2 (40.0)	2.22 (0.24-50.09, p=0.523)	-
	No	6 (27.3)	10 (45.5)	6 (27.3)	1.48 (0.34-6.37, p=0.593)	9.19 (0.65-544.96, p=0.174)
Sex	Female	2 (14.3)	7 (50.0)	5 (35.7)	-	-
	Male	10 (37.0)	9 (33.3)	8 (29.6)	0.28 (0.04-1.33, p=0.143)	0.19 (0.01-2.05, p=0.217)
Ancestry	Black-African	0 (0.0)	3 (100.0)	0 (0.0)	-	-
	Mixed	2 (25.0)	3 (37.5)	3 (37.5)	0.00 (NA-∞, p=0.994)	0.00 (NA-∞, p=0.996)
	White	9 (33.3)	9 (33.3)	9 (33.3)	0.00 (NA-∞, p=0.994)	0.00 (NA-∞, p=0.996)
	Indian	1 (33.3)	1 (33.3)	1 (33.3)	0.00 (NA-∞, p=0.994)	-

Number in data frame = 41, Number in model = 30, Missing = 11, AIC = 38.7, C-statistic = 0.878, H&L = Chi-sq (8) 6.82 (p=0.556)